

Article

Methods to Optimize Carbon Footprint of Buildings in Regenerative Architectural Design with the Use of Machine Learning, Convolutional Neural Network, and Parametric Design

Mateusz Płoszaj-Mazurek ^{*}, Elżbieta Ryńska and Magdalena Grochulska-Salak

Faculty of Architecture, Warsaw University of Technology (WUT), 00661 Warszawa, Poland;
elzbieta.rynska@pw.edu.pl (E.R.); magdalena.salak@pw.edu.pl (M.G.-S.)

* Correspondence: mateusz.mazurek@pw.edu.pl

Received: 25 August 2020; Accepted: 7 October 2020; Published: 12 October 2020



Abstract: The analyzed research issue provides a model for Carbon Footprint estimation at an early design stage. In the context of climate neutrality, it is important to introduce regenerative design practices in the architect's design process, especially in early design phases when the possibility of modifying the design is usually high. The research method was based on separate consecutive research works–partial tasks: Developing regenerative design guidelines for simulation purposes and for parametric modeling; generating a training set and a testing set of building designs with calculated total Carbon Footprint; using the pre-generated set to train a Machine Learning Model; applying the Machine Learning Model to predict optimal building features; prototyping an application for a quick estimation of the Total Carbon Footprint in the case of other projects in early design phases; updating the prototyped application with additional features; urban layout analysis; preparing a new approach based on Convolutional Neural Networks and training the new algorithm; and developing the final version of the application that can predict the Total Carbon Footprint of a building design based on basic building features and on the urban layout. The results of multi-criteria analyses showed relationships between the parameters of buildings and the possibility of introducing Carbon Footprint estimation and implementing building optimization at the initial design stage.

Keywords: life cycle assessment; parametric; optimization; artificial intelligence; AI; algorithms; GHG emissions; sustainable architecture; big data; machine learning; neural networks; computer vision; circular economy

1. Introduction

The built environment is a major contributor to the greenhouse gas emissions problem. It is estimated that up to 40% of the Global Greenhouse Gas (GHG) emissions result from the activity of the Architecture, Engineering, and Construction industry (AEC) [1]. A major part of these emissions stems from currently-applied design practices. Carbon Footprint “is a measure of the exclusive total amount of carbon dioxide emissions that is directly and indirectly caused by an activity or is accumulated over the life stages of a product.” [2] Precise information on how to calculate carbon emissions over the lifecycle of a building (under the indicator name of Global Warming Potential) has been provided in “Sustainability of construction works–Assessment of environmental performance of buildings–Calculation method” ([3]). Global Warming Potential, or Carbon Footprint are typically expressed in kilograms of CO₂ equivalent. The Total Carbon Footprint is the sum of the Carbon Footprints of all life cycle stages (Figure 1). The calculation of the Carbon Footprint of a single life

cycle stage i consists of multiplying all products and processes included in that life cycle stage i with their respective Global Warming Potential (Equation (1)).

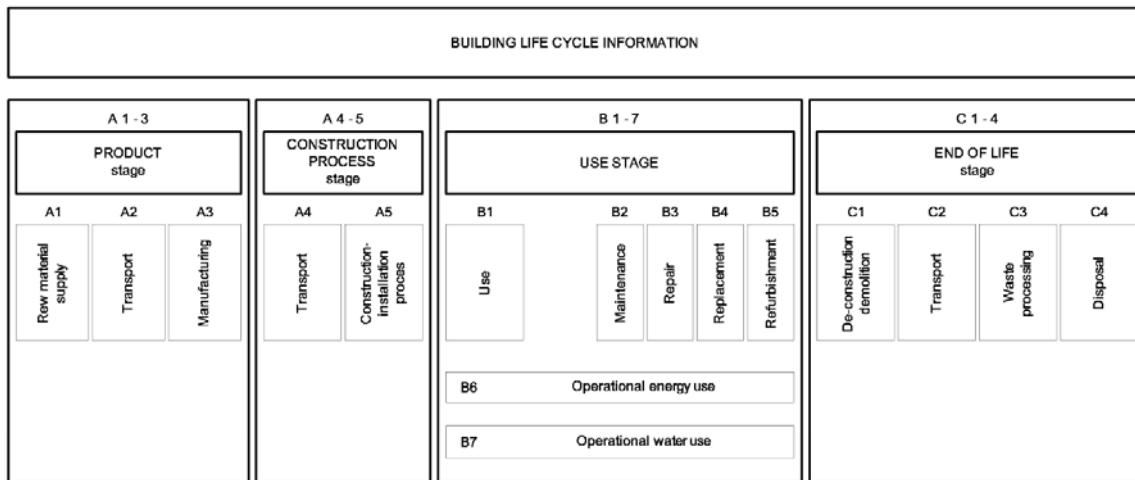


Figure 1. Building life cycle phases according to “Sustainability of construction works—Assessment of environmental performance of buildings—Calculation method” [3].

$$GWP_i = a_{1,i} \times GWP_{a1,i} + a_{2,i} \times GWP_{a2,i} + a_{3,i} \times GWP_{a3,i} + \dots + a_{N,i} \times GWP_{aN,i} \quad (1)$$

Architects need more tools that can help them during the decision process in the area of sustainable design as “analyses from the very initial stages allow the inclusion of smart energy choices influencing the massing, architectural features, proportions, flexibility of design, and economics” [4]. It has been observed that architecture students, when assisted with a tool that can facilitate their understanding of the effect their choices exert on the environment, create better architectural designs in terms of global warming and other environmental impacts [5]. A number of sustainability metrics, like for example the degree of application of “circular economy” practices (sometimes referred as “building circularity), still needs a proper measurement scale [6] and often a measurable feature like a carbon footprint is used in such situations. An important issue lies in the scientific development of potential methods for the analysis and dissemination of results, while also providing tools for designers who want to account for ecology-related aspects in their projects.

With the progressing degradation of the environment, growing emissions, and increasing depletion of resources, it is time not only to reduce our impact on the environment but also to start creating resilience and regenerative capacities for the built environment. The response to this issue may be seen in the concept of Regenerative Design [7]. It is necessary to proceed from sustainable architecture towards a restorative approach that can positively impact the environment. However, to achieve this goal, tools need to be developed that are open and allow for a high level of customization, one such example being the parametric approach.

1.1. Regenerative Design

Regenerative Design for Sustainable Development is the first handbook on regenerative design, published by John Tillman Lyle in 1994. The publication describes the objective to replace linear processes of energy and matter flow with circulation processes. This approach provides the basis for a change in the way of thinking about urban and architectural design, with a focus on the emergence of spatial systems that ensure the renewal and regeneration of resources used in the broadly understood processes of construction and consumption [8]. In regenerative design, integrated interdisciplinary action and user participation are important [9]. The regenerative design of processes related to the functioning of urban ecosystems involves the use of natural processes to improve the resilience of the built environment. It is important to raise awareness concerning the fact that, being part of

the urban ecosystem, people should act with deep respect and understanding of its processes [10]. The development of the regenerative design model in the context of transformed urbanized areas provides a basis for defining new urban standards and indicators, building parameters, and elements of the city's natural system, including green and blue infrastructure [11]. Regenerative design is a method to achieve climate neutrality and resistance in the built environment. Therefore, the carbon footprint analysis has been selected as the main focal point of the study.

1.2. Architectural Design Process

The typical approach to ecological design in the architectural design process is to use various kinds of assessment and simulation tools. Life Cycle Assessment (LCA) [3] is a common methodology, however such analyses are often neglected until the later phases of the design process. Therefore, such tools often fail to aid the designers during the design process, but are rather applied only to evaluate the final solution. At the same time, numerous early design phase decisions can vastly reduce the carbon footprint of the building. Early in the design process, the possibility to modify the design is usually very high, while it decreases in the later stages, as seen in Figure 2 [12]. The main features of the building, such as building dimensions or proportions, floor-to-floor height, construction technology, as well as the materials and parameters of window openings and glazing, which are typically set early, can exert a very big impact on the carbon footprint of a specific project and this early design phase can be called "the most pivotal phase in sustainable architectural design" [13].

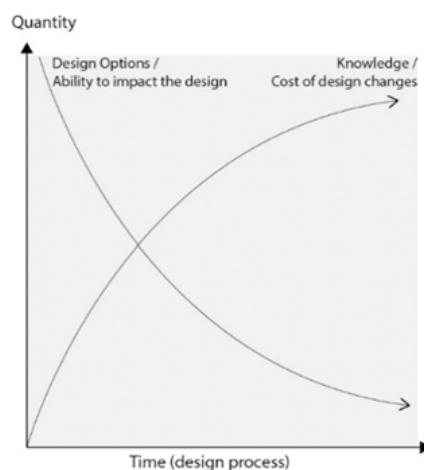


Figure 2. Impact of the information amount on the design process [12].

A need may be noticed for tools that can help architects to understand how their designs influence the environment, for example understanding the carbon footprint of a building design. Such a tool should be easy to use and not require in-depth knowledge or big amount of data so as to be applicable in the early phases of the design process.

1.3. Artificial Intelligence in the Architectural Design Process

The architects' design process is strongly related to their working method. Over the years, design processes have undergone profound changes. From sketching on a sheet of paper, through Computer Aided Design (CAD), to advanced computational methods [14]. Architectural design is more and more akin to the proper management of information [15]. The way the architects work constantly changes.

At the same time, a fast development of Artificial Intelligence (AI) approaches is observable in various scientific fields. Various statistical regression approaches are being used in finances, while image recognition supports the medical field. Machine Learning (ML), a subset of Artificial Intelligence, is a unique approach as it is based on statistical methods and is not programmed for a specific task. On the contrary, it is intended to detect repeating patterns in big sets of data and

to construct mathematical models that describe data behavior. Different types of machine learning exist, some are based on pre-made training data sets, while other types seek to learn without human supervision [16]. For tasks with high dimensional data, neural networks have been developed.

In the architectural design, however, the usage of AI is still limited [17] and there are some cases of using ML for design space exploration [18], plan generation, and style transfer [19], or for environmental assessment [20] and energy performance assessment and optimization [21], to provide just some examples [22]. Machine Learning approach could facilitate the assessment of environmental performance for architects and it has been suggested that Machine Learning may provide a tool that will greatly help to reduce the carbon footprint of buildings [23]. The application of Machine Learning in architectural practice is, however, limited by one crucial element, namely access to data. The algorithms require big amounts of labelled source data to be able to learn the patterns inside the datasets. So, for example, in order to train an algorithm to properly estimate the lifecycle carbon footprint of a projected building, one needs a lot of previously calculated examples. In a perfect situation, the data should come from real-world calculations, so that they are as close to reality as possible. However, in many situations, access to such data is not possible. In the case of carbon footprint, it is possible to simulate different building shapes, assess their Embodied and Operational Carbon and use those examples to train the ML algorithm.

2. Materials and Methods

Machine Learning requires big amounts of training data to be able to teach itself by recognizing specific patterns in the dataset. In the case of building simulation data, it is possible to create a parametric script that automates the process of dataset generation. The following sections describe three studies conducted on the application of Machine Learning to an architectural design issue. In the first study, the ability of the ML model to properly predict the Total Carbon Footprint of the building based on the features of the cuboid building has been tested, and then the created ML model has been used to arrive at the optimal solution. In the second study, attempts have been made to develop a method by which the algorithm can learn the information from more complicated prism shapes, while achieving high accuracy.

The building model has been created in Grasshopper, a parametric modelling plugin for Rhinoceros3D. The building has been described on the basis of specific relationships between building components. The main building features, such as height, width, or fenestration ratio on each façade have been described by variable parameters, as seen in Figure 3.

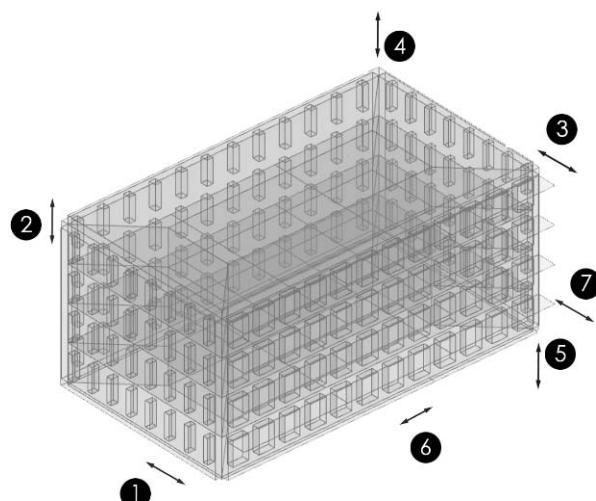


Figure 3. The parametric model used in Study 01. A total of seven parameters describe the building shape (1: width; 2: levels; 3: wall insulation thickness; 4: roof insulation thickness; 5: ground insulation thickness; 6: fenestration ratio on the southern façade; and 7: shading length).

3. Results

3.1. Case Study 01 Details

In all the studies, boundary conditions for the building were similar. The analyzed object has been a multi-family building, located in Warsaw, Poland. According to the local climate analysis, the local conditions demand both heating and cooling (the local climate is considered to be Dfb according to the Köppen–Geiger climate classification [24]). The climate data was used in the form of an EPW (EnergyPlus weather format) file [25]). The reference study period for the operation of the building was set to 30 years. Material data for Green House Gas emission factors was gathered from multiple sources [26,27]. In the case of the first study, the total area of the building was kept at a constant 1600 m², while in the second and third study, it was possible for the total floor area to reach between 1500 and 3000 m². In this article, the total area is defined as the area inside of the external wall perimeter, taking into account all the area under interior walls and other elements in accordance to the definition of “Interior Space” from PN-ISO 9836:1997 [28], and the limitation comes from the nature of simulation programs which often fail to take the wall thickness into consideration. The size of the buildings was selected on the basis of desired Floor Area Ratio (FAR) for the selected plot so as to keep it below 2 for the largest case scenario (3000 m²). The building height was limited to between 2 and 5 levels. The building structure is considered to be built of a concrete slab and column system, which is a common construction type in Poland. The building thermal envelope is presented in Table 1. The Heating, Ventilation, Air Conditioning (HVAC) system was considered only for Operational Carbon. The building uses mechanical ventilation with a heat recovery of 75%. The cooling need is reduced by natural ventilation during periods of the year and the program assumes an optimal schedule for opening the windows whenever it is favorable. The amount of staircases varies depending on the building length and floor area. The interior of the building (walls, doors, etc.) has been calculated on the ‘per square meter’ basis, as the influence of the amount of secondary elements on the final result in simplified LCA [29]. Five-floor plans of multi-family buildings in Warsaw were analyzed to extract the average amount of interior elements per square meter of the building, and the results were used to calculate the average amount of those elements per m² of the floorplan.

Table 1. External wall, roof, and ground floor compositions.

External Wall Composition	
Layer	Thickness
Finish–Plaster	2 cm
Insulation–Mineral Wool	13 cm–103 cm
Structure–Brick 90%	20 cm
Reinforced Concrete 10%	
Interior–Plaster	2 cm

Roof Composition	
Layer	Thickness
Insulation–Mineral Wool	18 cm–98 cm
Structure–Reinforced Concrete	15 cm

Ground Floor Composition	
Layer	Thickness
Finish–Wood	2 cm
Plywood	2 cm
Insulation	10 cm–90 cm
Slab on Grade–Reinforced Concrete	15 cm

3.2. Testing Machine Learning Approach—Study 01

The study presents a novel method of using Machine Learning for Carbon Footprint Estimation in the Architectural Design Process. A similar approach has been suggested by Theodore Galanos [20] for Energy Performance, Wind, and Sunlight studies, but no such solution for Carbon Footprint has been found. The abovementioned study has been described in greater detail in previous articles by M. Płoszaj-Mazurek [30].

Study 01 is based on the 1500 simulations performed using a Grasshopper script. The objective of this study was to test the level of accuracy of prediction conducted by a Machine Learning model. First, a list of all possible feature combinations of the 7 parameters—1: Width; 2: Levels; 3: Wall insulation thickness; 4: Roof insulation thickness; 5: Ground insulation thickness; 6: Fenestration ratio on the southern façade; and 7: Shading length (see Figure 2) has been made. The list consisted of 311,040 possible combinations. Using random sampling, 1500 cases were selected. The list was drawn up using a Python script and then exported to a csv file. Then, the Grasshopper script iterated over all of the rows on the list and generated a building shape for each one of them. Next, the script calculated the Embodied Carbon (A1-A3 Lifecycle stages [3]) using LCA Tool [31], and simulated the energy performance of the building (B6 Lifecycle stage [3]) using Ladybug + Honeybee plugins that exported the building geometry to EnergyPlus [32]. The simulation results were then imported back to Grasshopper in order to calculate the Operational Carbon Footprint with the use of coefficients for emission factors [33,34]), and then to calculate Total Carbon Footprint.

In the following step, the Machine Learning algorithm was trained on the previously generated data for 1500 simulated cases, using a Python script. A Supervised Learning algorithm—Gradient Boosting Regressor (GBR) [20], was selected. Two separate models for Embodied Carbon and Operational Carbon were trained and then tested. The Total Carbon Footprint yielded as a result of summing up the Embodied Carbon and Operational Carbon models. The combined model was able to explain over 99% of the variance.

The trained ML model was used to predict the Total Carbon Footprint for 100,000 randomly generated cases. This facilitated a better understanding of the relations between specific parameter values and the resulting Embodied Carbon, Operational Carbon, and Total Carbon Footprint. The results were plotted into various graphs that helped to determine the correlations between the variables and results.

The results scatterplot in Figure 4 clearly shows a Pareto frontier, which describes the relationship between Operational Carbon and Embodied Carbon results. Trying to minimize the Embodied Carbon Footprint will result in higher Operational Carbon Footprint and vice-versa. The height-related colored scatterplot in Figure 5 also highlights an important fact, namely the optimal height for this building seemed to be between 3 and 4 levels, while higher (5 levels) and lower (2 levels) buildings seemed to perform worse in regard to the Total Carbon Footprint.



Figure 4. Scatterplot of Carbon Footprint of 100,000 predicted building shapes. Black dots mark the cases with the lowest Embodied Carbon (**left**), lowest Operational Carbon (**right**), and lowest Total Carbon (**middle**).

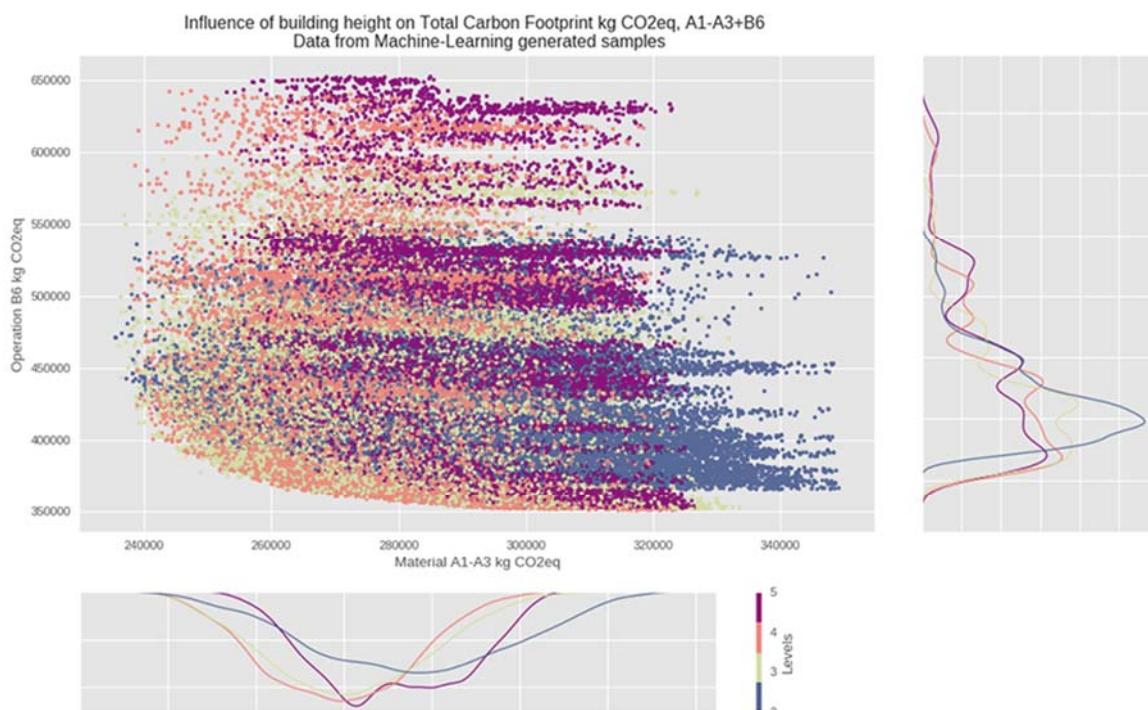


Figure 5. Scatterplot of Carbon Footprint with the coloring based on the building levels. The plot shows that lower buildings feature higher Embodied Carbon, while higher buildings feature higher Embodied Carbon, the optimal building height is nor the highest nor the lowest possible.

From all of the 100,000 generated cases, a building variant with the lowest Total Carbon Footprint was selected. The combination of the parameters for the lowest result is presented in Figure 5.

The optimal building shape has 3 levels, features a medium insulation thickness for walls (53 cm out of possible range of 13–103 cm) and roof (43 cm out of possible range of 18–98 cm), but the lowest possible thickness for the ground floor (10 cm out of possible range of 10–90 cm). The building shape is slightly longer on the west-east axis (with a proportion of 20 to 26.7 m), exposing southern windows to the sun to introduce solar gains, while at the same time it has the maximum shading length possible to reduce overheating (2 m out of possible 0–2 m). The fenestration ratio is also larger on the southern façade, but is still not at the maximum possible ratio (35% out of possible 15–50% range). The Machine Learning model prediction has been compared to an actual simulation of the same combination of parameters. The result from simulation diverged from prediction by only 1.3%. The ML model has predicted the Total Carbon Footprint at 626,368 kg CO₂eq (Figure 6). A simulation of the same building model resulted in a value of 634,777 kg.

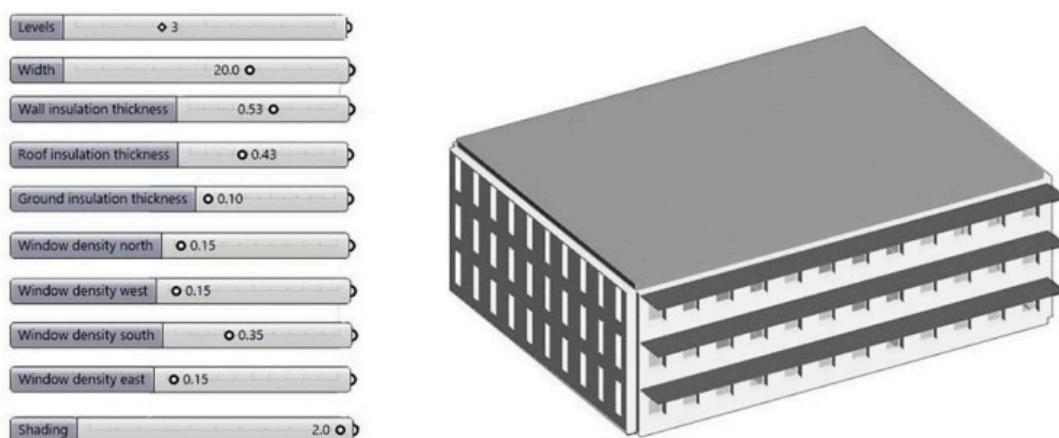


Figure 6. The optimal building shape with presented parameter values. This variant had the lowest Total Carbon Footprint as predicted by the ML (Machine Learning) model: 626,368 kg CO₂eq. The actual simulated value: 634,777 kg, which means an error of 1.3% for the Machine Learning Prediction.

The results of this study proved that it is possible to train a ML algorithm to correctly predict the building Total Carbon Footprint. However, the input features for the algorithm were selected in such a way that only cuboid buildings could be generated (for example measuring width value in a non-cuboid building is vague and it is not specified which value to measure). It became clear that the main limitation of this approach is the restriction to cuboid shapes. A method should be developed so that input of non-cuboid shapes into the ML model could be possible.

3.3. Improving the Tool for Different Building Shapes—Study 02

The second study focused on the creation of a plugin using a trained ML model, which can be applied in the preliminary design phase by architecture students. To simplify this approach, the insulation thicknesses and materials have already been defined during the model training. The factor that altered in each simulation was the building shape. The insulation thicknesses were selected at 2021 Technical Requirement levels in Poland (Table 2) [35].

Table 2. External wall, roof, and ground floor maximum U-values according to building regulations in force in Poland [33].

Building Component	U _c (max) [W/(m ² ·K)]
External Wall	0.20
Roof	0.15
Slab on Grade	0.3
Window	0.9

The objective of the second study focused on the possibility of extending the Machine Learning algorithm towards non-cuboid shapes. In this study, the possible building shape was limited to prisms, with a base shape formed of intersecting rectangles. This change required a completely different approach at selecting characteristic features of the building shape. For example, in an L-shaped prism, it is not clear which dimension should stand for width and length of the figure, as there are more than two different dimensions associated with the base shape. A need arose to select more abstract features, which could be extracted from any prism shape.

The response to this issue arrived in the form of calculating the total areas of various external building components. The external wall area, ground floor area, and roof area were chosen. Additionally, the area of windows facing each of the four main corners of the world (north, west, south, and east) were considered. Finally, the building height was considered. The below list of eight features created the input data to describe the geometry of the building. The eight parameters used to describe the shape of each building were as follows: Wall area, ground floor area, roof area, height, window area south, window area north, window area west, and window area east (Table 3).

Table 3. Input parameters in Study 02 and 03.

Parameter Name
Wall Area
Ground Floor Area
Roof Area
Height
Window Area South
Window Area North
Window Area West
Window Area East

To generate training data, a new script in Grasshopper was created. The script generated random building shapes based on the assumption that the shape should be a prism, with the base of two intersecting rectangles as shown in Figure 7. The script generated buildings with various fenestration ratios, various building heights, and floor areas.

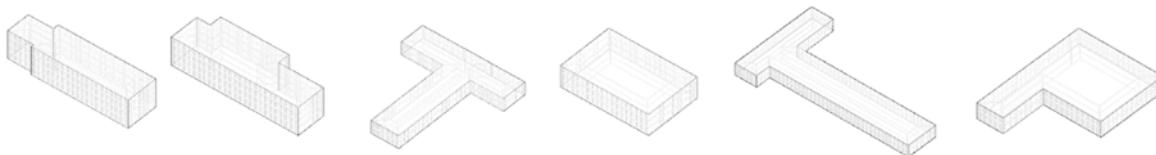


Figure 7. Examples of building shapes generated by script, used to train the Machine Learning model.

The script generated 1000 different building layouts. After sorting the data from the incorrect ones, 997 cases were then simulated in order to estimate the Total Carbon Footprint. Consecutively, Machine Learning models for Embodied Carbon and Operational Carbon were trained on the calculated training data. Again, a Supervised Learning algorithm—Gradient Boosting Regressor (GBR) was used for training.

The results proved that the prediction featured an average error of 1.6% on the Total Carbon Footprint value. The plots show that the model had a higher accuracy on Embodied Carbon than on Operational Carbon Footprint (98.7% vs. 97.5% explained variance), as seen in Figure 8.

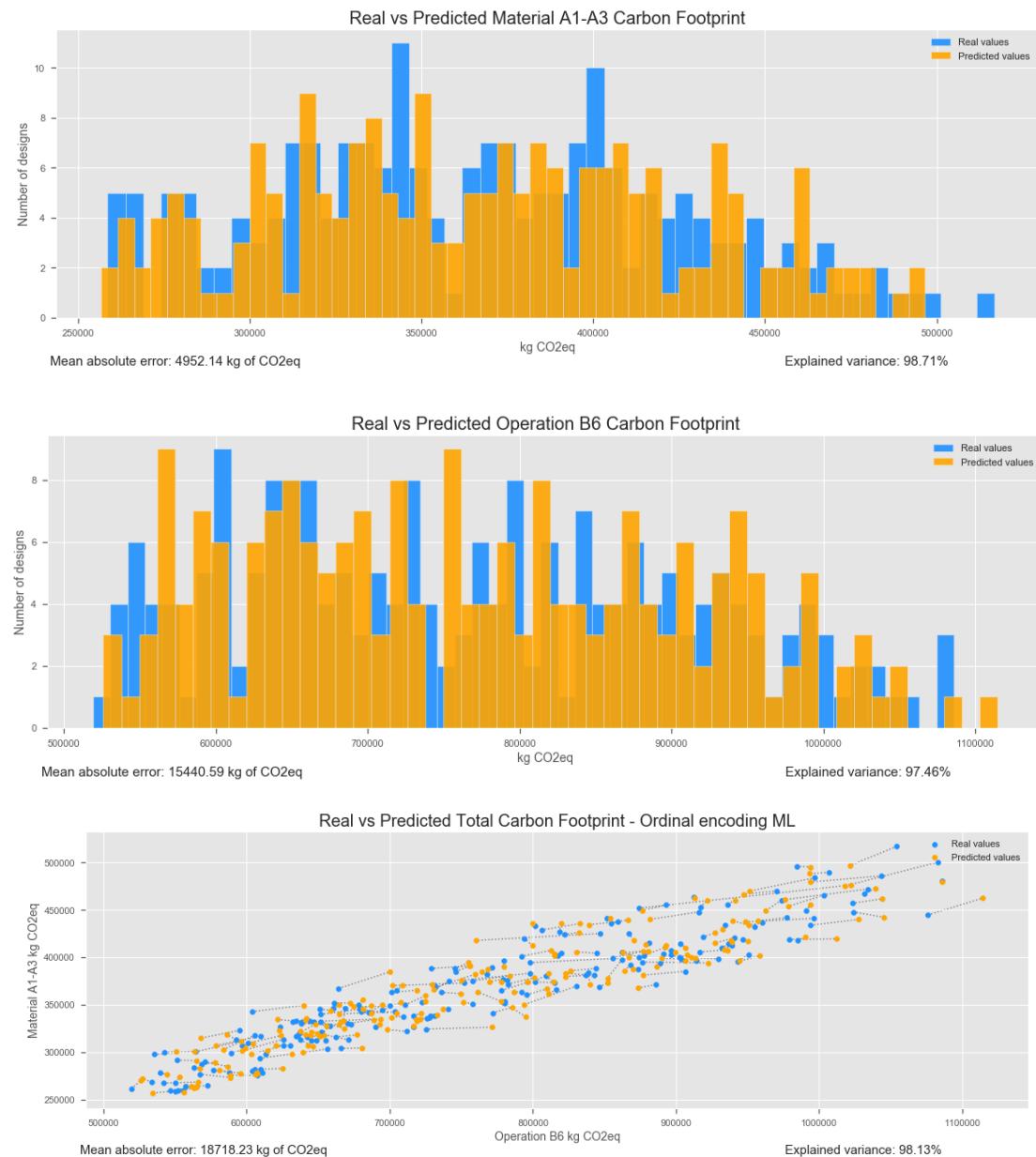


Figure 8. Results of the training process showing the better fit of the Embodied Carbon ML model, and slightly worse fit of Operational Carbon ML model.

3.4. Using the Trained Model from Study 02 in the Design Process

The trained ML model can be imported back into Grasshopper. There, it can be used to instantly estimate the proposed building Total Carbon Footprint. The input geometry can be modelled in Rhino, generated in Grasshopper, or imported from different programs, like Sketchup. The main requirement is that the geometry should be modelled as an analytical model (no thicknesses, no interior shapes), and the building components should be placed on specific layers. The model is read automatically by the script (Figure 9). The plugin then extracts eight characteristic features (external wall, roof and ground floor areas, window areas towards the main four corners of the world, and building height) of the building model automatically (by using layer names), and feeds them into the trained ML model. Finally, the predicted Total Carbon Footprint is displayed to the user.

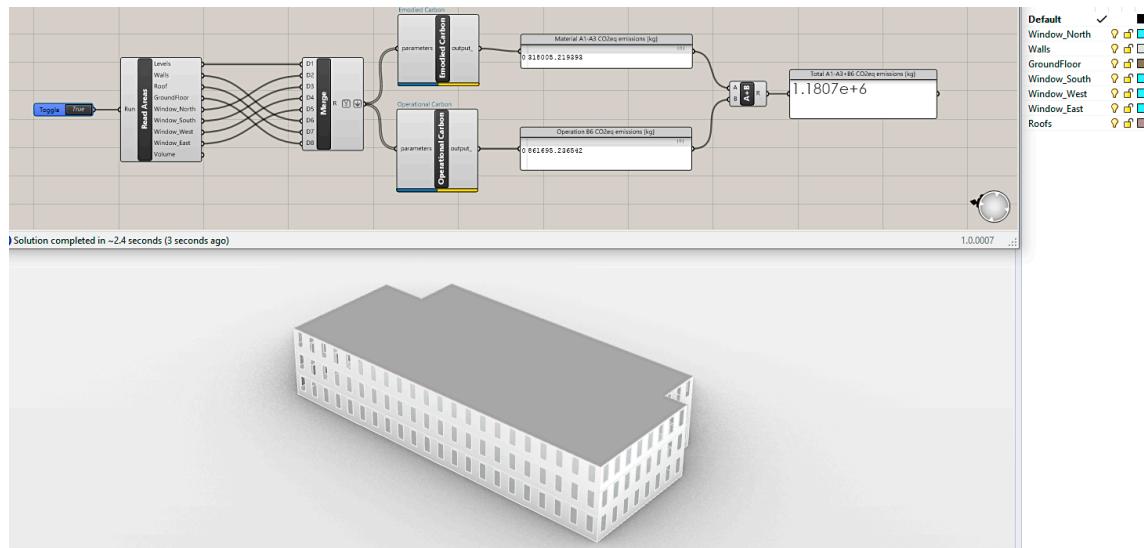


Figure 9. Using the trained ML model back in Grasshopper allows for a quick estimation of the Total Carbon Footprint of a given shape. The building model has to be separated into correctly named layers that represent the building components. The user just needs to click a button and then an estimation is instantly provided.

Such a tool can be used in the early phases of design so as to improve the understanding of the relationship between the building shape and its Total Carbon Footprint. The objective of this research is to provide a useful educational tool for students and architects in order to reduce carbon footprints yielded by their designs. In later phases, it is still advisable to make a full Life Cycle Assessment of the building design, in order to calculate a better Total Carbon Footprint estimate.

3.5. Convolutional Neural Networks—A Machine Learning Approach to Image Analysis

Neural Networks are a more advanced form of Machine Learning algorithms. A Convolutional Neural Network (CNN) is a specific type of Neural Network with at least one convolutional layer used and is especially good at analyzing images [29]. The CNN algorithm can analyze an input image by assigning importance to various objects in the image, and then performing other operations on the data.

A neural network typically consists of an input and output layer, as well as several hidden layers, most often fully connected layers (a layer that has all neurons connected to all the neurons of the previous layer). In Convolutional Neural Networks, the hidden layers are composed of different functions, mainly consisting of a Convolution and a Pooling layer. Convolution layer abstracts the image to a feature map. Pooling layer reduces the size of the map by applying a mathematical function, for example Max Pooling. Finally, the output of the Convolution Network is analyzed by several Fully Connected Layers, and then a loss function is applied. For classification, it can be a SoftMax function and for regression, it might be just a linear function.

3.6. Adding the Urban Layout Dimension—Study 03

The tool developed in the previous Studies 01 and 02 could help to develop a building with an optimized carbon footprint in an ideal situation, on an empty plot, without any shading objects. However in real life, such situations happen rarely. In typical situations, the architect has to face diverse urban environments with varying amounts of shading objects (buildings, foliage, and other objects). A practical tool should incorporate urban layout into consideration.

The next study focused on including the possibility to study the overshadowing effect on the Total Carbon Footprint of the building. To achieve this, a new Machine Learning model had to be created

and trained. This time, instead of simple regression model, a Deep Neural Network approach was selected. In this specific example, a CNN was used to analyze an image of the urban layout and then to consider its influence on the building's Total Carbon Footprint.

To train the CNN model, a set of 3000 simulations was performed. In addition to the aspects analyzed in Study 02, a randomized urban layout was generated for each case (Figure 10). Then, a grayscale height-map of the surroundings was saved for each simulation (Figures 11 and 12).

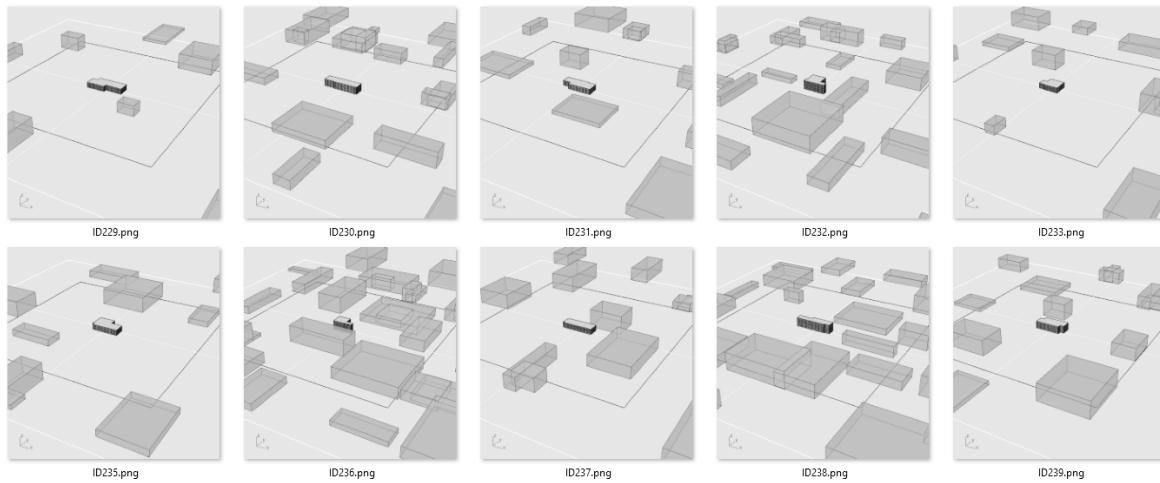


Figure 10. Randomized urban layouts for Convolutional Neural Network (CNN) training. A 3D view representing the model.

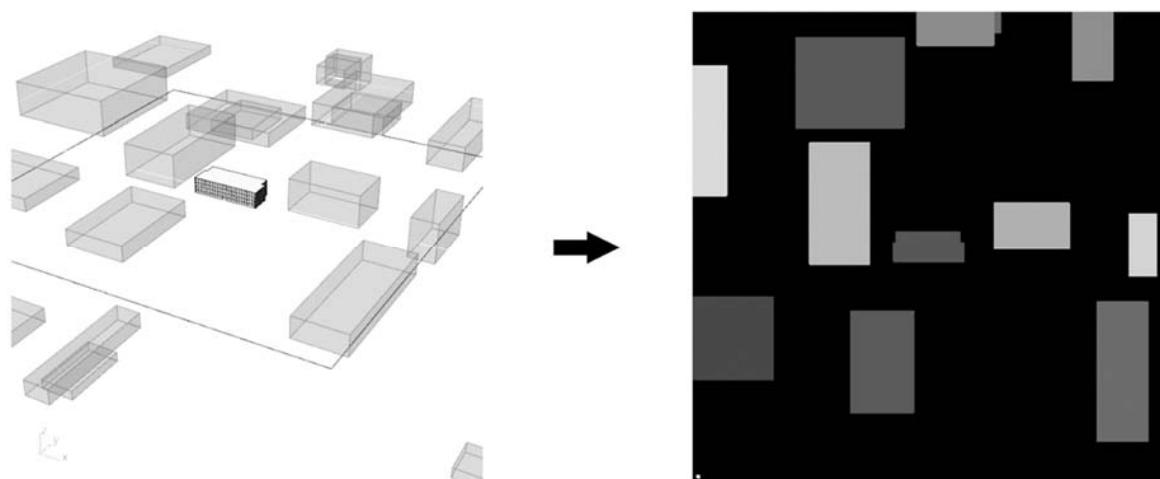


Figure 11. Example of translation of a 3D model into greyscale height map (Z-map).

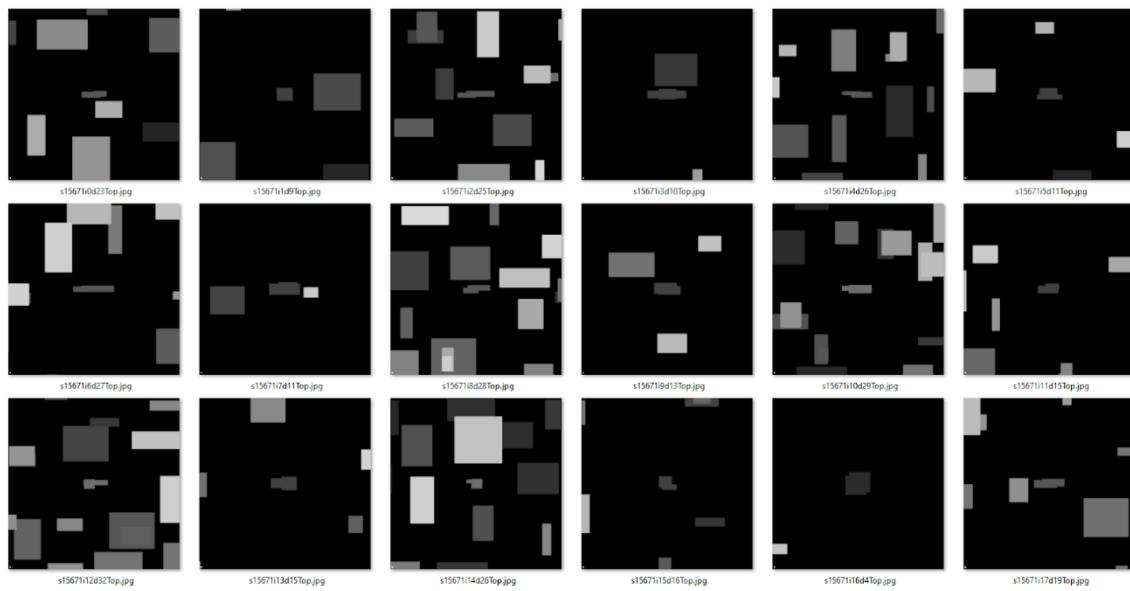


Figure 12. Randomized urban layouts for CNN training. In each case, the analyzed building is in the middle, while the surroundings represent various shading objects. The grayscale color code represents object height in linear scale, from black for the ground level (0 m), to white for 35 m and above.

The Neural Network has been designed according to suggestions by Rosebrock [30] on how to design a Convolutional Neural Networks for regression problems. The Network was built using Keras, an open-source neural-network library for Python [31]. The whole process was conducted in Python on a local computer.

The network architecture (Figure 13) consists of two sets of inputs: One for numerical data representing the analyzed building, and the second for image data representing the urban layout. The image is analyzed by a Convolutional Neural Network. The results from CNN are then flattened and concatenated with results from fully connected neural network that have analyzed the numerical data part. Finally, the network branches into two parts with fully connected layers. The two outputs from the network represent the Operational Carbon Footprint and Embodied Carbon Footprint.

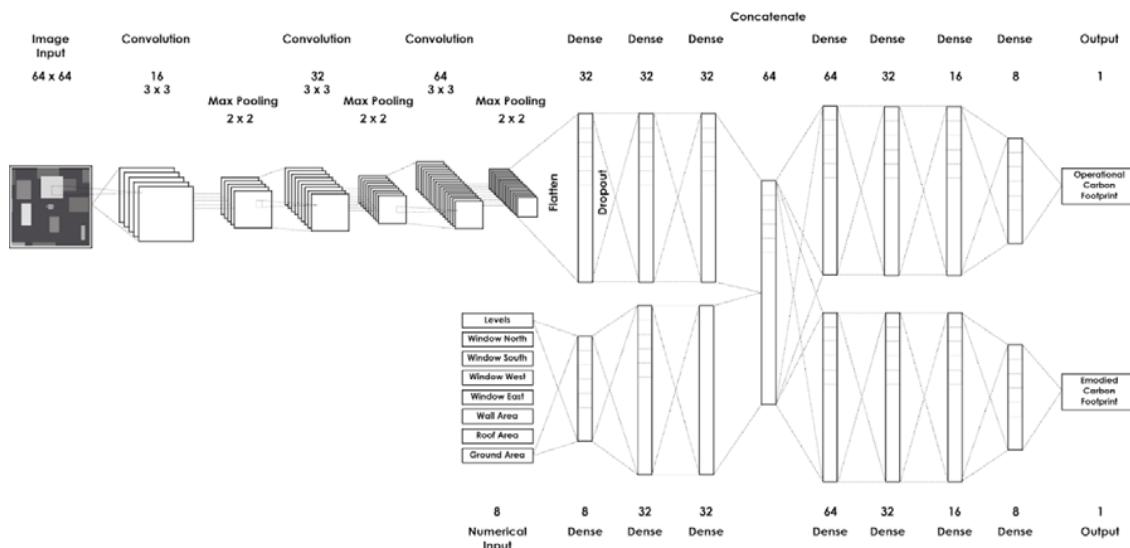


Figure 13. The network architecture combines graphical input analyzed by convolutional network and numerical input. Results of those partial neural networks are concatenated together. Then the network branches into two parts: One predicting the Operational Carbon Footprint and another predicting the Embodied Carbon Footprint of the building.

3.7. Model Performance and Results Discussion—Study 03

The network was trained for 300 epochs with a batch size of 10, on a 3000 row dataset with 75% training-test split. The training process took around an hour on a typical PC. The process of training was faster for Embodied Carbon as that part was not really influenced by the urban layout images (Figure 14). The Operation Carbon Footprint training required more epochs as this part was heavily influenced by the shading from the surroundings. The results of the training were very good, as the model had 99.6% explained variance for Operation Carbon Footprint and 99.7% explained variance for Embodied Carbon Footprint (see Figure 15). The average absolute percentage error was 0.6% for Embodied Carbon and 0.8% for Operational Carbon.

Running a dynamic energy simulation for a building of this size (on a typical Personal Computer) using the proposed tools (EnergyPlus) needed for precise Operational Carbon Footprint calculation, the part of Total Carbon Footprint would take approximately 1 min (30–90 s). The prediction generated using the final model from the study 03 took less than 1 s while running the script on the Python command prompt, or around 2–5 s while running the same script inside Grasshopper. This means that the user had almost instant feedback and could manipulate the building shape more freely, without waiting long for the results.

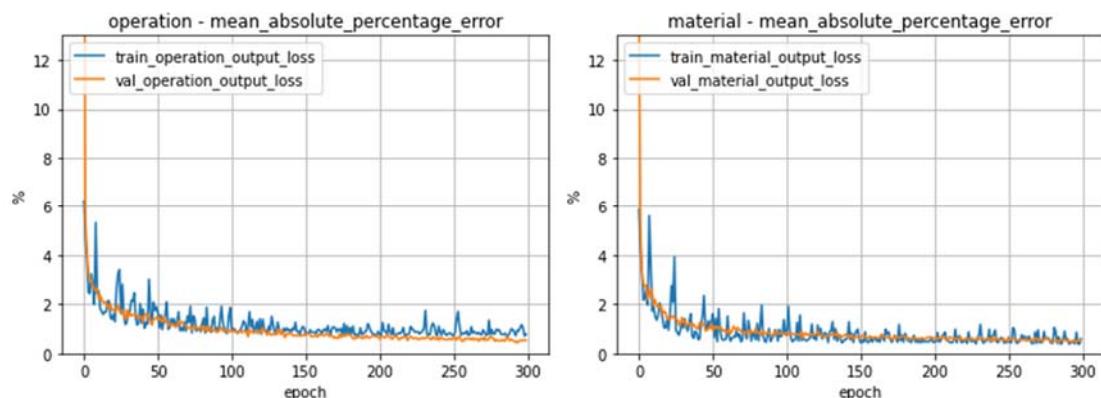


Figure 14. The plots visualize the Neural Network training process. It can be observed that the output for Material (Embodied Carbon) converges much faster. The Loss function value for Embodied Carbon almost stops improving after epoch 100. On the other hand, the output for Operation Carbon Footprint improves its Loss function value until the last epoch.

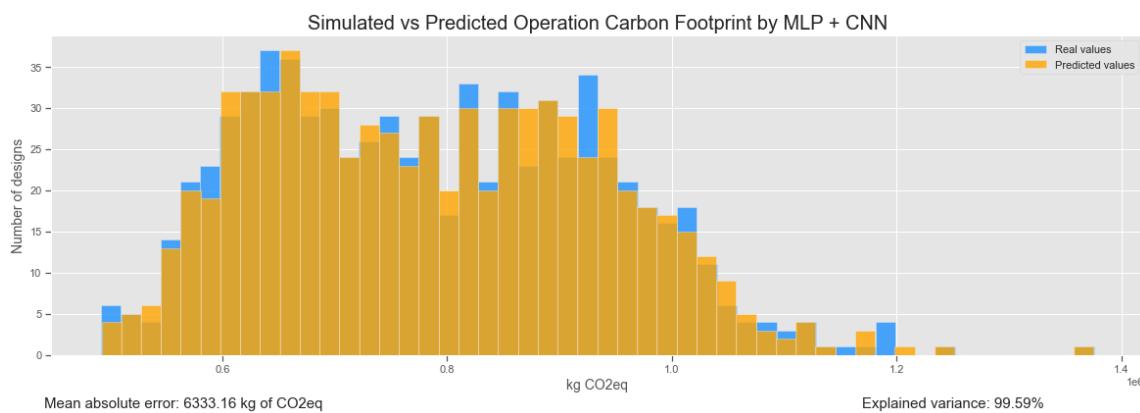


Figure 15. Cont.

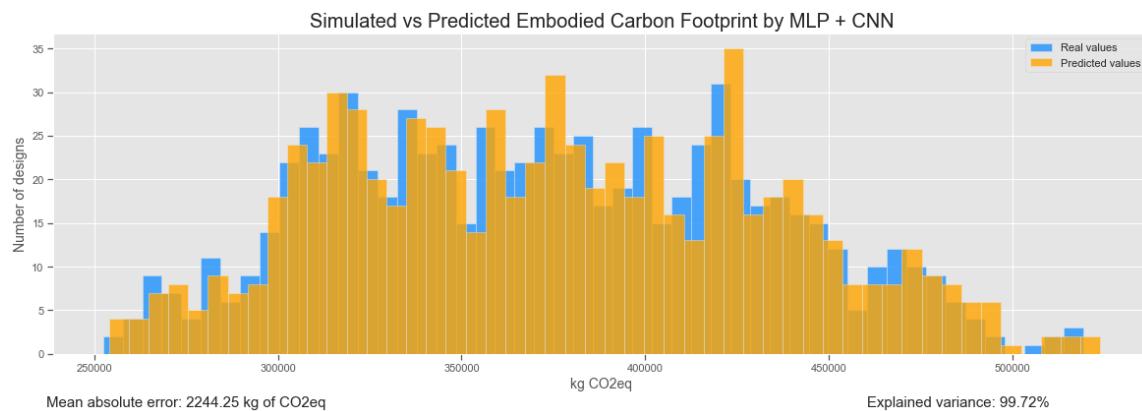


Figure 15. The plots confirm the predictions' low degree of error. The image shows the accuracy of the predictions in the validation set. The model has 99.6% explained variance for Operation Carbon Footprint and 99.7% explained variance for Embodied Carbon Footprint.

3.8. Testing the Model in a Design Situation—Study 03

To test the applicability of the Neural Network model in a real life situation, a following test has been taken. A set of buildings was modelled in Rhinoceros3D and then analyzed by a Grasshopper script that used the trained ML model. Subsequently, a simulation as performed for each case and the results were compared.

A total of 11 design options were modelled with a varying shape and size. The results are promising, as the mean absolute percentage error for the predictions appeared to be 0.4% for Embodied Carbon and 3.4% for Operational Carbon (see Figure 16). A much higher degree of error for the Operational Carbon could suggest that the model could still be improved by fine tuning the way the Neural Network is trained on the urban layout images, or by providing a bigger dataset.

Finally the trained model was linked to Grasshopper script, which allowed for the prediction of the building's Total Carbon Footprint in real time (see Figure 17). The script took into consideration the urban layout, including the influence of overshadowing on the Carbon Footprint of the building. The user could quickly verify how different plots or building positions influenced the environmental performance of the design.

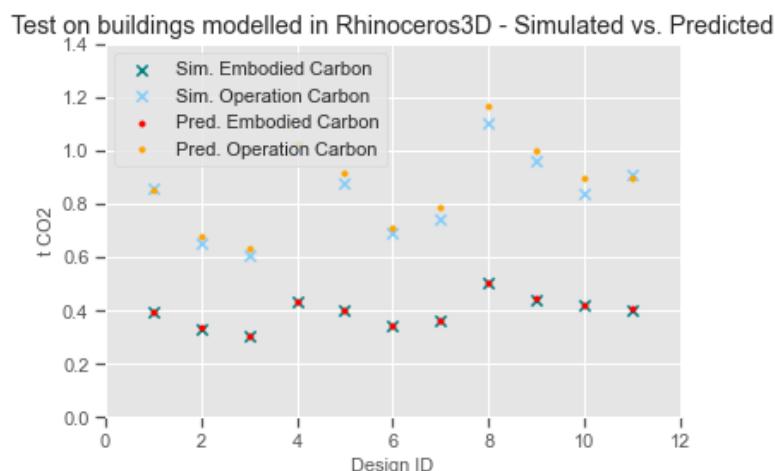


Figure 16. Testing the trained NN (Neural Network) model has proven that it can be used on building designs never previously seen by the algorithm. In this test, 11 buildings were modelled and then validated using prediction and simulation.

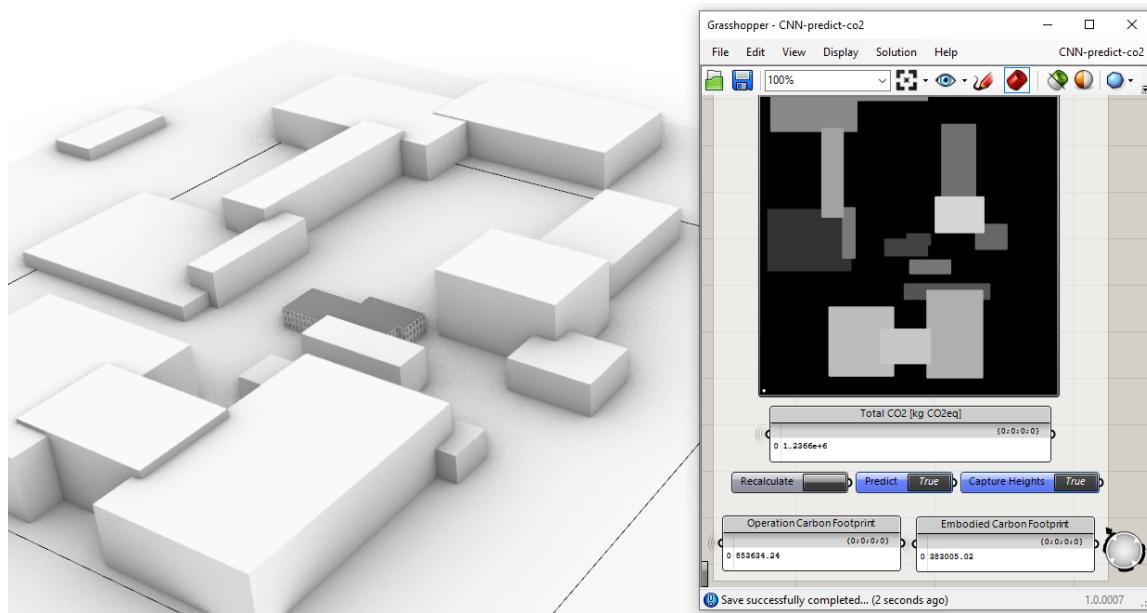


Figure 17. The updated script from Study 03 is able to predict the building’s Embodied and Operational Carbon Footprint based on the simplified model consisting of the building itself and simplified surroundings. The scripts take around 10–15 s to execute.

4. Discussion

The series of three studies addressed the issue of environmental analysis in the architectural design process. The task was to design a supplementary tool that could help designers to understand how their designs influence the environment. The Total Carbon Footprint of the building was selected as an environmental indicator. Starting from Study 1 to 3, each consecutive version of the tool accounted for more factors that influenced the environmental performance of the building. The first version of the tool could only work on cuboid-shaped buildings. The second version of the tool was modified to work on any prism-shaped building. Finally, in the third version, the influence of the urban surrounding on the building’s environmental performance was also considered.

As the tool is based on Machine Learning, it was important to verify the results against actual simulations. The tool achieved very high accuracy, with an error of no more than 3.4% for Operational Carbon in the third version of the tool.

The current version of the tool still has many limitations, but these can easily provide a topic for further research work. Firstly, the Machine Learning model has been trained only on buildings between a total area of 1500 and 3000 m². Outside of this range, the tool would have trouble calculating the proper value. Secondly, the tool has only been tested on single buildings, whereas no tests have been performed on groups of buildings. Thirdly, the tool has been trained on building simulations using preselected materials, structure, and components. The possibility to pick specific building components and see how it changes the result would be a big step forward.

All those limitations stem from the same problem: Machine Learning models need big amounts of data to properly train themselves. This has already been discussed by D’Amico et al. [23], who have suggested that the single problem with widespread adoption of Machine Learning in structural design lies in the lack of datasets to use for training.

It is important to compare this set of studies to other previous ones by different authors. For example, Seyedzadeh et al. [21] have made a review of different approaches of utilizing ML in relation to building energy consumption and performance (and also in relation to the CO₂ emissions reduction). The presented case studies however focus on engineering related solutions, while this article presents a tool, based on Machine learning, that enhances the architectural design process taking

into consideration the typical basic architectural features of the building (shape, fenestration and urban layout), while at the same time including not only Operational Carbon Footprint (which is directly linked to Energy Performance), but also Embodied Carbon footprint (which is related to material usage). The inclusion of both Operational and Embodied Carbon Footprint allowed one to observe interesting correlations. For example, the results show that the lowest sum of Total Carbon Footprint was found when none of the Embodied and Operational Carbon Footprints were lowest possible, which is an interesting conclusion of Study 01 (this can be observed in Figure 5). This result proves that a tool that quickly estimates Carbon Footprint could greatly help to reduce the environmental impact of the design, as a simple solution like “focusing on lowest Operational Carbon Footprint” (and thus the best Energy Performance) would not produce the best project option. The final tool is targeted at architects and architectural students, and should not produce 100% accurate Carbon Footprint estimation, but rather it should help designers to make conscious decisions.

Another important aspect is that this paper focuses on training a ML model based on a set of simulations, instead of real-world data. This approach has been suggested by Galanos [20], who has demonstrated how large datasets in the field of architectural design can be generated with the application of parametric design tools like Grasshopper for Rhinoceros3D. Such an approach enables the possibility of exploring the power of Machine Learning in architectural design, without access to big real-world datasets.

5. Conclusions

Artificial Intelligence and Machine Learning is already changing the way we design. Although this is currently happening at a slow pace, it can be predicted that this process will accelerate.

One of the areas in which ML models can help is simplifying the environmental assessment of the architectural designs. Trained ML models can be used to estimate the values of aspects like Total Carbon Footprint almost instantly, based on previously seen data. This can in turn help in evaluating other complex metrics, that are based on a collection of different measured features, for example when assessing the building circularity [6].

The three consecutive studies resulted in creating a working Machine Learning model that could aid architects in early design phases of developing environmentally conscious design models. The effect of the final Study 03 was a trained Neural Network that could predict the Total Carbon Footprint of a design proposal based on scarce information (wall area, ground floor area, roof area, height, window area south, north, west, east, and urban layout – all read automatically from the building and urban shape modelled by the user). Moreover, the model was tested on an actual concept design model modelled in Rhinoceros3D, and the predicted value was compared with an actual simulation, showing a very low degree of error. This proves that such a tool could be used in actual practice by qualified architects, but also architecture students and it could help make more decisions in early design stages with achieving lower carbon footprint, better regenerative design strategies, and with a lesser impact on the environment.

Author Contributions: Conceptualization, M.P.-M., E.R., and M.G.-S.; methodology, M.P.-M., E.R., and M.G.-S.; software, M.P.-M.; validation, M.P.-M.; formal analysis, M.P.-M.; investigation, M.P.-M.; resources M.P.-M., E.R., and M.G.-S.; data curation, M.P.-M.; writing—M.P.-M., E.R., and M.G.-S.; writing—review and editing, E.R. and M.G.-S.; visualization, M.P.-M.; supervision, E.R.; project administration, M.P.-M. and E.R. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. United Nations Environment Programme. *Building and Climate Change: Summary of Deicing-Makers*; United Nations Environment Programme: Washington, DC, USA, 2009.

2. Wiedmann, T.; Minx, J. A Definition of ‘Carbon Footprint’. In *Ecological Economics Research Trends*; Pertsova, C.C., Ed.; Nova Science Publishers Inc: New York, NY, USA, 2008; pp. 1–11.
3. CENELEC Management Centre. EN 15978:2011 *Sustainability of Construction Works—Assessment of Environmental Performance of Buildings—Calculation Method*; CENELEC Management Centre: Brussels, Belgium, 2011.
4. Ryńska, E.; Klimowicz, J.; Kowal, S.; Łyżwa, K.; Pierzchalski, M.; Rekosz, W. Smart Energy Solutions as an Indispensable Multi-Criteria Input for a Coherent Urban Planning and Building Design Process—Two Case Studies for Smart Office Buildings in Warsaw Downtown Area. *Energies* **2020**, *13*, 3757. [[CrossRef](#)]
5. Hollberg, A. *A Parametric Method for Building Design Optimization Based on Life Cycle Assessment*; Bauhaus-Universitätsverlag: Weimar, Germany, 2016.
6. Núñez-Cacho, P.; Górecki, J.; Molina-Moreno, V.; Corpas-Iglesias, F.A. What gets measured, gets done: Development of a circular economy measurement scale for building industry. *Sustainability* **2018**, *10*, 2340. [[CrossRef](#)]
7. Naboni, E.; Natanian, J.; Brizzi, G.; Florio, P.; Chokhachian, A.; Galanos, T.; Rastogi, P. A digital workflow to quantify regenerative urban design in the context of a changing climate. *Renew. Sustain. Energy Rev.* **2019**, *113*. [[CrossRef](#)]
8. Mang, P.; Haggard, B. *Regenerative Development & Design: A Framework for Evolving Sustainability*; Wiley: Hoboken, NJ, USA, 2012.
9. DeKay, M. *Integral Sustainable Design*; Taylor & Francis Ltd: London, UK, 2011.
10. Wahl, D.C. *Designing Regenerative Cultures*; Triarchy Press: Dorset, UK, 2016.
11. Grochulska-Salak, M.; ZInowiec-Cieplik, K. The Regenerative model of the City in view of Climatic Changes. In *Redefining Cities in View of Climatic Changes*; Warsaw University of Technology: Warsaw, Poland, 2019; pp. 135–136.
12. Boyd, P.C. Designing to Reduce Construction Costs. *J. Constr. Div.* **1976**, *102*, 587–592.
13. Ryńska, E. *Zintegrowany Proces Projektowania Prośrodowiskowego. Projektant a Środowisko*; Warsaw University of Technology: Warsaw, Poland, 2012.
14. Cudzik, J.; Radziszewski, K. Artificial Intelligence Aided Architectural Design. In *Computing for a Better Tomorrow*; Kępczyńska-Walczak Anetta, B.S., Ed.; Łódź Univniversity of Technology: Łódź, Poland, 2018; pp. 77–84.
15. Słyk, J. *Źródła Architektury Informacyjnej*; Warsaw University of Technology: Warsaw, Poland, 2012.
16. Burkov, A. *The Hundred-Page Machine Learning Book*; Kindle Direct Publishing: Seattle, WA, USA, 2019.
17. Belém, C.; Santos, L.; Leitão, A. On the Impact of Machine Learning Architecture without Architects. In Proceedings of the Hello, Culture! 18th International Conference, CAAD Futures, Daejeon, Korea, 26–28 June 2019.
18. İpek, E.; McKee, S.; Singh, K.; Caruana, R.; Supinski, B.; Schulz, M. Efficient architectural design space exploration via predictive modeling. *ACM Trans. Archit. Code Optim.* **2008**, *4*. [[CrossRef](#)]
19. Chaillou, S. *AI + Architecture/Towards a New Approach*; Harvard University: Harvard, MA, USA, 2019.
20. Galanos, T. Machine-Learned Regenerative Design. In *Regenerative Design in Digital Practice*; Naboni, E., Havinga, L., Eds.; Eurac Research: Bolzano, Italy, 2019; pp. 95–99.
21. Seyedzadeh, S.; Rahimian, F.; Glesk, I.; Roper, M. Machine learning for estimation of building energy consumption and performance: A review. *Vis. Eng.* **2018**. [[CrossRef](#)]
22. Newton, D. Generative Deep Learning in Architectural Design. *Technol. Archit. Des.* **2019**, *3*, 176–189. [[CrossRef](#)]
23. D’Amico, B.; Myers, R.J.; Sykes, J.; Voss, E.; Cousins-Jenvey, B.; Fawcett, W.; Richardson, S.; Kermani, A.; Pomponi, F. Machine Learning for Sustainable Structures: A Call for Data. *Structures* **2018**, *19*, 1–4. [[CrossRef](#)]
24. Peel, M.C.; Finlayson, B.L.; McMahon, T.A. Updated world map of the Köppen-Geiger climate classification. *Hydrol. Earth Syst. Sci.* **2007**, *11*, 1633–1644. [[CrossRef](#)]
25. EnergyPlus. Weather Data by Location—Warszawa Okecie 123750 (IMGW). Available online: https://energyplus.net/weather-location/europe_wmo_region_6/POL//POL_Warszawa.Okecie.123750_IMGW (accessed on 1 December 2019).
26. Bundesministerium des Innern für Bau und Heimat. ÖKOBAUDAT Datenbank. Available online: <https://www.oekobaudat.de/> (accessed on 1 December 2019).
27. ITB. ITB EPD Program. Available online: <https://www.itb.pl/epd.html> (accessed on 1 December 2019).

28. Polski Komitet Normalizacyjny. *PN-ISO 9836:1997 Performance Standards in Building—Definition and Calculation of Area and Space Indicators*; Polski Komitet Normalizacyjny: Warsaw, Poland, 1997.
29. Bonnet, R.; Hallouin, T.; Lasvaux, S.; Sibide, G. Simplified and reproducible building Life Cycle Assessment: Validation tests on a case study compared to a detailed LCA with different user's profiles. In Proceedings of the World SB14 Barcelona Conference, Barcelona, Spain, 28–30 October 2014; pp. 276–283.
30. Płoszaj-Mazurek, M. Machine Learning-Aided Architectural Design for Carbon Footprint Reduction. *Build. Sci.* **2020**, *276*, 35–39. [[CrossRef](#)]
31. Otovic, A.; Negendahl, K.; Bjerregaard Jensen, L. Expansion in Number of Parameters: Simulation of Energy and Indoor Climate in Combination with LCA. In Proceedings of the ASHRAE Annual Conference, St. Louis, MO, USA, 25–29 June 2016.
32. Roudsari, M.S.; Pak, M.; Smith, A. Ladybug: A Parametric Environmental Plugin for Grasshopper to Help Designers Create an Environmentally-Conscious Design. In Proceedings of the 13th Conference of International Building Performance Simulation Association, Chambery, France, 25–28 August 2013; pp. 3129–3135.
33. Veolia. Environmental Impact. 2018. Available online: https://energiadlawarszawy.pl/wp-content/uploads/sites/4/2019/04/wplyw_na_srodowisko_2018-1.pdf (accessed on 1 December 2019).
34. KOBiZE. Emission Factors. 2018. Available online: https://www.kobize.pl/uploads/materiały/materiały_do_pobrania/wskazniki_emisjynosci/Wskazniki_emisjynosci_2018.pdf (accessed on 1 December 2019).
35. Rozporządzenia Ministra Infrastruktury z dnia 12 Kwietnia 2002 r. w Sprawie Warunków Technicznych, Jakim Powinny Odpowiadać Budynki i Ich Usytuowanie; Warsaw, Poland 2002. Available online: https://www.polcen.com.pl/pliki/Prezentacja%20zmian_m.pdf (accessed on 1 December 2019).



© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).