# Room energy demand and thermal comfort predictions in early stages of design based on the Machine Learning methods

Forouzandeh Shahraki, N.; Zomorodian, Zahra Sadat; Tahsildoost, Mohammad; Shaghaghian, Zohreh

**DOI**
10.1080/17508975.2022.2049190

**Publication date**
2022

**Document Version**
Final published version

**Published in**
Intelligent Buildings International

Intelligent Buildings International

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/tibi20

# Room energy demand and thermal comfort predictions in early stages of design based on the Machine Learning methods

Nima Forouzandeh, Zahra Sadat Zomorodian, Zohreh Shaghaghian & Mohamad Tahsildoost

INTELLIGENT BUILDINGS INTERNATIONAL
https://doi.org/10.1080/17508975.2022.2049190

Taylor & Francis
Taylor & Francis Group

Check for updates

RESEARCH ARTICLE

# Room energy demand and thermal comfort predictions in early stages of design based on the Machine Learning methods

Nima Forouzandeh [a], Zahra Sadat Zomorodian [a], Zohreh Shaghaghian[b] and Mohamad Tahsildoost[a]

[a]Department of Construction, Shahid Beheshti University(SBU), Tehran, Iran; [b]Department of Architecture, Texas A&M, College Station, TX, USA

**ABSTRACT**
Recent studies have focused on data-driven methods for building energy efficiency, by using simulated or empirical data, for energy-based design assessment rather than the common physics-based techniques, which are mostly time-consuming. In this paper, the feasibility of using seven different Machine Learning models, including three single models and four ensemble ones, is studied to predict annual energy demand and thermal comfort of the model. For this purpose, 3024 synthetic samples of a single zone model with seven input features are simulated through the EnergyPlus engine for training in addition to 360 unseen samples as testing data for accuracy reporting. Heating and cooling demands, in addition to five annual thermal comfort indices, are calculated for each data point and used as target indices. Results show Extremely Randomized Trees and Random Forest models had the highest $R^2$ of 0.99 and 0.85 for cooling and heating demands respectively. Also, the $R^2$ of these models for predicting annual comfort was between 0.71 and 0.95. Results are then used to develop a prediction framework of thermal comfort and energy demand performance in the early stages of building design, where most of the information about building characteristics is not yet known.

**ARTICLE HISTORY**
Received 21 May 2021
Accepted 27 February 2022

**KEYWORDS**
Artificial intelligence; Machine Learning; Tehran; energy efficiency; energy simulation; data-driven model; early design phase; thermal comfort

## Introduction

Much effort has been focused on reducing energy consumption in buildings via retrofitting existing buildings or energy-efficient design methods in the pre-construction phase. There are two main approaches for building thermal calculations, white-box and black-box. White-box models use physics-based principles for such calculations. Different tools and simulation platforms are in this category, amongst which EnergyPlus, DOE-2, and TRNSYS are some of the most commonly used engines in building industry (Beckman et al. 1994; Crawley et al. 2001).

Black-box approaches use curve-fitting techniques to extract relationships between design variables and building performance indices. Such data-driven methods, including ML techniques, can be implemented for long-term energy and performance predictions (Amasyali and El-Gohary 2018). In the literature, ML methods have been implemented for the prediction of different building energy conservation indices such as energy demand in existing buildings (Amasyali and El-Gohary 2018), setpoint management (Brandi et al. 2020), HVAC system optimization or fault diagnosis (Han et al. 2020), and peak load prediction (Dietrich et al. 2020).

ML can also be used to evaluate the performance of building design alternatives with small sets of assumptions quickly compared to the standard building energy modeling tools to make informed decisions, hence, becoming more practical for non-professional users (Ciulla and D'Amico 2019).

**CONTACT** Zahra Zomorodian z_zomorodian@sbu.ac.ir

These methods are often reliable and relatively fast with a good approximation if models are trained efficiently. However, implementing ML approach can be tough in the design stage as the relation between inputs and target variables is not explicit and preparing the training data is a challenge (Wang and Srinivasan 2017).

## Aim and scope

Considering the drawbacks of white-box approach, and gaps in the existing studies implementing the black-box approach, this study aims to use data-driven ML techniques to predict long-term energy demand and thermal comfort for integrated design in the early stages. The outcome of this study provides an estimation tool for evaluating design alternatives at zone level, which could be utilized by architects without specific knowledge about the energy performance of buildings, and designers/building owners with minimum effort.

Towards these goals, the current study aims to find the most suitable ML models for energy demand and annual thermal comfort target indices by a comparative analysis between different ML approaches. Therefore, besides the common methods in literature (ANN, MR, SVM) the RF, Boosting, and ERT are utilized. Moreover, results are compared to the common physics-based (EnergyPlus) models, in terms of accuracy and calculation speeds. Results are used to develop an algorithmic framework for performance prediction in early design phases.

The content of this paper is organized as follows. The first section contains the review of existing literature. Next, the study's methodology is presented. The construction and optimization of ML models are explained afterwards. The results and accuracy of the models, and proposed framework, are then described and discussed in the next part. Lastly, conclusions and suggestions for future research are elucidated.

List of abbreviations.

| | |
|---|---|
| AB | Adaboost |
| ANN | Artificial Neural Networks |
| APE | Average Percentage Error |
| BPD | Building Performance Database |
| BR | Bagging Regressor |
| DhC | Degree-hours Criterion |
| DOE | Department of Energy |
| DT | Decision tree |
| EnerPro | Energy Profiling Tool |
| ERT | Extremely Randomized Trees |
| HVAC | Heating, Ventilation and Air Conditioning |
| KNN | K-Nearest Neighbors |
| LSTM | Long Short-Term Memory |
| MAE | Mean Absolute Error |
| MAPE | Mean Absolute Percentage Error |
| MeAE | Median Absolute Error |
| MI | Mutual Information |
| ML | Machine Learning |
| MR | Multi-Variate Regression |
| MRE | Mean Relative Error |
| MRPE | Mean Relative Percentage Error |
| NMAE | Normalized MAE |
| NRMSE | Normalized RMSE |
| PMV | Predicted Mean Vote |
| POR | Percentage Outside the Range |
| R | Linear Correlation Coefficient |
| $R^2$ | Coefficient of Determination |
| RF | Random Forest |
| RMSD | Root Mean Square Deviation |
| RMSE | Root Mean-Square Error |
| RSS | Residual Sum of Squares |
| SEE | Standard Error Of Estimate |
| SI | Synthesis Index |
| SSE | Total Square Error |
| TSS | Total Sum of Squares |
| SVM | Support Vector Machine |
| WWR | Window-to-Wall Ratio |

## Literature review

There exists a considerable body of literature on AI application in energy consumption prediction of buildings. Seyedzadeh and Rahimian have systematically presented a set of key topics pertaining the application of AI techniques for optimization of energy-related performance of non-domestic, large-scale buildings, and how such solutions can address modern problems (Seyedzadeh and Rahimian 2021). In another related attempt, Magoulès and Zhao explain data mining and Machine Learning techniques for solving prediction, analysis, or fault detection and diagnosis of building energy consumption (2016).

To find similar studies based on research objectives, a combination of related phrases including 'machine learning', 'artificial intelligence', 'building', 'energy', 'thermal', and 'comfort' were used to conduct a search in GoogleScholar and ScienceDirect. A total number of 67 papers were collected. Studies prior to 2010 were excluded from the review. Also, studies on systems level and urban level were excluded. All the studies with empirical or pre-simulated training datasets with heating, cooling, and comfort labels were included in this literature review. Studies implementing both single and ensemble ML methods are reviewed. The final list of 26 relevant papers is presented in Table 1.

In addition, ML model features, including the target variables and indices, analysis level (zone, building, etc.), as well as the architecture of the model in terms of their utilized ML algorithm, input features, type of dataset (real, simulated), and finally, their accuracy are presented.

The majority of the reviewed studies are concerned with heating and cooling energy demand predictions as target variables, while the feasibility of temperature and thermal comfort predictions are not studied except for the work done by Gelder et al. that predicts over-heating throughout the year (Van Gelder et al. 2014).

According to these studies, input features to predict target variables mainly include six groups of information:(1) Location/weather conditions, (2) Geometry/massing (i.e. perception of the general shape and form and size of a building)/ orientation, (3) Envelope properties (walls and windows U-values, WWR, shadings and blinds, airtightness, windows free apertures and vents, and glazing types), (4) HVAC system information (mainly setpoints), (5) Schedules, occupancy, and (6) Building type. The second and the third feature groups have been used more often.

The type of ML algorithms used for predictions can be divided into two main groups, single models and ensemble models. The former group predicts the target values based on the results of one ML model, while the latter method conducts a prediction based on the average accuracy of multiple models. ANN, MR, and SVM are the most commonly used single models in the literature (Walter and Sohn 2016), while implementing RF, Boosting, Bagging, and ERT are more common according to the literature review (Table 1).

Two types of datasets are used to train and test ML models, pre-simulated and empirical. For the generation of pre-simulated datasets, most studies used numerical energy simulation programs, e.g. EnergyPlus, TRACE 700, DOE-2, Ecotect, IES-VE, and TRNSYS, providing more flexibility upon the ranges of parameters and sample alternatives (Ngo 2019). Other groups of studies utilize real data mostly measured and collected by governmental authorities, supporting a more realistic estimation of building performance (Kontokosta and Tull 2017; Robinson et al. 2017).

Also, according to Table 1, $R^2$, RMSE, and MAE are the most commonly used accuracy/error metrics in literature with an accuracy of 0.4–1.00 for heating and cooling demand (Walter and Sohn 2016). It should be noted that the accuracy of results depends on validation method, training and testing datesets, and model architectures.

Overall, the literature lacks a comprehensive knowledge on feasibility of using ML algorithms in early stages of building design to predict building thermal comfort performance (Carlucci and Pagliano 2012; Geyer and Singaravel 2018; Sundaravelpandian Singaravel and Geyer 2016).

## Methodology

This section summarizes the dataset, data-driven statistical concepts, and the ML techniques used to predict model performance. Schematic outline of building energy demand and thermal comfort prediction method is presented in Figure 1.

### *Description of train and test datasets*

For this study, two separate datasets with sizes of 3024 and 360 and different sets of assumptions are generated parametrically as training and test data respectively in grasshopper. This ratio of testing and training

**Table 1.** Related research.

| | | | | ML algorithms | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | Single | | | | | | | Ensemble | | | | | | | | | | |
| Ref. | Indices | Resolution | level | ANNs | SVM | KNN | DT (CART/ CHAID) | GLR | multi-variate Regression | LR | Voting | Boosting | Bagging | Stacking | RF | ERT | Other | Data type | Validation strategy | Accuracy metrics | Accuracy / best model accuracy |
| Ngo (2019) | C, H | A | B | ✓ | ✓ | | ✓ | | | ✓ | ✓ | | ✓ | ✓ | | | | S | CV | MAE, RMSE, MAPE, R, SI | 0.78-0.99(R) |
| Chou and Bui (2014) | C, H | A | B | ✓ | ✓ | | ✓ | ✓ | | | | | | | | | | S | CV | RMSE, MAE, MAPE, R, SI | 2.98-4.96 (MAPE) |
| Tsanas and Xifara (2012) | C, H | A | Z, B | | | | | | | | | | | | ✓ | | ✓ | S | CV | MAE, MSE, MRE | 0.11-2.21(MAE) |
| Geyer and Singaravel (2018) Sundaravelpandian Singaravel and Geyer (2016) | C, H | A | B | ✓ | | | | | | | | | | | | | | S | Uns. | $R^2$ | 0.69-0.99($R^2$) |
| Kontokosta and Tull (2017) | C, H | A | U | | ✓ | | | | | | | | | | ✓ | | | R | Uns. | MAE, log accuracy ratio(LAR) | 0.17(MAE) |
| Li and Yao (2020) | C, H | A | B | | ✓ | | | | | | | | | | | | | S | Uns. | MAE, RMSE, NMAE, NRMSE | 0.61, 0.66(MAE) |
| Kumar, Pal, and Singh (2018) | C, H | A | B | ✓ | | | | | | | | | | | | | | R | – | MAE | 0.027,0.032 (MAE) |
| Walter and Sohn (2016) | E | A | B | | | | | | ✓ | | | | | | | | | R | – | $R^2$ | 0.4($R^2$) |
| Cheng and Cao (2014) | C, H | A | B | | | | | | | | | | | | | | ✓ | S | CV | RMSE, MAPE, MAE, $R^2$ | 0.99-1.00 |
| Turhan et al. (2014) | H | A | B | ✓ | | | | | | | | | | | | | | S | – | MSE, $R^2$, MAPE | 0.97($R^2$) |
| Van Gelder et al. (2014) | TE, H | A | B | ✓ | | | | | ✓ | | | | | | | | ✓ | S | Uns. | MSE, $R^2$, MAE | 0.91-1.00($R^2$) |
| Wei et al. (2015) | G, EI | A | U | | ✓ | | | | | | | ✓ | ✓ | | | | | R | CV | RMSE, $R^2$ | Gas:710, Electricity:280 (RMSE) |
| Sundaravelpandian Singaravel et al. (2017) | C, H | A | Z | ✓ | | ✓ | | | | | | | | | ✓ | ✓ | | S | CV | $R^2$ | 0.97-0.99($R^2$) |
| Pan and Zhang (2020) | E | A | B | | | | | | | | | ✓ | | | | | | R | CV | $R^2$ | 0.897($R^2$) |
| Sundaravelpandian Singaravel, Suykens, and Geyer (2019) | | A | B | ✓ | | | | | | | | | | | | | | S | CV | $R^2$, MAPE | -14%- 10% (MAPE) |

| | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Sundaravelpandian Singaravel, Suykens, and Geyer (2018) | C, H | M | B | ✓ | | | | | | | | | S | Uns. | $R^2$ | 0.96-0.99($R^2$) |
| Catalina, Virgone, and Blanco (2008) | H | M | B | | | | | | | | | | S | Uns. | SEE, $R^2$ | 2%(SSE) |
| Naji et al. (2016) | C, H | A | B | ✓ | | | | | | | | | S | Uns. | $R^2$ | 0.99($R^2$) |
| Robinson et al. (2017) | E | A | B | ✓ | ✓ | ✓ | | ✓ | ✓ | ✓ | ✓ | | R | CV | MAE, MeAE, $R^2$ | 0.82($R^2$) |
| Romani, Draoui, and Allard (2015) | C, H | A | B | | | | ✓ | | | | | | S | Uns. | MAE, RMSD, $R^2$ | Heating> 0.96, Cooling>0.91 ($R^2$) |
| Mottahedi et al. (2015) | C, H | A | B | | | | | | | | | | S | Uns. | $R^2$, F-test, RMSE | 0.95-1.00 |
| Hygh et al. (2012) | C, H | A | B | | | | ✓ | | | | | | S | Uns. | RMSE, $R^2$, APE | $R^2$>0.96 |
| Ekici and Teoman Aksoy (2009) | E | A | B | ✓ | | | | | | | | | S | – | MSE, SSE | 1.5-5.2%(SSE) |
| Bektas Ekici and Teoman Aksoy (2011) | C, H | A | B | | | | | | | | | ✓ | S | Uns. | $R^2$ | Heating:0.96, Cooling:0.83 ($R^2$) |
| Amiri, Mottahedi, and Asadi (2015) | C, H | A | B | | | | ✓ | | | | | | S | Uns. | RSS, $R^2$, RMSE | 0.95-0.98($R^2$) |

B: Building; Z: Zone; U: Urban and district scale; C: Cooling load; H: Heating load; E: total energy use; TE: temperatures exceeding 25 C; G: Gas; El: Electricity; A: Annual; M: Monthly Methods for data generation; R: Real data; S: Simulated data.

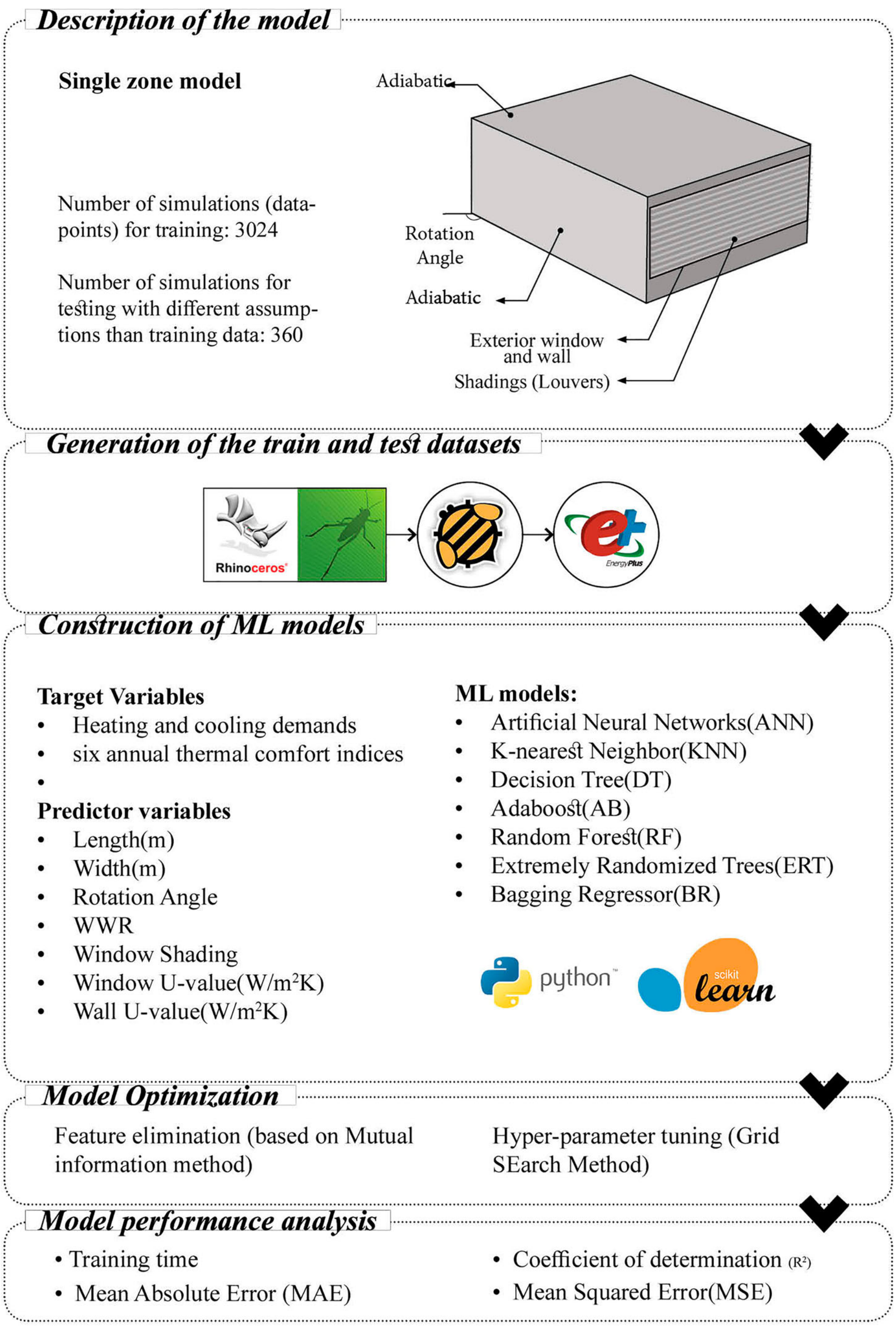


**Figure 1.** Outline for development of predictive models.

data is determined based on the literature review Table 1. Corresponding energy demand of each sample is calculated through Honeybee which uses EnergyPlus engine (Crawley et al. 2000) using Tehran annual weather file. The case study is a single office zone located in Tehran climate (Bsk, based on Köppen classification) with one exterior wall; other surfaces are considered adiabatic. This space is supposed to be representative of a common room in an office building with adjacency of similar spaces, thus only one surface is assumed not adiabatic. Key simulation assumptions are presented in Table 2. These assumptions are based on Iran's national building energy code (Iran National Building Code 2019).

Six independent variables including room length, width, rotation angle, WWR, window shading, window U-value, and wall U-value, are used as features of data samples with different values, assuming their importance in building energy consumption based on sensitivity analysis studies (Table 3). The second and the third column shows the number of steps and values for testing dataset with 3024 data points. The fourth and the fifth columns show the information for unseen testing data with 360 data points. values for each parameter are chosen based on common materials and options and relevant studies in Table 1. These ranges for parameters can be extended for more cases in future research.

## Description of target indices

Cooling and heating energy demands were calculated for each of the training and test cases as target variables besides five long-term annual thermal comfort indices selected from the literature (Carlucci and Pagliano 2012). Thermal comfort indices can be categorized into two main sections based on their calculation method introduced by ISO 7730-2005 and EN 16798-1 (BS EN 16798-1, 2019). Although EN standard is for European countries it is used because of the similar climatic context in Tehran (Standard, I. S. O. 2005).

### POR

According to Equation (1), this index is determined by the summation of the annual binary states of comfort for each occupancy hour. This method can be applied to both Fanger's model by calculating the PMV index (Equation (2)) and the Adaptive model by calculating the operative temperature for each occupancy hour (Equation (3)). A total number of five thermal comfort indices in this group are calculated for each simulation case explained in Table 3.

$$\mathrm{POR} = \frac{\sum_{i=1}^{Oh} (wf_i \cdot h_i)}{\sum_{i=1}^{Oh} h_i} \in [0; 1] \tag{1}$$

where POR is Percentage outside the range, $wf$ is Weighting Factor, $Oh$ is the number of operational hours, and $h$ is the time step (i.e. one hour).

$$\mathrm{POR}_{\mathrm{Fanger,\ PMV}} = f(wf_i): \begin{cases} wf_i = 1 \leftarrow (\mathrm{PMV} < \mathrm{PMV}_{\text{lower limit}}) \vee (\mathrm{PMV} > \mathrm{PMV}_{\text{upper limit}}) \\ wf_i = 0 \leftarrow (\mathrm{PMV}_{\text{lower limit}} \leq \mathrm{PMV} \leq \mathrm{PMV}_{\text{lower limit}}) \end{cases} \tag{2}$$

where $\mathrm{POR}_{\mathrm{Fanger,\ PMV}}$ is Percentage outside the range using the Fanger model and PMV values, $wf$ is Weighting Factor, Oh is the number of operational hours.

$$POR_{\mathrm{Adaptive}} = f(wf_i): \begin{cases} wf_i = 1 \leftarrow (\theta_{op,\,in} < \theta_{op,\,\text{lower limit}}) \vee (\theta_{op,\,in} > \theta_{op,\,\text{upper limit}}) \\ wf_i = 0 \leftarrow (\theta_{op,\,\text{lower limit}} \leq \theta_{op,\,in} \leq \theta_{op,\,\text{upper limit}}) \end{cases} \tag{3}$$

**Table 2.** Fixed simulation key assumptions and a summary of average Tehran climatic conditions in working hours (8–17).

| | | | | |
|---|---|---|---|---|
| Cooling setpoint(°C) | 25.00 | | | |
| Heating setpoint(°C) | 20 | | | |
| Cooling setback(°C) | 20 | | | |
| Heating setback(°C) | 0 | | | |
| Lighting target value(lux) | 300 | | | |
| Equipment load(w/m$^2$) | 14 | | | |
| Electric lighting power density(w/m$^2$) | 10.5 | | | |
| Occupants density (p/m$^2$) | 0.11 | | | |
| Occupancy | 8:00-17:00 | | | |
| Ventilation rate(m$^3$/s.m$^2$) | 3.05E-04 | | | |
| Infiltration load(w/m$^2$) | 3.00E-04 | | | |
| | Spring | Summer | Fall | Winter |
| Average DBT(°C) | 20.8 | 29.8 | 16.38 | 7.31 |
| Average RH(%) | 30.2 | 25.7 | 42.5 | 47.5 |
| Average Wind Speed(m/s) | 3.8 | 3 | 2.3 | 3.1 |

**Table 3.** Independent Variables utilized as parameters and their corresponding values and ranges to generate the training datasets.

| Variables | Number of steps (training) | Values/ ranges(training) | Number of steps (testing) | Values/ranges(testing) |
|---|---|---|---|---|
| Length(m) | 3 | 4.0, 7.0, 10.0 | 3 | 3.5, 6.0, 7.0 |
| Width(m) | | 3.0, 6.0, 8.0 | | 4.5, 5.0, 9.0 |
| Rotation Angle | 8 | 0°-315° with 45° steps | 5 | 22.5, 137.5, 275, 200, 300 |
| WWR(%) | 7 | 20–80 with 10 steps | 2 | 25, 65 |
| Window shading | 2 | With/without shading (louvers-15cm depth) | 2 | With/without shading (louvers-12cm depth) |
| Window U-value ($W/m^2K$) | 3 | 1, 7, 2.6, 3.4 | 2 | 2, 3 |
| Wall U-value($W/m^2K$) | 3 | 0.5, 0.7, 1.2 | 2 | 1.1, 1.67 |

where $POR_{Adaptive}$ is Percentage outside the range using the Adaptive thermal comfort model, *wf* is Weighting Factor, Oh is the number of operational hours, $\theta_{op,\,in}$ is indoor operative temperature.

### *DhC*

The absolute difference from upper and lower limits of the comfort temperature range is calculated for overheat and under-heat occupancy hours respectively (Equation (4)) and is multiplied by the number of hours. This method can be applied to both the Fanger's and Adaptive models as well. In this study, the Class II of EN comfort range is used based on the adaptive model as this class is proposed for office spaces. Equations (4) and (5) show the calculation steps for these indices. Also, the descriptions for the two thermal comfort models are presented in Table 4.

$$wf_i \begin{matrix} EN\ 16798-1 \\ \text{Adaptive} \end{matrix} \equiv |\theta_{op,i} - \theta_{op,\text{limit}}| \tag{4}$$

where $Wf_i$ is the weighting factor *according to EN standard and* $\theta_{op,\,in}$ is indoor operative temperature.

$$DhC = \sum_{i=1}^{oh} (wf_i \cdot h_i) \in [0;\ +\infty] \tag{5}$$

where DhC is Degree-hour criterion, Oh is the number of operational hours, and *h* is the time step (i.e. one hour).

## *ML models used in this study*

Although DT and KNN have shown good accuracies in similar studies they are not being widely used like ANNs (Singaravel 2020). However, in this study, both the common and less implemented methods are used in a single framework to compare their results on the prediction performance and accuracy of these algorithms on our test dataset. These models include both single models (ANN, KNN, DT) and ensemble ones (RF, BR, AB, ERT) which are developed with the decision tree models.

**Table 4.** Descriptions of the thermal comfort indices.

| Thermal Comfort Index | Description |
|---|---|
| $POR_{Fanger,80\%}$ | Percentage of occupancy hours with PMV value outside the 80% acceptability range |
| $POR_{Adaptive\,(ASHRAE),80\%}$ | Percentage of occupancy hours with the operative temperature outside the 80% acceptability range, according to ASHRAE 55 standard. |
| $POR_{Adaptive\,(ASHRAE),90\%}$ | Percentage of occupancy hours with the operative temperature outside the 90% acceptability range, according to ASHRAE 55 standard. |
| $POR_{Adaptive(EN),Class\ II}$ | Percentage of occupancy hours with the operative temperature outside the Class II of EN 16798-1 standard range. |
| $DhC_{overheat}$ | Degree-hours Criterion for overheat times, in the occupancy hours throughout a year, according to class II EN 16798-1 comfort range |

### *Development and optimization of the ML models*

Base models with typical hyperparameters are developed as a benchmark. To optimize the models, feature selection and hyperparameter tuning are conducted. Finally, the study framework is developed and the corresponding results are reported.

- *Feature Selection*: Feature selection is commonly used in ML methods to remove features with irrelevant or low correlation and has a positive impact on models' readability, accuracy, and training time. Models in this study don't suffer from high numbers of input features, however feature selection is conducted to identify significant features and increase models' accuracy.

MI index (Barraza et al. 2019) is used in this study to calculate the importance of each feature. $|U_i|$ is the size of the sample in $U_i$, and $|V_j|$ is the size of samples in $V_j$ (scikit-learn API Reference 2020).

$$\mathrm{MI}(U, V) = \sum_{i=1}^{|U|}\sum_{j=1}^{|V|}\frac{|U_i \cap V_j|}{N}\log\frac{N|U_i \cap V_j|}{|U_i||V_j|} \quad (6)$$

where $|U_i|$ is the number of the samples in cluster $U_i$ and $|V_j|$ is the number of the samples in cluster $V_j$.

MI values for each of the features are presented in Figure 2.

In Figure 2, MI scores, which reveal the importance of each feature(columns) in the calculation of each target variable(rows) are presented. MI values for each input feature with each output label are presented in Figure 2. Feature elimination is done in 3 steps:

(1) elimination of one feature with the least MI,
(2) elimination of two features with the least MI,
(3) elimination of three features with the least MI.

The average $R^2$ values are presented in Figure 3. The horizontal axis on the right side of the plot shows the $R^2$ values, and the horizontal axis on the left side shows the training time. The color gradient reveals the feature selection steps in sequence. According to these results, the accuracies of the models haven't improved through eliminating features, except for the DT models which performed better through the

| | Length | Width | Rotation Angle | WWR | Shading Depth | Wall U | Window U |
|---|---|---|---|---|---|---|---|
| Cooling Demand (kWh/m2) | 0.355 | 0.355 | 0.257 | 0.562 | 0.083 | 0 | 0 |
| Heating Demand (kWh/m2) | 0.277 | 0.277 | 0.169 | 0 | 0.017 | 0.029 | 0.112 |
| PORFanger, 80% | 0.059 | 0.059 | 0.658 | 0 | 0.099 | 0.026 | 0.108 |
| PORAdaptive(ASHRAE), 80% | 0.012 | 0.014 | 0.698 | 0.07 | 0.11 | 0 | 0.064 |
| PORAdaptive(ASHRAE), 90% | 0.025 | 0.024 | 0.645 | 0.094 | 0.12 | 0.014 | 0.052 |
| PORAdaptive(ASHRAE), 90% | 0.019 | 0.023 | 0.604 | 0.137 | 0.158 | 0.018 | 0.054 |
| PORAdaptive (EN), Class II | 0.057 | 0.053 | 0.721 | 0.078 | 0.058 | 0 | 0.019 |
| DhCoverheat | 0.004 | 0.004 | 0.663 | 0.073 | 0.069 | 0 | 0 |

**Figure 2.** Values of MI for each feature-output target variable.

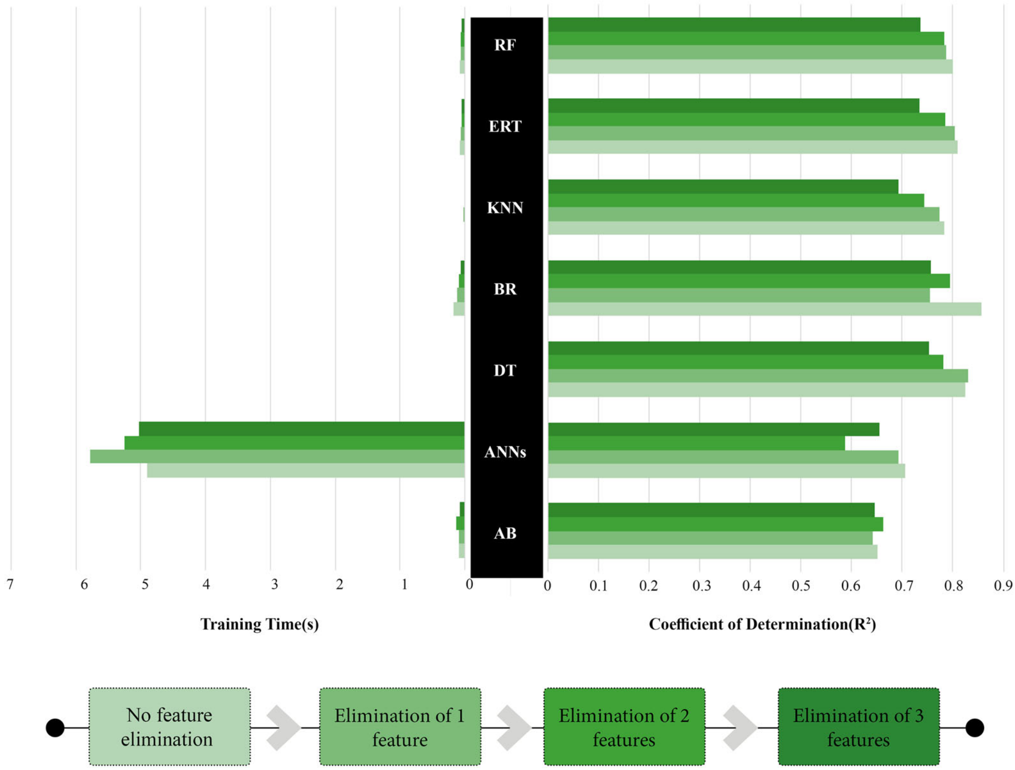


**Figure 3.** Validation with training data.

one-step feature elimination. However, reducing features shows a positive yet not significant impact on the training time with a trade-off in models' accuracies. Based on the results, the walls' U-value feature is eliminated from the DT model and no feature reduction is conducted for other prediction models. The same happened with the AB model with elimination of two features.

- *Hyperparameter tuning*: This step aims to increase the prediction accuracy. Table 5 shows the values and ranges of hyperparameters utilized for the tuning process. The grid search method is used to determine the best model and their corresponding optimum parameter values to avoid overfitting or underfitting phenomena. This method examines all the combinations of the parameters to find the optimum values (scikit-learn API Reference 2020).

**Table 5.** Base and optimum values for hyperparameters corresponding to each model.

| Model | Hyperparameters | Base value | Optimization range | Optimum values(Grid-Search tuning method) |
|---|---|---|---|---|
| ANN | hidden_layer_sizes=(100,) | 100 | 1–300 | 183 |
| | activation='relu' | relu | 'identity', 'logistic', 'tanh', 'relu' | 'relu' |
| | solver='adam' | adam | 'lbfgs', 'sgd', 'adam' | 'adam' |
| KNN | n_neighbors | 11 | 1–30 | 5 |
| | weights | 'uniform' | 'uniform', 'distance' | 'distance' |
| DT | max_depth | 7 | 1–30 | 15 |
| RF | n_estimators | 50 | 1–100 | 98 |
| | max_depth | 7 | 1–30 | 14 |
| ERT | n_estimators | 50 | 1–100 | 56 |
| | max_depth | 7 | 1–30 | 28 |
| AB | n_estimators | 50 | 1–100 | 8 |
| | learning_rate | 0.1 | 0.0-2.0 | 1.26 |
| BR | n_estimators | 50 | 1–100 | 94 |

- *Training time of the optimized models*: Prediction time differs in each model depending on their architecture. Training times are between 10 and 490 milliseconds, except for the ANNs model, which is trained in 6 s. Prediction time for none of the models exceeds 5 milliseconds. Compared to EnergyPlus, which took 10 s for each alternative on average, this is a meaningful reduction.

### *Model validation and error calculation*

Validation of the models is investigated using three main methods according to the literature.

- The first method is to use a random percentage of the training data to test the accuracy of the prediction models on the training set (Figure 4); results using this method commonly have higher variability in different runs due to the randomness of the validation set (S. Singaravel 2020).
- Another method, called cross-validation, is to repeat the previous method for all partitions of training data so that all the training set data points would be considered as the validation set in various runs. The final accuracy result would be the average results of all the steps (Figure 5) (Chou and Bui 2014).
- Neither of the above-mentioned approaches is able to measure the generalizability of models. To validate the model, unseen test data with different feature values is required. In this method, an unseen dataset is used to calculate the reliability of the prediction model (Singaravel 2020).

### *Error and accuracy calculation*

Three error calculation indices, including $R^2$, MAE, and MSE are used to identify the similarity of the predicted energy demand and comfort results with their actual values from EnergyPlus simulation. MAE has often been used in DT and RF studies. However, MSE has been generally used in domains relying on minimizing the least-squares. Low values of MAE and MSE indicate more accuracy while higher $R^2$ values, close to 1.0, demonstrate higher accuracy and show more similarity between the predicted and actual results.

The accuracy of the metrics, besides the training time of the model, is considered as performance indicator.

## Results

### *Prediction results of the base models*

Accuracy of base models in predicting the desired output indices of unseen data based on $R^2$, MSE, and MAE are presented in Figures 6-8, as well as the minimum and maximum values in the prediction of

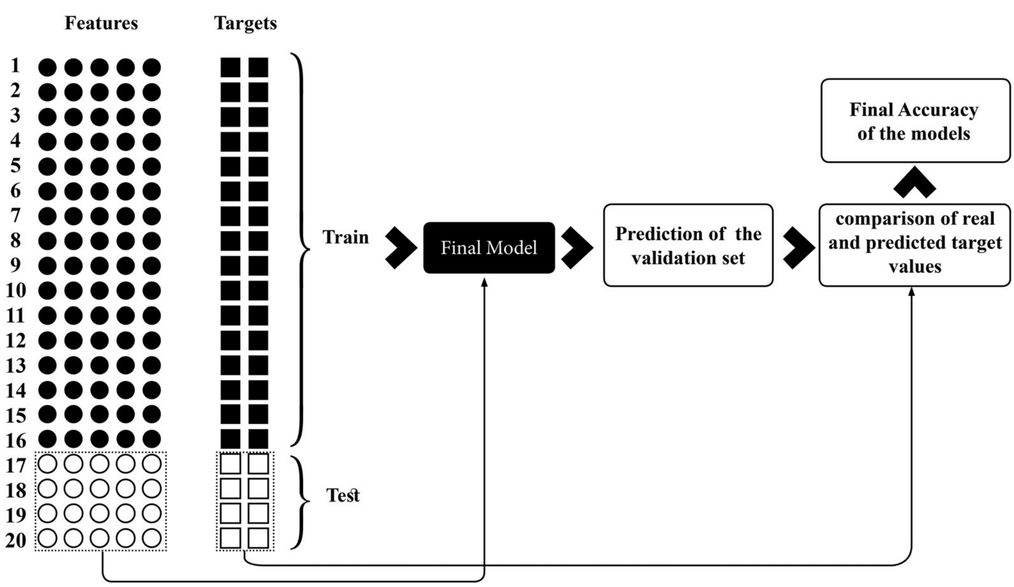


**Figure 4.** Validation using the training data.

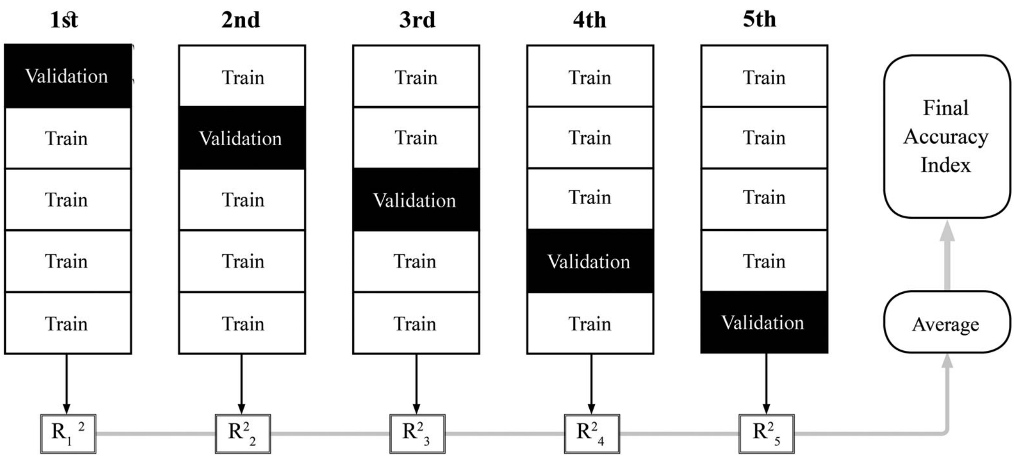


**Figure 5.** Cross-validation.

each output variable. According to the results, the average value of $R^2$ for heating and cooling demand is 0.62 and 0.83 respectively, and the accuracy of the models ranges between 0.25 and 0.84 for heating demand and 0.74–0.95 for cooling demand. For comfort, the average coefficients of determination are in the range of 0.59–0.91.

### *Prediction results of the optimized model*

In this section, performance improvements of each model are presented after optimization steps, including feature selection and hyperparameter tuning.

Detailed and final values of the hyperparameters for the optimized models are presented in Table 5. The final results of the optimized models are shown in Figures 9–11. According to these results, $R^2$ values related to the cooling and heating demands increased to 0.85 and 0.66, respectively. Also, the average $R^2$ values are in the range of 0.58–0.92 for the annual thermal comfort indices.

| | Cooling Demand (kWh/m2) | Heating Demand (kWh/m2) | $POR_{Fanger, 80\%}$ | $POR_{Adaptive(ASHRAE), 80\%}$ | $POR_{Adaptive(ASHRAE), 90\%}$ | $POR_{Adaptive (EN), Class II}$ | $DhC_{overheat}$ |
|---|---|---|---|---|---|---|---|
| Random Forest | 0.80 | 0.85 | 0.70 | 0.94 | 0.90 | 0.90 | 0.82 |
| Extremely Randomized Trees | 0.99 | 0.57 | 0.75 | 0.94 | 0.93 | 0.87 | 0.81 |
| K-Nearest Neighbor | 0.87 | 0.66 | 0.71 | 0.93 | 0.90 | 0.87 | 0.76 |
| Adaboost | 0.73 | 0.24 | 0.62 | 0.78 | 0.71 | 0.75 | 0.68 |
| Bagging | 0.81 | 0.85 | 0.70 | 0.94 | 0.90 | 0.90 | 0.82 |
| Decision Tree | 0.80 | 0.84 | 0.70 | 0.95 | 0.88 | 0.90 | 0.81 |
| Artificial Neural Network | 0.94 | 0.63 | 0.12 | 0.92 | 0.89 | 0.78 | 0.75 |
| Average | 0.85 | 0.66 | 0.61 | 0.91 | 0.87 | 0.85 | 0.78 |
| Minimum | 0.73 | 0.24 | 0.12 | 0.78 | 0.71 | 0.75 | 0.68 |
| Maximum | 0.99 | 0.85 | 0.75 | 0.95 | 0.93 | 0.90 | 0.82 |

**Figure 6.** Cross validation.

| | Cooling Demand (kWh/m2) | Heating Demand (kWh/m2) | $POR_{Fanger, 80\%}$ | $POR_{Adaptive(ASHRAE), 80\%}$ | $POR_{Adaptive(ASHRAE), 90\%}$ | $POR_{Adaptive (EN), Class II}$ | $DhC_{overheat}$ |
|---|---|---|---|---|---|---|---|
| Random Forest | 445.55 | 3.73 | 25.83 | 6.59 | 13.25 | 8.07 | 109208.90 |
| Extremely Randomized Trees | 25.27 | 10.45 | 21.48 | 6.51 | 9.38 | 10.59 | 110918.65 |
| K-Nearest Neighbor | 303.04 | 8.23 | 24.77 | 8.50 | 12.92 | 11.06 | 142236.77 |
| Adaboost | 613.85 | 18.41 | 31.94 | 26.09 | 37.53 | 20.73 | 191686.05 |
| Bagging | 438.74 | 3.75 | 25.58 | 6.57 | 13.12 | 8.10 | 108900.93 |
| Decision Tree | 448.87 | 3.97 | 25.34 | 6.13 | 15.86 | 7.93 | 111471.57 |
| Artificial Neural Network | 131.20 | 8.95 | 93.94 | 9.30 | 14.36 | 18.16 | 151829.56 |
| Average | 343.79 | 8.21 | 35.55 | 9.96 | 16.63 | 12.09 | 132321.78 |
| Minimum | 25.27 | 3.73 | 21.48 | 6.13 | 9.38 | 7.93 | 108900.93 |
| Maximum | 613.85 | 18.41 | 93.94 | 26.09 | 37.53 | 20.73 | 191686.05 |

**Figure 7.** Accuracy of base models in prediction of the unseen data based on $R^2$.

## Discussion

In this study, the accuracy of seven ML models in predicting thermal comfort and energy demands are studied and reported. Accuracy of models in similar studies in the literature (see Table 1) are between 0.4 and 1.00 for the prediction of heating and cooling demand in terms of $R^2$. In this study, the highest $R^2$ was 0.97 for ERT for prediction of cooling demand and 0.84 for the BR in predicting heating demand. Also, the accuracy values of the best models in this study for predicting comfort indices are between 0.74 and 0.96($R^2$) for BR and ERT(Figure 12).

In contrast with the energy demand that has been studied vastly in literature, little research is conducted for the annual thermal comfort performance. These studies are conducted on real case situations using field-measured datasets like ASHRAE Comfort Database II (Luo et al. 2020). The current approach differs from

| Random Forest | Cooling Demand (kWh/m2) | Heating Demand (kWh/m2) | $POR_{Fanger, 80\%}$ | $POR_{Adaptive(ASHRAE), 80\%}$ | $POR_{Adaptive(ASHRAE), 90\%}$ | $POR_{Adaptive (EN), Class II}$ | $DhC_{overheat}$ |
|---|---|---|---|---|---|---|---|
| Random Forest | 14.56 | 1.83 | 4.44 | 2.54 | 3.02 | 2.98 | 183.98 |
| Extremely Randomized Trees | 3.78 | 2.84 | 4.04 | 2.27 | 2.57 | 3.00 | 200.47 |
| K-Nearest Neighbor | 13.58 | 2.03 | 4.35 | 2.48 | 3.00 | 3.02 | 234.28 |
| Adaboost | 18.02 | 3.51 | 4.73 | 4.55 | 5.16 | 4.13 | 315.32 |
| Bagging | 17.18 | 1.44 | 4.43 | 2.21 | 2.90 | 2.55 | 175.99 |
| Decision Tree | 17.31 | 1.54 | 4.41 | 2.12 | 3.11 | 2.53 | 179.18 |
| Artificial Neural Network | 9.25 | 2.32 | 8.34 | 2.43 | 3.15 | 3.71 | 280.21 |
| Average | 13.38 | 2.22 | 4.96 | 2.66 | 3.27 | 3.13 | 224.20 |
| Minimum | 3.78 | 1.44 | 4.04 | 2.12 | 2.57 | 2.53 | 175.99 |
| Maximum | 18.02 | 3.51 | 8.34 | 4.55 | 5.16 | 4.13 | 315.32 |

**Figure 8.** Accuracy of base models in prediction of the unseen data based on MAE.

| | Cooling Demand (kWh/m2) | Heating Demand (kWh/m2) | $POR_{Fanger, 80\%}$ | $POR_{Adaptive(ASHRAE), 80\%}$ | $POR_{Adaptive(ASHRAE), 90\%}$ | $POR_{Adaptive (EN), Class II}$ | $DhC_{overheat}$ |
|---|---|---|---|---|---|---|---|
| Random Forest | 0.837 | 0.778 | 0.69 | 0.925 | 0.897 | 0.864 | 0.811 |
| Extremely Randomized Trees | 0.947 | 0.383 | 0.701 | 0.931 | 0.899 | 0.844 | 0.807 |
| K-Nearest Neighbor | 0.739 | 0.577 | 0.726 | 0.903 | 0.879 | 0.852 | 0.74 |
| Adaboost | 0.742 | 0.205 | 0.621 | 0.811 | 0.747 | 0.735 | 0.616 |
| Bagging | 0.805 | 0.844 | 0.698 | 0.944 | 0.898 | 0.9 | 0.818 |
| Decision Tree | 0.829 | 0.708 | 0.664 | 0.914 | 0.882 | 0.861 | 0.801 |
| Artificial Neural Network | 0.919 | 0.631 | 0.105 | 0.923 | 0.813 | 0.831 | 0.74 |
| Average | 0.831 | 0.589 | 0.683 | 0.907 | 0.859 | 0.841 | 0.762 |
| Minimum | 0.739 | 0.205 | 0.621 | 0.811 | 0.747 | 0.735 | 0.616 |
| Maximum | 0.947 | 0.844 | 0.726 | 0.944 | 0.899 | 0.900 | 0.818 |

**Figure 9.** Accuracy of optimized models in prediction of the unseen data based on $R^2$.

existing studies in that the simulated data is used instead of the measured data to compare the accuracy of different ML methods.

There are multiple factors affecting final accuracy, including the number of samples, feature and label data types, ML algorithms and model architectures, and validation strategy. Thus comparing different studies with different approaches may not be reasonable. For example, Singaravel et al. implemented Deep-learning neural network architecture with a component-based approach for a similar purpose (Sundaravelpandian Singaravel, Suykens, and Geyer 2018). In this approach, heat transfers through envelope were predicted using a dataset of 800 design combinations of a two-story building in Brussels with monthly energy data as the training set. This approach has resulted in an $R^2$ of 0.96-0.99 in predicting 201 unseen cases. This component-based approach also is investigated in other publications of these authors (Geyer and Singaravel 2018; Sundaravelpandian Singaravel and Geyer 2016)

| | Cooling Demand (kWh/m2) | Heating Demand (kWh/m2) | $POR_{Fanger, 80\%}$ | $POR_{Adaptive(ASHRAE), 80\%}$ | $POR_{Adaptive(ASHRAE), 90\%}$ | $POR_{Adaptive (EN), Class II}$ | $DhC_{overheat}$ |
|---|---|---|---|---|---|---|---|
| Random Forest | 366.35 | 5.35 | 26.25 | 8.74 | 13.57 | 11.15 | 112908.81 |
| Extremely Randomized Trees | 118.98 | 14.87 | 25.36 | 8.04 | 13.29 | 12.79 | 115419.43 |
| K-Nearest Neighbor | 587.10 | 10.20 | 23.22 | 11.35 | 15.89 | 12.15 | 155063.01 |
| Adaboost | 580.37 | 19.16 | 32.09 | 3.10 | 33.16 | 21.66 | 228965.52 |
| Bagging | 438.18 | 3.75 | 25.60 | 6.51 | 13.35 | 8.15 | 108660.71 |
| Decision Tree | 384.39 | 7.04 | 28.44 | 10.04 | 15.46 | 11.41 | 118623.70 |
| Artificial Neural Network | 182.84 | 8.91 | 84.75 | 8.95 | 24.62 | 13.86 | 155452.49 |
| Average | 379.75 | 9.90 | 35.10 | 8.10 | 18.48 | 13.02 | 142156.24 |
| Minimum | 118.98 | 3.75 | 23.22 | 3.10 | 13.29 | 8.15 | 108660.71 |
| Maximum | 587.10 | 19.16 | 84.75 | 11.35 | 33.16 | 21.66 | 228965.52 |

**Figure 10.** Accuracy of optimized models in prediction of the unseen data based on MSE.

| | Cooling Demand (kWh/m2) | Heating Demand (kWh/m2) | $POR_{Fanger, 80\%}$ | $POR_{Adaptive(ASHRAE), 80\%}$ | $POR_{Adaptive(ASHRAE), 90\%}$ | $POR_{Adaptive (EN), Class II}$ | $DhC_{overheat}$ |
|---|---|---|---|---|---|---|---|
| Random Forest | 14.56 | 1.83 | 4.44 | 2.54 | 3.02 | 2.98 | 183.98 |
| Extremely Randomized Trees | 8.24 | 3.16 | 4.23 | 2.49 | 2.96 | 3.29 | 204.26 |
| K-Nearest Neighbor | 18.81 | 2.30 | 4.16 | 2.84 | 3.36 | 3.18 | 253.28 |
| Adaboost | 18.02 | 3.56 | 4.75 | 0.84 | 4.78 | 4.23 | 351.93 |
| Bagging | 17.01 | 1.45 | 4.43 | 2.19 | 2.93 | 2.56 | 175.43 |
| Decision Tree | 15.19 | 2.13 | 4.58 | 2.68 | 3.19 | 2.99 | 190.91 |
| Artificial Neural Network | 10.82 | 2.42 | 8.13 | 2.33 | 3.97 | 3.17 | 290.08 |
| Average | 14.66 | 2.41 | 4.96 | 2.27 | 3.46 | 3.20 | 235.70 |
| Minimum | 8.24 | 1.45 | 4.16 | 0.84 | 2.93 | 2.56 | 175.43 |
| Maximum | 18.81 | 3.56 | 8.13 | 2.84 | 4.78 | 4.23 | 351.93 |

**Figure 11.** Accuracy of Optimized models in prediction of the unseen data based on MAE.

Present study has comparatively investigated the feasibility of using ML models for the prediction of building energy demand and annual thermal comfort. Results of this study can be used to develop a data-driven framework by determining the most accurate and reliable prediction model for each of the target variables. This algorithmic framework is presented in three steps, including 4 (1) Defining inputs by the users, (2) Importing user-defined inputs to the black-box models, and (3) Calculation of final results. This framework and two of the most accurate models for each of the target variables are presented in Figure 12.

This study is constrained to definite sets of design parameters (dimensions, WWR, window and walls U-values, shading depth, and rotation angle) with specific ranges as input features. For a more inclusive framework, all possible input parameters and value ranges need to be considered to increase the generalizability of the models and include more complex design options, such as different glazing systems, shadings, façade

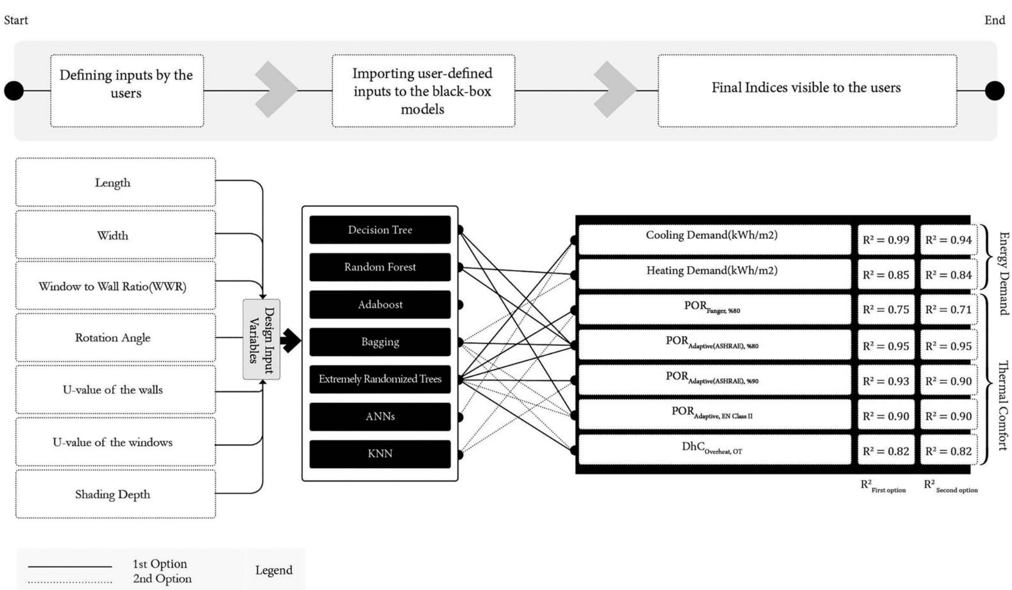


**Figure 12.** Framework for calculation of building energy demand and annual thermal comfort.

control strategies, etc. In addition, this study was limited to a single zone calculation. Building level frameworks should be studied for a more applicable framework. Other common building types, such as residential and educational, also should be included in the training data. Moreover, different weather conditions should be included for generating the training set. Other novel ML approaches, like deep-learning, should also be investigated and compared with other algorithms in terms of prediction accuracy. Finally, RMSE is not calculated as an accuracy metric; it is suggested that this metric be taken into account in the future research, considering its good performance in comparing different linear regression models.

## Conclusion

According to the results, ML methods can predict building energy demand and thermal comfort by up to the $R^2$ of 0.99 and 0.95 respectively by using the ERT model. Overall, single and ensemble models based on decision trees are shown to have relatively better performance.

Data-driven methods are implemented in some of the building energy modeling web-based toolkits such as Enerpro. These toolkits utilize hundreds of pre-simulated (DOE-2) model-based archetypes for its analysis, and Targeting Tool for Energy Retrofits (BETTER) that uses regression techniques to analyse a building's monthly energy use history and find the most cost-effective energy conservation measures (Szum Berkeley et al. 2018; Enersys Analytics Inc. n.d.). Moreover, reduction of simulation and computation time relative to physics-based models and elimination of the need for detailed energy modeling in the early stages are other advantages of this method. However, to achieve a suitable trained model capable of predicting unseen samples, much time and computation need to be dedicated to generating enough training data and conducting corresponding simulations.

## Disclosure statement

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

## Funding

This work was supported by Iran National Science Foundation [grant number 99008894].

## Notes on contributors

***Nima Forouzandeh*** is a PhD candidate in building physics at the Faculty of Architecture, Delft University of Technology. His research interest is in high performance buildings, daylighting, and building physics.

***Zahra Sadat Zomorodian*** is an Assistant Professor at the SBU Department of Construction. Her research expertise is in Thermal Comfort, Daylighting, Building Energy Simulation (BES), Indoor Environmental Quality (IEQ) assessments and Integrated Green Design Process (IGDP).

***Ms. Zohreh Shaghaghian*** is a PhD candidate in computational design at the Department of Architecture, Texas A&M University. She is a LEED AP certified for Building Design and Construction (BD+C) in the association of the US Green Building Council (USGBC). During her PhD program she has focused on emerging technologies and their application in architectural design and education.

***Mohammad Tahsildoost*** is an Associate Professor at the SBU Department of Construction. His research expertise is in Building Energy Simulation (BES), Integrated Green Design Process (IGDP), architectural technology, high-rise buildings, and construction 1technology.

## ORCID

*Nima Forouzandeh* http://orcid.org/0000-0001-7025-5509
*Zahra Sadat Zomorodian* http://orcid.org/0000-0003-2418-8856

## References

Amasyali, Kadir, and Nora M. El-Gohary. 2018. "A Review of Data-Driven Building Energy Consumption Prediction Studies." *Renewable and Sustainable Energy Reviews* 81 (September 2017): 1192–1205. doi:10.1016/j.rser.2017.04.095.

Amiri, Shideh Shams, Mohammad Mottahedi, and Somayeh Asadi. 2015. “Using Multiple Regression Analysis to Develop Energy Consumption Indicators for Commercial Buildings in the U.S.” *Energy and Buildings* 109: 209–216. doi:10.1016/j.enbuild.2015.09.073.

Barraza, Néstor, Sérgio Moro, Marcelo Ferreyra, and Adolfo de la Peña. 2019. “Mutual Information and Sensitivity Analysis for Feature Selection in Customer Targeting: A Comparative Study.” *Journal of Information Science* 45 (1): 53–67. doi:10.1177/0165551518770967.

Beckman, William A., Lars Broman, Alex Fiksel, Sanford A. Klein, Eva Lindberg, Mattias Schuler, and Jeff Thornton. 1994. “TRNSYS The Most Complete Solar Energy System Modeling and Simulation Software.” *Renewable Energy* 5 (1–4): 486–488. doi:10.1016/0960-1481(94)90420-0.

Bektas Ekici, Betul, and U. Teoman Aksoy. 2011. “Prediction of Building Energy Needs in Early Stage of Design by Using ANFIS.” *Expert Systems with Applications* 38 (5): 5352–5358. doi:10.1016/j.eswa.2010.10.021.

Brandi, Silvio, Marco Savino Piscitelli, Marco Martellacci, and Alfonso Capozzoli. 2020. “Deep Reinforcement Learning to Optimise Indoor Temperature Control and Heating Energy Consumption in Buildings.” *Energy and Buildings* 224: 110225. doi:10.1016/j.enbuild.2020.110225.

“BS EN 16798-1 : 2019”. 2019. *BSI Standards Publication Energy Performance of Buildings – Ventilation for Buildings*.

Carlucci, Salvatore, and Lorenzo Pagliano. 2012. “A Review of Indices for the Long-Term Evaluation of the General Thermal Comfort Conditions in Buildings.” *Energy and Buildings*. doi:10.1016/j.enbuild.2012.06.015.

Catalina, Tiberiu, Joseph Virgone, and Eric Blanco. 2008. “Development and Validation of Regression Models to Predict Monthly Heating Demand for Residential Buildings.” *Energy and Buildings* 40 (10): 1825–1832. doi:10.1016/j.enbuild.2008.04.001.

Cheng, Min Yuan, and Minh Tu Cao. 2014. “Accurately Predicting Building Energy Performance Using Evolutionary Multivariate Adaptive Regression Splines.” *Applied Soft Computing Journal* 22: 178–188. doi:10.1016/j.asoc.2014.05.015.

Chou, Jui Sheng, and Dac Khuong Bui. 2014. “Modeling Heating and Cooling Loads by Artificial Intelligence for Energy-Efficient Building Design.” *Energy and Buildings* 82 (2014): 437–446. doi:10.1016/j.enbuild.2014.07.036.

Ciulla, G., and A. D’Amico. 2019. “Building Energy Performance Forecasting: A Multiple Linear Regression Approach.” *Applied Energy* 253 (April): 113500. doi:10.1016/j.apenergy.2019.113500.

Crawley, Drury B, Linda K Lawrie, Curtis O Pedersen, and Frederick C Winkelmann. 2000. “Energy Plus: Energy Simulation Program.” *ASHRAE Journal* 42 (4): 49–56.

Crawley, Drury B., Linda K. Lawrie, Frederick C. Winkelmann, W. F. Buhl, Y. Joe Huang, Curtis O. Pedersen, Richard K. Strand, et al. 2001. “EnergyPlus: Creating a New-Generation Building Energy Simulation Program.” *Energy and Buildings* 33 (4): 319–331. doi:10.1016/S0378-7788(00)00114-6.

Dietrich, Bastian, Jessica Walther, Matthias Weigold, and Eberhard Abele. 2020. “Machine Learning Based Very Short Term Load Forecasting of Machine Tools.” *Applied Energy* 276 (June): 115440. doi:10.1016/j.apenergy.2020.115440.

Ekici, Betul Bektas, and U. Teoman Aksoy. 2009. “Prediction of Building Energy Consumption by Using Artificial Neural Networks.” *Advances in Engineering Software* 40 (5): 356–362. doi:10.1016/j.advengsoft.2008.05.003.

Enersys Analytics Inc. n.d. “Energy Profile Tool (Enerpro).”Accessed September 26, 2020. http://www.energyprofiletool.com/subscription/default.asp.

Gelder, Liesje Van, Payel Das, Hans Janssen, and Staf Roels. 2014. “Comparative Study of Metamodelling Techniques in Building Energy Simulation: Guidelines for Practitioners.” *Simulation Modelling Practice and Theory* 49: 245–257. doi:10.1016/j.simpat.2014.10.004.

Geyer, Philipp, and Sundaravelpandian Singaravel. 2018. “Component-Based Machine Learning for Performance Prediction in Building Design.” *Applied Energy* 228 (October 2017): 1439–1453. doi:10.1016/j.apenergy.2018.07.011.

Han, Hua, Zhan Zhang, Xiaoyu Cui, and Qinghong Meng. 2020. “Ensemble Learning with Member Optimization for Fault Diagnosis of a Building Energy System.” *Energy and Buildings* 226: 110351. doi:10.1016/j.enbuild.2020.110351.

Hygh, Janelle S., Joseph F. DeCarolis, David B. Hill, and S. Ranji Ranjithan. 2012. “Multivariate Regression as an Energy Assessment Tool in Early Building Design.” *Building and Environment* 57: 165–175. doi:10.1016/j.buildenv.2012.04.021.

Iran National Building Code(Article 19). 2019. BHRC.

Standard, I. S. O. 2005. *7730. Ergonomics of the Thermal Environment—Analytical Determination and Interpretation of Thermal Comfort using Calculation of the PMV and PPD Indices and Local Thermal Comfort Criteria*. Geneva: International Organization for Standardization.

Kontokosta, Constantine E., and Christopher Tull. 2017. “A Data-Driven Predictive Model of City-Scale Energy Use in Buildings.” *Applied Energy* 197: 303–317. doi:10.1016/j.apenergy.2017.04.005.

Kumar, Sachin, Saibal K. Pal, and Ram Pal Singh. 2018. “A Novel Method Based on Extreme Learning Machine to Predict Heating and Cooling Load Through Design and Structural Attributes.” *Energy and Buildings* 176: 275–286. doi:10.1016/j.enbuild.2018.06.056.

Li, Xinyi, and Runming Yao. 2020. “A Machine-Learning-Based Approach to Predict Residential Annual Space Heating and Cooling Loads Considering Occupant Behaviour.” *Energy* 212: 118676. doi:10.1016/j.energy.2020.118676.

Luo, Maohui, Jiaqing Xie, Yichen Yan, Zhihao Ke, Peiran Yu, Zi Wang, and Jingsi Zhang. 2020. “Comparing Machine Learning Algorithms in Predicting Thermal Sensation Using ASHRAE Comfort Database II.” *Energy and Buildings* 210: 109776. doi:10.1016/j.enbuild.2020.109776.

Magoulès, Frédéric, and Hai Xiang Zhao. 2016, February. “Data Mining and Machine Learning in Building Energy Analysis.” *Data Mining and Machine Learning in Building Energy Analysis* Chapter 4: 1–164.

Mottahedi, Mohammad, Atefeh Mohammadpour, Shideh Shams Amiri, David Riley, and Somayeh Asadi. 2015. “Multi-Linear Regression Models to Predict the Annual Energy Consumption of an Office Building with Different Shapes.” *Procedia Engineering* 118: 622–629. doi:10.1016/j.proeng.2015.08.495.

Naji, Sareh, Afram Keivani, Shahaboddin Shamshirband, U. Johnson Alengaram, Mohd Zamin Jumaat, Zulkefli Mansor, and Malrey Lee. 2016. “Estimating Building Energy Consumption Using Extreme Learning Machine Method.” *Energy* 97: 506–516. doi:10.1016/j.energy.2015.11.037.

Ngo, Ngoc Tri. 2019. "Early Predicting Cooling Loads for Energy-Efficient Design in Office Buildings by Machine Learning." *Energy and Buildings* 182: 264–273. doi:10.1016/j.enbuild.2018.10.004.
Pan, Yue, and Limao Zhang. 2020. "Data-Driven Estimation of Building Energy Consumption with Multi-Source Heterogeneous Data." *Applied Energy* 268 (February): 114965. doi:10.1016/j.apenergy.2020.114965.
Robinson, Caleb, Bistra Dilkina, Jeffrey Hubbs, Wenwen Zhang, Subhrajit Guhathakurta, Marilyn A. Brown, and Ram M. Pendyala. 2017. "Machine Learning Approaches for Estimating Commercial Building Energy Consumption." *Applied Energy* 208 (May): 889–904. doi:10.1016/j.apenergy.2017.09.060.
Romani, Zaid, Abdeslam Draoui, and Francis Allard. 2015. "Metamodeling the Heating and Cooling Energy Needs and Simultaneous Building Envelope Optimization for Low Energy Building Design in Morocco." *Energy and Buildings* 102: 139–148. doi:10.1016/j.enbuild.2015.04.014.
Scikit-learn API Reference. 2020. "Scikit-Learn API Reference." 2020. https://scikit-learn.org/stable/modules/classes.html.
Seyedzadeh, Saleh, and Farzad Pour Rahimian. 2021. *Data-Driven Modelling of Non-Domestic Buildings Energy Performance: Supporting Building Retrofit Planning*.
Singaravel, S. 2020. *Machine Learning for Energy Performance Prediction in Early Design Stage of Buildings*.
Singaravel, Sundaravelpandian, and Philipp Geyer. 2016. "Simplifying Building Energy Performance Models to Support an Integrated Design Workflow."*23rd International Workshop of the European Group for Intelligent Computing in Engineering, EG-ICE* 2016, no. June.
Singaravel, Sundaravelpandian, Philipp Geyer, Johan Suykens, and K U Leuven. 2017. "Component-Based Machine Learning Modelling Approach For Design Stage Building Energy Prediction : Weather Conditions And Size Architectural Engineering Division, KU Leuven, Belgium Case-Based Development and Evaluation of Component-Based MLM Proposed Me," no. August: 2617–26.
Singaravel, Sundaravelpandian, Johan Suykens, and Philipp Geyer. 2018. "Deep-Learning Neural-Network Architectures and Methods: Using Component-Based Models in Building-Design Energy Prediction." *Advanced Engineering Informatics* 38 (June): 81–90. doi:10.1016/j.aei.2018.06.004.
Singaravel, Sundaravelpandian, Johan Suykens, and Philipp Geyer. 2019. "Deep Convolutional Learning for General Early Design Stage Prediction Models." *Advanced Engineering Informatics* 42: 100982. doi:10.1016/j.aei.2019.100982.
Szum Berkeley, Carolyn, Han Li Berkeley, Steven C Snyder, Ahmed Bekhit, Clay G Nesler, and Sara Lisauskas. 2018. "Energy Efficiency Targeting Tool v1.0." *United States*. doi:10.11578/dc.20190108.2.
Tsanas, Athanasios, and Angeliki Xifara. 2012. "Accurate Quantitative Estimation of Energy Performance of Residential Buildings Using Statistical Machine Learning Tools." *Energy and Buildings* 49: 560–567. doi:10.1016/j.enbuild.2012.03.003.
Turhan, Cihan, Tugce Kazanasmaz, Ilknur Erlalelitepe Uygun, Kenan Evren Ekmen, and Gulden Gokcen Akkurt. 2014. "Comparative Study of a Building Energy Performance Software (KEP-IYTE-ESS) and ANN-Based Building Heat Load Estimation." *Energy and Buildings* 85: 115–125. doi:10.1016/j.enbuild.2014.09.026.
Walter, Travis, and Michael D. Sohn. 2016. "A Regression-Based Approach to Estimating Retrofit Savings Using the Building Performance Database." *Applied Energy* 179: 996–1005. doi:10.1016/j.apenergy.2016.07.087.
Wang, Zeyu, and Ravi S. Srinivasan. 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*. doi:10.1016/j.rser.2016.10.079.
Wei, Lai, Wei Tian, Elisabete A. Silva, Ruchi Choudhary, Qingxin Meng, and Song Yang. 2015. "Comparative Study on Machine Learning for Urban Building Energy Analysis." *Procedia Engineering* 121: 285–292. doi:10.1016/j.proeng.2015.08.1070.