Energy & Buildings 352 (2026) 116788

Contents lists available at ScienceDirect

Energy & Buildings

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

A comparative analysis of prototype and building-by-building methods for urban building energy modeling in a hot-summer and warm-winter climate zone

Chengjin Wu [a], Chengcheng Song [a], Fangyu Liu [a], Anni Xu [a], Yixing Chen [a,b,*]

[a] College of Civil Engineering, Hunan University, Changsha 410082, China
[b] Key Laboratory of Building Safety and Energy Efficiency of Ministry of Education, Hunan University, Changsha 410082, China

ARTICLE INFO

Keywords:
Urban building energy modeling
Prototype buildings
City-scale building dataset
Retrofit analysis
Building-by-building

ABSTRACT

Urban building energy assessment and retrofit analysis are crucial for formulating effective energy conservation and emission reduction policies. However, there is still a lack of research on the variability of different approaches to urban building energy modeling (UBEM) and their applicability in specific climate zones. This study uses Haikou City, located in a region with hot summers and warm winters, as a case example to investigate the differences between the prototype building method and the building-by-building method in urban-scale energy consumption simulation. Firstly, by integrating multi-source geographic information data and combining spatial analysis, cluster analysis and deep learning algorithms, this paper successfully identifies the types and construction years of 119,784 buildings. Subsequently, 57 prototype buildings were developed for Haikou City in accordance with national standards. Using the AutoBPS tool, two UBEM benchmark models covering 100,636 buildings were constructed. The application effectiveness of the two methods in retrofitting lighting systems, air-conditioning systems, and photovoltaic (PV) systems was analysed and compared. The results indicate that the total electricity consumption simulated using the prototype building method was 14,980 GWh, differing by approximately 7 % from the total simulated using the building-by-building method. This demonstrates that at the urban scale, the prototype building method enables rapid and practical assessment of overall building energy performance. Further energy retrofit analysis indicates that PV systems offer the most significant potential for upgrades, yielding an overall energy savings rate of 30.7 %. Upgrades to lighting systems and air conditioners would achieve savings rates of 8.2 % and 7.9 %, respectively. This study not only provides valuable reference for selecting building energy consumption modeling methods at the urban scale, but also offers data support for analyzing building energy characteristics and renovation strategies in other cities within the same climate zone.

1. Introduction

Urban areas account for 75 % of global primary energy consumption and 70 % of global greenhouse gas emissions [1]. Additionally, over half of the global population resides in urban areas, with the urban population share projected to reach 68 % by 2050 [2]. Rapid urbanization and population growth in cities have resulted in substantial energy consumption and increased carbon emissions. Buildings, as a significant component of cities, account for 60–80 % of total urban carbon emissions [3]. Reducing energy consumption in existing urban buildings is crucial for the building sector to achieve carbon peaking and carbon neutrality [4]. Therefore, by conducting a detailed assessment of the energy needs of urban buildings, city managers can formulate more informed local energy policies, thereby reducing energy consumption and carbon emissions.

The UBEM methodology was developed to assess the energy demand of urban buildings. This methodology is broadly categorized into two primary approaches: top-down and bottom-up[5]. The top-down approach typically uses building stock data, technology adoption models, and economic models to establish statistical and regression models, evaluating urban building energy levels from a macro perspective [6]. However, this approach lacks physical interpretability and cannot analyze building-level energy consumption or energy-saving retrofits. Bottom-up approaches apply statistical or physical modeling to

* Corresponding author at: College of Civil Engineering, Hunan University, Changsha 410082, China.
E-mail address: yixingchen@hnu.edu.cn (Y. Chen).

https://doi.org/10.1016/j.enbuild.2025.116788
Received 21 September 2025; Received in revised form 11 November 2025; Accepted 25 November 2025
Available online 26 November 2025

characterize each building [7]. Statistical methods use regression models or machine learning algorithms to link building characteristics (e.g., geometry, building type, and occupancy) with measured energy data to rapidly estimate urban building stock energy consumption [8]. For example, Choi et al. [9] used government data to obtain actual energy consumption data for 15,654 buildings in Seoul, South Korea, for modeling. Subsequently, data fusion techniques extracted three new building features from open-source data. Five scenarios were set up and combined with a multi-layer perceptron algorithm to validate the optimal combination. Finally, the best-performing model was used to predict the energy consumption of 31,737 buildings in two additional districts of Seoul. Sun et al. [10] collected Energy Performance Certificates and Google Street View image data for 30,000 buildings in Glasgow, UK, and utilized a dense convolutional neural network combined with a fully connected neural network to predict the buildings' energy efficiency ratings. Statistical methods can effectively evaluate urban building energy consumption, but acquiring large-scale local building energy data remains difficult. Although governments in some developed countries (e.g., the United States [11], Europe [12], and South Korea [9]) provide publicly accessible building energy data, most cities in China and other developing countries lack surveys and statistics on actual building energy data. For cities and regions lacking actual building energy databases, bottom-up physical modeling has become a primary method for calculating UBEM.

Physical modeling methods calculate building stock energy consumption using physical laws and thermodynamic principles [13]. The modeling workflow primarily involves data preprocessing and input, model generation, simulation, result output, and application [14]. Data inputs include building geometry data (e.g., building area, height, coordinate location), non-geometric parameters (e.g., envelope parameters, HVAC systems, schedules), and weather data [15]. Acquiring building geometric parameters is relatively easy. In recent years, many studies have used image recognition algorithms and remote sensing technologies to obtain geometric parameters such as building outlines [16], building heights [17], and building roofs [18] on a large scale. However, acquiring non-geometric parameters for large-scale buildings is very difficult. To address this situation, the building archetype method was developed [19]. A building archetype refers to a group of buildings with similar attributes, such as construction time and building type. However, only some countries and cities currently possess comprehensive archetype databases, such as the U.S. Department of Energy's prototype building data [20], Italian building archetype data [21], and the Swiss energy performance certificate database [22]. For most cities, it is still necessary to develop building archetype libraries that suit local conditions. The development of building archetypes generally involves two steps: classification and characterization [19]. Classification requires dividing buildings into categories based on building type, construction vintage, geometric size, etc. Characterization assigns corresponding building attributes, such as envelope thermal properties and HVAC systems, to the classified archetypes based on pre-definitions or local regulations. With the development of artificial intelligence, most current research uses machine learning or deep learning for building classification. Bandam et al. [23] utilized a random forest algorithm to classify 29 million buildings in Germany into 15 building types. Hoffman et al. [24] collected aerial and street-view image data for 56,259 buildings in the US and used a CNN deep learning framework to classify them into 4 types. Deng et al. [25] obtained 68,966 buildings in Changsha via satellite imagery and GIS data and successfully identified 22 building types using random forest and K-means algorithms. Further research [26] used the Mask-R-CNN algorithm to identify building construction time and created 66 building archetypes based on construction time and building type.

After obtaining classified building archetypes and assigning predefined thermal attributes, two main methods are used to evaluate UBEM performance: the prototype building method and the building-by-building method. The prototype building method classifies a city's building stock into different prototype buildings and calculates the Energy Use Intensity (EUI) for each. By multiplying the total area of the prototype buildings by their corresponding EUI, the method can quickly estimate the energy consumption of a city's buildings. Mohammadiziazi et al. [27] obtained geometric parameters for 209 commercial buildings in Pittsburgh using LiDAR, Google Street View imagery, and local data centers. They created 20 archetypes based on ASHRAE standards and U. S. DOE commercial reference building data. The results showed a 7 % difference between simulated total energy consumption and actual data. Buckley et al. [28] used 9,000 residential buildings in Dublin, Ireland, as a case study, creating a geographic database from GIS building footprint data. Subsequently, each building was classified through visual inspection and building age data, and corresponding archetypes were assigned based on the Tabula archetype library. The study derived 10 residential building archetypes and used UMI combined with EnergyPlus to obtain the EUI for each. Song et al. [29] obtained GIS data for Shanghai, such as building footprints, Points of Interest (POI), and Areas of Interest (AOI), and combined this with the K-means clustering algorithm to obtain 21 building types. Subsequently, they used satellite imagery combined with a deep learning algorithm to obtain building vintage information. The study created 63 building archetypes for 539,119 buildings in Shanghai based on building type and vintage. Finally, they used AutoBPS software to simulate the EUI data for the archetypes and assessed the overall energy consumption level of Shanghai. The prototype building method can quickly estimate the energy consumption and energy use characteristics of urban building complexes. However, the prototype building method does not consider the mutual influence between individual buildings, such as shadowing and long-wave radiation heat exchange [30], which is crucial for energy consumption at the individual building level.

The building-by-building calculation method takes into account factors such as shading, shadows, and heat exchange, providing more detailed simulation results and enabling building-level energy consumption analysis. HosseiniHaghighi et al. [31] collected three GIS datasets for Kelowna, Canada (Model City Dataset, Building Footprint Dataset, and Address Point Dataset) and created building geometry models. Subsequently, they used Canadian Survey of Household Energy Use data and the HOT2000 internal database to assign 8 building archetypes to 2,950 residential buildings. The study also considered inter-building shading and simulated the heating energy consumption of the 2,950 buildings using the building-by-building method. Davila et al. [7] utilized official GIS datasets and a building archetype library to create 52 archetypes for 83,541 buildings in Boston. Subsequently, the study used the UMI platform with the EnergyPlus engine on two 16-core computers to calculate the energy consumption of the 83,541 buildings, with a total simulation time of approximately 60 h. Deng et al. [32] selected a central area in Changsha with 3,633 buildings as the study area. They assigned archetypes to these buildings based on 66 archetypes previously created from 59,332 buildings in Changsha. Subsequently, they used the AutoBPS tool to simulate the EUI for the 3,633 buildings individually and evaluated the area's energy consumption level and energy-saving retrofit potential. As the above studies show, the building-by-building method can provide detailed, building-level energy analysis, but the resource and time cost for urban-scale calculation is enormous. To reduce data processing and automate large-scale simulations, many tools and platforms have been developed, such as SimStadt [33], CityBES [34], CitySim [35], CEA [36], TEASER [37], and the UMI [38] and AutoBPS [32] tools used in the previously mentioned studies. Even so, the resource and time consumption for building-by-building simulation of large-scale building stocks remains considerable.

In conclusion, UBEM provides crucial reference points and data support for urban energy consumption analysis. However, when analysing urban energy consumption, most studies employ either the prototype building approach for computational efficiency or the building-by-building method for high fidelity. Few studies systematically quantify and analyse the differences between these two approaches in terms of building energy consumption characteristics at the urban scale. Furthermore, building energy-saving potential analysis is crucial for cities to achieve energy saving and emission reduction. Most current studies also analyze this potential using only the prototype building method or the building-by-building method. For energy-saving technologies unaffected by inter-building shading (e.g., HVAC system retrofits) and technologies that need to consider shading impacts (e.g., rooftop photovoltaic retrofits), the differences between the two modeling methods remain unknown. By quantitatively comparing the analysis of the two modeling methods on urban energy consumption and energy-saving potential, this can provide a reference for engineers when selecting an appropriate modeling method for analyzing urban energy consumption. On the other hand, current research on urban building energy consumption and saving potential in China has largely focused on Hot-Summer-Cold-Winter regions [26,29,39], while energy consumption modelling and Energy Conservation Measures (ECMs) analysis for cities in Hot-Summer-Warm-Winter regions remain scarce. Additionally, Hot-Summer-Warm-Winter regions have distinct energy

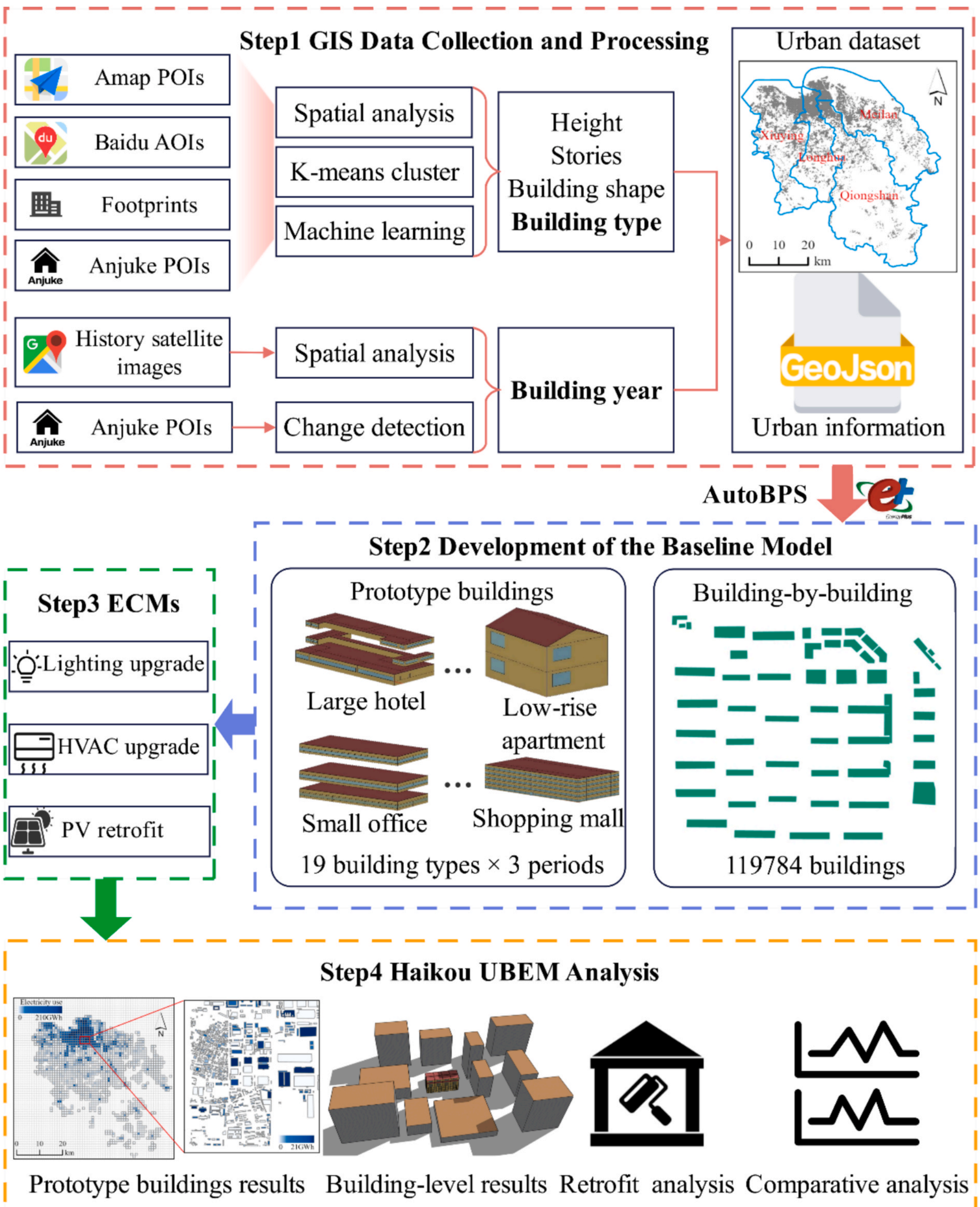


Fig. 1. The workflow of this research.

consumption patterns, dominated by cooling demands, which necessitate the development of corresponding UBEMs to analyze their unique building energy characteristics. Furthermore, their geographical advantages provide abundant solar resources, presenting significant, and largely un-analyzed, potential for rooftop photovoltaic (PV) applications.

Therefore, this paper takes Haikou City, a typical region with hot summers and warm winters, as a case study. Based on GIS data and local building characteristics, the AutoBPS tool was employed to establish both a prototype building model and building-by-building models for Haikou. Building upon this foundation, the energy consumption characteristics of different building types in Haikou were analyzed. Subsequently, the prototype method and building-by-building method were applied to conduct retrofit analyses of lighting systems, COP performance, and PV systems, thereby quantifying the energy-saving potential of Haikou's building stock under various measures. Additionally, this study compares the overall energy outcomes of buildings at the urban scale using the prototype building method and the building-by-building method, providing valuable reference for selecting modeling approaches at the urban scale in future research.

2. Methods

Fig. 1 shows the overall process of this study's development of UBEM for Haikou City and the comparative analysis of energy-saving renovations. First, basic GIS data for Haikou City was collected, primarily including POIs, AOIs, community information, building footprints, and historical satellite imagery. GIS spatial analysis, clustering analysis, and machine learning methods were employed to determine building height, the number of stories, and the building type. For the acquisition of building construction years, this study referenced the research by Song et al. [40], collecting historical satellite imagery from 2000 to 2024. Changes in satellite imagery were analyzed annually to estimate the specific construction years of buildings. Based on the above data, a comprehensive building dataset for Haikou City was established in GeoJSON format. Based on the GeoJSON data, following national and local standards, AutoBPS was used to establish both prototype building UBEMs and building-by-building UBEMs for Haikou City, simulating baseline energy consumption at the city scale. Subsequently, three ECMs were applied to the UBEMs created using these two methods: lighting system upgrade, COP performance upgrade, and addition of a PV system. Simulation analyses of the three ECMs were conducted using the AutoBPS-Retrofit and AutoBPS-PV modules. Finally, the applicability of the Haikou City prototype modeling method and the building-by-building modeling method was compared in terms of baseline energy consumption models and ECMs. This provides a specific reference basis for energy-saving renovation methods and measures in Haikou and cities in the same climate zone.

2.1. Case study

Haikou is the provincial capital of Hainan Province, located between $19°31' - 20°04'$ north latitude and $110°07' - 110°42'$ east longitude. Haikou City is situated at the northern edge of the low-latitude tropical zone and has a tropical monsoon climate. The city is characterized by high annual solar insolation and receives substantial levels of solar radiation year-round. The annual average sunshine duration exceeds 2000 h, and the average yearly temperature is 24.4 °C. According to the national standard GB 50176–2016, Haikou is classified as a region with hot summers and warm winters, categorized as 4A.

Haikou City is divided into four districts: Xiuying, Longhua, Qiongshan, and Meilan. It has a land area of 2,296.82 square kilometers and a permanent population of 3 million. Fig. 2 shows a map of Haikou, along with POI kernel density information and footprint information. The POI kernel density reveals the distribution of buildings and the level of urban development. A higher POI kernel density indicates a relatively higher building density, suggesting that the area is more developed.

2.2. Establishment of the urban building database

2.2.1. Data Collection and Introduction

Before conducting energy consumption simulations for urban building complexes, it is necessary to collect building complex data for the city. As shown in Fig. 1, the urban building complex database primarily includes geometric information for each building (building shape, height, number of floors, etc.), building type, and construction year. However, since most countries and regions currently lack comprehensive city-level building databases, making it extremely challenging to directly obtain building types and construction years at the city scale. With the advancement of big data, it is possible to indirectly infer building types and construction years by accessing GIS data such as POI, AOI, building footprints, and satellite imagery provided by

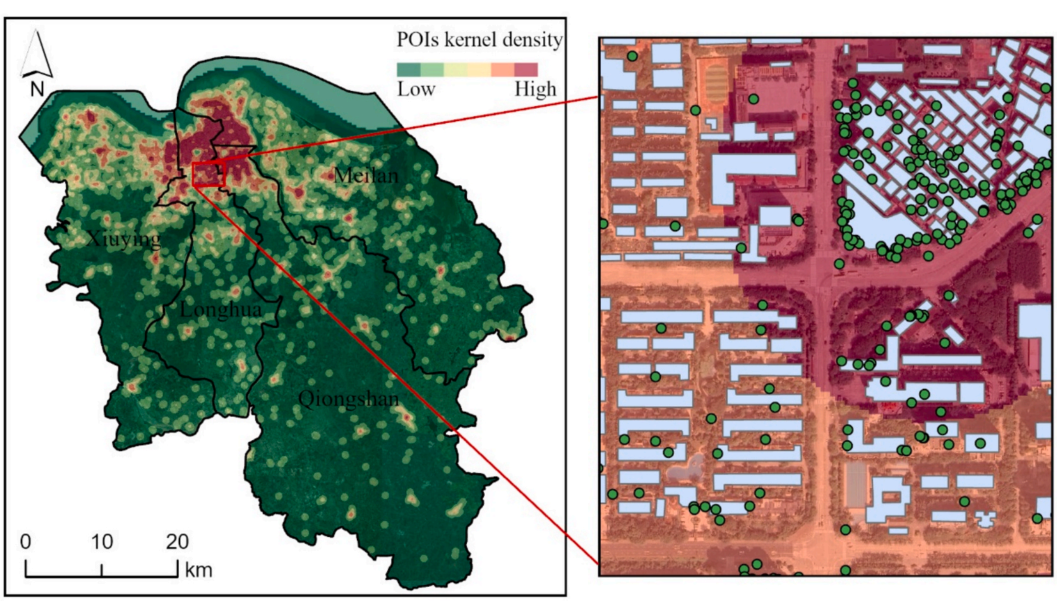


Fig. 2. POIs kernel density and footprint information in Haikou.

map service providers. Based on this approach, this paper collected POI, AOI, footprint, community information, and satellite imagery data for Haikou City. Specific details are shown in Table 1. Table 1 lists the basic information for each data source. Haikou City has a total of 119,784 building footprints, covering a building area of 462.28 square kilometers. Building footprints correspond to the geometric data of structures. The building footprints data sourced from Amap used in this study includes geometric shapes and the number of floors, but lacks height information. Building height can be calculated by multiplying the number of floors by the corresponding floor height. This paper assumes a floor height of 3 m for residential buildings and 4 m for public buildings. This is crucial for subsequent statistical analysis of prototype buildings and modeling. Among these, POI, AOI, and residential community information were primarily used for building type identification, while residential community information and satellite imagery were used mainly for determining building age. Sections 2.2.2 and 2.2.3 will detail the steps for building type identification and construction year identification.

2.2.2. Building type identification

This study identifies building types based on collected POI, AOI, and building footprints. A POI is a point in online maps containing name, coordinates, and category information. POI category information is divided into primary and secondary attributes. Primary attributes reflect a building's main function, while secondary attributes further refine its functional classification. An AOI is a bounded area formed by connecting multiple coordinate points. Its function is similar to that of a POI, primarily used to determine the primary function of the building complex within the bounded area. Fig. 3 shows the specific process of building type identification in this study. Based on the collected POI, AOI, and community information, spatial analysis was performed using ArcGIS Pro software to obtain buildings with main attributes, buildings without main attributes, and buildings without any attributes. For buildings with primary attributes and those without primary attributes, the building type is determined using the K-means clustering algorithm. For buildings without any attributes, machine learning methods are used to infer their building type based on the distinct features of their footprints. The following sections will provide a detailed explanation of these methods.

(1) GIS spatial analysis

Based on the acquired GIS basic information, ArcGIS Pro software was used to perform spatial analysis and associate GIS data (such as POI and AOI) with building footprints. To enhance the coverage of POI data for buildings while ensuring the accuracy of information association, this study sets the tolerance for POIs to 5 m based on the coverage of POIs in Haikou City and references previous research [29]. If POI data falls within a 5-meter radius of a building, the POI attributes are assigned to the corresponding building. After assigning POI attributes to building outlines, AOI attributes must also be assigned to the corresponding building outlines. When assigning AOI attributes, if a building outline occupies 70 % or more of an AOI's area, it is considered a building within that AOI, and the attributes are assigned to the building outline within the AOI. After assigning both POI and AOI attributes, a total of 46,456 buildings possess attributes. Through classification, nine primary categories are identified: residential, school, office building, hotel, factory, hospital, shopping mall, cultural and art museum, and university. These primary categories directly reflect the primary functions of buildings. Additionally, four secondary categories are identified: retail, restaurant, corporate, and leisure and entertainment. To further classify building types in detail, this study made the following assumptions: 1) Buildings containing only one primary category are classified according to that primary category. 2) Buildings containing two primary categories, each accounting for over 1 %, are classified as mixed buildings of the corresponding type. 3) Buildings containing three or more primary categories are classified as other mixed buildings. 4) Buildings containing no primary categories but containing secondary categories are classified using clustering methods. Using the aforementioned method, a total of 32,461 buildings were assigned their architectural types, accounting for 27 % of the total number of buildings.

(2) Clustering analysis

For buildings lacking a primary attribute but possessing numerous secondary POI attributes (e.g., leisure and entertainment, food and beverage services), the secondary attributes cannot directly determine the building's function. Their function is therefore inferred through clustering. Similarly, clustering is applied to buildings that have both primary and secondary POI attributes to identify additional mixed-use types. This study employs the K-means algorithm, one of the most widely used clustering methods due to its simplicity, speed, and interpretable results [41]. Since K-means requires predefined cluster counts, the elbow method is utilized to assist in determining the optimal number of clusters. Fig. 4 shows the clustering results for buildings without any primary attribute. This paper employs the elbow method to categorize residential buildings without primary attributes into five distinct groups. These categories represent the proportion of POI across five building types, thereby inferring the primary functions of the structures. Based on the POI percentage distribution, Cluster 1 was identified as Office, Clusters 2 and 3 as Restaurant, Cluster 4 as the Other Non-Residential, and Cluster 5 as Retail. Fig. 5 shows the results of further categorization using secondary attributes with Hotel as the main attribute., This analysis, also using the elbow method, yielded six clusters. Based on the POI proportions, Cluster 1 was classified as Mixed Office-Hotel, Clusters 2 and 3 as Mixed Hotel-Retail, and the remaining three clusters as Hotel. Overall, this study employed the K-means clustering method to supplement the classification of an additional 13,995 buildings, accounting for 12 % of the total building count.

(3) Machine learning methods

For the remaining 73,328 buildings with no attributes, supervised learning methods were used to identify building types based on geometric features, as these features often differ significantly between types. This study utilized ArcGIS Pro software to extract geometric features of buildings, including building footprint area, perimeter, number of floors, approximate rectangularity ratio, approximate rectangular width, and approximate rectangular aspect ratio. In fact, for these footprints with no GIS information, the vast majority are residential buildings. Therefore, this paper manually labeled 3036 buildings into residential and non-residential categories and trained six commonly used machine learning algorithms: decision trees, support vector machines, artificial neural networks, random forests, KNN, and logistic regression. These machine learning algorithms are implemented using Python and the scikit-learn third-party library, with each algorithm's hyperparameters set to the default values provided by the scikit-learn library. This study compared the performance of six different machine learning algorithms and adopted the random forest algorithm, which

Table 1
Data sources and quantitative information.

Name	Source	Data type	Quantity	Year
POI	Amap (https://ditu.amap.com/)	Shapefile	121,517	2024
AOI	Baidu map (https://map.baidu.com/)	Shapefile	6458	2024
Community information	Anjuke (https://cs.anjuke.com/)	csv	3657	2024
Footprints	Amap (https://ditu.amap.com/)	Shapefile	119,784	2024
Historical satellite imagery	Google map (http://www.gditu.net/hainan.html)	tiff	24	2000–2024

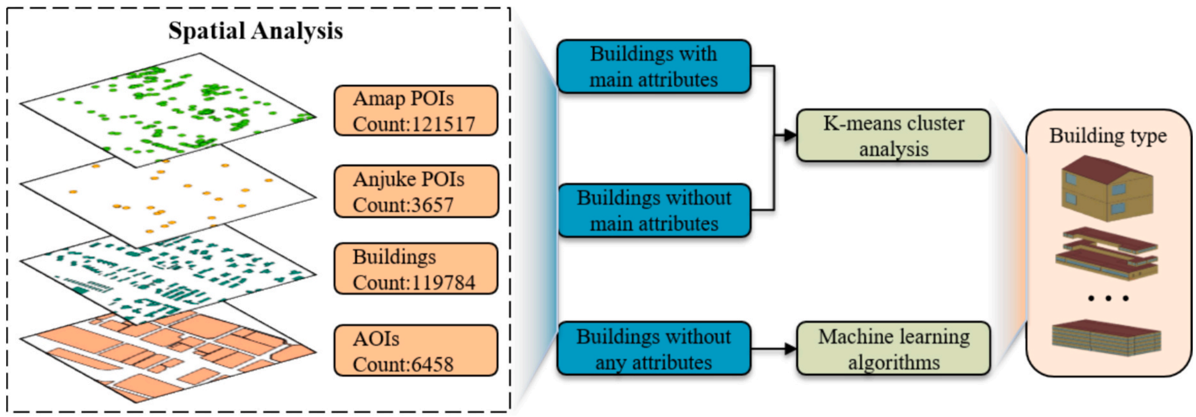


Fig. 3. The process of identifying building type.

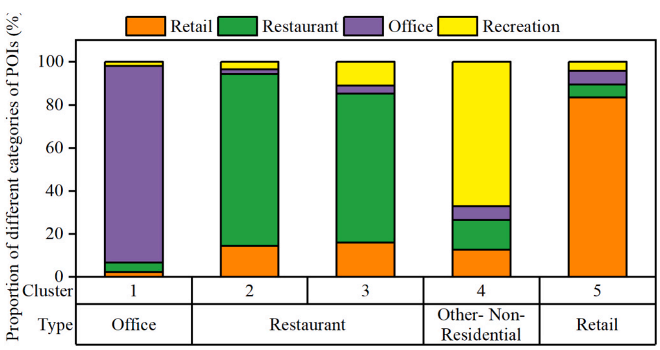


Fig. 4. Clustering results of buildings without main attributes.

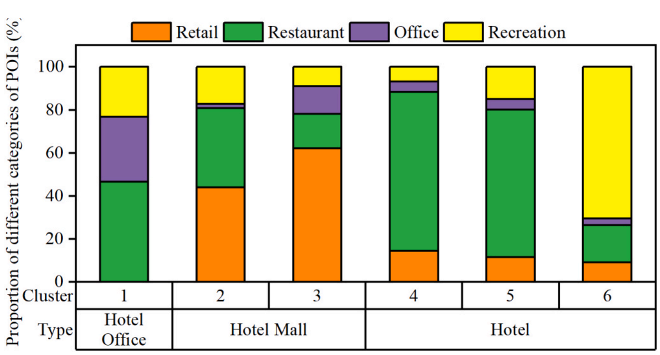


Fig. 5. Clustering result with hotel buildings.

achieved the highest accuracy (81 %), for the classification prediction of the remaining buildings. For more detailed steps, please refer to our previous research [29].

2.2.3. Building year identification

Determining the construction year is essential for UBEM to accurately incorporate the national and local standards applicable during different periods. Fig. 6 shows the process for determining the construction year. This study employs two primary methods to identify building construction dates. The first method utilizes information from third-party websites (e.g., the Anjuke platform) to obtain construction dates for some residential buildings. For the remaining buildings, the second method combines time-series satellite imagery and deep learning algorithms to detect their construction dates. The specific implementation of these two methods is detailed below.

(1) Anjuke's POI attributes assigned to buildings

This study first sourced 3,657 residential community records from the third-party website Anjuke. These records, classified as Point of Interest (POI) data, contain attributes such as the community's construction year and price. A two-step spatial analysis in GIS was used to assign these construction years to the buildings. First, the construction year attribute from each Anjuke POI was assigned to the nearest AOI using a spatial join. Subsequently, this construction year was allocated to all corresponding residential buildings located within that AOI. Through this method, construction year information was obtained for 16,368 buildings (13.67 % of the total).

To validate the construction dates from the residential community website (Anjuke), this study performed a manual comparison against annually collected satellite imagery. A sample of 237 residential buildings with Anjuke-provided dates in the central Longhua District was selected for validation. The validation criteria were based on building performance standards. As codes released in different periods correspond to different performance parameters, this study divided construction periods based on residential building codes JGJ75-2003 and JGJ75-2012 into three ranges: 2002 and earlier, 2003–2012, and 2013 and later. A date was considered 'correct' if both the website-provided year and the satellite-observed year fell within the same regulatory range. For example, if a building's website date was 2004 and its satellite-observed date was 2007, both fall within the 2003–2012 range, and the Anjuke data was deemed correct. Table 2 shows the accuracy for the 237 validated buildings. The results indicate that the accuracy of the Anjuke data is higher for more recently constructed buildings. Based on this manual validation, the data from the Anjuke third-party website is considered reliable, with an overall accuracy of 90.3 %.

(2) Year Identification Process Based on Multi-Year Satellite Imagery

For buildings lacking construction dates, this study employs historical satellite imagery (2000 to 2024) combined with deep learning algorithms to determine construction periods. The core methodology uses annual satellite images and image classification algorithms to detect temporal changes from construction or renovation. Taking a single building as an example, satellite images from 2000 to 2024 are cropped to its outline. Image classification algorithms then determine if these images depict a building. As shown in Fig. 6, images for this building captured in 2015 and later were all classified as "building," and the roof color remained unchanged. Therefore, the building's construction year is determined to be approximately 2015. The specific steps are as follows: First, based on the IDs, coordinates, and outline information for 119,784 buildings, Google satellite imagery from 2000 to 2024 was cropped annually, yielding 119,784 images per year. Subsequently, a portion of these images was randomly sampled from each year to construct a database for detecting building age changes.

Ultimately, 1536 building images and 1529 non-building images

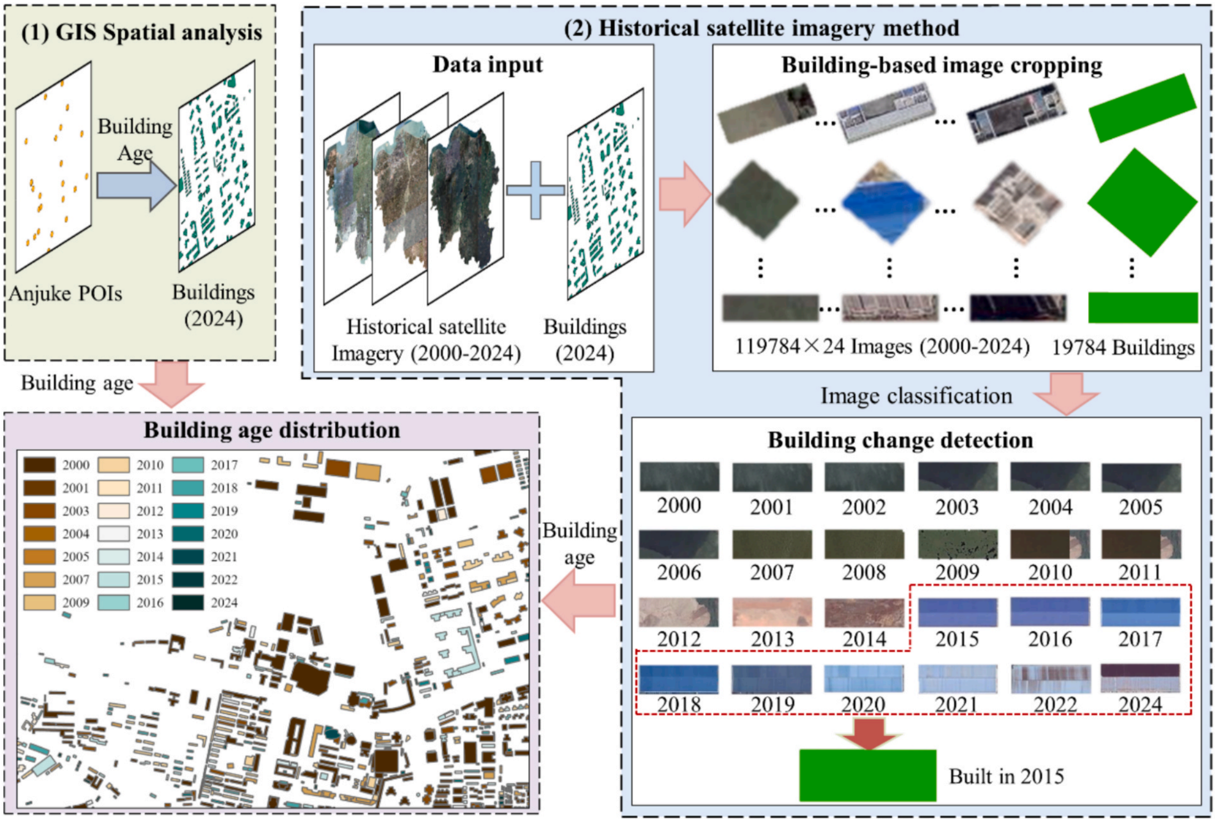


Fig. 6. Building age identification process.

Table 2
Accuracy validation of residential community construction dates.

Construction period	Residential Community Data	Manually Verified Correct Count	Accuracy rate
Pre-2002	14	9	64.3 %
2003–2012	181	163	90.1 %
Post-2013	42	42	100 %
Total	237	214	90.3 %

were obtained, totaling 3065 images. To simulate real-world image variations and enhance the model's robustness, the existing image dataset was subjected to various transformations, including rotation, brightness adjustments, cloud coverage, and shadow effects, expanding the dataset to 22,388 images.

Second, after establishing the dataset for training, an appropriate image classification algorithm must be employed to determine whether each image within a given year depicts a building. This paper employs the lightweight network architecture EfficientNet-b0 combined with transfer learning to achieve efficient and accurate annual classification predictions of satellite images of Haikou City. The EfficientNet series of models is an efficient convolutional neural network model proposed by the Google team in 2019 [42]. This model innovatively introduces a composite scaling method to uniformly scale three dimensions: network depth, network width, and input image resolution. By balancing each dimension, it achieves optimal classification performance. EfficientNet-b0 is the baseline architecture of the EfficientNet series models, featuring the fewest parameters and highest efficiency among them. Compared to the ResNet-50 and ResNet-101 models, EfficientNet-b0 has 1/5 and 1/8 of their parameter counts, respectively. Subsequently, the data-enhanced dataset is divided into training, validation, and testing sets in a 75 %, 15 %, and 15 % ratio. For loss calculation during model training, this paper adopts the binary cross-entropy loss function. It is a commonly used loss function in deep learning for binary classification tasks, with its calculation formula shown in (1).

$$Loss_{BCE} = -\frac{1}{N}\sum_{i=1}^{N}[y_i \times \log(\hat{y}_i) + (1-y_i) \times \log(1-\hat{y}_i)] \tag{1}$$

In this context, N represents the total number of pixels in the image, y_i denotes the true value of the i-th pixel, and $\hat{y}_i$ refers to the model's predicted value for the i-th pixel. The training process employs a two-stage transfer learning strategy. In the first stage, the pre-trained model EfficientNet-b0 is utilized to learn general feature representations and initialize weights. In the second stage, the pre-trained model was transferred to the building classification and detection task, and fine-tuned end-to-end by unfreezing all layers of the pre-trained model. Subsequently, the initial learning rate was set to 0.0003, the batch size to 16, and the model was trained for 30 epochs.

To evaluate the model's performance, this paper adopted commonly used deep learning model metrics, primarily including four metrics: accuracy, precision, recall, and F1 score. Accuracy refers to the proportion of correctly predicted samples, precision refers to the ratio of actual positive samples to predicted positive samples, recall refers to the proportion of actual positive samples that are correctly predicted, and the F1 score is the harmonic mean of precision and recall. After training, the performance of the prediction model on the test set was as follows: the accuracy was 0.965, the precision was 0.946, the recall was 0.987, and the F1 score was 0.966. This indicates that the model performs well in building change detection tasks.

Finally, the trained model was used to classify and predict satellite images of each building on an annual basis. By analyzing the prediction results of these satellite images, it was possible to determine the construction or renovation dates of the buildings. Ultimately, through these two approaches, the specific construction dates of 119,784 buildings in Haikou City were obtained. Since the performance of building systems constructed in different periods is related to locally issued mandatory standards, it is necessary to further classify each building according to its construction period after obtaining the specific year of construction for each building. This paper follows the residential building energy conservation design standards JGJ 75–2003 and JGJ 75–2012 for the summer-hot and winter-warm regions, dividing residential building construction periods into three categories: 2002 and earlier, 2003–2012, and 2013 and later. For other non-residential buildings, the construction periods are classified according to the National Energy Conservation Standards for Public Buildings GB 50189–2005 and GB 50189–2015, dividing public building construction periods into three categories: before 2005, 2006–2014, and 2015 and later.

2.3. Establishment of UBEM in Haikou

2.3.1. Establishment of the baseline energy model

This section aims to establish baseline energy consumption models using both the prototype building method and the building-by-building method in conjunction with the AutoBPS tool, based on the Haikou building database established in Section 2.2. AutoBPS is a recently developed UBEM tool by Hunan University [32]. AutoBPS utilizes EnergyPlus as its computational engine to simulate building energy consumption, energy retrofits, and rooftop photovoltaic analyses for urban building complexes [43–45]. Its workflow comprises three main steps: GIS data input, building energy modeling, and application analysis. AutoBPS accepts input files in GeoJSON format, an open-source standard based on JSON that features a simple data structure and flexible usage.

Prototype building methods and building-by-building methods differ in GIS data input. For the prototype building method, statistics are first compiled for the length, width, height, and number of floors of various building types based on their classification, and average values are calculated. Subsequently, the GeoJSON input file for the prototype method includes this averaged geometric data, along with climate zone, building type, and construction era for each category. Building type and construction year are used to determine applicable building codes, as shown in Table 3. This table details non-geometric parameter settings for different building types and eras, including thermal properties of building envelopes, indoor thermal disturbances, and HVAC systems. Based on prior research [26], this study assumes split-type air conditioners for residential buildings and chillers with boilers for public buildings. AutoBPS then converts this GeoJSON data into individual prototype building JSON files, performs geometric modeling, and generates corresponding EnergyPlus models for simulation. Finally, the resulting EUI for each prototype is multiplied by the total floor area of that category, and these values are aggregated to determine the city's total building energy consumption.

Table 3
Basic settings for building parameters.

Parameters	Residential buildings			Public buildings		
	Pre-2002	2003–2012	Post-2013	Pre-2005	2006–2014	Post-2015
Exterior wall U-value (W/(m^2·K))	2.47	2	1.5	2.35	1.5	0.8
Roof U-value (W/(m^2·K))	1.8	1	0.9	1.55	0.9	0.5
Window U-value (W/(m^2·K))	5.84	4.09	4.09	5.83	3.53	2.96
Window SHGC	0.62	0.36	0.36	0.62	0.41	0.39
Lighting power density (W/m^2)	8.3	6.6	5.6	11	10	9
Equipment power density (W/m^2)	8	4.3	3.8	20	17	15
Occupancy (m^2 /person)	35.3			18.5		
Cooling/heating setpoints (°C)	26/18			26/20		
Cooling/heating COP	2.7/1.9	2.7/1.9	2.9/2.2	4.2	4.4	5.9

Notes: The parameter settings for residential buildings primarily follow the Chinese national standards JGJ75-2003 and JGJ75-2012, while those for public buildings mainly adhere to GB50189-2005 and GB50189-2015.

The workflow for the building-by-building approach is simpler. In this method, the GeoJSON file containing data for all 119,784 individual buildings is imported directly into AutoBPS. This file includes each building's ID, length, width, height, number of floors, building type, construction year, climate zone, and orientation. Crucially, the building-by-building method also requires extra consideration of inter-building shading effects.

Based on this, this paper referenced the calculation method for building shadowing proposed by Deng et al. [32], adding potential building shadowing and building orientation information for each building. During the modeling process, similar to the prototype building method, a JSON file is generated for each building and a corresponding EnergyPlus model is obtained. The EUI for each building is ultimately derived, enabling the assessment of the city's overall energy consumption.

2.3.2. Establishment of the retrofit energy model

To better compare the applicability of different ECMs under the prototype building method and the building-by-building method, this paper selected three commonly used ECMs, primarily involving retrofits to lighting, COP performance, and PV systems. The energy-saving retrofit parameters for the Haikou building are listed in Table 4. The specific retrofit parameter settings in this case primarily reference two mandatory national standards in China: the Technical Standards for Nearly Zero-Energy Buildings (GB/T51350-2019) and the General Specifications for Building Energy Conservation and Renewable Energy Utilization (GB55015-2021). For lighting system renovations, the primary focus is on reducing lighting power density by adopting high-efficiency LED lights to achieve energy savings. For COP performance renovations, the primary focus is on improving system energy efficiency. By replacing equipment with higher energy efficiency, the system's COP is improved, thereby reducing the building's energy consumption.

The calculation of rooftop PV systems relies on the AutoBPS-PV module, which employs the EnergyPlus equivalent diode power generation model. This module utilizes EnergyPlus to calculate the potential of rooftop PV systems based on a JSON file that contains additional PV parameters. Table 5 presents the key parameters of PV modules used in calculating PV power generation potential. The optimal tilt angle of the PV panels can be approximated as the latitude value of the local location. Therefore, in this case, the tilt angle of the PV panels is set to the latitude of Haikou. Additionally, based on relevant research by Song et al. [29], this paper assumes that the area of the rooftop PV system is 60 % of the total roof area.

After formulating ECMs, the retrofit information was added to each baseline model's JSON file. The AutoBPS-Retrofit module was then used to batch-generate retrofitted EnergyPlus models for both the prototype and building-by-building approaches, yielding the retrofitted energy consumption results. To measure the impact of various renovation measures on building energy consumption, this paper utilizes energy savings rates to calculate the energy-saving effects of different renovation models. The specific calculation method for energy savings rates is shown in Eq. (2):

$$ESR = \frac{BEC_{\text{baseline}} - BEC_{retrofit}}{BEC_{\text{baseline}}} \times 100\% \tag{2}$$

Table 4
The ECMs for residential and public buildings.

Construction	Indicator	Residential building			Public building		
		Year	Baseline	Retrofit	Year	Baseline	Retrofit
Lighting	W/(m^2·K)	Pre-2002	8.3	5	Pre-2005	11	8/6
(Upgrade lighting system)		2003–2012	6.6	5	2006–2014	10	8/6
		Post-2013	5.6	5	Post-2015	9	8/6
HVAC	Cooling/	Pre-2002	2.7/1.9	3.5	Pre-2005	4.2	6.2
(Replace with high-performance air conditioner)	Heating	2003–2012	2.7/1.9	3.5	2006–2014	4.4	6.2
	COP	Post-2013	2.9/2.2	3.5	Post-2015	5.9	6.2

Notes: When upgrading lighting in public buildings, the lighting energy density for office areas is 8 W/(m^2·K), and 6 W/(m^2·K) for other regions.

Table 5
Parameter settings for rooftop PV modules.

Parameters	Value
Cell type	Crystalline Silicon
Number of cells in series	60
Current at maximum power	7.5A
Voltage at maximum power	30 V
Short circuit current	8.3A
Open circuit voltage	36.4 V
PV module area	1.6 m^2
Tilt angle from horizontal	19.2°N
Orientation	South
Covered roof area percentage	60 %

where BEC_{baseline} represents the building energy consumption of the baseline model, and $BEC_{retrofit}$ represents the building energy consumption after energy-saving retrofits.

3. Results

3.1. Identification results for Haikou prototype buildings

Based on the identification and detection of building types and construction dates, the study presents the classification results for 119,784 buildings in Haikou City, as shown in Fig. 7. This paper primarily considered residential buildings, commercial buildings, mixed-use buildings, and other public buildings based on their usage characteristics. As shown in the figure, 89.89 % of the 119,784 buildings were successfully identified. Residential buildings account for 73.82 %, commercial and mixed-use buildings account for 11.74 %, and other public buildings account for 4.33 %. The remaining 10.11 % of building types were unidentified. Through the building age detection algorithm, specific construction year results were obtained and grouped into periods based onrelevant standards. Most residential buildings were constructed in 2013 or later, accounting for 40.89 % of the total. Among public buildings, those built between 2006 and 2014 were the most numerous, accounting for 10.15 % of the total. This also suggests that Haikou City experienced rapid development during this period.

Due to the complex nature of universities, factories, and cultural and art museums, which require extensive on-site surveys, this study does not currently model these three building types. Other-Non-Residential buildings were also excluded. This study selected 19 typical building types for modeling, comprising 6 categories of residential buildings and 13 categories of public buildings. Based on different building types, the average geometric shapes were statistically analyzed, and corresponding typical buildings were established, as shown in Table 6. Furthermore, considering that each typical building corresponds to three construction eras, this study established a total of 57 building archtypes for modeling.

3.2. Results of baseline model in Haikou

This section primarily analyzes and compares the results of EUI and total building energy consumption in the baseline models of the prototype building method and the building-by-building method. For the prototype building method, based on the findings in Section 3.1, this paper employs EnergyPlus to simulate the baseline models and retrofit models of 57 prototype buildings. For the building-by-building approach, AutoBPS was used to create and run 402,544 EnergyPlus models across five servers, covering 100,636 buildings (each building comprising one baseline model and three energy-saving retrofit models). The simulation process took approximately 7 days, with each EnergyPlus model running for an average of about 3 min.

3.2.1. Analysis of EUI results for the baseline model

This study used the AutoBPS tool to simulate the established Haikou prototype buildings via the EnergyPlus engine. The energy consumption results for each archetype were shown in Fig. 8. Fig. 8 illustrates the electricity and gas consumption of the prototype building over various periods. To facilitate comparison of the results, the gas consumption units were converted to kWh/m^2. Overall, as mandatory regulations continue to be updated, building energy consumption shows a gradual decreasing trend across three time periods: Pre-2002/Pre-2005, 2003–2012/2006–2014, and Post-2013/Post-2015. For residential buildings, the total EUI for low-rise buildings ranges from 40.4 to 59 kWh/m^2, while that for mid-to-high-rise buildings ranges from 30 to 64.8 kWh/m^2. Residential buildings with commercial spaces have a higher EUI than pure residential buildings due to their commercial functions. Among public buildings, shopping malls have the highest total EUI, with total energy consumption ranging from 156.8 to 295.7 kWh/m^2.

In contrast, school buildings and mixed-use residential and commercial buildings have the lowest energy consumption, at 55.8–80.8 kWh/m^2 and 37.8–58.8 kWh/m^2, respectively. Retail, food service, and shopping centers have higher energy consumption due to the use of specialized equipment and the need to maintain a comfortable environment. Generally, mixed-use buildings consume more energy than single-use buildings.

For the building-by-building approach, this paper calculated the EUI distribution of baseline models for 100,636 buildings, as shown in Fig. 9. This figure illustrates the EUI distribution for electricity and natural gas consumption across different building categories during three construction periods. The color intensity represents the years Pre-2002/Pre-2005, 2003–2012/2006–2014, and Post-2013/Post-2015, respectively. To further analyze the EUI differences between the two methods, Fig. 10 was created. Fig. 10 displays the energy consumption intensity distribution for both modeling approaches across different building types and the three construction periods for each building. The figure indicates that the energy intensity values simulated by the prototype building method mostly fall between the lower and upper quartiles of the distribution obtained by the building-by-building method, validating the effectiveness of the prototype building method. Additionally, Fig. 10 includes baseline energy consumption indicators from existing research for comparison. Due to the lack of baseline data for Haikou City in existing research, this study referenced data from other regions with hot summers and warm winters. The specific data sources are listed in

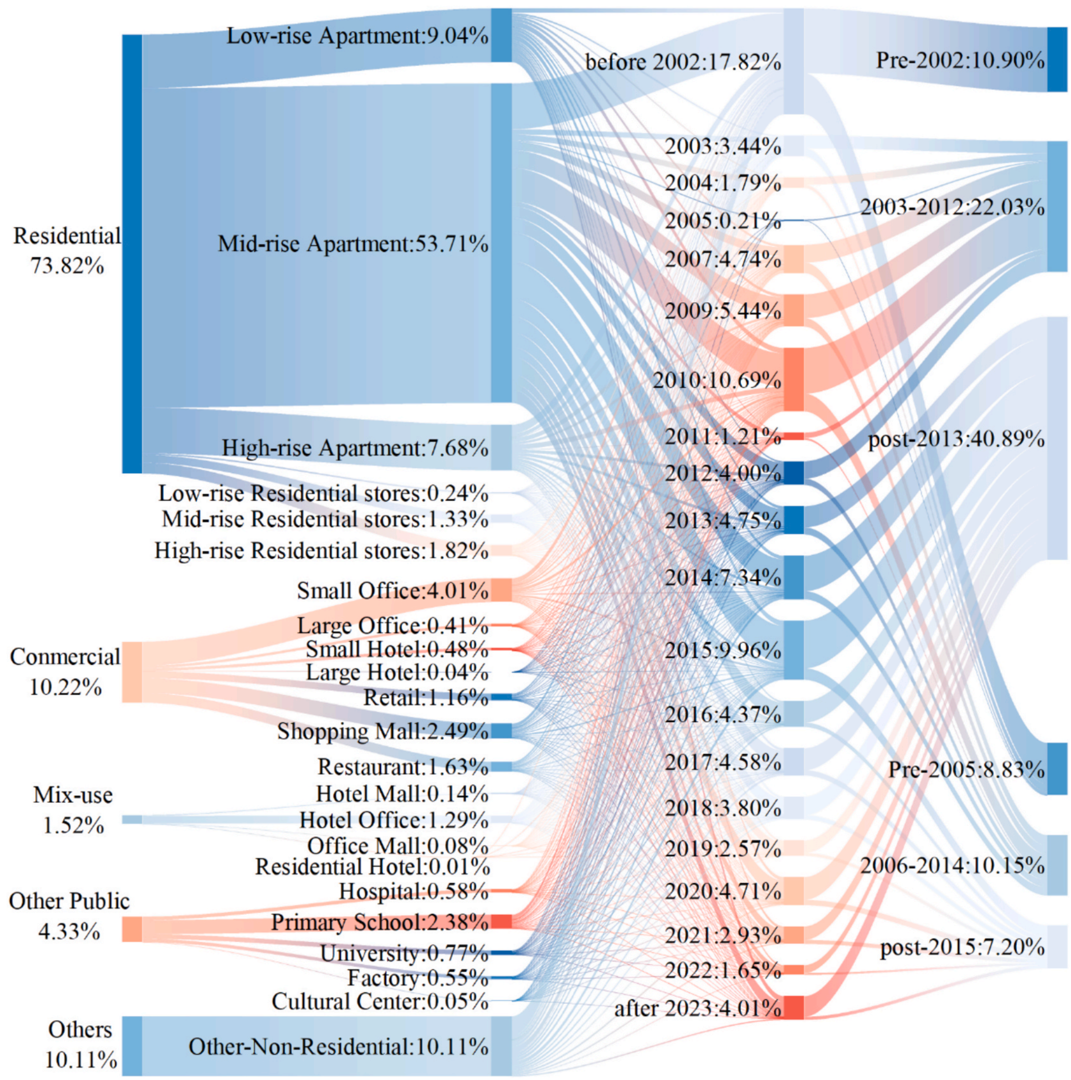


Fig. 7. Proportion of building types and construction years in Haikou.

Table 7.

3.2.2. *Analysis of total energy consumption for the baseline model*

This study calculated the total energy consumption for each building type using the EUI from both methods, as shown in Fig. 11, to further investigate the differences between the two modeling approaches. The figure reveals that for most building types, the total energy consumption calculated using the building-by-building method is lower than that derived from the prototype building method. For the entire Haikou building stock, the total energy consumption calculated by the building-by-building method is 17,520 GWh (electricity: 14,114 GWh), while the prototype method yields 18,802.1 GWh (electricity: 14,980 GWh). The building-by-building method resulted in a 7 % reduction in total energy consumption compared to the prototype method. This reduction is primarily attributed to the building-by-building method, which accounts for inter-building shading effects and building orientation, as inter-building shading significantly impacts the overall energy consumption of urban buildings [54]. Overall, while the prototype building method exhibits substantial differences in EUI distribution at the building level compared to the building-by-building approach, the variance in total energy consumption across building groups within the baseline model remains acceptable at the city scale.

This study divided Haikou into 1-square-kilometer grids to illustrate the energy consumption distribution of its building stock, as shown in Fig. 12. The figure shows the overall distribution, highlighting that high-energy-consumption areas are primarily concentrated in the northern region near the port. This concentration is due to the city's northward development, driven by transportation networks and port clusters, which has led to a large number of high-energy-consuming buildings (e.g., hospitals, shopping malls, mixed-use) in the central urban area. In contrast, the southern region retains extensive ecological green spaces and rural areas, resulting in relatively lower energy consumption.

3.3. *Analysis of energy-saving renovations*

This study analyzed the energy-saving potential of different measures across Haikou's building stock by applying three ECMs to each building: lighting system upgrades, COP performance upgrades, and rooftop PV installations. Energy retrofit analyses were conducted for 100,636 buildings in Haikou using a building-by-building approach, with results shown in Fig. 13. Fig. 13 displays total building energy consumption under the baseline model and under the three individual retrofit measures. Compared to the baseline model, the overall energy savings achieved through lighting upgrades, COP performance upgrades, and PV retrofits were 8.2 %, 7.9 %, and 30.7 %, respectively. The figure indicates that rooftop PV retrofits offer the most significant energy-saving potential in Haikou, resulting in an additional 5,381.3 GWh of electricity generation. This high potential is primarily due to

Table 6
Basic information about the Haikou prototype building.

Prototype building	Floor	Height (m)	Length (m)	Width (m)	Building area (m^2)
Low-rise Apartment	2	6	18.8	10.33	377.33
Mid-rise Apartment	4	12.27	22.48	12.34	1168.03
High-rise Apartment	20	58.51	36.03	15.86	10882.75
Low-rise Residential stores	2	6	32.63	14.08	919.82
Mid-rise Residential stores	4	13.39	54.47	22.01	3940.46
High-rise Residential stores	23	68.45	45.19	21.12	17586.15
Small Office	5	20.36	37.4	17.12	3048.22
Large Office	21	85.46	86	48.46	38571.6
Small Hotel	7	27.31	35.18	17.89	3522.44
Large Hotel	27	109.23	80.7	42.43	52038.67
Retail	5	19.04	28.96	15.39	2115.38
Shopping Mall	5	18.88	32.72	15.6	2135.67
Restaurant	5	18.99	26.97	14.29	1715.57
Hotel Mall	6	23.06	33.54	16.8	2809.87
Hotel Office	6	25.78	28.88	14.6	2424.53
Office Mall	5	21.69	32.42	15.15	2192.34
Commercial Residential	24	97	57.5	36.09	39718.8
Hospital	5	19.72	31.92	14.11	1988.9
Primary School	6	24.11	39.84	17.77	4180.77

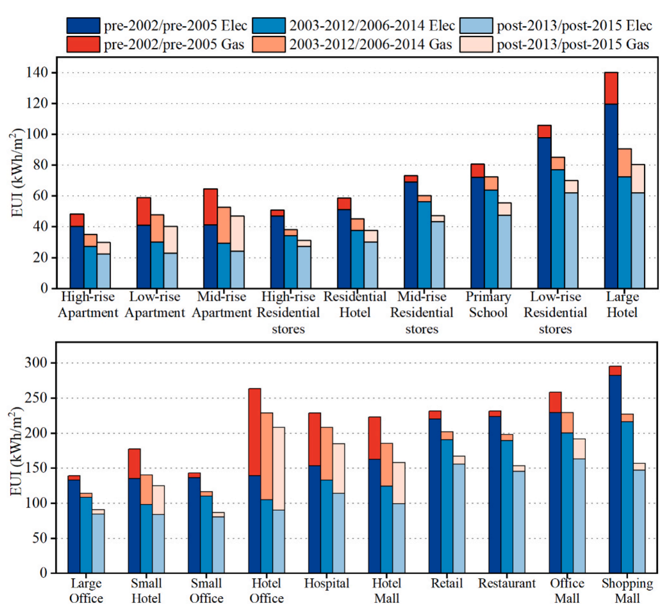


Fig. 8. Electricity and gas consumption per unit area of prototype buildings.

Haikou's location in a region with hot summers, warm winters, and abundant solar resources.

The energy-saving potential from different retrofit measures varies significantly across building types due to their different uses. The energy-saving potential results for various building types are shown in Fig. 14. Fig. 14 displays the baseline model and EUI distribution after energy-saving retrofits for 19 building types. Residential buildings show similar EUI reductions from both lighting and COP retrofits. In contrast, for commercial buildings, the improvement in COP performance exceeds that of lighting system upgrades. This is because residential buildings primarily use split-type heat pumps, while commercial buildings are mostly equipped with chillers. PV system retrofits offer the most significant energy-saving potential across all building types, with the most critical energy savings observed in low-to-mid-rise residential buildings, schools, restaurants, and other low-story structures.

Fig. 15 further illustrates the energy savings achievable for each retrofit type across different building categories. The chart reveals that for lighting system upgrades, residential buildings with retail spaces, hospitals, schools, restaurants, shopping centers, and retail buildings exhibit greater energy-saving potential, with savings exceeding 10 % across all categories. This stems primarily from their higher lighting demands, which translate to greater energy-saving potential. For COP performance upgrades, energy savings exceeding 10 % are achievable in small office buildings, small hotels, schools, retail, and restaurant buildings. These smaller commercial structures possess significant energy-saving potential. Regarding rooftop PV upgrades, since PV power generation primarily correlates with available roof area, buildings with fewer stories or larger usable roof areas achieve higher energy savings rates. Notably, low-rise residential buildings in this case exhibit harmful energy consumption after PV retrofitting. This occurs because these buildings average two stories with large roof areas, enabling annual electricity generation to exceed their energy consumption. Additionally, mid-rise residential buildings demonstrate substantial energy-saving potential through PV retrofits and constitute the largest building category. They generate the highest PV electricity output, accounting for 52 % of total generation.

4. Discussion

4.1. Baseline energy model accuracy analysis

This study established the UBEM for Haikou using both the prototype building method and the building-by-building method, analyzing the energy-saving potential of various energy-saving measures. However, the accuracy of the established UBEM requires further analysis of its reliability. Actual large-scale building energy consumption data for Haikou City was unavailable, so comparisons were limited to reports provided by the local government. Simulations using the prototype building method and the building-by-building method yielded total electricity consumption figures of 14,980 GWh and 14,114 GWh, respectively, for Haikou's 100,636 buildings. According to Haikou's 2024 statistical report, building-related electricity consumption was 10,062 GWh[55]. This represents discrepancies of 48 % and 40 % compared to the prototype building method and building-by-building method, respectively.

Table 8 shows the discrepancies between simulated values and official reported values for building electricity consumption in this study and previous research. The table indicates the primary cause of this study's large discrepancy: a substantial gap between the simulated residential energy consumption and the officially reported data. Studies in Changsha and Shanghai showed discrepancies remaining within 10 % for both residential and total electricity consumption. This study's residential discrepancy, in contrast, approaches 100 %, consequently leading to the substantial differences between total simulated energy and official reports.

The primary reasons for this are likely as follows: First, Haikou's population is significantly lower than that of other cities, resulting in low total residential electricity consumption. Second, Haikou has a developed tourism industry and a large building stock, while the actual permanent resident population is relatively small. This combination of a lower population and a high building count results in low actual occupancy rates, with some buildings remaining idle. In fact, occupancy rates show a significant positive correlation with actual building electricity consumption [56]. Energy consumption simulations, however, typically set occupancy rates based on national or local building codes, representing a relatively ideal scenario. This assumption is likely the primary reason for the significant discrepancy between the simulated results and

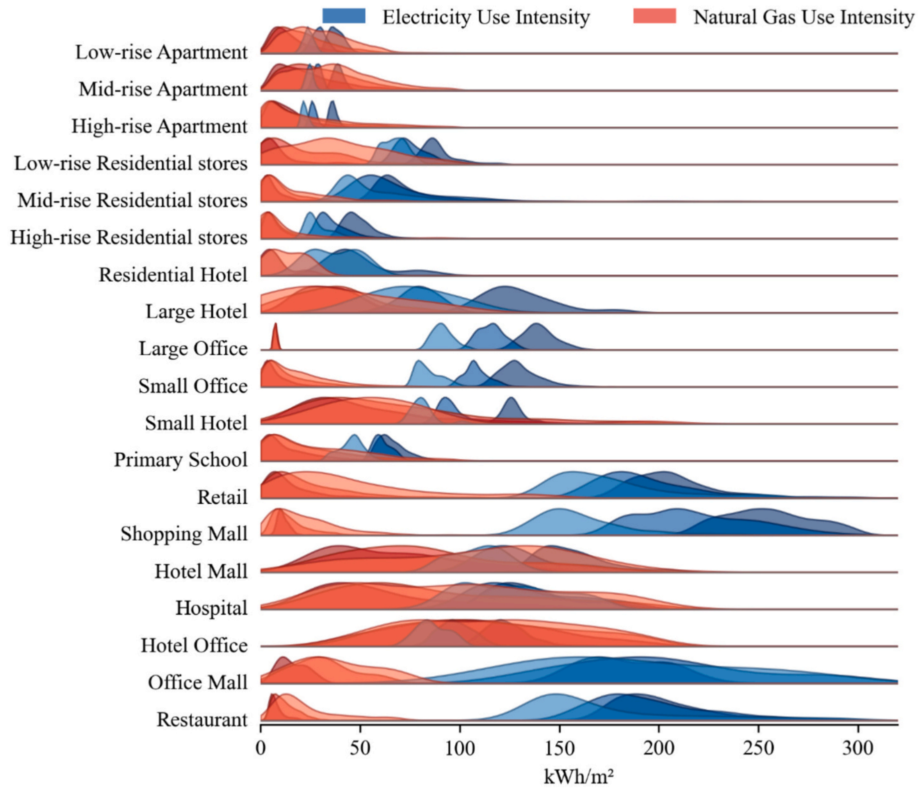


Fig. 9. EUI distribution of typical buildings in Haikou.

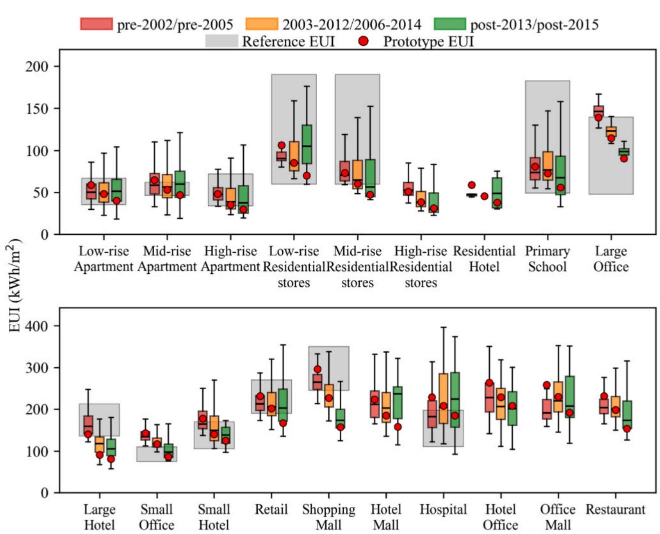


Fig. 10. Comparison of building EUI between the prototype building method and the building-by-building method.

the official values. Future assessments of urban building energy consumption must therefore consider the impact of vacancy rates, particularly in tourism-driven cities or those with an imbalance between building stock and resident population.

4.2. Comparative analysis of the two modeling methods

Section 3.2 analyzed the urban-scale energy consumption results from both the prototype and building-by-building methods, demonstrating the prototype method's applicability at this level. Previous research findings provide strong support for this conclusion. Deng et al.

Table 7
Reference EUI data comparison for some building types.

Building type	Building-by-building method (kWh/m^2)	Reference EUI (kWh/m^2)	Reference
Low-rise Apartment	18.5–100.15	35.33–66.91	[46]
Mid-rise Apartment	19.35–116.58	46.31–61.82	
High-rise Apartment	19.62–92.82	33.8–71.8	[47]
Low-rise Residential stores	59.86–176.19	60–190	[48]
Mid-rise Residential stores	41.4–140.99	60–190	
Primary School	33.19–141.44	49.22–182.73	[49]
Large Office	84.43–174.06	47.9–139.5	[50]
Large Hotel	57.57–247.93	136–213	[51]
Small Office	76.52–177.68	75–110	[52]
Small Hotel	96.85–256.96	105–170	
Retail	136.04–318.74	190–270	
Shopping Mall	124.46–360.15	245–350	
Hospital	92.2–397.47	110.83–197.5	[53]

[32] first employed the prototype building method to simulate the aggregate energy consumption of 59,332 buildings in Changsha. The study then selected a 3633-building neighborhood and used the building-by-building method for simulation. Finally, they calculated the total energy consumption for these 3,633 buildings using the EUI results from the prototype method. The results showed a 5.6 % difference between the two methods for that neighborhood, a figure that closely aligns with this paper's conclusions. This study further employed the prototype building method to calculate the results for the three retrofit measures. The total building energy consumption after lighting, COP performance, and PV retrofits amounted to 17,353.6 GWh, 17,144.5 GWh, and 12,866.5 GWh, respectively. Compared to the building-by-building method, the differences were approximately 8 %, 6 %, and 6 %, respectively, with all discrepancies below 10 %. This further demonstrates that the prototype building method enables rapid and

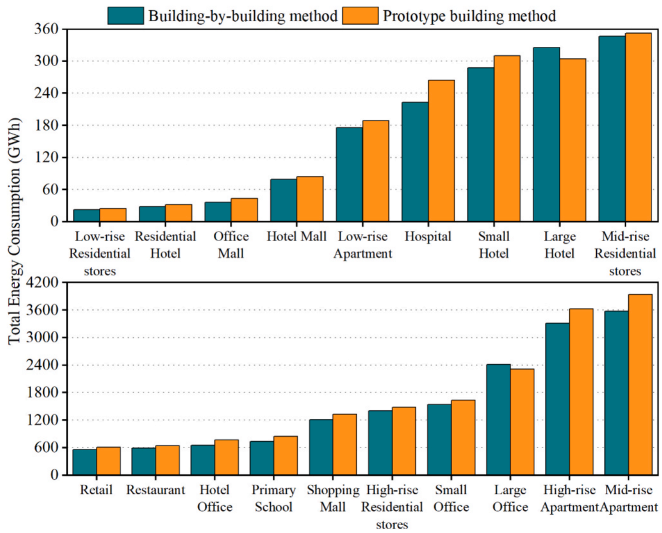


Fig. 11. Comparison of two modeling methods in terms of total building energy consumption.

reasonably accurate assessment of city-wide energy consumption at the urban scale.

This alignment validates that the prototype method's rapid and

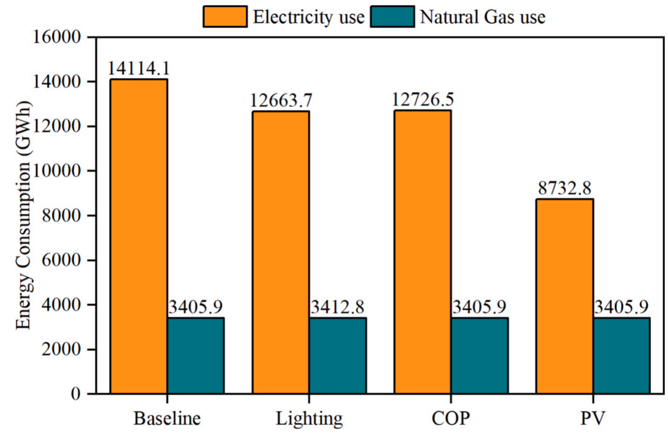
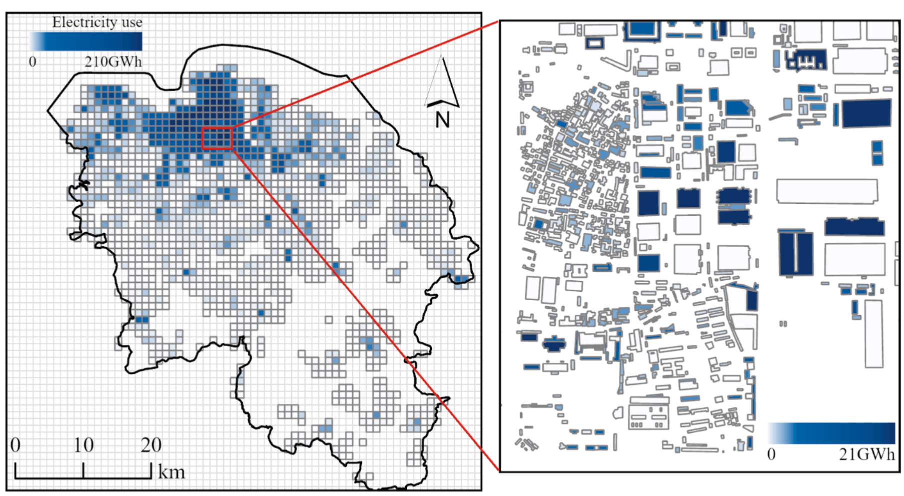


Fig. 13. Total energy consumption of buildings under different ECMs.



(a) Distribution of electricity use and detailed building electricity use

(b) Distribution of natural gas use and detailed building natural gas use

Fig. 12. Energy consumption distribution of buildings in Haikou.

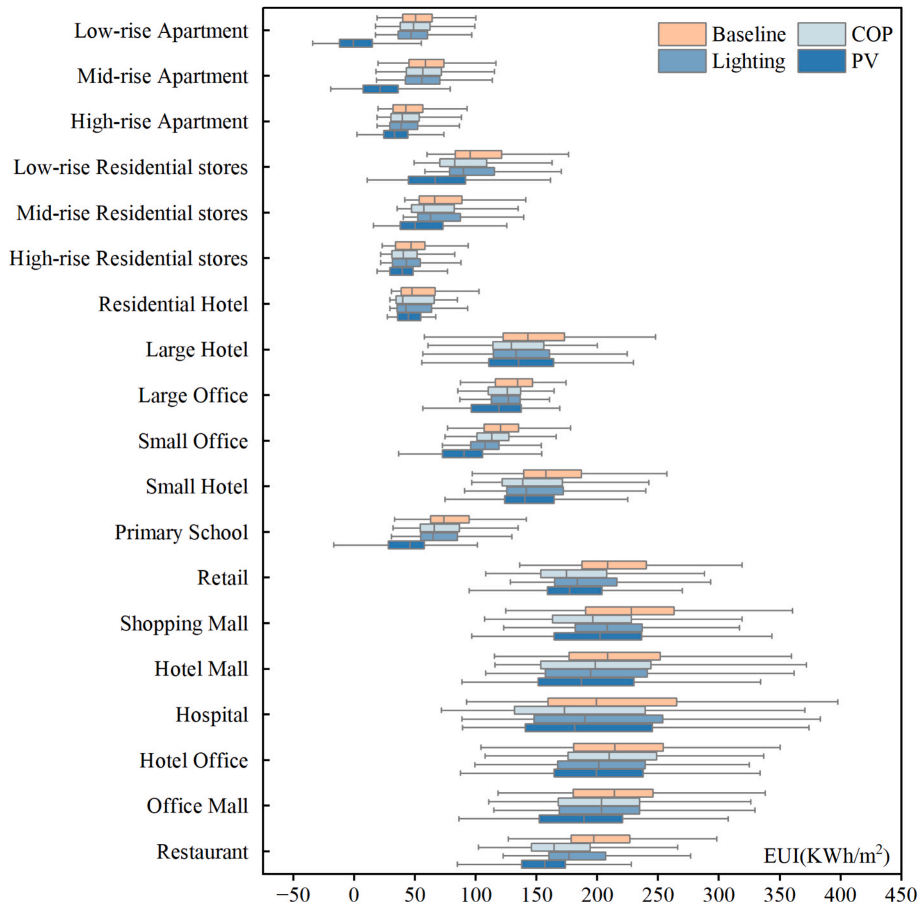


Fig. 14. Comparison of EUI Under Different ECMs.

accurate characteristics make it a strong alternative to the building-by-building method for calculating overall energy consumption. The prototype building method can accurately assess a city's overall energy-saving potential for different ECMs, regardless of whether inter-building shading is a factor (as with lighting/COP retrofits) or must be considered (as with PV retrofits). Additionally, this study compared simulation results between the two methods at the building level, as shown in Fig. 16. Fig. 16 displays the percentage differences in EUI at the building level for the baseline model. As shown in the figure, approximately 70 % of buildings exhibit an EUI difference of less than 30 % between the two modeling approaches. Additionally, about 30 % of buildings show an EUI difference exceeding 30 %. The figure indicates that for the building level, some buildings exhibit significant variations in EUI results. This is primarily due to factors such as inter-building shading effects and building orientation.

5. Conclusion

This study uses Haikou City, a typical region with hot summers and warm winters, as a case example. It establishes a baseline energy consumption model for Haikou's building stock using both the prototype building method and the building-by-building method, utilizing the AutoBPS tool. The two method were subsequently applied at the city scale to analyze retrofitting scenarios for lighting, COP performance, and PV systems. Key findings are summarized as follows:

(1) This study evaluated and compared energy consumption outcomes between the prototype building method and the building-by-building method under the baseline model and three energy-saving measures. Results indicate that the prototype building method achieved total energy consumption of 18,802 GWh, 17,354 GWh, 17,145 GWh, and 12,867 GWh for the baseline model, lighting system upgrades, COP performance upgrades, and PV retrofits, respectively. Compared to results from the building-by-building approach, the differences were all within 10 %. The prototype building method is more reliable for assessing overall urban building energy consumption and total energy-saving potential. However, at the building scale, approximately 30 % of buildings exhibit energy consumption differences exceeding 30 % between the prototype and building-by-building methods. Therefore, the prototype building method holds advantages for evaluating overall urban energy performance and energy-saving potential, while the building-by-building method is more suitable for detailed analysis at the building level.

(2) The energy-saving potential from rooftop PV retrofits reached 30.7 %, nearly double the combined potential from lighting upgrades and COP performance improvements. For cities like Haikou with abundant solar resources, PV development should be prioritized for subsidies. Low-rise residential buildings in Haikou could generate more electricity through PV retrofits than their annual electricity consumption. This provides data support for promoting net-zero energy and zero-carbon residential buildings. Additionally, mid-rise residential buildings contribute 52 % of all PV generation, making them a key focus for government-led large-scale clean energy retrofits.

(3) By comparing simulated building energy consumption with official report results, this study reveals that in tourism-dominated cities like Haikou (where building counts and actual occupancy rates are mismatched), simulating urban building energy consumption requires additional consideration of vacancy rates' impact on building energy use.

In summary, this study quantified and compared the prototype building and building-by-building methods for assessing urban building

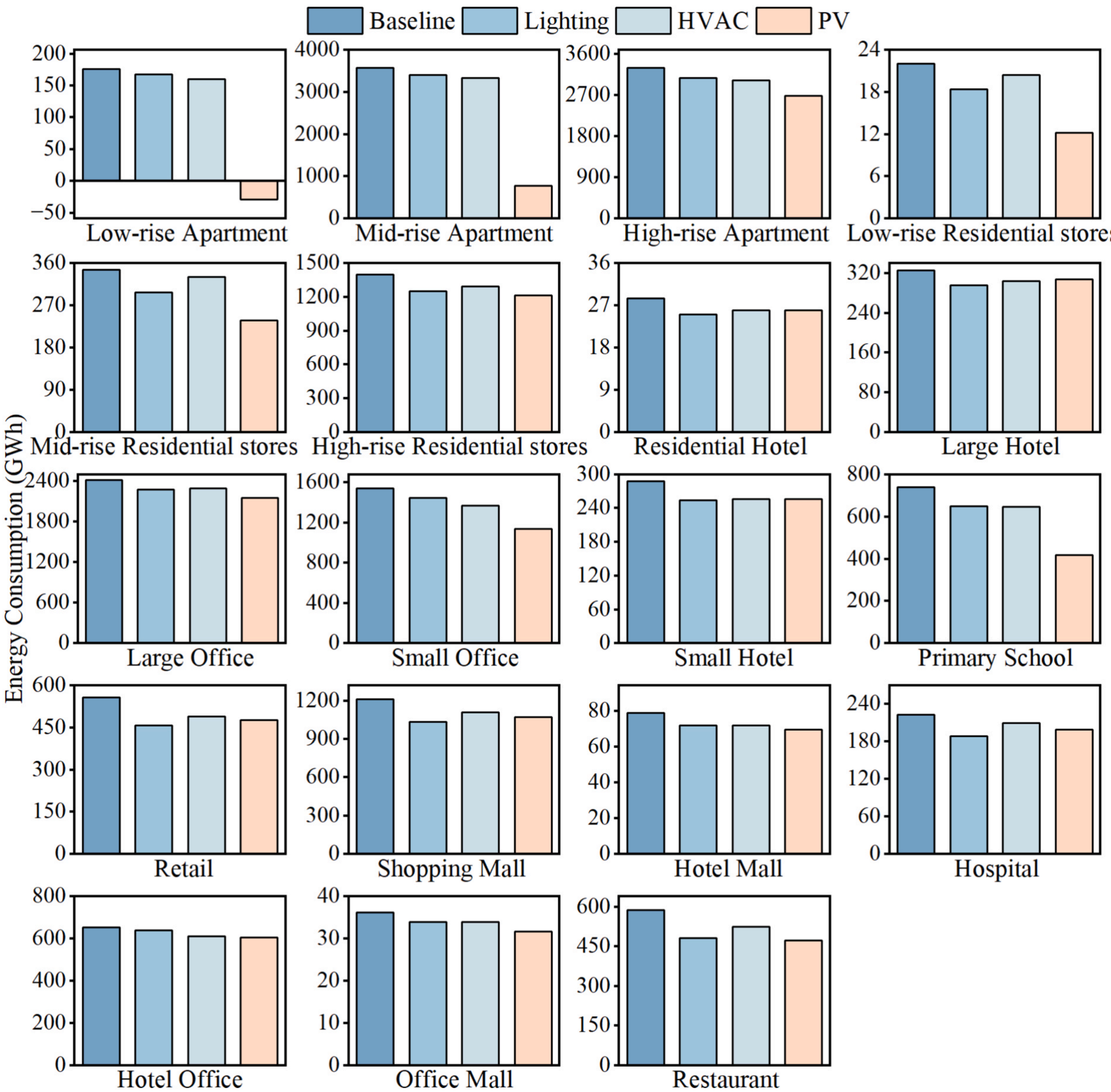


Fig. 15. Comparison of Energy Consumption Under Different ECMs.

Table 8
Comparison of building electricity consumption simulations and official report.

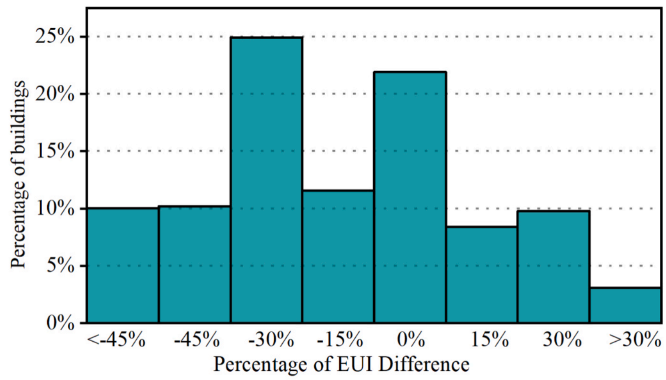
City	Number of buildings	Population (Million)	Simulated electricity consumption (GWh)		Government Report on Electricity Consumption (GWh)	
			Residential	Total	Residential	Total
Changsha [32]	59,332	9.28	7013	13,864	7373	15,184
Shanghai [29]	539,119	24.8	35,027	82,471	30,140	78,079
Haikou	100,636	2.87	6881	14,980	3381	10,062

Notes: The energy consumption data from the aforementioned studies and this study were all calculated using the prototype building method.

energy consumption and the applicability of different energy-saving measures. These conclusions provide reference for selecting modeling approaches in subsequent urban building energy consumption modeling projects and offer data support for analyzing energy retrofit potential in Haikou and other regions with hot summers and warm winters.

CRediT authorship contribution statement

Chengjin Wu: Writing – original draft, Visualization, Software, Methodology, Data curation. **Chengcheng Song:** Writing – review & editing, Visualization, Supervision, Data curation. **Fangyu Liu:** Writing – review & editing, Methodology, Formal analysis. **Anni Xu:** Writing – review & editing, Project administration, Formal analysis. **Yixing Chen:** Writing – review & editing, Supervision, Project administration, Funding acquisition.



Notes: The formula for calculating the percentage difference of EUI is as follows: $(EUI_{prototype}-EUI_{buidling\text{-}by\text{-}buidling})/EUI_{buidling\text{-}by\text{-}buidling}\times100\%$

Fig. 16. Percentage difference in EUI between the building-by-building method and the prototype building method.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This paper was supported by the National Natural Science Foundation of China (NSFC) through Grant No. 52478088. This paper is also supported by the Key Research and Development Project of Hunan Province of China through Grant No. 2025JK2072 and Grant No. 2024AQ2011.

Data availability

The authors do not have permission to share data.

References

[1] IEA, Empowering Urban Energy Transitions, IEA, Paris, 2024. https://www.iea.org/reports/empowering-urban-energy-transitions.
[2] R. Ma, Q. Li, B. Zhang, H. Huang, C. Yang, An ontology-driven method for urban building energy modeling, Sustainable Cities Soc. 106 (2024) 105394, https://doi.org/10.1016/j.scs.2024.105394.
[3] A. Anand, C. Deb, The potential of remote sensing and GIS in urban building energy modelling, energy, Built Environ. 5 (2024) 957–969, https://doi.org/10.1016/j.enbenv.2023.07.008.
[4] A. Huovila, H. Siikavirta, C. Antuña Rozado, J. Rökman, P. Tuominen, S. Paiho, Å. Hedman, P. Ylén, Carbon-neutral cities: critical review of theory and practice, J. Cleaner Prod. 341 (2022) 130912, https://doi.org/10.1016/j.jclepro.2022.130912.
[5] G. Salvalai, Y. Zhu, M. Maria Sesana, From building energy modeling to urban building energy modeling: a review of recent research trend and simulation tools, Energy Build. 319 (2024) 114500, https://doi.org/10.1016/j.enbuild.2024.114500.
[6] T. Hong, Y. Chen, X. Luo, N. Luo, S.H. Lee, Ten questions on urban building energy modeling, Build. Environ. 168 (2020) 106508, https://doi.org/10.1016/j.buildenv.2019.106508.
[7] C. Cerezo Davila, C.F. Reinhart, J.L. Bemis, Modeling Boston: a workflow for the efficient generation and maintenance of urban building energy models from existing geospatial datasets, Energy 117 (2016) 237–250.
[8] J. Ma, J.C.P. Cheng, Estimation of the building energy use intensity in the urban scale by integrating GIS and big data technology, Appl. Energy 183 (2016) 182–192, https://doi.org/10.1016/j.apenergy.2016.08.079.
[9] S. Choi, D.H. Yi, D.-W. Kim, S. Yoon, Multi-source data fusion-driven urban building energy modeling, Sustainable Cities Soc. 123 (2025) 106283, https://doi.org/10.1016/j.scs.2025.106283.
[10] M. Sun, C. Han, Q. Nie, J. Xu, F. Zhang, Q. Zhao, Understanding building energy efficiency with administrative and emerging urban big data by deep learning in glasgow, Energ. Buildings 273 (2022) 112331, https://doi.org/10.1016/j.enbuild.2022.112331.
[11] H. Wan, R. Meyer, A machine learning method of modern urban building energy modeling: a case study of chicago, Sustain. Cities Soc. 134 (2025) 106886, https://doi.org/10.1016/j.scs.2025.106886.
[12] U. Ali, S. Bano, M.H. Shamsi, D. Sood, C. Hoare, W. Zuo, N. Hewitt, J. O'Donnell, Urban building energy performance prediction and retrofit analysis using data-driven machine learning approach, Energy Build. 303 (2024) 113768, https://doi.org/10.1016/j.enbuild.2023.113768.
[13] H. Zhao, F. Magoulès, A review on the prediction of building energy consumption, Renew. Sustain. Energy Rev. 16 (2012) 3586–3592, https://doi.org/10.1016/j.rser.2012.02.049.
[14] Y.Q. Ang, Z.M. Berzolla, C.F. Reinhart, From concept to application: a review of use cases in urban building energy modeling, Appl. Energy 279 (2020) 115738, https://doi.org/10.1016/j.apenergy.2020.115738.
[15] E. Kamel, A systematic literature review of physics-based urban building energy modeling (UBEM) tools, data sources, and challenges for energy conservation, Energies 15 (2022) 8649, https://doi.org/10.3390/en15228649.
[16] W. Li, C. He, J. Fang, J. Zheng, H. Fu, L. Yu, Semantic segmentation-based building footprint extraction using very high-resolution satellite images and multi-source GIS data, Remote Sens. 11 (2019) 403.
[17] G. Liasis, S. Stavrou, Satellite images analysis for shadow detection and building height estimation, ISPRS J. Photogramm. Remote Sens. 119 (2016) 437–450.
[18] Z. Qian, M. Chen, T. Zhong, F. Zhang, R. Zhu, Z. Zhang, K. Zhang, Z. Sun, G. Lü, Deep roof refiner: a detail-oriented deep learning network for refined delineation of roof structure lines using satellite imagery, Int. J. Appl. Earth Obs. Geoinf. 107 (2022) 102680.
[19] Z. Li, J. Ma, Y. Tan, C. Guo, X. Li, Combining physical approaches with deep learning techniques for urban building energy modeling: a comprehensive review and future research prospects, Build. Environ. 246 (2023) 110960, https://doi.org/10.1016/j.buildenv.2023.110960.
[20] S. Heiple, D.J. Sailor, Using building energy simulation and geospatial modeling techniques to determine high resolution building sector energy consumption profiles, Energy Build. 40 (2008) 1426–1436, https://doi.org/10.1016/j.enbuild.2008.01.005.
[21] L. Carnieletto, M. Ferrando, L. Teso, K. Sun, W. Zhang, F. Causone, P. Romagnoni, A. Zarrella, T. Hong, Italian prototype building models for urban scale building performance simulation, Build. Environ. 192 (2021) 107590.
[22] A. Pongelli, Y.D. Priore, J.-P. Bacher, T. Jusselme, Definition of building archetypes based on the swiss energy performance certificates database, Buildings 13 (2023) 40.
[23] A. Bandam, E. Busari, C. Syranidou, J. Linssen, D. Stolten, Classification of building types in Germany: a data-driven modeling approach, DAta 7 (2022) 45.
[24] E.J. Hoffmann, Y. Wang, M. Werner, J. Kang, X.X. Zhu, Model fusion for building type classification from aerial and street view images, Remote Sens. 11 (2019) 1259.
[25] Z. Deng, Y. Chen, X. Pan, Z. Peng, J. Yang, Integrating GIS-based point of interest and community boundary datasets for urban building energy modeling, Energies 14 (2021) 1049, https://doi.org/10.3390/en14041049.
[26] Z. Deng, Y. Chen, J. Yang, Z. Chen, Archetype identification and urban building energy modeling for city-scale buildings based on GIS datasets, Build Simul-China 15 (2022) 1547–1559, https://doi.org/10.1007/s12273-021-0878-4.
[27] R. Mohammadiziazi, S. Copeland, M.M. Bilec, Urban building energy model: database development, validation, and application for commercial building stock, Energ. Buildings 248 (2021) 111175, https://doi.org/10.1016/j.enbuild.2021.111175.
[28] N. Buckley, G. Mills, C. Reinhart, Z.M. Berzolla, Using urban building energy modelling (UBEM) to support the new european union's green deal: case study of Dublin ireland, Energy Build. 247 (2021) 111115, https://doi.org/10.1016/j.enbuild.2021.111115.
[29] C. Song, Z. Deng, W. Zhao, Y. Yuan, M. Liu, S. Xu, Y. Chen, Developing urban building energy models for shanghai city with multi-source open data, Sustain. Cities Soc. 106 (2024) 105425, https://doi.org/10.1016/j.scs.2024.105425.
[30] C.F. Reinhart, C. Cerezo Davila, Urban building energy modeling – a review of a nascent field, Build. Environ. 97 (2016) 196–202, https://doi.org/10.1016/j.buildenv.2015.12.001.
[31] S. HosseiniHaghighi, P.M.Á. De Uribarri, R. Padsala, U. Eicker, Characterizing and structuring urban GIS data for housing stock energy modelling and retrofitting, Energy Build. 256 (2022) 111706, https://doi.org/10.1016/j.enbuild.2021.111706.
[32] Z. Deng, Y. Chen, J. Yang, F. Causone, AutoBPS: a tool for urban building energy modeling to support energy efficiency improvement at city-scale, Energ. Buildings 282 (2023) 112794, https://doi.org/10.1016/j.enbuild.2023.112794.
[33] R. Nouvel, A. Mastrucci, U. Leopold, O. Baume, V. Coors, U. Eicker, Combining GIS-based statistical and engineering urban heat consumption models: towards a new framework for multi-scale policy support, Energy Build. 107 (2015) 204–212, https://doi.org/10.1016/j.enbuild.2015.08.021.
[34] T. Hong, Y. Chen, S.H. Lee, M.A. Piette, CityBES: a web-based platform to support city-scale building energy efficiency, (2016).
[35] D. Robinson, F. Haldi, J. Kämpf, P. Leroux, D. Perez, A. Rasheed, U. Wilke, Citysim: comprehensive micro-simulation of resource flows for sustainable urban planning, (2009).
[36] J.A. Fonseca, T.-A. Nguyen, A. Schlueter, F. Marechal, City energy analyst (CEA): integrated framework for analysis and optimization of building energy systems in neighborhoods and city districts, Energy Build. 113 (2016) 202–226, https://doi.org/10.1016/j.enbuild.2015.11.055.
[37] P. Remmen, M. Lauster, M. Mans, M. Fuchs, T. Osterhage, D. Müller, TEASER: an open tool for urban energy modelling of building stocks, (2017).
[38] C. Reinhart, T. Dogan, A. Jakubiec, T. Rakha, A. Sang, Umi – An Urban Simulation Environment For Building Energy Use, Daylighting And Walkability, in: 2013. https://doi.org/10.26868/25222708.2013.1404.

[39] Z. Lin, T. Hong, X. Xu, J. Chen, W. Wang, Evaluating energy retrofits of historic buildings in a university campus using an urban building energy model that considers uncertainties, Sustainable Cities Soc. 95 (2023) 104602.
[40] C. Song, Y. Chen, Energy-saving potential of cool roofs at the urban scale: a case study of xiamen city, Energy Build. 344 (2025) 116034, https://doi.org/10.1016/j.enbuild.2025.116034.
[41] J.A. Hartigan, M.A. Wong, Algorithm AS 136: a K-means clustering algorithm, J. r. Stat. Soc., C (appl. Stat.) 28 (1979) 100–108, https://doi.org/10.2307/2346830.
[42] M. Tan, Q.V. Le, EfficientNet: rethinking model scaling for convolutional neural networks, (2020). https://doi.org/10.48550/arXiv.1905.11946.
[43] Z. Wang, J. Yang, G. Li, C. Wu, R. Zhang, Y. Chen, Development of rooftop photovoltaic models to support urban building energy modeling, Appl. Energy 378 (2025) 124811, https://doi.org/10.1016/j.apenergy.2024.124811.
[44] J. Yang, Z. Wang, G. Li, C. Wu, F. Liu, Y. Chen, AutoBPS-carbon: a tool to estimate multi-year dynamic embodied and operational carbon emissions in the building sector, J. Build. Eng. 101 (2025) 111942, https://doi.org/10.1016/j.jobe.2025.111942.
[45] A. Xu, C. Song, W. Zhao, Y. Chen, Demand flexibility of pre-cooling strategies for city-scale buildings through urban building energy modeling, Buildings 15 (2025) 1051, https://doi.org/10.3390/buildings15071051.
[46] M. Wang, J. Zhou, Y. Liang, H. Yu, R. Jing, Climate change impacts on city-scale building energy performance based on GIS-informed urban building energy modelling, Sustain. Cities Soc. 125 (2025) 106331, https://doi.org/10.1016/j.scs.2025.106331.
[47] V.J.L. Gan, H.K. Wong, K.T. Tse, J.C.P. Cheng, I.M.C. Lo, C.M. Chan, Simulation-based evolutionary optimization for energy-efficient layout plan design of high-rise residential buildings, J. Cleaner Prod. 231 (2019) 1375–1388, https://doi.org/10.1016/j.jclepro.2019.05.324.
[48] S. Zhang, L. Chen, L. Xu, Z. Wang, GeoBEM: a geospatial computing empowered framework for urban-scale building energy modeling, Sustainable Cities Soc. 121 (2025) 106203, https://doi.org/10.1016/j.scs.2025.106203.
[49] W. Chung, I.M.H. Yeung, A study of energy consumption of secondary school buildings in Hong Kong, Energy Build. 226 (2020) 110388, https://doi.org/10.1016/j.enbuild.2020.110388.
[50] P. Xu, J. Huang, P. Shen, X. Ma, X. Gao, Q. Xu, H. Jiang, Y. Xiang, Commercial building energy use in six cities in southern China, Energy Policy 53 (2013) 76–89, https://doi.org/10.1016/j.enpol.2012.10.002.
[51] Y. Sheng, Z. Miao, J. Zhang, X. Lin, H. Ma, Energy consumption model and energy benchmarks of five-star hotels in China, Energy Build. 165 (2018) 286–292, https://doi.org/10.1016/j.enbuild.2018.01.031.
[52] Shenzhen Municipal Bureau of Housing and Construction, Shenzhen Public Building Energy Consumption Standards, (2017). https://zjj.sz.gov.cn/attachment/0/999/999571/9934368.pdf.
[53] J. Wang, W. Yang, Y. Zhang, T. Ren, S. Liang, Research on energy consumption evaluation and energy saving and carbon reduction measures for typical general hospitals in hot summer and warm winter regions, Energy Sustainable Dev. 79 (2024) 101381, https://doi.org/10.1016/j.esd.2024.101381.
[54] C. Yu, W. Pan, Inter-building effect on building energy consumption in high-density city contexts, Energy Build. 278 (2023) 112632, https://doi.org/10.1016/j.enbuild.2022.112632.
[55] Haikou Statistics Bureau, 2024 Haikou Bureau of Statistics Yearbook, 2024. https://tjj.haikou.gov.cn/hkstjj/tjyb/list.shtml.
[56] Y.-S. Kim, J. Srebric, Impact of occupancy rates on the building electricity consumption in commercial buildings, Energy Build. 138 (2017) 591–600, https://doi.org/10.1016/j.enbuild.2016.12.056.