CSCE 2022 Annual Conference
Whistler, British Columbia
May 25th – 28th, 2022

Implementing surrogate modeling techniques for designing optimal building envelops: A case study

Shahrzad Monshet[1]*, Thomas M. Froese[1], Ralph Evins[1]
[1]Depeartment of Civil Engineering, University of Victoria, Victoria, BC, Canada
* Corresponding author, monshet@uvic.ca

Abstract: Buildings are known to have significant environmental impacts.The life cycle approach to measuring CO_2 emission and life cycle costs of buildings is getting more important in the building design process. However, due to the complexity of the design process and the computational time of simulations and data processing, such methods are difficult to implement within optimization processes. This paper aims to apply surrogate modeling techniques as a solution to resolve the computational difficulties in the optimization process of building envelopes. The paper will describe the methods applied and will evaluate several aspects of the process, including the impact of the size of the training set on the prediction accuracy and the impact of different energy system efficiencies on the final optimum envelope design concerning seven objectives related to the economic and environmental performance of the building.
The results showed that the size of the sampling test has a significant effect on the prediction accuracy; however, a balance between increasing the precision and computational time can be found to select an adequate number of samples. Moreover, it is found that to achieve the lowest total equivalent cost corresponding to the highest economic and environmental performance of the building, the minimum allowed the window to wall ratio (15 %) and the maximum permitted wall insulation thickness (0.02 m) is realized to be the final optimal solution and, therefore, recommended in design. The surrogate model was also shown to be efficiently capable of finding the optimum results according to the other objectives, including both economic and pure environmental aspects. Furthermore, the results provide insightful information on how the variation of energy systems' efficiency might affect the optimum solutions in the optimization process.

Keywords: Optimization, Surrogate models, Building Envelope, Economic and environmental performance

Acronyms

COP	Coefficient of Performance
EER	Energy Efficiency Ratio
EPD	Environmental Product Declaration
GWP	Global Warming Potential
LCA	Life Cycle Assessment
MARE	Mean Absolute Relative Error
MAE	Mean Absolute Error
MSE	Mean Squared Error
RMSE	Root Mean Square Error
R^2	Coefficient of determination

1 INTRODUCTION

Buildings play a significant role regarding energy consumption and environmental impact. It is stated that 40 percent of energy consumption in Europe is attributed to the building sector. This fact and the need to reduce the impact make the building sector an essential target to focusing on the environmental energy efficiency program. (Aste et al., 2016) (Amini Toosi et al., 2020)

The main focus for reducing energy consumption and improving the environmental performance of buildings has been on the operational phase, although more comprehensive methods such as life cycle approaches are now becoming established to design sustainable buildings. The most common methodology is Life Cycle Assessment (LCA) which aims to evaluate the environmental impact of a building within its whole life cycle from the cradle to the grave. One of the most used and agreed on environmental impact categories in LCA is the global warming potential (GWP), a factor established to allow comparisons of the global warming impacts of different assets. In other words, it is an indicator to measure the potential of a product or process to emit CO_2 equivalent to the environment for each functional unit of the asset under evaluation (Zieger et al., 2020)

One of the difficulties of calculating the GWP is the massive amount of data and the extensive time required for the computational analysis (Amini Toosi et al., 2021). We face a vast complexity in the building design scope due to many data and several variables and the complication of the design process and data analysis. Therefore, some automated techniques like machine learning can be beneficial at this point. The help of automation of analytical computation makes the simulation process and analysis of the building performances faster, preserving its accuracy (Westermann & Evins, 2019) (Amini Toosi et al., 2022).

Computational analysis can be processed by using a surrogate model, which is a function that prepares the detailed simulation models. Surrogate models, or meta-models, are promising to provide building performance assessment that is physical knowledge-based but much faster than simulation-based design analysis. The idea of surrogate modeling is to rival an expensive high-fidelity model, in the case of this paper, a building simulation model using a statistical model. Only a small set of simulation data (inputs and outputs) are needed to train the surrogate model. However, the synthetic data's reliability is based on the accuracy of the simulation program itself, and the range of error provided by the surrogate model is directly associated with the provided information for training the surrogate model. (De Wilde, 2014)

Therefore, it is recommended to take a sufficient amount of data. In the end, when the model is validated to approximate the detailed simulation model well enough, it can be used to almost instantly predict outcomes of the high-fidelity simulation given an appropriate set of building design information

(Westermann & Evins, 2019). The surrogate model can be beneficiary in four stages of the building design process: 1. Conceptual design stage, 2. Sensitivity analysis, 3. Uncertainty analysis, 4. Optimization.

In this paper, we concentrate on the first stage. In the early design phase, the designer needs to constantly change their proposal and see how it affects the total concept. Surrogate modeling can give design feedback in much less time than using simulation-based parametric analysis. It can also have a fast analysis of the impact of design decisions on design variability.

Surrogate model derivation involves the following steps. First, a problem should be defined to target the whole approach of the process. Then a design base model should be implemented by the designer, and the samples should be generated using different samples strategies (it can be both from statistics or adaptive sampling). Each variable defined through the samples would be simulated to create a stored database with inputs and outputs, so that in the next step, a surrogate model could be fitted to this data. Lastly, the model is validated by computing the model precision. It measures the deviation of surrogate predictions from simulation outcomes for the same set of inputs. (Westermann & Evins, 2019)

This paper examines the importance of window-to-wall ratio and wall insulation thickness in the energy and environmental performance of buildings (Troup et al., 2019) (Annibaldi et al., 2019). The first objective is to realize a surrogate energy model and to implement it for achieving the optimum solution for designing a building envelope considering the exterior wall insulation and the window-to-wall ratio of the building. Second, the study aims to define the adequate sample size of the input to the surrogate model by testing multiple statistical indices, including Mean Absolute Error (MAE), Mean Squared Error (MSE), Root Mean Square Error (RMSE), Coefficient of determination (R^2), and Mean Absolute Relative Error (MARE). The third goal is to find the optimum results considering the economic and environmental aspects of the building, both regarding the operational and embodied impacts.

2 METHODOLOGY

The thermal test zone that is considered in this paper is a cube with the dimensions of 3m*3m*5m located in Milan, Italy. The building is analyzed assuming a life span of 30 years. A window is located on the south side of the building. It is a UPVC double glazed window with a thermal transmittance (U) value of 1.3 w/m2k. The opaque wall is a cement prefabricated wall whit EPS insulation. A code in python will evaluate the energy consumption by executing an Energy Plus simulation.

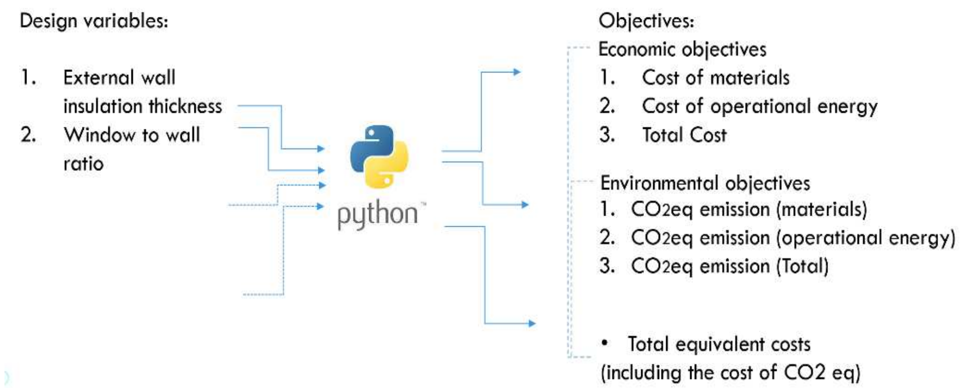


Figure 1. Design variables and optimization objectives.

The variables defined in this case are the window to wall ratio (0.15-0.9) and the thickness of the external wall insulation (0.05m-0.2m). The goal is to determine the optimum combination of the variables to rich the minimum value of the objectives, which is possible with the help of python code to train the surrogate modeling by defining a sample set.

The objectives are defined based on three different categories. The first one is the economic cost. In this group, the results can be calculated based on the cost of the materials only, the cost of energy consumption during the operation phase of the building, or a sum of both. The other category is environmental objectives. In this group, the target is to reach the minimum GWP of materials, operational energy, and the sum of both embodied and operational impacts. The seventh and the last objective is the total equivalent cost, which combines the total cost and the total GWP that is converted to cost, based on the European standard. It is established that one kilogram of CO_2 is equal to 0.02 euros. (Ristimäki et al., 2013). Figure 1 represents the schematic concept of the analysis showing the python code as the computational engine, design variables, and optimization objectives.

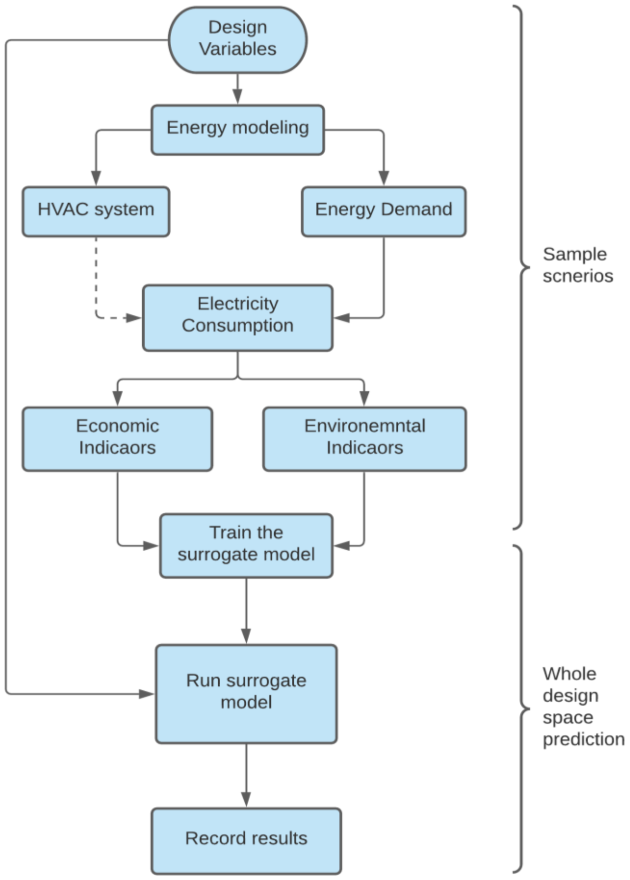


Figure 2. The surrogate modelling steps

As shown in figure 2, the analysis starts with the definition of the design variables. The parametric energy analysis is carried out using the BESOS library to run energy plus parametrically for producing the test zone's heating and cooling energy demand for a set of combination of design variables as the sample set. The sample set of design variables is produced randomly using python. Then having the energy demand of the sample set, the electricity consumption of each design scenario will be calculated by implementing the HVAC system's efficiency. In his case, electric heat pumps are considered to provide heating and cooling energy, the baseline efficiency of the heat pump units in heating mode (COP) and cooling mode (EER) are considered equal to 3 and 2, respectively.

To calculate the energy consumption of the design scenario based on energy system efficiency, a set of python codes are written and linked to the BESOS codes to run energy plus automatically.

Then, the electricity consumption and the quantity of the material are converted to economic and environmental impacts using the economic data such as the materials and energy prices and environmental life cycle inventories such as Ecoinvent and Environmental Product Declaration (EPD) to calculate the GWP of materials. (Ecoinvent, 2021) (Epd, 2021). The seven objectives as separated indicators for each design scenario are calculated using python codes provided for this analysis.

In the final step, the calculated indicators are used to train the surrogate model, and the surrogate model will predict the results of the indicator for the whole design space, including all possible (allowed) combinations of design variables. The surrogate models' results can then be used to find the optimum scenarios in which the economic and environmental indicators are minimized, corresponding with the low economic cost and environmental impacts.

Given the steps of training and implementing the surrogate models, the analysis carried out in this paper will first aim to realize the minimum size of the training set (samples) using the error metrics such as MAE, MSE, RMSE, R2, and Mean Absolute Relative Error (MARE). To do so, different training set sizes, including

10, 15, 20, 25, and 30 samples, are analyzed. Then, using the minimum training set size to achieve accuracy over 99 percent (MARE < 1%), the surrogate model predicts the results separately for the seven objectives shown in figure 1. Then the impact of different heat pump efficiency on the results will be analyzed by altering the heat pump efficiency. Therefore, the model will be implemented several times with different energy system efficiency.

3 RESULTS

Increasing the size of the training set will increase the computation time and is expected to increase the model accuracy by providing a more comprehensive training input to the model. The results showed that increasing the size of the training sample set from 5 to 25 affects the accuracy significantly. For the cases with more than 25 samples, the results accuracy is then less sensitive to the size of the training set, as shown in figure 3 and table 1.

Figure 3, The Mean Absolute Relative Error (MARE) for the predicted results with different sizes of the training set

Table 1 The error metrics for the predicted results by different sizes of the training set

	5 samples	10 Samples	15 Samples	20 Samples	25 Samples	30 Samples
MARE	3.37	1.66	1.27	1	0.88	0.83
MAE	247	128	98	77	69	65
MSE	67584	23854	13449	7684	7314	6703
RMSE	259	154	115	87	85	81
R^2	0.963	0.989	0.994	0.996	0.996	0.996

Figure 4 represent the actual (simulation-based) and predicted values by the surrogate model for different size of the training set, which shows the actual and predicted values become closer while increasing the size of the training set.

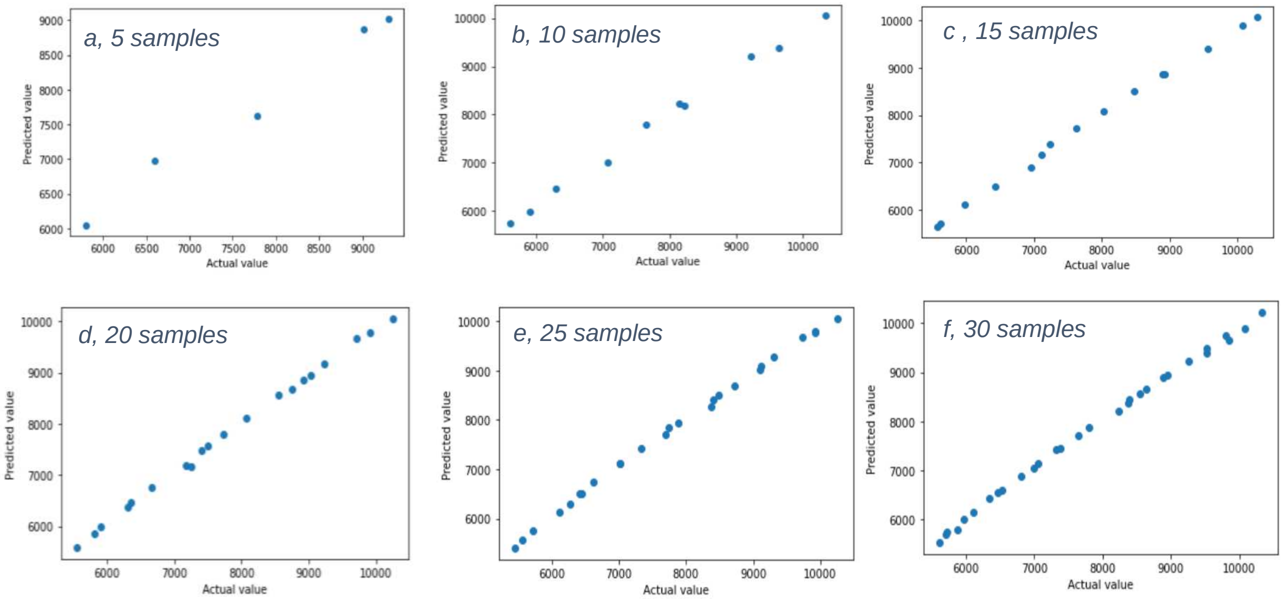
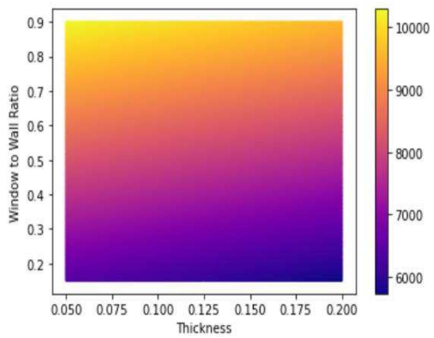
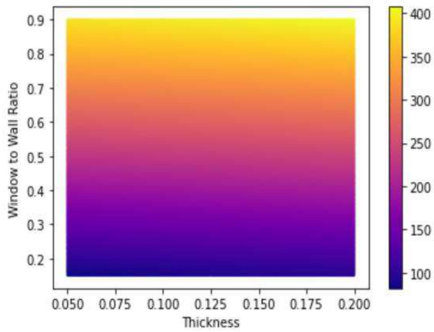


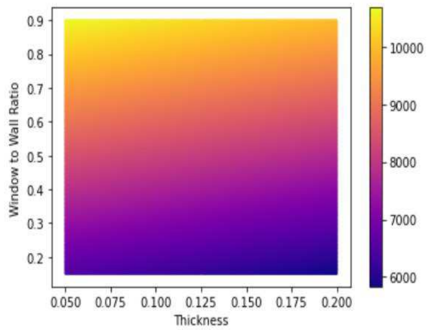
Figure 4, Actual versus predicted results by the surrogate model with different size of training set



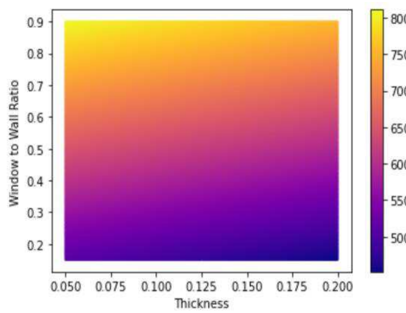
a, GWP of operation energy (kg CO_2 equivalent)



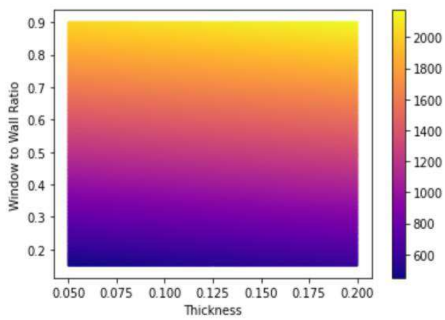
b, GWP of materials (kg CO_2 equivalent)



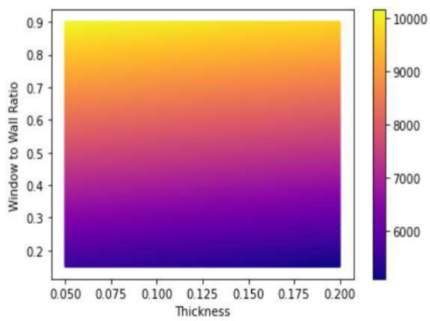
c, total GWP (kg CO_2 equivalent)



d, Cost of operation energy (Euro)



e, Cost of materials (Euro)



f, total Cost (Euro)

Figure 5, The predicted results by the surrogate model for GWP and the cost of materials and energy consumption. Colors indicate the life cycle cost in Euro (d,e,f) and intensity of life cycle kg CO_2 eq emission (a,b,c) of the building considering different combination of design variables.

A training set with a size of 25 samples is found to be accurate enough to predict the results with a MARE less than 1 percent; then it is used to predict the seven objectives described before for a case study in which the COP and EER of heat pump units are considered to be equal 3 and 2 accordingly. As shown in Figures 5 and 6, the results show that according to the cost and GWP of materials, the best scenario are those with the lowest allowed value of design variables, which are 15 percent of the window to wall ratio and 0.05 m of wall insulation thickness. Although regarding the cost and GWP of operational energy consumption (heating and cooling), the maximum size of wall insulation thickness (0.2 m) and the minimum window to wall ratio (15 %) are found to be the optimum solution since it significantly decreases the heating energy consumption as the most critical energy demand in Milan.

Given the difference between the price and environmental impacts of the material and energy, the optimum results based on the total cost and GWP are found to be achieved while the maximum insulation thickness and the minimum window to wall ratio are installed. These results indicate that the final results are affected mainly by the cost and impact of energy consumption.

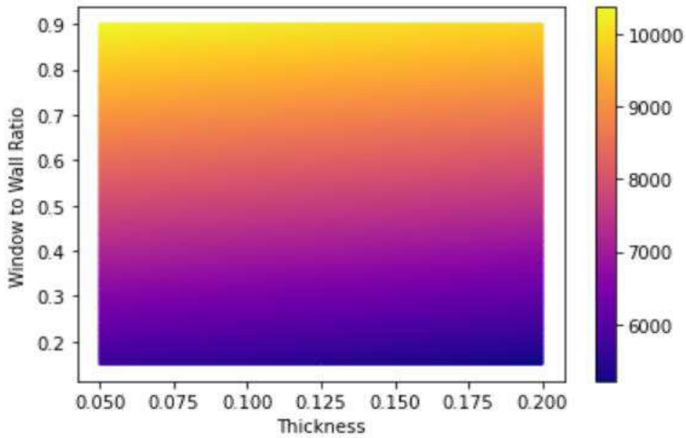


Figure 6 The predicted results for the total equivalent cost (Euro). Colors indicate the total equivalent cost in Euro of the building considering different combination of design variables.

Figure 6 represents the optimum results for the 7th objective in which the CO_2-equivalent emission is monetized by a factor of 0.02 Euro per kg CO_2 equivalent. Therefore, it represents the total equivalent cost as the optimization objective and shows that the optimum solution as expected is the minimum allowed window-to-wall ratio and the maximum allowed insulation thickness.

Finally, the results showed that in case studies with higher energy systems' efficiency, the impact of cost and embodied GWP of materials will get a higher share in the total equivalent cost due to decreasing the operational energy consumption. However, the final optimum results will remain affected mainly by the operational energy performance of the buildings.

Figure 7 illustrates the total cost of different cases with different COP and EER of heat pump units. As is shown in all four scenarios, the final optimum solution based on the predicted results by the surrogate model is similar and is achieved at the minimum allowed window to wall ratio and the maximum allowed wall insulation thickness. However, increasing the efficiency of energy systems will decrease the impact of operational energy consumption on the total cost and environmental impact. Therefore, it is found that for the buildings with high energy efficiency, the total impact of the building is being shifted toward the embodied impacts. For instance, in this case study, we showed that the highest wall insulation thickness provides better performance in terms of total equivalent cost since it significantly decreases the heating energy consumption.

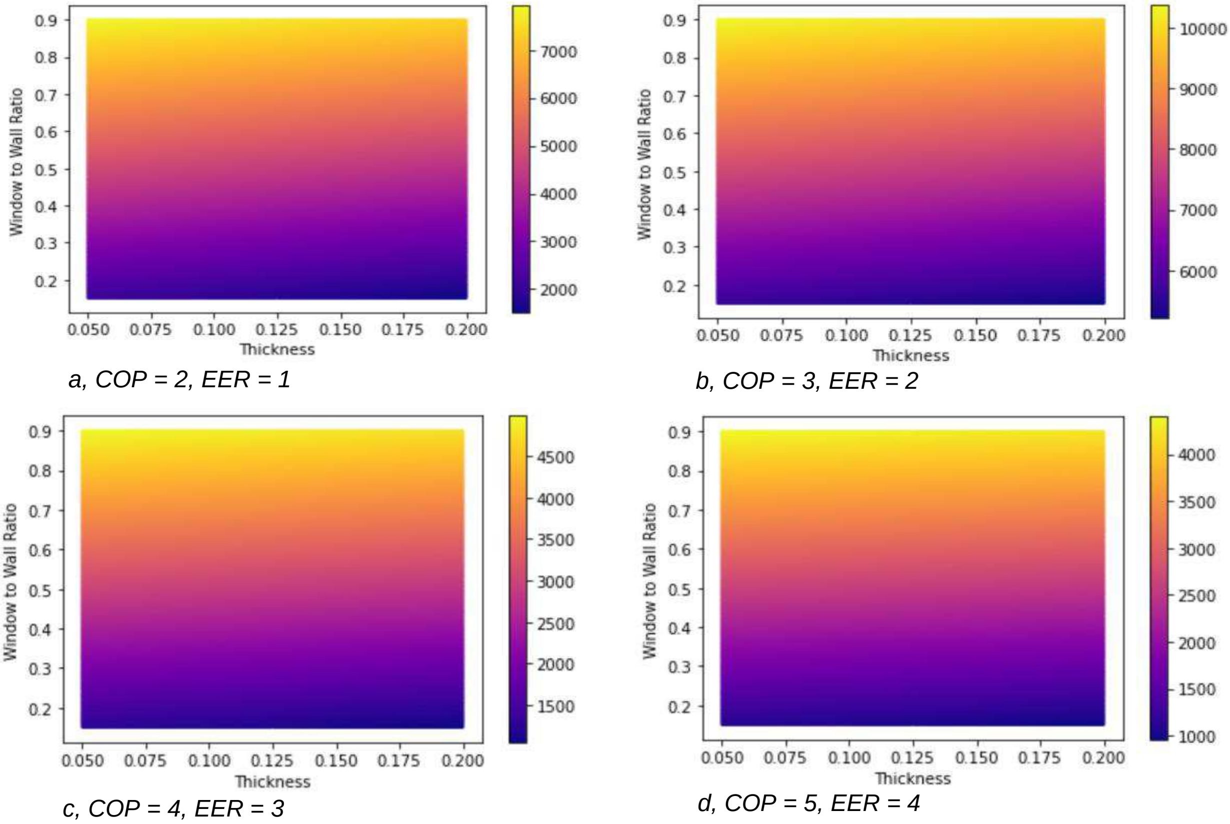


Figure 7, The predicted results for the cases with different heat pump efficiency.). Colors indicate the total equivalent cost in Euro of the building considering different combination of design variables.

While increasing the energy system's efficiency, the difference of indicator for a total equivalent cost between the lowest and highest wall insulation thickness will decrease noticeably. As a piece of evidence, the results showed that if the COP and EER of the heat pump units improved from 2 and 1 to the values equal to 5 and 4, the difference of the total equivalent cost for the lowest and highest wall insulation thickness will change from 37% to 15%, which indicates that the benefits of installing thicker wall insulation will decrease gradually by increasing the efficiency of the energy system.

4 CONCLUSIONS

This paper implements surrogate modeling techniques to design the optimum scenarios of building envelopes. We analyzed the effect of the size of the training set on the prediction accuracy by analyzing different error metrics by which we found that a training set with a minimum size of 25 samples in this case study will provide accuracy over 99 percent. Moreover, we implement the model to find the optimum design scenarios based on seven objectives related to the economic and environmental performance of the buildings. We concluded that in order to achieve the lowest total equivalent cost corresponding to the highest economic and environmental performance of the building, the minimum allowed window to wall ratio (15 %) and the maximum allowed wall insulation thickness (0.02 m) was found to be the final optimal solution and therefore recommend in design. Furthermore, by altering the energy system's efficiency, we showed that in the cases with the highest efficiency of energy systems, the final results tend to be more oriented towards embodied impacts of materials rather than the operational energy consumption. However, the operation energy consumption remains the main cause of total cost and environmental impacts in this case study.

The results in this paper also indicate and reaffirm that the share of embodied impacts in high-energy-efficient buildings such as net-zero energy buildings will become of paramount importance in achieving economic and environmental sustainability in buildings.

As future research, implementing this study under different energy price and decarbonization scenarios of the national electricity grid can be interesting to realize how different strategies for future electricity mix decarbonization and different energy prices can impact the optimum sustainable building design economically and environmentally.

References

Amini Toosi, H., Lavagna, M., Leonforte, F., Del Pero, C., & Aste, N. (2020). Life Cycle Sustainability Assessment in Building Energy Retrofitting; A Review. *Sustainable Cities and Society*, *60*(November 2019), 102248. https://doi.org/10.1016/j.scs.2020.102248

Amini Toosi, H., Lavagna, M., Leonforte, F., Del Pero, C., & Aste, N. (2021). Implementing Life Cycle Sustainability Assessment in Building and Energy Retrofit Design—An Investigation into Challenges and Opportunities. *Environmental Footprints and Eco-Design of Products and Processes*, 103–136. https://doi.org/10.1007/978-981-16-4562-4_6

Amini Toosi, H., Lavagna, M., Leonforte, F., del Pero, C., & Aste, N. (2022). A novel LCSA-Machine learning based optimization model for sustainable building design-A case study of energy storage systems. *Building and Environment*, 209, 108656. https://doi.org/10.1016/j.buildenv.2021.108656

Annibaldi, V., Cucchiella, F., De Berardinis, P., Rotilio, M., & Stornelli, V. (2019). Environmental and economic benefits of optimal insulation thickness: A life-cycle cost analysis. *Renewable and Sustainable Energy Reviews*, *116*(October), 109441. https://doi.org/10.1016/j.rser.2019.109441

Aste, N., Caputo, P., Buzzetti, M., & Fattore, M. (2016). Energy efficiency in buildings: What drives the investments? the case of Lombardy Region. *Sustainable Cities and Society*, *20*, 27–37. https://doi.org/10.1016/j.scs.2015.09.003

De Wilde, P. (2014). The gap between predicted and measured energy performance of buildings: A framework for investigation. *Automation in Construction*, *41*, 40–49. https://doi.org/10.1016/j.autcon.2014.02.009

Ecoinvent. (2021). *LCA database*. https://ecoinvent.org/

Epd. (2021). *No Title*. https://www.environdec.com/home

Ristimäki, M., Säynäjoki, A., Heinonen, J., & Junnila, S. (2013). Combining life cycle costing and life cycle assessment for an analysis of a new residential district energy system design. *Energy*, *63*(2013), 168–179. https://doi.org/10.1016/j.energy.2013.10.030

Troup, L., Phillips, R., Eckelman, M. J., & Fannon, D. (2019). Effect of window-to-wall ratio on measured energy consumption in US office buildings. *Energy and Buildings*, *203*, 109434. https://doi.org/10.1016/j.enbuild.2019.109434

Westermann, P., & Evins, R. (2019). Surrogate modelling for sustainable building design – A review.

Energy and Buildings, *198*, 170–186. https://doi.org/10.1016/j.enbuild.2019.05.057

Zieger, V., Lecompte, T., & Hellouin de Menibus, A. (2020). Impact of GHGs temporal dynamics on the GWP assessment of building materials: A case study on bio-based and non-bio-based walls. Building and Environment, 185(August), 107210. https://doi.org/10.1016/j.buildenv.2020.107210