

Adaptive Approaches In Urban Architectural Design For Enhancing Energy-Conscious Performance

Hefang Xu*

Henan Technical College of Construction, Zhengzhou, Henan, 450000, China

* Corresponding author. E-mail: 18336303921@163.com

Received August 6, 2025; accepted July 15, 2026

Estimation of energy consumption in urban areas, particularly in relation to the Universal Thermal Climate Index, and related environmental factors associated with architectural design, has remained an essential challenge in the related field of urban planning. This study, holding interest in the first three advanced algorithms, namely HGBR, CatBoost, and LightGBM, was proposed to explore related possibilities in using machine learning algorithms for estimation. The used algorithms were analyzed using three different architectural setups (types A, B, and C) that were optimized using the HGS optimization algorithm. The obtained results clearly highlight that the HGBR-HGS system was significantly accurate for type A buildings for cooling load calculations and estimates related to the Universal Thermal Climate Index. The CatBoost-HGS system performed better than the HGBR-HGS system for type B buildings, again forming the top-performing system for cooling loads. Type C buildings also followed similar lines, where again the CatBoost-HGS system performed better for calculations related to Universal Thermal Climate Index, while the HGBR-HGS system again formed the top-performing system for cooling loads. The results demonstrate that the HGBR-HGS model achieved superior performance, with R^2 values reaching up to 0.9994 for cooling load prediction and 0.9880 for UTCI estimation, along with the lowest RMSE values across most scenarios. Additionally, the CatBoost-HGS model exhibited strong performance in predicting UTCI, achieving R^2 values up to 0.9875 for specific building types.

Keywords: Cooling loads, Universal Thermal Climate Index, LightGBM, CatBoost, HGBR, HGS.

© The Author(s). This is an open-access article distributed under the terms of the [Creative Commons Attribution License \(CC BY 4.0\)](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are cited.

http://dx.doi.org/10.6180/jase.202701_36.022

1. Introduction

The meteorological elements, including winds, temperatures, humidity, and others, that change from time to time significantly affect human life in various ways. Accurate predictions for future meteorological conditions are, therefore, important and find numerous applications in various disciplines. The ever-rising use of Earth observation satellites and improvements in climate modeling allow meteorological researchers to gain access to abundant meteorological information. The analysis of atmospheric conditions, from the perspective that is applicable to human body chronophysiological activity, is important for biometeorol-

ogy. Air temperature alone is not important, but it is also not the only consideration. During the past 150 years, more than 100 basic thermal parameters, mainly two-variable parameters, were proposed to describe human-environment complex heat interactions [1–3].

The present outcome has implications for making decisions on whether to continue operations, affecting the strategic planning process for business people in various industries, along with stock, financial, and promotional matters. The accurate predictions made will enable the effective use of space, resources, and promotional services, thereby increasing retailers' financial income. Nowadays, through digitization in the sales process and the quick ad-

vancement in machine learning algorithms, businessmen in small, medium, or big ventures can perform the estimates, thereby attracting considerable attention in research studies for their applications [4, 5]. They can also be used for the prevention of conditions like excess merchandise stocked, thereby causing popular merchandise, leading to lower production costs.

For human biometeorological analysis, it is important to evaluate the thermal environment realistically, particularly when analyzing heat comfort, heat discomfort, and health impacts. The expression 'biothermal environment' covers both factors that cause heat interactions in the atmosphere and the associated body responses. An intensely effective autonomous system for heat regulation plays a major role in restoring heat balance in human bodies by adjusting the body's heat regulation according to varying heat interactions in the atmosphere. This process is also supported by modifications caused by human responses related to discomfort factors [6–8].

The project for developing a Universal Thermal Climate Index (UTCI) using advanced models of human thermophysiological responses was initiated by the International Society of Biometeorology (ISB) in 1999. Later, in 2005, the European COST Action 730 supported achieving the same by pooling experts from different disciplines, and finally, in 2009, it resulted in the successful development of the UTCI [9]. To enable a comprehensive expression of human physiological responses through the complex outdoor temperature environment, the Universal Thermal Climate Index, a tool for evaluating outdoor thermal conditions concerned with human biometeorology [10], has overcome the limitations in human heat budget methods by integrating an adaptive clothing model along with a multi-node human model for thermoregulation [11–13].

This concept has been utilized by Hope et al. [14] to respond to the difference in assessing indoor versus outdoor climatic conditions. The outdoor conditions are characterized by broad ranges in dynamic responses for human reaction to different outdoor environments. This is especially the case when reacting to cold climates, wherein attaining a stable state could be difficult, thereby requiring several hours. Fiala et al. [15] examined multi-segmental thermoregulatory models, albeit concentrating more intensely on them in their publication. Among various human body thermoregulatory models, the complex multi-node "Fiala model" was chosen. This model also estimates dynamic thermal responses, utilizing information from body core and skin temperatures. Under the COST Action 730, the biological model was thoroughly validated, modified, and improved according to the project objectives [15,

16]. Havenith et al. [17] brought in an advanced model for adapting to clothing systems in the framework. The process of continuously running the thermoregulatory model for calculating the equivalent temperatures for UTCI could pose challenges, requiring expert knowledge in managing complex simulation software.

The calculation process for the values of the Universal Thermal Climate Index (UTCI) by iterative calculations can be time-consuming. This can affect modeling for climates and forecasting the weather. Among the methods used to overcome such limitations, the relevance for classifying the values for the Universal Thermal Climate Index on the basis of thermal stress has been discussed by Koppe and Jendritzky [18] and Staiger et al. [19]. Krüger et al. [20] state that there is a constraint on UTCI measurements. When they looked into two subtropical areas, they found lower UTCI values. These values did not meet the established thermal comfort band thresholds.

Ibrahim et al. [5] optimized the courtyard blocks' factors for Cairo's dry and hot climate using Grasshopper evolutionary search. The study points towards the optimized solution, leading to a possible cooling load reduction by 31.7% and Universal Thermal Climate Index by 1.6 °C. The study also proposed that the number of surfaces could improve by 4% in optimized designs, beneficial for heating loads during winter. Xu et al. [21] analyzed the carbon-neutral architectural design strategies used by the Jin'an shopping mall in Harbin, particularly from the perspective of energy conservation, whereby the shape of the atrium was utilized. The use of computational fluid dynamics enabled them to assess thermal comfort, using both mean radiant temperature and UTCI. The resilient ventilation system, without using air-conditioned units, caused a significant energy saving of 7.34% to 9.64% when turned on.

Zhong et al. [22] looked into how to enhance outdoor environment prediction models using hybrid-driven approaches. In the context of a simulation of the Universal Thermal Climate Index (UTCI), they looked into how machine learning algorithms interpret spatial data. For the study, a number of neural network architectures were created, and the impact of model parameters on performance was investigated. They found that multilayer networks with tiny convolutional kernels and more training data improved prediction performance. Interpretability analysis also indicated that this approach might be extended to wind and heat environment forecasting. In their analysis of East Chinese worker villages, Zhong et al. [23] focused on optimizing thermal comfort and wind. They were accurate in predicting wind ratios and values for UTCI in the settlement using machine learning methods, specifi-

cally the Adaboost model. The findings revealed that there were varying experiences for communities in Zhejiang and Jiangsu provinces when it came to the climate. Machine learning was also beneficial in evaluating large outdoor zones in urban renewal schemes.

Though having ubiquitous use, these methods face a loss of sensitivity when used in regions with extreme climates. Findings from various research studies indicate that the reasons for such loss of sensitivity could be related to discrepancies in thermal comfort factors, mostly experimental and human factors [24]. To assess outdoor conditions in regions with harsh climates, it is essential to use a reliable thermal index that can help create better outdoor spaces in regions with harsh climates [25].

The forecasting of actual building cooling loads for a short period is highly significant for various energy management operations in buildings, such as fault detection, demand-side management, and optimization. The traditional methods, that is, principle-based methods, lack accuracy since they are based on assumptions. The increasing popularity can be largely justified by the increasing amount of building data, along with the flexibility in the model used. The significance of energy load forecasting for buildings, particularly for cooling-load buildings in southern Poland, was illustrated by a study carried out by Borowski et al. [26]. This study used time, occupation, and meteorological factors, along with neural networks, support vector machines, and support vector machines, respectively, for the reasonable model, showing a WAPE of 19.93% and CV of 27.03%; it states the superiority of neural networks for accuracy. The recent advancement in the domain of data science has made better and more efficient methods possible for problems like forecasting [27]. An innovative model for building energy forecasting using artificial neural networks that can provide forecasts for different user preferences was proposed by Li et al. [28]. This method incorporates real-time cooling load and performance prediction along with model recommendations. To take user preferences into account, the study used data from 40 cases in five factory buildings and multi-objective decision-making algorithms. The study concluded that it is often best to combine prediction with a monitoring system and identified which prediction models performed best in different scenarios. This work lays the groundwork for energy prediction applications based on user preferences. The XGBoost model outperformed the others with an R^2 of roughly 97%. Feng et al. [29] used building details, weather data, and utility data to develop an accurate prediction model for residential cooling energy usage. With the help of this user-friendly model, homeowners can make

energy-saving decisions by better understanding energy performance without the need for additional hardware.

The current study continues the research process performed by Ibrahim et al. [5] for multi-objective optimization using an evolutionary algorithm for urban courtyard blocks. Contrary to the objectives proposed by Ibrahim et al., instead of optimizing design variables for energy efficiency, the proposed study will bring in a paradigm shift in favor of modeling through advanced machine learning algorithms. The objectives for using machine learning algorithms will include predicting energy consumption variables such as UTCI and cooling load for three different architectural designs. The theme for the proposed study will include the determination of the accuracy level for the proposed models, comparative analysis for the different types, and applicability using hybrid optimization strategies (HGS) for energy forecasting. This proposed study will form a basis for future implementations for energy-efficient designs through similar models for real-time adaptation in designing smarter, energy-efficient cities. This proposed study will also highlight the importance of predicting instead of optimizing through the validation process for efficient forecasting in energy optimization through advanced machine learning strategies.

1.1. Main Novelty

Urban structures are ever-changing, so a shift from traditional, static building design strategies to more dynamic, real-time architectural strategies is required. Though the research presented in this paper focuses on modeling and validation using a static dataset, the modeling has its applications in future real-time architectures that can adapt to different situations. The research presents a framework for using real-time environmental inputs to dynamically change building structures for improved energy efficiency. With access to advanced machine learning algorithms, it would be a blend of HGBR, CatBoost, and LightGBM algorithms that this system will intelligently respond to various situations. For example, look at the case when an unpredictable heat wave emerges in the city, and the system will dynamically change architectural structures that will enable comfortable interior temperatures without putting too much load on the air-conditioned system. The system is unique because it will blend the use of modeling that is optimized by the use of the HGS algorithm, along with the potential for future adjustments for real-time responses. Additionally, it can make buildings in urban areas resistant to various weather conditions.

2. Methodology

This research simplifies urban configurations by modeling buildings using straightforward design variables for assessing architectural functionality. During early design phases, this technique helps identify how urban shape affects its surroundings. Multi-zone buildings' energy simulations are more time-consuming than optimization calculations, leading to the “dimensionality dilemma” where the number of possible design variables increases, making it difficult to design. The energy efficiency analysis for three types of courtyard designs, consisting of 8, 10, and 12 buildings (A, B, and C), will also be carried out in the study. The advanced algorithms, HGBR, CatBoost, and LightGBM, were applied for energy consumption predictions in the aforementioned designs. The HGS optimizer was utilized for tuning algorithmic parameters for improved accuracy. The correlation matrix will allow assessing the relationship among the input parameters and the output variables, namely, the Urban Terrain Correction Index (UTCI) and the Cooling output.

The dataset was divided evenly for training and testing (80%-20%). The statistical measures will follow for model validation, and the algorithm that performs best in predicting energy usage will be identified.

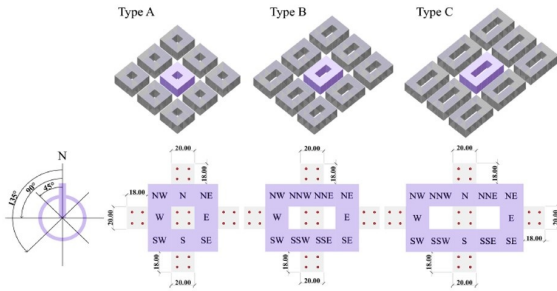


Fig. 1. Types of courtyards investigated in this study [5]

2.1. Data collection

The dataset used in this analysis was obtained from a reference [5]. Even though direct involvement in the actual operation of the equipment or in assessing the level of uncertainties in the dataset was absent, it is guaranteed that the actual source has validated the dataset. Preprocessing for each entry has also been carried out to further enhance the reliability of the dataset. The time formats and units were carefully checked to ensure consistency.

The dataset for the study consists of 13 input variables and two output variables—UTCI and Cooling. The input variables include the heights of the building blocks in different directions, the distances between them, and their

orientations, as can be clearly seen in Figure 1.

2.2. Machine learning methods

This section reviews the machine learning methodologies used and the optimization approach implemented to boost their efficiency. The study integrates the CatBoost, LightGBM, and HGBR algorithms, utilizing HGS as the chosen optimization technique.

2.2.1. Light Gradient Boosting Machine (LightGBM)

Ke et al. [30] introduced LightGBM, an innovative Gradient Boost Decision Tree (GBDT) algorithm in 2017. This algorithm has also found its use in different data mining applications, for example, ordering, regression, and classification. The LightGBM algorithm has these two distinctive characteristics: gradient-based one-sided sampling and exclusive feature bundling.

2.2.2. Categorical Gradient Boosting (Catboost)

Capability to Deal with Categorical Features

The standard GBDT algorithm does not support categorical variables inherently. Therefore, if the user is interested in using GBDT, processing the categorical variables prior to using them in the algorithm is mandatory. One-hot encoding is the commonly used technique for processing categorical variables, but it may lead to sparse features. However, CatBoost, a variant of the gradient boosting algorithm, has the ability to process categorical variables without using the concept of one-hot encoding. This algorithm carries out the conversion process of the categorical variable to a numerical attribute by using a process called the “ordered target statistic”. CatBoost has a technique that can forecast the expected output for each class in a categorical variable [31, 32].

Feature Combinations

CatBoost handles combinations of categorical variables as additional attributes. This allows it to model more complex patterns. While searching for the next split, it adopts a greedy algorithm, exploring all possible combinations for additional attributes. This technique helps CatBoost perform better than GBDT regarding model accuracy, and it helps reduce the risk of overfitting too [33].

2.2.3. Histogram gradient boosting regressor (HGBR)

HGBR is an ML algorithm that combines gradient boosting and DTs, utilizing a process called histogram binning, which deals well with big datasets [34, 35]. The process involves creating decision trees to approximate a function that relates the input space to the gradient space, thereby enhancing the accuracy of the learning process by iterating

several times. The main goal behind HGBR is to optimize the decision trees so that the loss function is reduced. With every tree added, efforts are made to improve the residuals, and finally, a combination model consisting of optimized trees is formed, ensuring that it can provide a good regression forecast and perform well on huge datasets [36].

2.2.4. Hunger Games Search (HGS)

HGS (Hunger Games Search) is a nature-inspired optimization algorithm that simulates the behaviors of animals driven by hunger [37]. It starts by generating an initial population of potential solutions. The fitness of each solution is evaluated using an objective function, and the optimal solution is identified. The algorithm then adjusts the solutions based on cooperative behavior among the individuals in the population, mimicking the search for food in nature. The objective of HGS is to optimize solutions by iteratively refining them through a process that mimics the competition and cooperation observed in animals. The algorithm aims to find the best solution by adjusting the population in a way that maximizes fitness, inspired by natural hunger-driven behavior.

It is worth noting that the extensive formula derivations of Section 2.2. has been moved to the supplementary file attached to the paper.

2.3. Model evaluation criteria

To ensure the trustworthiness of the suggested models, a variety of evaluation metrics are employed. These metrics measure the difference between the actual and estimated values, focusing primarily on their error margins [38]. The utilized metrics encompass Mean Absolute Error (MAE), Root Mean Square Error (RMSE), Relative Absolute Error (RAE) and Coefficient of Determination (R^2).

$$\text{MAE} = \frac{\sum_{i=1}^n |y_i - \hat{y}_i|}{n} \quad (1)$$

$$\text{RMSE} = \sqrt{\frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{n}} \quad (2)$$

$$\text{RAE} = \frac{[\sum_{i=1}^n (\hat{y}_i - y_i)^2]^{\frac{1}{2}}}{[\sum_{i=1}^n (y_i)^2]^{\frac{1}{2}}} \quad (3)$$

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \quad (4)$$

3. Results and discussion

This section focuses on the outcome analysis related to proposed models for energy estimation for three different types. As previously discussed, to predict the values for UTCI and Cooling, three different hybrid models, namely

CatBoost-HGS, HGBR-HGS, and LightGBM-HGS, have been used. An in-depth discussion related to the aforementioned will be carried out, and finally, by making use of different statistical measures, the best models will also be determined.

To study the correlations of the input parameters and their impacts on the output, a correlation matrix has been made for both input and output values. The input parameters include building block heights in different directions, distances, orientations, and the Floor Area Ratio (FAR). The types covered by this study, as identified in the methodology section, include three different types, and each type has its corresponding correlation matrix. The range for the values in the matrices is from -1 to $+1$. The positive values indicate that the factors are directly proportional, and the negative values indicate that the factors are inversely related to each other. The values in the matrices are also coded from blue to red, where red represents positive factors and blue represents negative factors.

The figure for Type A is Figure 2. Analyzing the figure, it can be identified that “Distance” has the highest positive impact on UTCIs. The factors for negative correlations were found to be FAR and Height(w). However, for Cooling, a majority of factors are directly related, and Height(w) has the highest, whereas only “Cooling” has negative factors, which is absolutely sensible since it opposes “UTCI”. Observing the UTCI row reveals a negative correlation with Cooling.

Figure 3 relates to Type B. In this type, as is apparent, the FAR parameter has increased, and its area has changed, considerably influencing the outputs. Based on the matrix for this type, the FAR parameter exhibits the strongest inverse relationship with UTCI, even surpassing Cooling. This effect is also noticeable in the Cooling output, where its influence is almost direct, and most parameters demonstrate a direct effect. For UTCI, the effects of the parameters are largely inverse, except for the Distance parameter, which displays a direct positive correlation.

Figure 4 pertains to type C. According to this representation, the influence of the FAR parameter on the outputs is quite evident. The FAR parameter exhibits the most pronounced negative correlation with UTCI and a direct positive relationship with Cooling. In type C, the effects of parameters Height(N) and Height(S) on UTCI have amplified. Concerning the Cooling output, the FAR parameter has demonstrated the most significant impact. While most parameters have rendered a positive effect, the Orientation parameter has exhibited a slight inverse influence, reversing its overall effect on the output.

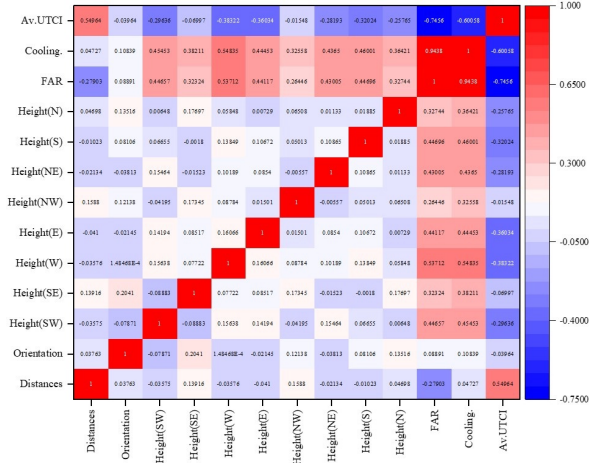


Fig. 2. The correlation matrix of features for type A

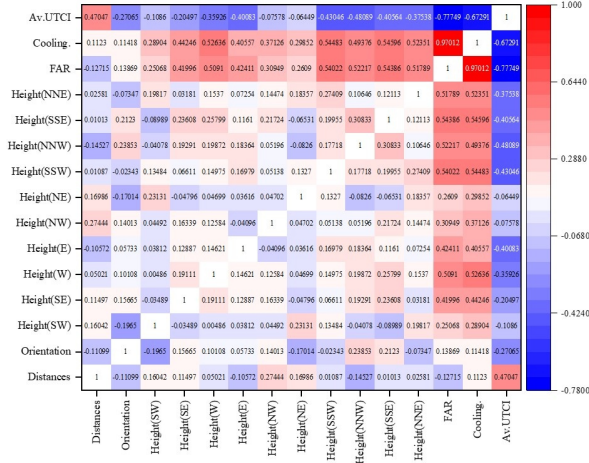


Fig. 3. The correlation matrix of features for type B

To thoroughly and accurately assess the performance of the proposed hybrid models, scatter plots have been generated for each. Additionally, the R^2 index has been specified for the test set of each model. Figure 5 displays the related plots for type A. Based on this figure, for Cooling, the hybrid model HGBR-HGS with $R^2=0.9990$ and for UTCI, the same model with $R^2=0.9880$ stands out as the best. Figure 6 shows the scatter plots for type B. According to this representation, for Cooling, the hybrid model HGBR-HGS exhibits an R^2 of 0.9992, while for UTCI, the CatBoost-HGS model achieves an R^2 of 0.9868.

$R^2=0.9994$ and the hybrid CatBoost-HGS model with $R^2=0.9875$ emerge as the most optimal for predicting Cooling and UTCI, respectively.

In addition to the provided charts and indices, the model's accuracy has been assessed using a range of statistical metrics. Table ?? displays the error metrics for types A, B, and C. In these charts, errors for both outputs, Cool-

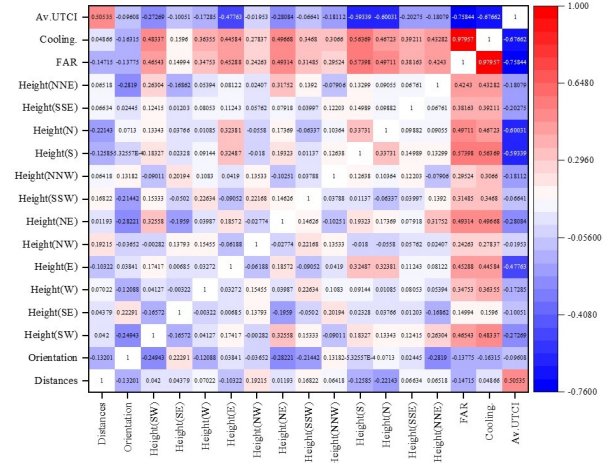


Fig. 4. The correlation matrix of features for type C

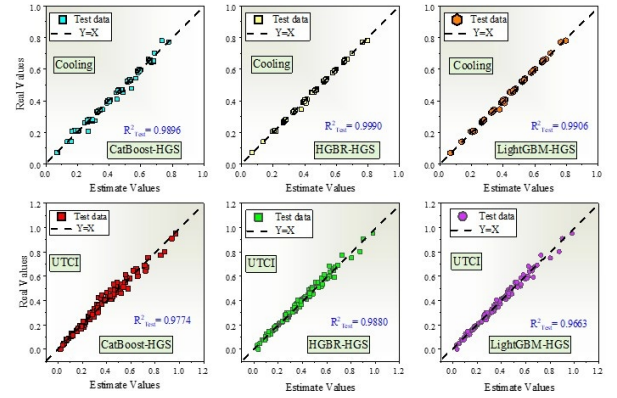


Fig. 5. Scatter plot of the observation-prediction for CatBoost, HGBR, and LightGBM hybrid models in type A

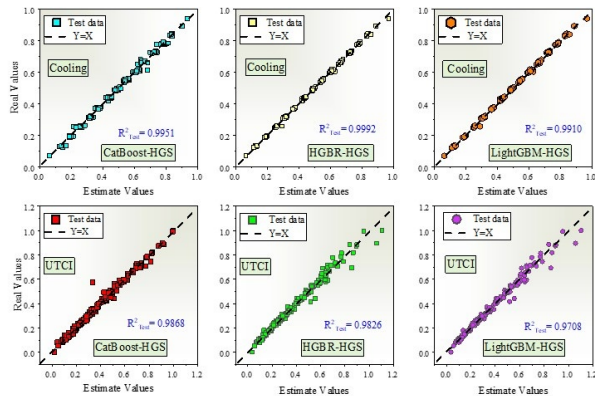
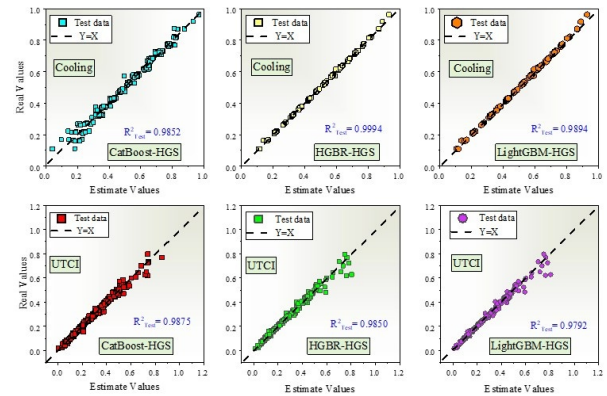
ing and UTCI, are clearly delineated. For type A, from Type A, when analyzing the metric RMSE, it is clear that for both Cooling and UTCI, the hybrid model that has the minimum amount of error is the hybrid HGBR-HGS model. This same conclusion can also be arrived at for type B, from Type B, wherein the hybrid HGBR-HGS model has the minimum amount of error for 'Cooling,' while the hybrid CatBoost-HGS model has the minimum amount of error for 'UTCI.' Moving on to type C, from Type C, it can once again be noticed that the same pattern has been followed.

4. Discussion

The analysis of the results has made it clear that the HGBR-HGS model performed better, ensuring that the minimum RMSE and maximum R^2 values were achieved in almost all situations, especially when it comes to the cooling load calculation. This can largely be attributed to the numerical variable processing technique adopted by the HGBR algo-

Table 1. Error metrics for proposed hybrid models.

		RMSE		MAE		RAE		R ²	
		Cooling	UTCI	Cooling	UTCI	Cooling	UTCI	Cooling	UTCI
Type A	CatBoost	0.0189	0.0440	0.0120	0.0295	0.1010	0.1950	0.9680	0.9320
	HGBR	0.0061	0.0249	0.0030	0.0180	0.0320	0.1090	0.9685	0.9450
	LightGBM	0.0208	0.0398	0.0153	0.0305	0.0995	0.1690	0.9550	0.9150
	CatBoost-HGS	0.0162	0.0303	0.0104	0.02113	0.0797	0.1325	0.9896	0.9774
	HGBR-HGS	0.0049	0.0214	0.0025	0.01595	0.0196	0.1007	0.9990	0.9880
	LightGBM-HGS	0.0150	0.0365	0.0117	0.023916	0.0928	0.1582	0.9906	0.9663
Type B	CatBoost	0.0152	0.0275	0.0115	0.01598	0.0723	0.0950	0.9540	0.9140
	HGBR	0.0066	0.0305	0.0032	0.01788	0.0291	0.1130	0.9590	0.9220
	LightGBM	0.0201	0.0370	0.0145	0.02379	0.0980	0.1950	0.9510	0.9010
	CatBoost-HGS	0.0130	0.0250	0.0091	0.013767	0.0598	0.0771	0.9951	0.9868
	HGBR-HGS	0.0051	0.0292	0.0027	0.016691	0.0175	0.0918	0.9992	0.9826
	LightGBM-HGS	0.0176	0.0367	0.0133	0.022920	0.0871	0.1321	0.9910	0.9708
Type C	CatBoost	0.0241	0.0205	0.0153	0.019100	0.1021	0.0892	0.9442	0.9662
	HGBR	0.0065	0.0255	0.0029	0.012200	0.0145	0.0795	0.9591	0.9651
	LightGBM	0.0198	0.0269	0.0142	0.016900	0.0991	0.1260	0.9492	0.9590
	CatBoost-HGS	0.0210	0.0181	0.0140	0.010532	0.0985	0.0816	0.9852	0.9875
	HGBR-HGS	0.0040	0.0205	0.0019	0.010241	0.0139	0.0784	0.9994	0.9850
	LightGBM-HGS	0.0182	0.0227	0.0131	0.014826	0.0949	0.1154	0.9894	0.9792

**Fig. 6.** Scatter plot of the observation-prediction for CatBoost, HGBR, and LightGBM hybrid models in type B**Fig. 7.** Scatter plot of the observation-prediction for CatBoost, HGBR, and LightGBM hybrid models in type C

algorithm. The HGBR algorithm is extremely useful for numerical variables since it adopts a histogram-based gradient boost algorithm, thereby ensuring that it can explore and analyze the trends in the numerical dataset in an efficient manner. The algorithm also divides the dataset into several bins so that it can readily identify the variations in the dataset, ensuring that accuracy in calculation is achieved. The gradient boost algorithm, wherein new models can be developed to compensate for the residual errors created by the previous model, is also a beneficial process for estimating cooling loads since, in architecture, slight fluctuations

act as a nonlinear trigger for energy loads.

On the other hand, CatBoost-HGS performed exceptionally well in predicting the value for UTCI, particularly for Type B and Type C buildings. This can be related to the unique strength that CatBoost has in handling categorical variables, regardless of the numerical nature of the inputs used for analysis. The strength that CatBoost has comes from its ability to handle categorical variables by converting them to a form that preserves numerical values, without causing any form of overfitting or bias. Although the dataset used for analysis contains numerical values, it

can be theorized that the analysis capabilities that CatBoost has in handling interactions, whether it is building type, environmental, or time, could prove to be its strength when used for predicting the values for UTCI.

The discrepancies noticed in the performance for various models may be related to the advantages that each model holds for the type and nature of the concerned data. Even if HGBR-HGS has an upper hand when it comes to continuous numerical regression problems, like the prediction of cool load, it has a restrictive outlook when it comes to predicting values related to environmental parameters, such as the prediction of UTCI, for which CatBoost-HGS has proven to perform better. This analysis will also form a conclusion that both models can prove to be effective when used judiciously, for predicting energy consumption patterns using HGBR-HGS and for predicting environmental factors using CatBoost-HGS. Such conclusions tend to establish the importance of making use of machine learning algorithms judiciously, depending upon the type of problem, along with the nature and characteristics of the dataset chosen for analysis. Regarding energy forecasting for buildings in urban regions, the level of accuracy for cool load forecasting using HGBR-HGS is found to be significantly higher, whereas the use of CatBoost-HGS will prove beneficial for environmental factors like UTCI.

It is important to note that for the purpose of relevance, the dataset used in evaluating the study has been obtained from a single source, and that could raise implications for the generalization of the final results for other buildings constructed in urban regions. The obtained results may not fully manifest diversity in building designs, geographical location, and local settings. Future studies can apply a different set of datasets for evaluation to assess the applicability of proposed models for different regions, climates, and urban settings.

5. Conclusion

Urban energy consumption forecast, particularly in relation to Universal Thermal Climate Index (UTCI) and the environmental factors that arise from architectural designs, still stands as a major challenge in urban planning. To explore in depth the use of machine learning algorithms that could appropriately identify and solve this challenge, the study placed emphasis mainly on three comprehensive algorithms, namely HGBR, CatBoost, and LightGBM. The algorithmic models were carefully examined for three different types (A, B, and C) of building designs using the HGS optimization approach. The results were clear, pointing to the effectiveness and accuracy of the HGBR-HGS model in predicting cooling loads and UTCI for Type A build-

ings, the superior performance of CatBoost-HGS in Type B for predicting UTCI, although still better than HGBR-HGS in cooling load forecasts, and Type C showing a similar pattern, where CatBoost-HGS again outperformed HGBR-HGS in predicting Type C's UTCI, while HGBR-HGS remained superior for predicting Type C's cooling loads.

This study, therefore, expands Ibrahim et al.'s study [3] in using machine learning algorithms in predicting energy consumption parameters, particularly for residential buildings' Universal Thermal Climate Index (UTCI) and cooling load in urban architectural design. While Ibrahim et al.'s contribution aimed at designing an optimized urban architecture for energy efficiency, the significance and value derived from this study arise from using advanced algorithms such as CatBoost, LightGBM, and HGBR, proving that accurate energy forecasting can be achieved for improved future urban planning and energy use.

Taking all factors into consideration, the study clearly proves how well hybrid ML algorithms can forecast energy consumption for construction architecture designs in accurate and dependable ways. The implications and conclusions drawn from this study can contribute to cleaner and greener cities by paving the way for more environmentally responsive and sustainably designed architecture. Future research can build upon these findings by exploring emerging algorithms and integrating real-time data processing to further enhance energy forecasting capabilities.

Acknowledgements

We would like to express our gratitude to the authors of the reference [5] Ibrahim et al. (2001) for providing the dataset used in this study. Their contribution has been invaluable in advancing our research in urban energy prediction.

Nomenclature

Acronyms and Abbreviations

R^2	Coefficient of Determination
CatBoost	Categorical Gradient Boosting
COST	Cooperation in Science and Technical Development
GBDT	Gradient Boosting Decision Tree
HGBR	Histogram Gradient Boosting Regressor
HGS	Hunger Games Search
ISB	International Society of Biometeorology
LightGBM	Light Gradient Boosting Machine
MAE	Mean Absolute Error
RAE	Relative Absolute Error
RMSE	Root Mean Square Error
SH	The summation of hunger feelings across all solutions
XGBoost	eXtreme Gradient Boosting

Mathematical Symbols

γ	The learning rate
Γ_k	The scoring function
Fit_b	The best fitness value
Fit_i	The fitness value
Fit_w	The worst fitness value

E	Signifies the control of variation
H_i	The hunger level of each solution
H_n	New hunger level
J	The number of leaves
$L(y, f(x))$	Loss function
M	The quantity of trees
q	The decision rules of the tree
r_1, r_2	Random numbers
S, G	The negative gradients of the loss function
W_1, W_2	The hunger weights
w_j^*	Optimal leaf weight
$w_{q(x)}$	The regression trees
X_b	The best solution
y_k	Represents the corresponding label value

References

- [1] C. Li, J. Yao, Y. Li, Y. Wei, M. Li, Y. Shen, G. Shao, H. Feng, and H. Zheng, (2025) "The application and development of Pennes' biological heat transfer model in human comfort and thermal therapy" **International Communications in Heat and Mass Transfer** 169: 109724. DOI: <https://doi.org/10.1016/j.icheatmasstransfer.2025.109724>.
- [2] M. Motamedi, S. Iranfar, F. W. K. Khudhur, and A. Hassanpouryouzband, (2026) "When heat meets water: Environmental trade-offs of underground thermal energy storage" **Journal of Energy Storage** 149: 120265. DOI: <https://doi.org/10.1016/j.est.2025.120265>.
- [3] Z. Xie, Y. Wang, G. Tan, J. Wang, J. Zhao, Y. Zhu, and B. Cao, (2026) "Physical characteristics, human perceptions, and energy-saving performance of dynamic airflow in built environments-A comprehensive review" **Renewable and Sustainable Energy Reviews** 235: 116936. DOI: <https://doi.org/10.1016/j.rser.2026.116936>.
- [4] K. Błażejczyk and A. Kunert, (2011) "Bioklimatyczne uwarunkowania rekreacji i turystyki w Polsce= Bioclimatic principles of recreation and tourism in Poland" **Monografie-Polska Akademia Nauk. Instytut Geografii i Przestrzennego Zagospodarowania im. Stanisława Leszczyckiego**.
- [5] Y. Ibrahim, T. Kershaw, P. Shepherd, and H. Elkady, (2022) "Multi-objective optimisation of urban courtyard blocks in hot arid zones" **Solar Energy** 240: 104–120. DOI: <https://doi.org/10.1016/j.solener.2022.05.024>.
- [6] T. Edition, (2001) "Glossary of terms for thermal physiology" **The Japanese Journal of Physiology** 51: 245.
- [7] G. Havenith, (2001) "Individualized model of human thermoregulation for the simulation of heat stress response" **Journal of applied physiology** 90: 1943–1954. DOI: <https://doi.org/10.1152/jappl.2001.90.5.1943>.
- [8] K. Parsons. *Human Thermal Environments: The Effects of Hot, Moderate, and Cold Environments on Human Health, Comfort, and Performance*. 3rd. Boca Raton: CRC Press, 2014. DOI: [10.1201/b16750](https://doi.org/10.1201/b16750).
- [9] G. Jendritzky, G. Havenith, P. Weihs, E. Batchvarova, and R. DeDear, (2007) "The universal thermal climate index UTCI goal and state of COST action 730" **Environmental ergonomics XII. Biomed, Ljubljana**: 509–512.
- [10] Z. Liu, Q. Xu, and H. Jin, (2025) "The influence of urban residential block morphology on seasonal outdoor thermal comfort in severely cold regions: A case study of Harbin, China" **Frontiers of Architectural Research**: DOI: <https://doi.org/10.1016/j.foar.2025.09.009>.
- [11] J. K. Vanos, J. S. Warland, T. J. Gillespie, and N. A. Kenny, (2010) "Review of the physiology of human thermal comfort while exercising in urban landscapes and implications for bioclimatic design" **International journal of biometeorology** 54: 319–334. DOI: <https://doi.org/10.1007/s00484-010-0301-9>.
- [12] N. A. Kenny, J. S. Warland, R. D. Brown, and T. G. Gillespie, (2009) "Part B: Revisions to the COMFA outdoor thermal comfort model for application to subjects performing physical activity" **International journal of biometeorology** 53: 429–441. DOI: <https://doi.org/10.1007/s00484-009-0227-2>.
- [13] P. Höppe, (1999) "The physiological equivalent temperature—a universal index for the biometeorological assessment of the thermal environment" **International journal of Biometeorology** 43: 71–75. DOI: <https://doi.org/10.1007/s004840050118>.
- [14] P. Höppe, (2002) "Different aspects of assessing indoor and outdoor thermal comfort" **Energy and buildings** 34: 661–665. DOI: [https://doi.org/10.1016/S0378-7788\(02\)00017-8](https://doi.org/10.1016/S0378-7788(02)00017-8).
- [15] D. Fiala, G. Havenith, P. Bröde, B. Kampmann, and G. Jendritzky, (2012) "UTCI-Fiala multi-node model of human heat transfer and temperature regulation" **International journal of biometeorology** 56: 429–441. DOI: <https://doi.org/10.1007/s00484-011-0424-7>.
- [16] A. Psikuta, D. Fiala, G. Laschewski, G. Jendritzky, M. Richards, K. Błażejczyk, I. Mekjavič, H. Rintamäki, R. de Dear, and G. Havenith, (2012) "Validation of the Fiala multi-node thermophysiological model for UTCI application" **International journal of biometeorology** 56: 443–460. DOI: <https://doi.org/10.1007/s00484-011-0450-5>.

- [17] G. Havenith, D. Fiala, K. Błażejczyk, M. Richards, P. Bröde, I. Holmér, H. Rintamaki, Y. Benshabat, and G. Jendritzky, (2012) "The UTCI-clothing model" **International journal of biometeorology** 56: 461–470. DOI: <https://doi.org/10.1007/s00484-011-0451-4>.
- [18] C. Koppe and G. Jendritzky, (2005) "Inclusion of short-term adaptation to thermal stresses in a heat load warning procedure" **Meteorologische Zeitschrift** 14: 271–278. DOI: [10.1127/0941-2948/2005/0030](https://doi.org/10.1127/0941-2948/2005/0030).
- [19] H. Staiger, K. Bucher, and G. Jendritzky, (1997) "Gefühlte temperatur. Die physiologisch gerechte bewertung von wärmebelastung und kältestress beim aufenthalt im freien in der maßzahl grad celsius" **Annalen der Meteorologie** 33: 100–107.
- [20] E. L. Krüger, T. J. V. Silva, S. Q. da Silveira Hirashima, E. G. da Cunha, and L. A. Rosa, (2021) "Calibrating UTCI'S comfort assessment scale for three Brazilian cities with different climatic conditions" **International journal of biometeorology** 65: 1463–1472. DOI: <https://doi.org/10.1007/s00484-020-01897-x>.
- [21] A. Xu, R. Zhang, J. Yu, and Y. Dong, (2023) "Energy saving optimization of commercial complex atrium roof with resilient ventilation using machine learning" **Smart Cities** 6: 2367–2396. DOI: <https://doi.org/10.3390/smartcities6050108>.
- [22] G. Zhong, (2022) "Convolutional neural network model to predict outdoor comfort UTCI microclimate map" **Atmosphere** 13: 1860. DOI: <https://doi.org/10.3390/atmos13111860>.
- [23] X. Zhang, L. Yang, R. Luo, H.-Y. Wu, J. Xu, C. Huang, Y. Ruan, X. Zheng, and J. Yao, (2022) "Estimating the outdoor environment of workers' villages in East China using machine learning" **Building and Environment** 226: 109738. DOI: <https://doi.org/10.1016/j.buildenv.2022.109738>.
- [24] Z. Fang, X. Feng, J. Liu, Z. Lin, C. M. Mak, J. Niu, K.-T. Tse, and X. Xu, (2019) "Investigation into the differences among several outdoor thermal comfort indices against field survey in subtropics" **Sustainable cities and society** 44: 676–690. DOI: <https://doi.org/10.1016/j.scs.2018.10.022>.
- [25] S. Zare, N. Hasheminejad, H. E. Shirvan, R. Hemmatjo, K. Sarebanzadeh, and S. Ahmadi, (2018) "Comparing Universal Thermal Climate Index (UTCI) with selected thermal indices/environmental parameters during 12 months of the year" **Weather and climate extremes** 19: 49–57. DOI: <https://doi.org/10.1016/j.wace.2018.01.004>.
- [26] M. Borowski and K. Zwolińska, (2020) "Prediction of cooling energy consumption in hotel building using machine learning techniques" **Energies** 13: 6226. DOI: <https://doi.org/10.3390/en13236226>.
- [27] C. Fan, F. Xiao, and Y. Zhao, (2017) "A short-term building cooling load prediction method using deep learning algorithms" **Applied energy** 195: 222–233. DOI: <https://doi.org/10.1016/j.apenergy.2017.03.064>.
- [28] W. Li, G. Gong, H. Fan, P. Peng, and L. Chun, (2020) "Meta-learning strategy based on user preferences and a machine recommendation system for real-time cooling load and COP forecasting" **Applied Energy** 270: 115144. DOI: <https://doi.org/10.1016/j.apenergy.2020.115144>.
- [29] Y. Feng, Q. Duan, X. Chen, S. S. Yakkali, and J. Wang, (2021) "Space cooling energy usage prediction based on utility data for residential buildings using machine learning methods" **Applied energy** 291: 116814. DOI: <https://doi.org/10.1016/j.apenergy.2021.116814>.
- [30] G. Ke, Q. Meng, T. Finley, T. Wang, W. Chen, W. Ma, Q. Ye, and T.-Y. Liu, (2017) "Lightgbm: A highly efficient gradient boosting decision tree" **Advances in neural information processing systems** 30:
- [31] A. V. Dorogush, V. Ershov, and A. Gulin, (2018) "CatBoost: gradient boosting with categorical features support" **arXiv preprint arXiv:1810.11363**. DOI: <https://doi.org/10.48550/arXiv.1810.11363>.
- [32] M. Entezami and S. Goodarzi, (2025) "Cooling Load Prediction in Residential Buildings: Machine Learning Approaches and Comparative Analysis" **Advances in Engineering and Intelligence Systems** 4: 86–100. DOI: <https://doi.org/10.22034/aeis.2025.537166.1344>.
- [33] L. Prokhorenkova, G. Gusev, A. Vorobev, A. V. Dorogush, and A. Gulin, (2018) "CatBoost: unbiased boosting with categorical features" **Advances in neural information processing systems** 31:
- [34] R. S. Benemaran, E. Khajavi, and A. R. T. Khanghah, (2026) "Construction of multi-parameter intelligent comprehensive estimation algorithms for soil compression modulus" **Scientific Reports**. DOI: <https://doi.org/10.1038/s41598-026-43812-1>.
- [35] B. Naeim, A. J. Khiavi, E. Khajavi, A. R. T. Khanghah, A. Asgari, R. Taghipour, and M. Bagheri, (2025) "Machine learning approaches for fatigue life prediction of steel and feature importance analyses" **Infrastructures** 10: 295. DOI: <https://doi.org/10.3390/infrastructures10110295>.

- [36] H. Hang, T. Huang, Y. Cai, H. Yang, and Z. Lin, (2021) "Gradient boosted binary histogram ensemble for large-scale regression" **arXiv preprint arXiv:2106.01986**: DOI: <https://doi.org/10.48550/arXiv.2106.01986>.
- [37] Y. Yang, H. Chen, A. A. Heidari, and A. H. Gandomi, (2021) "Hunger games search: Visions, conception, implementation, deep analysis, perspectives, and towards performance shifts" **Expert Systems with Applications** **177**: 114864. DOI: <https://doi.org/10.1016/j.eswa.2021.114864>.
- [38] E. Khajavi, A. R. T. Khanghah, and A. J. Khiavi. "An efficient prediction of punching shear strength in reinforced concrete slabs through boosting methods and metaheuristic algorithms". In: *Structures*. **74**. Elsevier, 2025, 108519. DOI: <https://doi.org/10.1016/j.istruc.2025.108519>.