mstas 2021
İstanbul Teknik Üniversitesi, Lisansüstü Eğitim Enstitüsü Bilişim Anabilim Dalı, Mimari Tasarımda Bilişim Lisansüstü Programı; Mimarlık Fakültesi; Hesaplamalı Tasarım Araştırmaları ve Eğitimi Ağı

MİMARLIKTA SAYISAL TASARIM
XV. ULUSAL SEMPOZYUMU

Artificial Neural Networks to Predict Performance of Classroom Spaces

Ayça Duran [1]; İpek Gürsel Dino [2]
[1, 2] Middle East Technical University
[1]aycaduran.com, [1]aycad@metu.edu.tr; [2]ipekg@metu.edu.tr

Abstract

Educational facilities account for approximately 12% of the energy consumed by buildings in the US and UK. Classrooms should provide their occupants' satisfactory indoor environments as indoor conditions play a determinant role in the performance, productivity, attendance, and health of students and teachers. Indoor air quality and thermal comfort are two major determinants of healthy classrooms. Generally, classrooms operate at full capacity, leading to severe indoor overheating degrees (IOD) and high carbon dioxide (CO_2) concentrations if not adequately ventilated. To assess classroom design alternatives in the design development phase and retrofit scenarios, building energy simulation is a widely used method to estimate performance indicators. However, consideration of a high number of design alternatives increases computational cost and requires tedious modeling efforts. Research in building performance predictions with machine learning methods gained increasing interest in recent years. Artificial neural networks (ANNs) are reported to yield satisfactory performance in the prediction of non-linear patterns of building performance. This study presents a data-driven framework to estimate heating energy demand, IOD, and CO_2 concentration of naturally ventilated classrooms with ANNs. Five input variables are selected to predict specified performance indicators. 200 classrooms with varying orientations, values of shape factor, glazing area, occupant density, and outdoor surface area are simulated. The ANNs are trained with a subset of EnergyPlus simulation results. Prediction models for three performance indicators are individually built, and prediction performances are evaluated. While regression coefficients range between 0.986 and 0.993, the average root means square error calculated is between %2 and 9%, implying high predictive capacity.

Keywords: *Artificial neural networks, building performance prediction, building energy simulation.*

1. Introduction

Buildings constitute 40% of Europe's energy consumption and 36% of the greenhouse gas emissions (Europeran Commission, 2020). Notably, educational facilities account for 13% and 10% of the energy use of the US and UK buildings, respectively (Pérez-Lombard et al., 2008). Energy consumption of buildings is growing worldwide. Heating, ventilation, and air conditioning, the main factors of controlling indoor air quality, are responsible for the majority of building energy demand. In other words, providing thermal comfort indoors is the main demand behind energy consumption (Yang, et al., 2014). Specifically, children spend around 85% of their time indoors (Langevin, et al., 2016). Until the age of 18, students are reported to spend more time at school than any other place but at home (Bluyssen et al., 2018).

Performance, productivity, attendance, and health of both students and teachers depend significantly on the indoor conditions in educational facilities (Zomorodian, et al., 2016). A study based on test scores implies that the increase in satisfaction with the indoor environmental quality enhances the learning performance of students' (Mumovic, et al., 2009). Several studies have found an inverse correlation between the CO_2 concentration levels and pupils' annual school attendance (Gaihre, et al., 2014; Shendell, et al., 2004). Similarly, it is found that students' performance in math exams is also significantly related to classroom-level ventilation concerning CO_2 concentrations (Shaughnessy, et al, 2006).

Indoor environmental quality influences the well-being in school buildings since pupils are more sensitive to indoor climate conditions because of the nature of children's anatomical structures. Children are more susceptible to certain environmental pollutants than adults as the amount of air intake proportional to their body weight is more significant (Faustman, et al., 2000). As two major determinant factors of healthy classrooms for children, indoor air quality and thermal comfort are broadly emphasized in many studies on educational facilities (Zomorodian, et al., 2016). One of the key drivers is indoor air quality through CO2 concentration. It is also recommended by ASHRAE Standard 62 (ASHRAE, 2019) and STM D6245 (ASTM, 1998). Therefore, the evaluation of room ventilation due to CO_2 generated by its occupants becomes a standard evaluation method (Bartlett, et al., 2004). Additionally, the link between indoor air and thermal quality is a prominent study subject in the field (Fabi, et al., 2013). The indoor overheating degree (IOD) is studied as thermal comfort is an important determinant of students' learning performance. Air temperature, one of the significant factors influencing thermal comfort, has a considerable impact on learning (Heschong Mahone Group, 2003). On the other hand, while providing comfort conditions in classrooms, energy demand may increase. Notably, in school buildings, space heating accounts for 47% of the total energy demand (NREL, 2013). There is a tradeoff between occupant comfort and resource consumption, and the annual heating energy demand ($Q_{heating}$) of naturally ventilated classrooms is studied as the third performance indicator.

Building performance simulation has been an accurate and widely used method for quantifying performance indicators that enabling the design and operation of energy-efficient buildings (Yezioro, et al. 2008). Simulated results give an insight into the real-world data's underlying changes and trends (Wan et al., 2011). Even though using such advanced simulation tools give reliable results, it can be time-consuming and requires the user to learn and become an expert on the tool. When a number of design alternatives are needed to be evaluated, and various building parameters are involved in the design, machine learning tools are trusted broadly by the researchers for many years (Yu, et al., 2010; Tsanas & Xifara, 2012; Catalina, et al., 2008).

Various machine learning techniques have been involved in the research for the prediction of building performance indicators. For instance, multiple linear regression, artificial neural networks (ANNs) (Kumar, et al., 2013), decision trees, and support vector machines have been explored in several studies (Seyedzadeh, et al., 2018). In some studies, estimations with ANNs have been found more reliable compared to other data-driven models because of their higher performance in the prediction of non-linear patterns (Walker, et al., 2020; Yalcintas & Ozturk, 2006; Seyedzadeh, et al., 2018). For instance, Ascione et al. (2017) accurately predicted building performance through energy consumption and thermal comfort of occupants with ANNs and offered a framework to be applied independently of the building type. Namely, space heating and cooling and the ratio of yearly discomfort hours are calculated. Later, they improved the whole-building parameters such as geometry, envelope, operation, etc., with ANNs for retrofitting purposes. Similarly, there have been several studies employing ANNs for performance predictions of educational buildings (Sterman & Baglione, 2012; Fan, et al., 2019; Neto & Fiorelli, 2008; Kusiak, et al., 2010). However, in these studies, selected educational buildings were only university buildings, and predictions were based on real-world data. There is a research gap in the development of predictive models applied to primary and secondary school buildings, particularly for the classroom spaces. Therefore, this study aims to predict the performance of classroom spaces based on the proposed framework for the analysis of synthetic data and the building of an ANNs.

2. Methodology

This paper proposes a data-driven framework for the analysis of the synthetic data and building of ANNs. Firstly, a naturally ventilated classroom in Ankara, Turkey, is modeled and simulated with EnergyPlus engine through Ladybug Tools (Roudsari & Pak, 2013) for data generation. Then the data is explored through the analysis of the simulation outputs. Finally, after data preprocessing and exploration, ANNs are built and their performances are evaluated. The described ANNs are generated in an Anaconda (*Anaconda Software Distribution*, 2020) environment in a python programming language with the TensorFlow (Abadi, et al., 2016) library.

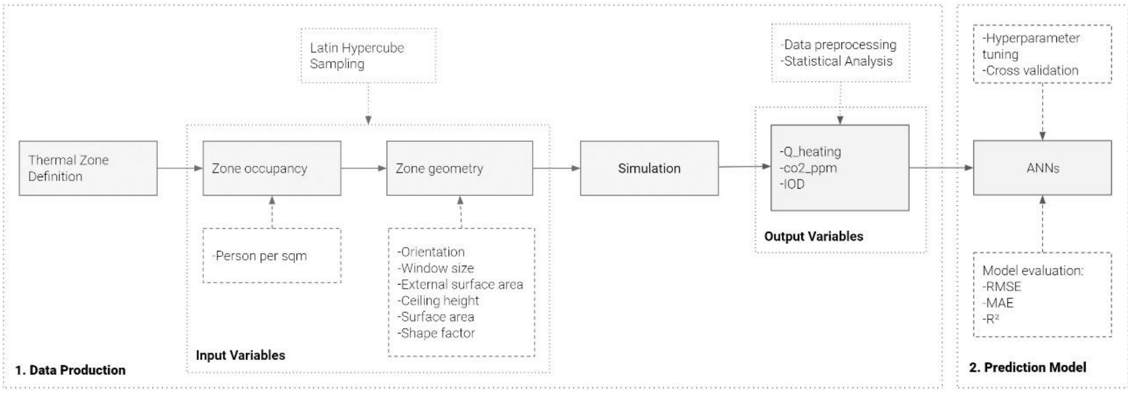

Figure 1: The proposed framework

2.1. Data

Variables and sampling: The impact of five input variables (**Table 1**); namely orientation, shape factor, glazing area, occupant density, and exposed surface area per floor area, is explored to determine the output variables of $Q_{heating}$, CO2 concentration, and IOD of classroom space. The majority of the indicated input variables have been associated widely with building performance in the energy performance of buildings literature (Tsanas & Xifara, 2012; Pessenlehner & Mahdavi, 2003). Furthermore, since studies show that CO_2 concentration in classrooms is significantly related to the number of students sharing the same space (Yalçın, et al., 2018), occupant density is also included in the input space. Input variables contain both categorical and continuous data. For the continuous variables, the *Latin hypercube sampling (LHS)* method, a generalized stratified sampling technique (Shields & Zhang, 2016) is applied. For the performance simulations, the weather file for Ankara is used. Building materials chosen for the simulated classroom space is based on a previously studied school building (Akköse, et al., 2021). Following the applied framework, selected performance indicators can be easily predicted with classrooms in other cities with different building materials.

$Q_{heating}$ (kWh/m^2), indoor CO_2 exceedance (ppm), and indoor overheating degree (°C) are recorded as output variables. Total heating energy demand normalized by floor area is recorded as the first output variable. The second output variable is quantified by the difference between indoor temperatures and the indoor operative temperature limit of 28°C and is named as indoor overheating degree (IOD) (Hamdy, et al, 2017). IOD is the cumulative sum of these hourly temperature differences. Finally, indoor CO_2 exceedance is calculated in the same manner to understand the intensity and frequency of concentration levels above the threshold specified by ASHRAE standards (ASHRAE, 2019). The outdoor CO_2 concentration is assumed to be 300 parts per million (ppm); therefore, with respect to ASHRAE standards, the highest acceptable CO_2 concentration limit is selected as 1000 ppm (ASHRAE, 2019). For the occupied hours, concentrations above 1000 pm is considered and aggregated annually.

Table 1: Input variables

Mathematical representation	Input variable	Unit	Distribution	Range	Sampling
X1	Orientation	-	uniform	$\in \{N, E, S, W\}$	-
X2	Shape factor	m^{-1}	uniform	$\in R: (1.0 - 1.3)$	LHS
X3	Glazing area	m^2	uniform	$\in R: (6.0 - 8.0)$	LHS
X4	Occupant density	ppl/m^2	uniform	$\in R: (16 - 30)$	LHS
X5	Exposed surface area per m^2	m^{-1}	uniform	$\in R: (18.0 - 36.4)$	LHS

Data preprocessing and statistical analysis: 200 simulations were performed. First missing and faulty data points that might be obtained due to errors occurred during the simulations are checked and eliminated from the dataset. In order to understand the data, descriptive statistics is used to analyze data distributions. Following, a correlation matrix is generated to capture variable dependency, and the data points are visualized. The Spearman correlation coefficient (ρ) is calculated to evaluate the monotonic relationship between variables.

2.2. Prediction Model

ANN: The human brain is composed of massive neural networks that enable humans to accomplish many complex tasks, including but not limited to face recognition, speaking, body movement, etc. More than 10 billion interconnected neurons in the human brain receive, process, and transmit information through biochemical reactions (Kumar, et al., 2013). In reference to the brains' nervous system, artificial neural networks have been developed as generalizations of underlying mathematical models of the interconnected neurons in the brain.

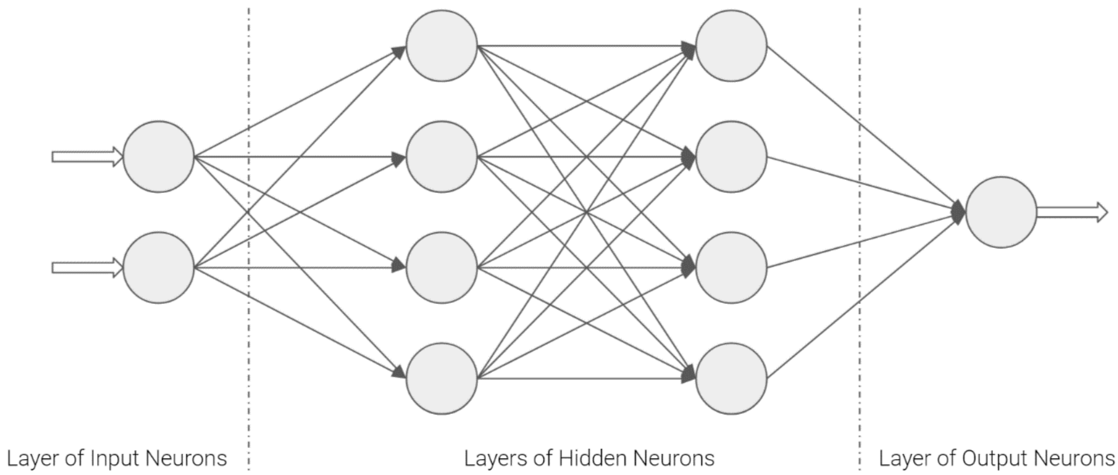

Figure 2: Feed-forward multi-layer perceptron (MLP) architecture.

An ANN is a processing data system that learns the input-output relationship from the obtained data. The most popular and simple architecture of ANN is the feed-forward multi-layer perceptron (MLP), and it is used in this study. It is composed of one input layer, single or multiple layers of hidden neurons, and one output layer (**Figure 2**). A neural network takes in inputs, multiplies them by weights and adds biases. Following, the results are passed to activation functions and outputs are received. An activation function defines the output of a node given an input or number of inputs. The last output is the ANN's prediction for the given input or inputs.

ANN Performance: The performance of ANN is related to both input and output data together with the model's architecture and parameters (Kalogirou & Bojic, 2000). The number of hidden layers is one of the most critical factors affecting the prediction model performance, and is typically detected by trial and error. Three prediction models are trained for each output variable, and different numbers of layers of hidden neurons are explored for better prediction performance. Although the number of layers changes, neurons in each layer is set to 5, based on the number of input variables. The activation function for hidden layers is also critical for ANNs' performance. Rectified Linear Unit (ReLU) activation function is a commonly preferred activation function in the literature, and it is chosen in this study. RMSE is selected for the loss function, while the regression optimizer is Adam, which is an algorithm for stochastic gradient descent for training ANNs. The batch size and epochs of training are also critical parameters of an ANN. Batch size is the number of training samples worked through in each iteration, i.e., epochs. It is found that larger batch sizes degrade the model quality in terms of the model's ability to generalize (Keskar, et al., 2017). Therefore, batch sizes of 1 and 4 are chosen with the adjusted number of epochs. Additionally, the proportion of data used for training and testing is also one of the critical decisions. The complete and preprocessed dataset is split into train and test subsets. Since input variables vary in terms of units and quantities, train and test sets are scaled separately to prevent data leakage before the training process. The ratio of training data to testing data is chosen as 8/2 for the CO_2 prediction model and 9/1 for IOD and $Q_{heating}$ after trials with 9/1, 8/2 and 7/2 proportions.

Model Evaluation: The networks' performance has been evaluated with the calculation of three different metrics: root mean square error (RMSE), mean absolute error (MAE), and R-squared (R^2). RMSE is the standard deviation of the prediction errors. It is the measure of how far the predictions deviate from the actual values (1). MAE is the average magnitude of errors in the prediction data set (2). Lastly, R^2, called the coefficient of determination in statistics, is the measure of how well the ML model predicts the actual values. It is based on the ratio of the total variation of outcomes predicted by the prediction model (3).

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n} \quad (p_i - a_i)^2}{n}} \tag{1}$$

$$MAE = \frac{1}{n}\sum_{i=1}^{n} \quad |p_i - a_i| \tag{2}$$

$$R^2 = 1 - \frac{\sum_{i=1}^{n} \quad (p_i - a_i)^2}{\sum_{i=1}^{n} \quad (p_i - \bar{a_i})^2} \tag{3}$$

where p_i is the predicted value of the performance indicator, a_i is the actual value of the indicator, and n is the number of data points in the dataset.

3. Results

3.1. Data Analysis

Two observations are removed from the dataset due to missing values found in some of the output variables. First, the statistical description of the data is examined (**Table 2**). The minimum energy demanding case is almost one third of the most demanding with the specified input parameters. The minimum-maximum ratio is significant in IOD results. Standard deviation is greater than the mean of the values, and the data points' distribution for IOD is right-skewed; in other words, there are extreme values significantly less in number and resulted uneven distribution of data points in the dataset.

Table 2: Output variable description

Mathematical Representation	Output Variable	Units	Mean	Std.	Min	Max
y1	$Q_{heating}$	kWh/m^2	20.706	8.327	5.514	49.188
y2	IOD	°C	124.652	184.075	9.005	1226.964
y3	CO_2 exceedance	ppm	2293987	545962	1391631	3878137

The relationship between input and output variables is visualized with the parallel coordinates graph (**Figure 3**). It is noticeable that the highest value of $Q_{heating}$ is observed when the glazing area is maximized. The energy loss from the larger glazing surface implies an increased $Q_{heating}$ demand. South-facing classrooms demand less $Q_{heating}$, unsurprisingly. Likewise, maximum IOD values are observed in south-facing classrooms. Lastly, the upper bound of obtained CO_2 exceedance is observed strictly with high occupant densities.

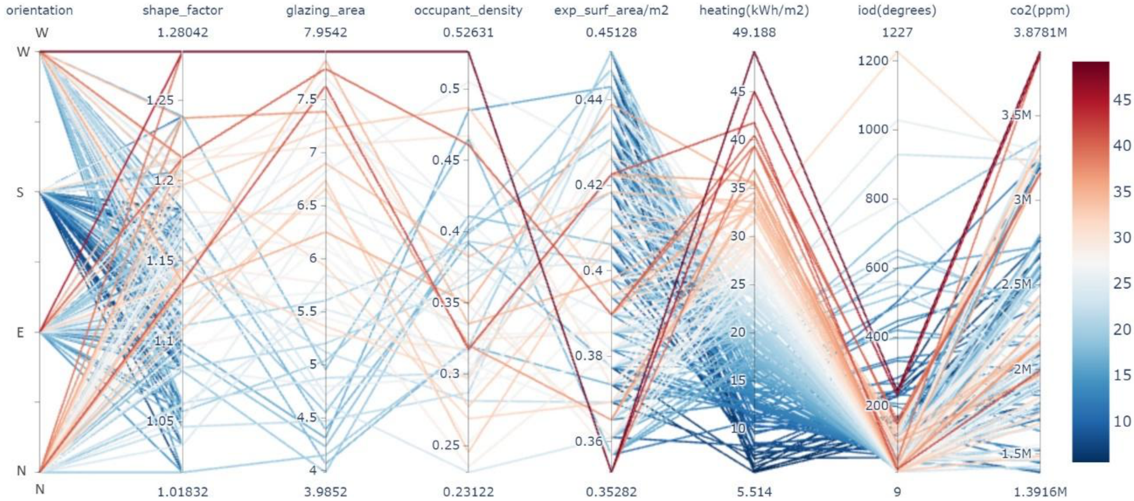

Figure 3: Parallel coordinates graph, colored by $Q_{heating}$.

The correlations among the five input variables and three output variables are calculated (**Figure 4**). Between CO_2 concentrations and occupant density, a very high level of correlation is observed. Exposed surface area per m^2 is negatively correlated with shape factor. $Q_{heating}$ also has a moderate positive relation with the glazing area. IOD does not have a direct correlation with any of the input variables. However, it exhibits weak positive correlations with orientation, shape factor, glazing area, and occupant density. Only between exposed surface area per m2 and shape factor are correlated in input parameters. Output variables of IOD and CO_2 exceedance also have a week positive correlation.

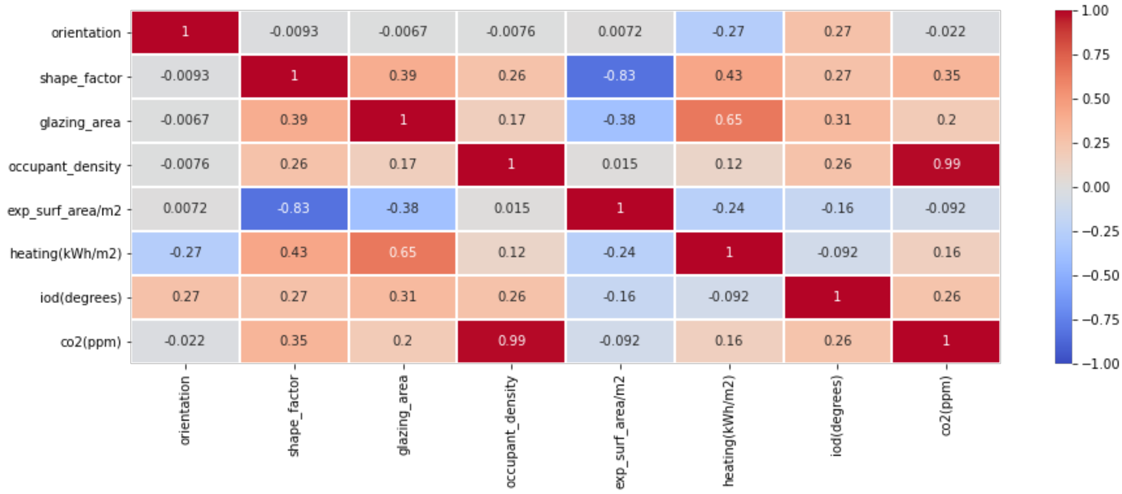

Figure 4: Correlation matrix.

3.2. Model Performances

ANNs are evaluated based on the performance metrics of R^2 and RMSE and MAE errors (**Table 3**). While $Q_{heating}$ and CO_2 concentration are predicted successfully with the average error rate of 4% and 2%, respectively, IOD is predicted with 9% RMSE error. It can be stated that the almost perfect linear relationship between occupant density and CO_2 concentration enabled the CO_2 prediction model to outperform. On the other hand, the IOD prediction model underperforms when compared to the other ANN models. This can be explained with the skewed distribution of data points and the IOD's insignificant linear relationships between input variables. Additionally, it should be reminded that model performances are dependent of the data distribution. Unless containing a missing attribute, none of the observed data is removed from the dataset in this study. However, there are extreme values observed in the data distributions. Those values could be detected as outliers. However, they are left in the datasets since these extreme values are legitimate observations that are a natural part of the population.

Table 3: Prediction performances based on R^2, RMSE, and MAE.

Output Variable	R^2	RMSE	MAE
$Q_{heating}$	0.992	0.762	0.579
IOD	0.986	12.18	8.176
CO_2 Concentration	0.993	38072.179	31699.001

It can be concluded that with the defined and demonstrated framework, selected input variables of occupant density, shape factor, glazing area, exposed surface area per m^2, and orientation are sufficient in predicting performance indicators of CO_2 exceedance and $Q_{heating}$. The RMSEs of specified performance indicators are 38072.179 ppm and 0.762 kWh/m^2, respectively. On the other hand, selected input variables give relatively poor results for predicting IOD. It can be due to skewed data distribution resulting from the nature of design parameters. Therefore, a larger space of input attributes or a higher number of data points has the potential to improve prediction performances together with the tuning of ANN parameters.

3.3. Implications on Classroom Design

Simple design parameters selected and used in this study give insights into classroom spaces' performance. Although ANN is reported to be a black box method due to any insights approximated from the result, preliminary data analysis gives insights for classroom design. It is observed that $Q_{heating}$ is positively correlated with the glazing area. Increased energy loss from the glazing area implies and higher energy demand. Similarly, larger exterior surface area and shape factor indicate greater energy demand. For CO_2 concentration levels, occupant density is a vital parameter. In smaller classrooms, the student capacity should be controlled to maintain indoor air quality. On the other hand, for IOD, a direct relation between inputs and outputs is not evident, and the black box prediction method plays an essential role in predicting IOD accurately.

4. Discussion and Conclusion

Prediction performances of built ANN can be improved with larger data sets. Moreover, with the described framework, input features can be expanded, and the ANNs can be better tuned to achieve more detailed and accurate predictions. This study focuses on only the physical parameters of a classroom space together with the occupant density. Occupants play an essential role in building performance, and further research can be carried on with the inclusion of occupant-centric controls to input space. Furthermore, the selected performance indicators can vary together with the input variables. We have presented a data-driven framework to predict $Q_{heating}$, IOD, and CO_2 concentration levels of classroom spaces with ANN.

We have predicted three performance indicators with the input variables of orientation, shape factor, glazing area, occupant density, and exposed surface area per m^2. After the data exploration, three MLP models are built for each prediction of each output variable. With the selected input parameters, satisfactory results are obtained. The ANNs for predicting $Q_{heating}$ and CO_2 concentration are provided better results than for predicting IOD. The test set was predicted with the average RMSE rate of 4% and 2% concerning the mean of actual values. IOD is relatively poorly predicted with an average error rate of 9%. Model performances can be improved as discussed. The described framework can be applied for the easy and fast assessment of design alternatives during the design development phase or retrofit scenarios.

REFERENCES

Abadi, M, Barham, P., Chen, J., Chen, Z., Davis, A., Dean, J., . . . Zheng, X. (2016). Tensorflow: A system for large-scale machine learning. *12th Symposium on Operating Systems Design and Implementation* (pp. 265–283). https://www.usenix.org/conference/osdi16/technical-sessions/presentation/abadi

Akköse, G., Meral Akgül, Ç., & Gürsel Dino, İ. (2021). Educational building retrofit under climate change and urban heat island effect. *Journal of Building Engineering, 40*, 102294. https://doi.org/10.1016/j.jobe.2021.102294

Anaconda Software Distribution. (2020). *Anaconda Documentation*. Anaconda Inc. Retrieved from https://docs.anaconda.com/

Asadi, E., Silva, M. G., Antunes, C. H., Dias, L., & Glicksman, L. (2014). Multi-objective optimization for building retrofit: A model using genetic algorithm and artificial neural network and an application. *Energy and Buildings, 81*, 2014. https://doi.org/10.1016/j.enbuild.2014.06.009

Ascione, F., Bianco, N., Stasio, C. D., Mauro, G. M., & Vanoli, G. P. (2017). Artificial neural networks to predict energy performance and retrofit scenarios for any member of a building category: A novel approach. *Energy, 118*, 999-1017. https://doi.org/10.1016/j.energy.2016.10.126

ASHRAE. (2019). *ANSI/ASHRAE Standard 62.1-2019 Ventilation for Acceptable Indoor Air Quality.* Atlanta: American Society of Heating, Refrigerating and Air-Conditioning Engineers, Inc.

ASTM. (1998). *Standard Guide for Using Indoor Carbon Dioxide Concentration to Evaluate Indoor Air Quality and Ventilation.* West Conshohocken, PA: ASTM.

Bartlett, K. H., Martinez, M., & Bert, J. (2004). Modeling of Occupant-Generated CO2 Dynamics in Naturally Ventilated Classrooms. *Journal of Occupational and Environmental Hygiene, 1*(3), 139-148. https://doi.org/10.1080/15459620490424393

Bluyssen, P. M., Zhang, D., Kurvers, S., Overtoom, M., & Ortiz-Sanchez, M. (2018). Self-reported health and comfort of school children in 54 classrooms of 21 Dutch school buildings. *Building and Environment, 138*, 106-123. https://doi.org/10.1016/j.buildenv.2018.04.032

Catalina, T., Virgone, J., & Blanco, E. (2008). Development and validation of regression models to predict monthly heating demand for residential buildings. *Energy and Buildings, 40*(10), 1825-1832. https://doi.org/10.1016/j.enbuild.2008.04.001

European Commission. (2020, February 12). *In Focus: Energy Efficiency in Buildings.* Retrieved April 7, 2021 from ec.europa.eu/info/news/focus-energy-efficiency-buildings-2020-feb-

Fabi, V., Andersen, R. V., & Corgnati, S. P. (2013). Influence of occupant's heating set-point preferences on indoor environmental quality and heating demand in residential buildings. *HVAC&R Research, 19*(5), 635-645. https://10.1080/10789669.2013.789372

Fan, C., Sun, Y., Zhao, Y., Song, M., & Wang, J. (2019). Deep learning-based feature engineering methods for improved building energy prediction. *Applied Energy, 240*, 35-45. https://doi.org/10.1016/j.apenergy.2019.02.052

Faustman, E. M., Silbernagel, S. M., Fenske, R. A., Burbacher, T. M., & Ponce, R. A. (2000). Mechanisms Underlying Children's Susceptibility to Environmental Toxicants. *Environmental Health Perspective, 108*(Supplemen 1), 13-21. https://doi.org/10.1289/ehp.00108s113

Gaihre, S., Semple, S., Miller, J., Fielding, S., & Turner, S. (2014). Classroom Carbon Dioxide Concentration, School Attendance, and Educational Attainment. *Journal of School Health, 84*(9), 569-574. https://doi.org/10.1111/josh.12183

Hamdy, M., Carlucci, S., Hoes, P.-J., & Hensen, J. L. (2017). The impact of climate change on the overheating risk in dwellings—A Dutch case study. *Building and Environment, 122*, 307-323. https://doi.org/10.1016/j.buildenv.2017.06.031

Heschong Mahone Group. (2003). *Windows and classrooms: a study of student performance and the indoor environment.* Fair Oaks, CA: Californian Energy Commission.

Kalogirou, S. A., & Bojic, M. (2000). Artificial neural networks for the prediction of the energy consumption of a passive solar building. *Energy, 25*(5), 479-491. https://doi.org/10.1016/S0360-5442(99)00086-9

Keskar, N. S., Mudigere, D., Nocedal, J., Smelyanskiy, M., & Tang, P. T. (2017). On Large-Batch Training for Deep Learning: Generalization Gap and Sharp Minima. *ICLR 2017.* arXiv:1609.04836v2

Kumar, R., Aggarwal, R., & Sharma, J. (2013). Energy analysis of a building using artificial neural network: A review. *Energy and Buildings, 65*, 352-358. https://doi.org/10.1016/j.enbuild.2013.06.007

Kusiak, A., Li, M., & Zhang, Z. (2010). A data-driven approach for steam load prediction in buildings. *Applied Energy, 87*, 925–933. https://doi.org/10.1016/j.apenergy.2009.09.004

Langevin, J., Wen, J., & Gurian, P. L. (2016). Quantifying the human–building interaction: Considering the active, adaptive occupant in building performance simulation. *Energy and Buildings,, 117*, 372-386. https://doi.org/10.1016/j.enbuild.2015.09.026

Mumovic, D., Palmer, J., Davies, M., Orme, M., Ridley, I., Oreszczyn, T., . . . Way, P. (2009). Winter indoor air quality, thermal comfort and acoustic performance of newly built secondary schools in England. *Building and Environment, 44*(7), 1466-1323. https://doi.org/10.1016/j.buildenv.2008.06.014

Neto, A. H., & Fiorelli, F. v. (2008). Comparison between detailed model simulation and artificial neural network for forecasting building energy consumption. *Energy and Buildings, 40*, 2169–2176. https://doi.org/10.1016/j.enbuild.2008.06.013

NREL. (2013). *Advanced Energy Retrofit Guide – K 12 schools.* Department of Energy US. https://www.nrel.gov/docs/fy14osti/60913.pdf

Pérez-Lombard, L., Ortiz, J., & Pout, C. (2008). A review on buildings energy consumption information. *Energy and Buildings, 40*(3), 394-398. https://doi.org/10.1016/j.enbuild.2007.03.007

Pessenlehner, W., & Mahdavi, A. (2003). Building Morphology, Transparence, and Energy Performance. *Eighth International IBPSA Conference .* Eindhoven, Netherlands. http://www.ibpsa.org/proceedings/BS2003/BS03_1025_1032.pdf

Roudsari, M. S., & Pak, M. (2013). Ladybug: A Parametric Environmental Plugin for Grasshopper to Help Designers Create an Environmentally-Conscious Design. *Proceedings of BS2013: 13th Conference of International Building Performance Simulation Association.* Lyon, France. http://www.ibpsa.org/proceedings/bs2013/p_2499.pdf

Seyedzadeh, S., Rahimian, F. P., Glesk, I., & Roper, M. (2018). Machine learning for estimation of building energy consumption and performance: a review. *Visualization in Engineering, 6*(5). https://doi.org/10.1186/s40327-018-0064-7

Shaughnessy, R. H.-S., Nevalainen, A., & Moschandreas, D. (2006). A preliminary study on the association between ventilation rates in classrooms and student performance. *Indoor Air, 16*(6), 465-468. https://doi.org/10.1111/j.1600-0668.2006.00440.x

Shendell, D., Prill, R., Fisk, W., Apte, M., Blake, D., & Faulkner, D. (2004). Associations between classroom CO2 concentrations and student attendance in Washington and Idaho. *Indoor Air, 14*(5), 333-341. https://doi.org/10.1111/j.1600-0668.2004.00251.x

Shields, M. D., & Zhang, J. (2016). The generalization of Latin hypercube sampling. *Reliability Engineering & System Safety, 148*, 96-108. https://doi.org/10.1016/j.ress.2015.12.002

Sterman, M., & Baglione, M. (2012). Design of Artificial Neural Network Using Solar Inputs for Assessing Energy Consumptıon in a High Performance Academic Building. *Proceedings of the ASME 2012 International Mechanical Engineering Congress & Exposition.* Texas, USA. https://doi.org/10.1115/IMECE2012-86976

Tsanas, A., & Xifara, A. (2012). Accurate quantitative estimation of energy performance of residential buildings using statistical machine learning tools. *Energy and Buildings, 49*, 560-567. https://doi.org/10.1016/j.enbuild.2012.03.003

Walker, S., Khan, W., Katic, K., Maassen, W., & Zeiler, W. (2020). Accuracy of different machine learning algorithms and added-value of predicting aggregated-level energy performance of commercial buildings. *Energy and Buildings, 209*. https://doi.org/10.1016/j.enbuild.2019.109705

Wan, K. K., Li, D. H., Liu, D., & Lam, J. C. (2011). Future trends of building heating and cooling loads and energy consumption in different climates. *Building and Environment, 46*(1), 223-234. https://doi.org/10.1016/j.buildenv.2010.07.016

Yalçın, N., Balta, D., & Özmen, A. (2018). A modeling and simulation study about CO2 amount with web-based indoor air quality monitoring. *Turkish Journal of Electrical Engineering and Computer Sciences, 26*, 1390 – 1402. https://doi.org/10.3906/elk-1612-57

Yalcintas, M., & Ozturk, U. A. (2006). An energy benchmarking model based on artificial neural network method utilizing US Commercial Buildings Energy Consumption Survey (CBECS) database . *International Journal of Energy Research, 31*(4), 412-421. https://doi.org/10.1002/er.1232

Yang, L., Yan, H., & Lam, J. C. (2014). Thermal comfort and building energy consumption implications – A review. *Applied Energy, 115*, 164-173. https://doi.org/10.1016/j.apenergy.2013.10.062

Yezioro, A., Dong B., & Leite, F. (2008). An applied artificial intelligence approach towards assessing building performance simulation tools. *Energy and Buildings, 40*(4), 612-620. https://doi.org/10.1016/j.enbuild.2007.04.014

Yu, Z., Haghighat, F., Fung, B. C., & Yoshino, H. (2010). A decision tree method for building energy demand modeling. *Energy and Buildings, 42*(10), 1637-1646. https://doi.org/10.1016/j.enbuild.2010.04.006

Zomorodian, Z. S., Tahsildoosta, M., & Hafezi, M. (2016). Thermal comfort in educational buildings: A review article. *Renewable and Sustainable Energy Reviews, 59*, 895-906. https://doi.org/10.1016/j.rser.2016.01.033