engineering proceedings

MDPI

Proceeding Paper

Optimizing Daylight and Energy Consumption for Climate Change Adaptation: Integrating an Artificial Neural Network Model with a Multi-Objective Optimization Approach †

Nur Nasuha Abd Salam [1,2]*, Simon Lannon [1]

[1] Welsh School of Architecture, Cardiff University, Cardiff CF10 3NB, United Kingdom; Abd.SalamN@cardiff.ac.uk, Lannon@cardiff.ac.uk
[2] Faculty of Civil Engineering and Built Environment, Universiti Tun Hussein Onn Malaysia, Batu Pahat 86400, Malaysia.
* Correspondence: Abd.SalamN@cardiff.ac.uk
† Presented at the 1st International Online Conference on Buildings, 24-26 October 2023.

Abstract: Machine learning models have been proven for their capability to improve the computational efficiency of building performance simulations. However, studies on their reliability to produce Pareto front solutions for multi-objective optimization are limited, particularly for climate adaptation studies. This study proposed a dependable workflow to integrate an artificial neural network (ANN) model with energy consumption and daylight multi-objective optimization for climate change adaptation. The trained ANN model attained high R^2 scores with RMSE scores of 2.23 and 4.52 for UDI and cooling EUI respectively. Statistical hypothesis analysis of the Pareto front solutions produced by conventional simulation-based and ANN-based optimization shows that the two models have no significant difference indicating the reliability of the proposed workflow.

Keywords: Neural Network; Multi-Objective Optimization; Climate Change; Daylight; Energy

Citation: Salam, N.N.A.; Lannon, S. Optimizing Daylight and Energy Consumption for Climate Change Adaptation: Integrating an Artificial Neural Network Model with a Multi-Objective Optimization Approach. **2023**, *5*, x. https://doi.org/10.3390/xxxxx
Published: 24 October

1. Introduction

Most studies aimed at enhancing building efficiency and crafting climate-resilient design have employed parametric analyses. Nevertheless, the intricate interactions among design variables and conflicting objectives of realistic building requirements limit the maximal potential solutions achievable through this approach [1]. Multi-objective optimization (MOO) has garnered considerable attention within the realm of building performance simulations due to its capacity to pinpoint a collection of Pareto-optimal solutions, embodying the delicate balance among numerous objectives [2]. However, most optimization studies have previously concentrated on enhancing building efficiency under existing climate conditions [1].

One of the challenges of adopting MOO for analyzing building performance is high computational demands. The computational intensity escalates when working with higher resolutions, larger and more intricate scenes, and heightened levels of detail, particularly when it involves energy modeling and daylight ray-trace calculations [2]. To alleviate the computational challenges associated with optimizing daylight and energy performance, the utilization of machine learning surrogate models holds significant promise. Surrogate modelling captures the intricate relationships between design variables and objective functions, enabling faster evaluations of design alternatives in contrast to resource-intensive simulations. In the context of building simulation, artificial neural networks (ANN) represent a commonly employed algorithm for predicting building performance due to their ability to capture the nonlinear aspects of building behaviour [3].

A study conducted by Han et al. introduced a workflow for constructing an ANN surrogate model to forecast daylight performance within an office space. This novel modelling approach yields highly accurate daylight predictions while operating at a remarkable speed, performing 250 times faster than conventional Radiance simulation tools [4]. Additionally, Lu et al. harnessed the power of a generative adversarial network to predict both daylight and thermal comfort within a commercial building located in Tokyo. The conventional optimization method consumed 1500 hours, while the new procedure completed the task in 105 hours of simulation time, representing a 14-fold increase in efficiency over the traditional method [5]. Although the application of the machine learning model has been extensively studied in the realm of building simulation and MOO, there remains a notable gap in its application in the context of future climate considerations. To address the gap and challenges discussed above, this study proposes a dependable workflow to integrate an ANN-based model with multi-objective optimization (ANN-MOO) for climate change adaptation.

2. Materials and Methods

2.1 Parametric Model and Synthetic Data Generation

In this study, the ANN surrogate model was developed using a synthetic database based on a parametric model of a double-storey terrace house in Kuala Lumpur, Malaysia. Specifically, the study focused on the ground floor's open dining and living area, which typically suffers from low daylight levels due to the installation of large overhangs aimed at mitigating solar heat gain. Given the conflicting nature of daylight and solar heat gain in a hot and humid climate, this study aimed to explore alternative optimized façade designs. The investigation primarily centered on design parameters that are difficult to modify or considered fixed elements to avoid costly and disruptive changes for climate adaptation. Table 1 provides a summary of the investigated design parameters, their ranges, and the precision employed to construct the synthetic database for training the ANN surrogate model.

Table 1. Range and steps of the input data used for the machine learning modelling.

Parameters	Min.	Max.	Steps	Counts	Explanation
RO	0	359	1	360	The angle of the room starting from South (^{0})
RW	3.0	10.0	0.1	70	Width of room (m)
RD	3.0	10.0	0.1	70	Depth of room (m)
RH	3.0	5.0	0.1	20	Height of room (m)
GR	0.1	0.9	0.1	9	The window-to-wall ratio of the front façade
GT	0.1	0.9	0.1	9	Transmissivity of the window
SD	0.0	5.0	0.1	60	Depth of the overhang at the front façade (m)
ELSD	0.0	1.0	0.1	11	Depth of external light shelf (m)
ILSD	0.0	1.0	0.1	11	Depth of internal light shelf (m)
The total number of design alternatives:					5.19e+16

To enhance workflow reproducibility while managing time constraints, the Latin Hypercube Sampling (LHS) method was employed. LHS strikes a balance between simple random and stratified sampling techniques, ensuring a uniform sample distribution [6]. The LHS procedure was performed using Python 3.0 with the SALib module version 3.0 [7]. Subsequently, the samples were simulated using Radiance and OpenStudio version 4.5 by The Ladybug Tool version (1.5.0). The simulations were conducted using future climate data for Kuala Lumpur in the 2080s generated by the CCWorldWeatherGen tool [8].

To optimize building performance, this study investigated two daylight-related metrics, Useful Daylight Illuminance (UDI) and Uniformity Ratio (UR), along with three energy-related metrics: lighting Energy Use Intensity (LTE), cooling Energy Use Intensity (CLE), and solar gain Energy Use Intensity (SGE), all expressed in kWh/m^2/year. UDI was categorized into three levels: acceptable (UDIa, 100-2000 lux), supplementary (UDIs, <100 lux), and excessive (UDIe, >2000 lux). Annual simulations were conducted for all metrics, yielding a single average value per metric due to constraints imposed by the optimization tool, which allows only one value per fitness objective.

2.3 Model training and validation

A data-driven approach was used to obtain a satisfactory number of samples where an additional 200 samples were simulated; a new model was trained. The final amount of data that was satisfactory for the training of the surrogate model was stopped at 2000 samples. These data samples were checked for null values or invalid data and were normalized using the min-max scaler in the Scikit Learn module by Python [9]. The min-max scaler scales and translates the data between zero to one range. The normalized data were then split into 3 datasets which were the training set, the validation set and the testing set (unseen data) with the ratio of 60%, 20%, and 20% respectively.

Following data normalization and processing, the next step is to train the model using the ANN algorithm using the training dataset. The proposed structure of the ANN surrogate model is composed of 1 input layer with 11 neurons (design parameters), N hidden layers with M neurons, and 1 output layer with 1 neuron (output metrics). The neural network uses Adam Solver, an algorithm designed to optimize stochastic objective functions using first-order gradients [10].

The performance of an ANN heavily depends on the values chosen for hyperparameters. Optimization of hyperparameters in this study was performed using a 5-fold (k) cross-validation technique on the validation dataset. The resultant optimized hyperparameters for this study were different for each metric. Finally, the performance of the trained model against the unseen dataset (testing dataset) was validated with mean absolute error (MAE), root mean square error (RMSE) and coefficient of determination (R^2). Apart from evaluating the accuracy of the model, the latency (in milliseconds) or speed of prediction was also considered as one of the model performance metrics which correlates to the reduction of simulation time during the deployment stage.

2.4 Model deployments

The model that demonstrated the best performance in the training stages was integrated and imported into the Grasshopper interface through GH-C Python in the form of pickle files. A Python script was created in GH-C Python to build a Grasshopper custom component that represents the ANN surrogate model which receives design parameters as the input and produces predicted daylight and energy metrics as its output. It's crucial to emphasize that the range and precision of the input (design parameters) must align with those used during the model training to guarantee the precision of the projected daylight and energy performance.

The subsequent steps mirrored the conventional optimization process, where two essential sets of data were required from the optimizer tool: the Genes and the Objectives. The difference between the ANN-MOO workflow with the conventional method is the predicted daylight and energy performances from the surrogate model were connected as the input for the Objectives. The multi-objective optimization was performed using the Grasshopper plug-in Wallacei version 2.5 [11] which employs the Non-dominated Sorting Genetic Algorithm II (NSGA-II). Three validation methods were employed to assess the surrogate model's reliability in predicting Pareto front solutions: Pareto front analysis, statistical hypothesis analysis, and computational efficiency analysis. These analyses compare results from the conventional optimization method to the ANN-MOO workflow.

3. Results and Discussion

This section summarizes and discusses the results of model training and deployment of the study. Figure 2 shows the regression plot of each surrogate model for present and future models. In general, the R^2 for all plots were in the range of 0.786 to 0.997 which suggests that the models capture a significant portion of the target variable's variance and provide reliable predictions. The plots also depict the results of RMSE and MAE of each surrogate model on an unseen dataset. The observations of MAE for UDI metrics indicate that the model predictions deviate from the actual value by 1.22 to 1.638. Han et al. reported a similar observation in their experiment where the MAE error was between 0.79 to 1.75 for the UDI metric [4].

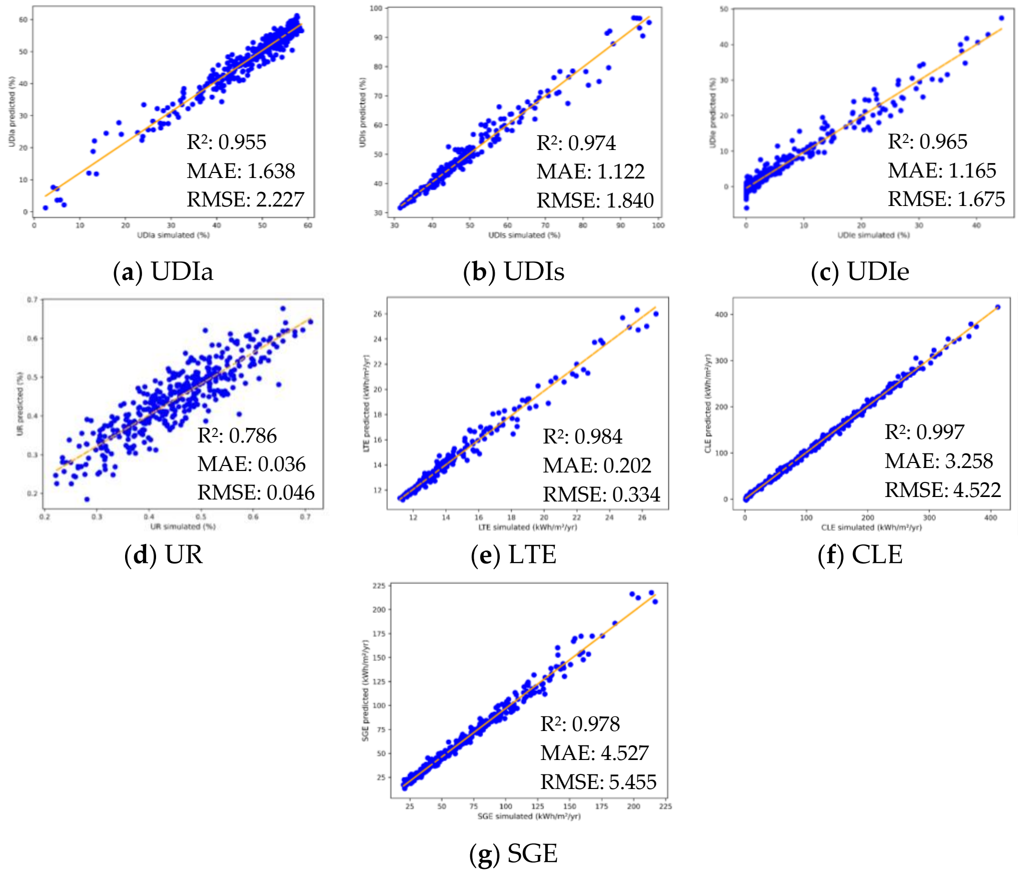

Figure 2. The regression plot of daylight and energy metrics.

Figure 3 depicts the 2D Pareto fronts graph for 7 objectives arranged in a 3 by 4 matrix. The y-axis represents energy-related metrics while the x-axis represents daylight-related metrics. The figure compares the Pareto front solutions produced by the simulated model (blue markers) and the surrogate model (red markers). The purpose of this graph is to analyze the agreement in terms of the distribution of the Pareto front solutions generated by the two models The graph depicts an agreement between the two models where in most cases the trend distribution relatively overlapped each other.

To quantify the agreement between the simulated model and the surrogate model, statistical hypothesis analysis using a two-tailed t-test was conducted. The t-test calculates the significant difference (p-value) between the means of the Pareto front's fitness value produced by the simulated model and the surrogate model. In this test, the means of the two models indicate no significant difference if the p-value is equal to or greater than 0.05 ($\alpha \geq 0.05$) [12], which is the aim of this analysis. The result of the t-test in Table 3 showed that the significance values for all performance metrics were found to be greater than 0.05. This indicates that there is no significant difference between the means of the two models further emphasising the ability of the ANN-MOO workflow to represent the behaviour of the conventional optimization method.

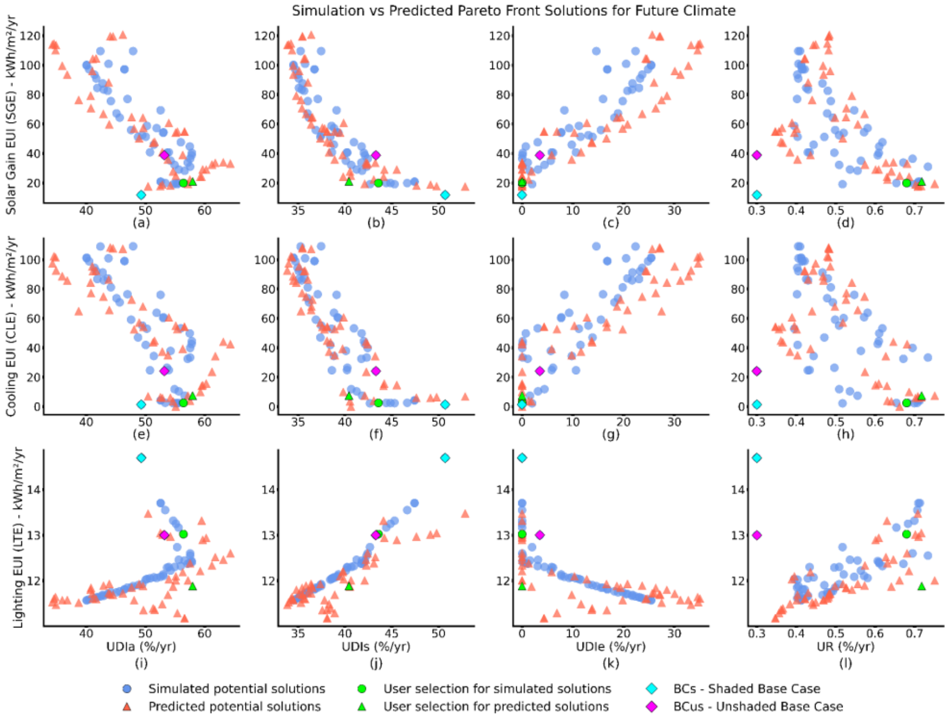

Figure 3. The comparison of Pareto front solutions produced by simulation and surrogate model for the future climate.

Table 3. Statistical comparison between Pareto front solutions produced by the simulated model and surrogate model using a two-tail t-test.

Daylight				Energy			
Metrics	**P-Value**	**Metrics**	**P-Value**	**Metrics**	**P-Value**	**Metrics**	**P-Value**
UDIa	0.810	UDIe	0.135	LTE	0.052	SGE	0.721
UDIs	0.834	UR	0.920	CLE	0.961		

The computational efficiency of the ANN surrogate model was evaluated to determine the efficiency of the proposed ANN-MOO workflow. The evaluation time per solution for the conventional method was approximately 6 minutes, while the surrogate model remarkably reduced this time to just 7 seconds. As a result, the entire simulation process for the conventional method took 213.3 hours while the ANN-MOO method took 4.6 hours to achieve Pareto front solutions. Specifically, the ANN-MOO method was 46.2 times faster (97.8% simulation time decrease) than the conventional method. In comparison, the GPU-accelerated simulation demonstrates a speed boost of only up to 10 times for annual simulation [13].

4. Conclusions

Multi-objective optimization tools commonly used for optimizing daylight and energy performances in the early design stage are often slow and time-consuming. To effectively employ these design tools, this study proposes an ANN-based alternative approach to optimized daylight and energy performance for climate adaptation strategies.

The results from model testing indicated that the ANN-based surrogate model can predict daylight and energy performance with a high level of accuracy for the future climate, leading to dependable deployment for multi-objective optimization. The statistical hypothesis analysis enhances the confidence of the proposed workflow to predict Pareto front solutions. The findings of this study also demonstrated the substantial time-saving advantage offered by the surrogate model, making it an efficient alternative for multi-objective optimization tasks.

For future work, this workflow could be employed in other climates or building typologies by updating the input samples for the training of the surrogate model to the required climate weather files and design parameters. Other performance metric could be explored such as thermal comfort for a more comprehensive understanding of the design problem. Every investigation into the influence of climate change on buildings inherently involves various uncertainties resulting from different sources, including the selection of global climate models, emissions scenarios, and downscaling methods. Additional research is recommended to explore how various climate change scenarios affect daylight and energy optimization.

Author Contributions: N.N.A.S played a key role in the conceptualization of the study, conducted the performance simulations and optimization, trained and validated the ANN models, analyzed, and interpreted the results and wrote the original draft for publication. S.L. provided expertise on multi-objective optimization and development of machine learning models, contributed to the research design, reviewed the methodology and data analysis, and provided guidance throughout the study. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Ministry of Higher Education Malaysia and the Universiti Tun Hussein Onn Malaysia, under the 'Skim Latihan Akademik Bumiputera (SLAB)' grant scheme.

Institutional Review Board Statement: Not applicable

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

References

1. K. Bamdad: M. E. Cholette, S. Omrani, and J. Bell. Future energy-optimised buildings — Addressing the impact of climate change on buildings. *Energy Build* **2020**, *vol. 231*.
2. F. Yu, R. Wennersten, and J. Leng. A state-of-art review on concepts, criteria, methods and factors for reaching 'thermal-daylighting balance. *Building and Environment* **2020**, *vol. 186*.
3. M. Ayoub. A review on machine learning algorithms to predict daylighting inside buildings. *Solar Energy* **2020**, *vol. 202*.
4. Y. Han, L. Shen, and C. Sun. Developing a parametric morphable annual daylight prediction model with improved generalization capability for the early stages of office building design. *Build Environ* **2021**, *vol. 200*.
5. Y. Lu, W. Wu, X. Geng, Y. Liu, H. Zheng, and M. Hou. Multi-Objective Optimization of Building Environmental Performance: An Integrated Parametric Design Method Based on Machine Learning Approaches. *Energies (Basel)* **2022**, *vol. 15*.
6. Z. Pang, Z. O'Neill, Y. Li, and F. Niu. The role of sensitivity analysis in the building performance analysis: A critical review. *Energy and Buildings* **2020**, *vol. 209*.
7. T. Iwanaga, W. Usher, and J. Herman. Toward SALib 2.0: Advancing the accessibility and interpretability of global sensitivity analyses. *Socio-Environmental Systems Modelling* **2022**, *vol. 4*.
8. M. F. Jentsch, A. B. S. Bahaj, and P. A. B. James. Climate change future proofing of buildings-Generation and assessment of building simulation weather files. *Energy Build* **2008**, *vol. 40*, page 2148–2168.
9. Pedregosa et al. Scikit-learn: Machine Learning in Python. Available online: http://scikit-learn.sourceforge.net (accessed on 21 June 2023).
10. D. P. Kingma and J. Ba. Adam: A Method for Stochastic Optimization. Available online: http://arxiv.org/abs/1412.6980 (accessed on 8 September 2023).
11. Wallacei Primer 2.0. Available online: https://www.wallacei.com/ (accessed on 15 May 2023).
12. E. M. Abdelkader, A. Al-Sakkaf, and R. Ahmed. A comprehensive comparative analysis of machine learning models for predicting heating and cooling loads. *Decision Science Letters* **2020**, *vol. 9*, 409–420.
13. N. L. Jones and C. F. Reinhart. Speedup potential of climate-based daylight modelling on GPUs. In Building Simulation Conference Proceedings, International Building Performance Simulation Association, 2017.