# Machine Learning-Aided Architectural Design for Carbon Footprint Reduction

mgr inż. arch.
**MATEUSZ PŁOSZAJ-MAZUREK**
Politechnika Warszawska
Wydział Architektury
**ORCID: 0000-0002-4618-7849**

The built environment is considered to be responsible for at least 20-40% of greenhouse gases emission. As architects and engineers, we can strive to overcome this issue in our everyday practice. The way we design buildings can greatly influence our society's carbon emissions – by means of minimizing the carbon footprint of a building over its entire lifecycle. But where should the emphasis be placed? Performing Life Cycle Assessment may help designers observe the factors that cause the highest levels of emission in a project. Thus, such an assessment may also facilitate decision-making when choosing between different alternatives. Machine Learning is an emerging trend in many fields of science. It can accurately predict outcomes of processes, actions or detect and measure trends. What are the possibilities to incorporate ML into Integrated Design Process?

## Machine Learning in Architectural Design

Machine Learning is beginning to revolutionize various fields of science. This fact may be noticed in areas, such as finance or medicine, yet its application in architectural design is limited [1]. The unique factor related to ML approach to designing lies in the fact that the algorithms are not programmed for specific tasks. The algorithms create mathematical models with the use of input data (known as Training Set), and apply the models to learn the ways so as to correctly predict the outcome. Three main branches of Machine Learning may be listed [2]. Firstly, supervised learning is related to the algorithm trained on the already labelled data, which means that the algorithm immediately knows if the prediction is correct or wrong. Secondly, the unsupervised learning is applied in order to find patterns in data, and such algorithms are often used in image analysis and generation. Finally, the reinforcement learning is based on a different approach - the algorithm is training itself continuously by receiving feedback on its accuracy. Currently, the main applications of Machine Learning in architectural design include: those related to functional plan generation [3] or design space exploration [4] among many other uses [5]. Environmental assessment with the use of the ML approach was first presented by Theodore Galanos [6], while a suggestion that ML can help reduce carbon footprint was made by D'Amico et al. [7].

## Integrated Design Process and Simplified Life Cycle Assessment (LCA)

Integrated Design Process is one of the design methods that can help minimize the impact of architectural design on the environment [8] by incorporating various analyses into the iterative design process, which covers entire life cycle of the building. Carbon footprint estimation (typically conducted with application of Life Cycle Assessment [9]) may be provided as one example of such analyses that may lead to creation of more environmentally-friendly buildings. In early design phases, a simplified version of LCA can be applied in order to assess the impacts of the building sooner [10].

## Parametric Approach and Optimization in Integrated Design Process

The utility of Integrated Design Process is based on fact that it offers an ability to test particular design variants, by analysing them and then assessing the applicability of each variant. With parametric approach, however, a building model is created not based of exact values, but rather on the basis of specific relationships between building components. For example, such as the width, length and heighted a building is defined by parameters that can easily be modified. The parametric design approach requires more planning at the initial phase of the design process, as all the relationships between individual elements are defined by these parameters, but at the same time the approach allows for much greater flexibility when it comes to changing the parameter values. The outcome of such modifications is almost instantly noticeable. It also offers a possibility for easy automation of the process of creating particular variants. A great benefit of using parametric model may be sought in the possibility to easily apply various optimisation methods. Making use of Genetic Algorithms

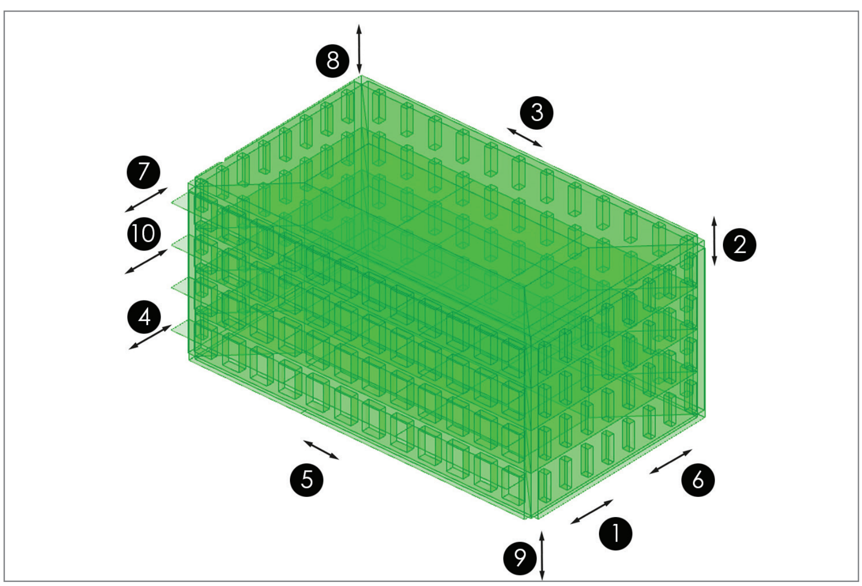


Fig. 1.: Parameters of the building model (described in Table 1)

Fig. author's archives

that can arrive at the optimal parameter combination for the selected fitness variable may be seen as a popular approach. A range of studies indicate the advantages of parametric approach in architectural optimisation [11, 12], some of which use parametric approach for Life Cycle Assessment [13, 14].

## Case study details

To test the possibility of applying Machine Learning in research and design process, the author has chosen a multifamily building for a case study. The building shape and features have been described by 10 parameters (as presented in fig. 1 and described in Table 01). The parameters could be changed within predefined ranges in order to test various combinations resulting in a unique multifamily building model in each case.

The building covered a constant area of 1600m2. The material composition of the walls, roofs and floors was to remain a constant value as well whereas it was possible to modify the thicknesses of insulation layers. Material GHG emissions data was gathered by the author from multiple sources, mainly by means of analysing products available in polish market, for which an EPD was available [15], or from similar products available abroad with declared EPDs [16]. The Carbon Emission factors for electricity were derived from the 2018 report by KOBiZE[17], and data for district heating was collected from the report by Veolia[18]. Local climate conditions were considered in the study using the EPW file that contains climate data (Warszawa Okęcie 123750 IMGW).

## Methods for finding the minimal optimized Total Carbon Footprint

Theoretically, it would be possible to simulate all possible parameter values and select the most beneficial combination (the lowest Total Carbon Footprint), but this approach seems rather impracticable. The number of possible combinations increases exponentially with each additional dimension (parameter) added. With higher numbers of possible solutions, it is often hard to simulate each case. Therefore, an approach that takes advantage of Machine Learning prediction power in order to search through the entire design space has been suggested. Given a trained Machine Learning model, it is possible to generate hundreds of thousands of variants in the matter of seconds. A different approach to the issue, namely the one applying a Genetic Algorithm to find the optimal solution, had been tested by the author in a previous study [19].

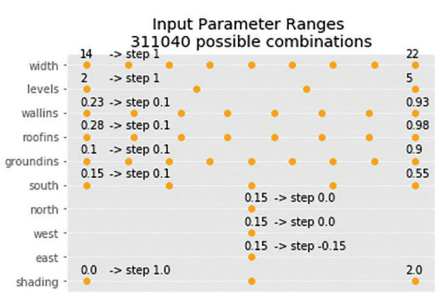


Fig. 2. The input parameters for data generation for Machine Learning Training Set.

## Machine Learning applied in Design Exploration

The first step to apply a supervised Machine Learning algorithm is to find or generate the data that can be used in the training process. The full design space of the problem (all possible combinations of parameters) consists in over twenty seven quadrillion possible combinations. Design of Experiment was planned on a less numerous set of parameters that represents the design space (fig. 2). Each case has been generated using values from the predefined range in a random manner.

Simulating various fenestration ratios along non-south oriented facades (west, east and north oriented) was omitted in the present study, as the previous analyses [20] confirmed that the correlation is almost linear – larger fenestration areas along non-south oriented facades resulted in higher level of Carbon Footprint.

Table 1. – Parameters considered in the Case Study

| Parameter number | Parameter | Range |
|---|---|---|
| 1 | Width | 12.0 – 24.0m |
| 2 | Levels | 2-5 |
| 3,4,5,6 | Window density north, west, south, east | 0.15-0.60 |
| 7 | Wall insulation thickness | 0.13-0.83m |
| 8 | Roof insulation thickness | 0.18-0.88m |
| 9 | Ground insulation thickness | 0.10-0.80m |
| 10 | Shading length | 0.0-2.0m |

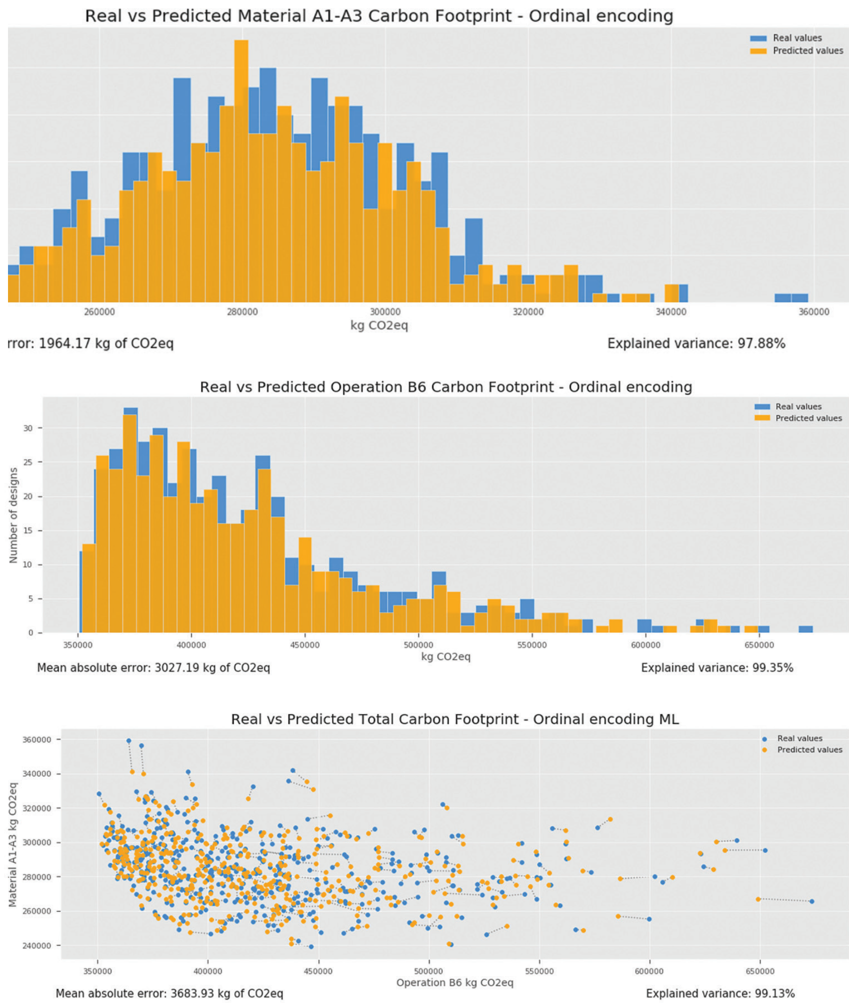


Fig. 3.: The results of the machine learning model training. First two diagrams indicate histograms of predicted and actual values from the test set. The third diagram shows the relationship between actual (simulated values) and the predicted ones.

The following step was to generate all the possible iterations of the parameters. The list was generated with the use of Python code. This resulted in a tally of 311040 possibilities. 1500 randomly selected cases with different parameter combinations were saved to a.csv file and subjected to simulation with the use Grasshopper Programme. For each case, the script led to evaluation of the design by calculating the Embodied Carbon using LCA Tool plugin for Grasshopper [14], and Operational Carbon level using Dynamic Simulation. Dynamic Simulation was performed in EnergyPlus, using Honeybee plugin for the Grasshopper Programme [21].

### Training the Machine Learning Model

The process of training Machine Learning model was performed in the Python environment. The approach proposed by Theodore Galanos [6], analysed in the present study was followed by two consecutive steps. Two separate machine learning models (for Embodied and Operational Carbon) were trained in Python using a supervised learning algorithm - Gradient Boosting Regressor (GBR) from Sci-Kit Learn library. The results were summarised and resulted in Total Carbon Footprint. The combined model explains over 99% of the variance. Mean absolute error is estimated to be at the value of about 3600 kg of $CO_2$eq (fig. 3).

### Exploring the Design Space

The following step was to predict the Total Carbon Footprint for much greater number of designs using the prediction offered by Machine Learning. A much bigger list of features was created that consisted of 100,000 cases (feature combinations). The results were then predicted using previously trained model in less than 2 minutes.

The results were displayed on a scatterplot (Embodied Carbon with regards to Operational Carbon, with colour coding Total Carbon Footprint) which presents general distribution of the outcomes (fig. 4). The optimal solution is the one that properly balances between Embodied Carbon and Operational Carbon, it is neither a solution with the lowest Embodied Carbon nor with the lowest Operational Carbon. The influence of the building height on the carbon footprint also seems to be important (fig. 5). Higher buildings (dark red) are positioned further to the right than the rest of the variables. The lowest buildings (violet) seem to have higher level of Operational Carbon Footprint. The optimal solution then should consist of 3 or 4 levels. The density sideplots confirm observations from the main scatterplot.

An optimal thickness value for Wall Insulation and Roof Insulation may be provided which is neither the maximum nor the minimum value (fig. 6), but rather consist in opti-

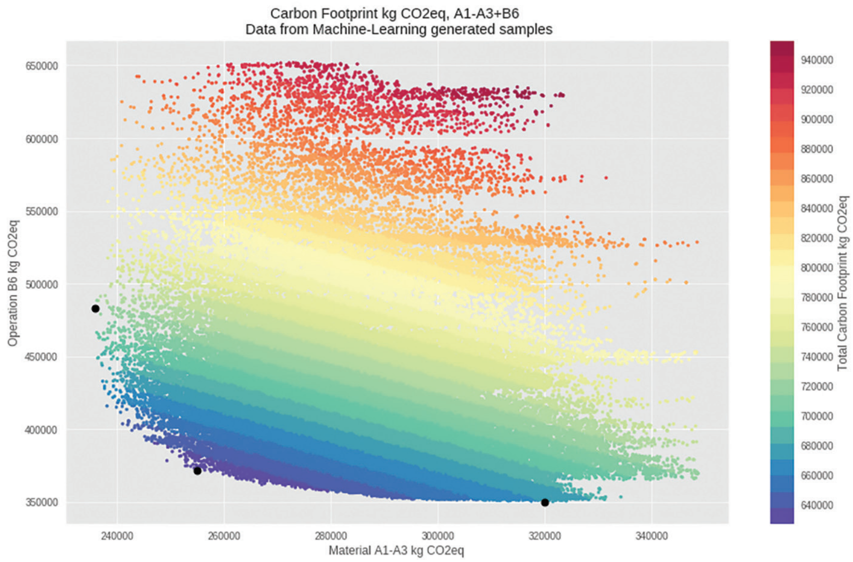


Fig. 4. The plot illustrates the distribution of results versus their Embodied and Operational Carbon Footprint. Black dots marks the positions of the lowest levels of Embodied Carbon, Total Carbon and Operational Carbon respectively.

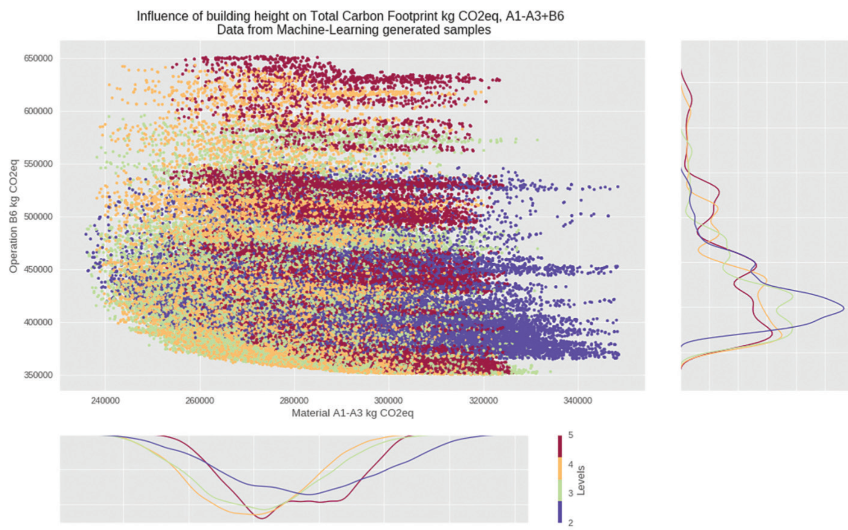


Fig. 5. The influence of building height on Total Carbon Footprint. The lowest Total Carbon Footprint can be observed in the cases of 3 and 4 level buildings. The 2 level buildings have higher Embodied (Material) Carbon Footprint, while the 5 level buildings have higher Operational Carbon Footprint.

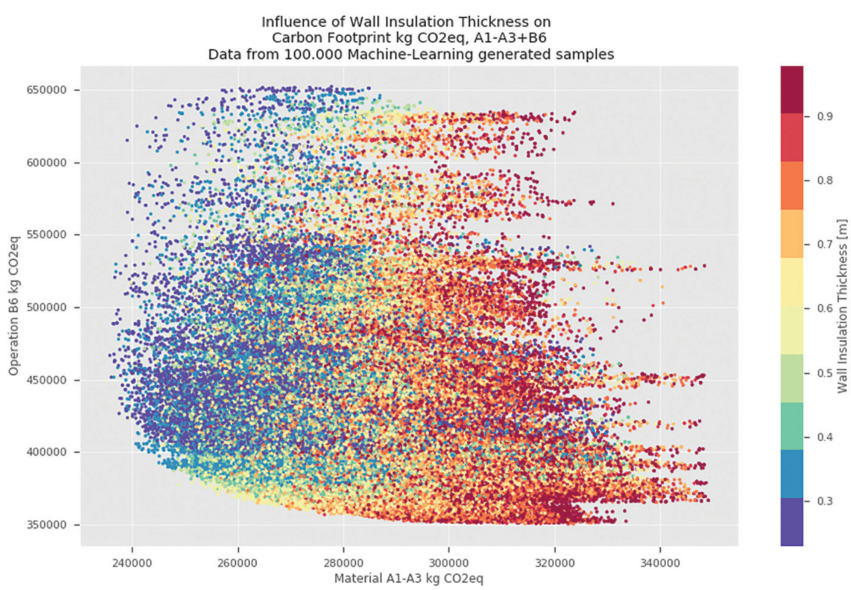


Fig.6. 4 Scatterplots illustrate the influence of Wall Insulation, Roof Insulation, Ground Insulation and Glazing Percentage along southern facade on the Total Carbon Footprint. The area with lowest Total Carbon Footprint is located in the bottom left part of each of the diagrams. In each of the building components analysed the situation is slightly different: for Wall and Roof Insulation the middle values seem to have the lowest Total Carbon Footprint, while for the Ground Insulation the lowest values have lower Total Carbon Footprint.

mal balance between Embodied Carbon of producing the insulation and the carbon saved in the Operational phase. On the other hand, the Ground insulation seems to have too insignificant effect on the Operational Carbon to be worth mentioning, so it is advisable to use the lowest thickness ssible (the correlation matrix on fig. 7 shows the correlation between different parameters and the values of Operational and Embodied Carbon Footprint). Similarly, the amount of glazing should also be minimized, because higher percentages of it lead to cooling needs, which can hardly be balanced by reducing the need for heating.

It can be observed that the Operational Carbon exerts a much greater influence on the Total Carbon Footprint than the Embodied Carbon. Some of the features, however, exert a nonlinear impact on the results – for example as it has been observed from fig. 6. - the optimal (minimal) Total Carbon Footprint is achieved at medium Wall Insulation Thicknesses.

The optimal solution selected out of all the cases generated by the Machine Learning algorithm yielded the value of 626368 kg of Total Carbon Footprint. The parametric 3d model has been created using the parameter values selected by the algorithm (fig. 8).

This was further compared to an actual simulation with the same parameters. The result amounted to 634777 kg, which denotes that the error for the Machine Learning Prediction was at the level of 1,3%.

## Applying Machine Learning Prediction in Integrated Design Process

Another possible usage of a trained Machine Learning model can be discussed by using a more real-life application. The trained ML model can be imported back to Grasshopper [6] to immediately estimate the Embodied Carbon Footprint and Operational Carbon Footprint of a case generated in the program (fig. 9). This provides architects with the possibility to sketch designs based on their intuition with instant feedback. Thus, ML can be seen as a useful educational and productivity-related tool.

## Conclusions and further research

The Machine Learning approach is a promising method that can be used in the Design Process. However, the tool still needs improvements, allowing for taking into consideration more features: different building shapes, different construction materials and urban layout. The data used for training the model should also be gathered in a database, that would allow for further reuse in different case studies.

The experiments have shown that Machine Learning can be a useful research tool for exploring vast design spaces in the field of Su-

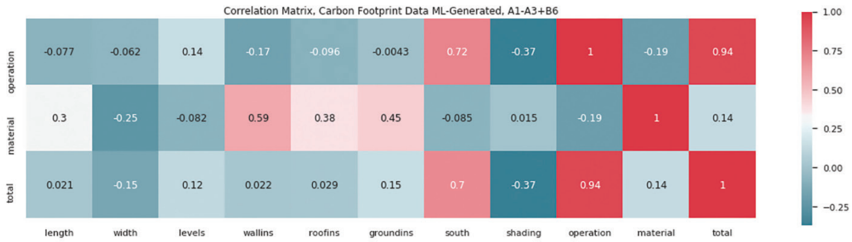


Fig.7. Correlation matrix shows the relationship between changing model parameters and the resulting output.

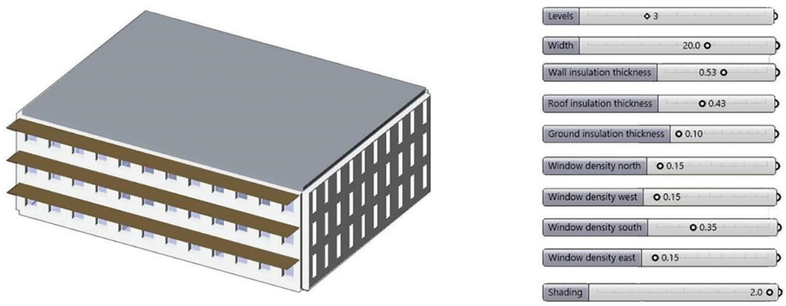


Fig. 8. Optimal solution selected from ML generated cases.

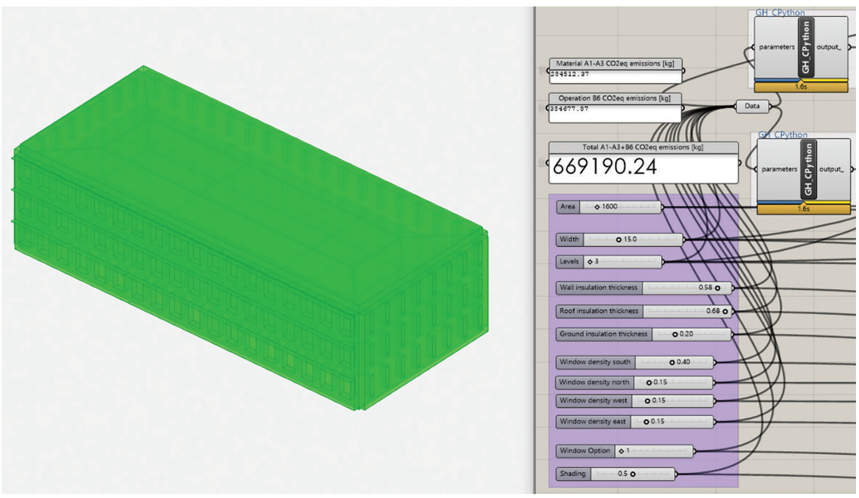


Fig.9. ML model predicts Total Carbon Footprint based on the values of the input sliders.

stainable Architectural Design. Additionally, it can offer the possibility for architects to freely sketch building designs, while being provided with instant information feedback. The current situation requires, unfortunately, a lot of data to properly apply the power of ML approach. However, in the near future, given the possibility of web-based platforms that share ML models, it might be possible to use the instant feedback of ML algorithms in everyday work.

## Literature

[1] C. Belém, L. Santos, and A. Leitão, "On the Impact of Machine Learning. Architecture without Architects?" in CAAD Futures, no. June, 2019.
[2] A. Burkov, The 100-Page Machine Learning Book. 2019.
[3] P. Merrell, E. Schkufza, and V. Koltun, "Computer-Generated Residential Building Layouts," ACM Trans. Graph., vol. 29, no. 6, pp. 1–12, 2010.
[4] C. Sjoberg, C. Beorkrem, and J. Ellinger, "Emergent Syntax: Machine Learning for the Curation of Design Solution Space," Proc. 37th Annu. Conf. Assoc. Comput. Aided Des. Archit., pp. 552–561, 2017.
[5] J. Cudzik and K. Radziszewski, "Artificial Intelligence Aided Architectural Design," vol. 1, no. October, pp. 77–84, 2018.
[6] T. Galanos, "Machine-Learned Regenerative Design," in Regenerative design in the digital practice, In press., E. Naboni and L. Havinga, Eds. Eurac Research, 2019, pp. 95–99.
[7] B. D'Amico et al., "Machine Learning for Sustainable Structures: A Call for Data," Structures, vol. 19, no. November, pp. 1–4, 2019.
[8] E.D. Ryńska, Zintegrowany proces projektowania prośrodowiskowego. Warszawa: Oficyna Wydawnicza Politechniki Warszawskiej, 2012.
[9] PKN, "PN-EN 15978 - Sustainability of construction works - Assessment of environmental performance of buildings - Calculation method" 2011.
[10] A. Säynäjoki, J. Heinonen, S. Junnila, and A. Horvath, "Can life-cycle assessment produce reliable policy guidelines in the building sector?," Environ. Res. Lett., vol. 12, no. 1, pp. 1–16, 2017.
[11] G. Lobaccaro, A. H. Wiberg, G. Ceci, M. Manni, N. Lolli, and U. Berardi, "Parametric design to minimize the embodied GHG emissions in a ZEB," Energy Build., 2018.
[12] E. Naboni, A. Malcangi, Y. Zhang, and F. Barzon, "Defining the energy saving potential of architectural design," Energy Procedia, vol. 83, no. December, pp. 140–146, 2015.
[13] A. Hollberg, "A parametric method for building design optimization based on Life Cycle Assessment," no. November, 2016.
[14] K. Negendahl, A. P. Otovic, and L. M. B. Jensen, "Expansion in Number of Parameters : Simulation of Energy and Indoor Climate in Combination with LCA" 2016 ASHRAE Annu. Conf., no. September, 2016.
[15] ITB, 2019, ITB EPD Program, accessed on August,12th,2019 <https://www.itb.pl/epd.html>
[16] ÖKOBAUDAT, 2019, accessed on August,12th,2019 <https://www.oekobaudat.de/>
[17] KOBiZE, "Emission Factors 2018," 2018. accessed on August,12th,2019 <https://www.kobize.pl/uploads/materialy/materialy_do_pobrania/wskazniki_emisyjnosci/Wskazniki_emisyjnosci_2018.pdf>
[18] Veolia, "Environmental Impact 2018," 2018. view accessed on August,12th,2019 <https://energiadlawarszawy.pl/wp-content/uploads/sites/4/2019/04/wplyw_na_srodowisko_2018-1.pdf%0D>
[19] M. Płoszaj-Mazurek, "Parametric Optimisation of Carbon Footprints," in Regenerative design in the digital practice, In press., E. Naboni and L. Havinga, Eds. Eurac Research, 2019, pp. 256–261.
[20] M. Płoszaj-Mazurek, "Parametric Optimization of Carbon Footprint," in iiSBE Forum of Young Researchers in Sustainable Building 2019, 2019, no. July.
[21] M. S. Roudsari, M. Pak, and A. Smith, "Ladybug: a Parametric Environmental Plugin for Grasshopper To Help Designers Create an Environmentally-Conscious Design," 13th Conf. Int. Build. Perform. Simul. Assoc., pp. 3129–3135, 2013.

DOI: 10.5604/01.3001.0014.1615

CORRECT METHOD OF QUOTATION
Płoszaj-Mazurek Mateusz, 2020, Machine Learning-Aided Architectural Design for Carbon Footprint Reduction, „Builder" 07 (276).
DOI: 10.5604/01.3001.0014.1615

**Abstract:** The built environment is considered responsible for at least 20-40% of greenhouse gases emission. The way we design may exert an impact on this percentage. A new paradigm, namely artificial intelligence, is arriving. More and more tasks are becoming automated via algorithms. How could this power be applied in order to strengthen our knowledge about the ways we design buildings? The author of the following paper presents a study in which carbon footprint yielded by a multifamily building is analysed. ML has been used to generate an extensive overview of the possible design solutions. This, in turn, made it possible to observe correlations between various parameters that resulted in a reduced carbon footprint.
**Keywords:** life cycle assessment, parametric optimization, artificial intelligence, AI, algorithms, ghg emissions, sustainable architecture, big data, machine learning (ML)

**Streszczenie:** WSPOMAGANE UCZENIEM MASZYNOWYM PROJEKTOWANIE ARCHITEKTURY W CELU ZMNIEJSZENIA ŚLADU WĘGLOWEGO. Środowisko zabudowane odpowiada za co najmniej 20 do 40% emisji gazów cieplarnianych, a sposób, w jaki projektujemy, może wpłynąć na tę wartość. Coraz więcej zadań zostaje zautomatyzowanych za pomocą algorytmów. Jak możemy wykorzystać to narzędzie, aby wspomóc naszą wiedzę na temat sposobów projektowania budynków? Autor przedstawia badanie analizujące ślad węglowy budynku wielorodzinnego. Algorytm uczenia maszynowego został wykorzystany do wygenerowania obszernego przeglądu możliwych rozwiązań projektowych. Umożliwiło to zaobserwowanie korelacji między różnymi parametrami, co pozwoliło na wybór kombinacji parametrów o najniższym śladzie węglowym.
**Słowa kluczowe:** ocena cyklu życia, optymalizacja parametryczna, ślad węglowy, uczenie maszynowe