energies

MDPI

Article

Assessing the Influence of Occupancy Factors on Energy Performance in US Small Office Buildings

Seddigheh Norouziasl, Sorena Vosoughkhosravi , Amirhosein Jafari * and Zhihong Pang

Bert S. Turner Department of Construction Management, Louisiana State University, Baton Rouge, LA 70803, USA; snorou1@lsu.edu (S.N.); svosou1@lsu.edu (S.V.); zpang@lsu.edu (Z.P.)
* Correspondence: ajafari1@lsu.edu

Abstract: Office buildings are responsible for about 35% of the total electricity in the US and over 70% of building energy consumption occurs during occupancy periods. Therefore, understanding occupancy behavior is crucial for reducing building energy consumption. However, given the stochastic nature of occupant behavior, identifying which occupancy parameters have the most impact on energy consumption poses a considerable challenge. This study aims to investigate and quantify the impact of various occupancy parameters on the energy performance of a US small-sized office building using an EnergyPlus-based nationwide energy simulation. First, dynamic occupancy schedules are created based on different occupancy parameters using an agent-based model. Next, the generated dynamic occupancy schedules are integrated into a small office building model from the Department of Energy's prototypes. This creates a dataset of occupancy parameters and building energy performance across various climate zones. Finally, various feature selection and statistical analysis methods are applied to the generated dataset. This helps identify significant occupancy parameters and quantify their impact on building energy performance across different climate zones. According to the results of the study, buildings located in cool marine, mixed marine, and warm marine climate zones had lower total energy consumption compared to other zones. Additionally, feature selection methods identified "Occupant Density" as the primary significant variable impacting energy consumption, across all climate zones. These findings offer valuable insights into the influential occupancy parameters across various climate zones, highlighting the importance of tailoring occupancy schedules to enhance energy efficiency. They provide practical guidance that can be directly applied to optimize energy consumption and achieve significant energy savings in small office settings with different weather conditions.

Keywords: building simulation; occupant behavior; energy consumption; dynamic occupancy schedules

check for updates

Citation: Norouziasl, S.; Vosoughkhosravi, S.; Jafari, A.; Pang, Z. Assessing the Influence of Occupancy Factors on Energy Performance in US Small Office Buildings. *Energies* **2024**, *17*, 5277. https://doi.org/10.3390/en17215277

Academic Editors: Joanna Ferdyn-Grygierek, Krzysztof Grygierek and Agnes Psikuta

Received: 19 September 2024
Revised: 11 October 2024
Accepted: 19 October 2024
Published: 23 October 2024

1. Introduction

Office buildings account for about 35% of overall electricity consumption in the US [1]. Crucially, over 70% of building energy consumption occurs during occupancy periods, underlining the importance of understanding and optimizing occupancy behavior for improving building energy efficiency [2]. Analyzing the various occupancy parameters and their impacts on building energy performance is essential for moving toward smart and energy-efficient buildings [3]. Therefore, it is necessary to identify which occupancy parameters have a significant impact on building energy performance and quantify their impact on building energy consumption [4].

Some studies have focused on the role of occupant behavior in building energy performance simulation and presented a review of the existing literature. For instance, Yan et al. [5] reviewed the literature in terms of monitoring occupant behavior and collecting the occupancy data, modeling occupant behavior, evaluating the occupancy models, and integrating the models into building performance simulation tools. They emphasized the importance of developing a standardized framework for describing and modeling occupant behavior in buildings. Yang et al. [6] explored occupancy sensing technologies and methods for modeling occupant behavior in institutional buildings. They highlighted the challenges associated with implementing occupancy sensing and monitoring in institutional buildings, primarily because of the substantial number of occupants, their considerable variability, and the diverse functions of these buildings for certain techniques and methodologies. Norouziasl et al. [7] conducted a systematic review of the literature regarding modeling and simulation tools for human-building energy-related interaction. They established a framework for inputs and outputs in modeling occupant behavior and outlined the most effective techniques for simulating occupant behavior in building energy performance. Bäcklund et al. [8] focused on campus buildings, highlighting the evolving behaviors of occupants influenced by smart building systems. Their semisystematic literature review emphasized the significant impact of such systems on energy use, promoting a shift towards more energy-aware behaviors. This research underscores the potential for integrating smart systems into building management to optimize energy consumption and enhance educational environments. Additionally, Vosoughkhosravi et al. [9] provided a systematic review of the use of the American Time Use Survey (ATUS) in modeling occupant behavior. In this review, the authors investigated occupant behavior models and approaches developed based on ATUS. They offered a comprehensive analysis of modeling methods, required inputs and outputs, as well as the most practical occupant behavior methods.

Among the studies aiming to simulate occupant energy-related behavior, Chen et al. [10] developed an agent-based occupancy simulator to simulate the stochastic behavior of occupants, including occupants' presence and movement. They employed a homogeneous Markov chain model to simulate the stochastic occupancy schedules for each office room and the whole building. Then, the generated occupancy schedules were used in the EnergyPlus and obFMU simulation to evaluate the impacts of occupant behavior on building energy performance [11]. Putra et al. [12] also developed an agent-based model (ABM) to study building occupant behavior during load shedding, simulating occupants' adaptive actions and their impact on building energy consumption. In a similar study, Jia et al. [13] investigated the impact of actual and modeled occupant behavior information on building performance simulation. They used an agent-based modeling approach to simulate occupant behavior and conducted a cosimulation with a building energy model. Their study highlighted the significant influence of different occupant behavior inputs on building energy performance. Another study by Parys et al. [14] focused on integrating stochastic models of occupant behavior with dynamic building simulations. The authors reviewed various methods for this integration, emphasizing the importance of accurately modeling occupant behavior to improve the precision of energy performance predictions in office buildings. By coupling dynamic building simulations with stochastic occupant behavior models, the study aimed to address the variability and unpredictability of human actions, which significantly impact energy consumption. This integrated methodology helped in creating more realistic and reliable simulations, ultimately leading to better-informed decisions for energy-efficient building design and operation. Almeida et al. [15] also studied the uncertainty in occupant behavior in building energy models. They found that energy consumption could vary significantly based on different occupancy schedules and environmental preferences, highlighting the importance of accurate occupant behavior modeling. In their 2019 study, Gunay et al. [16] developed an occupancy learning algorithm for terminal heating units. They investigated how occupant behavior impacts the energy performance of HVAC systems by utilizing both field data and simulation models. Their findings highlighted the importance of accurately modeling occupant behavior to optimize HVAC system performance and improve energy efficiency in buildings. In a similar study by Li et al. [17], they explored the use of radio frequency identification (RFID) technology to measure and monitor occupancy in buildings. The authors developed an RFID-based system designed to provide real-time occupancy data, which can be used to optimize HVAC operations. By accurately tracking the presence and movement of occupants, the system allows for demand-driven HVAC control, which adjusts heating and cooling based on actual occupancy rather than predefined schedules. This approach can significantly improve energy efficiency by reducing unnecessary heating or cooling of unoccupied spaces, leading to potential energy savings and enhanced comfort for building occupants. In addition, Chen et al. [18] proposed two stochastic Markov chain models using real data to simulate the occupancy schedule in commercial buildings. These models simplified transition probability calculations and offered occupancy-based energy models for single-zone and multi-zone offices. Page et al. [19] developed an algorithm for simulating occupant presence in buildings using an inhomogeneous Markov chain model. The model was then integrated with building energy simulation as an input to account for future occupant behavior. By applying this model to occupancy data from private offices, the study demonstrated the key aspects of occupant presence, including arrival and departure times, and intermediate periods of absence in energy consumption patterns. A number of studies have concentrated on enhancing energy efficiency in buildings through improved occupancy modeling and predictive analysis. For instance, Oldewurtel et al. [20] investigated the potential of using occupancy information to realize a more energy efficient building climate control. In their research, a model predictive control (MPC) framework was employed to assess the energy savings potential of office buildings with different occupancy types. This comparative analysis considered different building types, HVAC systems, seasonal variations, and occupancy patterns to evaluate their respective effects on energy-saving potential. In another study, Rafsanjani et al. [21] conducted research on the influence of occupants' energy-consuming behaviors, such as arrival, departure and electricity-use patterns, in commercial buildings and quantified their potential for energy savings. The proposed study combined occupancy sensing with building energy data to assess the feasibility of the developed approach in identifying occupant-specific energy consumption information. Erickson et al. [22] addressed the inefficiencies of existing climate control systems that rely on maximum occupancy numbers, often resulting in unnecessary heating or cooling of infrequently used rooms. They utilized the occupancy data to develop multivariate Gaussian and agent-based models for predicting occupancy patterns and then implemented optimal control strategies to reduce the energy consumption of the HVAC system. Recent developments in urban-building energy modeling (UBEM) underscore the significant impact of occupant behavior on energy consumption within urban environments. Banfi et al. [23], in their comprehensive review, emphasized the limitations of static occupant profiles often utilized in current modeling practices. Advocating for dynamic and stochastic models, the study examines the integration challenges and the need for more sophisticated occupant behavior models to enhance the accuracy and relevance of urban energy simulations.

Existing research has primarily focused on the general effects of occupancy schedules on building energy performance. However, these studies often do not fully explore the influence of specific occupancy-related parameters, such as arrival and departure times, lunch breaks, and the frequency and duration of meetings on energy consumption, particularly within office settings. This gap limits the applicability of such studies for creating accurate, actionable energy management strategies tailored to routine human behaviors. Moreover, the body of research considering occupancy schedules rarely extends its analysis to compare these effects across diverse climatic conditions. The United States presents a unique landscape with a wide range of climate zones, each presenting distinct challenges and opportunities for energy management in office buildings. Comparative analysis across different climate zones is crucial but has been rarely covered in the research literature. Gaining insight into these differences is important. Understanding how occupancy schedules influence energy consumption in different climates can significantly enhance the development of localized, climate-specific energy conservation measures. Such detailed and comparative research is critical not only for advancing theoretical knowledge but also for informing policymakers and building managers. Tailored strategies could subsequently be developed to optimize energy use in office buildings nationwide, potentially leading to substantial reductions in energy costs and environmental impacts. This research gap presents a significant opportunity for a pioneering study that could set new directions for future energy efficiency initiatives and policies.

To address the mentioned gaps, this study aims to analyze the impacts of different occupancy parameters (e.g., the time occupants arrive at or leave the workplace, the time and number of meetings, and the time and duration of lunch breaks, among others) on energy consumption in office buildings across various climate zones in the US. In this study, an agent-based model (ABM) [24] is used to generate dynamic occupancy schedules from various sets of occupancy parameters to reflect stochastic occupancy behavior. In addition, the small-sized office building in the Department of Energy prototype Commercial Building Prototype Model (CBPMs) [25] is used for energy simulation. The generated stochastic occupancy schedules, as well as the office model, are integrated into the EnergyPlus simulation model to create a dataset of occupancy parameters and building end-use energy performance in different climate zones in the US. This dataset is used to select the most significant occupancy variables impacting building energy consumption using feature selection techniques. Furthermore, this research provides key insights that are invaluable for building designers, facility managers, and policymakers by delineating the critical occupancy parameters that substantially affect energy consumption in office buildings. This knowledge authorizes stakeholders to formulate specialized, climate-responsive strategies that not only optimize energy efficiency but also promote broader sustainability goals. Additionally, a thorough comprehension of these occupancy influences is crucial for crafting effective policies aimed at diminishing energy consumption and enhancing the energy efficiency of design and operational practices in American office buildings. By identifying these key parameters, the study equips stakeholders with the necessary tools to implement strategic interventions that can lead to significant energy savings and operational efficiencies, particularly in diverse climatic conditions across the United States.

2. Research Method

This study adopts a four-step methodology to analyze the impacts of different occupancy parameters on energy consumption, as illustrated in Figure 1. First, several dynamic occupancy schedules were created based on specific occupancy parameters using an agent-based model. Then, the generated dynamic occupancy schedules are integrated into the DOE prototype small-sized office building model. This prototype was developed by the Pacific Northwest National Laboratory (PNNL) [26] and the US Department of Energy's Building Energy Codes Program (BECP) to estimate how changes in energy codes and standards can lead to energy savings [27]. This prototype provides EnergyPlus IDF models for different office buildings designed based on the energy codes (i.e., The American Society of Heating, Refrigerating and Air-Conditioning Engineers (ASHRAE) Standard 90.1-2019 Energy Efficiency Standard for d Buildings Except Low-Rise Residential Buildings).

A nationwide energy simulation was conducted to generate a database of occupancy parameters and building end-use energy performance in different US climate zones. Then, the impact of different occupancy parameters on building energy performance was analyzed using a sensitivity analysis and four different feature selection methods. In the subsequent sections, this study investigates the occupancy parameters utilized to generate the occupancy schedule. Next, the integration process of the occupancy schedule, considering various climate zones, into the EnergyPlus simulation model is elaborated. Finally, this study presents the employed feature selection methods to analyze the outcomes of the energy performance simulation and identify the most influential variables.

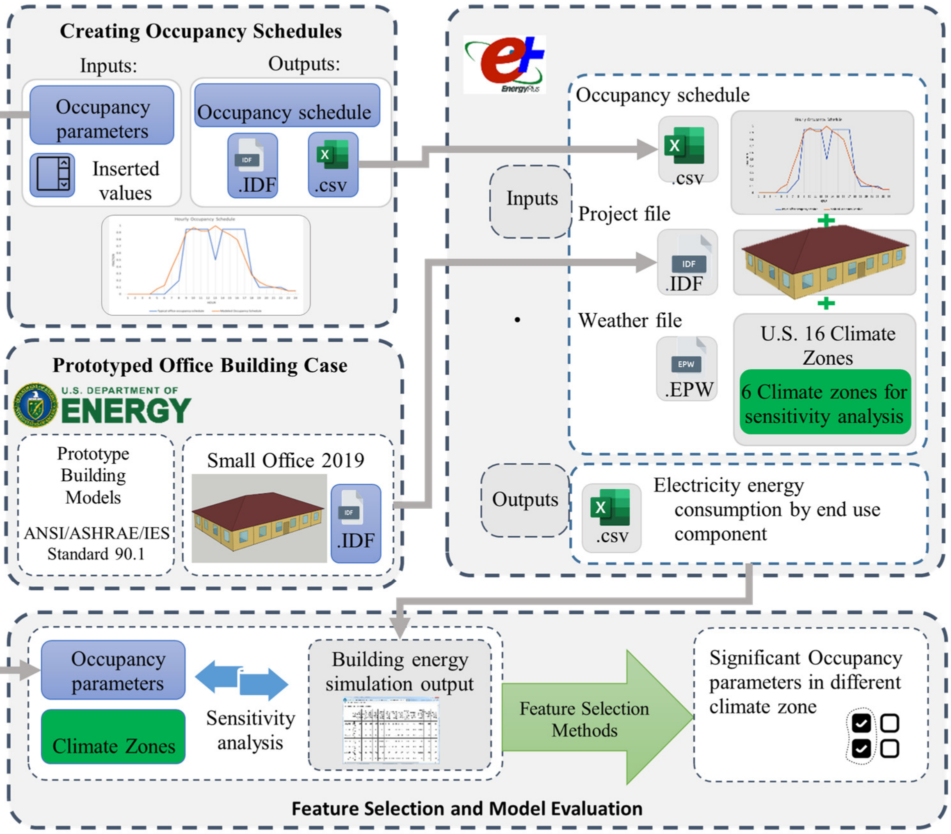

Figure 1. Dynamic occupancy schedule and building energy performance simulation framework.

2.1. Creating Occupancy Schedules

Occupancy in an office building can be defined with various parameters, such as the time occupants arrive at or leave the workplace, the time and number of meetings, and the time and duration of lunch breaks throughout the working hours for a specific day [28]. In the first step, different occupancy parameters are used to create various occupancy schedules using an agent-based model. This simulation approach offers a significant advantage over traditional case studies by allowing for the exploration of multiple scenarios, thereby providing a more comprehensive understanding of how different occupancy dynamics can influence building energy consumption. For this purpose, a web-based occupancy simulator tool is used, which was developed by Lawrence Berkeley National Laboratory (LBNL) [10]. The web-based occupancy simulator tool utilizes a Markov-chain model to simulate occupant movements and generate stochastic schedules. This innovative tool has been validated in real-world scenarios. The validation studies, as outlined in research by Luo et al. [29], demonstrate the tool's efficacy in accurately reflecting occupant behavior in commercial buildings, thereby providing reliable data for building simulation to optimize energy usage and operational efficiency. To closely mirror real-world occupancy patterns, the occupancy schedule is crafted by meticulously integrating as many relevant parameters as possible. The occupancy schedule can be generated by integrating the occupancy parameters and office layout parameters in the web-based occupancy simulator application. This simulator is a user-friendly application that uses a Markov Chain (MC) model to simulate occupancy in buildings. It takes in high-level inputs on occupant density, space area, event arrangements, etc., and simulates occupancy movements inside the building to generate stochastic and dynamic occupancy schedules for each space. This detailed simulation ensures that the generated occupancy schedules closely approximate actual daily activities and interactions within office environments. The generated occupancy schedule may vary daily, reflecting the inherent nature of occupants and their energy-related behavior in office buildings. The output results of these schedules can be downloaded in

CSV format to facilitate integration with energy simulation software, such as EnergyPlus 22.1.0 [30].

Table 1 presents the occupancy parameters utilized in this study to generate the occupancy schedules (the parameters are selected based on the advancement of the current LBNL occupancy simulator), including variable names, definitions, units, and values. As illustrated in this table, three types of occupants are considered in the simulation: regular staff, administrators, and managers. Each type of occupant has different arrival and departure times, and the duration that each occupant type spends in various spaces varies.

Table 1. Considered occupancy parameters for the considered office building.

No.	Variable Name	Variable Definition	Variable Unit	Values *	
1	Occupant_Density	Number of people per area	person/m^2	1 = 0.05 *** 2 = 0.06** 3 = 0.10	4 = 0.14 5 = 0.20 6 = 0.30
2	Occupant_Percent	Percentage of each occupant type: Regular staff/Manager/Administrator	Percentage	*** 1 = 40/30/30** 2 = 20/40/40 3 = 30/35/35 4 = 50/25/25 5 = 60/20/20 6 = 50/10/40 7 = 45/20/35	8 = 35/40/25 9 = 30/50/20 10 = 50/40/10 11 = 45/35/20 12 = 35/25/40 13 = 30/20/50
3	Meeting_Count	Number of meetings per day (Min–Max)	Count	1 = 1–3 2 = 2–4 *** 3 = 3–5**	4 = 4–6 5 = 5–7
4	Meeting_Attend	Number of people per meeting (Min–Max)	People	1 = 2–4 2 = 3–5 *** 3 = 4–6**	4 = 5–7 5 = 6–8
5	Meeting_Duration	Probability of duration of meeting for the following numbers (30, 60, 90, 120)	Percentage	1 = 5, 60, 20, 15 2 = 10,65, 15,10 *** 3 = 15, 70, 10, 5**	4 = 20, 65, 10, 5 5 = 25, 60, 10, 5 6 = 30, 55, 10, 5
6	Staff_Arriv_Depar	Regular Staff: Arrival time/Departure time	Time	1 = 6:30/15:30 2 = 7:00/16:00 *** 3 = 7:30/16:30**	4 = 8:00/17:00 5 = 8:30/17:30
7	Admin_Arriv_Depar	Administrator: Arrival time/Departure time	Time	1 = 7:00/16:00 2 = 7:30/16:30 *** 3 = 8:00/17:00**	4 = 8:30/17:30 5 = 9:00/18:00
8	Manag_Arriv_Depar	Manager: Arrival time/Departure time	Time	1 = 8:00/16:30 2 = 8:30/17:00 *** 3 = 9:00/17:30**	4 = 9:30/18:00 5 = 10:00/18:30
9	Arriv_Depar_Vari	Arrival/departure time variation	Minutes	1= 0 min 2 = 15 min *** 3 = 30 min**	4 = 45 min 5 = 60 min
10	Lunch_Time	Lunch or short-term leaving start time	Time	1 = 11:00 2 = 11:30 *** 3 = 12:00**	4 = 12:30 5 = 13:00
11	Lunch_Start_Vari	Lunch or short-term leaving Start time variation	Minutes	1= 0 min 2 = 15 min *** 3 = 30 min**	4 = 45 min 5 = 60 min
12	Lunch_Duration	Lunch or short-term leaving duration	Minutes	1 = 30 min 2 = 45 min *** 3 = 60 min**	4 = 75 min 5 = 90 min

Table 1. *Cont.*

No.	Variable Name	Variable Definition	Variable Unit	Values *	
13	Lunch_duration_Vari	Lunch or short-term leaving duration variation	Minutes	1 = 5 min 2 = 10 min *** 3 = 15 min**	4 = 20 min 5 = 25 min
14	Staff_Room_Stay	(Regular Staff) Percentage of time that occupants stay in each space	Percentage	1 = 50, 20, 10, 10, 10 2 = 55, 20, 10, 10, 5 3 = 60, 15, 10, 10, 5 4 = 65, 15, 10, 5, 5	*** 5 = 70, 10, 10, 5, 5** 6 = 75, 10, 10, 5, 0 7 = 80, 10, 5, 5, 0
15	Admin_Room_Stay	(Administrator) Percentage of time that occupants stay in each space	Percentage	1 = 35, 10, 35, 10, 10 2 = 40, 10, 30, 10, 10 3 = 45, 10, 30, 5, 10 *** 4 = 50, 10, 30, 5, 5**	5 = 55, 10, 25, 5, 5 6 = 60, 10, 25, 5, 0 7 = 65, 10, 20, 5, 0
16	Manag_Room_Stay	(Manager) Percentage of time that occupants stay in each space	Percentage	1 = 35, 10, 40, 5, 10 2 = 40, 10, 35, 5, 10 3 = 45, 5, 35, 5, 10 *** 4 = 50, 5, 35, 5, 5**	5 = 55, 5, 30, 5, 5 6 = 60, 5, 30, 5, 0 7 = 65, 5, 25, 5, 0
17	Own_Stay_Duration	Average stay time at Own office	Minutes	1 = 30 min 2 = 45 min *** 3 = 60 min**	4 = 75 min 5 = 90 min
18	Other_Stay_Duration	Average stay time at Other offices	Minutes	1 = 10 min 2 = 15 min *** 3 = 20 min**	4 = 25 min 5 = 30 min
19	Meeting_Stay_Duration	Average stay time at Meeting rooms	Minutes	1 = 30 min 2 = 45 min *** 3 = 60 min**	4 = 75 min 5 = 90 min
20	Auxiliary_Stay_Duration	Average stay time at Auxiliary room	Minutes	1 = 10 min 2 = 15 min *** 3 = 20 min**	4 = 25 min 5 = 30 min
21	Outdoor_Stay_Duration	Average stay time at Outdoor	Minutes	1 = 10 min 2 = 20 min *** 3 = 30 min**	4 = 40 min 5 = 50 min
22	Time_Step	Simulation time step	Minutes	1 = 5 min *** 2 = 10 min**	3 = 15 min 4 = 20 min

* Baseline scenario.

In this study, the baseline scenario was established by considering the default values for occupancy parameters according to the LBNL occupancy simulator tool [31] which is also illustrated in Table 1. As an example, the average square meters per person in the baseline scenario is 0.05 person/m^2, and 40% of occupants in the baseline are regular staff, 30% are administrators, and the other 30% of occupants are managers. To enhance the realism of the simulation, occupancy schedules are meticulously generated in the web-based occupancy simulator using the one-at-a-time (OAT) method [32] by varying individual occupancy parameters while keeping the remaining parameters at their baseline values. This process results in the creation of a total of 104 distinct annual occupancy schedules for each office space. These schedules are designed to closely replicate real-world occupant behavior, and will be used in EnergyPlus model to analyze the impact of each parameter on building energy consumption with a simulation time step of 10-min as illustrated in the baseline scenario.

2.2. Prototyped Office Building Case

This study used the US DOE Commercial Prototype. Building Models (CPBM) for small offices to analyze the impact of various occupancy parameters. The prototype includes the building model for different climate locations (16 US climate zones) based on

ASHRAE Standard 90.1-2019 [33]. The small office has a rectangular shape layout with a total area of 5500 ft2 (511 m^2) and consists of five thermal zones (four perimeter zones and one core zone), as illustrated in Figure 2. It has to be noted that one thermal zone can have one or more subspaces, though the thermal condition is maintained the same for the subspaces within the same thermal zone [34]. The office layout that is designed for this study is as follows: Zones 1, 2, and 3 contain the individual office spaces, Zone 4 contains the meeting room and lobby space, and the core zone contains the auxiliary and corridor area. This layout is shown in Figure 2. The generated stochastic occupancy schedules for spaces within each zone will be combined for use within that specific zone. For example, the occupancy schedules for Office 1, Office 2, and Office 3 will be aggregated to form the combined occupancy schedule for Zone 1.

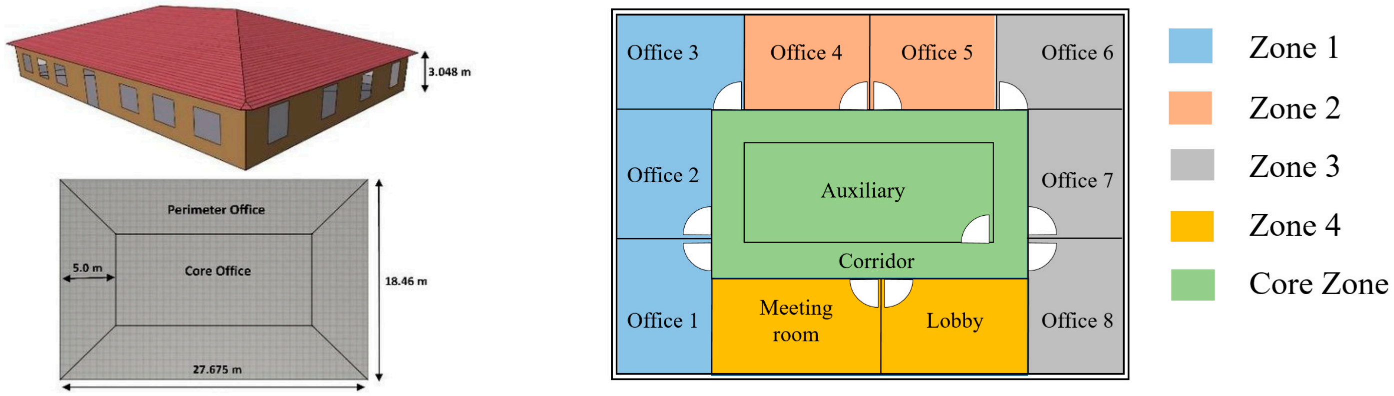

Figure 2. Modified DOE prototype model layout for a small office.

The small office model adheres to the energy efficiency standards established by ASHRAE 90.1, with the specific version (e.g., ASHRAE 90.1-2019) determining the envelope properties. For instance, the exterior wall U-value is approximately 0.057 Btu/h·ft^2·°F, reflecting an insulation level consistent with an R-value of about 17.5 ft^2·h·°F/Btu. This ensures that the building minimizes heat loss through the walls, contributing to overall thermal performance. The roof has an even lower U-value, around 0.027 Btu/h·ft^2·°F (equivalent to an R-value of approximately 37 ft^2·h·°F/Btu). This higher insulation level is essential in reducing heat transfer and limiting heat gain from the environment, especially in warmer climates where the roof is often exposed to direct sunlight. Windows in the small office model are another important component of the building envelope. The U-value for the windows is about 0.38 Btu/h·ft^2·°F. Additionally, the g-value, which indicates the amount of solar radiation admitted through the glazing, is 0.30. This balance allows beneficial daylight into the building while limiting unwanted solar heat gain, which can significantly impact cooling loads during the summer months.

2.3. Integrated Energy Simulation Model

To investigate the impact of each occupancy parameter on building energy performance, a total of 104 occupancy schedules were created to perform sensitivity analysis [35]. To incorporate occupancy schedules as inputs into the EnergyPlus model, a systematic approach is followed. First, a total of 104 annual occupancy schedule files are formatted as CSV files. These files are then organized within the EnergyPlus IDF project folder. Subsequently, a custom Python script is developed to automate the EnergyPlus simulation process. This script dynamically modifies the EnergyPlus IDF file, iterates through the schedule files one by one from the designated folder, and stores the results of each simulation in dedicated output folders, one for each unique occupancy schedule. This automatic workflow ensures the systematic integration of occupancy schedules into the EnergyPlus model, offering an organized approach for conducting simulations and managing their outputs.

To account for the effect of climate zones on energy performance simulation, 16 International Energy Conservation Code (IECC) climate zones in the US are integrated into EnergyPlus as well. These climate zones vary from "Very Hot Humid" to "Very Cold" and "Arctic" climate zone [36]. Table 2 presents the list of the 16 climate zones in the US, their climate zone ID, and the representative city. The energy performance of various building energy systems is then simulated for each climate zone, considering the various occupancy schedules. Finally, a dataset of 1664 data points (104 occupancy schedules and 16 climate zones) with 22 independent variables (occupancy parameter variables), and various systems' energy consumption variables consisting of heating, cooling, lighting, equipment, and fans are generated. The generated dataset is used as input for feature selection methods.

Table 2. US 16 climate zones [36].

No.	Climate Zone	Thermal Climate Zone Name	Weather Location
1	1A	Very Hot Humid	Honolulu, HI, USA
2	2A	Hot Humid	Tampa, FL, USA
3	2B	Hot Dry	Tucson, AZ, USA
4	3A	Warm Humid	Atlanta, GA, USA
5	3B	Warm Dry	El Paso, TX, USA
6	3C	Warm Marine	San Diego, CA, USA
7	4A	Mixed Humid	New York, NY, USA
8	4B	Mixed Dry	Albuquerque, NM, USA
9	4C	Mixed Marine	Seattle, WA, USA
10	5A	Cool Humid	Buffalo, NY, USA
11	5B	Cool Dry	Denver, CO, USA
12	5C	Cool Marine	Port Angeles, WA, USA
13	6A	Cold Humid	Rochester, MN, USA
14	6B	Cold Dry	Great Falls, MT, USA
15	7	Very Cold	International Falls, MN, USA
16	8	Subarctic/Arctic	Fairbanks, AK, USA

2.4. Feature Selection and Model Evaluation

Feature selection is an important task in identifying variables that can significantly impact the performance of the simulation model [37]. Previous studies on building energy performance simulation have used different methods for feature selection, including expert knowledge and judgment [38], correlation matrix [39], boosting tree algorithm to rank variables [40], and linear and monotonic correlation [41], among others. The generated dataset regarding different sets of occupancy schedules and climate zones is used as input for the feature selection models to determine the significant independent variables (occupancy schedules) to predict the dependent variables (energy consumption in different climate zones). In this study, four feature selection methods are utilized to identify the significant features regarding the dependent variables of building energy consumption:

- Multivariate linear regression (MVLR): A multivariate linear regression model expresses a d-dimensional continuous response vector as a linear combination of predictor terms plus a vector of error terms with a multivariate normal distribution. The "mvregress" function can be used to create a multivariate linear regression model [42]. While MVLR assumes linearity and can be influenced by multicollinearity and outliers, these challenges—multicollinearity and the influence of outliers—are prevalent in many types of statistical modeling, not just in MVLR. We have addressed this by careful variable selection and data preprocessing to minimize their impact.

- Least absolute shrinkage and selection operator (LASSO): LASSO constructs a dataset with redundant predictors and identifies those predictors. The "LASSO" function finds the coefficients of a regularized linear regression model using 10-fold cross-validation and the elastic net method [43]. LASSO may prioritize simpler models potentially at the cost of excluding some correlated predictors. However, this characteristic helps in enhancing model interpretability and reducing overfitting, which are crucial for the predictive robustness of the approach.
- Neighborhood component analysis (NCA) feature selection method: Neighborhood component analysis (NCA) is a supervised learning algorithm for choosing features with the goal of increasing the predictive power of regression and classification algorithms. The "fscnca" and "fsrnca" functions of the Statistics and Machine Learning Toolbox perform neighborhood component analysis feature selection with regularization to develop feature weights for the objective function that reduces the average leave-one-out classification or regression loss over the training data [44]. Despite NCA's computational demand, it is chosen for its effectiveness in smaller, well-defined datasets where feature interdependencies are critical, aligning well with our study's scope.
- Feature ranking method using the Relief algorithm: Relief is a feature selection technique that uses a filter-method approach to identify significant variables and is highly sensitive to feature interactions. Each feature in Relief is given a feature score, which can be used to rank and choose the highest scoring features for feature selection. These scores can also be used as feature weights to direct further modeling. The algorithm penalizes the predictors that result in different values to neighbors of the same class, and rewards predictors that provide different values to neighbors of different classes [45,46]. Although Relief's performance may be affected by noisy data, it is highly effective for datasets like ours where interaction among features is a significant factor. Proper parameter setting, based on extensive testing, ensures optimal feature selection.

The results of the four feature selection (also known as feature ranking) methods for various building energy systems (i.e., heating, cooling, lighting, equipment, and fans) are identified. In addition, the identified significant features, as results of feature selection methods, are integrated into a multiple linear regression model to compute the corresponding R-square values across different climate zones using the below equation:

$$R^2 = 1 - \frac{SSR}{SST} \tag{1}$$

$$SSR = \sum (\hat{y}_i - \overline{y})^2 \tag{2}$$

$$SSE = \sum (\hat{y}_i - y_i)^2 \tag{3}$$

where R^2, the coefficient of determination, is the proportion of the variation in the dependent variable that is predictable from the independent variables. SSR is the sum of squares of residuals, and SST is the sum of the distance the data are away from the mean all squared [47,48]. This value is utilized to compare the performance of feature selection methods, where a higher R^2 value represents a better fitting of the algorithm, and, on the other hand, a lower R^2 value represents a larger discrepancy between the actual and predicted results.

3. Results and Discussion

3.1. Building Energy Performance on Baseline Occupancy

In this study, the baseline scenario was established by considering the average occupancy parameters. Figure 3 compares the baseline stochastic occupancy schedule with ASHRAE Standard 90.1-2019 [49], and simulated occupancy. According to the results (for the baseline occupancy schedule), this small office building accommodates up to 56 (considering the minimum required space for each person in an office area, every office room can occupy up to 7 occupants) occupants who typically arrive at the office around 8:00 a.m. and depart around 5:00 p.m. As demonstrated in Figure 3, the occupancy schedule aligns with the office schedule recommended by ASHRAE Standard 90.1-2019, demonstrating a remarkable similarity to real-world occupancy patterns; however, the stochastic occupancy schedule provides more realistic occupancy patterns in office buildings [28,49]. This close alignment not only validates the simulation's accuracy but also highlights its potential to predict actual building usage with high fidelity. The occupancy levels fluctuate throughout the day, with specific hours experiencing higher occupancy rates. Notably, between 9:00 a.m. and 11:00 a.m. and again from 2:00 p.m. to 4:00 p.m., the occupancy schedule reaches its peak.

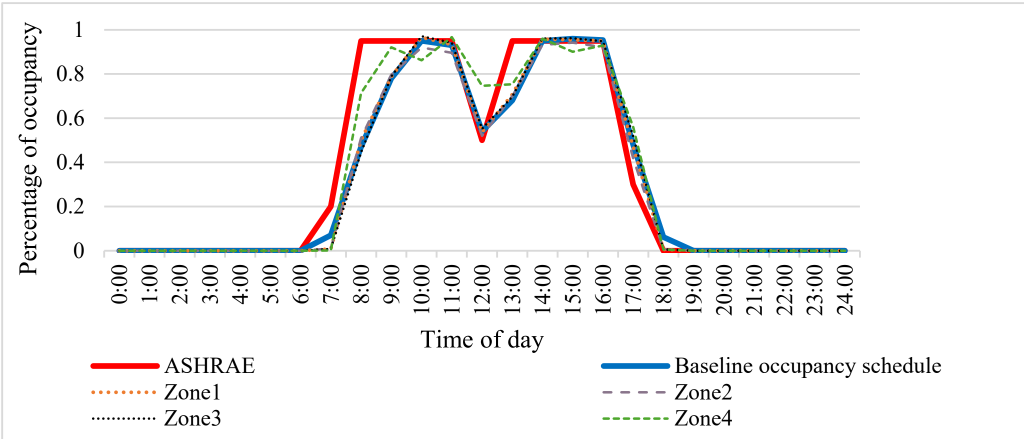

Figure 3. Baseline occupancy schedule vs. ASHRAE schedule.

Figure 4 illustrates the baseline occupancy schedule of each zone. The core zone and Zone 4, which include auxiliary, corridor, meeting room, and lobby, generally maintain lower occupancy levels compared to other zones. In addition, as shown in Figure 4, Zones 1, 2, and 3, which contain the office spaces, have more occupants since the average time spent by occupants in office spaces is longer than in other areas. It can be seen that the number of occupants decreases at 12:00 p.m. due to the lunch break.

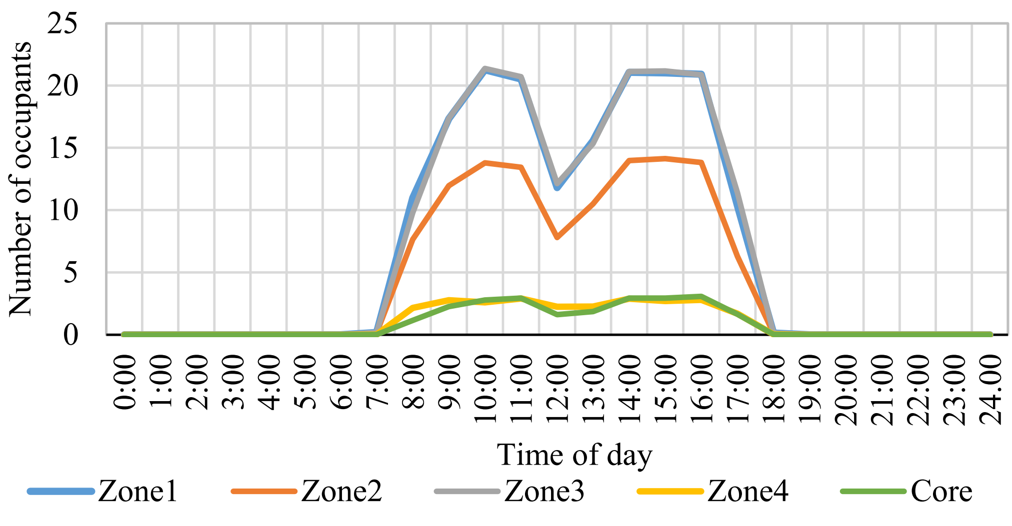

Figure 4. The zone occupancy of the simulated office.

To illustrate the average energy consumption during different seasons, we selected Tampa, FL (CZ 2A), as an example case due to the high building energy consumption in this climate zone. Figure 5 displays the average daily total electricity consumption of the simulated office building (including the electricity consumption for lighting, heating, cooling, equipment, and fans, in accordance with the baseline occupancy schedule) for the hot–humid climate zone (2A), for spring (1 March to 31 May), summer (1 June to 31 August), fall (1 September to 30 November), and winter (1 December to 28 February) seasons.

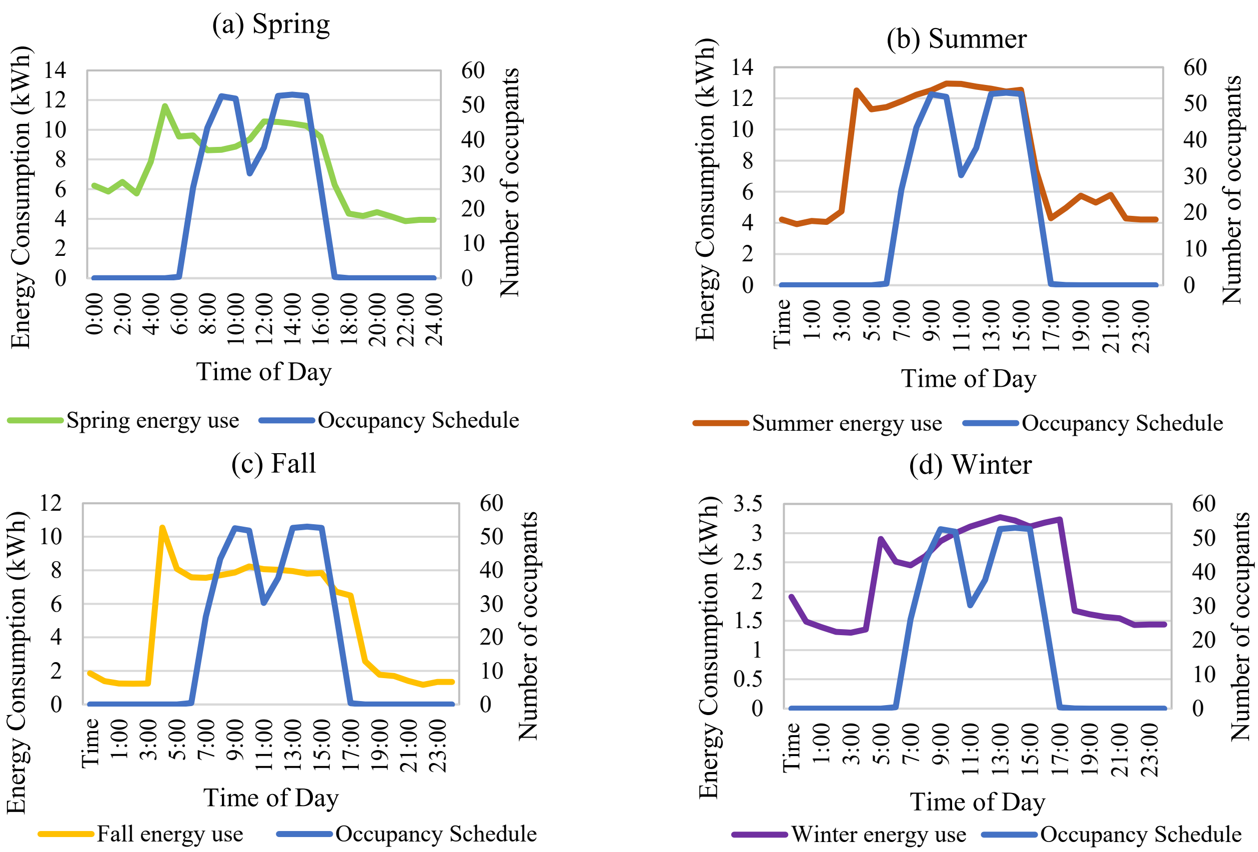

Figure 5. Simulated energy consumption and baseline occupancy schedule in (**a**) spring, (**b**) summer, (**c**) fall, and (**d**) winter seasons.

The graph reveals variations in the building's energy consumption across seasons, ranging from 1.5 to 13.5 kWh with respect to the baseline occupancy schedule. As shown, the building exhibits lower energy consumption in winter and fall seasons compared to summer and spring, primarily due to the need for higher cooling and air conditioning in the hot–humid climate zone.

Figure 6 illustrates the energy performance of various building energy systems (i.e., heating, cooling, lighting, equipment, and fans) in different climate zones. It can be seen that the total building energy can range from 35,000 to 47,000 kWh. According to this figure, buildings in 5C (cool marine), 4C (mixed marine), and 3C (warm marine) climate zones have the lowest total energy consumption among the other climate zones. One reason is the moderate and relatively stable temperatures characteristic of marine climates. The absence of extreme temperature fluctuations reduces the demand for heating or cooling, resulting in lower energy consumption. On the other hand, the energy consumption of buildings in hot–humid climate zones is higher than in other climate zones. This is due to the high demand for cooling and the necessity of using air conditioning, dehumidifying, and circulating the air in these climate zones.

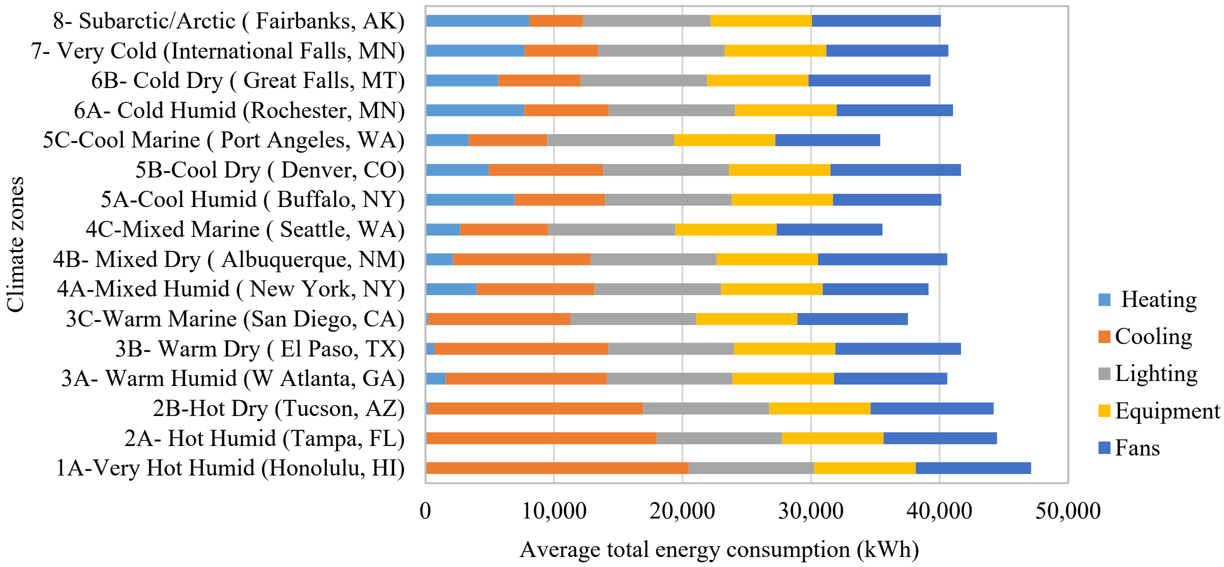

Figure 6. Energy consumption of the simulated office buildings in various climate zones.

As shown in Figure 6, the energy consumption of heating and cooling is very sensitive to climate conditions, resulting in significant variations. The average heating energy consumption in Fairbanks, AK (CZ 8), exceeds 6000 kWh, while in hot climate zones (e.g., 1A and 2A), heating energy consumption is negligible. In contrast, cooling energy consumption accounts for a large share of energy consumption, roughly 17,600 kWh, in very hot and humid climate zones and around 4400 kWh in cold climate zones. In addition, the energy consumption of lighting systems is mainly determined by indoor lighting needs, which are relatively consistent across various climate zones. Regardless of the climate, buildings require lighting for adequate illumination, resulting in comparable energy consumption for lighting systems. Similarly, the energy consumption of equipment, such as office equipment and appliances, is influenced more by occupant behavior and usage patterns rather than climate. Therefore, equipment energy consumption remains consistent across all climate zones. Additionally, fan energy consumption is about 9000 kWh with minimal variations in different climate zones. However, the demand for fans is slightly higher in dry climates to circulate the air and create a perceived cooling sensation for occupants, attributable to the low moisture content in the air in these dry climates.

The performance of the various end-use energy consumption in 16 climate zones considering the baseline occupancy parameters is illustrated in Figure 7. According to Figure 7a, heating electricity consumption is significantly high in humid climate zones (4A, 5A, and 6A). As shown in Figure 7b, cooling accounts for the largest share of electricity consumption in office buildings in many climate zones, ranging from very hot (e.g., 1A) to cool climate zones (e.g., 5B) compared to other end uses. Figure 7c,d demonstrate the electricity consumption for lighting and equipment in various climate zones. Although there are some slight differences regarding the comparison of lighting and equipment between climate zones, the results for these categories suggest that climate zones do not significantly impact lighting and equipment usage. This consistency across different climates can be attributed to the standardized nature of lighting and equipment operation in office settings. Unlike heating or cooling systems, which are directly influenced by external temperature variations and climate-specific requirements, lighting and equipment demands are predominantly driven by fixed office hours and internal activities that do not vary substantially with climate. Lighting consumption is largely determined by the daily work schedules that remain constant regardless of external weather conditions. Similarly, the usage of office equipment such as computers, printers, and other peripherals is aligned more closely with personnel presence and operational requirements rather than environmental factors. This explains the relative uniformity in energy usage for these categories across various climate zones observed in the study. Lastly, Figure 7e displays the electricity consumption for fans, revealing that dry climate zones (3B, 4B, and 5B) exhibit the highest electricity consumption for fan usage.

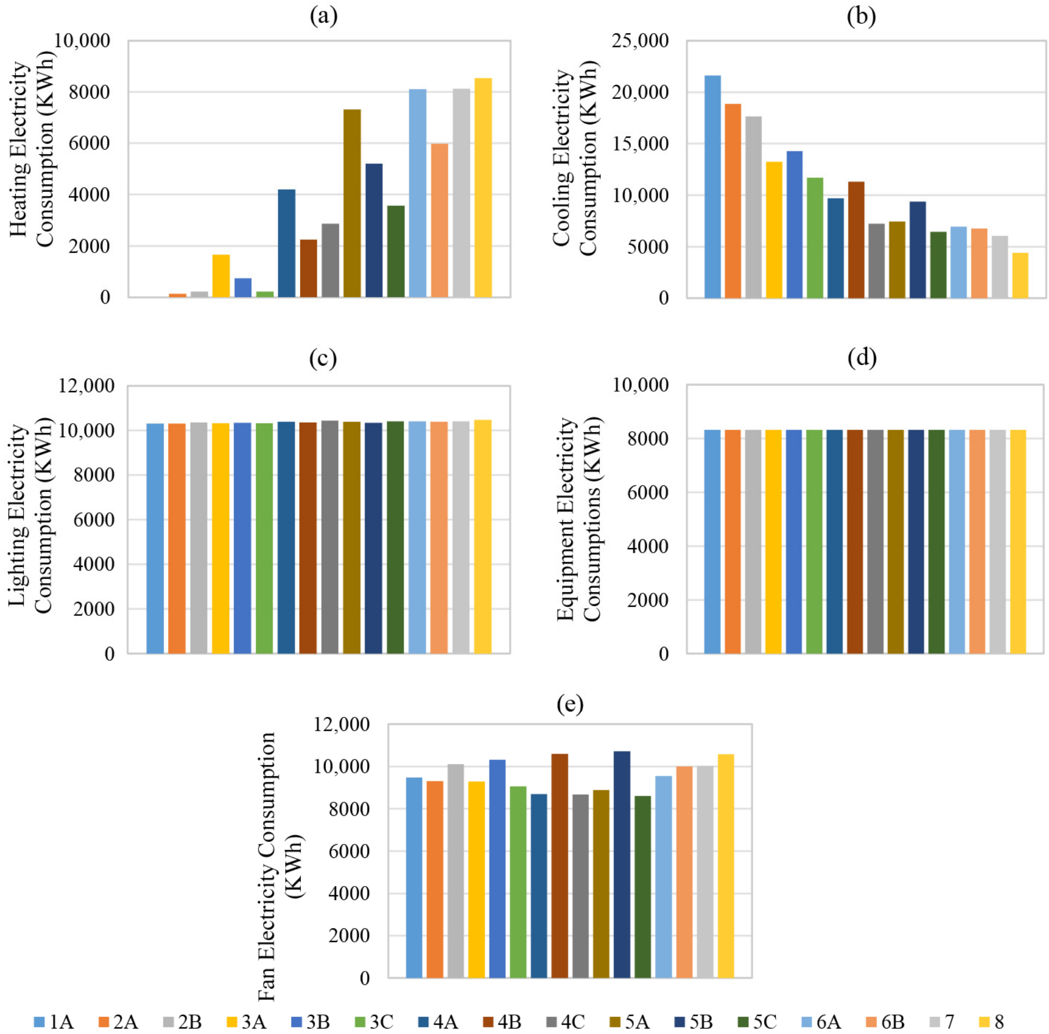

Figure 7. Average annual energy consumption of end-use systems for various climate zones. (**a**) Heating electricity consumption; (**b**) Cooling electricity consumption; (**c**) Lighting electricity consumption; (**d**) Equipment electricity consumption; (**e**) Fan electricity consumption.

Our findings indicate significant variations in building energy consumption across different climate zones, primarily driven by occupancy patterns and climatic conditions. This observation aligns with Chen et al. [50], who investigated the impact of future climate changes on office buildings across diverse climate zones in China, highlighting a similar trend of varying energy demands due to climatic differences. However, our study extends these findings by incorporating a wider range of climate zones in the US and employing dynamic occupancy schedules, which offer a more detailed understanding of energy consumption patterns. Unlike the static models used in this study, our dynamic approach reflects real-world variability and provides a new perspective on energy optimization strategies tailored to specific climate conditions.

Additionally, Meng et al. [51] investigated heating energy consumption in office buildings across various climate zones. Their findings confirm our results, which suggest that heating demands vary considerably, emphasizing the need for region-specific energy management strategies. Our study contributes further by examining both heating and cooling demands in a dynamic context, thereby improving the understanding of total energy consumption dynamics under changing climatic conditions.

3.2. Sensitivity Analysis of Occupancy Parameters

In this study, given the constraints on time and resources, six climate zones ranging from hot to cold were chosen to analyze occupancy parameters and assess the effectiveness of feature selection methods. The selected climate zones are presented in Table 3.

Table 3. Selected climate zones for sensitivity analysis.

No.	Climate Zone	Thermal Climate Zone Name	Weather Location
1	2A	Hot Humid	Tampa, FL, USA
2	3B	Warm Dry	El Paso, TX, USA
3	3C	Warm Marine	San Diego, CA, USA
4	4A	Mixed Humid	New York, NY, USA
5	5A	Cool Humid	Buffalo, NY, USA
6	6A	Cold Humid	Rochester, MN, USA

A sensitivity analysis was performed to analyze the impact of changing each occupancy schedule parameters on building energy performance. These parameters are analyzed by changing them while keeping all the other variables constant. The results are presented here.

Occupant Density: One of the occupancy parameters that may greatly impact building energy consumption is occupant density. Occupant density can be defined as the number of people per area (person/m^2). In this study, six different scenarios were considered to analyze the impact of occupant density on building energy performance:

- Scenario 1: 0.05 person/m^2, equal to 8 occupants using the office.
- Scenario 2: 0.06 person/m^2, equal to 16 occupants using the office.
- Scenario 3: 0.1 person/m^2, equal to 24 occupants using the office.
- Scenario 4: 0.14 person/m^2, equal to 40 occupants using the office.
- Scenario 5: 0.2 person/m^2, equal to 56 occupants using the office.

Figure 8 illustrates the annual building energy consumption across various climate zones, considering different occupant densities. Energy consumption for cooling, fans, and lighting tends to increase with higher occupant densities, while equipment and heating consumption show relatively smaller fluctuations. In warmer climate zones such as 2A, 3B, and 3C, the cooling and fan energy consumption exhibit high sensitivity to occupant density. Increasing the occupant density or number of occupants in these zones results in higher energy consumption for cooling and fans. This is primarily because a larger occupant density generates more heat and increases the cooling demand to maintain thermal comfort.

In addition, increasing the occupant density contributes to increased fan energy consumption to enhance air circulation in more crowded spaces. On the other hand, colder climate zones like 6A and 5A show higher heating energy consumption, with a negative correlation to occupant density. This can be attributed to the heat generation from occupants in more crowded spaces, resulting in reduced reliance on heating systems for maintaining comfortable temperatures.

This observed trend aligns with the findings of Zhao Dong et al. [52], who explored the impact of occupant behavior, particularly density, on energy consumption in office buildings. Their study provides a basis for understanding how increased occupancy contributes to higher energy demands. By referencing this broader scope of research, we can highlight that our findings are consistent with established trends in the field, suggesting that occupant density is a crucial factor in building energy dynamics. This also emphasizes the importance of incorporating occupant density considerations in the design and operation of HVAC systems, potentially through strategies such as demand-controlled ventilation and occupancy-based HVAC operation to optimize energy efficiency in office spaces.

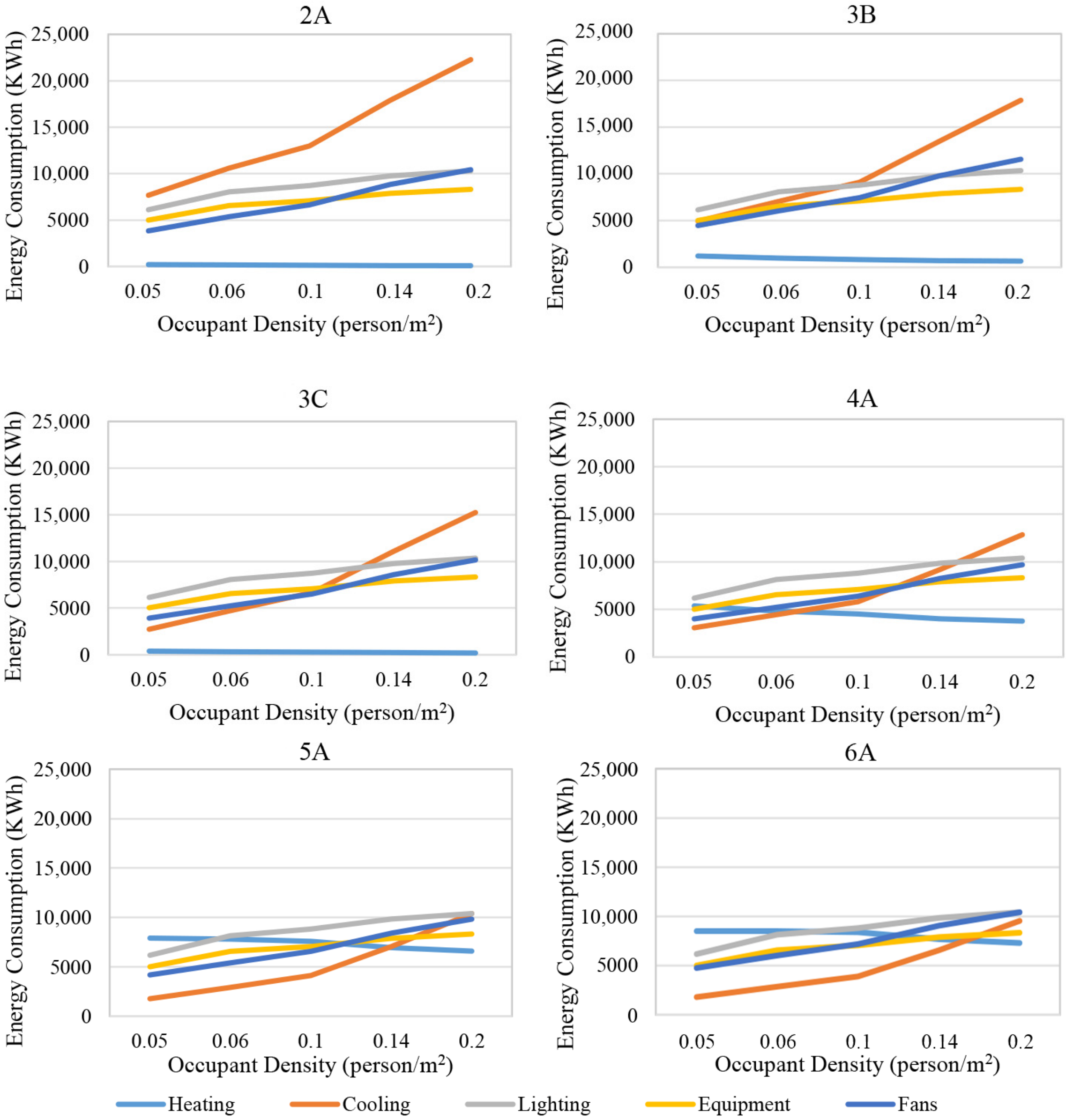

Figure 8. The impact of occupant density on building energy performance.

Arrival and departure: To analyze the impact of the arrival and departure time of occupants on building energy consumption, five distinct scenarios were considered:

- Scenario 1: Regular staff arrive and depart at 6:30 and 15:30, respectively.
- Scenario 2: Regular staff arrive and depart at 7:00 and 16:00, respectively.
- Scenario 3: Regular staff arrive and depart at 7:30 and 16:30, respectively.
- Scenario 4: Regular staff arrive and depart at 8:00 and 17:00, respectively.
- Scenario 5: Regular staff arrive and depart at 8:30 and 17:30, respectively.

Considering the average working hours of 8 h per day with a 1 h lunch break, regular staff work 9 h in the office building. Figure 9 presents the annual energy consumption of building systems in six climate zones considering different arrival and departure times for regular staff. Although the total working hours of regular staff are the same in all scenarios, the annual energy consumption is different. For example, in scenarios 1 and 2, the regular staff arrive at the office between 6:30 a.m. and 7:00 a.m. and leave the building between 3:30 p.m. and 4:00 p.m. It results in lower energy consumption due to taking advantage of cooler morning temperatures. In most of the climate zones, the early morning hours tend to be relatively cooler compared to the later part of the day. By starting work earlier, occupants can benefit from the cooler temperature and natural ventilation, reducing the need for extensive cooling and fan usage during the day. Consequently, in all climate zones, the cooling and fan energy consumption increases from scenario 1 to scenario 5, when shifting the arrival and departure times from 6:30 a.m. to 8:30 a.m. and 3:30 p.m. to 5:30 p.m., respectively. In contrast, there is a slight reduction in lighting and equipment energy consumption in all climate zones from scenario 1 to 5.

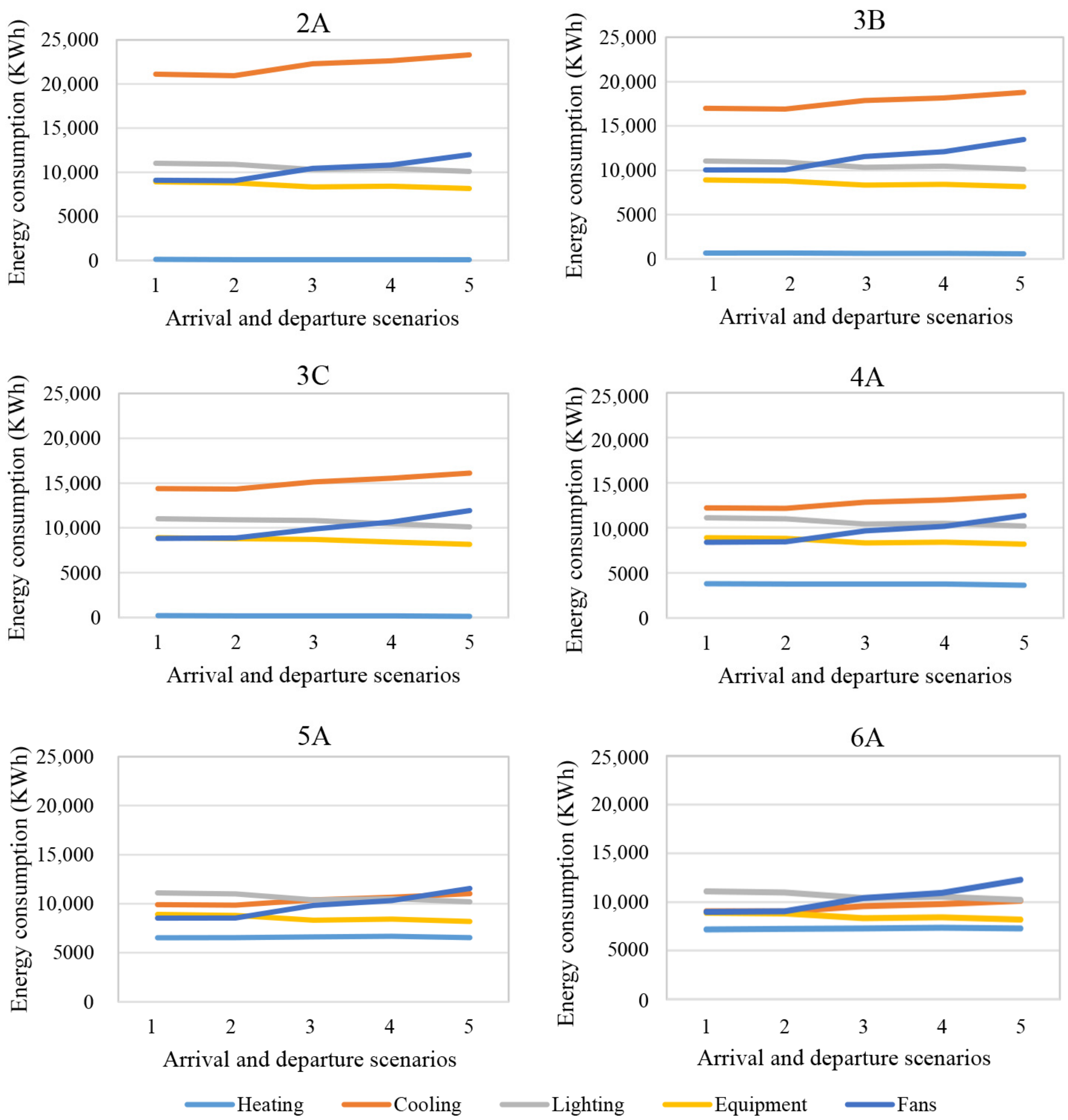

Figure 9. The impact of arrival and departure times of regular staff on building energy performance.

This observation aligns with findings from a study conducted by Gu et al. [53], which demonstrated that energy consumption varies significantly with different occupancy levels due to the presence and movement of occupants within the building. Specifically, the study highlighted how adjustments in the arrival and departure times can influence the operational schedules of building systems, thereby affecting the energy consumption for lighting, cooling, and heating. While the study of Gu et al. provided crucial insights, our study advances this by incorporating a broader analysis across multiple climate zones and more varied work schedule scenarios. This comprehensive approach not only allows for a detailed comparison across regions but also enhances the applicability of our findings to diverse environmental conditions and work patterns. These results highlight the potential for strategic scheduling to optimize energy efficiency in office buildings, corroborating the importance of considering occupant behavior patterns in energy management strategies.

Occupants' stay-time: Occupants' stay time in their own office is selected for analysis, where five different scenarios were considered to investigate the impact of occupants' stay-time in six climate zones:

- Scenario 1: Occupants stay in their own office for 30 min, on average.
- Scenario 2: Occupants stay in their own office for 45 min, on average.
- Scenario 3: Occupants stay in their own office for 60 min, on average.
- Scenario 4: Occupants stay in their own office for 75 min, on average.
- Scenario 5: Occupants stay in their own office for 90 min, on average.

Figure 10 presents the energy performance of various energy systems across six different climate zones, considering different durations of occupants' stay-time in their offices. As shown in the figure, for all six climate zones, there are minor changes in the energy consumption of the cooling system and fans when the duration of occupants' stay-time increases from 30 min to 90 min. As occupants stay in a space for a longer duration, their body heat gradually increases the room's temperature. Therefore, the cooling system and fans may need to operate for a slightly extended period to maintain air quality and temperature.

Conversely, the remaining energy systems, including lighting, equipment, and heating systems, remain unchanged, indicating that an increase in the duration of occupants' stay-time did not have a significant impact on these systems. This is because lighting and equipment energy consumption is related to occupancy presence and specific activities rather than the duration alone. Within the considered range of 30 to 90 min, the occupancy duration does not significantly affect the usage patterns or energy consumption of lighting and equipment. Similarly, heating systems respond to the set-point temperature and occupancy needs in colder climates. Therefore, increasing the occupants' stay-time does not demand an adjustment in these systems, resulting in unchanged energy consumption. This observation indicates that for the range of stay times considered, the direct impact on energy consumption is limited. This stability offers potential for energy management strategies that focus more on occupancy-based controls rather than adjustments based on duration of stay alone. Future studies might explore more granular time increments or different types of activities to further refine our understanding of occupancy impact on energy use. Such insights could inform more targeted energy efficiency measures that align with actual usage patterns and contribute to broader energy sustainability goals.

Time-step: To analyze the impact of simulation time-step on the energy performance of various energy systems in different climate zones, four different time-steps are considered:

- Scenario 1: Time-step size of 5 min.
- Scenario 2: Time-step size of 10 min.
- Scenario 3: Time-step size of 15 min.
- Scenario 4: Time-step size of 20 min.

Figure 11 presents the simulation performance of heating, cooling, lighting, and other energy consumptions using different time-step sizes. As shown, when employing a 5 min time-step, the energy performance of cooling systems and fans exhibits a significant increase across all six climate zones compared to larger step sizes. The utilization of a smaller time-step allows for a more detailed representation of the dynamic behavior of cooling systems and fans. With a smaller time-step, the simulation captures shorter fluctuations in cooling demand and the need for increased fan operation to maintain thermal comfort.

Consequently, the energy consumption of cooling systems and fans appears higher when using a 5 min time-step, reflecting their more responsive nature to varying conditions. This responsiveness underscores the importance of selecting an appropriate simulation time-step in energy modeling to capture the nuanced effects of environmental changes and occupant interactions within a building.

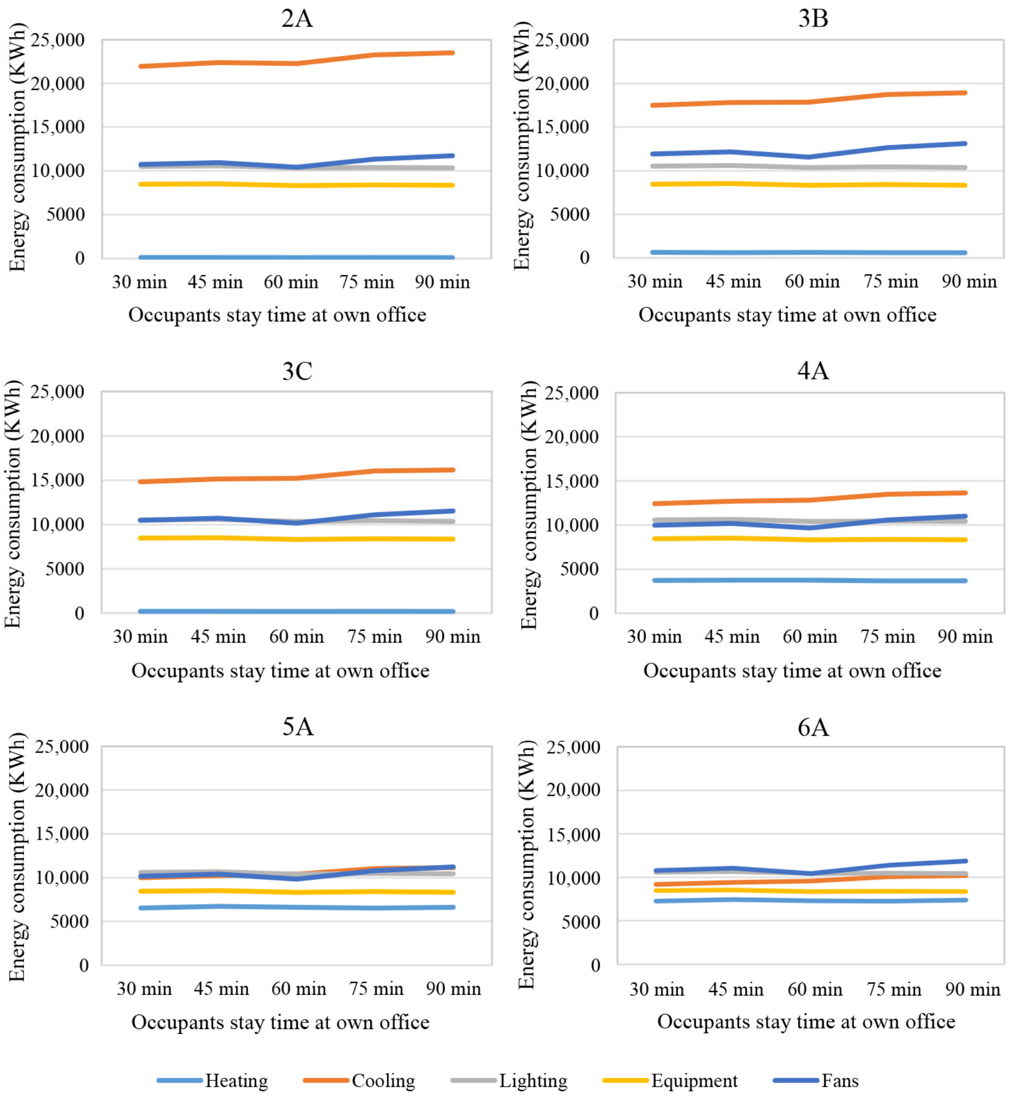

Figure 10. The impact of occupants' stay-time duration at their own office on building energy performance.

Additionally, the smaller time-step provides a clearer insight into the intermittent nature of heating requirements, thereby enhancing the ability to fine-tune heating operations to actual needs rather than sustained assumptions. This precision leads to energy conservation, particularly in climates where heating demands fluctuate significantly throughout the day. The energy consumption of lighting and equipment remains relatively constant regardless of the simulation time-step. This stability indicates that these systems' operational demands are less sensitive to shorter time fluctuations and more dependent on fixed schedules of occupant activity. Understanding these dynamics is crucial for building managers and designers aiming to optimize energy usage without compromising occupant comfort and productivity. Looking more closely at different time-steps could help us find new ways to save energy, especially in systems that automatically adjust based on real-time data.

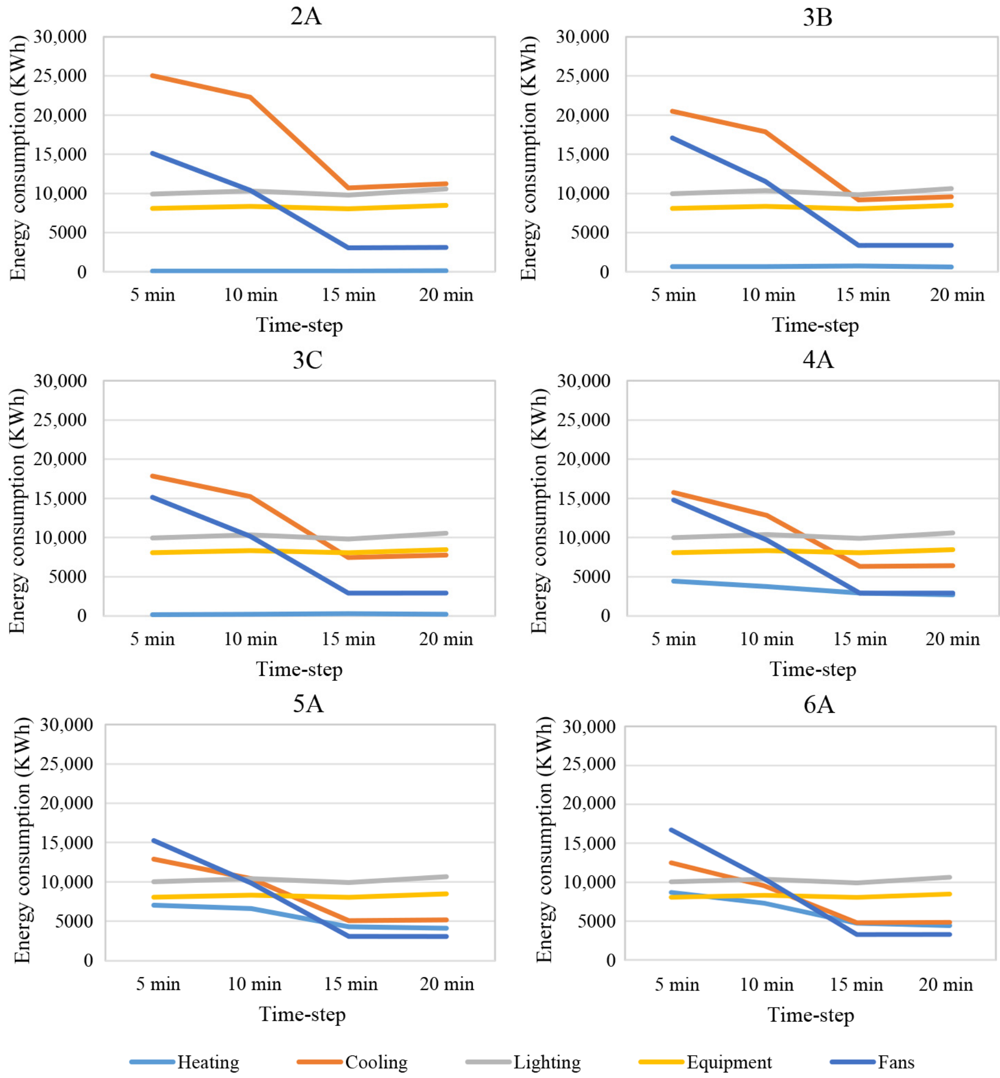

Figure 11. The impact of occupancy simulation time-step size on building energy performance.

The sensitivity analysis of four occupancy parameters was presented. Although a total of 22 parameters exist, it is important to note that the selection of these four parameters does not diminish the significance of the other parameters. The choice to analyze these particular parameters was driven by two reasons. Firstly, these four parameters were identified as having a high potential for influencing building energy consumption based on prior research and expert knowledge in the field. Secondly, considering the constraints of limited resources and time, focusing on a subset of parameters allowed for a more in-depth and targeted analysis.

3.3. Feature Selection Results

The generated dataset of building energy performance based on various climate zones and occupancy parameters served as input for feature selection methods to identify the most significant occupancy parameters affecting building energy consumption. In this regard, MVL, LASSO, NCA, and ReliefF feature selection methods were applied to the generated dataset. To assess the results of these feature selection methods, the selected features for each building energy system (i.e., heating, cooling, lighting, equipment, and fans) were incorporated into a multiple linear regression model to compute the corresponding R-square values. Subsequently, these calculated R-square values were evaluated against a predefined threshold of 0.05 as established in this study [54]. Features that demonstrated contributions greater than this threshold in the calculation of the R-square value were retained and considered significant variables. Conversely, features that exhibited contributions below the threshold in the calculation of the R-square value were eliminated from the list. This systematic process facilitated the identification of significant variables, allowing for a focused analysis of the features that exhibited a significant impact on the performance of the building energy systems.

Heating: This section focuses on the performance of heating energy consumption and presents the results of four feature selection methods to identify the significant occupancy parameters impacting heating energy consumption in six climate zones, as presented in Table 4. As shown in the table, all four methods consistently identified "Occupant-Density" as the primary significant variable across all six climate zones. Furthermore, "Time-Step" emerged as the second significant parameter in most cases, except for the "2A" climate zone where it did not exhibit the same significance.

Cooling: This section focuses on the performance of cooling energy consumption and presents the results of four feature selection methods to identify the significant occupancy parameters impacting cooling energy consumption in six climate zones, as presented in Table 5.

As shown in the table, all four methods consistently identified "Occupant-Density" and "Time-step" as the first two primary significant variables across all six climate zones. According to Table 5, it can be seen that the LASSO feature selection method presented a better performance across all six climate zones in identifying the significant features influencing cooling energy consumption. This method generated higher R-squared values compared to other methods. The selected features in the LASSO method include "Time-step", "Occupant_Density", "Staff_Arriv_Depar", "Lunch_Duration", and "Own_Stay_Duration". Conversely, the ReliefF method presented limited effectiveness in identifying significant features. This analysis focuses on influential occupancy parameters affecting cooling energy consumption in small office buildings across different climate zones and highlights the LASSO feature selection method as the preferred approach for identifying the features that have a significant impact on cooling energy consumption.

Lighting: This section focuses on the analysis of lighting energy performance and presents the results of four feature selection methods to identify the significant occupancy parameters impacting lighting energy consumption in six climate zones as presented in Table 6. As shown in the table, all four methods consistently identified "Occupant-Density" as the primary significant variable across all six climate zones. According to Table 6, it can be seen that both the NCA and LASSO feature selection methods presented a better performance across all six climate zones in identifying the significant features influencing lighting energy consumption. These methods generated higher R-squared values compared to other methods.

Conversely, the ReliefF method presented a low performance in identifying significant features. This analysis focuses on influential occupancy parameters affecting lighting energy consumption in small office buildings across different climate zones and highlights that both NCA and LASSO feature selection methods are preferable for identifying the features that have a significant impact on lighting energy consumption.

Equipment: This section focuses on the analysis of equipment energy performance and presents the results of four feature selection methods to identify the significant occupancy parameters impacting equipment energy consumption in six climate zones, as presented in Table 7. As shown in the table, all four methods consistently identified "Occupant-Density" as the primary significant variable across all six climate zones. According to Table 7, it can be seen that the LASSO feature selection method presented a better performance across all six climate zones in identifying the significant features influencing equipment energy consumption. This method generated higher R-squared values compared to other methods. The selected features in the LASSO method include "Occupant_Density", "Staff_Arriv_Depar", "Manag_Arriv_Depar", "Arriv_Depar_Vari", and "Time_Step". Conversely, the ReliefF method presented limited effectiveness in identifying significant features. This analysis focuses on influential occupancy parameters affecting equipment energy consumption in small office buildings across different climate zones and highlights the LASSO feature selection method as a preferred approach for identifying the features that have a significant impact on equipment energy consumption.

Table 4. Identified features for heating energy performance across different climate zones.

Climate Zone	MVLinear	LASSO	NCA	ReliefF
2A	• Occupant_Density • Staff_Arriv_Depar • Lunch_Time • Lunch_Duration	• Occupant_Density • Staff_Arriv_Depar • Admin_Arriv_Depar • Manag_Arriv_Depar • Own_Stay_Duration	• Occupant_Density • Admin_Arriv_Depar • Manag_Arriv_Depar	• Occupant_Density • Meeting_Stay_Duration • Manag_Room_Stay
R-squared	0.89596	0.90338	0.88389	0.87748
3B	• Occupant_Density • Time_Step • Staff_Arriv_Depar • Lunch_Duration • Lunch_Time	• Occupant_Density • Time_Step • Staff_Arriv_Depar • Admin_Arriv_Depar • Manag_Arriv_Depar	• Occupant_Density • Time_Step • Staff_Arriv_Depar	• Occupant_Density • Time_Step
R-squared	0.83587	0.83982	0.83773	0.82429
3C	• Occupant_Density • Time_Step • Staff_Arriv_Depar	• Occupant_Density • Time_Step • Staff_Arriv_Depar • Manag_Arriv_Depar • Own_Stay_Duration	• Occupant_Density • Time_Step • Manag_Arriv_Depar	• Occupant_Density • Time_Step
R-squared	0.89596	0.89345	0.88437	0.86468
4A	• Occupant_Density • Time_Step • Lunch_Duration	• Occupant_Density • Time_Step • Staff_Arriv_Depar • Manag_Arriv_Depar • Admin_Arriv_Depar	• Occupant_Density • Time_Step • Staff_Arriv_Depar • Manag_Arriv_Depar	• Occupant_Density • Time_Step
R-squared	0.88695	0.89287	0.89161	0.88692
5A	• Occupant_Density • Time_Step • Lunch_Duration	• Occupant_Density • Time_Step • Manag_Arriv_Depar • Staff_Arriv_Depar	• Occupant_Density • Time_Step • Manag_Arriv_Depar • Staff_Arriv_Depar	• Occupant_Density • Time_Step • Lunch_Duration_Vari
R-squared	0.87708	0.88016	0.87937	0.87729
6A	• Occupant_Density • Time_Step • Lunch_Duration • Outdoor_Stay_Duration	• Occupant_Density • Time_Step • Manag_Arriv_Depar • Auxiliary_Stay_Duration	• Occupant_Density • Time_Step • Manag_Arriv_Depar • Auxiliary_Stay_Duration	• Occupant_Density • Time_Step • Admin_Room_Stay • Own_Stay_Duration
R-squared	0.9034	0.90748	0.90694	0.90336

Fans: This section focuses on the performance of fan energy consumption and presents the results of four feature selection methods to identify the significant occupancy parameters impacting fan's energy consumption in six climate zones, as presented in Table 8. As shown in the table, three methods of MVLinear, NCA, and LASSO consistently identified "Time-Step" as the primary significant variable across all six climate zones. Furthermore, "Occupant-Density" emerged as the second significant parameter, except in the "ReliefF" method, which is the opposite. According to Table 8, both MVLinear and LASSO feature selection methods presented better performances across all six climate zones in identifying the significant features influencing fan energy consumption.

These methods generated higher R-squared values compared to other methods. The selected features in the MVLinear and LASSO methods include "Time_Step", "Occupant_Density", "Staff_Arriv_Depar", "Admin_Arriv_Depar", and "Own_Stay_Duration", and, conversely, the ReliefF method presented limited effectiveness in identifying significant features. The higher performance of MVLinear and LASSO in our study points to their robustness in handling diverse datasets and their capability in effectively isolating key factors that influence fan energy consumption. These methods prove particularly valuable in scenarios where predictive accuracy is important to developing energy management solutions that are both effective and scalable across different climate zones.

Table 5. Identified features for cooling energy performance across different climate zones.

Climate Zone	MVLinear	LASSO	NCA	ReliefF
2A	• Time_Step • Occupant_Density • Own_Stay_Duration	• Time_Step • Occupant_Density • Staff_Arriv_Depar • Lunch_Duration • Own_Stay_Duration	• Time_Step • Occupant_Density • Manag_Room_Stay • Lunch_Duration	• Occupant_Density • Time_Step • Own_Stay_Duration • Meeting_Stay_Duration
R-squared	0.85107	0.85221	0.85143	0.84634
3B	• Time_Step • Occupant_Density • Staff_Arriv_Depar • Own_Stay_Duration	• Time_Step • Occupant_Density • Staff_Arriv_Depar • Lunch_Duration • Own_Stay_Duration	• Occupant_Density • Time_Step • Staff_Arriv_Depar	• Occupant_Density • Time_Step
R-squared	0.83984	0.84153	0.84052	0.83586
3C	• Time_Step • Occupant_Density • Staff_Arriv_Depar	• Time_Step • Occupant_Density • Staff_Arriv_Depar • Lunch_Duration • Own_Stay_Duration	• Occupant_Density • Time_Step • Manag_Room_Stay • Staff_Arriv_Depar	• Occupant_Density • Time_Step • Own_Stay_Duration
R-squared	0.82795	0.83965	0.83164	0.82541
4A	• Time_Step • Occupant_Density • Staff_Arriv_Depar	• Time_Step • Occupant_Density • Staff_Arriv_Depar • Lunch_Duration	• Occupant_Density • Time_Step • Staff_Arriv_Depar • Lunch_Duration	• Occupant_Density • Time_Step • Own_Stay_Duration • Auxiliary_Stay_Duration
R-squared	0.81269	0.8148	0.81346	0.81127
5A	• Occupant_Density • Time_Step • Own_Stay_Duration	• Time_Step • Occupant_Density • Lunch_Duration • Own_Stay_Duration	• Occupant_Density • Time_Step • Staff_Arriv_Depar • Lunch_Duration • Own_Stay_Duration	• Occupant_Density • Time_Step • Own_Stay_Duration • Auxiliary_Stay_Duration
R-squared	0.78551	0.78816	0.78816	0.78456
6A	• Occupant_Density • Time_Step • Staff_Arriv_Depar • Own_Stay_Duration	• Time_Step • Occupant_Density • Lunch_Duration • Staff_Arriv_Depar • Own_Stay_Duration	• Time_Step • Occupant_Density • Staff_Arriv_Depar • Lunch_Duration • Own_Stay_Duration	• Occupant_Density • Time_Step • Own_Stay_Duration
R-squared	0.77699	0.77968	0.77968	0.77546

This analysis focuses on influential occupancy parameters affecting fan energy consumption in small office buildings across different climate zones and highlights the LASSO and MVLinear feature selection methods as preferred approaches for identifying the features that have a significant impact on fan energy consumption. By systematically identifying these key parameters, our approach aids facility managers and designers in implementing precision-driven energy conservation measures. Specifically, understanding the detailed impacts of occupancy parameters such as "Occupant-Density" and "Time-Step" allows for the optimization of HVAC operations and other energy systems to better match actual building use patterns, thus avoiding both overuse and underuse of energy resources.

Table 6. Identified features for lighting energy performance across different climate zones.

Climate Zone	MVLinear	LASSO	NCA	ReliefF
2A	• Occupant_Density • Manag_Arriv_Depar • Arriv_Depar_Vari • Time_Step	• Occupant_Density • Staff_Arriv_Depar • Manag_Arriv_Depar • Arriv_Depar_Vari • Time_Step	• Occupant_Density • Arriv_Depar_Vari • Staff_Arriv_Depar • Manag_Arriv_Depar • Time_Step	• Occupant_Density • Time_Step • Manag_Room_Stay
R-squared	0.83369	0.84654	0.84654	0.81792
3B	• Occupant_Density • Manag_Arriv_Depar • Arriv_Depar_Vari • Time_Step • Lunch_Start_Vari	• Occupant_Density • Staff_Arriv_Depar • Manag_Arriv_Depar • Arriv_Depar_Vari • Time_Step	• Occupant_Density • Arriv_Depar_Vari • Staff_Arriv_Depar • Manag_Arriv_Depar	• Occupant_Density • Time_Step • Manag_Room_Stay
R-squared	0.83252	0.84592	0.84592	0.81617
3C	• Occupant_Density • Manag_Arriv_Depar • Arriv_Depar_Vari • Time_Step • Lunch_Start_Vari	• Occupant_Density • Staff_Arriv_Depar • Manag_Arriv_Depar • Arriv_Depar_Vari	• Occupant_Density • Arriv_Depar_Vari • Staff_Arriv_Depar • Manag_Arriv_Depar • Admin_Arriv_Depar	• Occupant_Density • Time_Step • Manag_Room_Stay
R-squared	0.83225	0.846712	0.845331	0.81572
4A	• Occupant_Density • Manag_Arriv_Depar • Arriv_Depar_Vari • Time_Step • Lunch_Start_Vari	• Occupant_Density • Staff_Arriv_Depar • Manag_Arriv_Depar • Arriv_Depar_Vari	• Occupant_Density • Arriv_Depar_Vari • Staff_Arriv_Depar • Manag_Arriv_Depar	• Occupant_Density • Time_Step • Manag_Room_Stay • Auxiliary_Stay_Duration
R-squared	0.83203	0.84544	0.84544	0.81546
5A	• Occupant_Density • Manag_Arriv_Depar • Arriv_Depar_Vari • Time_Step • Lunch_Start_Vari	• Occupant_Density • Staff_Arriv_Depar • Manag_Arriv_Depar • Arriv_Depar_Vari	• Occupant_Density • Arriv_Depar_Vari • Staff_Arriv_Depar • Manag_Arriv_Depar • Admin_Arriv_Depar	• Occupant_Density • Time_Step • Manag_Room_Stay • Auxiliary_Stay_Duration
R-squared	0.83351	0.84682	0.84647	0.81456
6A	• Occupant_Density • Manag_Arriv_Depar • Arriv_Depar_Vari • Time_Step • Lunch_Start_Vari	• Occupant_Density • Staff_Arriv_Depar • Manag_Arriv_Depar • Arriv_Depar_Vari	• Occupant_Density • Arriv_Depar_Vari • Staff_Arriv_Depar • Manag_Arriv_Depar	• Occupant_Density • Time_Step • Manag_Room_Stay • Auxiliary_Stay_Duration • Outdoor_Stay_Duration
R-squared	0.83316	0.84647	0.84635	0.81632

The results of the feature selection analyses highlight critical occupancy parameters such as “Occupant-Density” and “Time-Step”, which are consistently significant across multiple climate zones. This consistency across zones not only validates our feature selection methods but also reinforces the importance of these parameters in the energy efficiency profiling of office buildings. These parameters offer valuable insights into the design and operation of energy-efficient buildings, tailored to the unique characteristics of each climate zone. For instance, the consistent significance of occupant density suggests that strategies focused on optimizing space usage can lead to substantial energy savings. This is particularly relevant in denser office environments where the effective management of space and occupant schedules can reduce unnecessary energy expenditure during peak and off-peak hours. Moreover, the distinction in significant parameters across climate zones allows for the development of region-specific energy management strategies. In colder climates, during high heating demands, understanding the impact of occupancy timings can help in better scheduling of heating systems to align with actual office hours, thus avoiding wastage. Similarly, in hotter regions, cooling systems can be optimized based on detailed occupancy schedules to ensure that energy is not wasted cooling unoccupied spaces.

Table 7. Identified features for equipment energy performance across different climate zones.

Climate Zone	MVLinear	LASSO	NCA	ReliefF
2A	• Occupant_Density • Manag_Arriv_Depar • Arriv_Depar_Vari • Lunch_Start_Vari	• Occupant_Density • Staff_Arriv_Depar • Manag_Arriv_Depar • Arriv_Depar_Vari • Time_Step	• Occupant_Density • Arriv_Depar_Vari • Staff_Arriv_Depar • Manag_Arriv_Depar	• Occupant_Density • Time_Step • Manag_Room_Stay • Auxiliary_Stay_Duration • Outdoor_Stay_Duration
R-squared	0.83041	0.84465	0.84444	0.8087
3B	• Occupant_Density • Manag_Arriv_Depar • Lunch_Start_Vari	• Occupant_Density • Staff_Arriv_Depar • Manag_Arriv_Depar • Arriv_Depar_Vari • Time_Step	• Occupant_Density • Arriv_Depar_Vari • Staff_Arriv_Depar • Manag_Arriv_Depar	• Occupant_Density • Time_Step • Outdoor_Stay_Duration
R-squared	0.83041	0.84465	0.84444	0.8087
3C	• Occupant_Density • Manag_Arriv_Depar • Arriv_Depar_Vari • Lunch_Start_Vari	• Occupant_Density • Staff_Arriv_Depar • Manag_Arriv_Depar • Arriv_Depar_Vari • Time_Step	• Occupant_Density • Arriv_Depar_Vari • Staff_Arriv_Depar • Manag_Arriv_Depar	• Occupant_Density • Time_Step • Manag_Room_Stay • Auxiliary_Stay_Duration • Outdoor_Stay_Duration
R-squared	0.83041	0.84465	0.84444	0.8087
4A	• Occupant_Density • Manag_Arriv_Depar • Arriv_Depar_Vari • Lunch_Start_Vari	• Occupant_Density • Staff_Arriv_Depar • Manag_Arriv_Depar • Arriv_Depar_Vari • Time_Step	• Occupant_Density • Staff_Arriv_Depar • Manag_Arriv_Depar	• Occupant_Density • Time_Step • Manag_Room_Stay • Auxiliary_Stay_Duration • Outdoor_Stay_Duration
R-squared	0.83041	0.84465	0.84444	0.8087
5A	• Occupant_Density • Manag_Arriv_Depar • Arriv_Depar_Vari • Lunch_Start_Vari	• Occupant_Density • Staff_Arriv_Depar • Manag_Arriv_Depar • Arriv_Depar_Vari • Time_Step	• Occupant_Density • Arriv_Depar_Vari • Staff_Arriv_Depar • Manag_Arriv_Depar	• Occupant_Density • Time_Step • Manag_Room_Stay

Table 7. *Cont.*

Climate Zone	MVLinear	LASSO	NCA	ReliefF
R-squared	0.83041	0.84465	0.84444	0.8087
6A	• Occupant_Density • Manag_Arriv_Depar • Lunch_Start_Vari	• Occupant_Density • Staff_Arriv_Depar • Manag_Arriv_Depar • Arriv_Depar_Vari • Time_Step	• Occupant_Density • Arriv_Depar_Vari • Staff_Arriv_Depar • Manag_Arriv_Depar	• Occupant_Density • Time_Step • Manag_Room_Stay • Auxiliary_Stay_Duration • Outdoor_Stay_Duration
R-squared	0.83041	0.84465	0.84444	0.8087

In this study, we applied a variety of feature selection methods to enhance the accuracy and efficiency of predicting building energy consumption. This approach is supported by recent research, including the work by Henriques et al. [55], who utilized advanced feature-selection methods like recursive feature elimination and random forests to uncover atypical energy-consumption patterns in households. Their findings demonstrate the effectiveness of comprehensive feature selection in identifying significant predictors of energy usage, which aligns with our methodology for improving model performance. Additionally, the application of feature selection for support vector regression models, as discussed in a study by Zhao and Magoulès [56], highlights the reduction in model complexity and the improvement in predictive accuracy for building energy consumption. This reference further validates our choice of feature selection methods by illustrating their benefits in similar applications, thereby highlighting the relevance of our methodological choices in the context of current research trends.

Table 8. Identified features for fan energy performance across different climate zones.

Climate Zone	MVLinear	LASSO	NCA	ReliefF
2A	• Time_Step • Occupant_Density • Staff_Arriv_Depar • Admin_Arriv_Depar • Own_Stay_Duration	• Time_Step • Occupant_Density • Staff_Arriv_Depar • Admin_Arriv_Depar • Own_Stay_Duration	• Time_Step • Occupant_Density • Staff_Arriv_Depar • Auxiliary_Stay_Duration • Manag_Room_Stay	• Occupant_Density • Time_Step • Manag_Room_Stay • Own_Stay_Duration
R-squared	0.83267	0.83267	0.83092	0.81171
3B	• Time_Step • Occupant_Density • Staff_Arriv_Depar • Admin_Arriv_Depar • Own_Stay_Duration	• Time_Step • Occupant_Density • Staff_Arriv_Depar • Admin_Arriv_Depar • Own_Stay_Duration	• Time_Step • Occupant_Density • Staff_Arriv_Depar • Auxiliary_Stay_Duration • Manag_Room_Stay	• Occupant_Density • Time_Step • Manag_Room_Stay • Own_Stay_Duration
R-squared	0.82586	0.82586	0.82374	0.80272
3C	• Time_Step • Occupant_Density • Staff_Arriv_Depar • Admin_Arriv_Depar • Own_Stay_Duration	• Time_Step • Occupant_Density • Staff_Arriv_Depar • Admin_Arriv_Depar • Own_Stay_Duration	• Time_Step • Occupant_Density • Staff_Arriv_Depar • Auxiliary_Stay_Duration • Manag_Room_Stay	• Occupant_Density • Time_Step • Own_Stay_Duration

Table 8. *Cont.*

Climate Zone	MVLinear	LASSO	NCA	ReliefF
R-squared	0.82016	0.82016	0.81748	0.80914
4A	• Time_Step • Occupant_Density • Staff_Arriv_Depar • Admin_Arriv_Depar • Own_Stay_Duration	• Time_Step • Occupant_Density • Staff_Arriv_Depar • Admin_Arriv_Depar • Own_Stay_Duration	• Time_Step • Occupant_Density • Staff_Arriv_Depar • Auxiliary_Stay_Duration • Admin_Arriv_Depar	• Occupant_Density • Time_Step
R-squared	0.81928	0.81928	0.81827	0.79416
5A	• Time_Step • Occupant_Density • Staff_Arriv_Depar • Own_Stay_Duration	• Time_Step • Occupant_Density • Staff_Arriv_Depar • Admin_Arriv_Depar • Own_Stay_Duration	• Time_Step • Staff_Arriv_Depar • Occupant_Density • Auxiliary_Stay_Duration • Admin_Arriv_Depar	• Occupant_Density • Time_Step • Own_Stay_Duration
R-squared	0.81505	0.81505	0.81381	0.78929
6A	• Time_Step • Occupant_Density • Staff_Arriv_Depar • Admin_Arriv_Depar	• Time_Step • Occupant_Density • Staff_Arriv_Depar • Admin_Arriv_Depar • Own_Stay_Duration	• Time_Step • Staff_Arriv_Depar • Occupant_Density • Auxiliary_Stay_Duration	• Occupant_Density • Time_Step • Own_Stay_Duration
R-squared	0.81444	0.81444	0.81307	0.78684

These insights not only enhance the ability to create more adaptive and intelligent building management systems but also support the development of policies and standards that promote energy efficiency. By integrating these findings into the regulatory frameworks, it is possible to set more realistic and attainable energy usage benchmarks that reflect the real-world conditions of office buildings. This approach can lead to more sustainable energy practices and significant reductions in operational costs, contributing to broader environmental and economic benefits.

4. Conclusions

This study utilizes a nationwide energy simulation to analyze the impact of occupancy parameters on building energy performance across different US climate zones. Dynamic occupancy schedules are generated based on identified occupancy parameters using an ABM. The generated schedule, along with the DOE small office prototype model, are integrated into the BPS tool (i.e., EnergyPlus). The simulation results provide a dataset of occupancy parameters and building energy performance in various climate zones. This study employs sensitivity analysis and feature selection methods to assess the influence of occupancy parameters on the energy consumption of various building energy systems (i.e., heating, cooling, lighting, equipment, and fans).

The energy performance simulation across 16 climate zones revealed that buildings located in cool marine, mixed marine, and warm marine climate zones had lower total energy consumption compared to other zones. This lower energy consumption can be attributed to the moderate and stable temperatures characteristic of marine climates. On the other hand, buildings in hot–humid climate zones demonstrated significantly higher energy consumption, primarily due to the high demand for cooling and air conditioning systems. As a result of the sensitivity analysis, heating and cooling energy consumption were found to be sensitive to climate zones. However, lighting and equipment energy consumption remained relatively constant across climate zones, due to their dependence on occupancy presence and specific activities rather than climate zones. Fan energy consumption exhibited minimal variation across climate zones but was slightly higher in dry climates to create a perceived cooling sensation for occupants.

The sensitivity analysis of occupancy parameters revealed the impact of occupant density, arrival and departure times, and occupants' stay-time and simulation time-step on energy consumption. For instance, reducing occupancy density from 0.2 person/m^2 to 0.05 person/m^2 if possible (e.g., transition to a hybrid work environment) can result in annual energy savings ranging from 9000 to 12,000 kWh. In other words, an underoccupied office building can reduce energy consumption by 60 percent. Modifying arrival and departure times, particularly by starting work earlier at 6:30 a.m. instead of 8:00 a.m., resulted in lower cooling and fan energy consumption due to taking advantage of cooler temperatures. Also, in hot climate zones, by shifting the arrival and departure time to the early morning, an annual energy savings of up to 3000 kWh can be achieved. In addition, changing the duration of occupants' stay-time in their own office from 30 to 90 min had a minimal impact on building energy consumption across all climate zones, with only slight changes observed in cooling and fan energy consumption. Moreover, the simulation time-step size influenced the energy performance of cooling systems and fans, with smaller time-steps resulting in higher energy consumption for these systems. In other words, changing the simulation time-step from 20 min to 5 min can result in a 50% discrepancy, which is around 12,000 to 15,000 kWh in annual cooling and fan energy consumption.

Feature selection methods applied in this study effectively identified significant occupancy parameters impacting energy consumption, with "Occupant Density" consistently identified as the primary significant variable across all climate zones. Other variables such as "staff's arrival and departure time", "stay-time in own office", "lunch break time", and "simulation time-step" were also found to have a significant impact on the simulation results. These findings offer valuable insights into the influential occupancy parameters for different small offices across various climate zones, underscoring the importance of tailoring occupancy schedules to enhance energy efficiency. This study provides practical guidance that can be directly applied to optimize energy consumption and achieve significant energy savings in small office settings with different weather conditions.

The novel insights gained from this research highlight the critical role of detailed occupancy parameter analysis in understanding and optimizing building energy performance across diverse climate conditions. By employing advanced modeling techniques and comprehensive sensitivity analyses, this study improves our knowledge of how occupancy behaviors impact building energy consumption and provides a pathway to more effective energy management strategies tailored to specific climatic and operational contexts. Overall, this study not only clarifies the effects of various occupancy parameters on energy consumption but also underscores the importance of detailed modeling in crafting energy-efficient buildings. The findings emphasize the necessity of considering both the micro (occupancy behavior) and macro (climate influences) elements in building energy assessments to foster sustainability in the built environment. This research also provides actionable insights that could significantly refine current building energy standards and practices. Specifically, the findings suggest that ASHRAE standards could be updated to include more nuanced models of occupant behavior, particularly in relation to occupant density and dynamic occupancy schedules. We recommend that ASHRAE Standard 90.1 incorporate specific guidelines on adjusting HVAC operation based on real-time occupancy data, which our research shows can lead to substantial energy savings.

One limitation of this study is the focus solely on small office buildings, neglecting the analysis of other types of office buildings prevalent in the US. Additionally, the study uses a single building model for all 16 climate zones, which may limit the reflection of specific architectural and environmental adaptations typically required for optimal energy efficiency in diverse climates. This narrow scope restricts the generalizability of our findings to broader office building contexts. In addition, while we extensively investigated the influence of occupancy parameters on building energy consumption, our analysis did not encompass the impact of other factors such as heating, cooling, lighting, and energy control strategies. Moreover, the assumption that subspaces within each thermal zone maintain similar thermal conditions may not fully capture the variability in actual office environments, where different uses or solar exposures can affect zone-specific energy demands. Future research could explore these aspects comprehensively to provide a more holistic understanding of energy efficiency in diverse office building settings.

Author Contributions: Conceptualization, S.N. and A.J.; methodology, S.N. and AJ; software, S.N.; validation, S.N., A.J., S.V. and Z.P.; formal analysis, S.N.; investigation, S.N.; resources, S.N.; data curation, S.N.; writing—original draft preparation, S.N., S.V. and A.J.; writing—review and editing, S.N., S.V., A.J. and Z.P.; visualization, S.N. and S.V.; supervision, A.J. and Z.P.; project administration, A.J. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: The data presented in this study are available on request from the corresponding author.

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

References

1. Alazazmeh, A.; Asif, M. Commercial building retrofitting: Assessment of improvements in energy performance and indoor air quality. *Case Stud. Therm. Eng.* **2021**, *26*, 100946. [CrossRef]
2. Li, J.; Yu, Z.J.; Haghighat, F.; Zhang, G. Development and improvement of occupant behavior models towards realistic building performance simulation: A review. *Sustain. Cities Soc.* **2019**, *50*, 101685. [CrossRef]
3. Vassiljeva, K.; Matson, M.; Ferrantelli, A.; Petlenkov, E.; Thalfeldt, M.; Belikov, J. Data-Driven Occupancy Profile Identification and Application to the Ventilation Schedule in a School Building. *Energies* **2024**, *17*, 3080. [CrossRef]
4. Ouf, M.M.; O'Brien, W.; Gunay, H.B. Improving occupant-related features in building performance simulation tools. *Build. Simul.* **2018**, *11*, 803–817. [CrossRef]
5. Yan, D.; O'Brien, W.; Hong, T.; Feng, X.; Gunay, H.B.; Tahmasebi, F.; Mahdavi, A. Occupant behavior modeling for building performance simulation: Current state and future challenges. *Energy Build.* **2015**, *107*, 264–278. [CrossRef]
6. Yang, J.; Santamouris, M.; Lee, S.E. Review of occupancy sensing systems and occupancy modeling methodologies for the application in institutional buildings. *Energy Build.* **2016**, *121*, 344–349. [CrossRef]
7. Norouziasl, S.; Jafari, A.; Zhu, Y. Modeling and simulation of energy-related human-building interaction: A systematic review. *J. Build. Eng.* **2021**, *44*, 102928. [CrossRef]
8. Bäcklund, K.; Molinari, M.; Lundqvist, P.; Palm, B.; Occupants, B. Their Behavior and the Resulting Impact on Energy Use in Campus Buildings: A Literature Review with Focus on Smart Building Systems. *Energies* **2023**, *16*, 6104. [CrossRef]
9. Vosoughkhosravi, S.; Jafari, A. American Time Use Survey in Modeling Occupant Behavior: A Systematic Review. In *Computing in Civil Engineering 2023*; American Society of Civil Engineers: Corvallis, OR, USA, 2024; pp. 77–84. [CrossRef]
10. Chen, Y.; Hong, T.; Luo, X. An agent-based stochastic Occupancy Simulator. *Build. Simul.* **2018**, *11*, 37–49. [CrossRef]
11. Chen, Y.; Luo, X.; Hong, T.; Chen, Y.; Luo, X.; Hong, T. *An Agent-Based Occupancy Simulator for Building Performance Simulation*; Lawrence Berkeley National Laboratory; LBL Publications: Berkeley, CA, USA, 2016; Available online: https://escholarship.org/uc/item/0047c6c3 (accessed on 1 July 2024).
12. Putra, H.C.; Andrews, C.J.; Senick, J.A. An agent-based model of building occupant behavior during load shedding. *Build. Simul.* **2017**, *10*, 845–859. [CrossRef]
13. Jia, M.; Srinivasan, R.; Ries, R.; Bharathy, G.; Weyer, N. Investigating the Impact of Actual and Modeled Occupant Behavior Information Input to Building Performance Simulation. *Buildings* **2021**, *11*, 32. [CrossRef]
14. Parys, W.; Saelens, D.; Hens, H. Coupling of dynamic building simulation with stochastic modelling of occupant behaviour in offices—A review-based integrated methodology. *J. Build. Perform. Simul.* **2011**, *4*, 339–358. [CrossRef]
15. Almeida, L.; Tam, V.W.Y.; Le, K.N.; She, Y. Effects of occupant behaviour on energy performance in buildings: A green and non-green building comparison. *ECAM* **2020**, *27*, 1939–1962. [CrossRef]
16. Gunay, H.B.; O'Brien, W.; Beausoleil-Morrison, I. Development of an occupancy learning algorithm for terminal heating and cooling units. *Build. Environ.* **2015**, *93*, 71–85. [CrossRef]
17. Li, N.; Calis, G.; Becerik-Gerber, B. Measuring and monitoring occupancy with an RFID based system for demand-driven HVAC operations. *Autom. Constr.* **2012**, *24*, 89–99. [CrossRef]
18. Chen, Z.; Xu, J.; Soh, Y.C. Modeling regular occupancy in commercial buildings using stochastic models. *Energy Build.* **2015**, *103*, 216–223. [CrossRef]
19. Page, J.; Robinson, D.; Morel, N.; Scartezzini, J.L. A generalised stochastic model for the simulation of occupant presence. *Energy Build.* **2008**, *40*, 83–98. [CrossRef]

20. Oldewurtel, F.; Sturzenegger, D.; Morari, M. Importance of occupancy information for building climate control. *Appl. Energy* **2013**, *101*, 521–532. [CrossRef]
21. Rafsanjani, H.N.; Ahn, C.R.; Chen, J. Linking building energy consumption with occupants' energy-consuming behaviors in commercial buildings: Non-intrusive occupant load monitoring (NIOLM). *Energy Build.* **2018**, *172*, 317–327. [CrossRef]
22. Erickson, V.L.; Lin, Y.; Kamthe, A.; Brahme, R.; Surana, A.; Cerpa, A.E.; Sohn, M.D.; Narayanan, S. Energy efficient building environment control strategies using real-time occupancy measurements. In *BuildSys '09: Proceedings of the First ACM Workshop on Embedded Sensing Systems for Energy-Efficiency in Buildings*; Association for Computing Machinery: New York, NY, USA, 2009; pp. 19–24. [CrossRef]
23. Banfi, A.; Ferrando, M.; Li, P.; Shi, X.; Causone, F. Integrating Occupant Behaviour into Urban-Building Energy Modelling: A Review of Current Practices and Challenges. *Energies* **2024**, *17*, 4400. [CrossRef]
24. Dziedzic, J.W.; Yan, D.; Sun, H.; Novakovic, V. Building occupant transient agent-based model–Movement module. *Appl. Energy* **2020**, *261*, 114417. [CrossRef]
25. Thornton, B.A.; Rosenberg, M.I.; Richman, E.E.; Wang, W.; Xie, Y.; Zhang, J.; Cho, H.; Mendon, V.V.; Athalye, R.A.; Liu, B. *Achieving the 30% Goal: Energy and Cost Savings Analysis of ASHRAE Standard 90.1-2010 (No. PNNL-20405)*; Pacific Northwest National Lab. (PNNL): Richland, WA, USA, 2011. [CrossRef]
26. Goel, S.; Rosenberg, M.; Athalye, R.; Xie, Y.; Wang, W.; Hart, R.; Zhang, J.; Mendon, V. *Enhancements to ASHRAE Standard 90.1 Prototype Building Models (No. PNNL-23269)*; Pacific Northwest National Lab. (PNNL): Richland, WA, USA, 2014.
27. Bae, Y.; Yoon, Y.; Malhotra, M.; Jung, S. *Prototype College Building Energy Model: Building Characteristics and Energy Simulation Results (No. ORNL/TM-2022/2532)*; Oak Ridge National Lab. (ORNL): Oak Ridge, TN, USA, 2022. Available online: https://www.osti.gov/biblio/1878673 (accessed on 1 July 2024).
28. Norouziasl, S.; Jafari, A.; Wang, C. An agent-based simulation of occupancy schedule in office buildings. *Build. Environ.* **2020**, *186*, 107352. [CrossRef]
29. Luo, X.; Lam, K.P.; Chen, Y.; Hong, T. Performance evaluation of an agent-based occupancy simulation model. *Build. Environ.* **2017**, *115*, 42–53. [CrossRef]
30. Alonso, M.J.; Dols, W.S.; Mathisen, H.M. Using Co-simulation between EnergyPlus and CONTAM to evaluate recirculation-based, demand-controlled ventilation strategies in an office building. *Build. Environ.* **2022**, *211*, 108737. [CrossRef]
31. Chen, Y.; Liang, X.; Hong, T.; Luo, X. Simulation and visualization of energy-related occupant behavior in office buildings. *Build. Simul.* **2017**, *10*, 785–798. [CrossRef]
32. Hernandez, H.; Ochoa, S. Adaptive Step-size One-at-a-time (OAT) Optimization. *ForsChem Res. Rep.* **2022**, *7*, 1–44. [CrossRef]
33. Standard 90.1. Available online: https://www.ashrae.org/technical-resources/bookstore/standard-90-1 (accessed on 1 July 2024).
34. EnergyPlusTM Input Output Reference. 2024. Available online: https://energyplus.net/assets/nrel_custom/pdfs/pdfs_v24.1.0/InputOutputReference.pdf (accessed on 1 July 2024).
35. Blight, T.S.; Coley, D.A. Sensitivity analysis of the effect of occupant behaviour on the energy consumption of passive house dwellings. *Energy Build.* **2013**, *66*, 183–192. [CrossRef]
36. Commercial Reference Buildings | Department of Energy. Available online: https://www.energy.gov/eere/buildings/commercial-reference-buildings (accessed on 1 July 2024).
37. Zhang, L.; Wen, J. A systematic feature selection procedure for short-term data-driven building energy forecasting model development. *Energy Build.* **2019**, *183*, 428–442. [CrossRef]
38. Leung, M.C.; Tse, N.C.F.; Lai, L.L.; Chow, T.T. The use of occupancy space electrical power demand in building cooling load prediction. *Energy Build.* **2012**, *55*, 151–163. [CrossRef]
39. Murrieum, M.; Jafari, A.; Akhavian, R. Building Energy Performance Prediction Using Machine Learning: A Data-Driven Decision-Making Framework for Energy Retrofits. In *Construction Research Congress 2020: Computer Applications—Selected Papers from the Construction Research Congress 2020*; American Society of Civil Engineers (ASCE): Reston, VA, USA, 2020; pp. 436–447. [CrossRef]
40. Kusiak, A.; Li, M.; Zhang, Z. A data-driven approach for steam load prediction in buildings. *Appl. Energy* **2010**, *87*, 925–933. [CrossRef]
41. Kapetanakis, D.S.; Mangina, E.; Finn, D.P. Input variable selection for thermal load predictive models of commercial buildings. *Energy Build.* **2017**, *137*, 13–26. [CrossRef]
42. Asadi, S.; Amiri, S.S.; Mottahedi, M. On the development of multi-linear regression analysis to assess energy consumption in the early stages of building design. *Energy Build.* **2014**, *85*, 246–255. [CrossRef]
43. Jain, R.K.; Damoulas, T.; Kontokosta, C.E. Towards Data-Driven Energy Consumption Forecasting of Multi-Family Residential Buildings: Feature Selection via The Lasso. In *Computing in Civil and Building Engineering—Proceedings of the 2014 International Conference on Computing in Civil and Building Engineering*; American Society of Civil Engineers: Reston, VA, USA, 2014; pp. 1675–1682. [CrossRef]
44. Yang, W.; Wang, K.; Zuo, W. Neighborhood Component Feature Selection for High-Dimensional Data Detection and Segmentation View project Human Recognition View project Neighborhood Component Feature Selection for High-Dimensional Data. *J. Comput.* **2012**, *7*, 161–168. [CrossRef]
45. Kowshalya, A.M.; Madhumathi, R.; Gopika, N. Correlation Based Feature Selection Algorithms for Varying Datasets of Different Dimensionality. *Wirel. Pers. Commun.* **2019**, *108*, 1977–1993. [CrossRef]

46. Zhang, Y.; Ding, C.; Li, T. Gene selection algorithm by combining reliefF and mRMR. *BMC Genomics* **2008**, *9*, S27. [CrossRef]
47. Chen, Y.; Guo, M.; Chen, Z.; Chen, Z.; Ji, Y. Physical energy and data-driven models in building energy prediction: A review. *Energy Rep.* **2022**, *8*, 2656–2671. [CrossRef]
48. Sun, Y.; Haghighat, F.; Fung, B.C.M. A review of the-state-of-the-art in data-driven approaches for building energy prediction. *Energy Build.* **2020**, *221*, 110022. [CrossRef]
49. Energy Standard for Buildings Except Low-Rise Residential Buildings. 2010. Available online: https://www.ashrae.org/technical-resources/standards-and-guidelines (accessed on 1 July 2024).
50. Chen, Y.; Li, M.; Xiong, M.; Cao, J.; Li, J. Future Climate Change on Energy Consumption of Office Buildings in Different Climate Zones of China. *Pol. J. Environ. Stud.* **2018**, *27*, 45–53. [CrossRef]
51. Meng, F.; Li, M.; Cao, J.; Li, J.; Xiong, M.; Feng, X.; Ren, G. The effects of climate change on heating energy consumption of office buildings in different climate zones in China. *Theor. Appl. Clim.* **2018**, *133*, 521–530. [CrossRef]
52. Dong, Z.; Zhao, K.; Hua, Y.; Xue, Y.; Ge, J. Impact of occupants' behaviour on energy consumption and corresponding strategies in office buildings. *IOP Conf. Ser. Earth Environ. Sci.* **2019**, *294*, 012076. [CrossRef]
53. Gu, J.; Xu, P.; Ji, Y. A Fast Method for Calculating the Impact of Occupancy on Commercial Building Energy Consumption. *Buildings* **2023**, *13*, 567. [CrossRef]
54. Prasetiyowati, M.I.; Maulidevi, N.U.; Surendro, K. Determining threshold value on information gain feature selection to increase speed and prediction accuracy of random forest. *J. Big Data* **2021**, *8*, 84. [CrossRef]
55. Henriques, L.; Lima, F.P.; Castro, C. Combining Advanced Feature-Selection Methods to Uncover Atypical Energy-Consumption Patterns. *Future Internet* **2024**, *16*, 229. [CrossRef]
56. Zhao, H.; Magoules, F. Feature selection for support vector regression in the application of building energy prediction. In Proceedings of the 2011 IEEE 9th International Symposium on Applied Machine Intelligence and Informatics (SAMI), Smolenice, Slovakia, 27–29 January 2011; pp. 219–223. [CrossRef]