The current issue and full text archive of this journal is available on Emerald Insight at:
https://www.emerald.com/insight/0168-2601.htm

A geometry-based decomposition method for energy prediction in early design stages of residential buildings

Open House International

İbrahim Agah Taştemir and Erdem Koymen
Faculty of Engineering and Natural Sciences, Istanbul Sabahattin Zaim University, Istanbul, Turkey, and
Enes Yasa
Faculty of Architecture, Istanbul Technical University, Istanbul, Turkey

Received 17 May 2024
Revised 9 October 2024
21 November 2024
21 November 2024
Accepted 21 November 2024

Abstract

Purpose – The purpose of the study is to develop a geometry-based energy estimation method for surrogate and metamodels to be used in the early design phase of buildings.

Design/methodology/approach – Optimizing building form and design variables in the early stages of the architectural design process, particularly during the conceptual phase, can significantly enhance overall design performance and energy efficiency at minimal cost. This study introduces a novel decomposition method for evaluating building energy performance by simplifying complex building forms into basic geometric shapes.

Findings – The developed method is applied to certain cases of design variation under specified boundary conditions, and the accuracy of heating and cooling energy loads is calculated with simulated energy models of these cases. As a result of the calculation, accuracy rates between 84.30 and 99.98% were founded.

Originality/value – This paper proposes a prediction model with a geometric identification method for an innovative geometry-based surrogate modeling method. This method also provides a way for artificial intelligence-based prediction models used in surrogate models to create a dataset and can be used in the training in future works.

Keywords Decomposition, Building form, Building energy efficiency, Energy prediction

Paper type Research paper

1. Introduction

Architectural designers are crucial for providing appropriate indoor comfort conditions for user actions in building spaces. A significant portion of global energy consumption is driven by the heating and cooling systems of buildings, which are used to maintain thermal comfort conditions. The increasing demand for energy leads to the depletion of natural energy resources such as oil, natural gas, and coal, further exacerbating the strain on these limited resources. The carbon emissions and greenhouse gases produced through the consumption of these resources contribute to air pollution, global warming, and consequently, climate change, creating a negative impact on the global ecosystem. Residential building construction and the operation of buildings account for a significant portion of global energy consumption (IEA, 2018). The early design is an important design phase for energy-efficient design, as decisions about buildings have the highest impact on final performance at the lowest cost (Depecker *et al.*, 2021). In the architectural design process for residential buildings, building mass design is a fundamental step to determine the aesthetic and functional context of the building and to investigate the appropriate functionality of the three-dimensional external shaping (form, size, or orientation) and structural states. However, many studies show that the building form determined in the

Declaration of competing interest: The authors declare that they have no known financial or non-financial competing interests in any material discussed in this paper.

Funding information (11 pt, Sentence case): This work was funded by Istanbul Sabahattin Zaim University Scientific Research Project (IZU BAP 2023–03), Istanbul Sabahattin Zaim University.

Open House International

e-ISSN: 2633-9838
p-ISSN: 0168-2601
DOI 10.1108/OHI-05-2024-0178

early design phase of building mass-form design greatly impacts energy efficiency (Mehrotra and Yi, 2023; Kahraman and Koymen, 2023).

Building energy simulation tools (such as Energy Plus) play an important role in the design phase of building energy assessment. These programs offer significant capabilities, such as the ability to simulate daily occupancy and consumption patterns, and to accurately predict heating/cooling loads. However, despite significant advances in building energy modeling tools, several limitations remain, such as the complexity of algorithms, the need for multidisciplinary expertise (e.g. HVAC systems), the cost of software and time management issues. In the early stages of design, when rapid decision making is critical, data-driven surrogate modeling is a complementary technique that provides a faster, but less accurate, tool for assessing comparative energy performance. Combining the capabilities of building simulation tools and surrogate modeling can offer a method to provide a more complete assessment of building energy performance.

1.1 Building energy performance simulation in early design stage

Building design an iterative and evolutionary process, wherein designers explore various design alternatives through different methods, considering numerous factors (Singh and Smith, 2023). The early design stage refers to the phase prior to detailed design. In this phase, crucial decisions such as building geometry, massing, area-to-volume ratios, and orientation are made (İşeri and Dursun, 2022). Enhancements in building orientation and envelope design at this stage have the potential to reduce the building's energy demand by up to 40% (Wang *et al.*, 2005). Therefore, to better understand the relationship between design decisions and building performance, technical investigations such as building energy simulations and statistical data analysis should be carried out to increase the importance given to early design phases.

Building performance optimization and prediction studies play a critical role during the early design phase (Konis *et al.*, 2016). However, due to the time-consuming nature of performance calculations and the need for customized parameter settings, environmental performance assessments are typically conducted during the later stages of design, often by professionals from various disciplines (Sing, 2022). Performance-based architectural design, on the other hand, allows designers to achieve enhanced energy efficiency and environmental performance by enabling the measurement and visualization of building performance throughout the design process (Yasa, 2016).

Designers should receive rapid and iterative performance feedback during the early design stages, when decisions have the most substantial impact on building performance and occupant comfort, rather than assessing whether a predetermined design meets compliance requirements in the later stages of development (Wilde, 2018). Key aspects such as building geometry, orientation, ventilation configurations, and thermal comfort management strategies should be integral to the building concept in order to achieve energy efficiency and optimal climatic comfort. To this end, designers must optimize the environmental services provided by natural systems, rather than focusing solely on incremental improvements to mechanical systems. Consequently, design workflows should facilitate the examination and optimization of passive environmental strategies in the early stages of design (Konis *et al.*, 2016).

The increasing adoption of machine learning algorithms across various professional fields has extended to architecture, where these methods are now employed in building energy prediction models. This approach, known as the "black box" method, has recently gained prominence in the evaluation of building energy performance (Hamdaoui *et al.*, 2021). The primary objective of constructing such forecasting models is to establish mathematical relationships between independent parameters and target variables, using historical data as a basis (Zhang *et al.*, 2022). Black box models share the common goal of determining the impact of various factors on building energy consumption (Sun *et al.*, 2020). With the growing availability of building energy consumption datasets and reduced dependency on extensive building parameters, data-driven energy performance models have emerged as one of the most effective tools for predicting energy use.

Pittarello *et al*. developed an artificial intelligence-based tool designed for rapid and comparable assessments of energy-efficient design during the early stages of building design, without directly relying on traditional building energy simulation tools. The researchers created an automated analysis and evaluation system by utilizing energy model outputs, in which various design criteria are processed through deep feed-forward artificial neural networks. This tool was tested on an accurate building model, and the results were compared with data derived from conventional building energy simulation tools, demonstrating its effectiveness (Pittarello *et al*., 2021).

Oluj-Ajayi *et al*. underscored the importance of developing a machine learning-based performance evaluation model that can be utilized by designers during the early design stages of buildings to ensure energy efficiency and climatic comfort. The study explores the effectiveness of various machine learning techniques, including Artificial Neural Networks (ANN), Gradient Boosting (GB), Deep Neural Networks (DNN), Random Forest (RF), Stacking, and K-Nearest Neighbors (KNN), in energy prediction and performance evaluation. The analysis revealed that the Deep Neural Network (DNN) method outperformed the other techniques, yielding more efficient results (Oluj-Ajayi *et al*., 2021).

Forouzandeh *et al*. (2022) developed a prediction model to estimate annual energy demand and thermal comfort performance using seven different machine learning methods. This study was based on simulations using the EnergyPlus engine. The room models consisted of single-zone office rooms, and their energy demand and thermal comfort performances were calculated. R^2 values ranging from 0.71 to 0.95 were reported for the annual thermal comfort performance, indicating high accuracy for the models (Forouzandeh *et al*., 2022).

Singh and Smith (2023) developed a web-based decision support model using CNN (Convolutional Neural Network) that predicts energy consumption from 2D images based on the geometric features of buildings. The model was tested on German office buildings and used multi-layered neural networks to capture geometric complexity and improve the accuracy of energy consumption predictions. With an accuracy rate above 90%, the model showed successful results; however, authors recommend that further discussion on the building models and methods used is recommended (Singh and Smith, 2023).

Li *et al*. proposed a method to develop an artificial neural network (ANN) based method to fast and accurately predict energy consumption in the early design phase of complex architectural forms. As part of this method, they introduced an architectural form decomposition methodology aimed at simplifying the complex geometry of building shapes during the early design stages. This approach transforms the energy consumption prediction problem for complex architectural forms into several smaller, more manageable energy consumption prediction tasks for simple blocks. This is achieved through the use of the Characterization Decomposition Method (MCD) and Spatial Homogenization Decomposition Method (MSHD). Accuracy tests in the study show a deviation of less than 10% in cooling and heating energy consumption, while high accuracy is achieved in total energy consumption (Li *et al*., 2019).

Ciardiello *et al*. conducted a multi-objective optimization of buildings with regular plan shapes, including I, L, O, C, Y, H, X, and T forms, by considering key shape parameters such as window-to-wall ratio and orientation. Through the use of a mass decomposition method, they divided complex building forms into several basic blocks. The total energy consumption for thermal comfort was then estimated by summing the partial energy uses of the divided geometrical forms, allowing for a more accurate assessment of energy performance in relation to the building's shape and orientation (Ciardiello *et al*., 2020).

Zhu *et al*. developed a hybrid metamodel-based approach aimed at rapidly predicting energy demand during the early design stages of buildings. Recognizing the critical impact of early design decisions on building performance, the authors sought to alleviate the high computational burden of traditional energy simulations by proposing a faster yet accurate alternative. This method involves breaking down complex building forms into simpler box units and utilizing machine learning algorithms to predict energy demand based on inputs such as solar radiation. GPU acceleration is employed to further speed up calculations. This study considered not only the building scale but also the neighborhood scale, accounting for the radiation effects of surrounding buildings on the main building. As a result of the case study analysis, an accuracy rate of 6% for cooling energy demand and 10% for heating energy demand was achieved. (Zhu *et al.*, 2021).

Li *et al*. developed an integrated building energy performance evaluation method that combines parametric modeling, automated energy simulation, and energy consumption prediction through a Genetic Algorithm-Neural Network (GA-NN) model. The study aimed to address the challenges of traditional energy evaluation methods, such as their time-consuming nature and high computational complexity, by proposing a more efficient and user-friendly approach. The methodology consists of three stages that are parametric modeling, automatic energy simulation and energy consumption prediction using the GA-NN model to provide accurate energy performance estimates. The results demonstrated high prediction accuracy, with deviations of less than 10% for cooling and heating loads. This approach enhances the operational efficiency of energy modeling, allowing designers to make informed decisions during the early stages of building design.

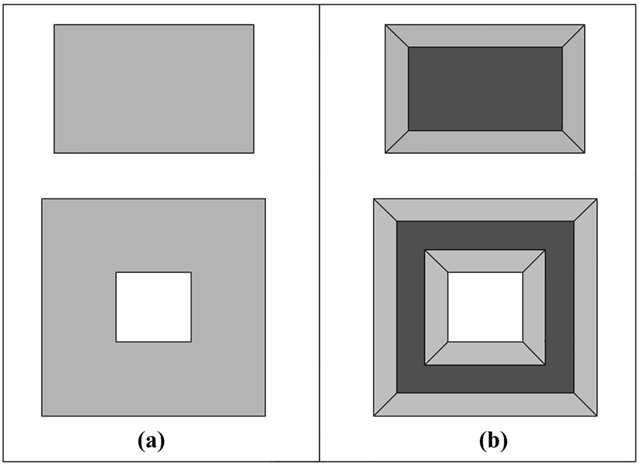
Salami *et al*. (2023) developed a sophisticated hybrid machine learning approach for predicting building energy loads by integrating Bayesian-based metaheuristic optimization with explainable tree-based models. The study focused on eight key building features—relative compactness, surface area, wall area, roof area, overall height, orientation, glazing area, and glazing area distribution. The model was tested on a dataset of 768 buildings, demonstrating high predictive accuracy, with the extreme gradient boosting algorithm performing best. Results showed deviations of less than 2% for both heating and cooling loads (Salami *et al.*, 2023).

Thermal zoning in building energy simulations is typically aligned with the spatial division of a building into individual rooms or spaces. This method facilitates the simulation process and the design of heating, ventilation, and air conditioning (HVAC) systems by enabling the grouping of multiple rooms with similar thermal load profiles—determined by factors such as space usage and solar gains—into a single zone (Bobenhausen, 1994). However, during the schematic design phase, the exact room layout is often not yet defined. In response to this challenge, ASHRAE 90.1 Appendix G provides guidance for such cases, recommending that floors be divided into core and perimeter zones (Figure 1) (ANSI/ASHRAE/IES Standard 90.1–(2019)). The perimeter zone is defined as the area extending five meters from the facade, and if it spans multiple orientations, it should be subdivided accordingly. The central area of the floor, left over after the perimeter is defined, forms the core zone. This guideline offers a streamlined approach for zoning during early-stage design, facilitating effective energy modeling and HVAC system design (Dogan *et al.*, 2016).

Many studies in the literature have categorized buildings into various building archetypes. However, in practice buildings can be complex shapes. This method is based on the theory that the adjacent surfaces of box units can be regarded as adiabatic boundaries for simplification and that internal wall heat conduction has an insignificant effect on the total building energy demand. In this context, the energy demand of a box unit in a building is considered equivalent to an isolated thermal zone with the same boundary conditions and environment (Chen *et al.*, 2017; Dogan *et al.*, 2016)

1.2 Aim of study

Building energy simulation is essential for evaluating energy consumption during the early design stages, but it faces challenges like uncertain design data and long computation times. Since building form and geometry significantly impact energy performance, early decisions



Source(s): Image adapted from the work of Dogan *et al.* (2016) and ANSI/ASHRAE/IESNA (2019)

Figure 1. Floor plans zoned according to ASHRAE 90.1 Appendix G(a) plan geometry (b) plan geometry thermal zones

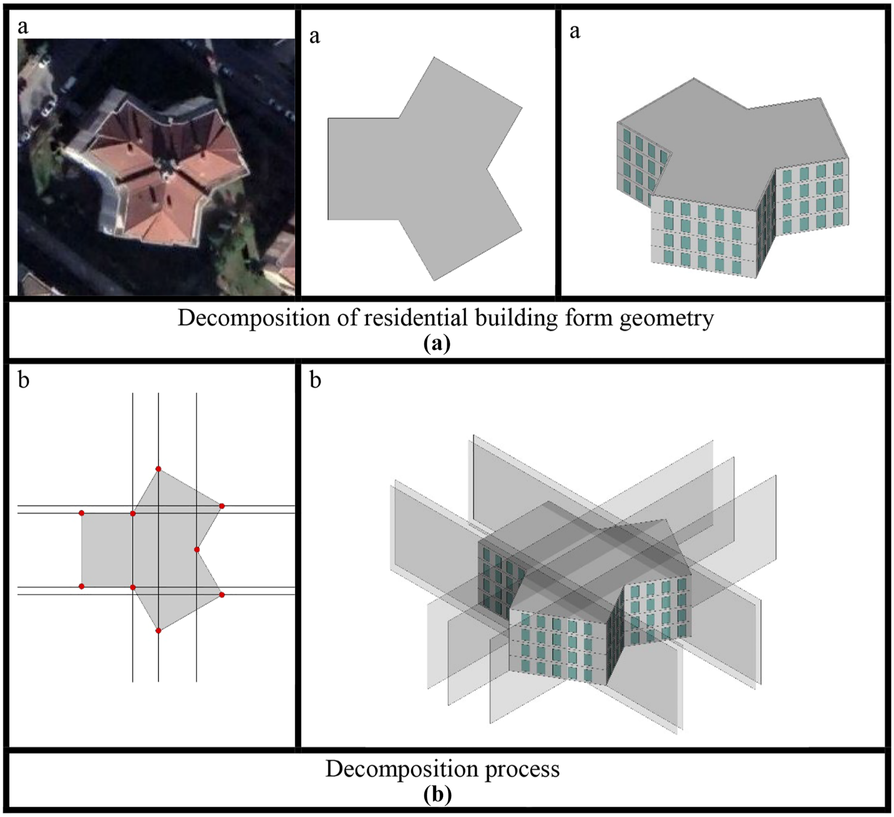
are key to improving efficiency. The aim of this research is to develop a more effective and faster workflow model for designers, particularly in the energy simulation phase of building form and geometry design during the early design stages. This study seeks to introduce a method that not only simplifies the process but also serves as a reference for machine learning (black-box) based building energy prediction models. The proposed method is designed to address the limitations of existing simulation tools by providing a structured geometric definition method, which can be applied in future AI-based surrogate and metamodel studies.

A significant challenge in this area is the complexity of architectural forms, which makes the extraction of geometric parameters difficult. To overcome this, a simplified geometric identification method has been developed. This method forms the foundation for a geometry-energy prediction framework that is applicable not only to standard building forms but also to more complex and irregular architectural designs. The goal is to create an evaluation method capable of handling diverse forms and shapes that do not fit within conventional geometric prototypes, thus providing a comprehensive approach to building energy performance analysis within a universal geometric context.

2. Material and method

The decomposition method developed in this study enables the analysis of buildings with different and complex plan geometries. In this study, the GeoBTUD (Geometric Basic Triangle and Unit Box Decomposition Method) is proposed as a novel framework for energy prediction based on building form (Taştemir *et al.*, 2024). The purpose of method is to decompose complex architectural forms for energy prediction. This method consists of dividing the geometry into triangular and rectangular-square forms by using long and short parts of the geometry of any structure. The building form in Figures 2 and 3, where the method is described, refers to a residential building form (Figures 2 and 3).

The study focuses on residential buildings due to their significant role in energy consumption, particularly in heating and cooling demand. By focusing on this building type, the study aims to validate the GeoBTUD methodology in a context where early design decisions can have a significant impact on energy performance. This study advances previous



Source(s): Created by author

Figure 2. Decomposition methodology: (a) decomposition plan geometry (b) decomposition process

research (Taştemir *et al.*, 2024) by incorporating residential case studies, thereby providing a more comprehensive and realistic assessment. In addition, the analytical methods have been extended to improve the scope of the analysis.

The GeoBTUD method was developed in Rhinoceros Grasshopper, a parametric design and modeling tool. The decomposition of building geometry using this method, as illustrated in Figure 2, involves the following steps (see Figure 4).

(1) Step 1: Separation pf plan geometry with vertical and horizontal direction.

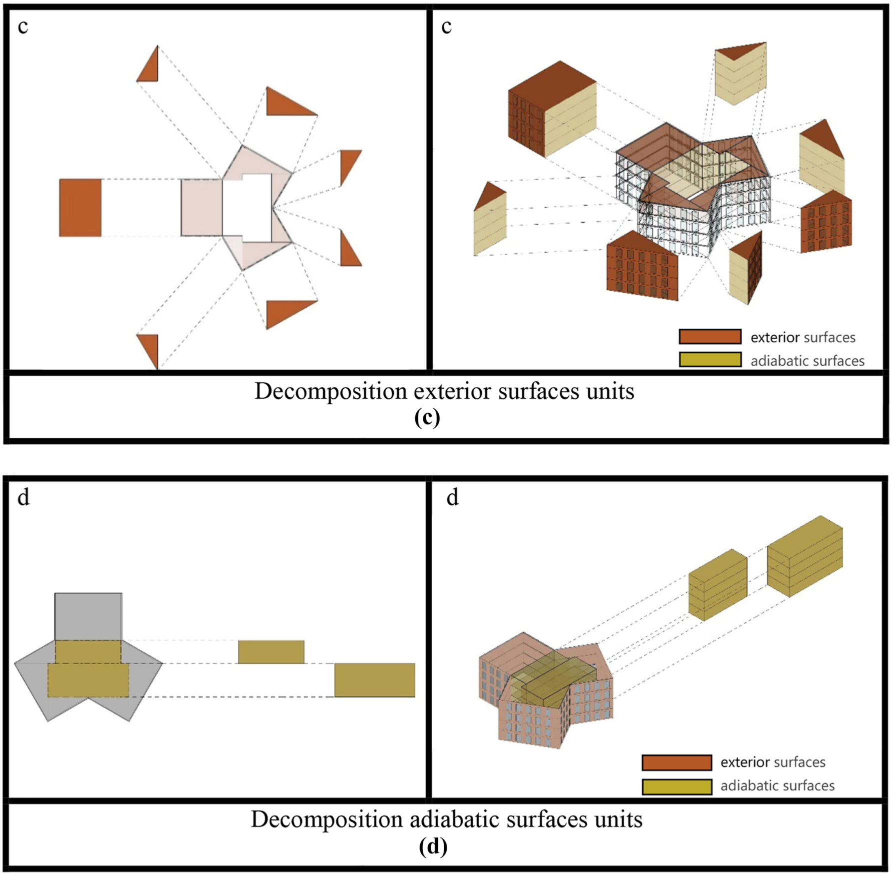
(2) Step 2: Cutting of plan geometry in vertical and horizontal direction with long edge side.

(3) Step 3: Decomposition of triangle units.

(4) Step 4: Decomposition of box-rectangle units.

2.1 GeoBTUD prediction workflow

In this study, an energy prediction method is developed using a data set generated by decomposing a complex structure using the GeoBTUD method and defining the shape and



Source(s): Created by author

Figure 3. Decomposition methodology: (c) decomposition exterior surfaces units (d) decomposition adiabatic surfaces units

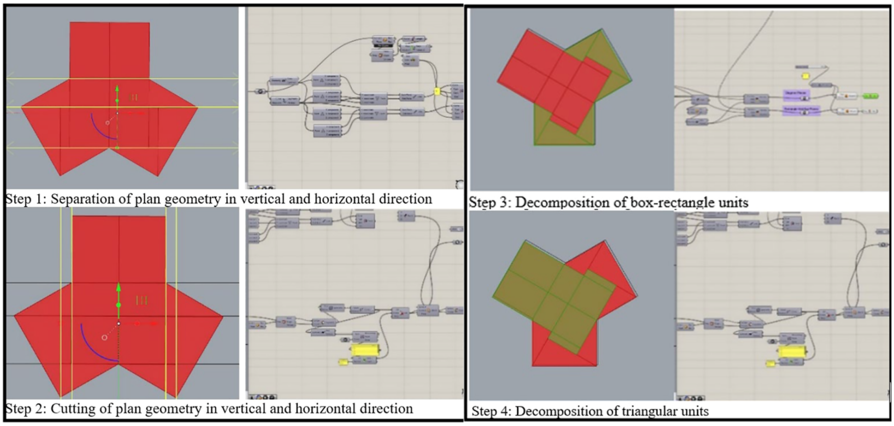
geometric properties of triangular and box-rectangular units regarding energy consumption. The workflow of this prediction framework is shown in Figure 5.

The application steps of the workflow consist of the following stages.

(1) Creating Basic Forms Energy Consumption Dataset: This step involves defining the length, width, orientation, and surface boundary values of the rectangular and triangular form variants for use in the GeoBTUD method. Once defined, energy consumption simulations for these form variants are conducted to generate the dataset.

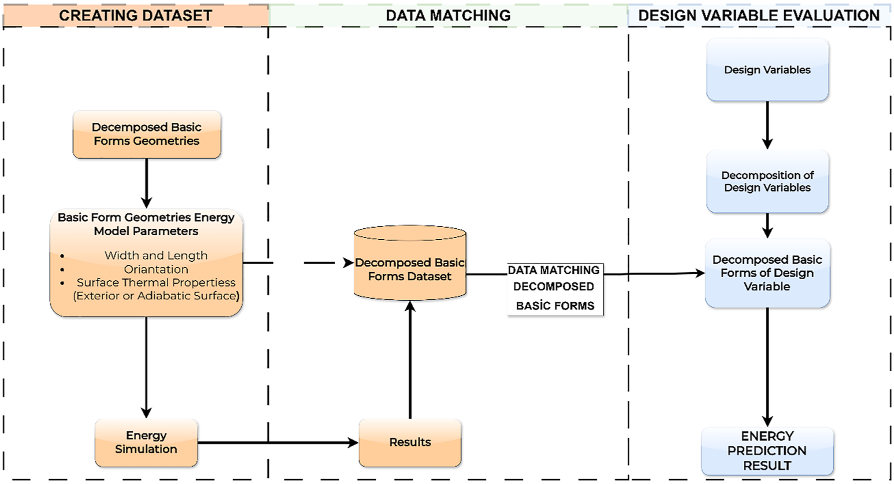
(2) Decomposition of Design Variable with GeoBTUD Method: The energy analysis generates a dataset by linking the energy consumption values of the form variants to their limit values.

(3) Data Matching Decompose Basic Forms: The design variant is broken down into simpler form parts, and the decomposition method transforms it into a structure compatible with the GeoBTUD dataset.



Source(s): Created by author

Figure 4. The decomposition of a building geometry with the method, Rhino Grasshopper



Source(s): Created by author

Figure 5. Workflow of GeoBTUD prediction method

(4) Energy Prediction Results: The energy consumption values of the relevant simple form parts are matched with the decomposed design variant, and energy consumption values are calculated according to the ASHRAE thermal zoning rules (ANSI/ASHRAE/IES Standard 90.1–(2019)).

2.2 GeoBTUD method energy simulation for dataset

According to the workflow presented in the method section of the study, building heating and cooling energy loads were calculated based on the thermal comfort condition of the simple forms, which is the first stage of the GeoBTUD decomposition method, and the dataset for the energy prediction model is created. The formation and boundary values of geometric part variations consisting of triangular and square-rectangular units are shown in Figure 4. The data related to the variable boundary values of the simulation are shown in Table 1, and the data related to the fixed boundary values are shown in Table 2.

As illustrated in Figure 6, the building geometry, decomposed using the GeoBTUD method, is categorized into two surface types: exterior surfaces and adiabatic surfaces.

Table 1. Dynamic conditions of simulation

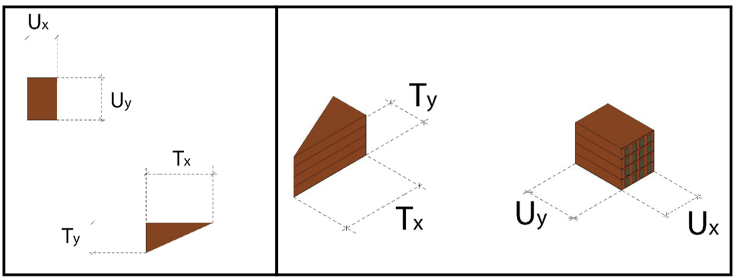
Time	Description	Value range	Value Range Change
Ux	Unit box plan width	1–10 m	1 m
Uy	Unit box plan length	1–10 m	1 m
Tx	Unit triangle plan width	1–10 m	1 m
Ty	Unit triangle plan width	1–10 m	1 m
Adiabatic surface ratio	Adiabatic surface ratio of intersect units	%0–%100	%5
Rotation	Rotation of units	0°–360°	45°

Source(s): Created by authors

Table 2. Steady conditions of model

Parameter name	Description	Value
U_wall	Total thermal transmittance of walls	0.350 W/m^{2}K
U_win	Total thermal transmittance of windows	1.96 W/m^{2}K
U_Top	Total thermal transmittance of top(roof)	0.250 W/m^{2}K
U_Floor	Total thermal transmittance of floors	2.90 W/m^{2}K
U_Ground	Total thermal transmittance of ground floor	0.250 W/m^{2}K
WWR	Window toWall Ratio of Exterior Surface	%30
Adiabatic Surfaces	The boundary conditions of adiabatic surfaces	0(full exterior)-1.0(full interior)
Activity: Residential	Activity conditions of building	Occupancy: 0.11 p/m^2 Clo: 1.0 Met: 0.90 met
HVAC settings	Building HVAC Conditions	Heating: Natural Gas Cooling: Electricity
Floor properties	Number of floors and storey heights in the building	Number of floors:6 Storey height:3m

Source(s): Created by authors



Source(s): Created by author

Figure 6. Decomposition method geometric properties

Depending on whether the building geometry is flat or diagonal, it is further divided into triangular units and box-rectangle units. The edge boundary values for the triangular units are denoted as “Tx” and “Ty,” while for the box-rectangle units, they are labeled as “Ux” and “Uy.” The floor boundary values are expressed as “U Ground” for the ground boundary of floors in contact with the earth, “U Floor” for the boundary between floors, and “U Top” for the boundary at the roof. The specific boundary values are provided in Table 2.

The energy simulations for this study were conducted using the EnergyPlus simulation engine. Istanbul climate data is utilized for the energy model simulations. The data set was generated using a parametric algorithm framework that had been tested on a variety of geometric configurations, with the capacity to reach up to 32,000 variations. This approach provided a robust and adaptable framework, yielding a consistent number of variations that were necessary for the effective evaluation of the GeoBTUD method. The data set used in the study was generated using the EnergyPlus calculation method. Which adheres to the ASHRAE 90.1 standard, ensuring that the computation approach aligns with the guidelines established by ASHRAE 90.1. Details regarding the key variables used in the energy simulation are provided in Table 1.

According to Table 1, the geometric boundary values for “Tx,” “Ty,” “Ux,” and “Uy,” which represent the edge boundaries of the units, were assigned a value range of 1 m–10 m for each edge. For the adiabatic surface ratio, the dataset was created using 20 variables with 5% increments, ranging from 0% to 100%. The rotation input was defined with 8 different angles, ranging from 0° to 360°, with 45° intervals. The data related to the basic variables in the energy simulation are shown in Table 1.

Every basic unit shape has its own unique energy consumption estimate. The lowest basic form spaces are in direct contact with the ground, while the uppermost surfaces are fully exposed to the atmosphere. In contrast, the interstory layer contains no basic form spaces that are either grounded or roofed. Furthermore, energy consumption is calculated separately for each basic form space within every layer, ensuring accurate estimations across diverse architectural forms.

In this study, the building forms were selected from residential buildings in Istanbul, Turkey, and the geometries of these forms were used in a case study to validate the proposed model. Each basic shape and space has a unique energy consumption estimate. The lowest basic form spaces are in direct contact with the ground, while the uppermost surfaces are fully exposed to the atmosphere. The interstory layer contains no basic form spaces that are either grounded or roofed. Furthermore, energy consumption is calculated separately for each basic form space in every layer, ensuring precise estimations across the different architectural forms.

The steady-state simulation conditions for the energy analysis using the GeoBTUD method are presented in Table 2. In this simulation study, the building activity function was set to “Residential Apartment Activity.” Additionally, electronic and equipment and other elements that contribute to the building’s thermal efficiency were disabled to isolate the core energy performance variables.

2.3 Limitation of study

This study employs a decomposition method that simplifies complex building geometries into basic geometric shapes. While this approach enhances computational efficiency, it may overlook finer architectural details that could significantly influence energy performance in real-world scenarios. As a result, the simplification may limit the model’s applicability to more intricate and non-standard building designs, reducing its accuracy when applied to complex architectural forms. However, in this study, the validation was conducted using a diverse range of geometries, which provided a balanced basis for assessing the model’s ability to predict energy performance for typical building forms. This approach ensures the model’s reliability within the scope of the selected geometric variations, while acknowledging potential limitations for more complex forms. The prediction model used in this study relied on a dataset with a defined limit, ensuring consistency in the evaluation process. Additionally, the energy simulations conducted in this research are based on climate data specific to Istanbul. This focus on a single climatic region may limit the generalizability of the findings to other climates or geographic areas, especially those with weather patterns that differ substantially from those in Istanbul. The outcomes derived from these simulations might not fully reflect the energy performance of buildings in other climatic contexts. Moreover, the scope of the study is primarily centered on residential buildings in Istanbul. This focus could restrict the broader application of the findings to other building types, such as commercial or industrial structures, which may have different architectural requirements and energy performance considerations. Consequently, further research involving diverse building types and regions is needed to confirm the wider applicability of the model.

3. Results and discussion

The first part of the results section presents the dataset, which was generated through energy simulations of simple form geometries performed under specific boundary conditions for the prediction model. The second part focuses on the validation of the GeoBTUD model, analyzing its accuracy and effectiveness.

3.1 Accuracy validation of GeoBTUD method

For the prediction validation of the GeoBTUD method, residential building variation cases selected (Table 3).

(1) Decomposition of Design Variable with GeoBTUD method.

(2) Data Matching Decompose Basic Forms.

(3) Energy Prediction Results.

(4) Energy Simulation of non-decompose design variables.

(5) Comparation energy loads predicted cases and simulated cases.

The underlying principle of the energy prediction model is based on aggregating the energy consumption values of the individual units created through the building variations decomposed by the GeoBTUD algorithm, similar to the methodology outlined in ASHRA 90.1 (ANSI/ASHRAE/IES Standard 90.1–(2019))

The total energy load is typically calculated using the follows (1)

$$E_{TOTAL} = \sum_{i_{unit}=1}^{n} E_i$$

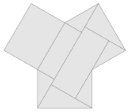
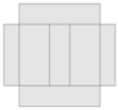
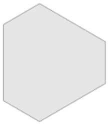
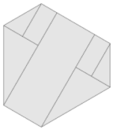
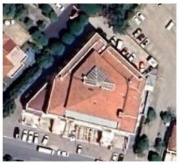
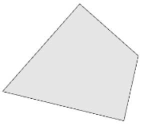
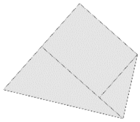
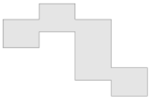
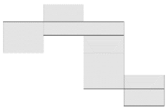
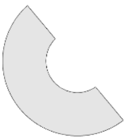
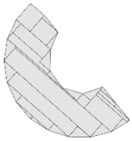
(1) E_{TOTAL}: Total energy consumption of the building.

(2) E_n: Number of decomposed thermal zone units within the decomposed building.

(3) $E_{i\,unit}$: Represents the energy consumption of each decomposed thermal zone units

Prediction model is developed using 36,000 variations through the data-matching method for energy prediction. The dataset, created with simple building geometries using the GeoBTUD method, was matched, and the heating and cooling energy loads were analyzed. Case studies were then simulated to verify the results of this analysis. The simulation inputs for the energy simulation process were consistent with those used in the dataset simulation process. The results of the energy simulation are presented in Table 4. In this study, two basic metrics are used to evaluate the prediction, accuracy rate and R2 (Cameron and Windmeijer, 1997). The accuracy rate and R2 rate calculated to compare the predicted and simulated results as follows (2) and (3):

Table 3. Case design variables properties for energy prediction validation

No	Cases Residential Buildings	Buiding Geometry	Decomposed Building Geometry	Floor Area (m^2)	Unit Number/ Unit Types
R_1				794	9/Rectangle-Box Unit Triangle Unit
R_2				916	7/Rectangle-Box Unit
R_3				1250	7/Rectangle-Box Unit Triangle Unit
R_4				2450	3/Rectangle-Box Unit Triangle Unit
R_5				6287	9/Rectangle-Box Unit
R_6				1378	28/Rectangle-Box Unit Triangle Unit

Source(s): Created by author

$$\text{Accuracy Rate} = 100\% - \left(\frac{\left|\mathrm{E}_{\text{predicted}} - \mathrm{E}_{\text{simulated}}\right|}{\mathrm{E}_{\text{simulated}}} x100\%\right) \tag{2}$$

(1) $E_{\text{predicted}}$: Total predicted energy consumption of the decomposed building variation

(2) $E_{\text{simulated}}$: Total simulation energy consumption of the whole building variation

Table 4. Building energy simulation and GeoBTUD energy

Building type	Predicted cooling energy (kWh/m^2)	Simulated cooling energy (kWh/m^2)	Cooling accuracy rate (%)	Predicted heating energy (kWh/m^2)	Simulated heating energy (kWh/m^2)	Heating accuracy rate (%)	Total predicted energy (kWh/m^2)	Total simulated energy (kWh/m^2)	Total accuracy rate (%)
R1	43.57	44.14	%97.50	69.57	70.42	%98.80	113.14	114.56	%98.74
R2	43.88	44.42	%98.80	59.59	61.24	%97.32	103.47	105.66	%97.88
R3	41.42	42.38	%97.75	46.21	45.95	%99.44	87.63	88.33	%99.20
R4	42.04	42.32	%99.44	41.08	41.02	%99.32	83.12	83.34	%99.73
R5	24.41	23.62	%96.80	49.74	47.80	%96.10	74.15	71.24	%95.91
R6	38.28	45.42	%84.30	61.53	59.53	%96.76	99.81	104.95	%94.85

Source(s): Created by authors

$$R^2 = \frac{\sum_{i=1}^{n} \left(y_{i,p} - y_{i,t}\right)^2}{\sum_{i=1}^{n} \left(y_{i,t} - \overline{y_{i,t}}\right)^2} \tag{3}$$

(1) *yi,t*: target value

(2) *yi,p*: predicted value

(3) *yi,t*: mean of the building energy respectively

As shown in Table 4 and Figure 7, the energy prediction model showed high accuracy for all building types, with cooling energy load predictions ranging from 84.30% to 99.94% and heating energy load predictions ranging from 96.10% to 99.98%. Total energy consumption prediction rates varied between the building forms, highlighting differences in thermal performance. In particular, the R6 building form had the lowest prediction accuracy (84.30% for cooling and 96.76% for heating), due to its complex geometry, which poses challenges for accurate energy performance modeling. In contrast, the R4 building form achieved the highest accuracy, with rates of 99.94% for cooling and 99.98% for heating.

These results indicate that the energy prediction model exhibits a high level of prediction accuracy. The cooling energy load prediction rate, while still high, is slightly lower than that of the heating energy load. Further analysis shows that the prediction accuracy for heating energy loads is generally higher than for cooling energy loads. Building forms with complex geometries and large surface areas have a substantial impact on energy consumption. (Figure 8)

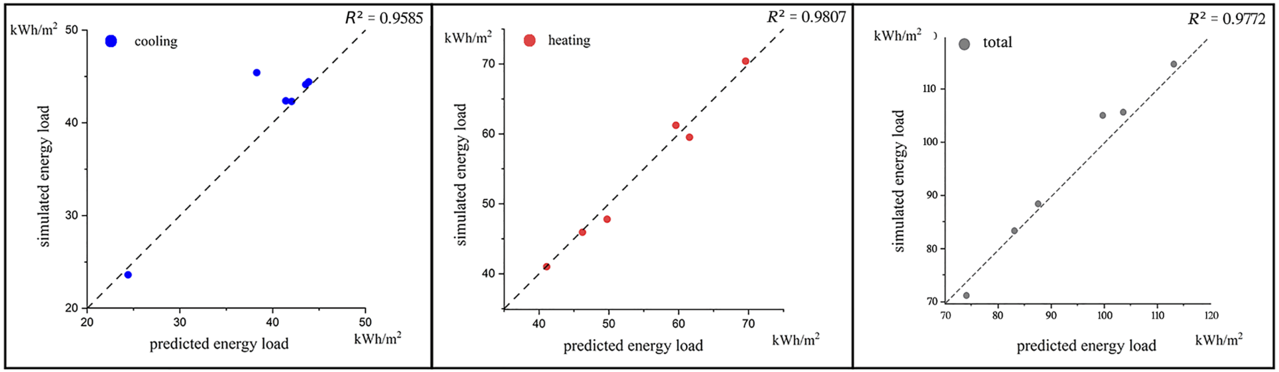
These factors influence the building's thermal performance, leading to increased energy demand for heating, cooling, and maintaining indoor comfort conditions. (D'Amico and Pompani, 2019) This reduced accuracy, particularly for cooling energy loads, can be attributed to the complex geometry of the R_6 building form, which poses challenges for accurate energy performance modeling.

4. Conclusion

The early design phase is critical for energy-efficient design, as decisions made during this stage have the greatest impact on the final performance at the lowest cost. A careful selection of building form and variables, particularly during the conceptual design phase, can enhance design performance early in the process with minimal expense. This study introduces a decomposition method for defining form geometry in building surrogate models during the early design phase. The accuracy of the decomposition method, GeoBTUD, is evaluated using design examples with various geometries. Validation analyses revealed that the method achieved a high accuracy rate, ranging from 85.40% to 99.98%. In this study, variations in accuracy were observed depending on whether the building forms were compact or fragmented. The geometric integrity and complexity of the building form emerged as significant factors affecting the accuracy of energy performance predictions. Based on these inferences, the following conclusions have been summarized:

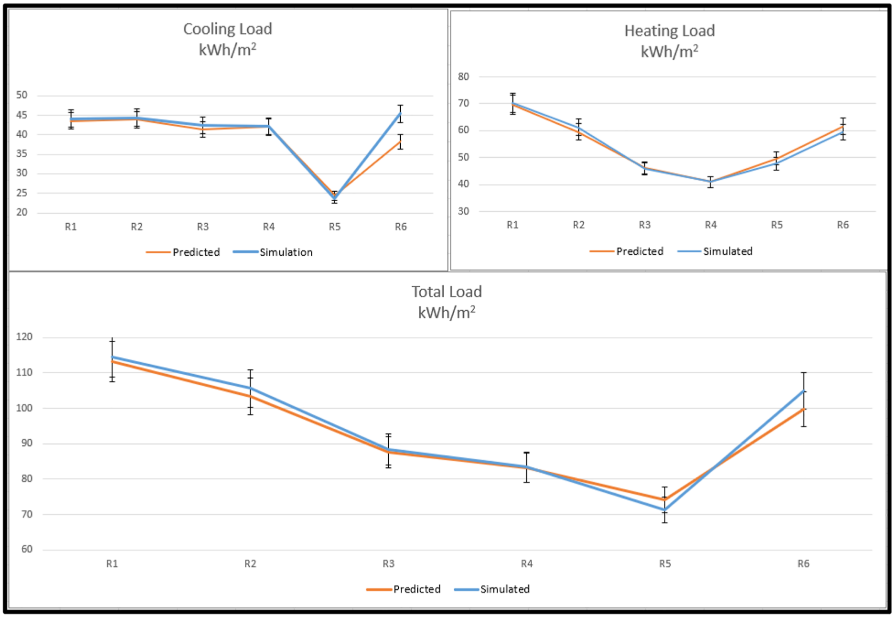
(1) As an alternative to traditional building energy simulations, energy performance prediction models based on the definition and decomposition of building geometry can provide highly accurate results in assessing comfort conditions.

(2) The validation study findings of the GeoBTUD technique, designed to assess energy loads using a simple geometric decomposition methodology, indicate that the method achieved a high level of accuracy. These results align with the initial objectives of the research, confirming the method's reliability and effectiveness. The difficulties of



Source(s): Created by author

Figure 7. Correlation analysis of cooling, heating and total energy load



Source(s): Created by author

Figure 8. Comparison of cooling, heating and total energy consumption values and prediction values

building energy simulation in the design phase of architectural design can be eliminated by using building energy prediction methods.

(3) In the validation studies conducted using the GeoBTUD method; it was observed that compact form geometries achieved higher accuracy rates compared to complex and fragmented geometries. Therefore, it is recommended that future research utilizing the developed GeoBTUD method should take into consideration the fact that simpler forms yield better results than complex geometries.

(4) In future research, the development of supplementary methods that can provide reference data for machine learning models used in building energy prediction could enhance the accuracy of energy performance forecasts. Additionally, alternative decomposition methods, as discussed in the literature, may serve as verification tools for the application and advancement of thermal comfort principles and building design strategies within energy prediction models. Moreover, the scope of future work could be broadened by employing larger data sets and incorporating a wider range of functions and variations. Such an approach would enable more extensive analyses and provide valuable insights into the versatility and robustness of the proposed method across diverse building contexts.

(5) In future studies, the development of additional methods that can provide reference and data to machine learning models used in building energy prediction models can increase the accuracy of building energy performance prediction. Different decomposition methods, which are also used in the literature, can be used as a verification method and tool for the application and development of thermal comfort and building design principles for building energy prediction models.

References

ANSI/ASHRAE/IES Standard 90.1–(2019) (n.d.), *Energy Standard for Buildings except Low-Rise Residential Buildings, Appendix G. 2019*, The American Society of Heating, Refrigerating and Air-Conditioning Engineers, Atlanta, GA.

Bobenhausen, W. (1994), *Simplified Design of HVAC Systems*, John Wiley & Sons, New York.

Ciardiello, A., Rosso, F., Dell'Olmo, J., Ciancio, V., Ferrero, M. and Salata, F. (2020), "Multi-objective approach to the optimization of shape and envelope in building energy design", *Applied Energy*, Vol. 280, 115984, doi: 10.1016/J.APENERGY.2020.115984.

Cameron, A.C. and Windmeijer, F.A.G. (1997), "An R-squared measure of goodness of fit for some common nonlinear regression models", *Journal of Economics*, Vol. 1997 No. 77, pp. 329-342.

Chen, Y., Hong, T. and Piette, M.A. (2017), "Automatic generation and simulation of urban building energy models based on city datasets for city-scale building retrofit analysis", *Applied Energy*, Vol. 205, pp. 323-335, doi: 10.1016/j.apenergy.2017.07.128.

D'Amico, B. and Pomponi, F. (2019), "A compactness measure of sustainable building forms", *Royal Society Open Science*, Vol. 6 No. 6, doi: 10.1098/rsos.181265.

Depecker, P., Menezo, C., Virgone, J. and Lepers, S. (2021), "Design of buildings shape and energetic consumption", *Building and Environment*, Vol. 36 No. 5, pp. 627-635, June 2001, doi: 10.1016/S0360-1323(00)00044-5.

Dogan, T., Reinhart, C. and Michalatos, P. (2016), "Autozoner: an algorithm for automatic thermal zoning of buildings with unknown interior space definitions", *Journal of Building Performance Simulation*, Vol. 9 No. 2, pp. 176-189, doi: 10.1080/19401493.2015.1006527.

Forouzandeh, N., Zomorodian, M., Shaghaghian, Z. and Tahsildoost, M. (2022), "Room energy demand and thermal comfort predictions in early stages of design based on the machine learning methods", *Intelligent Buildings International*, Vol. 15, pp. 1-18, doi: 10.1080/17508975.2022.2049190.

Hamdaoui, M.A., Benzaama, M.H., El Mendili, Y. and Chateigner, D. (2021), "A review on physical and data-driven modeling of buildings hygrothermal behavior: models, approaches and simulation tools", *Energy and Buildings*, Vol. 2021 No. 251, 111343, doi: 10.1016/j.enbuild.2021.111343.

IEA (2018), *World Energy Outlook 2018*, IEA, Paris, available at: https://www.iea.org/reports/world-energy-outlook-2018

İşeri, O.K. and Dursun, O. (2022), "The impacts of early architectural design decisions on building performance", *International Journal of Digital Innovation in the Built Environment*, Vol. 11 No. 2, pp. 1-21, doi: 10.4018/IJDIBE.301245.

Kahraman, Ö. and Köymen, E. (2023), "The influence of building form compactness on energy efficiency in accommodation structures: the case of Türkiye", *Sakarya University Journal of Science*, Vol. 27 No. 5, pp. 1055-1078, doi: 10.16984/saufenbilder.1253136.

Konis, K., Gamas, A. and Kensek, K. (2016), "Passive performance and building form: an optimization framework for early-stage design support", *Solar Energy*, Vol. 125, pp. 161-179, doi: 10.1016/j.solener.2015.12.020.

Li, Z., Dai, J., Chen, H. and Lin, B. (2019), "An ANN-based fast building energy consumption prediction method for complex architectural form at the early design stage", *Building Simulation*, Vol. 12 No. 4, pp. 665-681, doi: 10.1007/s12273-019-0538-0.

Mehrotra, A. and Yi, H. (2023), "Effect of adaptive intelligent sampling and machine-learning emulators in surrogate energy modeling of architectural massing", *Journal of Building Engineering*, Vol. 72, 106614, 2352-7102, doi: 10.1016/j.jobe.2023.106614.

Olu-Ajayi, R., Alaka, H., Sulaimon, I., Sunmola, F. and Ajayi, S. (2021), "Building energy consumption prediction for residential buildings using deep learning and other machine learning techniques", *Journal of Building Engineering*, Vol. 45, 103406, doi: 10.1016/j.jobe.2021.103406.

Pittarello, M., Scarpa, M., Ruggeri, A.G., Gabrielli, L. and Schibuola, L. (2021), “Artificial neural networks to optimize zero energy building (Zeb) projects from the early design stages”, *Applied Sciences*, Vol. 11 No. 12, p. 2021, doi: 10.3390/app11125377.

Salami, Babatunde Abiodun, Abba, Sani I., Adewumi, Adeshina A., Dodo, Usman Alhaji, Otukogbe, Ganiyu K. and Oyedele, Lukumon O. (2023), “Building energy loads prediction using bayesian-based metaheuristic optimized-explainable tree-based model”, *Case Studies in Construction Materials*, Vol. 19, p. e02676.

Sing, M. (2022), “Early Stage Design Support Using Machine Learning and Building Information Modelling”, PhD thesis, KU Leuven University.

Singh, M. and Smith, I. (2023), “Convolutional neural network to learn building-shape representations for early-stage energy design”, *Energy and AI*, Vol. 14, 100293, doi: 10.1016/j.egyai.2023.100293.

Sun, Y., Haghighat, F. and Fung, B.C.M. (2020), “A review of the-state-of-the-art in data-driven approaches for building energy prediction”, *Energy and Buildings*, Vol. 2020 No. 221, 110022, doi: 10.1016/j.enbuild.2020.110022.

Taştemir, İbrahim Agah, Köymen, Erdem and Yaşa, Enes (2024), “A Decomposition Method for Energy Prediction Metamodels and Surrogate Models in Early Design Stage of Buildings”, *Lecture Notes in Civil Engineering*, Springer Nature Singapore, Singapore, pp. 500-511.

Wang, W., Zmeureanu, R. and Rivard, H. (2005), “Applying multi-objective genetic algorithms in green building design optimization”, *Building and Environment*, Vol. 40 No. 11, pp. 1512-1525, doi: 10.1016/j.buildenv.2004.11.017.

Wilde, Pieter (2018), *Building Performance Analysis*, Wiley.

Yasa, E. (2016), “Microclimatic comfort measurements evaluation of building physics: the effect of building form and building settled area, on pedestrian level comfort around buildings”, *Journal of Building Physics*, Vol. 2 No. 5, pp. 1-29, doi: 10.1177/1744259115621979.

Zhang, L., Jin, G., Liu, T. and Zhang, R. (2022), “Generalized hierarchical expected improvement method based on black-box functions of adaptive search strategy”, *Applied Mathematical Modelling*, Vol. 2022 No. 106, pp. 30-44, doi: 10.1016/j.apm.2021.12.041.

Zhu, S., Ma, C., Zhang, Y. and Xiang, K. (2021), “A hybrid metamodel-based method for quick energy prediction in the early design stage”, *Journal of Cleaner Production*, Vol. 320, 128825, doi: 10.1016/J.JCLEPRO.2021.128825.

Corresponding author
İbrahim Agah Taştemir can be contacted at: ibrahim.tastemir@izu.edu.tr