buildings

MDPI

Article

An Operational Carbon Emission Prediction Model Based on Machine Learning Methods for Urban Residential Buildings in Guangzhou

Lintao Zheng [1], Kang Luo [2] and Lihua Zhao [2,*]

[1] School of Built Environment and Design, Guangdong Polytechnic of Water Resources and Electric Engineering, Guangzhou 510925, China; zhenglintao@gdsdxy.cn
[2] State Key Laboratory of Subtropical Building Science, School of Architecture, South China University of Technology, Guangzhou 510641, China
* Correspondence: lhzhao@scut.edu.cn; Tel.: +86-18819826874

Abstract: The carbon emissions of urban residential buildings are substantial. However, the standard operating conditions specified in current energy-saving standards significantly differ from the actual energy consumption under real operating conditions. Therefore, it is essential to consider the impact of residents' actual energy consumption behavior in carbon emission forecasts. To improve the accuracy of carbon emission predictions for urban residential buildings, this paper focuses on residential buildings in Guangzhou. Taking into account the energy consumption behavior of residents, parameterized modeling is carried out in the R language, and simulation is carried out using EnergyPlus software. Analysis revealed that the higher the comfort level of residential energy consumption behavior, the more it is necessary to encourage residents to adopt energy-saving behaviors. Combining carbon emission factors, air-conditioning energy efficiency, and the power consumption models of lighting and electrical equipment, a comprehensive operational carbon emission prediction model for urban residential operations in Guangzhou was developed. By comparing the prediction model with an actual case, it was found that the prediction deviation was only 4%, indicating high accuracy. The proposed operational carbon emission model can quickly assist designers in evaluating the carbon emissions of urban residential buildings in the early stages of design, providing an accurate basis for decision-making.

Keywords: operational carbon emission prediction; urban residential buildings; sensitivity analysis; energy consumption behavior; machine learning method

check for updates

Citation: Zheng, L.; Luo, K.; Zhao, L. An Operational Carbon Emission Prediction Model Based on Machine Learning Methods for Urban Residential Buildings in Guangzhou. *Buildings* **2024**, *14*, 3699. https://doi.org/10.3390/buildings14113699

Academic Editor: Pei Huang

Received: 24 October 2024
Revised: 7 November 2024
Accepted: 15 November 2024
Published: 20 November 2024

Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

With the rapid advancement of global urbanization, fuel consumption has significantly increased, leading to serious environmental issues and social concerns. Among all energy-consuming activities, the construction and maintenance of buildings account for over 40% of the world's total energy consumption and one-third of total carbon dioxide emissions [1]. Many countries have established carbon peaking and carbon neutrality goals based on their domestic conditions [2]. China has also steadily promoted various initiatives with carbon reduction as a key strategic direction, proposing a gradual shift from dual control of energy consumption to dual control of total carbon emissions and intensity. Urban and rural construction is one of the main sources of China's carbon emissions, with building energy consumption exceeding 25% of the country's total energy consumption [3]. Statistics show that in 2022, the energy consumption of urban residential buildings in China during the operation phase was 410 million tons of coal equivalent, accounting for 39% of the total energy consumption of national buildings during the operation phase. Corresponding carbon emissions were 900 million tons of CO_2, representing 42% of the total carbon emissions during the operation phase of national buildings [4]. Moreover, as

living standards continuously improve, people's requirements for living comfort are also increasing, leading to a continuous rise in total carbon emissions from building operations. Clearly, the energy-saving and carbon-reduction efforts for urban residential buildings have a long way to go.

Currently, carbon peaking and carbon neutrality in the construction sector have garnered global attention. However, most existing domestic forecasts for building carbon peaks are from a macro perspective, lacking in-depth consideration of individual energy consumption behaviors. Consequently, research results are often limited to an analysis of the overall construction sector [5–7]. This is particularly problematic given the significant variations in local carbon emissions and the heterogeneity of individual energy consumption behaviors. Existing research has yet to provide adequate support for the development of regional carbon peak and carbon neutrality pathways. By considering complex factors such as the physical characteristics of the building, climate models, and occupancy behavior, simulation can provide a comprehensive performance evaluation of the building automation and control system, which helps to predict energy efficiency savings more accurately [8]. Recent studies highlight the growing impact of individual energy consumption behavior on building energy use and carbon emissions [9–11], yet the focus remains on appliance usage, overlooking user behavior in the setting and operation of air conditioners. As a result, personal energy consumption patterns have not been fully integrated into existing carbon peak implementation plans.

When analyzing the energy-saving and emission-reducing potential of buildings, it is essential to conduct a detailed comparison of energy consumption and carbon emissions under different energy-saving measures. By establishing a building carbon emission prediction model, researchers can quickly obtain energy consumption and carbon emission data for various energy-saving programs, significantly reducing research costs and time consumption. Currently, various prediction methods such as multivariate linear regression, neural networks, support vector machines, and time series analysis are widely used. For example, multivariate linear regression can effectively predict long-term energy consumption trends by establishing a linear relationship model between multiple influencing factors and building energy consumption, making it suitable for large-scale building energy consumption analysis and prediction [12]. In contrast, neural networks are better at handling complex nonlinear problems, extracting patterns from historical data through adaptive learning, and can make short-term and high-precision energy consumption predictions [13,14]. Support vector machines perform well in minimizing both training error and the upper limit of the generalization error, making them suitable for high-precision building energy consumption predictions, such as predicting the cooling load of commercial buildings [15,16]. Time series methods can effectively predict short-term fluctuations in building energy consumption by analyzing the temporal patterns of historical data, particularly suitable for predicting and optimizing the energy consumption of single buildings [17,18]. Meanwhile, supervised machine learning models offer more nuanced predictions by accounting for multifactorial influences such as weather, production rates, and equipment operations [19]. Transfer learning, a relatively new approach, enhances model performance by applying knowledge from related fields [20]. These techniques collectively improve forecasting accuracy, providing essential insights for developing targeted strategies to enhance energy efficiency and reduce emissions in the building sector.

At present, although existing research has provided substantial theoretical and technical support for building energy conservation and emission reduction, there is still a need for in-depth studies on the actual operation of different regions and building types. This is particularly important given the significant variations in local carbon emissions and the heterogeneity in individual energy consumption behaviors. Such research is essential to formulate more targeted and effective carbon peaking and carbon neutrality pathways. Research on the factors influencing carbon emissions from building operations can be divided into macro-level and micro-level studies. At the macro level, key factors include residents' average wage, population density, electricity intensity, educational attainment,

and per capita household liquefied petroleum gas consumption in urban areas [21]. At the micro level, considerations include operation time, start-up time of the HVAC system, setpoint temperature level, geometry characterization, and building envelopes properties [22]. However, there are relatively few studies on the influence of indoor occupants' energy consumption behavior on residential energy consumption. Most building energy consumption predictions focus on the objective conditions of the buildings themselves, with fewer studies comprehensively considering subjective factors such as occupants' energy consumption behavior. This paper addresses this gap by establishing a carbon emission prediction model for urban residential buildings in Guangzhou based on the different energy-using behaviors of residents. In developing this model, careful selection of input parameters is essential due to the many influencing factors. Common approaches, such as parameterized methods or probability distribution functions combined with Monte Carlo or Latin hypercube sampling, generate extensive datasets for simulation. The Standardized Regression Coefficient (SRC) and Partial Correlation Coefficient (PCC) can then assess each parameter's influence, helping to isolate the most critical factors for prediction accuracy [22]. Recently, feature selection methods like Elastic Net have proven effective for optimizing prediction models by reducing redundant features and enhancing accuracy, particularly in building cooling load predictions [23,24]. This study uses SRC to filter key features [25], subsequently applying them as inputs to a machine learning model that predicts urban residential carbon emissions under varying energy-saving scenarios [26]. This model also incorporates changes in residents' energy-use behavior, providing insights into potential carbon reduction pathways, including carbon peak scenarios and strategies for achieving carbon neutrality.

The structure of this article is arranged as follows: first, it introduces the influencing factors and current situation of carbon emissions in the construction field, followed by an analysis of the shortcomings and challenges in current research. It then details the research methods and data sources, showcasing the process and technical details of model establishment. Finally, the experimental results are summarized and analyzed, and suggestions and policy recommendations for future research are put forward, aiming to contribute new ideas and methods to the planning of the carbon peaking and carbon neutrality pathways for urban residential buildings in China.

2. Materials and Methods

As the capital of Guangdong Province, Guangzhou is a central city in the Guangdong–Hong Kong–Macao Greater Bay Area. It has a high urbanization rate and a dense population, resulting in a large urban residential construction area. Consequently, its operational carbon emissions are significant and cannot be overlooked. Moreover, Guangzhou is located in a region with hot summers and mild winters, characterized by long summers and no real winters. As people's demands for indoor thermal comfort increase, energy consumption for comfort is becoming more prevalent. Therefore, the task of building energy conservation and emission reduction in Guangzhou is challenging, necessitating urgent action to achieve peak carbon emissions from building operations.

The purpose of this paper is to develop carbon emission prediction models for typical urban residential operations with different energy consumption behaviors. The aim is to enable rapid and accurate predictions of carbon emissions from urban residential operations and to assist policymakers and designers in achieving precise carbon emission forecasts at early stages. This will facilitate quick evaluations of the carbon emission characteristics of urban residential operations in Guangzhou in response to changes in residents' energy consumption behaviors in the future. To achieve these goals, data on energy-use behavior was collected through online questionnaires and on-site interviews. A total of 300 questionnaires were recovered, including 254 online questionnaires and 46 interview questionnaires, with 289 valid responses, resulting in an efficiency rate of 96.3%. The typical building model was based on data from 200 residential settlements across 11 administrative districts in Guangzhou. Basic data such as building layout, unit form, number of buildings, floor

area ratio, and outdoor site layout were collected and analyzed. The research process of this study is shown in Figure 1, and the process can be divided into four steps:

1. Analyzing typical energy consumption behavior: through questionnaires, typical energy consumption behaviors of urban residents in Guangzhou were analyzed, including indoor stay and rest, air conditioning use, lighting, and electrical appliance use. And typical energy consumption behavior patterns were extracted;
2. Parameterized modeling: using the R language [27], typical urban dwellings in Guangzhou were parameterized, taking into account geometric characteristics, internal heat sources, and envelope structures affecting energy consumption. The Latin hypercube sampling method was used to generate multiple sets of parameter combinations [28];
3. Energy consumption simulation: EnergyPlus software was used to simulate and calculate the energy consumption of these parameterized models, obtaining residential energy consumption data under different scenarios [29]. Key influencing factors were determined through SRC sensitivity analysis [30];
4. Carbon emission intensity prediction model: based on the typical established energy consumption behavior model and energy consumption model, a carbon emission intensity prediction model for typical urban residential operations in Guangzhou was developed. Different machine learning methods were screened to establish the most effective model.

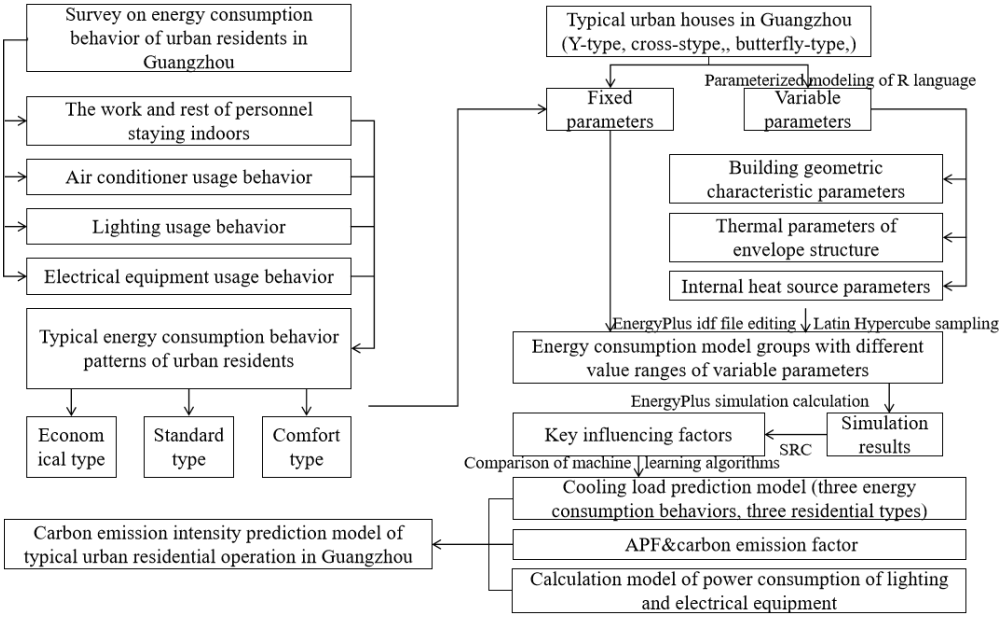


Figure 1. Technology roadmap.

This structured approach aims to provide a scientific basis for achieving carbon peaking and carbon neutrality in urban residential buildings in Guangzhou, offering new ideas and methods for the planning and implementation of energy-saving and emission-reducing strategies. The tools used in each process are described below:

1. Data collection tool—questionnaire survey: through a questionnaire survey, the energy consumption behavior of urban residents in Guangzhou was studied. The survey results were used to refine and summarize the typical energy consumption behavior patterns of these residents;
2. Drawing analysis tool—statistical method: resident information, building information, and especially the lifestyle of residents in the collected samples were analyzed. The urban residents of Guangzhou were divided into two groups: intermittent users and continuous users. The characteristics and differences in residential indoor stay,

rest, and air-conditioning use behavior of these two groups were analyzed separately. Their air-conditioning use behavior was clustered and analyzed, and typical air-conditioning use behavior patterns were refined and summarized. Combining the typical indoor stay, rest, and typical lighting and electrical equipment use behaviors obtained from the survey, three typical energy-use behavior patterns for intermittent users and two for continuous users were identified;

3. Modeling tool—parameter modeling: the LHS package in the R language was used for Latin hypercube sampling. Generally, the number of samples should be more than 10 times the number of variables to ensure the rationality of the sampling results. To obtain enough samples, this study set the number of samples to 100 times the number of variables. The sampling results were output to the edited idf file through the code, resulting in the final residential energy consumption calculation model;
4. Sensitivity analysis tool—SRC: by calculating the value of SRC, the key influencing variable parameters of residential energy consumption for the two groups of intermittent users and continuous users under different energy consumption behavior patterns were determined;
5. Simulation calculation tool—EnergyPlus: this study used the R language for parameterized modeling and EnergyPlus to simulate and calculate the energy consumption model. This approach provided a large number of energy consumption values, creating conditions for sensitivity analysis and the establishment of energy consumption prediction models;
6. Data analysis tool—generalized linear model: building energy consumption prediction is usually a linear problem. Therefore, different machine learning methods were introduced and compared to determine the best algorithm for the building carbon emission prediction model. This model was established to predict carbon emissions from urban residential operations in Guangzhou, accommodating the research of multiple variables affecting residential energy consumption in this study.

3. Results

3.1. Typical Energy Consumption Behavior Patterns of Urban Residential Residents in Guangzhou

The energy consumption and operational carbon emissions of residential buildings are influenced by various factors such as the enclosure structure, internal heat sources, occupant energy consumption behavior, and building geometric characteristics. Among these, the energy consumption methods and work–rest habits of indoor occupants are critical factors leading to differences in energy consumption and carbon emissions. To improve the accuracy and reliability of predictions regarding energy consumption and carbon emissions in residential buildings, it is essential to analyze residents' energy consumption behaviors and summarize their typical patterns. Therefore, this study investigated the energy consumption behavior of urban residential residents in Guangzhou.

First, a questionnaire on energy use behavior was designed according to the research objectives, and data were collected through online platforms and on-site interviews. A total of 300 questionnaires were distributed. The survey content primarily included residents' indoor stay and rest, air-conditioning usage and settings, lighting usage, and electrical equipment usage, as shown in Table 1.

Table 1. Questionnaire on energy consumption behavior of urban residents in Guangzhou.

Survey Items	Main Content
Basic resident information	Gender, age, number of permanent residents in the family, number of elderly people living together
Basic building information	Administrative district, residential type, apartment type
Lifestyle of residents	Indoor air conditioning work and rest, air conditioning work and rest and temperature setting, electrical equipment and lighting work and rest

The survey results indicate that the types of urban residential buildings in Guangzhou are predominantly middle- and high-rise (7–10 floors) and high-rise residential buildings (11–30 floors). The residential apartment types are mainly two-bedroom and three-bedroom units, with three-bedroom units being more common. Households of urban residential residents in Guangzhou typically consist of three-person and four-person households. The proportion of households without elderly members is close to that of households with elderly members, and in those with elderly members, there are generally one or two elderly individuals. Typical indoor stay and rest patterns for urban residents in Guangzhou are as follows: on weekdays, residents usually leave home around 8 a.m., return home around 7 p.m., and go to bed around 11 p.m. On holidays, residents typically get up around 9 a.m., stay at home for most of the day, and go to bed around 11 pm.

Analysis of the air-conditioning usage behavior of urban residents reveals significant differences in usage patterns and a diversity of air-conditioning schedules and temperature settings within the group. These differences impact the overall building energy consumption and operational carbon emissions. To provide representative input conditions for the prediction model, it is necessary to cluster and analyze these air-conditioning usage patterns and temperature settings and summarize the typical air-conditioning behavior patterns.

In this study, second-order clustering analysis is performed using log likelihood as the distance measure and the Schwarz Bayesian Criterion as the clustering criterion. After clustering, the most common air-conditioning usage pattern for each type of schedule is selected as the typical air-conditioning behavior. The typical air-conditioning usage characteristics and cooling temperatures are summarized, as shown in Table 2. The air-conditioning usage behavior patterns are categorized into three types: economical, standard, and comfortable.

Table 2. Three types of typical air-conditioning work and rest of personnel and air-conditioning set temperature.

Type	Is the Bedroom Heated?	Is the Living Room Heated?	Is the Bedroom Cooling/ Temperature?	Is the Living Room Cooling/ Temperature?
Economical	no	no	often/28 °C	no
Standard	no	no	often/26 °C	occasionally/26 °C
Comfort	occasionally	occasionally	often/24 °C	often/24 °C

Combining the above-mentioned typical indoor stay and rest patterns with the provisions of the latest energy-saving standards [31] regarding the hourly occupancy rate of personnel in each room of a residential building, we derived the typical hourly occupancy rate, typical lighting usage time, and typical electrical equipment usage rate and indoor air conditioning temperature for urban residential residents in Guangzhou, as shown in Figure 2. There are three types of air-conditioning usage and cooling temperature settings: economical, standard, and comfortable.

According to the research and analysis results, the cooling time in bedrooms generally lasts all night, while air-conditioning in living rooms is used predominantly when people are present and it feels hot. Most bedrooms are heated for 1 to 3 h before bedtime, averaging a 2 h heating duration. The preferred heating temperature is between 23 °C and 25 °C, with 24 °C being the most desirable. For the vast majority of elderly residents, bedroom air-conditioning is set for up to 3 h, typically around 2 h, with the consensus being that it is only needed when it feels hot. The air conditioner in the living room is turned on only when it is occupied. In summary, the typical air-conditioning usage behavior patterns for the available personnel are shown in Figure 2g,h. It can be seen that according to the research results, this study does not consider heating energy consumption.

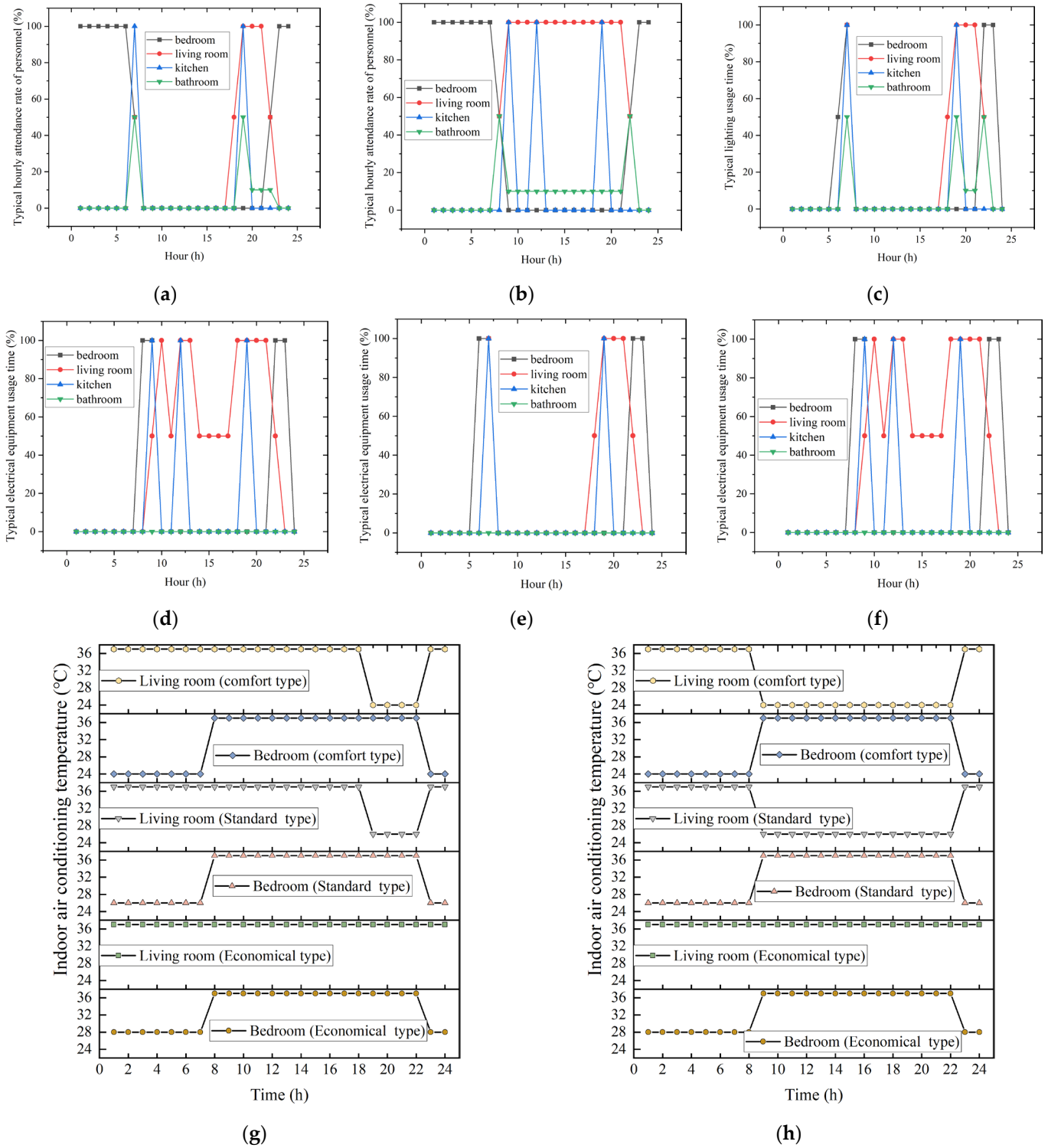
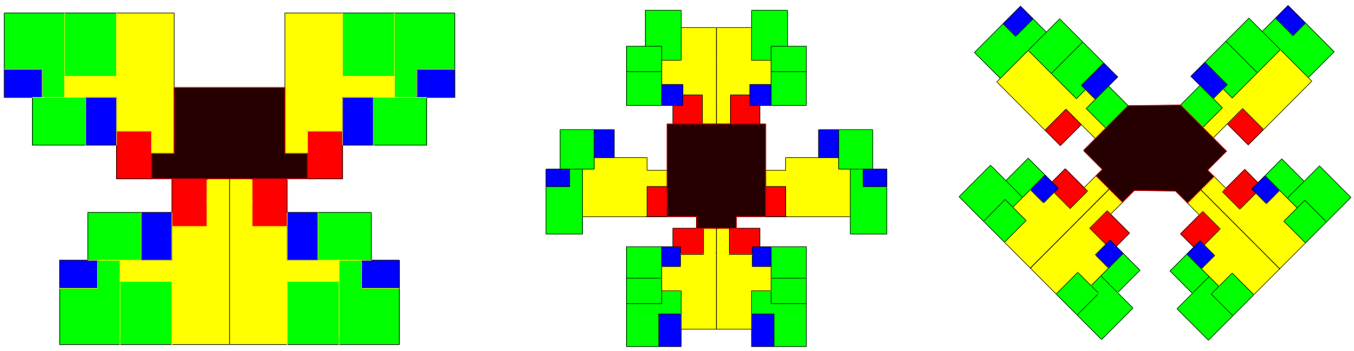


Figure 2. Input condition settings for three typical energy consumption modes, including the following: (**a**) weekday staff attendance hours, (**b**) holiday staff attendance hours, (**c**) weekday lighting usage duration, (**d**) holiday lighting usage duration, (**e**) weekday electrical equipment operation time, (**f**) holiday electrical equipment operation time, (**g**) weekday indoor air conditioning temperature settings, and (**h**) holiday indoor air conditioning temperature settings.

3.2. Carbon Emission Prediction Model for Typical Urban Residential Operations in Guangzhou

3.2.1. Typical Building Types of Urban Dwellings in Guangzhou

Based on the research results of 200 residential districts, urban residential buildings in Guangzhou can be categorized into three main types: Y-type, cross-type, and butterfly-type. This study uses these three types as the standard floor plan layouts, depicted in Figure 3. In the figure, yellow represents the living room, green represents the bedroom, blue represents the bathroom, red represents the kitchen, and black represents the elevator hall. The Y-type standard floor covers an area of 515.92 m^2, the cross-type standard floor covers 453.7 m^2, and the butterfly-type standard floor covers 658.4 m^2.



(**a**) (**b**) (**c**)

Figure 3. Typical plan types of urban residential buildings in Guangzhou, including the following: Y-type (**a**), cross-type (**b**), and butterfly-type (**c**).

3.2.2. Model Parameter Setting

The energy consumption of buildings is influenced by four main types of factors: the geometric characteristics of the building, the thermal properties of the building envelope, the internal heat sources, and the indoor air design parameters. For the Guangzhou area, the indoor air design parameters primarily refer to the set temperature for air conditioning used by residents. According to Table 2, these parameters are categorized into three types: economical, standard, and comfortable.

Regarding the geometric characteristics, the standard floor plan layout of the building is fixed, and the typical residential building height is approximately 3 m. Therefore, the geometric factors affecting the building load include the number of floors, the building orientation, and the window-to-wall area ratio in air-conditioned rooms such as bedrooms and living rooms.

The heat dissipation from occupants, a component of the internal heat source parameters, is set based on the literature [32]. The range of variability for these influencing factors is determined based on relevant standards [31–33] and research data, as shown in Table 3.

Table 3. Parameter setting of typical urban residential building model.

Category	Symbol	Value
the geometric characteristics of the building	NF	18~33
	OR	−180~180
	BR/LR	The minimum value is calculated and determined according to the window-to-wall area ratio of each room not being less than 1/7, and the maximum value is determined according to the limit of the window-to-wall area ratio of each room in the literature [31].

Table 3. *Cont.*

Category	Symbol	Value
the thermal properties of the building envelope	WK	0.3~1.5
	RK	0.25~0.4
	CK	2.5~3.5
	SHGC	0.15~0.35
	WA	0.7~0.8
	RA	0.7~0.8
the internal heat sources	OC	1 or 2
	LI	5~6
	EQ	3.8~6
	PH	Bedroom: 108 Living room: 134 Kitchen: 235
the indoor air design parameters	T	According to Table 2
	V	1.0

3.2.3. Model Group Generation and Load Simulation Results

The fixed parameters needed for the simulation model are initially configured in EnergyPlus to create the idf file. Using the Latin hypercube sampling method in the R language, random combinations of simulation models for Y-type, cross-type, and butterfly-type houses are generated. These combinations are created within the range of variable parameter values and account for three different energy-use behaviors: economical, standard, and comfortable.

This process produces 2100 Y-type residential simulation models, totaling 6300 models, and 3600 cross-type and butterfly-type residential simulation models, totaling 10,800 models. The results of the load simulations conducted using EnergyPlus 24.1.0 for these generated models are shown in Figure 4.

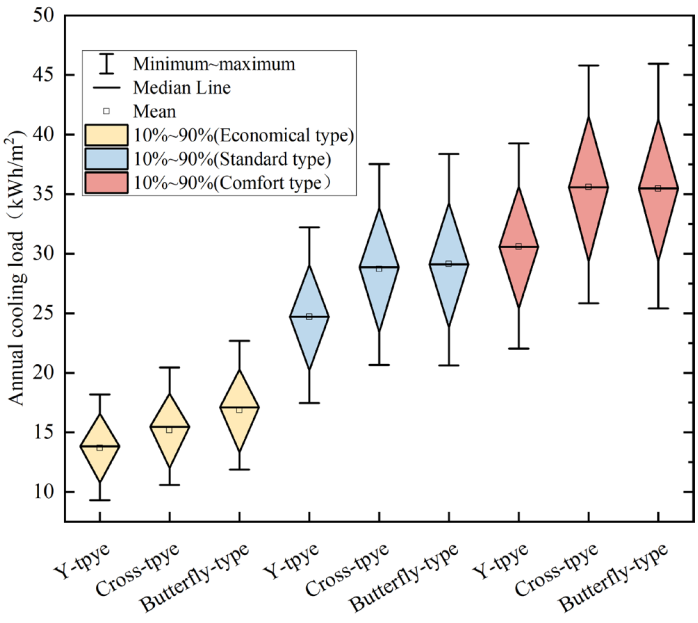


Figure 4. Simulation results of annual cooling load of three typical urban residential buildings under three different energy consumption behaviors.

The abscissa represents three different typical urban residential types in Guangzhou, and the ordinate is the annual cooling load value of each model under different energy consumption behaviors. The results show that the energy consumption behavior of resi-

dents has a greater impact on the annual cooling load of residential buildings. The annual cooling load of economical residential buildings is the lowest, with an average value of around 15 kWh/m^2, followed by the standard type between 25 and 30 kWh/m^2. The annual cooling load of residential buildings under the comfortable energy consumption behavior is the highest, with an average value of 30 kWh/m^2 and above. The economical cooling load is very different from the standard type and the comfortable type. The standard type is 1.6 times that of the economical type, and the comfortable type is 2 times that of the economical type. Economical users only turn on air conditioning and cooling in the bedroom in summer, while standard and comfortable users also cool living rooms in summer. Energy behavior has a huge impact on operating energy consumption, and ignoring the prediction of carbon emissions by energy consumption behavior may cause large errors.

Under the economical energy consumption behavior, the variable parameters of the model have the greatest impact on its annual cooling load, reaching about 65%, about 60% for the standard type, and 56% for the comfortable type. It can be seen that if you want to reduce residential energy consumption by optimizing the thermal performance of the envelope, the geometric characteristic parameters of the building, and other factors (in addition to energy consumption behavior), the effect on economical energy-using residents is relatively better. In addition, under the same energy consumption behavior model, the Y-type residential buildings have the lowest annual cooling load, and the cross-type and butterfly-type residential buildings have the same level. This is related to the type of residential floor layout. The standard floor area of Y-type residential buildings ranks second among these three types, but it can only accommodate four households. Compared with the six-family cross-type and butterfly-type residential buildings, the total energy demand is less, so the annual cooling load level is the lowest.

3.3. Sensitivity Analysis of Parameters Affecting Residential Load

To further clarify the key factors influencing the annual cooling load of urban dwellings, a global sensitivity analysis was conducted on the above load simulation results and corresponding variable parameters using the SRC value for evaluation. Figure 5 presents the sensitivity analysis results of the annual cooling load and variables for three typical urban residential models under different energy consumption behavior patterns. The ordinate represents the variable parameters of the input parameterized model, while the abscissa displays the SRC sensitivity analysis result value for each variable parameter.

Overall, the results indicate that BR and LR have minimal impact on the annual cooling load of residential buildings. This is because their variable range is relatively small. To meet the natural lighting requirements of residential buildings, regulations stipulate that the area ratio of the room to the window shall not be less than 1/7. The current energy-saving design standards in Guangzhou limit the window–wall area ratio to a maximum of 0.4 in each direction, with only one room per residential unit allowed with a window–wall area ratio not exceeding 0.6. Therefore, for newly built urban houses in Guangzhou, under the current energy-saving standards, the energy-saving potential of the room window–wall area ratio is very limited.

According to the sensitivity ranking in Figure 5, the primary factors impacting the annual cooling load of residences under the three energy consumption behaviors are OC, LI, EQ, and the thermal properties of the exterior walls and windows (including WK, WA, and SHGC). Among these, the number of people in the room has the greatest impact on the cooling load. This is because human metabolism generates 108 W of heat even when at rest. Consequently, the more people in a room, the greater the cooling load needed to reach the set temperature during summer.

Economical users typically do not use air conditioning in the living room during summer. As a result, the SHGC sensitivity of living room windows ranks higher in the economical type compared to the standard and comfort types, though still higher than that of the bedroom. This is due to the large area of living room window walls. During summer

days, a significant amount of heat enters the living room through these windows, which then transfers through partition walls into adjacent bedrooms, affecting the bedroom's cooling load at night.

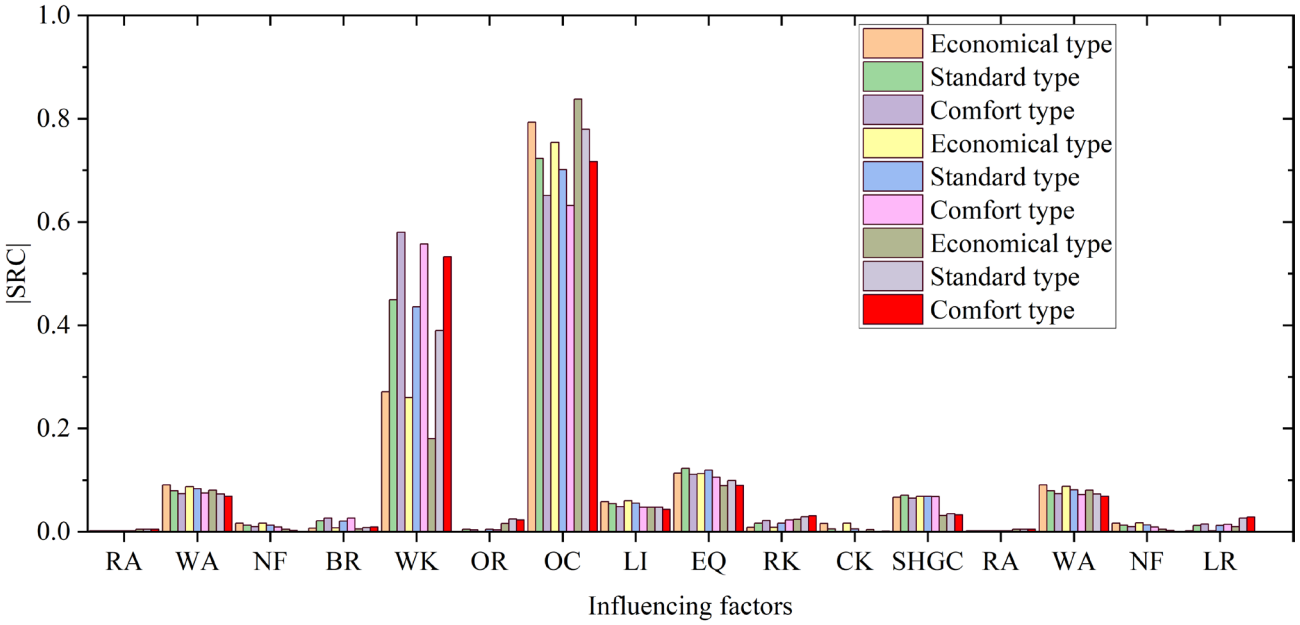


Figure 5. Sensitivity analysis of annual cooling load and variables of residential buildings under different energy consumption behaviors.

Cross-type houses differ from the other two types of houses. The SHGC of bedroom and living room windows has a minimal impact on the annual cooling load. Even under standard and comfort energy consumption behaviors, only some living room windows show greater sensitivity to SHGC. This indicates that for cross-type houses, the primary influencing factors for the annual cooling load are the disturbances within the room and the thermal performance of the exterior walls, with the energy-saving potential of windows being relatively low.

In summary, the main factors affecting the annual cooling load of urban dwellings in Guangzhou are internal disturbances, the thermal performance of exterior walls, and SHGC. Among these, OC is the most significant factor, followed by WK. For Y-type and butterfly-type homes, special attention should be given to the SHGC of all bedroom and living room windows.

3.4. Establishment of Carbon Emission Prediction Model for Urban Residential Operation

The carbon emissions of urban residential operations are mainly composed of emissions from air-conditioning systems, lighting, and electrical equipment. The carbon emissions from lighting and electrical equipment can be directly calculated based on power density, corresponding usage patterns, and carbon emission factors. Therefore, determining the carbon emissions from air-conditioning operation is the most critical step in predicting the overall carbon emissions from residential operations.

Based on the aforementioned cooling load sensitivity analysis, the key variables for the cooling load prediction model are selected. Appropriate algorithms are then employed to predict the cooling load of air conditioners. Given the common use of split-type air conditioners in residential buildings, their energy consumption is calculated using the Annual Performance Factor (APF) to obtain the carbon emissions of the air conditioners.

In this study, nine common machine learning models were selected. Datasets consisting of 2100 groups, 6300 groups, and 3600 groups generated by Y-type, cross-type, and butterfly-type houses, respectively, were used as input data. The k-fold cross-validation

method (k = 10) was utilized to calculate and predict cooling load. By comparing the generalization capabilities of different algorithms, the best prediction method for air-conditioning carbon emissions was determined. As shown in Table 4, the analysis found that the Gradient Boosting Regression Tree, Gradient Boosting Nearest Neighbors, and gradient boosting algorithms provided the best prediction results, with R^2 values higher than 0.98. This is significantly better than the prediction accuracy of linear regression and support vector machines. Gradient boosting algorithms typically outperform others in energy consumption simulation tasks due to their ability to handle complex nonlinear relationships, robustness to missing data and noise, and faster training speed, offering better interpretability and generalization. While neural networks and support vector machines excel in specific tasks, tree-based models often deliver superior efficiency and performance in practical applications like energy simulation, requiring less feature engineering and parameter tuning, and being less sensitive to noise. The prediction results for the three building layouts under three different energy-use patterns were consistent. Therefore, any of these three algorithms can be selected as a prediction model for the carbon emissions of air-conditioning operations.

Table 4. R^2 for predicting total energy consumption per unit area.

	Economical			**Standard**			**Comfort**		
	Y-Type	**Cross-Type**	**Butterfly-Type**	**Y-Type**	**Cross-Type**	**Butterfly-Type**	**Y-Type**	**Cross-Type**	**Butterfly-Type**
Linear Regression	0.779	0.759	0.765	0.818	0.789	0.793	0.851	0.820	0.828
Decision Tree	0.950	0.946	0.955	0.925	0.954	0.964	0.941	0.950	0.961
Random Forest	0.974	0.978	0.976	0.970	0.977	0.978	0.972	0.977	0.980
Support Vector Machine	0.731	0.734	0.747	0.746	0.751	0.769	0.774	0.781	0.804
Ridge Regression	0.775	0.758	0.762	0.814	0.788	0.791	0.848	0.819	0.826
Gradient Boosting	0.980	0.983	0.981	0.980	0.983	0.981	0.982	0.985	0.984
Extreme Random Trees	0.972	0.978	0.978	0.971	0.976	0.979	0.973	0.975	0.981
Gradient Boosting Nearest Neighbors	0.980	0.983	0.981	0.980	0.983	0.981	0.982	0.985	0.984
Gradient Boosting Regression Tree	0.980	0.983	0.981	0.980	0.983	0.981	0.982	0.985	0.984

Combined with the cooling load prediction model established above, and considering the power consumption of lighting and electrical equipment, a final model of carbon emissions from urban residential operations can be established, as shown in Equations (1)–(4).

$$CE = (Q_c \,/\, \mathrm{APF} + Q_l + Q_e) \times N \tag{1}$$

$$Q_c = f\,(OC, WK, EQ, WA, SHGC, LI) \tag{2}$$

$$Q_l = LI \times 1.045 \tag{3}$$

$$Q_e = LI \times 1.460 \tag{4}$$

where CE is the carbon emission from building operation, in $kgCO_2/(m^2{\cdot}a)$. APF is the annual operating energy efficiency of air-conditioning equipment. N represents the carbon emission factor of the power grid, in $kgCO_2/kWh$. Q_c is the annual cooling load of the building, in $kW/(m^2{\cdot}a)$. Q_l is the annual power consumption of building lighting, in $kW/(m^2{\cdot}a)$. Q_e is the annual power consumption of building electrical equipment, in $kW/(m^2{\cdot}a)$.

To verify the accuracy of the prediction model, the measured data of the carbon emissions from the operation of a residential building in Guangzhou are compared with the predicted values of the model. The information indicators of the residential building are shown in Table 5. Judging from the family structure and air-conditioning usage behavior, the energy consumption behavior pattern of the residents is close to the standard energy

consumption behavior pattern. Therefore, the prediction model under the standard energy consumption behavior pattern for Y-type residential buildings is used for calculation.

Table 5. Residential information.

Information Indicators	Indicator Value
Household population	3 people are the main
Family structure	Mainly urban white-collar workers
WK	1.38 W/(m^2·K)
LI	7.0 W/m^2
EQ	3.8 W/m^2
WA	0.48
SHGC	0.45
Air conditioner cooling set temperature	26 °C
Bedroom air conditioning on time	22:00~6:00
Living room air conditioning on time	17:00~22:00
Annual cooling load	29.40 kWh/m^2

The average annual air-conditioning cooling load of the residential building is 28.54 kWh/m^2, which is lower compared with the measured data of 29.40 kWh/m^2 in the table. The error is 4.0%, which is less than 10% and thus within the acceptable range. This demonstrates that the accuracy of the prediction model is good and also proves that the carbon emission prediction model of building operation proposed in this study has sufficient accuracy to assist designers and policymakers in program planning and decision-making.

4. Discussion

At present, rapid assessment tools for the carbon emissions of residential building operations exist, but they are often limited by the varied energy consumption characteristics of different residential buildings. This variability makes it challenging to predict carbon emissions simply and directly. To address this problem, it is essential to first determine the energy consumption characteristics of residential buildings through research and then classify them accordingly. Using machine learning methods, parameterized models, and the energy consumption simulation software EnergyPlus, an annual cooling load prediction model for urban residential buildings can be developed. By integrating air conditioning energy efficiency, lighting, and electrical equipment power consumption, a comprehensive carbon emission prediction model for urban residential operations is established. This model provides rapid evaluation tools for designers and policymakers, aiding in the transformation of old city residential buildings and the construction of new districts to support informed decision-making. However, there are still some limitations that need to be addressed in this model:

1. Building types: this study focuses solely on high-rise tower houses as a typical urban residential building type in Guangzhou. The model's applicability is limited due to this single building type. Future research can expand to include various building types such as low-rise buildings, multi-story buildings, and slab buildings;
2. Survey methodology: the energy consumption behavior survey of urban residents in Guangzhou faced certain challenges, primarily relying on online questionnaires for data collection. Supplementing this with follow-up data from summer field research can provide a more accurate energy consumption behavior pattern for Guangzhou urban residents;
3. Geographical scope: the carbon emission prediction model developed in this study is specific to the Guangzhou area. To make it applicable to other regions, further research can be conducted to include different climates. An in-depth analysis based on local climate conditions can offer relevant guidance for residential building planning in various climate zones.

By addressing these limitations, the model can be refined and its applicability broadened, thereby enhancing its utility for a wider range of residential building types and regions.

5. Conclusions

The establishment of a carbon emission prediction model for urban residential operations in Guangzhou involves the following three steps:

First of all, a questionnaire survey was conducted on the energy consumption behavior of urban residents in Guangzhou City, analyzing the characteristics and differences of people's indoor stay and rest, air-conditioning use behavior, and clustering analysis of their air-conditioning use behavior to obtain typical air-conditioning use behavior patterns. Combined with the investigation of typical indoor stay and rest, typical lighting, and electrical equipment use behavior, three typical energy-use behavior patterns of personnel were refined: economical, standard, and comfortable.

Secondly, the parameterized models of three typical urban dwellings in Guangzhou are established through the R language, and the various variables that affect the energy consumption of residential operations are entered into the parameterized model. Latin hypercube sampling is performed to randomly generate a large number of energy consumption calculation models. EnergyPlus is used to perform batch energy consumption simulation of the residential energy consumption calculation model, and the global sensitivity analysis method is used to determine the key influencing variables of residential energy consumption of personnel under different energy consumption behavior patterns. Regardless of the type of group, residence, and energy consumption behavior pattern, the number of rooms and the heat transfer coefficient of the exterior wall are the two parameters that have the greatest impact on the energy consumption of residential operations. Other more sensitive variables are equipment power density, lighting power density, solar radiation absorption coefficient, and the solar heat gain coefficient of the outer windows of the bedroom and living room. It can be seen that in order to reduce the operating energy consumption and carbon emissions of residential buildings, in addition to the energy consumption behavior of residents, urban residential buildings in Guangzhou also need to focus on indoor thermal disturbance, exterior wall thermal performance, and the sun's heat coefficient of the outer windows of air-conditioned rooms. For the heat transfer coefficient of the outer windows, the thermal performance of the roof and the area ratio of the window and wall, etc., it is enough to meet the mandatory requirements of the current energy-saving standards.

Finally, a database of residential air conditioning load samples of three layout types was established separately, and K-fold cross-verification was used for segmentation. The learning and modeling performance of nine common machine learning algorithms were compared. As a result, three prediction algorithms, Gradient Boosting Regression Tree, Gradient Boosting Nearest Neighbors, and gradient boosting algorithms, can be used as prediction models for the cooling load of urban residential air conditioners in Guangzhou. After combining carbon emission factors, air conditioning energy efficiency APF, lighting energy consumption, and electrical equipment energy consumption, a carbon emission prediction model for urban residential buildings in Guangzhou was finally established and compared with actual residential cases. The deviation is only 4.0%, proving that the carbon emissions of urban residential buildings in Guangzhou can be predicted very accurately.

In addition, through analysis, when other variables are the same, residents' energy consumption behavior changes from comfortable to standard and economical: the carbon emission intensity of Y-type residential operation is reduced by 17.5% and 52.4%, respectively; cross-type residential is reduced by 18.3% and 56.2%, respectively; butterfly-type residential is reduced by 16.7% and 50.9%, respectively. This indicates that in the process of realizing the carbon emission control of urban residential operations, focusing on guiding residents' energy consumption behavior is also a very effective method.

Author Contributions: All authors contributed to the study conception and design. Material preparation, data collection and analysis were performed by L.Z. (Lintao Zheng) and K.L. The first draft of the manuscript was written by L.Z. (Lintao Zheng) and all authors commented on previous versions of the manuscript. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by SZ02 Architectural Engineering Technology Specialty Group of Guangdong Polytechnic of Water Resources and Electric Engineering.

Data Availability Statement: Data is available on request due to privacy restrictions.

Conflicts of Interest: The authors declare no conflict of interest.

Nomenclature

SRC	Standardized regression coefficient
PCC	Partial correlation coefficient
NF	Number of floors (floors)
OR	Building orientation (°)
BR/LR	The window–wall area ratio of bedroom/living room
WK	Heat transfer coefficient of exterior wall ($W \cdot m^{-2} \cdot K^{-1}$)
RK	Heat transfer coefficient of roof ($W \cdot m^{-2} \cdot K^{-1}$)
CK	Heat transfer coefficient of exterior window ($W \cdot m^{-2} \cdot K^{-1}$)
SHGC	Solar heat gain coefficient
WA	Solar radiation absorption coefficient of exterior wall
RA	Solar radiation absorption coefficient of roof
OC	Number of people in the room
LI	Lighting power density ($W \cdot m^{-2}$)
EQ	Equipment power density ($W \cdot m^{-2}$)
PH	Personnel heat dissipation ($person^{-1}$)
T	Air conditioner cooling set temperature (°C)
V	Number of ventilations (h^{-1})
CE	The carbon emission from building operation ($kgCO_2/(m^2 \cdot a)$)
APF	Annual operating energy efficiency of air-conditioning equipment
N	The carbon emission factor of the power grid ($kgCO_2/kWh$)
Q_c	Annual cooling load of the building ($kW/(m^2 \cdot a)$)
Q_l	Annual power consumption of building lighting ($kW/(m^2 \cdot a)$)
Q_e	Annual power consumption of building electrical equipment ($kW/(m^2 \cdot a)$)

References

1. Liu, Z.; Fan, G.; Sun, D.; Wu, D.; Guo, J.; Zhang, S.; Yang, X.; Lin, X.; Ai, L. An innovative distributed energy system that combines hybrid energy storage with a multi-objective optimization approach for nearly zero-energy communities and buildings. *Energy* **2022**, *239*, 122577. [CrossRef]
2. Wang, S.J. Corporate Green Development Strategy Under the Carbon Peaking and Carbon Neutrality Goals. *Learn. Educ.* **2022**, *10*, 105–106. [CrossRef]
3. Du, X. Thoughts on strategies and paths to achieve carbon peaking and carbon neutrality in China. *Front. Energy* **2023**, *17*, 324–331. [CrossRef]
4. China Building Energy Conservation Association Building Energy Consumption and Carbon Emission Data Professional Committee. A Series of Research Reports on Carbon Emissions in the Field of Urban and Rural Construction in China in 2022. Available online: https://www.cabee.org/upload/file/20230104/1672820934145324.pdf (accessed on 2 January 2024).
5. Tan, X.C.; Lai, H.P.; Gu, B.H.; Zeng, Y.; Li, H. Carbon emission and abatement potential outlook in China's building sector through 2050. *Energy Policy* **2018**, *118*, 429–439. [CrossRef]
6. Zhou, N.; Khanna, N.; Feng, W.; Ke, J.; Levine, M. Scenarios of energy efficiency and CO_2 emissions reduction potential in the buildings sector in China to year 2050. *Nat. Energy* **2018**, *3*, 978–984. [CrossRef]
7. Yang, T.; Pan, Y.Q.; Yang, Y.Q.; Lin, M.; Qin, B.; Xu, P.; Huang, Z. CO_2 emissions in China's building sector through 2050: A scenario analysis based on a bottom-up model. *Energy* **2017**, *128*, 208–223. [CrossRef]
8. Vandenbogaerde, L.; Verbeke, S.; Audenaert, A. Optimizing building energy consumption in office buildings: A review of building automation and control systems and factors influencing energy savings. *J. Build. Eng.* **2023**, *76*, 107233. [CrossRef]
9. Ge, S.; Li, J.; Liu, H.; Wang, Y.; Zhou, H. Domestic Energy Consumption Modeling per Physical Characteristics and Behavioral Factors. *Energy Procedia* **2019**, *158*, 2512–2517. [CrossRef]

10. Ge, S.; Li, J.; Liu, H. Modelling of household energy consumption characteristics considering physical features and behavior factors. *Electr. Power Autom. Equip.* **2019**, *39*, 36–44.
11. Yuan, P.; Gao, K.; Zhao, X.; Liu, X.; Kong, W. Electricity consumption and load prediction method for Chinese rural residences based on the randomness and seasonality in electricity usage behavior. *Energy Build.* **2023**, *279*, 112647. [CrossRef]
12. Catalina, T.; Virgone, J.; Blanco, E. Development and validation of regression models to predict monthly heating demand for residential buildings. *Energy Build.* **2008**, *40*, 1825–1832. [CrossRef]
13. Olinger, M.S.; de Araújo, G.M.; Dutra, M.L.; Silva, H.A.D.; Júnior, L.P.; de Macedo, D.D. Metamodel Development to Predict Thermal Loads for Single-family Residential Buildings. *Mob. Netw. Appl.* **2022**, *27*, 1977–1986. [CrossRef]
14. Chen, Z.; Zhao, Z.; Deng, Q.; Tang, P.; Yang, C.; Li, X.; Gui, W. A knowledge embedded graph neural network-based cooling load prediction method using dynamic data association. *Energy Build.* **2023**, *278*, 112635. [CrossRef]
15. Deb, C.; Eang, L.S.; Yang, J.; Santamouris, M. Forecasting diurnal cooling energy load for institutional buildings using Artificial Neural Networks. *Energy Build.* **2016**, *121*, 284–297. [CrossRef]
16. Senthil Kumar, D.; George Washington, D.; Reshmy, A.K.; Noorunnisha, M. Energy Performance Prediction of Residential Buildings Using Nonlinear Machine Learning Technique. In *International Conference on Innovative Computing and Communications: Proceedings of ICICC 2021*; Springer: Singapore, 2022.
17. Li, G.; Tian, W.; Zhang, H.; Fu, X. A novel method of creating machine learning-based time series meta-models for building energy analysis. *Energy Build.* **2023**, *281*, 112752. [CrossRef]
18. Li, L.; Su, X.; Bi, X.; Lu, Y.; Sun, X. A novel Transformer-based network forecasting method for building cooling loads. *Energy Build.* **2023**, *296*, 113409. [CrossRef]
19. Kapp, S.; Choi, J.K.; Hong, T. Predicting industrial building energy consumption with statistical and machine-learning models informed by physical system parameters. *Renew. Sustain. Energy Rev.* **2023**, *172*, 113045. [CrossRef]
20. Pinto, G.; Wang, Z.; Roy, A.; Hong, T.; Capozzoli, A. Transfer learning for smart buildings: A critical review of algorithms, applications, and future perspectives. *Adv. Appl. Energy* **2022**, *5*, 100084. [CrossRef]
21. Liu, S.; Zhang, Z.; Yang, J.; Hu, W. Exploring Increasing Urban Resident Electricity Consumption: The Spatial Spillover Effect of Resident Income. *Energies* **2022**, *15*, 4249. [CrossRef]
22. Prataviera, E.; Vivian, J.; Lombardo, G.; Zarrella, A. Evaluation of the impact of input uncertainty on urban building energy simulations using uncertainty and sensitivity analysis. *Appl. Energy* **2022**, *311*, 118691. [CrossRef]
23. Yousaf, S.; Bradshaw, C.R.; Kamalapurkar, R.; San, O. Investigating critical model input features for unitary air conditioning equipment. *Energy Build.* **2023**, *284*, 112823. [CrossRef]
24. Fan, C.; Xiao, F.; Zhao, Y. A short-term building cooling load prediction method using deep learning algorithms. *Appl. Energy* **2017**, *195*, 222–233. [CrossRef]
25. Li, X.; Yao, R. Modelling heating and cooling energy demand for building stock using a hybrid approach. *Energy Build.* **2021**, *235*, 110740. [CrossRef]
26. Ma, J.; Cheng, J.C. Estimation of the building energy use intensity in the urban scale by integrating gis and big data technology. *Appl. Energy* **2016**, *183*, 182–192. [CrossRef]
27. R Language. Available online: https://www.rstudio.com/ (accessed on 16 September 2024).
28. Tian, W.; Heo, Y.; De Wilde, P.; Li, Z.; Yan, D.; Park, C.S.; Feng, X.; Augenbroe, G. A review of uncertainty analysis in building energy assessment. *Renew. Sustain. Energy Rev.* **2018**, *93*, 285–301. [CrossRef]
29. EnergyPlus. Available online: https://energyplus.net/ (accessed on 30 September 2024).
30. Saltelli, A.; Ratto, M. *Global Sensitivity Analysis*; John Wiley: Hoboken, NJ, USA, 2008.
31. *GB 55015-2021*; General Code for Energy Efficiency and Renewable Energy Application in Buildings. Energy Efficiency Design Standards for Public Buildings. China Construction Industry Press: Beijing, China, 2021.
32. *JGJ/T449—2018*; Standard for Green Performance Calculation of Civil Buildings. Energy Efficiency Design Standards for Public Buildings. China Construction Industry Press: Beijing, China, 2018.
33. *GB/T 51350-2019*; Technical Standard for Nearly Zero Energy Buildings. China Construction Industry Press: Beijing, China, 2019.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.