buildings

Article

Daily Carbon Assessment Framework: Towards Near Real-Time Building Carbon Emission Benchmarking for Operative and Design Insights

Mingyu Zhu * and Philip James

School of Engineering, Newcastle University, Newcastle upon Tyne NE1 7RX, UK; philip.james@ncl.ac.uk
* Correspondence: m.zhu14@newcastle.ac.uk

Abstract: The energy consumption and its related carbon emission of non-domestic complex buildings in an urban context are complicated due to their wide variety of functions and services. A detailed assessment of the carbon emission of such buildings can contribute to decision making for in-operation building management and schematic designs of future proposals. Concurrently, advances in smart meter data analytics and sensor-enabled operational data streams offer the opportunity to investigate this problem at a finer temporal resolution. This research developed a daily carbon emission benchmarking system of a mixed-use building in a UK university. The research period was set at an annual range from 1 January 2019 to 31 December 2019 and was segmented by strategic periods in line with the operation schedule of the building. The daily benchmark revealed the fluctuation of the building's energy consumption and associated carbon emissions. Based on this, a digital twin framework was developed to identify the possible time periods when the building is less carbon efficient and potential building characters that can lead to increased carbon emission in the operational stage compared with what originally expected at the design stage. We discuss how these insights can offer actionable knowledge for user groups such as asset managers and architects.

Keywords: carbon assessment; smart meter; digital twin

Citation: Zhu, M.; James, P. Daily Carbon Assessment Framework: Towards Near Real-Time Building Carbon Emission Benchmarking for Operative and Design Insights. *Buildings* **2022**, *12*, 1129. https://doi.org/10.3390/buildings12081129

Academic Editor: Xi Chen

Received: 16 June 2022
Accepted: 26 July 2022
Published: 29 July 2022

1. Introduction

Buildings account for a significant contributor to GHGs and account for 28% of global energy-related carbon dioxide emissions; they also have the largest energy consumption compared with other sectors in the urban environment as they are responsible for about 30% of final energy consumption [1,2]. Given the significant impact of carbon emissions of buildings, there is a substantial motivation to assess the carbon emissions in existing buildings. The insight acquired can support the adoption of retrofit actions, intelligent management systems and design guidelines for new proposals. An ideal assessment of the energy efficiency and associated carbon emissions of a building normally spans across its life cycle, from pre-construction prediction to operational performance review. For energy efficiency modelling, necessary methodologies at each stage are generally available and have been integrated into national or public evaluation programs and have therefore been used to estimate carbon emissions [3]. However, there is a need to develop frameworks to facilitate interactions between these modelling techniques for validation and calibration and therefore towards a more effective system for planning and management [3].

Concurrently, smart meter and IoT tools such as wireless sensor networks can provide data at the building level and at a time resolution of one hour or shorter related to the building's energy usage, occupants' behavior and surrounding environmental information [4]. They are appearing to have wide-spread impacts on the way that physical/engineering assets are managed in their life cycles [5]. They offer a unique opportunity to understand the complex energy dynamics of a building in operation with lower capital and labor costs. The emergence of such new data streams is enabling the new data-driven approaches to assess, compare and improve traditional benchmarking methods of a building's energy performance.

The recent endeavor of the urban digital twin (DT) enables a framework to host these data to create a dynamic digital representation to mimic buildings' real-world performance [6]. It offers a comprehensive approach to manage, plan, predict and demonstrate buildings [7]. For the urban DT, its vision is improving the monitoring, control and decision making by unitizing the physical, social and operational data acquired from different data sources and various temporal periods throughout the asset's life cycle. With the exploitation of digitalization technologies such as the BIM, sensor networks and GIS, the urban DT offers a coherent platform for the whole-life management of a single building or wider urban system. For instance, [8] explored the framework of DT providing increased visibility into cities' human–infrastructure interaction and predictive insights into a city's smarter performance and growth. The authors of [9] provided a practical investigation into the process of developing DT using real-time data from IoT technologies. Such research developed specific applications for enhanced collaboration, improved visualization and optimized workflow. Despite the urban DT capabilities and promises for improving the understanding of buildings throughout their life cycle, current research has focused mostly on the data integration and acquisition mechanism and asset and facility management system at the operational stage.

Given the increasing data availability related to building energy consumption and operational information, a digital twin paradigm offers a platform to create a framework for benchmarking a building's energy and carbon efficiency in further detail. Different energy modelling techniques can be integrated and compared on the platform. The insight gained from this can help the asset management team to improve operational plans. As for architects and consultants, it works as a strategic tool and provides knowledge-based feedback opportunities to improve the schematic design of future projects commissioned.

Based on the aforementioned discussion, this research aims to develop a digital twin framework to benchmark daily carbon efficiency of a single building based on its energy performance. The adoption of smart-meter and other near real-time data streams enables the development of energy modelling across temporally segmented periods of the in-operation building. It can also be compared with a building's pre-construction energy simulation result to reveal the formation of the gap between the actual and simulated performance. As a case study, we used an annual period of datasets at 15-min intervals collected from a UK university building and its energy supplier. The methods, experimental datasets and the results are explained in the following sections.

2. Background

2.1. Building Energy Modelling

The use of building energy modelling to estimate carbon emissions has been a hot topic. The major methods of building energy modelling can be categorized into two sets which are purely statistical/machine learning models, also known as 'top–down' approaches and physical-based simulation, also known as 'bottom–up' approaches.

The first one applies statistic or machine learning algorithms to standardize the consumption based on characters across the portfolio of buildings to assess their efficiency and predict the consumption under certain conditions. It is commonly used in the operational stage and can be performed on entire building pollution, given aggregate statistics or assumptions about the distribution of archetype buildings or on an individual building, given enough repeated measurements [10]. The authors of [11] developed benchmarks for supermarkets with central air conditioning systems considering nine variables in a regression analysis. The authors of [12] used a multiple regression model to predict the electricity consumption of an administration building of a higher education sector. The authors of [13] developed and validated a regression model for detecting building energy efficiency in the Spanish banking sector. The explanatory variables adopted in these studies are categorized and outlined in Table 1. Once the energy consumption of a building is predicted, its carbon emission can be estimated based on the usage amount and type of fuel. It can also be used to compare the effectiveness between different energy saving methods, such as installing thermal insulation and replacing inefficient appliances [14].

Table 1. Summarization of explanatory variables.

Study	Explanatory Variables
[11]	Building age Occupancy Chiller equipment Lighting equipment Light control
[12]	Ambient temperature Solar radiation Relative humidity Wind speed Occupancy
[13]	Winter climatic severity Summer climatic severity Office surface area Number of employees Glazed surface in facade HAVC installed power Office height Building age

The 'bottom–up' approach is a physical based method using theories of thermal diffusion and heat transfer to simulate energy usage and carbon emission. It requires both the geometry of the building which can be BIM, a point cloud or GIS shapefile, and its non-geometric properties, including equipment specification and usage/occupancy schedule. The field has flourished in recent years, leading to increasingly robust urban data streams that start from geographic information systems (GIS), light detection and ranging (LiDAR) and tax assessor databases and end in synthetic hourly building energy demand profiles for current and potential future conditions [15]. After the simulation is complete, the result is compared with performance criteria such as benchmarks or regulatory requirements to assess its efficiency [3]. The architects and policy makers can leverage simulation models to gain understanding of the energy and carbon trade-offs between different design options. This can be studied in conjunction with a full range of urban planning considerations and technical design aspects, such as land use, zoning, urban economics, transportation, standards, codes, safety and security

In practice, these two methods are often used to inform one another in a complementary way. For top–down methods, physical-based theory and assumption are necessary to determine the importance of variables, how they should be defined and combined, and in what functional model specification, with linear or nonlinear relationships, or distribution of random errors, and so on [16]. As for bottom–up methods, statistical methods are used for the evaluation of simulation models and the identification of key factors that reduce accuracy and confidence. The performance gap between predicted and actual consumption is predominately caused by operation practice, occupancy behavior and building system specification. Statistical analyses of these variables contribute to the calibration of simulation models and therefore reduce the impact of performance gaps.

2.2. Digital Twin for Building Energy and Carbon Assessment

Proposition and evolution of urban DTs are rooted in the concept of a smart city which originated in the 1980s to deal with the increased amount, type and complexity of data generated from urban activities. Cities, as adaptive complex systems, experience constant spatial–temporal flux with respect to individuals' activities. Accurate connection and reliable representation with physical counterparts are the foundation of a digital twin. With the development of BIM, IoT and other technologies, realistic 3D representation and real-time operational information of built assets are increasingly available. By utilizing these, DT can simulate the behavior of physical entities in the real environment with digital virtual models through virtual and real interactive feedback, data fusion analysis, and decision-making iteration optimization, thereby playing a role as a bridge connecting the physical model and the information model [17]. Although the attempt to develop full-scale urban DT remains relatively novel, a variety of researchers have developed a wide range of applications for building energy load analyses, forecasting and management by assessing the real-time energy data [18]. They examine the structural and behavioural determinants of residential electricity consumption [19], reveal the relationship between occupant behaviour and energy usage [20] and develop efficiency benchmarking methods across multiple metrics at a daily time scale [21]. The insights gained from these studies enable urban DTs to provide optimized outcomes for where to focus, which methods to use and whom to be concerned about for sustainable urban development and preservation in terms of carbon emissions [22].

Built on these findings, this research utilized BIM, energy supplier and sensor data as the foundation of the DT framework to assess carbon emissions. In the data modelling layer, the study first developed temporally segmented benchmarking methods based on the "top–down" energy model and evaluated the difference between in-operational emissions and "bottom–up" simulation results. Furthermore, it explored how these evaluations can offer actionable insight via the service layer of the DT framework, including detection of temporal periods with inefficient performance and identifying the major contributor to the gap between the two modelling methods.

3. Methods

3.1. Building Description

Adopting a case study approach, this research selected the Urban Science Building (USB) of Newcastle University as the object of study. The USB is located at the Helix campus in Newcastle upon Tyne, northeast England. It is a complex building with multiple functions and long opening hours. It includes teaching facilities such as computer clusters, seminar rooms and lecture theatres for the students and is also the home of several research groups. The choice is determined by the following factors:

1. The availability of data including electricity consumption, on-site weather and room occupancy at a 15-min interval.
2. The multi-functionality of the building is a representation of the urban dynamic at a wider context.
3. It is a relatively newly constructed building that opened in September 2017. It can provide insight into the current design technology and strategy adopted.

The total floor area of the USB is approximately 12,500 m^2 and is known as the demonstration of best practices for 'digitally enabled urban sustainability'. There are more than 4000 digital sensors and computing technology systems embedded throughout its structure. They provide intensive data to create and study the building's digital replica. The USB has a total of seven levels as shown in Figure 1. It houses 1200 students, 55 academic staff, 120 post-doctoral researchers and numerous visitors. The building is composed of two wings which are perpendicular to each other and joined by a southeast-facing atrium. The atrium is enveloped by a curtain wall façade and sedum roof with skylights allowing natural light to flood into the building. The east–west wing has seven levels and is mainly used as cafés, open-plan offices and seminar and meeting rooms. They receive sufficient daylight from the south-facing façade. The other five levels in the north–south wing mainly accommodate teaching spaces such as computer clusters and lecture rooms and research facilities such as laboratories. Generally, they require limited daylight and some need a mechanically controlled environment. There are two elevator groups and each of them serves each respective wing. The USB is also equipped with automatic double-layer sliding doors at the entrance to reduce heat loss during cold seasons and automated windows on the curtain wall of the atrium to assist the natural ventilation of lower levels.

As stated above, the architectural design provides an optimized solution to achieve energy efficiency with regards to the comfort and convenience of the users. It received the BREEAM Innovation Credit that is awarded to buildings beyond best practice in sustainability. Its electricity consumption during an annual period from 1 January 2019 to 31 December 2019 was used as experimental data.

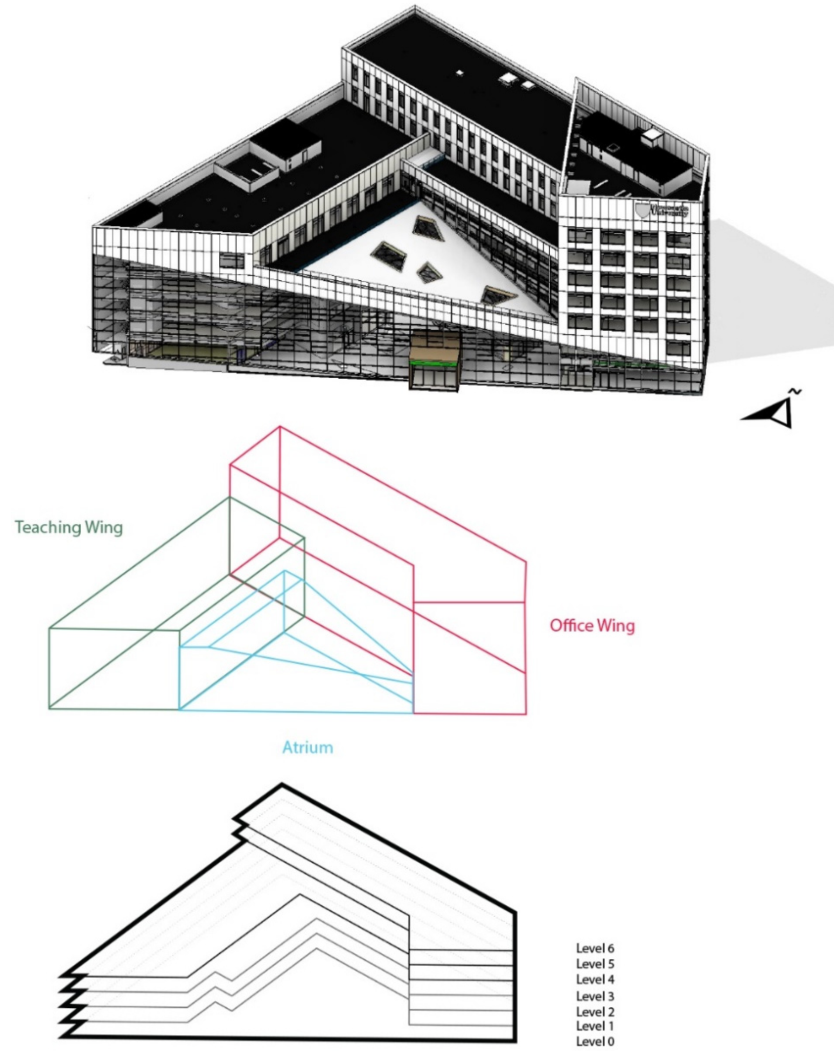

Figure 1. Urban Science Building (USB).

3.2. Digital Twin Framework

Using the insights gained from the literature review of an urban digital twin, the proposed DT framework is composed of three layers and is shown as Figure 2. The data acquisition layer is the foundation of the framework and collects necessary data. In this project, it includes smart meter data, weather and building occupation data collected from sensing devices, the building information model and energy supplier's carbon intensity data. The data modelling layer includes both top–down regression and bottom–up simulation methods, and the results from these models are further converted to benchmarks to demonstrate the building's carbon performance. Lastly, the service layer is the implementation layer that provide insights for asset managers and architects.

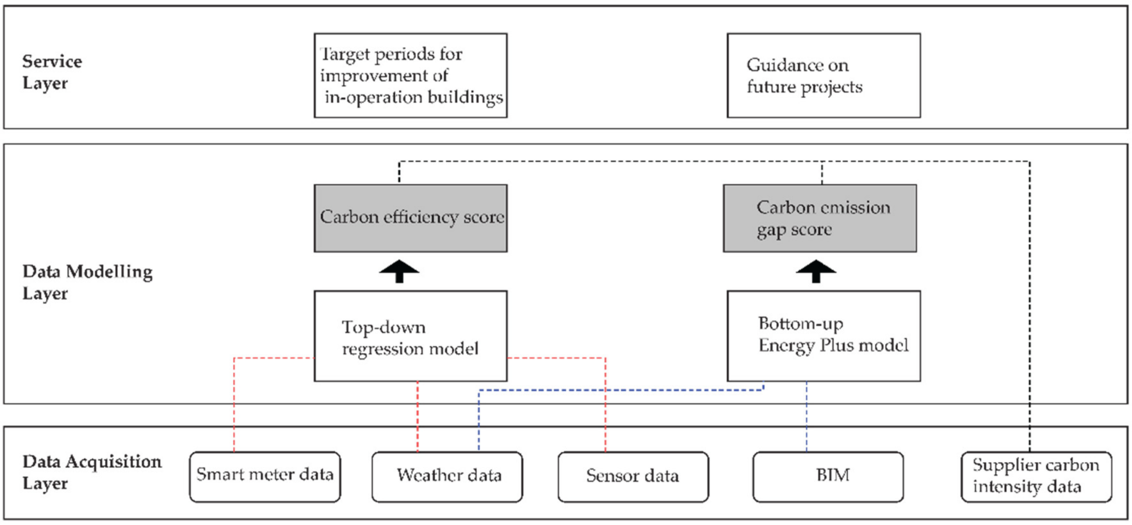

Figure 2. Digital twin framework.

3.3. Segmented Time Period

To acquire insights into the carbon emission under different operational conditions, the whole year is further segmented as follows: term occupied period (a), term unoccupied period (b), non-term occupied period (c) and non-term unoccupied period (d). An all-time period (e) is also adopted for comparison. Such periods are selected based on their alignment with the operational schedule of the building. The usage of equipment such as appliances and lighting is presumably determined by the number of occupants and room activities [23]. Exact hours and dates covered by each period stated above are outlined in Table 2. Occupied periods are set between 8 am to 5 pm from Monday to Friday when the building has open access. Unoccupied periods are set to cover the remaining hours when a pass is required to enter and some appliances and equipment such as lighting are operated in energy conservation mode. These hours are assumed to have lower occupancy with fewer teaching and event activities, so they are referred to as unoccupied periods in this research. Term and non-term periods are separated due to a similar reason. Term and non-term dates are defined by the academic calendar of Newcastle University. Lastly, the all-time period covers all 24 h of both term and non-term dates. All weekends are excluded from the analysis due to the less consistent operational states, thus leading to less reliable estimates.

Table 2. Details of segmented periods.

Period	Dates	Hours
Term occupied (a)	7 January–29 March 29 April–17 June 23 September–13 December	8:00–17:00 (M–F)
Term unoccupied (b)	7 January–29 March 29 April–17 June 23 September–13 December	0:00–8:00 17:00–24:00 (M–F)
Non-term occupied (c)	1 January–4 January 1 April–26 April 17 June–21 September 16 December–31 December	8:00–17:00 (M–F)
Non-term unoccupied (d)	1 January–4 January 1 April–26 April 17 June–21 September 16 December–31 December	0:00–8:00 17:00–24:00 (M–F)
All-time (e)	1 January–31 December	0:00–24:00 (M–F)

3.4. Datasets

Newcastle University's Urban Observatory is a research group that works on development and testing of real-time smart technologies for urban sustainability. It collects and stores real-time data from sensors located in the USB. Firstly, historical smart meter data of electricity consumption from 1 January 2019 to 31 December 2019 were accessed to generate daily electricity consumption for each period. In the following step, the consumption was further divided by the total floor area to generate the daily Energy Usage Intensity (EUI) of the whole building during occupied, unoccupied hours and the whole day as the dependent variable, such as the traditional annual benchmark approach. Secondly, weather related variables including daily average ambient temperature, solar radiation, relative humidity and wind speed over each period were chosen as independent variables as weather factors can significantly affect the energy consumption of HAVC systems. Secondly, the independent variables were selected. They represented the explanatory variables of energy consumption, which were used to normalize the energy use between each day [11]. The weather-related variables were chosen as the independent variables since the weather factor could significantly affect the energy consumption of the HAVC system. Adopting a similar method as [12], the average values of ambient temperature, solar radiation, relative humidity and wind speed were adopted in the model in this research. The other independent variable used was the occupancy of the building. As stated previously, the number of users affects the energy consumption mainly due to the change in equipment usage patterns. Since the top–down model was developed based on repeated measurements of an individual building, other independent variables such as building age, equipment type and installed power, which vary across a portfolio of different buildings, are not applicable. The other independent variable used was the daily occupancy of the building as the number of users affects the energy consumption mainly due to changes in equipment usage patterns. As mentioned in the previous section, the smarter meter and all explanatory variables were collected every 15 min and were transmitted to a cloud database. They were accessed via the REST API provided by the Urban Observatory.

The bottom–up energy simulation approach was developed using the building scheme and design including 2D drawings and a 3D BIM; they were provided by the asset management service team of Newcastle University. Lastly, carbon intensity data of the energy supplier were accessed via the National Grid ESO's carbon intensity API. It provides daily localized carbon intensity related to electricity generation only, which includes emissions from all large metered power stations, interconnector imports and transmission and distribution losses.

3.5. Carbon Emission Modelling

3.5.1. Top–Down Regression Model and Carbon Emission Score

A multivariate linear regression approach, such as the one used by [24], is adopted in this research to benchmark the energy efficiency which leads to a carbon emission score. Owing to their characteristics, such as straightforward form, ease of use and generally high level of statistical significance, the multivariate linear methodology is commonly used for the top–down modelling and is the foundation of industry benchmarking applications such as the Energy Star score [3,12]. The size of the sample dataset determines the accuracy of the regression model significantly. Thus, it is suitable for analysing high-frequency data due to the large volume. To judge whether a model fits the data well enough, its coefficient of determination (R^2) was calculated. All statistical analysis and modelling were performed using R version 4.0.1.

For each of the four segmented periods and the all-time period, a regression model was created. EUIs and occupancy inputs were log-transformed to account for the skewed distribution characteristics. To help with the interpretation of the coefficient results, all explanatory variables were re-scaled to have mean of 0 and standard deviation of 1. The form of the regression model is given by the following equation:

$$EUI = a + b_1T + b_2WS + b_3RH + b_4S + b_5Occ \quad (1)$$

In the equation, a = intercept; b_1 to b_5 = model coefficients; T = ambient temperature; WS = wind speed; RH = relative humidity; S = solar radiation and Occ = occupancy. For all five models covering each period, the relative humidity was identified as an insignificant variable, and there was high collinearity between the ambient temperature and solar radiation. Thus, the relative humidity and solar radiation were excluded from the final model adopted. The adjusted R^2 values of the final five models range between 0.65 and 0.73.

Next, daily normalized energy consumption and carbon emission were calculated as follows:

$$EUIn = EUIo - EUIp + a \quad (2)$$

$$CE = EUIn \times CI \quad (3)$$

EUIn = normalized EUI value; EUIp = predicted EUI based on the coefficient from regression Equation (1), EUIo = observed EUI of the building and a = model intercept. CE = carbon emissions, and CI = carbon intensity of the energy supplier.

Since the explanatory variables were rescaled with a mean of 0, the intercept represents the average EUI over the segmented period. The normalized value was the sum of the residual between the observed and predicted value (EUIo − EUIp) and the intercept. If the predicted EUIp was lower than the observed EUIo, the difference would be positive. Then, the difference was added to the average EUI over the segmented period, leading to a higher normalized value. In this case, the building is using more energy than predicted, which means it is operating less efficiently and producing more carbon.

Lastly, daily carbon emission values within each period were rescaled between 0 and 1 as the carbon emission score. The date with the highest carbon emission is given a score of 0, as it is the least efficient, while a score of 1 is given to the date with the lowest emission and is the most efficient. The result was a distribution of daily scores over the segmented periods.

3.5.2. Bottom–Up Simulation and Carbon Emission Gap Score

With the use of the Revit model and other inputs, an energy simulation model was developed just as the architects and design teams did in the preliminary design stage. Its input parameters and data sources are shown in Table 3. Firstly, a 3D massing model of each room of the building was created in Rhino 6.0 based on the BIM (Figure 3) [25]. Then, it was further converted to an energy simulation model using Ladybug 0.067 and Honeybee 0.064 tools in Grasshopper (Figure 4) [26]. The simulation consisted of four steps. Modelling: the massing of each analytic space and glazed openings including windows, curtain walls and skylights on each surface was constructed. Simulation conditions: Each analytic space was assigned its actual function such as open-plan office, IT room and lecture theatre according to the design scheme, and a standardized operational program was adopted. Weather files and building orientation were also imported. Simulation: The simulation process generated hourly electricity consumption values of small appliances, lighting and cooling—heating for each space. The time range of the simulation was the same as that of the top–down methods. Collection: numerical data of the consumption values were saved as .CSV files for further analyses.

Table 3. Input parameters for the simulation model.

Input Parameter	Data Source
Building geometry	Revit Model CAD plan and elevation
Construction material	Revit Model
Weather condition	EPW file
Operational load	Building functional program

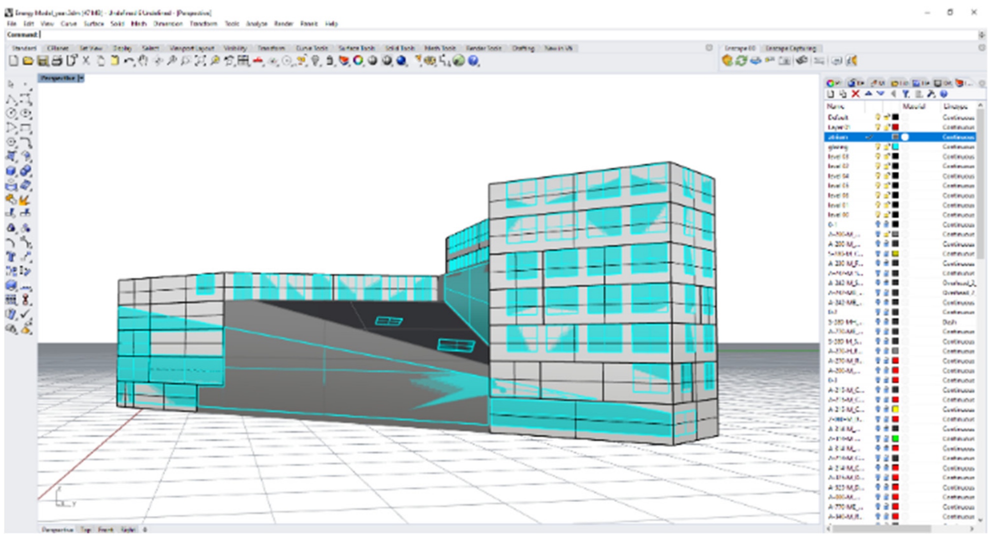

Figure 3. Model in Rhino.

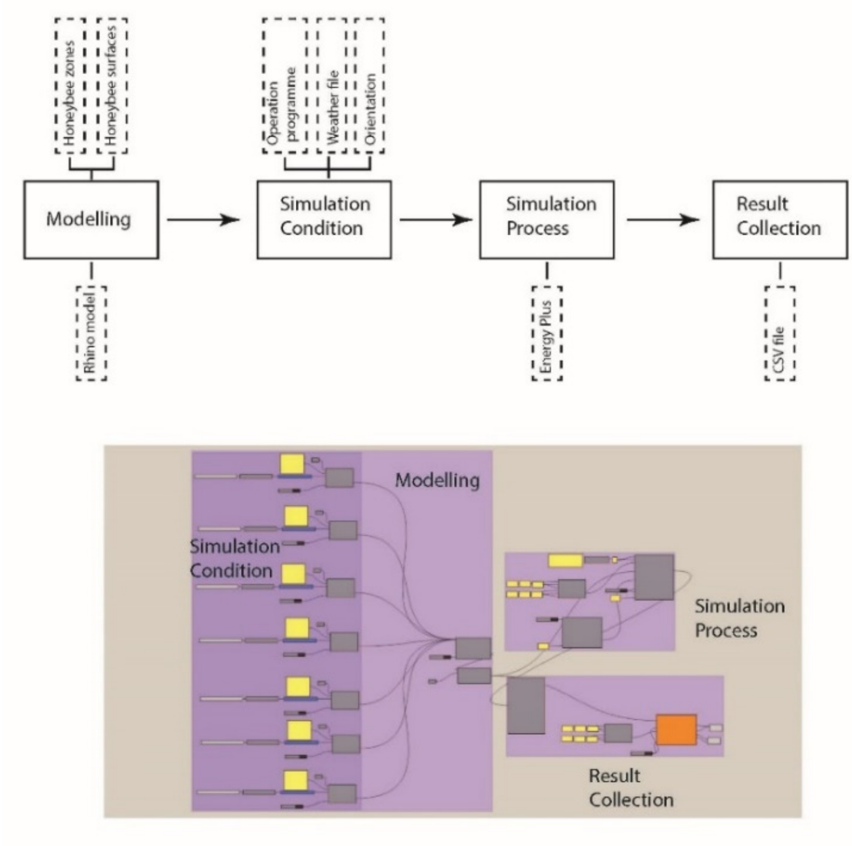

Figure 4. Energy simulation using Grasshopper.

Next, the result of the simulation model was compared with the normalized energy consumption gained from Equation (2) to generate the energy performance gap between the in-operational and pre-construction simulation usage using Equation (4). The carbon emission gap between the actual and simulated data was calculated using Equation (5). Dates and time periods with higher positive carbon emission gap values showed that the building used an excessive amount of energy than originally simulated, indicating that the simulation result was less accurate during such periods.

$$\mathrm{EPG} = \mathrm{EUIn} - \mathrm{EUIsim} \tag{4}$$

$$\mathrm{CEG} = \mathrm{EPG} \times \mathrm{CI} \tag{5}$$

EPG = energy performance gap; EUIsim = simulated EUI value; EUIn = normalized EUI value calculated using the top–down model of each segmented period; CEG = carbon emission gap; and CI = carbon intensity of the energy supplier.

Lastly, the daily carbon emission gap value within the same segmented period was also scaled to a range between 0 and 1. The higher the gap value, the lower the score given as it means the simulation model was less accurate and trustworthy; on the contrary, dates with lower gap values were given higher scores, which means the simulation model was more reliable for these dates and times.

4. Results

4.1. Carbon Emission Score

Figure 5 presents the 15-day rolling average of carbon emission scores over the year. They were compiled into three graphs to distinctly present patterns of efficiency scores in the occupied period (term time + non-term time), unoccupied period (term time + non-term time) and all-time period. This figure highlights that carbon emission scores of different periods had similar fluctuations over the year. However, there were also obvious opposite moving trends between the occupied and unoccupied periods during some specific months. It can be observed that the efficiency scores of both occupied and unoccupied periods were considerably low in January and February compared with the rest of the year. This is because the winter exam season of the university took place in January. During this time, the teaching activities were minimal, and the students tended to prepare for exams at home or in university libraries. In this case, the occupancy was relatively low, but the building still operated on the term time programme. In addition, the winter weather also caused a higher heating load compared with other seasons of the year. The same reason can possibly explain the decrease in the occupied period score during the summer exam season in May and June. During the summer period, most of the students and staff were away on holiday, and the occupied and unoccupied period scores decreased along a synchronised gradient in July. Later, they all increased again when the building was closed entirely and operated on a different programme in August. From September until early November, the unoccupied period score continued to increase and then decreased again around the beginning of November. The occupied period score decreased sharply in September and then followed a similar fluctuation such as the unoccupied period score in the rest of the year. It can also be observed that the occupied period score was consistently lower than the unoccupied period score during the term time. During the non-term time, it was right the opposite. During the spring term (January to March), Easter break (April) and summer break (July to September), the gap between the occupied and unoccupied efficiency score was relatively narrow and constant. Further, all scores increased and decreased at a similar value during these periods. In the summer term (May to June) and autumn term (October to December), the gap was widened significantly due to the increase in the unoccupied period score and decrease in the occupied period score. A sharp change of the score's moving trend may indicate that an operational shift has occurred in the building. It can highlight the need to investigate users' behavior around these turning points.

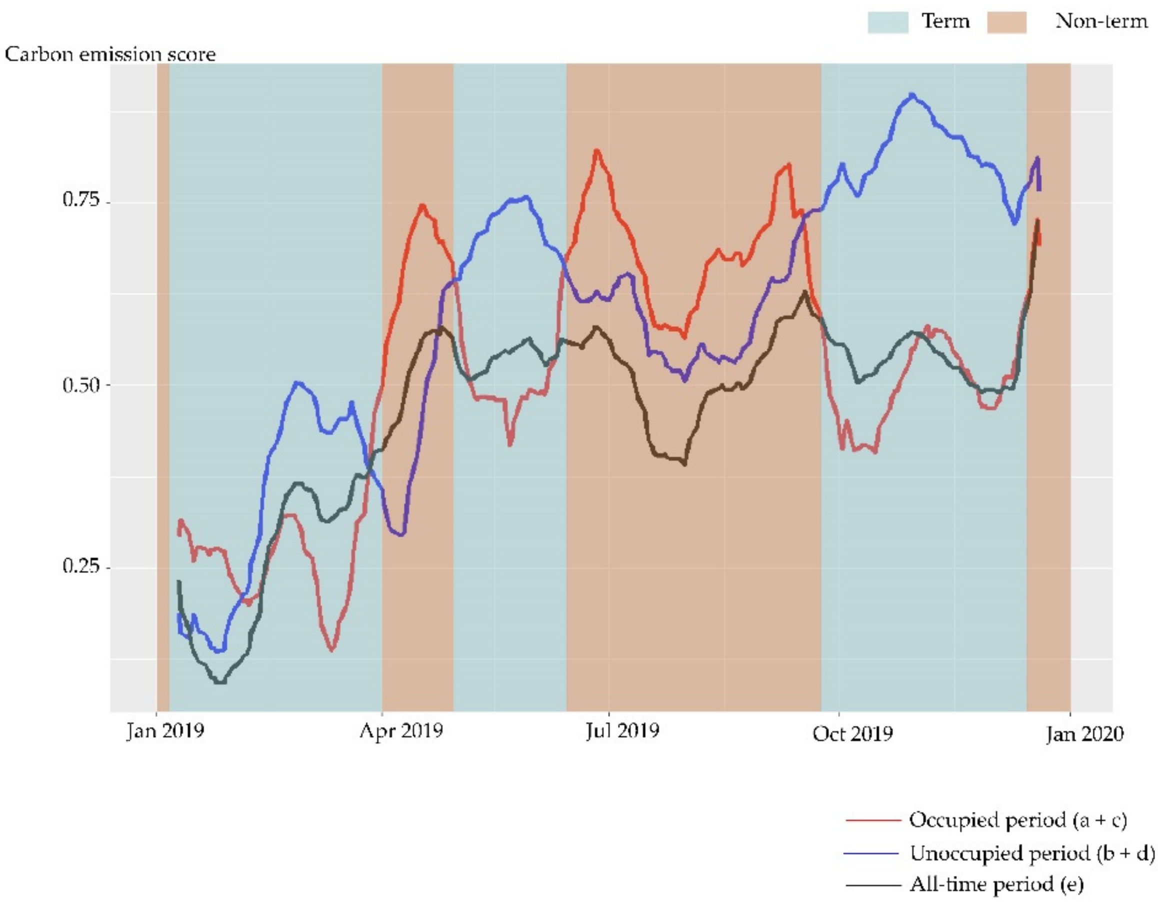

Figure 5. The 15-day rolling average of the carbon emission score.

4.2. Carbon Emission Gap Score

Figure 6 presents the 15-day rolling average of the carbon emission gap scores across the year. From January until September, both occupied and unoccupied period scores fluctuated, and the occupied score was consistently higher than the unoccupied score. The potential reason is that despite the term unoccupied as the building is closed to the public, the students and staff can still gain access during these hours. The extra energy load from lighting and heating especially during the evening time was not adequately considered in the simulation. After the commencement of the autumn term, scores increased drastically, reaching 0.8 and 0.7 around the end of the year. Despite fluctuations of exact score values, the deviation between two scores remained relatively large from the beginning of the year (January) until the end of the summer break (September). The largest deviation across the year appeared in the middle of the Easter break (April) and middle of the summer break (July). The result of the carbon emission gap score provided clearer evidence on the accuracy of the simulation model, which can contribute to the calibration of the model at the design stage.

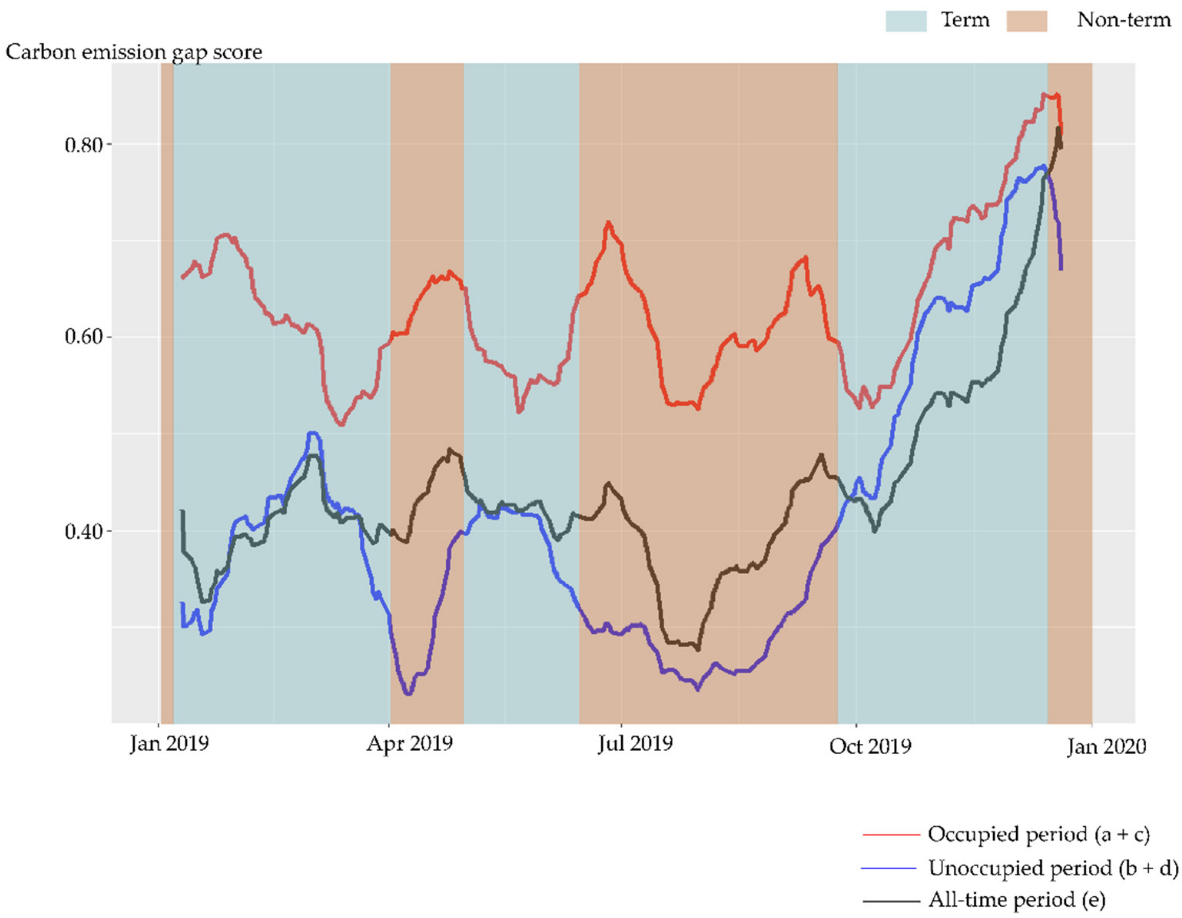

Figure 6. The 15-day rolling average of the carbon emission gap score.

5. Service Layer

Previous sections described the methodology used to develop the benchmark system. From the digital twin perspective, it completes data acquisition and data modelling layers of the DT framework. In this section, the development and usage of the service layer are presented and discussed.

5.1. Deviation between Segmented and All-Time Periods

Using segmented period scores can provide detailed information on the variation that occurred in carbon emissions. It can be observed from Figure 4 that there was a distinct difference between the segmented period scores (a + c, b + d) and the all-time score (e). The evaluation of these differences can provide insights masked by judging the raw score value only.

The average weekly difference between the segmented and all-time period performance scores across the whole year is illustrated in Figure 7. In this figure, each cell represents one week, and the filled colour is determined by the difference value. The left panel illustrates the difference between the occupied and all-time scores, while the right panel illustrates the difference between the unoccupied and all-time scores. In the spring term, if judged only by the raw score value shown in Figure 4, the building appeared to not perform well all day since both occupied and unoccupied period scores were low. However, for the unoccupied period score, it was consistently higher than the all-time score, which means it was less carbon intensive than initially thought. The occupied period score was higher than the all-time score at the beginning of the term, but the difference gradually descended to negative. This indicated a large potential improvement during the occupied period.

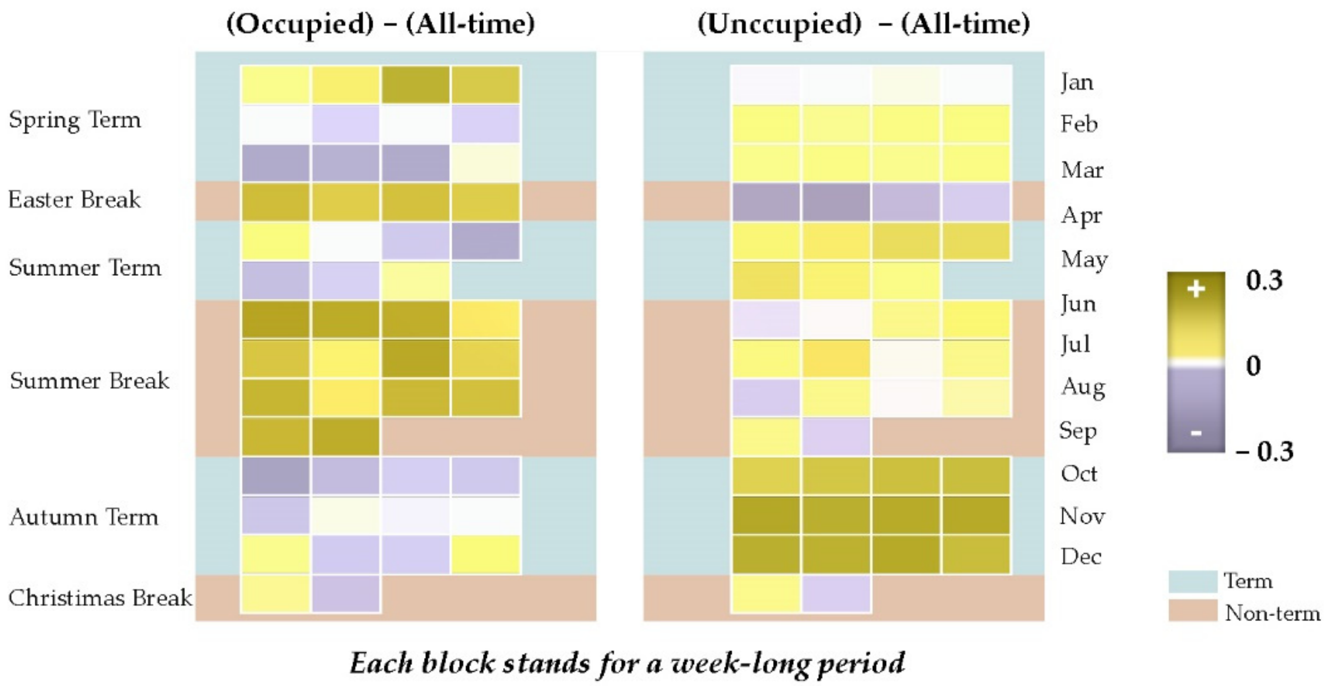

Figure 7. Difference in carbon emission scores between the segmented and all-time periods.

For the asset management team, a DT service layer built on such difference values can help the team members assess the performance of each temporal period more accurately. If the score difference value is constantly negative or positive during a certain period, the team can be more assertive on whether it should prioritize the potential retrofit strategy. On the other hand, during some periods when the difference value changes quite swiftly and differs from week to week, the team may need to review the building mechanical system such as the automated controlled equipment and investigate any potential errors causing such instability.

5.2. Insights for Future Projects

As previously shown in Section 3.2, the carbon emission gap score during the unoccupied period was consistently lower than that during the occupied period. This indicates the simulation result was less accurate during the unoccupied period and it has great potential to be improved.

Table 4 presents the average carbon emission gap score for each segmented period. It can be observed that the summer term (0.39) and summer break (0.28) had lowest unoccupied scores for the term and non-term periods, respectively. Further analyses were conducted to understand how the energy performance gap formed and varied in different parts of the building, leading to the carbon emission gap during these two periods. The building's smart meter data recorded the usage of each distribution zone from level 3 to 6. Thus, the daily energy usage of each level can be gained, and the daily value of the energy performance gap of these levels was computed. Figure 8 presents the average daily energy performance gap value of level 3 to 6 during these two periods. Levels with mixed plan layouts of teaching and office zones (levels 3 and 4) had smaller gaps compared with the levels mainly used as office spaces (levels 5 and 6), and between the top two levels, the gap was always larger in level 5. A higher office building usage ratio (BUR) could be the major contributor to the inaccuracy of a simulation model. This is further evaluated in the following step.

Table 4. Average carbon emission gap score in segmented periods.

Period	Average Carbon Emission Gap Score (Occupied)	Average Carbon Emission Gap Score (Unoccupied)
Spring term (7 January–29 March)	0.62	0.40
Summer Term (29 April–17 June)	0.56	0.39
Autumn Term (23 September–13 December)	0.66	0.59
Easter break (1 April–26 April)	0.56	0.33
Summer break (17 June–21 September)	0.62	0.28
Christmas break (1 January–4 January and 16 December–31 December)	0.71	0.69

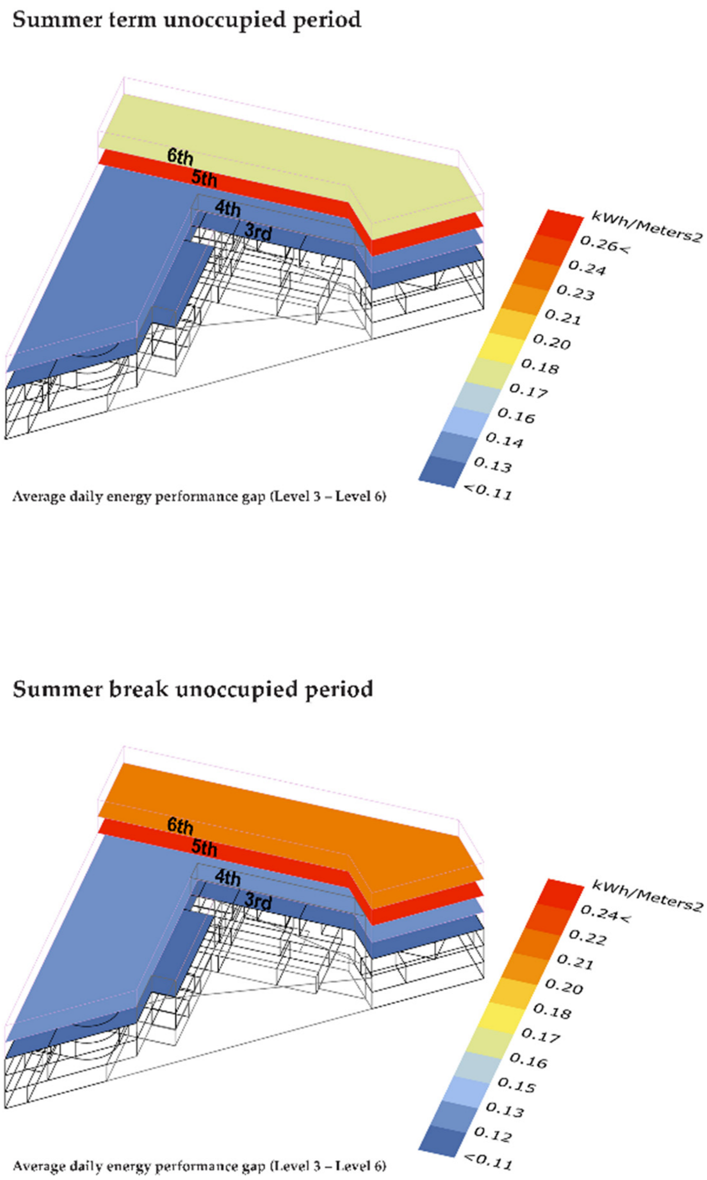

Figure 8. Average daily energy performance gap.

Smilar to the office BUR, there are other physical factors such as the window–wall ratio (WWR), thermal material property and functional layout that can affect the results of the simulation model. For architects and consultants, they need to produce different design options related to these characters and test them during the schematic design stage. The relationship between the gap value and these physical characters will help them to evaluate the simulation result and assess the design schemes more precisely. Table 5 lists the physical characters of level 3 to 6 extracted from the BIM. For the calculation of the window–wall ratio, the window area includes all fixed windows and curtain wall glazing. For the building usage ratio (BUR), the teaching area includes all lecture rooms and computer labs, the office area includes all open-plan offices and meeting rooms, the service area includes all bathrooms and tearooms and the circulation areas includes all corridor and stairs.

Table 5. Building characters for level 3 to 6.

Level	WWR%	Circulation BUR%	Office BUR%	Teaching BUR%	Service BUR%	Plant BUR%
3	32	25	21	38	11	5
4	31	22	25	37	10	6
5	37	2	65	9	20	4
6	37	1	79	0	16	4

From analyzing the daily energy performance gap values of each level and their building characters, we found out that if judged by the numerical values, levels with higher office and service ratios and lower window–wall ratios were likely to have a larger energy performance gap. The relationship between the values is shown in Figure 9. The other characters showed less direct impact. As introduced previously in Figure 1, levels 3 and 4 span the teaching and office wings of the building. In addition, they both connect with the atrium, which has mechanically operated windows on the curtain wall to assist with ventilation. Levels 5 and 6 consist of the office wing only and are cut off from the atrium. The windows are generally smaller to align with the individual office space and are opened and closed manually. For the architects and design team, the results show that the simulation model is more reliable when applied to mixed use spaces. If a large atrium with a curtain wall front, which is becoming more common in contemporary architectural design, is adopted in the design scheme, the result of simulation model is closer to the in-operation usage. As for single function spaces such as an office zone with less natural lighting and ventilation, users tended to have a higher energy demand than originally expected.

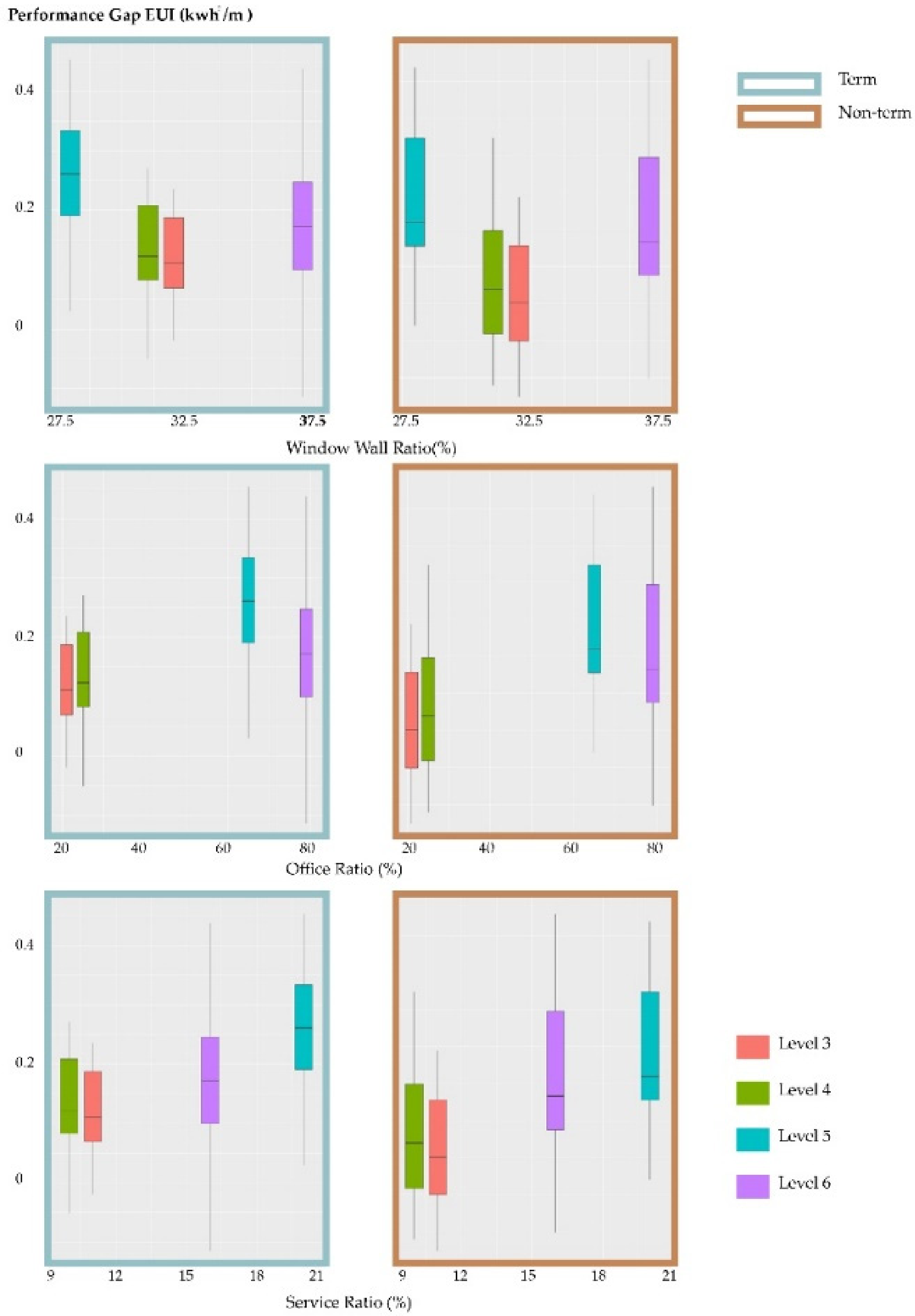

Figure 9. Building characters and energy performance gap.

6. Discussion

6.1. Summarising the Benchmark Systems

This study developed daily carbon emission benchmarking scores segmented by strategic periods for an individual building using two types of energy models. For owners of individual buildings, their keen interest is primarily how much a certain upgrade to a particular building or portfolio buildings is going to improve their energy and carbon efficiency and increase the return on investment. For a large multi-function building such as the USB, there are increasing challenges of managing carbon emission. The complexity of building systems may lead to higher maintenance costs, rapid obsolescence of systems and a constant need for re-training of operations staff and deterioration of performance [3]. As studies have shown, buildings equipped with complex systems generally run "more liberally and wastefully" despite the fact the design of these systems is generally driven by sustainability certifications and rating schemes [27]. The benchmarking system can support identification of the exact temporal period, building system and spatial segmented part that has abnormal energy performance.

Recent studies have leveraged smart-metered data to explore how temporally segmented energy benchmarks can gain more specific insights, as conventional annual benchmarks are limited in their ability to identify areas for improvement [3,24]. The concept of an ensemble model was introduced by [28] as a group of calibrated models for an individual building that mimic the building's measured energy use using different combinations of possible input parameters. In that paper, [28] tracked historical building energy data and explored potential scenarios for retrofit. It was applied as a continuous energy performance planning system for a university campus in Cambridge, Massachusetts. For the future of building energy and carbon modeling, a potential tendency could be an ensemble modelling framework across strategically segmented temporally periods. In this research, such a paradigm is set. This approach can be further applied to a portfolio of buildings that shares similar external contextual and building characters, such as buildings located within the same campus, retails chains across a city or residential blocks which are closely situated to each other. The outcome of such models will be a spread of possible and potential more accurate saving strategies for each building [15].

However, the granularity of the parameters and fidelity of recorded energy usage data (in this research, the smart meter data) directly affect the accuracy of the model, with finer time-steps typically displaying greater uncertainties [15]. Further, it could be beneficial to include data pertaining to the building such as construction assemblies, mechanical specification and operational schedules. Thus, there are additional technical requirements for the modeler. Such factors will restrict the model's applicability.

6.2. Limitation and Future Directions

There are several limitations in this research. Firstly, in the data acquisition layer, the most evident limitation is that intensive 15-min interval data such as occupancy are not universally available. Most buildings currently in use do not have many sensors installed and are unlikely to acquire similar data sets. Thus, it is not quite possible for them to adopt the DT framework seamlessly. However, in the report Data for the Public Good by the National Infrastructure Commission, a key scope is the development of a so-called national digital twin [29]. One of the fundamental concepts of the national digital twin is that it is not a single monolithic model of whole nation's infrastructure but consists of digital twins that are constructed at different scales and built for various purposes using different approaches. More specifically, occupancy data can be substituted with other suitable resources. More widely accessible data sources such as CCTV cameras, mobile connections and consumer transactions can be investigated as they possibly provide even more detailed information of occupants' behavior.

Secondly, in the data modelling layer, the carbon emission score system relies on regression-based techniques. Regression residuals were assumed to reflect the lower efficiencies of the building only. In reality, there are many other factors such as statistical noise and measure errors that need to be considered for optimization of the model [30]. However, such weakness in the data modelling process is unavoidable as other modelling techniques have their own limitations as well. Considering the core of this study is to present a complete framework of the digital twin model which can be reproduced for other buildings, the multivariate regression model was adopted due to its wide implementation as an industry standard such as the Energy Star score. One possible avenue for future study is the use of other modelling techniques and to compare their results consistently. Particularly for buildings with different functions and usage patterns from the USB, other methods are potentially more appropriate for their own purposes. In fact, such speculation echoes the national digital twin concept as any DT framework should not simply be replicated for different built assets.

In addition, another major limitation is that the user comfort experience was not assessed in the development of the benchmark system. Especially for a building such as the USB which has long opening hours and after-hours access for students and staff, extra loads of heating, air conditioning, lighting and appliance usage are needed to keep the building at a suitable comfort level even when the occupancy is low. Such requirements lead to the conflict between reducing carbon emissions and user comfort. One possible approach for improvement is adopting explanatory variables such as the indoor temperature and air quality for the consideration of the occupants' comfort and health.

7. Conclusions

In sum, this research explains a method and implementation for using smart meter and other near real-time sensor data to benchmark the daily carbon emissions of a complex building and the carbon emission gap between the actual and pre-construction prediction across strategic segmented periods. A case study was used to examine the method. Our findings demonstrate that between the segmented periods, the carbon efficiency scores fluctuated significantly due to changes in user occupancy and weather condition. The fluctuation provided the overall building performance over each period and identified those with constant low scores for potential improvement. Within the same time framework, we found that the occupied and unoccupied period scores may demonstrate different or even opposite moving patterns, and it can provide further detailed insight into the root cause of the inefficiency. The changes in the carbon emission gap score showed how the accuracy of the simulation result changed across each period. The analysis of the building characters and the gap contributed to the assessment of the simulation result and calibration of the simulation model.

A digital twin-based framework was adopted as the platform for the implementation of the benchmark system. The structure of the DT framework was shown in the case study. The insight gained from the benchmark system can contribute to the decision-making process related to carbon strategy for in-operation buildings and design scheme improvement for future projects. The increasing accessibility of smart meter and other building operational data provides potential applications towards smarter operation and design and helps the realization of the national digital twin advocated by [29].

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

Funding: This research was funded by EPSRC, grant number EP/S023577/1.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Publicly available datasets were analyzed in this study. This data can be found here: https://api.usb.urbanobservatory.ac.uk/, https://carbon-intensity.github.io/api-definitions/#regional.

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

References

1. Li, X.; Yao, R. Modelling heating and cooling energy demand for building stock using a hybrid approach. *Energy Build.* **2021**, *235*, 110740. [CrossRef]
2. Salihi, M.; El Fiti, M.; Harmen, Y.; Chhiti, Y.; Chebak, A.; Alaoui, F.E.M.; Achak, M.; Bentiss, F.; Jama, C. Evaluation of global energy performance of building walls integrating PCM: Numerical study in semi-arid climate in Morocco. *Case Stud. Constr. Mater.* **2022**, *16*, e00979. [CrossRef]
3. Borgstein, E.H.; Lamberts, R.; Hensen, J.L.M. Evaluating energy performance in non-domestic buildings: A review. *Energy Build.* **2016**, *128*, 734–755. [CrossRef]

4. Metallidou, C.K.; Psannis, K.E.; Egyptiadou, E.A. Energy Efficiency in Smart Buildings: IoT Approaches. *IEEE Access* **2020**, *8*, 63679–63699. [CrossRef]
5. Pärn, E.A.; Edwards, D.J.; Sing, M.C. The building information modelling trajectory in facilities management: A review. *Autom. Constr.* **2017**, *75*, 45–55. [CrossRef]
6. Bolton, A. *The Gemini Principles*; Centre for Digital Built Britain: Cambridge, UK, 2018. [CrossRef]
7. Lu, Q.; Parlikad, A.K.; Woodall, P.; Don Ranasinghe, G.; Xie, X.; Liang, Z.; Konstantinou, E.; Heaton, J.; Schooling, J. Developing a Digital Twin at Building and City Levels: Case Study of West Cambridge Campus. *J. Manag. Eng.* **2020**, *36*, 05020004. [CrossRef]
8. Mohammadi, N.; Taylor, J.E. Smart city digital twins. In Proceedings of the 2017 IEEE Symposium Series on Computational Intelligence, SSCI, Honolulu, HI, USA, 21 November–1 December 2017; pp. 1–5. [CrossRef]
9. Dawkins, O.; Dennett, A.; Hudson-Smith, A. Living with a Digital Twin: Operational management and engagement using IoT and Mixed Realities at UCL's Here East Campus on the Queen Elizabeth Olympic Park. In Proceedings of the 26th Annual GIScience Research UK Conference: GISRUK 2018, Leicester, UK, 17–20 April 2018.
10. Swan, L.G.; Ugursal, V.I. Modeling of end-use energy consumption in the residential sector: A review of modeling techniques. *Renew. Sustain. Energy Rev.* **2009**, *13*, 1819–1835. [CrossRef]
11. Chung, W.; Hui, Y.V.; Lam, Y.M. Benchmarking the energy efficiency of commercial buildings. *Appl. Energy* **2006**, *83*, 1–14. [CrossRef]
12. Amber, K.; Aslam, M.W.; Hussain, S.K. Electricity consumption forecasting models for administration buildings of the UK higher education sector. *Energy Build.* **2015**, *90*, 127–136. [CrossRef]
13. Aranda, A.; Ferreira, G.; Mainar-Toledo, M.D.; Scarpellini, S.; Sastresa, E.L. Multiple regression models to predict the annual energy consumption in the Spanish banking sector. *Energy Build.* **2012**, *49*, 380–387. [CrossRef]
14. Balaras, C.A.; Gaglia, A.G.; Georgopoulou, E.; Mirasgedis, S.; Sarafidis, Y.; Lalas, D. European residential buildings and empirical assessment of the Hellenic building stock, energy consumption, emissions and potential energy savings. *Build. Environ.* **2007**, *42*, 1298–1314. [CrossRef]
15. Ang, Y.Q.; Berzolla, Z.M.; Reinhart, C.F. From concept to application: A review of use cases in urban building energy modeling. *Appl. Energy* **2020**, *279*, 115738. [CrossRef]
16. Hsu, D. Identifying key variables and interactions in statistical models ofbuilding energy consumption using regularization. *Energy* **2015**, *83*, 144–155. [CrossRef]
17. Kaewunruen, S.; Rungskunroch, P.; Welsh, J. A Digital-Twin Evaluation of Net Zero Energy Building for Existing Buildings. *Sustainability* **2018**, *11*, 159. [CrossRef]
18. Wang, Y.; Chen, Q.; Hong, T.; Kang, C. Review of Smart Meter Data Analytics: Applications, Methodologies, and Challenges. *IEEE Trans. Smart Grid* **2019**, *10*, 3125–3148. [CrossRef]
19. Kavousian, A.; Rajagopal, R.; Fischer, M. Determinants of residential electricity consumption: Using smart meter data to examine the effect of climate, building characteristics, appliance stock, and occupants' behavior. *Energy* **2013**, *55*, 184–194. [CrossRef]
20. 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]
21. Roth, J.; Jain, R.K. Data-driven, multi-metric, and time-varying (DMT) building energy Benchmarking using smart meter data. In *Advanced Computing Strategies for Engineering. EG-ICE 2018. Lecture Notes in Computer Science*; Springer: Berlin/Heidelberg, Germany, 2018; Volume 10863, pp. 568–593. [CrossRef]
22. Park, J.; Yang, B. GIS-Enabled Digital Twin System for Sustainable Evaluation of Carbon Emissions: A Case Study of Jeonju City, South Korea. *Sustainability* **2020**, *12*, 9186. [CrossRef]
23. Gul, M.S.; Patidar, S. Understanding the energy consumption and occupancy of a multi-purpose academic building. *Energy Build* **2015**, *87*, 155–165. [CrossRef]
24. Francisco, A.; Mohammadi, N.; Taylor, J.E. Smart City Digital Twin–Enabled Energy Management: Toward Real-Time Urban Building Energy Benchmarking. *J. Manag. Eng.* **2020**, *36*, 04019045. [CrossRef]
25. McNeel, R. *Rhinoceros 3D*; Robert McNeel & Associates: Seattle, WA, USA, 2010.
26. Roudsari, M.S. Ladybug + Honeybee. 2014. Available online: https://www.ladybug.tools/ (accessed on 30 June 2020).
27. Bordass, B.; Cohen, R.; Standeven, M.; Leaman, A. Assessing building performance in use 3: Energy performance of the Probe buildings. *Build. Res. Inf.* **2001**, *29*, 114–128. [CrossRef]
28. Nagpal, S.; Mueller, C.; Aijazi, A.; Reinhart, C.F. A methodology for auto-calibrating urban building energy models using surrogate modeling techniques. *J. Build. Perform. Simul.* **2019**, *12*, 1–16. [CrossRef]
29. National Infrastructure Commission. Data for the public good. 2017, p. 3. Available online: https://nic.org.uk/app/uploads/Data-for-the-Public-Good-NIC-Report.pdf (accessed on 15 June 2022).
30. Chung, W. Review of building energy-use performance benchmarking methodologies. *Appl. Energy* **2011**, *88*, 1470–1479. [CrossRef]