Taylor & Francis
Taylor & Francis Group

Advances in Building Energy Research

ISSN: 1751-2549 (Print) 1756-2201 (Online) Journal homepage: www.tandfonline.com/journals/taer20

Comparative assessment of machine learning approaches to predict building annual cooling load intensity in urban environments

Alireza Attarhay Tehrani, Saeideh Sobhaninia, Aurora Bertini, Shady Attia, Zeynep Duygu Tekler, Shiva Manshour & Deepak Amaripadath

To cite this article: Alireza Attarhay Tehrani, Saeideh Sobhaninia, Aurora Bertini, Shady Attia, Zeynep Duygu Tekler, Shiva Manshour & Deepak Amaripadath (06 Jul 2026): Comparative assessment of machine learning approaches to predict building annual cooling load intensity in urban environments, Advances in Building Energy Research, DOI: 10.1080/17512549.2026.2686610

To link to this article: https://doi.org/10.1080/17512549.2026.2686610

Published online: 06 Jul 2026.

ADVANCES IN BUILDING ENERGY RESEARCH
https://doi.org/10.1080/17512549.2026.2686610

Taylor & Francis
by informa

Check for updates

Comparative assessment of machine learning approaches to predict building annual cooling load intensity in urban environments

Alireza Attarhay Tehrani [a], Saeideh Sobhaninia [b], Aurora Bertini [c,d], Shady Attia [c], Zeynep Duygu Tekler [e], Shiva Manshour [f] and Deepak Amaripadath [g,h]

[a]Department of Architecture, Faculty of Art and Architecture, Islamic Azad University, Tehran, Iran; [b]Department of Urban Design and Planning, University of Washington, Seattle, USA; [c]Sustainable Building Design Lab, Department UEE, Universite of Liege, Liège, Belgium; [d]Louvain Research Institute for Landscape, Architecture, Built Environment (LAB), Faculty of Architecture, Architectural Engineering, and Urban Planning (LOCI), UC Louvain, Liège, Belgium; [e]Department of Engineering Science, University of Oxford, Oxford, UK; [f]University of Nevada, Las Vegas, NV, USA; [g]School of Geographical Sciences and Urban Planning, Arizona State University, Tempe, AZ, USA; [h]Urban Climate Research Center, Arizona State University, Tempe, AZ, USA

ABSTRACT

Reducing energy consumption has become a critical priority in sustainable urban development, with accurate prediction of building cooling loads essential for efficient energy management systems. However, challenges such as inconsistent data and model generalizability restrict the seamless integration of energy modelling into urban planning. As such, this study introduces a novel hybrid modelling framework to forecast annual cooling load intensity in Phoenix, Arizona. The proposed hybrid approach integrates Extreme Gradient Boosting (XGBoost), Light Gradient Boosting Machine (LightGBM), and Categorical Boosting (CatBoost) in a stacking ensemble, and uses a GA to optimize hyperparameters and the meta-learner. The results demonstrate that this optimized framework achieved a better accuracy with a root mean square error (RMSE) of 1.09 kWh/m^2, a coefficient of determination (R^2) of 0.99, a mean absolute error (MAE) of 0.26 kWh/m^2, and a mean square error (MSE) of 1.19 kWh/m^2. Additionally, SHAP interpretability analysis indicates the window-to-wall ratio as the most influential driver of cooling load in Phoenix, Arizona. The model's robustness was further validated on 100 unseen buildings in Scottsdale, Arizona, demonstrating strong generalizability. These findings highlight the potential of the proposed framework to enhance building energy forecasting and to provide actionable insights that support urban energy management.

ARTICLE HISTORY
Received 3 December 2025
Accepted 1 June 2026

KEYWORDS
Energy consumption, energy management, machine learning, optimization, cooling loads

Nomenclature

Abbreviation	Description
AdaBoost	Adaptive Boosting

CONTACT Alireza Attarhay Tehrani allentehrani1995@gmail.com Department of Architecture, Faculty of Art and Architecture, Islamic Azad University, Tehran, Iran

AI	Artificial Intelligence
ANN	Artificial Neural Network
CatBoost	Categorical Boosting
DL	Deep Learning
EPW	Energy Plus Weather
GA	Genetic Algorithm
HB	Honeybee
KNN	K-Nearest Neighbours
LightGBM	Light Gradient Boosting Machine
MAE	Mean Absolute Error
MAPE	Mean Absolute Percentage Error
ML	Machine Learning
MLR	Multiple Linear Regression
MSE	Mean Squared Error
PReg	Polynomial Regression
QR	Quantile Regression
R^2	R-Squared
RMSE	Root Mean Square Error
RR	Ridge Regression
SHAP	SHapley Additive exPlanations
SVR	Support Vector Regression
XGBoost	Extreme Gradient Boosting

1. Introduction

The rapid growth of urban populations, coupled with the ongoing effects of climate change, is significantly transforming cities worldwide. According to the World Bank projections, nearly 70% of the global population will live in urban areas by 2050 (Hoornweg et al., 2010). While urbanization presents opportunities for economic growth and technological advancements, it also brings immense challenges, particularly related to the pressure it places on urban infrastructure, resources, and public services (Ferrer et al., 2018; Salimi & Al-Ghamdi, 2020; Shahidehpour et al., 2018; Veisi et al., 2025). The rising building cooling energy demand is a growing energy challenge, intensified by climate shifts and more frequent extreme weather events, such as heat waves (Babu et al., 2024; Stasi et al., 2024). As global temperatures rise, cities, especially in hot regions, are becoming increasingly susceptible to extreme heat, leading to greater reliance on cooling and ventilation systems (Peng et al., 2022) and more frequent power outages. This not only drives up energy consumption but also increases risks associated with occupants' personal thermal comfort, health risks, and reduced productivity (Amaripadath et al., 2026; Tehrani et al., 2025; Tekler et al., 2024). Addressing energy inefficiencies in buildings has thus become a critical priority for urban sustainability, as buildings are significant energy consumers and major contributors to greenhouse gas emissions. Improving energy efficiency, particularly in terms of cooling, is vital to managing the growing urban energy demand (Amaripadath et al., 2024; Xu et al., 2024). Forecasting building energy use supports these goals by enabling optimized operation, better resource planning, and demand-reduction strategies. Accurate forecasts can inform building design, operation, and retrofitting, thereby reducing energy consumption and promoting sustainable urban growth (Ekici & Aksoy, 2009; Sadaghat et al., 2024). Beyond energy conservation, forecasting also aids policymakers in their efforts to reduce emissions and ensure energy security in rapidly expanding cities (Ben-Nakhi & Mahmoud, 2004).

The main methods for predicting building energy performance can be categorized into three types: the physical or 'white box' method, the statistical or 'black box' method, and the hybrid or 'grey box' method (Foucquier et al., 2013; Wei et al., 2018). White-box models employ detailed heat and mass transfer equations to predict building loads accurately. These models offer a comprehensive understanding of the underlying physics governing energy transfer within buildings. Commercially available software tools such as EnergyPlus and DesignBuilder provide sophisticated platforms for configuring and simulating white-box models (Li & Wen, 2014). Grey-box models simplify building thermal dynamics by representing them with reduced-order resistance and capacity (RC) models. Parameter values (Rs and Cs) in these models are often inferred from measured data through parameter identification techniques, effectively minimizing prediction errors (Braun & Chaturvedi, 2002; Das et al., 2024). By integrating both physical principles and empirical data, grey-box models strike a balance between accuracy and computational efficiency (Ahmad et al., 2017).

Black-box models, on the other hand, are exclusively data-driven, relying solely on historical data to predict building thermal load. With advancements in data collection technologies, these models can leverage large datasets to identify complex patterns in building energy consumption. Early black-box models often utilized statistical methods such as linear regression (Forrester & Wepfer, 1984) and autoregressive integrated moving-average (ARIMA) to forecast building load (Spethmann, 1989). However, modern black-box models can employ more advanced ML algorithms (Li et al., 2016), such as neural networks (Tehrani et al., 2024a) and time-series models (Tehrani et al., 2024b), to improve prediction accuracy and adaptability to various building types and conditions (Le et al., 2019). The significant advancements in deep learning have established deep neural networks as powerful tools for predictive modelling across diverse domains. While classical ML models often rely on manually extracted features and deep learning with hierarchical feature extraction, ensemble-based models possess a more complex architecture capable of comprehending patterns without manual feature engineering. This flexibility allows ensemble-based models to excel in various domains like computer vision and sequential data analysis (Tan et al., 2022), enabling end-to-end learning (Low et al., 2021) and achieving remarkable performance outcomes (Ali et al., 2024).

The data-driven approach to predicting building energy consumption has garnered significant research interest and has been extensively applied (Amasyali & El-Gohary, 2018; Ilojianya et al., 2024; Kapp et al., 2023; Wei et al., 2018). Compared with physics-based approaches, data-driven models offer several advantages, including reduced computational resource burden and faster speed, while maintaining acceptable prediction accuracy (Ahmad et al., 2018; Wang et al., 2018). In recent years, there has been a proliferation of studies dedicated to modelling building energy demand (Zhou et al., 2024). These studies have utilized a variety of methods, including traditional regression techniques (Ghiaus, 2006), artificial neural networks (Zhang & Haghighat, 2010; Zhou & Haghighat, 2009), and building simulation methods (Al-Ajmi & Hanby, 2008; Zhou et al., 2008), among others. Catalina et al. (2013) used multiple regression analysis to predict the building heating energy demand. Ahmad et al.

(2017) compared the performance of feed-forward back-propagation ANN with random forest (RF) for forecasting hourly HVAC energy consumption. Kwok et al. (2011) utilized a multi-layer perceptron (MLP) model to estimate the cooling load intensity of a building. Ngo (2019) developed an ensemble ML model to predict cooling load intensity in early design, utilizing a dataset of 243 buildings. Pino-Mejías et al. (2017) utilized linear regression and neural network models to predict cooling and heating energy demands, along with overall energy consumption. Deb et al. (2016) investigated the effectiveness of neural network-based solutions in forecasting diurnal cooling energy loads. Li et al. (2009) conducted a comparative analysis of various ML techniques, including radial basis function neural network (RBFNN), general regression neural network (GRNN), traditional backpropagation neural network (BPNN), and support vector machine (SVM), for predicting the hourly cooling load intensity of a typical residential building. Olu-Ajayi et al. (2022) developed a ML model that forecasts annual average energy use based on building design features in the initial development stages. Tehrani et al. (2025) conducted a comparative analysis of various ML techniques, including SVM, GBR, RF, XGBoost, and deep neural network (DNN) to predict cooling load intensity based on building characteristics.

To further enhance prediction accuracy, existing literature have proposed hybrid prediction models that combine machine learning, deep learning, and optimization techniques. Moazzami et al. (2013) proposed a hybrid framework for day-ahead peak load forecasting by combining two different ANNs with genetic optimization. Xiao et al. (2016) developed a hybrid model combining a general regression neural network (GRNN) with a multi-objective firefly algorithm (MOFA) for power load forecasting, using MOFA to optimize the initial weights and thresholds of the GRNN model. Almaleck et al. (2024) developed an ensemble model combining an ANN with an ARIMAX model to predict electrical consumption in sports venues. Chen et al. (2025) designed a hybrid DL model combining MKDCN and LSTM architectures to achieve more accurate building electrical load forecasting. He et al. (2019) developed a hybrid short-term load forecasting method combining variational mode decomposition (VMD) with LSTM networks, and employed the bayesian optimization algorithm (BOA) to optimize the model's hyperparameters. Zhao et al. (2025) combined opposition-learning golden-sine grey wolf optimization (OGGWO) with a GRU network to obtain accurate short-term heating load predictions, using the OGGWO algorithm to adjust hyperparameters and thereby enhance the model's prediction accuracy. Table 1 summarizes the types of ML models used in other studies.

Despite broad research on data-driven building energy prediction, several gaps remain. As demonstrated in Table 1, previous studies employing data-driven methodologies have primarily concentrated on forecasting the energy consumption of single buildings and have used only a limited number of parameters to predict potential energy consumption. Furthermore, most existing studies focus on single-model or hybrid approaches, with limited attention given to combining multiple algorithms alongside optimization techniques for improved generalization and robustness. To address these limitations, this study aims to develop a novel hybrid framework that integrates parametric simulations, ensemble-based ML approaches, an optimization technique, and SHapley Additive exPlanations (SHAP) analysis to predict annual cooling load intensity across urban residential buildings and to interpret the

Table 1. ML model analyzed in existing load prediction studies.

Authors	Type	ML Method	Dependent Variables	Study scale	SHAP
Álvarez-Sanz et al. (2024)	Single model approach	XGBOOST, ET, RF	Heating load	Single building	No
Ahmad et al. (2017)	Single model approach	ANN, RF	Cooling load and heating loads	Single building	No
Amber et al. (2017)	Single model approach	MLR	Overall energy	Single building	N0
Amber et al. (2018)	Single model approach	MLR, GP, SVM, ANN, DNN	Overall energy	Single building	N0
Fan et al. (2014)	Single model approach	AR, DT, MLR, MARS, KNN, SVM, ANN, RF, BDT	Overall energy	Single building	N0
Chou and Bui (2014)	Single model approach	GLR, DT, ANN, BPNN, SVM	Cooling load and heating load	Single building	No
Tehrani et al. (2025)	Single model approach	XGBOOST, SVM, GPR, RF, DNN	Cooling load	Urban scale	Yes
Lachut et al. (2014)	Single model approach	AR, NB, SVM	Overall energy	Single building	No
Kondath et al. (2024)	Single model approach	DNN, CNN, RNN, LSTM, CNN-LSTM	Cooling load	Single building	No
Hua et al. (2024)	Single model approach	MLR, ANN	Heating load	Urban scale	No
Fu et al. (2015)	Single model approach	AR, DT, SVM, ANN	Overall energy	Single building	No
Mehdizadeh Khorrami et al. (2024)	Single model approach	DT, LR, ANN, DT, LR, ANN	Cooling load and heating load	Single building	No
Seo et al. (2023)	Single model approach	ANN, LSTM	Cooling load	Single building	No
Xu et al. (2022)	Single model approach	MLP	Cooling load and heating load	Single building	No
Amber et al. (2018)	Single model approach	MLR, GP, SVM, ANN, DNN	Overall energy	Single building	No
Moazzami et al. (2013)	Hybrid model approach	ANN-GA	Overall energy	Single building	No
Alobaidi et al. (2018)	Hybrid model approach	ANN, MLP-FFNN, ANN-based Boosting	Overall energy	Single building	No
Xiao et al. (2016)	Hybrid model approach	GRNN-MOA	Overall energy	Single building	No
Almaleck et al. (2024)	Hybrid model approach	ANN-AR	Overall energy	Single building	No
Chen et al. (2025)	Hybrid model approach	MKDCN- LSTM	Overall energy	Single building	No
Zhao et al. (2025)	Hybrid model approach	OGGWO- GRU	Heating load	Single building	No
He et al. (2019)	Hybrid model approach	VMD-LSTM-BOA	Overall energy	Single building	No

importance of input features, thereby providing practical support during the early architectural design phase for optimizing building energy performance. Parametric simulation techniques can create synthetic data encompassing a wide range of relevant urban-scale scenarios for stakeholders. This study implements an ensemble-based ML algorithm that uses a genetic algorithm (GA) optimization technique to adjust hyperparameters and thereby enhance the model's prediction accuracy for the annual cooling load intensity at the urban scale. Furthermore, this research identifies the main building characteristics and performs feature interpretability analysis using SHAP, demonstrating enhanced forecast precision and delivering interpretable insights into feature contributions. The goal is to present stakeholders in the building sector with a reliable and insightful tool that encourages greater engagement in performance-driven decision-making from the earliest stages of design, thereby supporting more effective energy policies and practical interventions.

2. Methodology

Urban-scale energy forecasting requires integrating diverse datasets that capture both building characteristics and climatic conditions (Tehrani et al., 2025). This study develops a comprehensive dataset of annual cooling load intensity for residential buildings in Phoenix, Arizona, by combining parametric energy simulations with optimized ML methods. The methodology consists of four key stages: data acquisition and integration, parametric simulation, ML model development, and performance evaluation. The study conceptual framework is shown in Figure 1.

2.1. Data acquisition and integration

The data acquisition process for data-driven approaches is a critical step that integrates a diverse array of inputs, including urban morphology, building characteristics, and microclimate data (Reinhart & Davila, 2016). These inputs are essential for performing

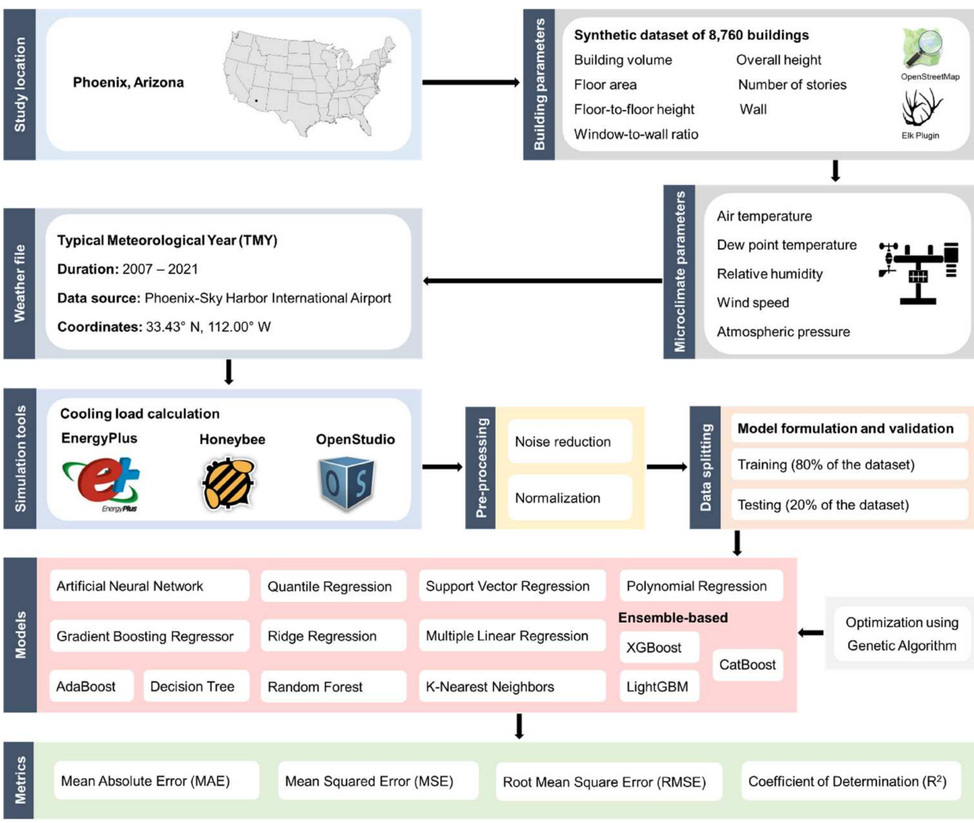


Figure 1. Study conceptual framework for efficient energy management in smart cities through urban cooling load prediction.

comprehensive simulations that account for the geometric parameters of urban areas and buildings. The main factors include building volume, building height, floor-to-floor height, window-to-wall ratio, building shapes, number of floors, building types, and geographical coordinates (Johari et al., 2020; Tehrani et al., 2024a). This research employs a parametric approach and utilizes open-access repositories, such as OpenStreetMap (OSM) (Haklay & Weber, 2008), to accurately gather and model urban environments (Tehrani et al., 2024B). The reference weather data used in this study covers the typical meteorological year 3 (TMY3) from 2007 to 2021. Specifically, the data were sourced from Phoenix–Sky Harbor International Airport, located at a latitude of 33.43 and a longitude of −112.00, with an elevation of 337.40 m. The average dry bulb temperature, relative humidity, and wind speed recorded during this period were 24 °C, 30%, and 2.8 m/s, respectively.

2.2. Parametric energy simulation

Urban buildings share similar physical and visual characteristics and can be categorized based on their function (Ali et al., 2024). In this study, accurately modelling the city of Phoenix, the selected case study, was crucial since different building types and shapes significantly impact precise energy estimations. This research focuses on residential buildings and integrates OSM data with the Elk plugin to develop urban layouts. OSM provides detailed data on building physics and urban morphology, while the Elk plugin extracts maps and topographic models from OSM to generate a precise city model (Kaitouni et al., 2025). This study developed a detailed parametric model of Phoenix, combining urban layouts and building characteristics to support energy simulations. For this study, Honeybee, a plugin within the Ladybug Tools suite, which serves as an interface for reputable energy simulation engines such as EnergyPlus and OpenStudio (Tehrani et al., 2025) was utilized. Parametric simulation of building cooling loads using EnergyPlus and OpenStudio provided a robust method for evaluating building energy performance. EnergyPlus enabled the simulation of cooling load parameters, offering insights into their impact on energy consumption, thermal comfort, and other performance metrics. Additionally, it facilitated in-depth analysis and optimization of building designs.

This research focused on a specific region concerning cooling energy consumption. Phoenix, classified as climate zone 1B (hot and dry) (Tehrani et al., 2025), was selected due to its unique climatic conditions, where high annual cooling load intensities present a significant challenge for sustainable urban development. Table 2 presents Phoenix's climatic characteristics. For the energy simulation, the developed parametric model

Table 2. Phoenix climate parameters influencing research objectives. Data were sourced from the weather file provided by the EnergyPlus database for Phoenix 2007–2021.

Weather data	unit	Average	Max	Min
Dry-bulb temperature	C	24.6	45.6	0.4
Relative humidity	%	29.9	100	10
Dew point temperature	C	3.4	32	−19.3
Wind speed	m/s	2.8	18.5	0
Direct normal radiation	Wh/m^2	323.13	1007	0
Diffuse horizontal radiation	Wh/m^2	47.9	407	0
Global horizontal radiation	Wh/m^2	245.7	1063	0
Total sky cover	tenth	4.3	10	0
Barometric pressure	Pa	97293.9	98911	96242

of Phoenix was imported into the Honeybee component, where the cooling and heating setpoints were set to 24 and 21 °C, respectively, for all studied systems. Air ventilation was set to zero for all considered systems. The main assumptions related to the design parameters included a number of occupants of 0.02 people/m^2, heat gain from occupants of 120 W/person, lighting power density of 11 W/m^2, equipment power density of 4 W/m^2, and an infiltration rate of 0.5 ACH. These assumptions were applied uniformly across the selected HVAC model. The selected HVAC system was a packaged terminal air conditioner (PTAC), which is commonly used in residential applications (Toutou et al., 2018). Using the Honeybee 'Export to OpenStudio' component, the zones were exported to OpenStudio and subsequently analyzed using EnergyPlus. In addition, the

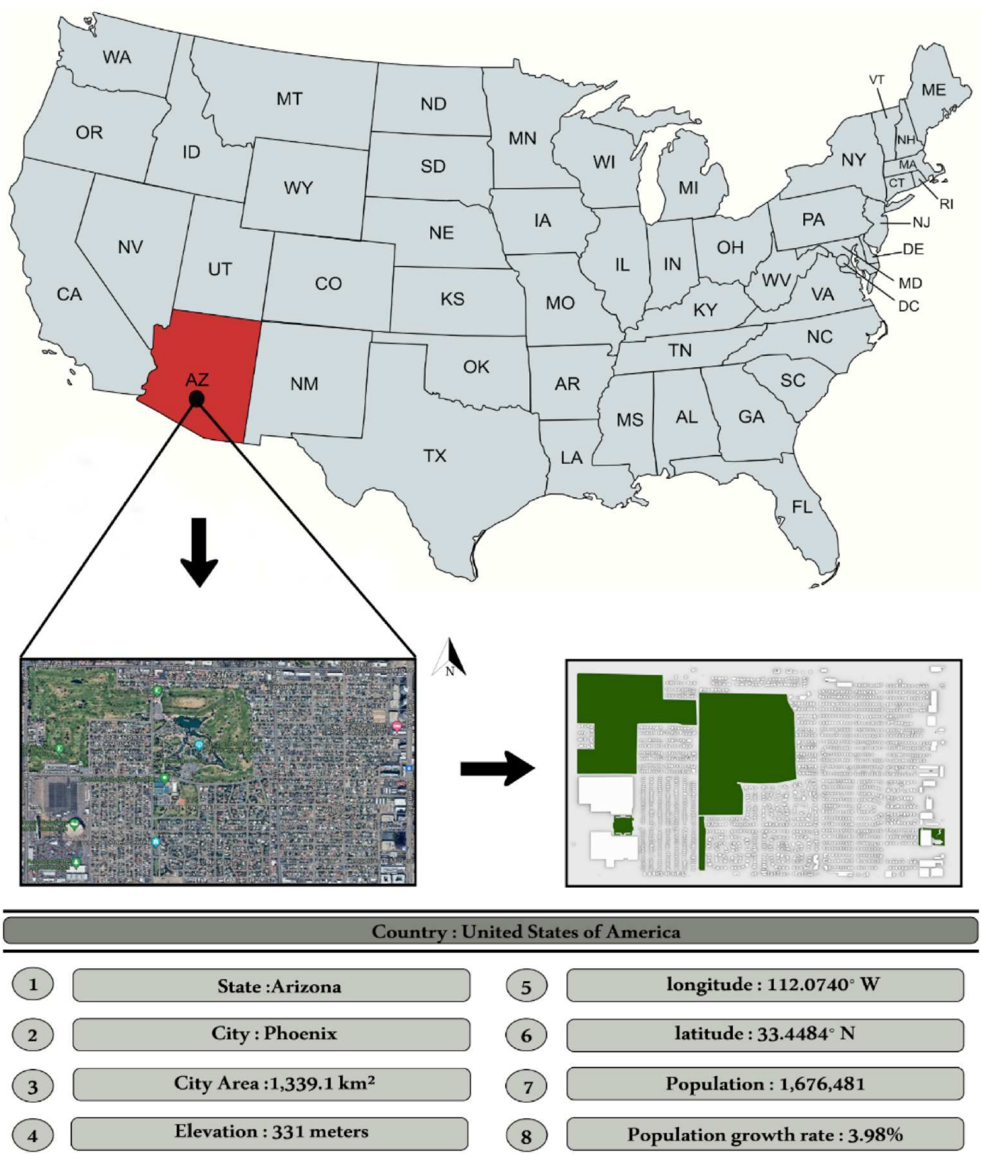


Figure 2. Aerial view of Phoenix city: Location map and 3D model of the study area.

surrounding areas was modelled for each building to understand the effects of shading from neighbouring buildings as shown in Figure 2. This workflow enabled precise simulation of building annual cooling load intensity. After completing the energy simulations, a database was constructed by developing a synthetic dataset consisting of 8,760 residential buildings for the data-driven analysis as shown in Figure 3. In this study, the input parameters derived from the dataset included building characteristics alongside microclimate variables as listed in Table 3. The goal was to establish correlations between these factors and the annual cooling load intensity, which served as the output data. The simulations assumed that the buildings were located in Phoenix, Arizona, as in Figure 2. Building shapes were generated using OpenStreetMap and the ELK plugin in Grasshopper within Rhino software, ensuring accurate representations of the city. Table 4 provides a comprehensive summary of the main descriptive statistics for various building features and environmental parameters derived from the dataset.

2.3. Algorithms

Data preprocessing, a crucial step in data-driven approaches that ensures data quality and consistency is described. The proposed framework and the ML models employed in this study is explained. The proposed approach enhances the accuracy and reliability of annual cooling load inte nsity predictions by leveraging single, ensemble, and hybrid ML techniques in conjunction with GA as shown in Figure 4.

2.3.1. Data pre-processing and splitting

The dataset, consisting of annual cooling load intensity outputs, was randomly split into 80% training and 20% independent hold-out test sets for reproducible evaluation. Preprocessing included removing duplicates, correcting inconsistencies, standardizing formats, and normalizing continuous variables to the 0–1 range using Min–Max scaling. To prevent data leakage, all preprocessing steps were fitted only on the training data: within each cross-validation fold, scalers were trained on the fold's training subset and applied to its validation subset, and for final testing, scalers were fitted on the full 80% training set and applied to the 20% hold-out set. The model development followed an out-of-fold (OOF) stacking framework within a nested cross-validation approach. Specifically, the GA optimization of base-learner hyperparameters was performed within each fold of the ten-fold cross-validation, ensuring that tuning was based solely on the fold's

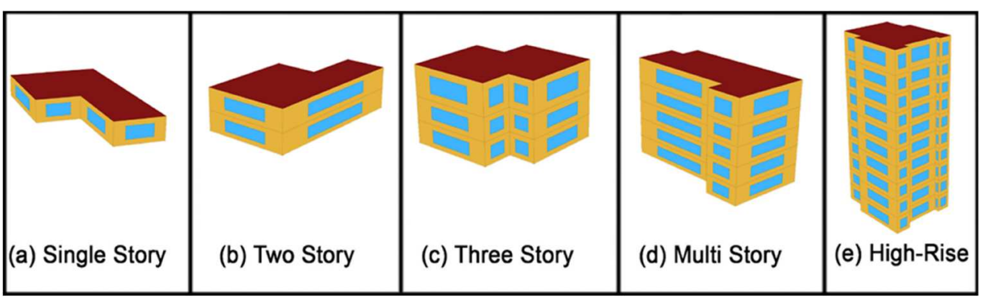


Figure 3. Various building configurations in terms of the number of stories, shape, and orientation included in the simulations.

Table 3. Input parameters used in the study.

Category	Parameters
Building characteristics	Building volume Overall height Floor area Floor-to-floor height Window-to-wall ratio Number of stories Wall
Microclimate variables	Air temperature Dew point temperature Wind speed Relative humidity Atmospheric pressure

training subset and validated on its corresponding validation subset. Ten-fold cross-validation within the training set generated OOF predictions from the base learners (CatBoost and LightGBM), which were used as inputs to the meta-learner (XGBoost). The meta-learner was trained exclusively on these validation-fold predictions. Afterward, base models were retrained on the full training set and applied to the hold-out test set, and the meta-learner combined these predictions to produce the final outputs. All reported performance metrics are based on the unseen 20% hold-out test set. To further validate the model's effectiveness, the study assesses 100 new buildings in Scottsdale, ensuring its applicability across different geographic locations. This step enhances the model's reliability and real-world usability.

2.3.2. Single machine learning model development

In this stage, various single ML models as methodological benchmarks were developed to assess their predictive accuracy, efficiency, and suitability for building energy performance modelling. These models were included to compare their performance with the proposed hybrid framework, as several of these algorithms have demonstrated exceptional results in energy forecasting (Zhang et al., 2021). Single models, such as multiple linear regression (MLR), ridge regression (RR), quantile regression (QR), polynomial regression (PReg), artificial neural networks (ANN), k-nearest neighbours (KNN), support vector regression (SVR), and decision trees (DT), operate independently to establish relationships between input features and the target variable. In addition, ensemble and gradient-boosting models were included because they combine multiple learners to enhance

Table 4. Descriptive statistics of building features and environmental parameters.

Feature	Unit	Mean	Std	Min	25%	50%	75%	Max
Building volume	m^3	1817.24	939.20	176.50	1140.40	1678.40	2360.80	5134.60
Overall height	m	9.55	14.39	4.10	6.30	8.50	10.60	197.00
Wall	count	6.89	2.99	4.00	4.00	6.00	8.00	26.00
Floor-to-floor height	m	3.20	0.35	2.80	2.90	3.20	3.40	4.00
Number of stories	count	2.15	0.82	1.00	2.00	2.00	3.00	4.00
Window-to-wall ratio	%	0.41	0.11	0.20	0.35	0.40	0.50	0.60
Floor area	m^2	215.59	95.95	24.90	153.90	195.50	264.70	632.20
Air temperature	°C	24.67	10.00	0.40	16.70	25.60	32.80	45.60
Dew Point temperature	°C	3.46	7.78	−19.30	−2.00	2.30	8.30	32.00
Relative humidity	%	29.91	18.59	10.00	15.00	25.00	39.00	100.00
Wind speed	m/s	2.82	1.98	0.00	1.50	2.60	3.60	18.50
Atmospheric pressure	Pa	97293.96	417.41	96242.00	96992.00	97246.00	97575.00	98911.00

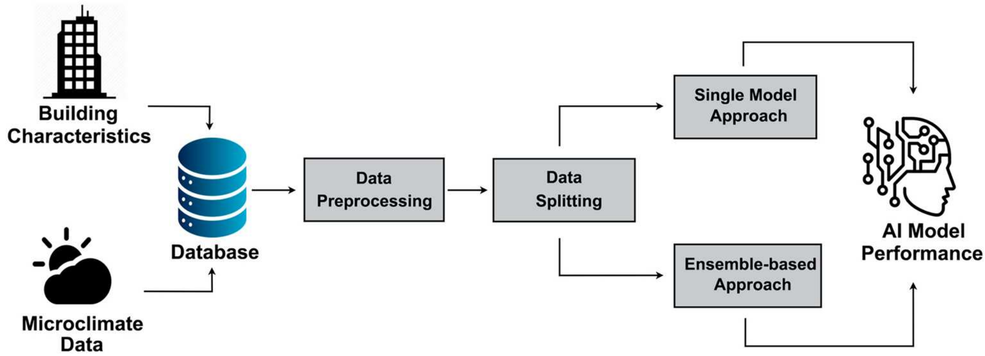


Figure 4. The steps from data collection (building characteristics and microclimate data) to model evaluation. Data is preprocessed and split, followed by the application of both single and ensemble-based machine learning approaches to predict annual cooling load intensity, with AI model performance as the final outcome.

predictive performance and robustness. These models included adaptive boosting (AdaBoost), gradient boosting regressor (GBR), random forest (RF), extreme gradient boosting (XGBoost), categorical boosting (CatBoost), and light gradient boosted machine (LightGBM). To maximize the predictive performance of each model and ensure a fair comparison, the hyperparameters of all single models were optimized using a GA with a consistent configuration, including a population size of 20, a maximum of 50 iterations, and a mutation probability of 0.1.

2.3.3. Proposed model

Hybrid techniques represent a sophisticated data-driven approach designed to create more robust models by combining two or more base learners, thereby enhancing the model accuracy by mitigating overfitting and increasing generalizability (Wang & Srinivasan, 2017). The model integrated three gradient boosting base learners optimized using a GA, and combined their outputs through a stacking ensemble with a separate XGBoost meta-learner. This framework leverages the strengths of each algorithm while addressing their individual limitations through a unified optimization approach. The GA was employed to fine-tune the hyperparameters of each base learner, ensuring that the models achieved optimal performance by balancing bias and variance. Figure 5 illustrates the step-by-step flowchart of the GA, showcasing the sequential processes involved in the optimization procedure. The framework was designed to handle complex, non-linear relationships in the data, making it particularly suitable for predicting annual cooling load intensity, which is influenced by multiple factors such as building characteristics and weather conditions. The model began with the initialization of hyperparameters for each base learner, as outlined in Table 5. XGBoost, CatBoost, and LightGBM were chosen for their complementary strengths: XGBoost excels in handling sparse data and providing robust regularization (Hakkal & Lahcen, 2024), CatBoost is particularly effective with categorical features and offers built-in handling of missing values, and LightGBM is optimized for speed and efficiency with large datasets. The GA was then applied to optimize key hyperparameters such as the number of estimators, learning rate, and maximum depth for each base learner.

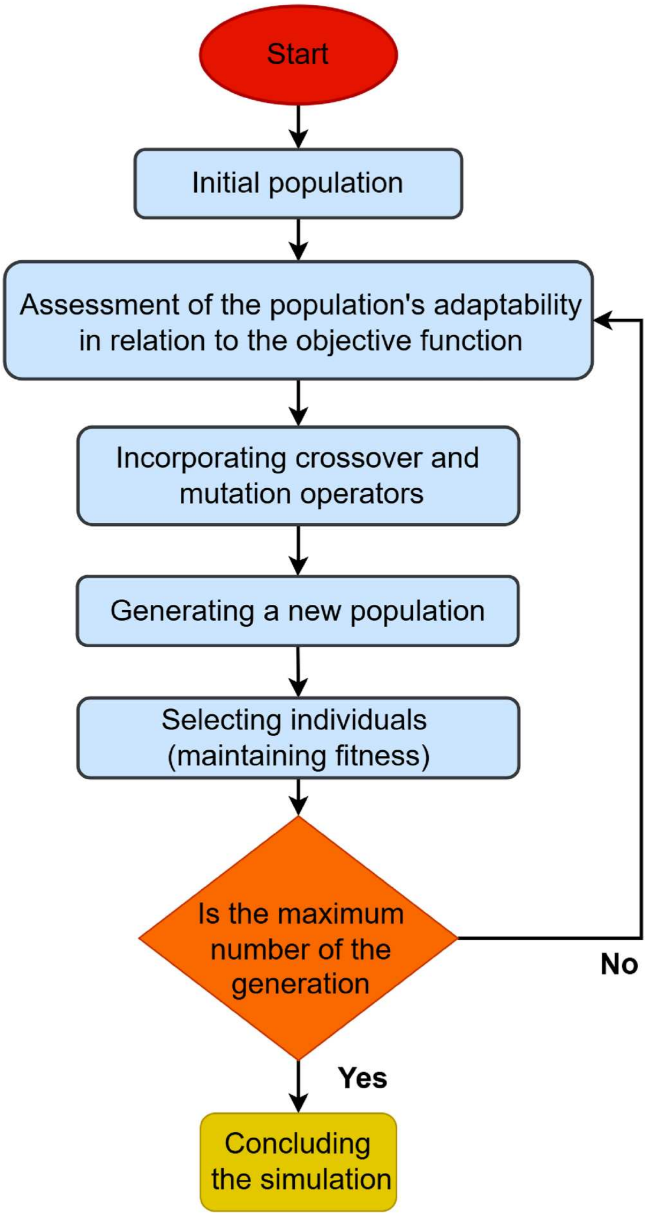


Figure 5. The flowchart of the genetic algorithm optimization process.

The final model integrated predictions from the three base learners using a stacking architecture, as illustrated in Figure 6. Within a cross-validation framework, ten-fold cross-validation within the training set generated out-of-fold predictions from the optimized base learners (XGBoost, CatBoost, and LightGBM). These OOF predictions served as input features to train a fourth, entirely separate model: an XGBoost meta-learner. This meta-learner learned to optimally weight and combine the base-learner outputs, and was trained exclusively on the validation-fold predictions rather than on the original feature space. This ensemble strategy reduced overfitting and improved

Table 5. Hyperparameters for the hybrid model with genetic algorithm optimization.

Hyperparameter	Description	XGBoost	CatBoost	LightGBM
*n*estimators	Number of boosting rounds.	50–500	50–500	50–500
learning _rate	Step size shrinkage to prevent overfitting.	0.001–0.2	0.001–0.2	0.001–0.2
Max depth	Maximum depth of trees to control model complexity.	3–12	3–12	3–12
Min child weight	Minimum sum of instance weight needed in a child node.	1–10	1–50	1–50
Subsample	Fraction of training samples used to grow trees.	0.5–1.0	–	0.5–1.0
Colsample bytree	Fraction of features used per tree.	0.5–1.0	–	0.5–1.0
gamma	Minimum loss reduction required to make a further partition on a leaf node.	0–1.0	–	–
Reg lambda	L2 regularization term on weights to prevent overfitting.	0–10	–	0–10
Reg alpha	L1 regularization term on weights to prevent overfitting.	0–10	–	0–10
Iterations	Number of boosting iterations.	–	0–10	–
l2 leaf reg	L2 regularization coefficient to prevent overfitting.	–	0–1.0	–
Bagging temperature	Controls the randomness of bagging. A higher value leads to more randomness.	–	8–255	–
Border count	Number of splits for numerical features.	–	–	8–128
Num leaves	Maximum number of leaves in one tree.	–	–	1–50
Min data in leaf	Minimum number of samples a leaf must have.	–	–	0.3–1.0
Teature fraction	Fraction of features to be used for each iteration.	–	–	0.3–1.0
Bagging fraction	Fraction of data to be used for each iteration.	50–500	50–500	50–500
Lambda l1	L1 regularization term on weights to prevent overfitting.	0.001–0.2	0.001–0.2	0.001–0.2
Lambda l2	L2 regularization term on weights to prevent overfitting.	3–12	3–12	3–12
Population size	Number of individuals in the population.		20	
Max iterations	Maximum number of GA iterations.		50	
Mutation probability	Probability of mutation.		0.1	
Crossover probability	Probability of crossover.		0.5	
Elitism ratio	Ratio of top-performing individuals retained.		0.01	
Parents portion	Portion of parents selected for mating.		0.3	
Crossover type	Type of crossover used.		Uniform	
Max no improve	Maximum iterations without improvement.		10	

generalization by leveraging the complementary strengths of diverse base models. The framework also employed early stopping criteria, halting training after 10 iterations without improvement, to minimize computational overhead. By coupling GA-driven hyperparameter tuning with a stacked ensemble, this hybrid approach delivered a robust, scalable, and efficient solution for predicting annual cooling load intensity, offering a powerful tool for energy management and optimization in building systems. The hyperparameters used for the proposed hybrid model with GA optimization is listed in Table 5.

2.4. Performance evaluation

The Mean Squared Error (MSE), Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), and R-squared (R^2) metrics are employed to assess the performance of the models. In these formulas, n represents the total number of observations, y_i denotes the actual value of the i-th data point, and $\hat{y}_i$ signifies the predicted value for the same point.

$$\mathrm{MSE} = \frac{1}{n}\sum\nolimits_{i=1}^{n}(y_i - \widehat{y_i}) \tag{1}$$

MSE calculates the squared residual, emphasizing larger errors, while 1/n averages the

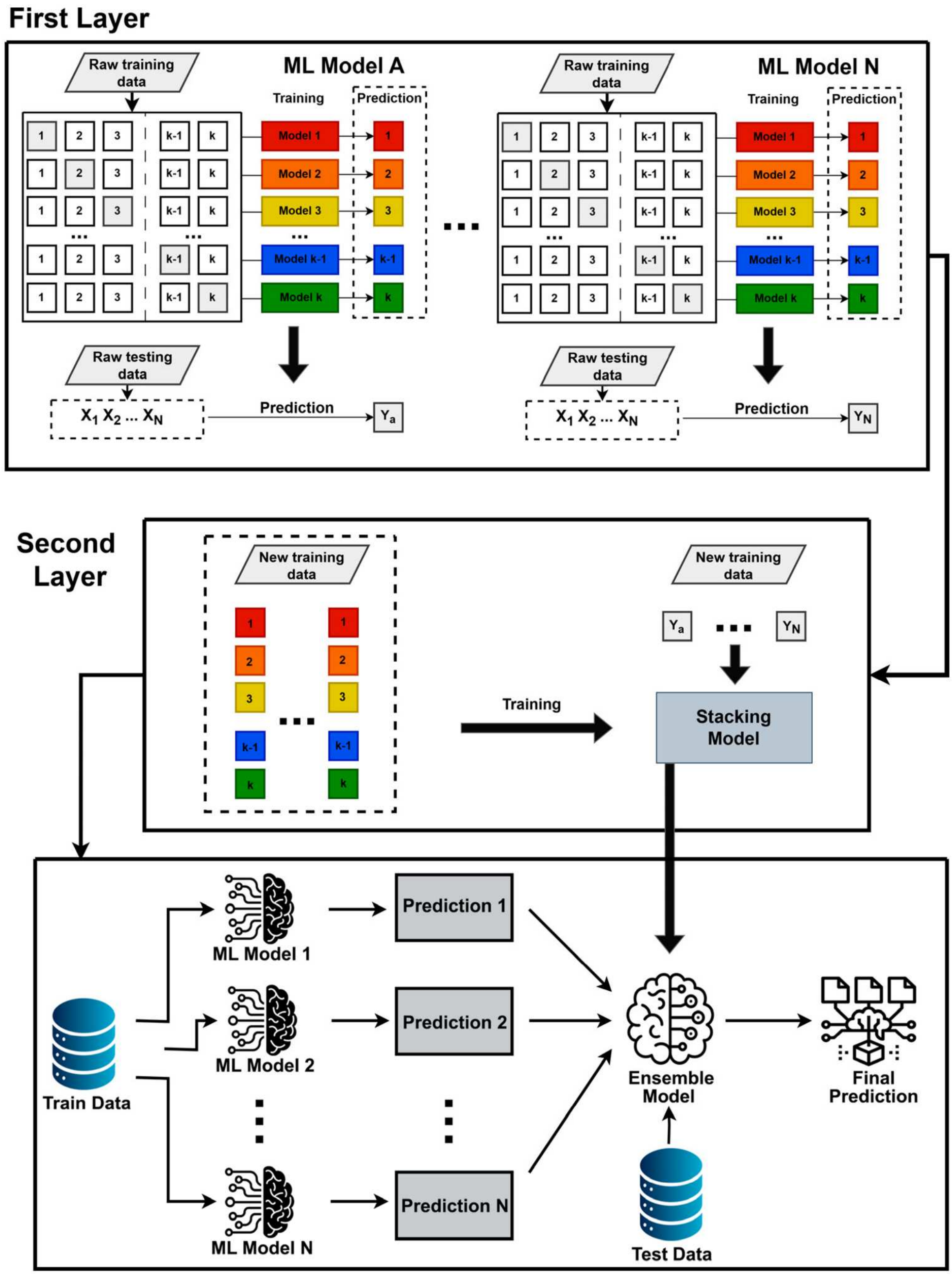


Figure 6. The architecture of the hybrid model.

sum over all data points.

$$\mathrm{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \widehat{y}_i)^2} \tag{2}$$

RMSE uses the same components but applies a square root to return the error to the original units, reflecting the standard deviation of residuals.

$$\mathrm{MAE} = \frac{1}{n}\sum_{i=1}^{n} |y_i - \hat{y}_i| \tag{3}$$

MAE uses the absolute difference to compute the average error magnitude, offering a robust measure less sensitive to outliers.

$$R^2 = \frac{\sum_{i=1}^{n}(X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\sum_{i=1}^{n}(X_i - \bar{X})^2 \sum_{i=1}^{N}(Y_i - \bar{Y})^2}} \tag{4}$$

R^2 represents the proportion of variance explained by the model. Xi and Yi are data points, and $\bar{X}$, $\bar{Y}$ are their means. The ratio when subtracted from 1, quantifies the proportion of variance explained by the model. Unlike MSE, RMSE, and MAE, R^2 is not an error metric but an indicator of model performance. Together, these metrics provide a comprehensive evaluation, with lower MSE, RMSE, and MAE values and an R^2 closer to 1 indicating superior model performance.

2.5. SHAP methodology

The SHAP algorithm is a powerful tool for analyzing how different input variables influence model predictions. It employs simplified explanation models that closely approximate the results of the original predictive model to interpret complex ML models. This process is mathematically represented by Equation 5 and Equation 6 from (Lundberg & Lee, 2017). In these equations, m denotes the simplified explanation model, while f represents the actual predictive model. The input feature vector is symbolized by x, which is mapped to $x' \in \{0, 1\}^N$. Here, N stands for the number of input variables, with 1 indicating a feature's inclusion in the simplified model and 0 indicating its exclusion. The sets F and S represent the complete set of features and a subset of features excluding the one with index i, respectively. x_S denotes the input feature values for the features in set S. The impact of a feature indexed by j is represented by ϕ_j, which is calculated based on the differences in model output when the feature is included or excluded from the model. For a more comprehensive explanation of the SHAP procedure, refer to (Lundberg & Lee, 2017).

$$m(x') = \phi_0 + \sum_{j=1}^{N} \phi_j x'_j \tag{5}$$

$$\phi_j = \sum_{S \subseteq F \setminus \{j\}} \frac{|S|!(|F| - |S| - 1)!}{|F|!} [\, f_{S\cup\{j\}}(x_{S\cup\{j\}}) - f_S(x_S)] \tag{6}$$

3. Results

This study focuses on formulating an urban-scale annual cooling load intensity prediction model using ML techniques. The process begins by generating synthetic data, which is then preprocessed to remove outliers and enhance data quality, ensuring robustness before applying ML models. Following preprocessing, the dataset is split into two subsets: 80% for training and 20% for testing. To mitigate the risk of overfitting and improve the model's generalizability, a 10-fold cross-validation method is employed instead of random selection for creating the training and testing sets. A proposed framework combining XGBoost, LightGBM, and CatBoost with a GA was applied alongside other single ML algorithms to evaluate their predictive capabilities for annual cooling load intensity. Each model was assessed using key performance metrics, including R^2, MAE, MSE, and RMSE. The superior models were identified as those with lower values of RMSE, MSE, and MAE, together with an R^2 value closer to 1, indicating higher accuracy and minimal prediction errors. The following subsections present the results of the analysis conducted in this research.

3.1. Statistical analysis of building characteristics

AI analysis is a statistical approach utilized to scrutinize data points collected at consistent intervals, with the aim of uncovering underlying patterns and trends. To asses the impact of individual features on single models and the ensemble-based model, the correlation coefficient was employed to estimate their performance. The correlation coefficient quantifies the degree to which one variable can predict the value of another variable. One of the most popular ways to measure the dependence between quantitative variables is by calculating the Pearson correlation coefficient. For instance, if $Y = a + bX$, the Pearson correlation coefficient between X and Y can be calculated using Equation 7.

$$\left[\rho(X,\ Y) = \frac{E[(X - E(X))(a + bX - E(X))]^2}{\left[\sum xi^2 - \left(\sum xi\right)^2\right]^{0.5}\left[\sum yi^2 - \left(\sum yi\right)^2\right]^{0.5}}\right] \tag{7}$$

where E denotes the mathematical expectation, and V denotes the variance. For a random sample of size n of random variables X and Y (X_i,Y_i), the Pearson sample correlation coefficient is calculated using Equation 8.

$$\left[r(X,\ Y) = \frac{n\sum xiyi - \sum(xi)\sum yi}{\left[\sum xi^2 - \left(\sum xi\right)^2\right]^{0.5}\left[\sum yi^2 - \left(\sum yi\right)^2\right]^{0.5}}\right] \tag{8}$$

In this formula, Σ denotes the sum of the values, and the square root of the variances is used to normalize the coefficient. Using the correlation coefficient in time series analysis enabled us to evaluate the performance of each feature and understand the relationship between the variables. This information was critical in predicting the dependent variable value based on the independent variables. The correlation coefficient was a valuable tool for analyzing the relationship between variables. The degree and direction of the linear relationship between annual cooling load intensity and the independent variables in the study are depicted in Figure 7.

3.2. Model performance

The performance of each model was assessed by comparing their results with those of similar solutions. Table 6 presents a summary of the results obtained after testing the new data. The results clearly demonstrate that the hybrid model, comprising XGBoost,

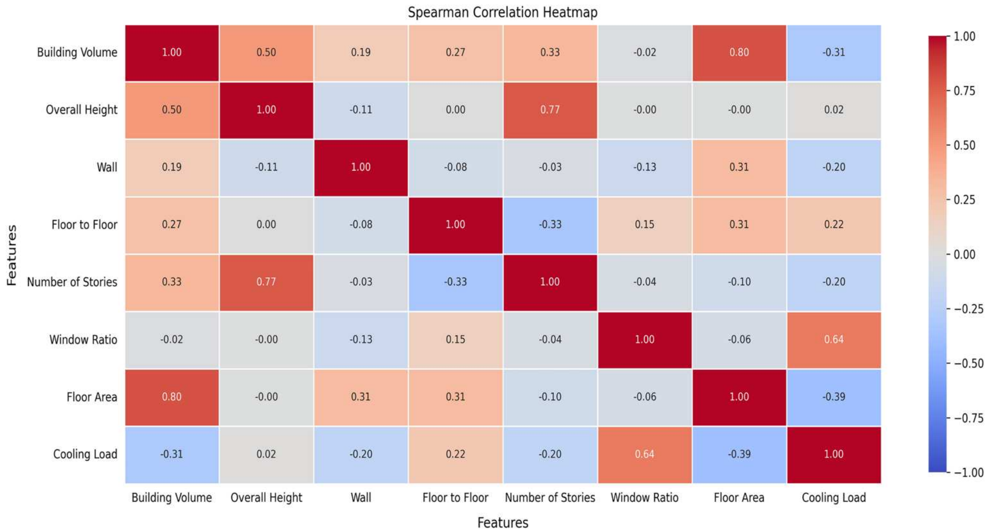


Figure 7. Correlation coefficient heat map.

LightGBM, CatBoost, and a GA, substantially outperforms all other models in predicting annual cooling load intensity. With an MAE of 0.26 kWh/m^2, an MSE of 1.19 kWh/m^2, an RMSE of 1.09 kWh/m^2, and an R^2 of 0.99, this ensemble method showcases high predictive accuracy and precision. The integration of these gradient-boosting algorithms with GA maximizes the strengths of each: XGBoost excels in handling missing values and offers powerful regularization, LightGBM stands out for its speed and efficiency with large datasets, and CatBoost effectively deals with categorical data and reduces overfitting. Meanwhile, GA enhances model optimization by fine-tuning hyperparameters, further boosting predictive performance. Together, these methods complement one another, enabling the model to capture non-linear relationships and complex interactions, yielding minimal residual errors and highly accurate predictions. Individually, each of these algorithms delivers competitive performance. XGBoost achieves an R^2 of 0.87, with an MAE of 6.32 kWh/m^2 and an RMSE of 7.99 kWh/m^2, while LightGBM records an R^2 of 0.85, an MAE of

Table 6. Comparison between machine learning models.

Models	MAE	MSE	RMSE	R^2
Hybrid model	0.26	1.19	1.09	0.99
XGBoost	6.32	63.87	7.99	0.87
LightGBM	6.86	74.61	8.64	0.85
CatBoost	5.98	71.88	8.48	0.85
ANN	6.66	91.54	9.57	0.81
KNN	7.17	98.94	9.95	0.80
DT	7.42	102.21	10.23	0.80
PReg	7.84	130.11	11.41	0.74
AdaBoost	10.26	164.99	12.84	0.67
GBR	11.16	209.94	14.49	0.57
Ridge	10.00	216.24	14.71	0.56
MLR	9.99	216.83	14.73	0.56
RF	9.45	227.65	14.84	0.54
QR	9.38	229.38	15.15	0.53
SVR	13.97	286.06	16.91	0.44

6.86 kWh/m^2, and an RMSE of 8.64 kWh/m^2. CatBoost performs similarly, with an R^2 of 0.85, an MAE of 5.98 kWh/m^2, and an RMSE of 8.48 kWh/m^2, demonstrating its strong ability to predict cooling loads. Among other tree-based models, DT achieves an R^2 of 0.79, with an MAE of 7.42 kWh/m^2 and an RMSE of 10.23 kWh/m^2, while RF records an R^2 of 0.54, an MAE of 9.45 kWh/m^2, and an RMSE of 14.84 kWh/m^2. Notably, RF underperforms relative to the boosting algorithms, suggesting that its bagging strategy is less effective than boosting for capturing the complex non-linear relationships present in cooling load data. These results underline the reliability of tree-based models in managing the intricacies of energy consumption patterns, as they handle high-dimensional data and variable interactions more effectively than simpler models.

On the other hand, single ML methods such as SVR and MLR perform less favourably in this context. The SVR model exhibits significant limitations, with an R^2 of 0.44, an MAE of 13.97 kWh/m^2, and a high RMSE of 16.91 kWh/m^2, indicating considerable prediction errors. This suggests that SVR struggles to model the non-linear relationships inherent in cooling load data. Similarly, MLR presents moderate performance with an R^2 of 0.56, an MAE of 9.99 kWh/m^2, and an RMSE of 14.73 kWh/m^2. Although it performs better than SVR, its simplicity in capturing complex patterns results in a noticeable gap in predictive power compared to the hybrid methods. Moreover, models like GBR and RR offer mixed results. The GBR model achieves an R^2 of 0.57, with an MAE of 11.16 kWh/m^2 and an RMSE of 14.49 kWh/m^2, while Ridge records an R^2 of 0.56, an MAE of 10.00 kWh/m^2, and an RMSE of 14.71 kWh/m^2. While both methods outperform simpler algorithms like SVR, their overall performance still falls short when compared to the more sophisticated hybrid techniques. This highlights the limitations of single-algorithm approaches, especially when dealing with highly complex data patterns. In contrast, more specialized algorithms such as QR and PReg show varying performance levels. QR delivers an R^2 of 0.53, an MAE of 9.38 kWh/m^2, and an RMSE of 15.15 kWh/m^2, reflecting its utility in handling skewed data but with limited overall accuracy. On the other hand, PReg demonstrates better performance, achieving an R^2 of 0.74, an MAE of 7.84 kWh/m^2, and an RMSE of 11.41 kWh/m^2, suggesting that non-linear transformations of features can capture more complex relationships in the data, albeit not as effectively as the hybrid model. Figure 8(a) to 8(c) show the annual cooling load intensity predictions on the models' test dataset. Overall, the hybrid model stands out as the most effective approach for predicting annual cooling load intensity, consistently achieving the lowest error rates and highest accuracy across all metrics. The ability to combine the strengths of multiple advanced tree-based methods with GA optimization provides a significant advantage in terms of both predictive power and model robustness.

3.3. SHAP analysis

Table 7 provides the ranking of features influencing ML models for annual cooling load intensity predictions across ML models, based on SHAP analysis. It indicates the average impact of various building characteristics and weather conditions on annual cooling load intensity. Several features exhibit distinct patterns in their influence on annual cooling load intensity. For instance, the window-to-wall ratio, floor area, and number of stories emerge as the most influential factors. Larger windows correlate strongly with higher cooling demands due to increased solar heat gain, while expansive

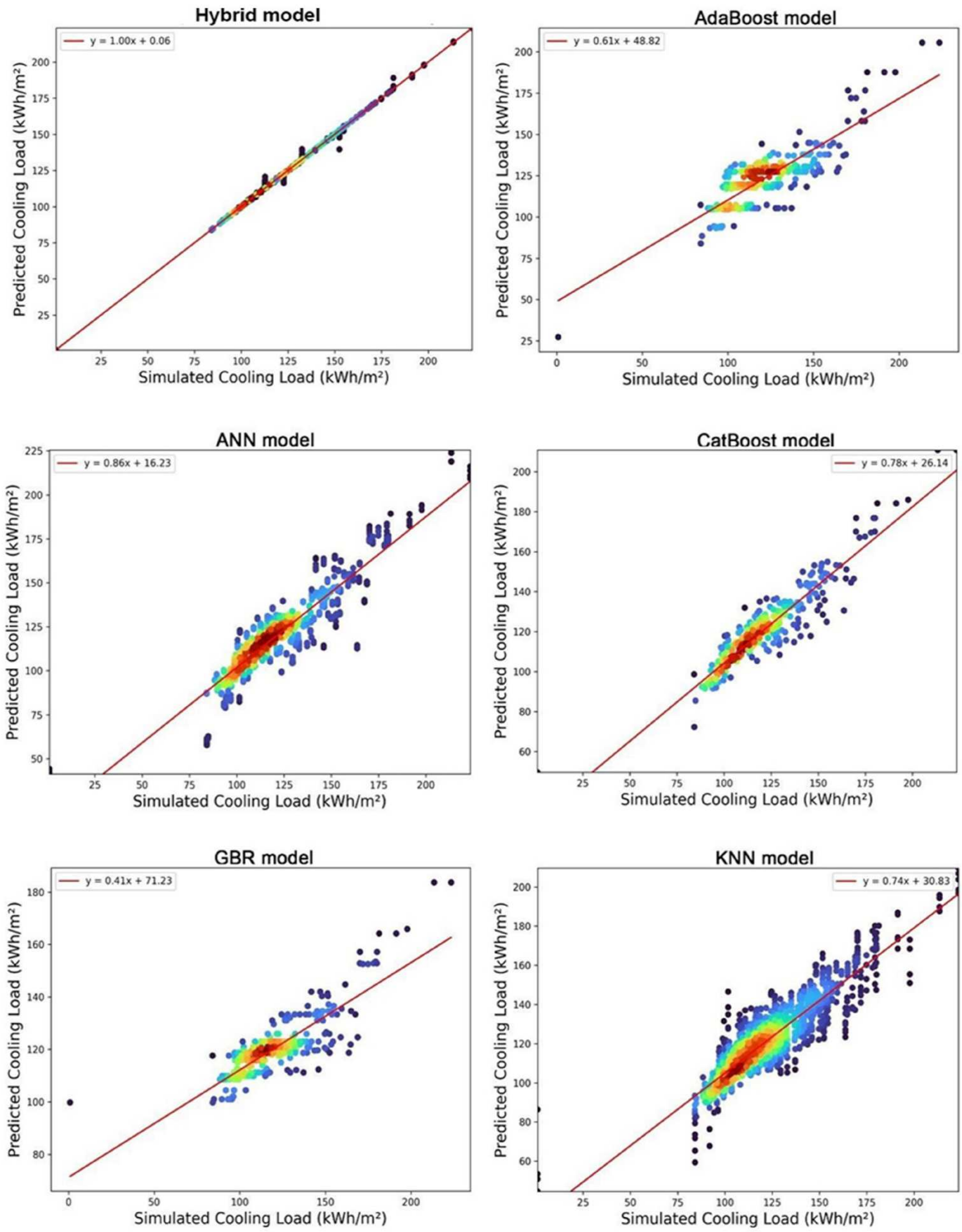


Figure 8. (a). Annual cooling load intensity predictions on the testing dataset of the models. (b) Annual cooling load intensity predictions on the testing dataset of the models. (c) Annual cooling load intensity predictions on the testing dataset of the models.

floor areas and multi-story designs amplify energy requirements by expanding the thermal zones needing climate control. Vertical dimensions, such as overall building height and floor-to-floor height, further compound these effects, as taller structures face greater exposure to external temperature fluctuations and convective heat transfer. In contrast, building volume and wall area exhibit more context- dependent impacts; for

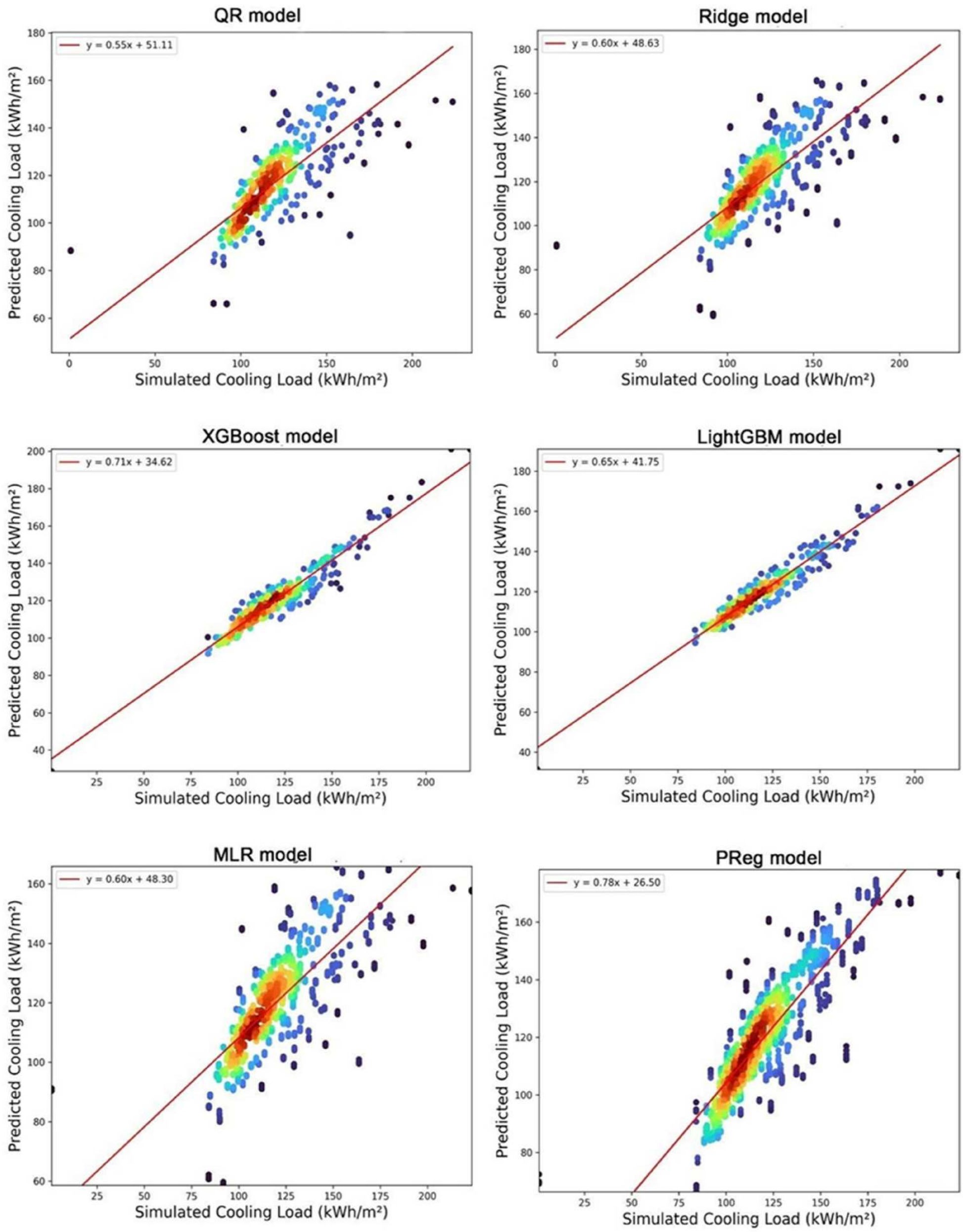


Figure 8. *Continued.*

instance, larger volumes may reduce annual cooling load intensity per unit area in well-designed spaces but escalate total energy consumption if airflow or insulation is inadequately optimized.

Environmental factors, particularly air temperature, surpass the influence of key building attributes in driving cooling demands. Air temperature dominates microclimate impacts due to its direct role in heat exchange, with higher temperatures intensifying cooling needs. Dew point temperature and relative humidity also play important roles,

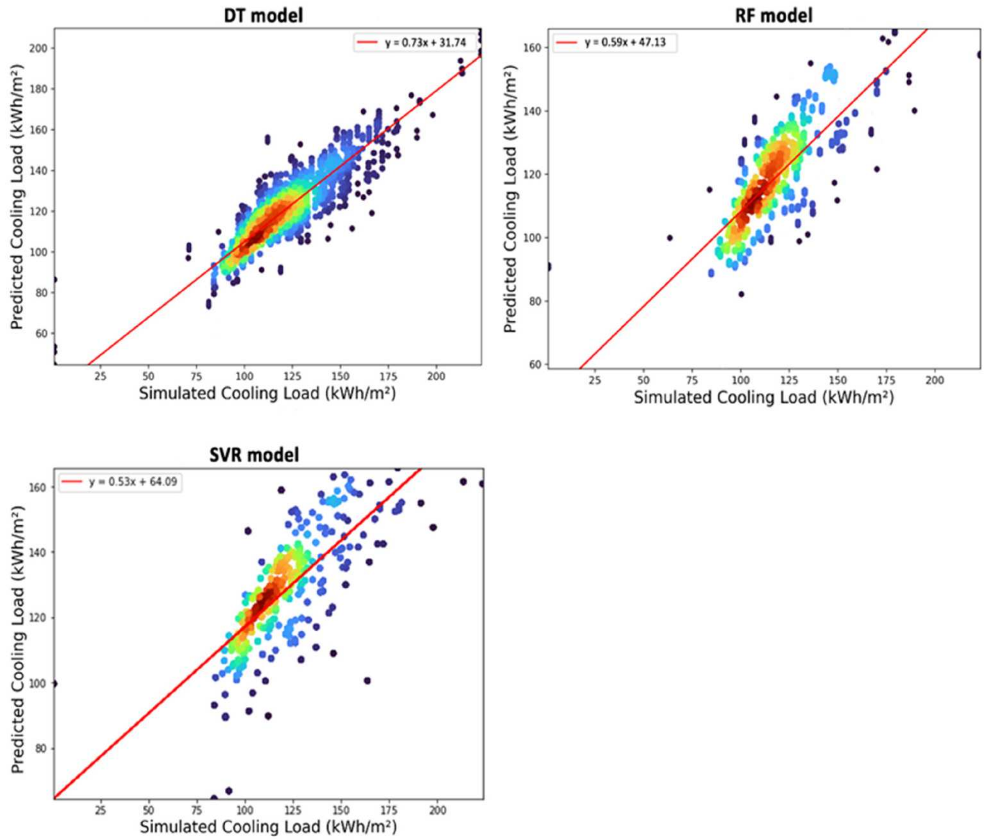


Figure 8. *Continued.*

Table 7. Ranking of features influencing machine learning models for annual cooling load intensity using SHAP analysis.

Rank	Building Parameter	Rank	Weather Parameter
1	Window-to-wall ratio	1	Air temperature
2	Floor area	2	Dew point temperature
3	Number of stories	3	Relative humidity
4	Overall height	4	Wind speed
5	Floor-to-floor height	5	Atmospheric pressure
6	Building volume		
7	Wall		

as elevated humidity levels force HVAC systems to expend additional energy on latent heat removal through dehumidification. Wind speed and atmospheric pressure, while less impactful, introduce nuanced effects: wind can either reduce annual cooling load intensity by enhancing natural ventilation or increase them by introducing warm air, depending on local climate patterns. These findings underscore that regions with extreme weather conditions, such as persistently high temperatures or humidity, may experience environmental factors overriding the influence of building design in determining peak energy use.

The relationship between building features and annual cooling load intensity is further nuanced by design choices and environmental adaptability. For example, floor-to-floor height demonstrates variable effects, as taller floor plates may improve airflow and reduce cooling needs in naturally ventilated buildings but increase energy consumption in sealed, mechanically cooled environments. Similarly, the weak ranking of wall area suggests that its influence is secondary to factors like insulation quality or window placement, which are not explicitly quantified in the analysis. This highlights the importance of holistic design strategies – such as minimizing excessive glazing or incorporating shading systems, to mitigate the drawbacks of high-impact features like large windows. Meanwhile, the modest role of atmospheric pressure implies its influence is likely indirect, correlating with broader weather systems rather than directly driving cooling demands.

Ultimately, the SHAP rankings emphasize that optimizing annual cooling load intensity requires balancing architectural design with climate-responsive strategies. While structural parameters like window-to-wall ratio and building scale set the baseline for energy consumption, environmental conditions, particularly air temperature and humidity, dictate operational intensity, especially in extreme climates. This duality suggests that sustainable solutions demand both passive design measures (e.g. optimizing building proportions, leveraging natural ventilation) and active systems tailored to local microclimates. By prioritizing high-impact features while accounting for regional weather patterns, designers and engineers can develop buildings that are both energy-efficient and resilient to fluctuating environmental demands.

To examine the robustness of the SHAP-based feature ranking, the relative importance of features was compared across different cross-validation folds. The results show that the top-ranked features, such as window-to-wall ratio, floor area, and number of stories, appear consistently among the most influential variables across training subsets, indicating limited sensitivity to data partitioning. In contrast, features with lower SHAP values show greater variation across folds, suggesting that their influence may depend more on the specific training data. The SHAP analysis was conducted on the final hybrid stacking ensemble model to reflect the combined behaviour of all base learners. To ensure the interpretation reflects the model's generalizable behaviour, SHAP values for global interpretation were computed using the held-out test set. This approach provides a stable and unbiased basis for explaining the model's predictions on unseen data. The difference between the correlation analysis and SHAP-based ranking is due to the different nature of these methods. Correlation analysis measures the direct linear relationship between each input parameter and annual cooling load intensity, whereas SHAP evaluates each feature's contribution within the trained nonlinear hybrid model, including interactions among variables. Therefore, although building volume may show a relatively strong correlation with cooling load, its SHAP ranking is lower because part of its influence overlaps with related geometric variables such as floor area, number of stories, and overall height. The results of the SHAP analysis is shown in Figure 9.

3.4. Validation of the proposed framework and the machine learning model

A crucial step in annual cooling load intensity prediction is assessing how well a model generalizes to unseen data. To rigorously evaluate their generalizability, robustness, and reliability, the ML models were tested them on 100 new data from Scottsdale,

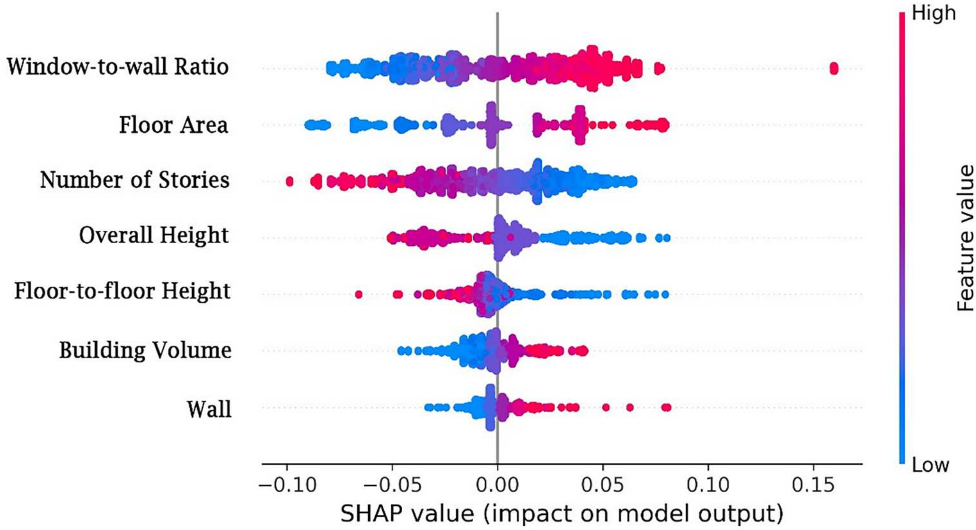


Figure 9. Feature importance.

Arizona, focusing on annual cooling load intensity predictions for residential buildings with varying shapes, orientations, and locations. This validation step is essential, as real-world energy performance is influenced by multiple external factors, including climate variations, urban morphology, and building design parameters. The Scottsdale dataset was generated using the same parametric and simulation framework as the main dataset, with Typical Meteorological Year (TMY) weather data specific to Scottsdale.

Figure 10(a) to 10 (e) illustrate the performance of different models when predicting annual cooling load intensity for these newly introduced data points. The comparative analysis offers valuable insights into how effectively each approach adapts to unseen cases beyond the training dataset. The comparative analysis of the models reveals that individual models struggled to generalize effectively to the new dataset. Across various algorithms, including QR, MLR, RR, KNN, PReg, and CatBoost, significant deviations between simulated and predicted annual cooling load intensities were observed. These deviations were especially prominent in buildings with higher annual cooling load intensities, suggesting that these models struggled to capture complex interactions between building geometry and energy demand. In contrast, the proposed hybrid framework demonstrated superior predictive accuracy and generalizability com- pared to the individual models. The predicted values from the hybrid model closely aligned with the simulated annual cooling load intensities, minimizing the error across the dataset. Table 8 presents a summary of the results obtained for the unseen data. This is evident in Figure 10(a), where the hybrid model consistently tracks the simulated values, reducing both over- and under-prediction issues seen in other models. The main advantages of the hybrid model over the individual models include:

a. Improved Accuracy: The hybrid framework achieved the lowest deviation from the simulated values, with minimal residual error between predicted and actual annual cooling load intensities.

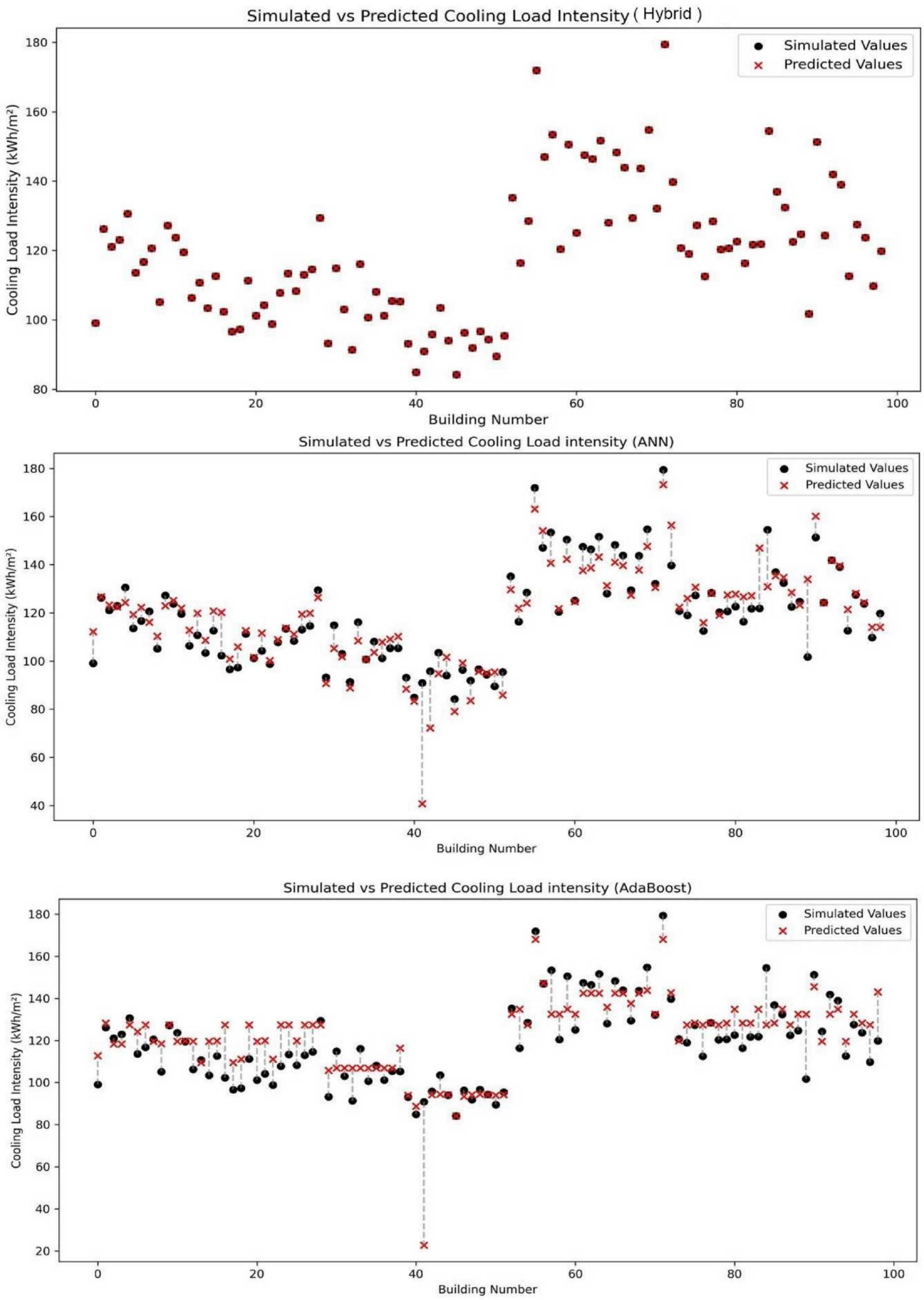


Figure 10. (a) Performance comparison of various machine learning models in predicting annual cooling load intensities for residential buildings using new data from Scottsdale, Arizona. (b) Performance comparison of various machine learning models in predicting annual cooling load intensities for residential buildings using new data from Scottsdale, Arizona. (c) Performance comparison of various machine learning models in predicting annual cooling load intensities for residential buildings using new data from Scottsdale, Arizona. (d) Performance comparison of various machine learning models in predicting annual cooling load intensities for residential buildings using new data from Scottsdale, Arizona. (e) Performance comparison of various machine learning models in predicting annual cooling load intensities for residential buildings using new data from Scottsdale, Arizona.

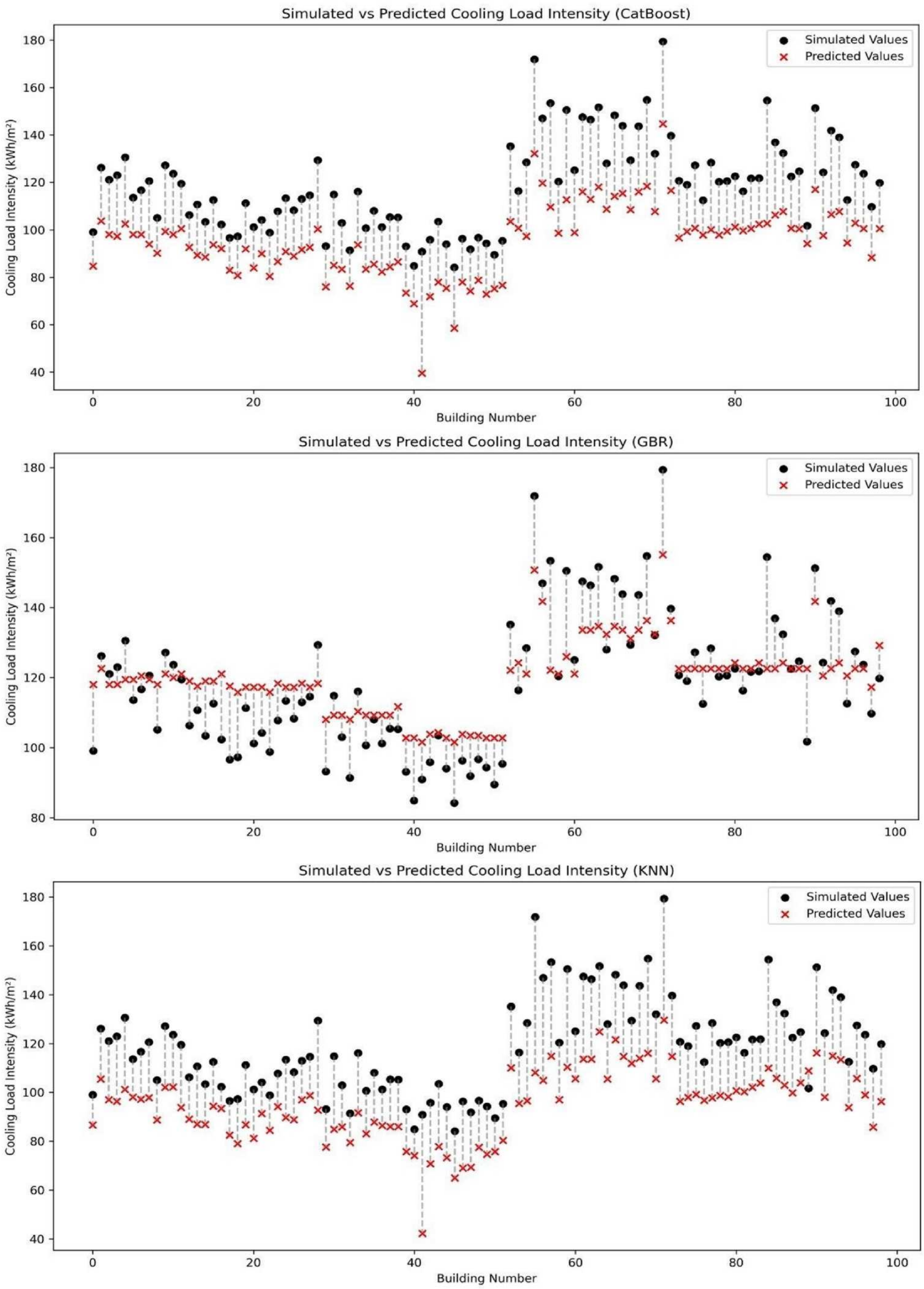


Figure 10. *Continued.*

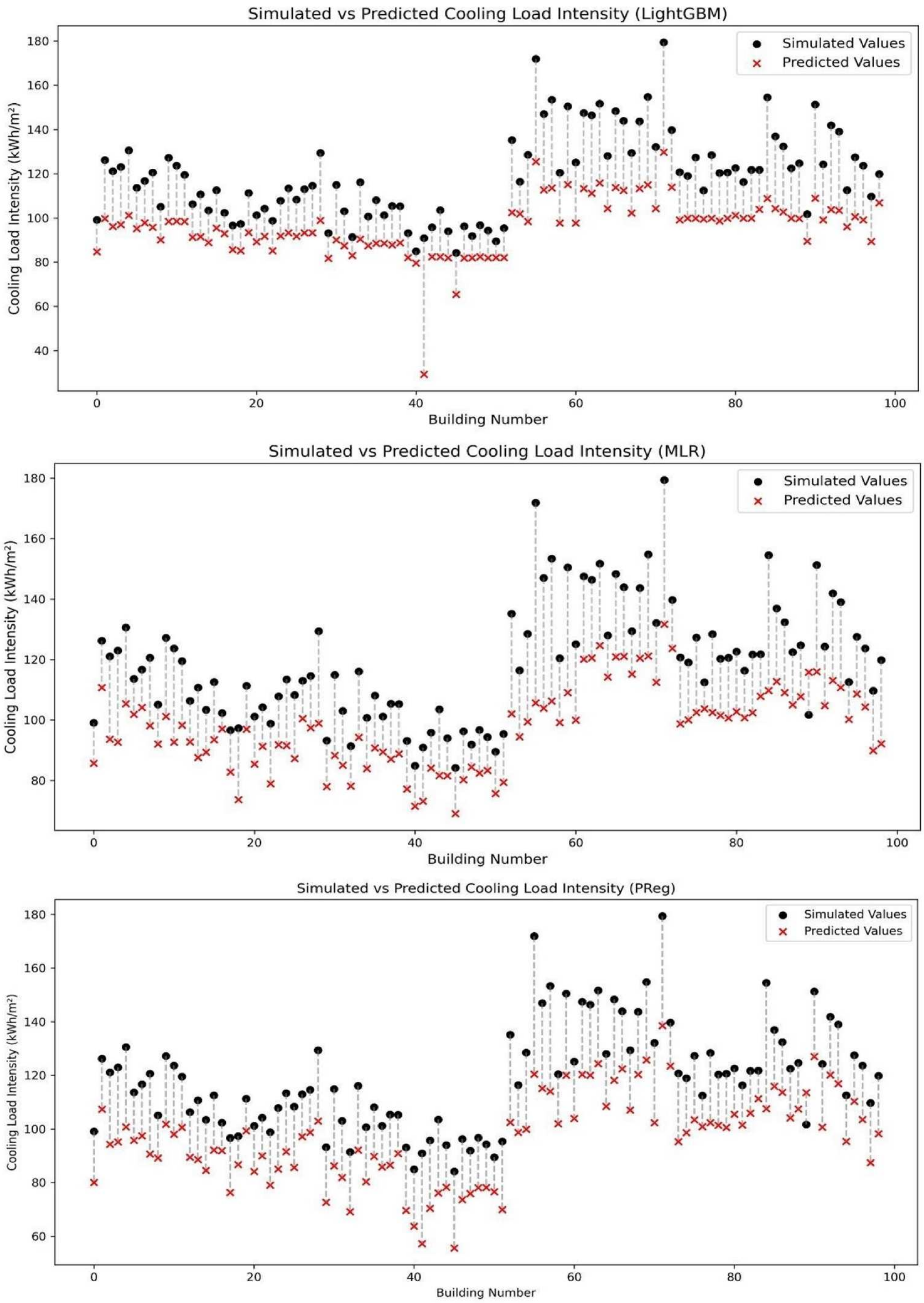


Figure 10. *Continued.*

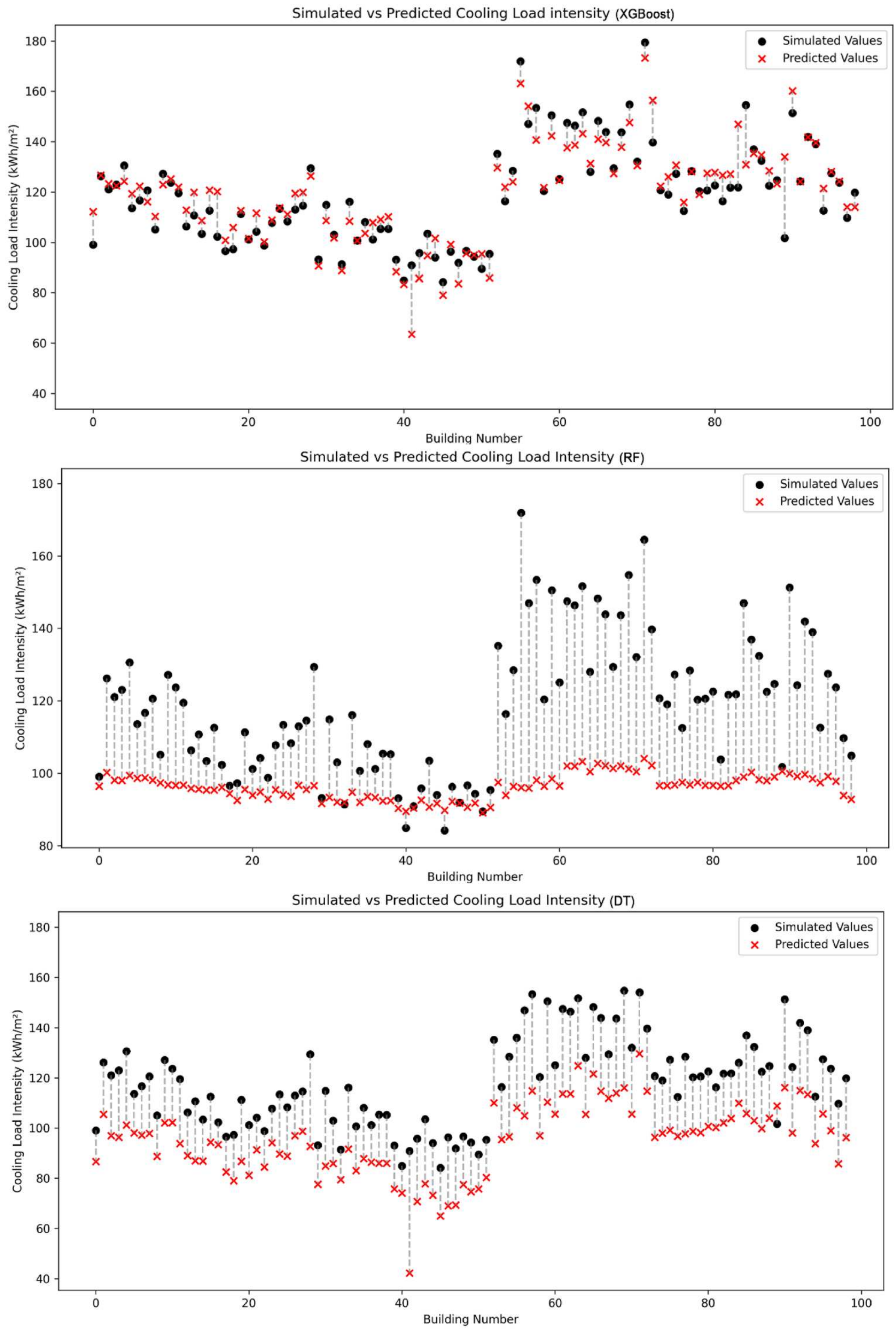
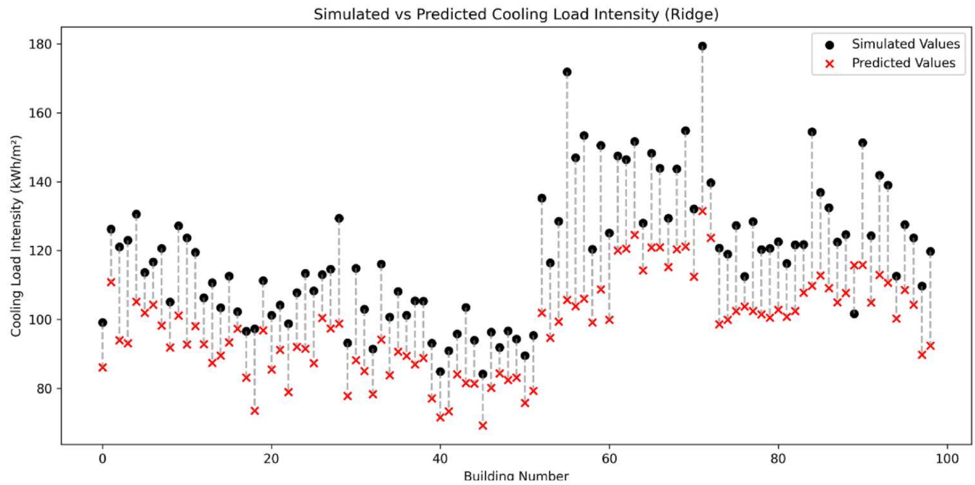
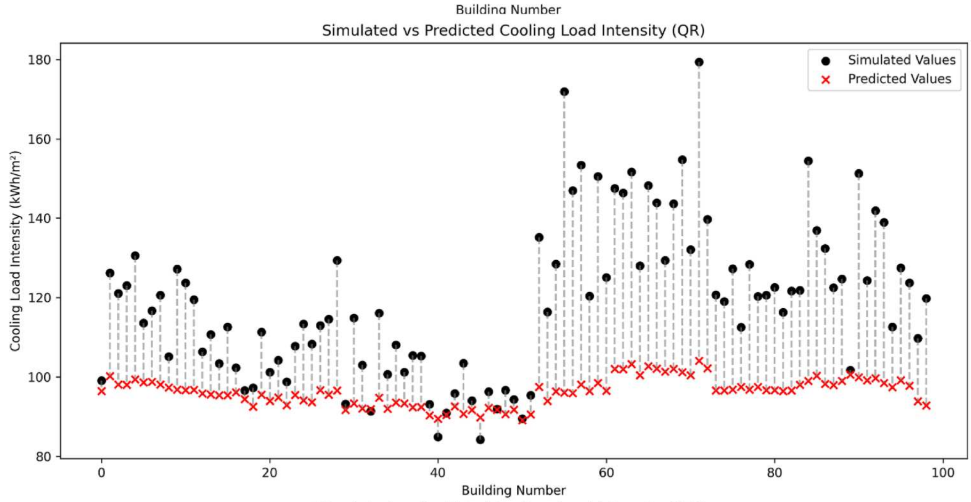


Figure 10. *Continued.*





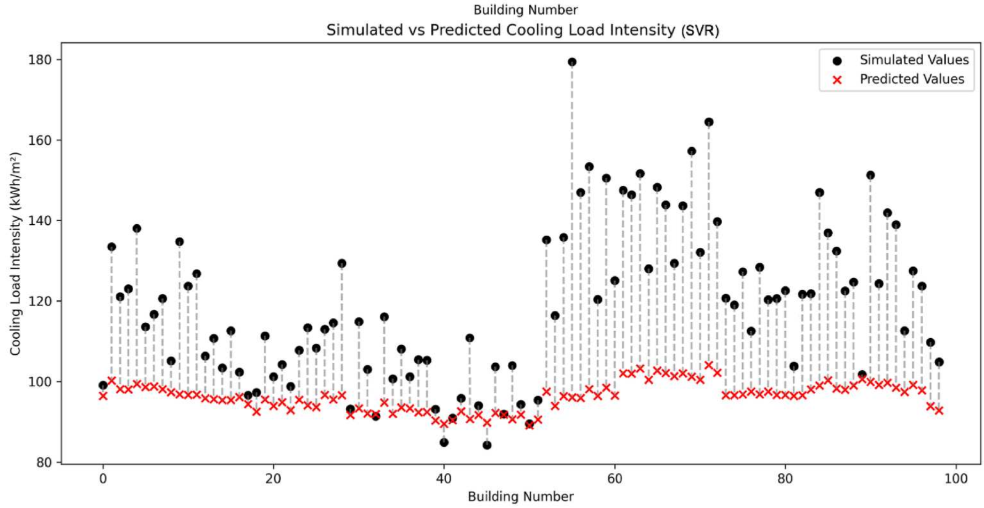


Figure 10. *Continued.*

Table 8. Comparison between machine learning models for unseen data.

Models	MAE	MSE	RMSE	R^2
Hybrid model	0.45	1.37	1.27	0.97
XGBoost	7.42	64.84	8.80	0.84
LightGBM	7.98	75.74	8.87	0.82
CatBoost	6.35	72.25	9.08	0.81
ANN	7.31	91.84	9.77	0.80
KNN	7.48	99.36	10.45	0.77
DT	8.32	120.23	11.03	0.75
Preg	9.14	131.52	11.74	0.72
AdaBoost	10.88	165.36	13.02	0.64
GBR	11.74	210.28	15.25	0.54
Ridge	10.65	216.78	15.44	0.53
MLR	10.58	217.36	15.85	0.50
RF	10.35	220.25	15.94	0.49
QR	10.14	229.99	16.02	0.47
SVR	14.36	286.78	17.41	0.39

b. Better Generalization: Unlike individual models, which exhibited large deviations in specific regions of the dataset, the hybrid approach-maintained stability across different building types and annual cooling load intensities.
c. Enhanced Robustness: The hybrid model successfully handled variations in architectural features, location, and energy profiles, making it a more reliable choice for large-scale urban energy predictions.

Although models such as ANN and AdaBoost performed relatively well, they still exhibited noticeable deviations in buildings with extreme annual cooling load intensities. This highlights the limitation of single-model approaches in capturing the intricate relationships among building parameters. The ensemble nature of the hybrid framework mitigated these shortcomings by leveraging the strengths of multiple algorithms, leading to more reliable and scalable predictions. Moreover, the applicability of the hybrid model extends beyond this specific dataset. Given its strong performance in this validation phase, it holds promise for large-scale deployment in urban energy planning, enabling policymakers and engineers to develop energy-efficient strategies with higher confidence. The ability to accurately estimate cooling loads across diverse building typologies enhances its relevance for practical implementation in sustainability-focused urban development initiatives.

4. Discussion

This study presents a robust data-driven framework for predicting annual cooling load intensities at an urban scale, utilizing a synthetic dataset of 8,760 residential buildings generated for the hot-dry climate of Phoenix, Arizona. By integrating advanced ensemble ML techniques, XGBoost, LightGBM, and CatBoost, with GA optimization, the model achieves predictive accuracy (RMSE of 1.09 kWh/m^2 and R^2 of 0.99), significantly surpassing traditional single-model approaches. This high performance underscores the potential of hybrid ML frameworks to tackle the complexities of urban energy modelling, where diverse building characteristics and microclimate variables intersect. To validate the developed framework and models, annual cooling load intensity predictions were performed on a different dataset of 100 buildings in Scottsdale, a city with varying urban fabric and building types not introduced during model training. The results, illustrated in Figure 10(a) to 10(e), demonstrated that the developed framework accurately estimated annual cooling load intensities in a different city under distinct conditions. Unlike traditional energy simulation methods like EnergyPlus, which are computationally intensive and data-heavy, this approach offers a faster, resource-efficient alternative, potentially reducing the time and cost of urban-scale analyses. In addition, SHAP analysis supports interpretability by identifying influential variables like window-to-wall ratio, floor area, and air temperature, providing actionable insights for architects and policymakers.

The framework simplifies the prediction of residential building energy consumption, highlighting significant potential for energy savings during the early design stage. By forecasting annual cooling loads before construction, design schemes can be optimized and retrofit strategies prioritized to minimize carbon emissions. This contributes directly to climate mitigation and sustainable urban development. As such, monitoring and managing energy consumption becomes more effective when predictions can be generated quickly and compared across design, retrofit, and operational scenarios. The workflow may also reduce the effort required to build and run large numbers of simulation models, particularly relative to white-box approaches that can be time-intensive and require specialized expertise. However, practical deployment depends on access to reliable input data. Emerging IoT and sensor technologies could help address this limitation by enabling the integration of real-time building and environmental data. With measured data streams, the framework could be extended from static, simulation-based assessment toward ongoing monitoring and optimization of energy performance. Future work could expand the framework to other building types and climates, with validations using measured building energy and occupancy data. Broader application will depend on improving data availability and maintaining transparency in the model's decision process to support stakeholder trust and use in sustainable urban planning.

4.1. Practical applications

The predictive framework offers a scalable approach to managing annual cooling load intensities in urban settings, particularly in extreme climates like Phoenix, Arizona. Its accuracy and ability to generalize across unseen data, demonstrated through validation in Scottsdale, position it as a versatile tool with wide-ranging practical applications. This hybrid model can drive real-world solutions for energy efficiency, urban planning, and climate resilience, addressing the escalating energy challenges posed by rapid urbanization and rising global temperatures. Below, innovative applications of this framework are explored, leveraging its strengths to enhance sustainability and operational efficiency in urban environments.

a. Optimizing Building Design and Retrofitting: The SHAP analysis identified variables such as window-to-wall ratio, floor area, and air temperature as influential drivers of annual cooling load intensity, which can help translate model outputs into design-relevant insights. In early-stage design, these insights can support screening of passive cooling strategies (e.g. reducing glazing on highly sun-exposed façades, improving shading, or adjusting orientation to limit solar heat gains). For existing buildings, the framework can support prioritization of retrofit measures (e.g. shading devices, glazing upgrades, or envelope improvements) by highlighting typologies and features associated with higher simulated annual cooling load intensities, particularly in hot-dry climates such as Phoenix.

b. Enhancing demand-Side energy management: Accurate annual cooling load intensity predictions enable utilities and building managers to implement demand-side management (DSM) strategies effectively. By estimating annual cooling load intensities at an urban scale, the model can support dynamic load shifting, encouraging off-peak usage through time-of-use pricing or smart HVAC controls. For example, in Phoenix, where summer temperatures often exceed 40°C, the framework could predict daily or hourly cooling peaks, allowing utilities to pre-cool buildings during off-peak hours or activate demand response programs to curtail usage during grid stress. This not only reduces strain on energy infrastructure but also lowers operational costs for consumers, aligning with smart city initiatives that prioritize grid stability and efficiency.

c. Supporting policymaking and urban sustainability: The workflow can also support planning and policy analysis by enabling scenario-based assessment across large building stocks. By estimating annual cooling load intensities for many buildings (as demonstrated with the synthetic dataset of 8,760 Phoenix residences), planners can identify areas with relatively high simulated cooling intensity and evaluate interventions such as cool roofs, shading policies, or urban greening strategies. Model's scalability suggests it could underpin municipal energy benchmarks, setting annual cooling load intensity thresholds for new constructions or incentivizing retrofits through tax credits. In extreme climates, this could translate to city-wide energy savings of 10–15%, based on studies of urban heat island mitigation (Tehrani et al., 2025), while enhancing resilience against heatwaves exacerbated by climate change.

d. Empowering community-level energy solutions: At a smaller scale, the framework can support community energy initiatives, such as microgrids or shared cooling systems. By predicting annual cooling load intensities for clusters of buildings, it can optimize the sizing and operation of localized energy resources – e.g. solar-powered chillers or thermal storage units. In a Phoenix suburb, for instance, a microgrid serving 50 homes could use the model to balance cooling demands with solar generation, storing excess energy in ice banks during the day for nighttime use. This could reduce reliance on fossil-fuel-based grids, cutting emissions by 25–30% per community, based on typical solar microgrid efficiency gains, while fostering energy independence.

e. Bridging gaps in data-scarce regions: Although this study used simulation-derived data, the workflow can be adapted for data-limited contexts by combining open-source building information (e.g. OpenStreetMap geometry) with local climate inputs. In cities with limited energy monitoring, proxy data (e.g. footprints, height estimates, construction archetypes, and climate zone assumptions) may enable preliminary cooling-demand screening. Such use should be treated as an initial estimate and strengthened through calibration and validation as measured data become available.

4.2. Limitations

The main limitation of this study is the absence of detailed occupant behaviour data. Occupancy patterns and occupant thermal preferences in energy simulations are typically represented in a simplified manner, which fails to capture the full complexity of real-world residential actions. Despite this limitation, the findings remain valuable, as they highlight the significant role of building characteristics and weather conditions in estimating annual cooling load intensity. Future studies should incorporate real occupant behaviour data to validate and refine the model, enabling it to better reflect actual building energy use. Additionally, in this study, the number of walls were included as a parameter to capture variations in external surface area and heat-exchange potential across different building typologies in the study area. We also recommend that parameters such as exposed wall area (Li & Gou , 2025) or surface-to-volume ratio (Li et al., 2024) be considered for this purpose in future studies. Another limitation was the access to real buildings for validation. Future research can use the proposed model to test on real buildings data.

4.3. Future research directions

The recommendations for future studies are as follows:

a. Real-Time monitoring and control systems with digital twin integration: Future research could focus on the development of real-time monitoring and control systems for managing cooling loads in smart buildings and cities, with the integration of digital twin technology. This would involve incorporating the developed models into Internet of Things (IoT) platforms to enable automated adjustment of cooling systems based on real-time data and environmental conditions (Zhuang et al., 2023), while leveraging digital twin simulations to enhance predictive capabilities and optimize system performance (Zhu & Lin, 2024).
b. Integration with building energy management systems (BEMS): Further exploration is needed into the integration of AI-based models and computational intelligence techniques with Building Energy Management Systems (BEMS) to optimize energy usage in buildings (Manic et al., 2016). Future research could develop algorithms for dynamic energy management, considering factors such as occupancy patterns, energy pricing, and renewable energy availability, to improve overall system efficiency.
c. User-friendly applications: Developing user-friendly software tools for stakeholders such as building owners, facility managers, and urban planners is crucial to facilitate the practical implementation of the models. These applications could feature intuitive interfaces for visualizing cooling load predictions, setting energy-saving goals, and optimizing building parameters for enhanced decision-making as part of smart energy management systems (Tekler et al., 2022a). Beyond their development, further investigation into user adoption is needed to inform policy and design implications, particularly from an urban energy management perspective, where user acceptance and organizational policies are central to successful implementation (Tekler et al., 2022b).

d. Long-term performance monitoring in urban heat island (UHI) contexts: Future studies should focus on the long-term monitoring of various important indicators and the evaluation of AI models deployed in real-world environments, particularly in urban heat island (UHI) contexts (Tehrani et al., 2025). Collecting and analyzing data on energy consumption, building performance, and user satisfaction over extended periods could provide valuable insights into the effectiveness of the models in mitigating UHI effects and improving urban microclimate resilience.

5. Conclusions

As energy policymakers increasingly prioritize city-scale assessment of building energy performance, the need for data-driven tools to inform sustainable policy development has never been more critical. Reducing energy consumption across urban environments is a cornerstone of global sustainability efforts, yet the traditional approaches to gathering and analyzing large-scale building energy data remain complex, time-consuming, and resource-intensive. These challenges often hinder timely decision-making and limit the ability of urban planners and policymakers to implement effective energy-saving strategies early in the design and planning process. In response to these obstacles, this study introduces a hybrid framework that provides a rapid, accurate, and integrated solution for assessing the energy performance of complex building systems at scale. By streamlining data processing and incorporating advanced computational techniques, the framework offers a transformative approach to urban energy management.

The proposed methodology integrated data collection and integration, parametric energy simulation, hybrid model development, optimization, and validation, predictive modelling and SHAP feature analysis, providing interpretable insights into the factors driving energy consumption. The framework was applied using a dataset synthetic dataset of 8,760 residential buildings in Phoenix, generated through parametric energy analysis. Comparative evaluation using traditional ML methods and ensemble models confirmed that the optimized hybrid model achieved an exceptional coefficient of determination of 99%, far surpassing the performance of conventional approaches. By providing precise and interpretable predictions annual cooling load intensity, the framework equips stakeholders, ranging from city planners to building developers with actionable insights.

The integration of ML techniques with SHAP analysis further enhances its value by identifying the significant drivers of energy consumption, such as building design parameters, occupancy patterns, and climatic conditions. Furthermore, the insights derived from SHAP analysis could guide the prioritization of retrofitting efforts, the optimization of building codes, or the deployment of passive cooling strategies in urban heat islands. More broadly, this research highlights the role of hybrid ML techniques in addressing the complexities of urban energy systems. The ability to predict annual cooling load intensity with such high fidelity facilitates more effective energy planning and conservation strategies, aligning with global objectives to mitigate climate change and enhance urban resilience. In the context of Phoenix, where extreme heat drives higher energy demand, this framework offers a scalable approach with its adaptability across a wide range of urban contexts, providing a practical tool to support data-driven energy planning and resilience.

Author contributions

CRediT: **Alireza Attarhay Tehrani:** Conceptualization, Data curation, Formal analysis, Methodology, Software, Visualization, Writing – original draft, Writing – review & editing; **Saeideh Sobhaninia:** Supervision, Validation, Writing – original draft, Writing – review & editing; **Aurora Bertini:** Writing – original draft, Writing – review & editing; **Shady Attia:** Investigation, Supervision, Writing – original draft, Writing – review & editing; **Zeynep Duygu Tekler:** Supervision, Validation, Writing – original draft, Writing – review & editing; **Shiva Manshour:** Validation, Writing – review & editing; **Deepak Amaripadath:** Validation, Visualization, Writing – original draft, Writing – review & editing.

Disclosure statement

No potential conflict of interest was reported by the author(s).

ORCID

Alireza Attarhay Tehrani https://orcid.org/0009-0005-9391-311X
Saeideh Sobhaninia https://orcid.org/0000-0003-1425-3060
Aurora Bertini https://orcid.org/0009-0002-2299-0079
Shady Attia http://orcid.org/0000-0002-9477-5098
Zeynep Duygu Tekler https://orcid.org/0000-0002-1858-0846
Shiva Manshour https://orcid.org/0009-0009-0374-413X
Deepak Amaripadath https://orcid.org/0000-0003-2822-5749

References

Ahmad, M. W., Mourshed, M., & Rezgui, Y. (2017). Trees vs neurons: Comparison between random forest and ANN for high-resolution prediction of building energy consumption. *Energy and Buildings*, *147*, 77–89. https://doi.org/10.1016/j.enbuild.2017.04.038

Ahmad, T., Chen, H., Guo, Y., & Wang, J. (2018). A comprehensive overview on the data driven and large scale based approaches for forecasting of building energy demand: A review. *Energy and Buildings*, *165*, 301–320. https://doi.org/10.1016/j.enbuild.2018.01.017

Al-Ajmi, F. F., & Hanby, V. (2008). Simulation of energy consumption for Kuwaiti domestic buildings. *Energy and Buildings*, *40*(6), 1101–1109. https://doi.org/10.1016/j.enbuild.2007.10.010

Ali, U., Bano, S., Shamsi, M. H., Sood, D., Hoare, C., & Zuo, W. (2024). Urban building energy performance prediction and retrofit analysis using data-driven machine learning approach. *Energy and Buildings*, *303*, 113768. https://doi.org/10.1016/j.enbuild.2023.113768

Almaleck, P., Massucco, S., Mosaico, G., Saviozzi, M., Serra, P., & Silvestro, F. (2024). Electrical consumption forecasting in sports venues: A proposed approach based on neural networks and ARIMAX models. *Sustainable Cities and Society*, *100*, 105019. https://doi.org/10.1016/j.scs.2023.105019

Alobaidi, M. H., Chebana, F., & Meguid, M. A. (2018). Robust ensemble learning framework for day-ahead forecasting of household based energy consumption. *Applied Energy*, *212*, 997–1012. https://doi.org/10.1016/j.apenergy.2017.12.054

Álvarez-Sanz, M., Satriya, F. A., Terés-Zubiaga, J., Campos-Celador, A., & Bermejo, U. (2024). Ranking building design and operation parameters for residential heating demand forecasting with machine learning. *Journal of Building Engineering*, 108817. https://doi.org/10.1016/j.jobe.2024.108817

Amaripadath, D., Joshi, M. Y., Hamdy, M., Petersen, S., Stone, B., & Attia, S. (2026). Thermal resilience in a renovated nearly zero-energy dwelling during intense heat waves. *Journal of Building Performance Simulation*, *19*(1), 86–105. https://doi.org/10.1080/19401493.2023.2253460

Amaripadath, D., Levinson, R., Rawal, R., & Attia, S. (2024). Multi-criteria decision support framework for climate change-sensitive thermal comfort evaluation in European buildings. *Energy and Buildings*, *303*, 113804. https://doi.org/10.1016/j.enbuild.2023.113804

Amasyali, K., & El-Gohary, N. M. (2018). A review of data-driven building energy consumption prediction studies. *Renewable and Sustainable Energy Reviews*, *81*, 1192–1205. https://doi.org/10.1016/j.rser.2017.04.095
Amber, K., Ahmad, R., Aslam, M., Kousar, A., Usman, M., & Khan, M. S. (2018). Intelligent techniques for forecasting electricity consumption of buildings. *Energy*, *157*, 886–893. https://doi.org/10.1016/j.energy.2018.05.155
Amber, K. P., Aslam, M. W., Mahmood, A., Kousar, A., Younis, M. Y., & Akbar, B. (2017). Energy consumption forecasting for university sector buildings. *Energies*, *10*(10), 1579. https://doi.org/10.3390/en10101579
Babu, A. D. C. A., Srivastava, R. S., & Rai, A. C. (2024). Impact of climate change on the heating and cooling load components of an archetypical residential room in major Indian cities. *Building and Environment*, *250*, 111181. https://doi.org/10.1016/j.buildenv.2024.111181
Ben-Nakhi, A. E., & Mahmoud, M. A. (2004). Cooling load prediction for buildings using general regression neural networks. *Energy Conversion and Management*, *45*(13-14), 2127–2141. https://doi.org/10.1016/j.enconman.2003.10.009
Braun, J. E., & Chaturvedi, N. (2002). An inverse gray-box model for transient building load prediction. *HVAC&R Research*, *8*(1), 73–99. https://doi.org/10.1080/10789669.2002.10391290
Catalina, T., Iordache, V., & Caracaleanu, B. (2013). Multiple regression model for fast prediction of the heating energy demand. *Energy and Buildings*, *57*, 302–312. https://doi.org/10.1016/j.enbuild.2012.11.010
Chen, W., Rong, F., & Lin, C. (2025). Short-term building electricity load forecasting with a hybrid deep learning method. *Energy and Buildings*, *330*, 115342. https://doi.org/10.1016/j.enbuild.2025.115342
Chou, J.-S., & Bui, D.-K. (2014). Modeling heating and cooling loads by artificial intelligence for energy-efficient building design. *Energy and Buildings*, *82*, 437–446. https://doi.org/10.1016/j.enbuild.2014.07.036
Das, H. P., Lin, Y.-W., Agwan, U., Spangher, L., Devonport, A., & Yang, Y. (2024). Machine learning for smart and energy-efficient buildings. *Environmental Data Science*, *3*, e1. https://doi.org/10.1017/eds.2023.43
Deb, C., Eang, L. S., Yang, J., & Santamouris, M. (2016). Forecasting diurnal cooling energy load for institutional buildings using artificial neural networks. *Energy and Buildings*, *121*, 284–297. https://doi.org/10.1016/j.enbuild.2015.12.050
Ekici, B. B., & Aksoy, U. T. (2009). Prediction of building energy consumption by using artificial neural networks. *Advances in Engineering Software*, *40*(5), 356–362. https://doi.org/10.1016/j.advengsoft.2008.05.003
Fan, C., Xiao, F., & Wang, S. (2014). Development of prediction models for next-day building energy consumption and peak power demand using data mining techniques. *Applied Energy*, *127*, 1–10. https://doi.org/10.1016/j.apenergy.2014.04.016
Ferrer, A. L. C., Thomé, A. M. T., & Scavarda, A. J. (2018). Sustainable urban infrastructure: A review. *Resources, Conservation and Recycling*, *128*, 360–372. https://doi.org/10.1016/j.resconrec.2016.07.017
Forrester, J., & Wepfer, W. (1984). Formulation of a load prediction algorithm for a large commercial building. *ASHRAE Transactions*, *90*(2), 536–551.
Foucquier, A., Robert, S., Suard, F., Stéphan, L., & Jay, A. (2013). State of the art in building modelling and energy performances prediction: A review. *Renewable and Sustainable Energy Reviews*, *23*, 272–288. https://doi.org/10.1016/j.rser.2013.03.004
Fu, Y., Li, Z., Zhang, H., & Xu, P. (2015). Using support vector machine to predict next day electricity load of public buildings with sub-metering devices. *Procedia Engineering*, *121*, 1016–1022. https://doi.org/10.1016/j.proeng.2015.09.097
Ghiaus, C. (2006). Experimental estimation of building energy performance by robust regression. *Energy and Buildings*, *38*(6), 582–587. https://doi.org/10.1016/j.enbuild.2005.08.014
Hakkal, S., & Lahcen, A. A. (2024). XGBoost To enhance learner performance prediction. *Computers and Education: Artificial Intelligence*, *7*, 100254. https://doi.org/10.1016/j.caeai.2024.100254
Haklay, M., & Weber, P. (2008). Openstreetmap: User-generated street maps. *IEEE Pervasive Computing*, *7*(4), 12–18. https://doi.org/10.1109/MPRV.2008.80

He, F., Zhou, J., Feng, Z., Liu, G., & Yang, Y. (2019). A hybrid short-term load forecasting model based on variational mode decomposition and long short-term memory networks considering relevant factors with Bayesian optimization algorithm. *Applied Energy*, *237*, 103–116. https://doi.org/10.1016/j.apenergy.2019.01.055

Hoornweg, D., Bhada, P., Freire, M., Trejos, C., & Sugar, L. (2010). Cities and climate change: An urgent agenda. The World Bank, Washington, DC. https://openknowledge.worldbank.org/server/api/core/bitstreams/3843aba9-9c34-5610-8a48-9e4b1e4d805b/content

Hua, P., Wang, H., Xie, Z., & Lahdelma, R. (2024). District heating load patterns and short-term forecasting for buildings and city level. *Energy*, *289*, 129866. https://doi.org/10.1016/j.energy.2023.129866

Ilojianya, V. I., Usman, F. O., Ibekwe, K. I., Nwokediegwu, Z. Q. S., Umoh, A. A., & Adefemi, A. (2024). Data-driven energy management: Review of practices in Canada, USA, and Africa. *Engineering Science & Technology Journal*, *5*(1), 219–230. https://doi.org/10.51594/estj.v5i1.745

Johari, F., Peronato, G., Sadeghian, P., Zhao, X., & Widén, J. (2020). Urban building energy modeling: State of the art and future prospects. *Renewable and Sustainable Energy Reviews*, *128*, 109902. https://doi.org/10.1016/j.rser.2020.109902

Kaitouni, S. I., Gargab, F.-Z., Es-sakali, N., Mghazli, M. O., El Mansouri, F., & Jamil, A. (2025). Digital workflow for nearly zero-energy high-rise office building design optimization at the district scale in Mediterranean context. *Energy and Built Environment*, *6(6)*, 1025–1038. https://doi.org/10.1016/j.enbenv.2024.04.008

Kapp, S., Choi, J.-K., & Hong, T. (2023). Predicting industrial building energy consumption with statistical and machine-learning models informed by physical system parameters. *Renewable and Sustainable Energy Reviews*, *172*, 113045. https://doi.org/10.1016/j.rser.2022.113045

Kondath, N., Myat, A., Soh, Y. L., Tung, W. L., Eugene, K. A. M., & An, H. (2024). Enhancing day-ahead cooling load prediction in tropical commercial buildings using advanced deep learning models: A case study in Singapore. *Buildings*, *14*(2), 397. https://doi.org/10.3390/buildings14020397

Kwok, S. S., Yuen, R. K., & Lee, E. W. (2011). An intelligent approach to assessing the effect of building occupancy on building cooling load prediction. *Building and Environment*, *46*(8), 1681–1690. https://doi.org/10.1016/j.buildenv.2011.02.008

Lachut, D., Banerjee, N., & Rollins, S. (2014). Predictability of energy use in homes. *In: International Green Computing Conference. IEEE*, 1–10. https://doi.org/10.1109/IGCC.2014.7039146

Le, L. T., Nguyen, H., Dou, J., & Zhou, J. (2019). A comparative study of PSO-ANN, GA-ANN, ICA-ANN, and ABC-ANN in estimating the heating load of buildings' energy efficiency for smart city planning. *Applied Sciences*, *9*(13), 2630. https://doi.org/10.3390/app9132630

Li, D., Hu, G., & Spanos, C. J. (2016). A data-driven strategy for detection and diagnosis of building chiller faults using linear discriminant analysis. *Energy and Buildings*, *128*, 519–529. https://doi.org/10.1016/j.enbuild.2016.07.014

Li, J., Liang, C., & Zhou, W. (2024). A review of building physical shapes on heating and cooling energy consumption. *Energies*, *17*(22), 5766. https://doi.org/10.3390/en17225766

Li, Q., Meng, Q., Cai, J., Yoshino, H., & Mochida, A. (2009). Predicting hourly cooling load in the building: A comparison of support vector machine and different artificial neural networks. *Energy Conversion and Management*, *50*(1), 90–96. https://doi.org/10.1016/j.enconman.2008.08.033

Li, X., & Gou, Z. (2025). Urban morphology and energy self-sufficiency: A comparative study of residential blocks in eight global cities. *Cities*, *166*, 106240. https://doi.org/10.1016/j.cities.2025.106240

Li, X., & Wen, J. (2014). Review of building energy modeling for control and operation. *Renewable and Sustainable Energy Reviews*, *37*, 517–537. https://doi.org/10.1016/j.rser.2014.05.056

Low, R., Tekler, Z. D., & Cheah, L. (2021). An end-to-end point of interest (POI) conflation framework. *ISPRS International Journal of Geo-Information*, *10*(11), 779. https://doi.org/10.3390/ijgi10110779

Lundberg, S. M., & Lee, S.-I. (2017). A Unified Approach to Interpreting Model Predictions, Advances in Neural Information Processing Systems, 30. https://proceedings.neurips.cc/paper_files/paper/2017/file/8a20a8621978632d76c43dfd28b67767-Paper.pdf

Manic, M., Wijayasekara, D., Amarasinghe, K., & Rodriguez-Andina, J. J. (2016). Building energy management systems: The age of intelligent and adaptive buildings. *IEEE Industrial Electronics Magazine*, *10*(1), 25–39. https://doi.org/10.1109/MIE.2015.2513749

Mehdizadeh Khorrami, B., Soleimani, A., Pinnarelli, A., Brusco, G., & Vizza, P. (2024). Forecasting heating and cooling loads in residential buildings using machine learning: A comparative study of techniques and influential indicators. *Asian Journal of Civil Engineering*, *25*(2), 1163–1177. https://doi.org/10.1007/s42107-023-00834-8

Moazzami, M., Khodabakhshian, A., & Hooshmand, R. (2013). A new hybrid day-ahead peak load forecasting method for Iran's National Grid. *Applied Energy*, *101*, 489–501. https://doi.org/10.1016/j.apenergy.2012.06.009

Ngo, N.-T. (2019). Early predicting cooling loads for energy-efficient design in office buildings by machine learning. *Energy and Buildings*, *182*, 264–273. https://doi.org/10.1016/j.enbuild.2018.10.004

Olu-Ajayi, R., Alaka, H., Sulaimon, I., Sunmola, F., & Ajayi, S. (2022). Building energy consumption prediction for residential buildings using deep learning and other machine learning techniques. *Journal of Building Engineering*, *45*, 103406. https://doi.org/10.1016/j.jobe.2021.103406

Peng, Y., Lei, Y., Tekler, Z. D., Antanuri, N., Lau, S. K., & Chong, A. (2022). Hybrid system controls of natural ventilation and HVAC in mixed-mode buildings: A comprehensive review. *Energy and Buildings*, *276*, 112509. https://doi.org/10.1016/j.enbuild.2022.112509

Pino-Mejías, R., Pérez-Fargallo, A., Rubio-Bellido, C., & Pulido-Arcas, J. A. (2017). Comparison of linear regression and artificial neural networks models to predict heating and cooling energy demand, energy consumption and CO 2 emissions. *Energy*, *118*, 24–36. https://doi.org/10.1016/j.energy.2016.12.022

Reinhart, C. F., & Davila, C. C. (2016). Urban building energy modeling – a review of a nascent field. *Building and Environment*, *97*, 196–202. https://doi.org/10.1016/j.buildenv.2015.12.001

Sadaghat, B., Afzal, S., & Khiavi, A. J. (2024). Residential building energy consumption estimation: A novel ensemble and hybrid machine learning approach. *Expert Systems with Applications*, *251*, 123934. https://doi.org/10.1016/j.eswa.2024.123934

Salimi, M., & Al-Ghamdi, S.-G. (2020). Climate change impacts on critical urban infrastructure and urban resiliency strategies for the Middle East. *Sustainable Cities and Society*, *54*, 101948. https://doi.org/10.1016/j.scs.2019.101948

Seo, B., Yoon, Y., Lee, K. H., & Cho, S. (2023). Comparative analysis of ANN and LSTM prediction accuracy and cooling energy savings through AHU-DAT control in an office building. *Buildings*, *13*(6), 1434. https://doi.org/10.3390/buildings13061434

Shahidehpour, M., Li, Z., & Ganji, M. (2018). Smart cities for a sustainable urbanization: Illuminating the need for establishing smart urban infrastructures. *IEEE Electrification Magazine*, *6*(2), 16–33. https://doi.org/10.1109/MELE.2018.2816840

Spethmann, D. (1989). Optimal control for cool storage. *ASHRAE Transactions*, *95*, 1189–1193.

Stasi, R., Ruggiero, F., & Berardi, U. (2024). Influence of cross-ventilation cooling potential on thermal comfort in high-rise buildings in a hot and humid climate. *Building and Environment*, *248*, 111096. https://doi.org/10.1016/j.buildenv.2023.111096

Tan, L., Tang, T., & Yuan, D. (2022). An ensemble learning aided computer vision method with advanced color enhancement for corroded bolt detection in tunnels. *Sensors (Basel, Switzerland)*, *22*(24), 9715. https://doi.org/10.3390/s22249715

Tehrani, A. A., Sobhaninia, S., Nikookar, N., Levinson, R., Sailor, D. J., & Amaripadath, D. (2025). Data-driven approach to estimate urban heat island impacts on building energy consumption. *Energy*, *316*, 134508. https://doi.org/10.1016/j.energy.2025.134508

Tehrani, A. A., Veisi, O., Delavar, Y., Bahrami, S., & Mehan, A. (2024b). Predicting urban heat island in European cities: A comparative study of GRU, DNN, and ANN models using urban morphological variables. *Urban Climate*, *56*, 102061. https://doi.org/10.1016/j.uclim.2024.102061

Tehrani, A. A., Veisi, O., Fakhr, B. V., & Du, D. K. (2024a). Predicting solar radiation in the urban area: A data-driven analysis for sustainable city planning using artificial neural networking. *Sustainable Cities and Society*, *100*, 105042. https://doi.org/10.1016/j.scs.2023.105042

Tekler, Z. D., Lei, Y., & Chong, A. (2024). Data-efficient comfort modeling: Active transfer learning for predicting personal thermal comfort using limited data. *Energy and Buildings*, *319*, 114507. https://doi.org/10.1016/j.enbuild.2024.114507

Tekler, Z. D., Low, R., & Blessing, L. (2022b). User perceptions on the adoption of smart energy management systems in the workplace: Design and policy implications. *Energy Research & Social Science*, *88*, 102505. https://doi.org/10.1016/j.erss.2022.102505

Tekler, Z. D., Low, R., Yuen, C., & Blessing, L. (2022a). Plug-Mate: An IoT-based occupancy-driven plug load management system in smart buildings. *Building and Environment*, *223*, 109472. https://doi.org/10.1016/j.buildenv.2022.109472

Toutou, A., Fikry, M., & Mohamed, W. (2018). The parametric based optimization framework daylighting and energy performance in residential buildings in hot arid zone. *Alexandria Engineering Journal*, *57*(4), 3595–3608. https://doi.org/10.1016/j.aej.2018.04.006

Veisi, O., Tehrani, A. A., Gharaei, B., Du, D. K., & Shakibamanesh, A. (2025). Towards universal thermal climate index prediction via machine learning approaches. *Renewable and Sustainable Energy Reviews*, *217*, 115680. https://doi.org/10.1016/j.rser.2025.115680

Wang, L., Kubichek, R., & Zhou, X. (2018). Adaptive learning based data-driven models for predicting hourly building energy use. *Energy and Buildings*, *159*, 454–461. https://doi.org/10.1016/j.enbuild.2017.10.054

Wang, Z., & Srinivasan, R. S. (2017). A review of artificial intelligence based building energy use prediction: Contrasting the capabilities of single and ensemble prediction models. *Renewable and Sustainable Energy Reviews*, *75*, 796–808. https://doi.org/10.1016/j.rser.2016.10.079

Wei, Y., Zhang, X., Shi, Y., Xia, L., Pan, S., & Wu, J. (2018). A review of data-driven approaches for prediction and classification of building energy consumption. *Renewable and Sustainable Energy Reviews*, *82*, 1027–1047. https://doi.org/10.1016/j.rser.2017.09.108

Xiao, L., Shao, W., Wang, C., Zhang, K., & Lu, H. (2016). Research and application of a hybrid model based on multi-objective optimization for electrical load forecasting. *Applied Energy*, *180*, 213–233. https://doi.org/10.1016/j.apenergy.2016.07.113

Xu, X., Dong, Q., & Zhen, M. (2024). Assessing effects of future climate change on outdoor thermal stress and building energy performance in severe cold region of China. *Building and Environment*, *251*, 111236. https://doi.org/10.1016/j.buildenv.2024.111236

Xu, Y., Li, F., & Asgari, A. (2022). Prediction and optimization of heating and cooling loads in a residential building based on multi-layer perceptron neural network and different optimization algorithms. *Energy*, *240*, 122692. https://doi.org/10.1016/j.energy.2021.122692

Zhang, J., & Haghighat, F. (2010). Development of artificial neural network based heat convection algorithm for thermal simulation of large rectangular cross-sectional area earth-to-air heat exchangers. *Energy and Buildings*, *42*(4), 435–440. https://doi.org/10.1016/j.enbuild.2009.10.011

Zhang, L., Wen, J., Li, Y., Chen, J., Ye, Y., & Fu, Y. (2021). A review of machine learning in building load prediction. *Applied Energy*, *285*, 116452. https://doi.org/10.1016/j.apenergy.2021.116452

Zhao, A., Zhang, M., Quan, W., & Sun, W. (2025). A hybrid prediction model for heating load of buildings within residential communities considering occupancy rates. *Energy and Buildings*, *329*, 115220. https://doi.org/10.1016/j.enbuild.2024.115220

Zhou, L., & Haghighat, F. (2009). Optimization of ventilation systems in office environment, part II: Results and discussions. *Building and Environment*, *44*(4), 657–665. https://doi.org/10.1016/j.buildenv.2008.05.010

Zhou, Y., Wu, J., Wang, R., Shiochi, S., & Li, Y. (2008). Simulation and experimental validation of the variable-refrigerant-volume (VRV) air-conditioning system in EnergyPlus. *Energy and Buildings*, *40*(6), 1041–1047. https://doi.org/10.1016/j.enbuild.2007.04.025

Zhou, Y., Zheng, S., & Hensen, J. L. M. (2024). Machine learning-based digital district heating/cooling with renewable integrations and advanced low-carbon transition. *Renewable and Sustainable Energy Reviews*, *199*, 114466. https://doi.org/10.1016/j.rser.2024.114466

Zhu, H., & Lin, B. (2024). Digital twin-driven energy consumption management of integrated heat pipe cooling system for a data center. *Applied Energy*, *373*, 123840. https://doi.org/10.1016/j.apenergy.2024.123840

Zhuang, D., Gan, V. J. L., Tekler, Z. D., Chong, A., Tian, S., & Shi, X. (2023). Data-driven predictive control for smart HVAC system in IoT-integrated buildings with time-series forecasting and reinforcement learning. *Applied Energy*, *338*, 120936. https://doi.org/10.1016/j.apenergy.2023.120936