# Machine learning for sustainable structures: a call for data

B. D'Amico[a,b,c], R. J. Myers[a,d], J. Sykes[e], E. Voss[e], B. Cousins-Jenvey[e], W. Fawcett[f], S. Richardson[g], A. Kermani[c], F. Pomponi*[a,b]

* Corresponding author: f.pomponi@napier.ac.uk

[a] REBEL (Resource Efficient Built Environment Lab) - Edinburgh (UK) – [b] School of Engineering and the Built Environment - Edinburgh Napier University, Edinburgh (UK) – [c] Centre for Timber Engineering - Edinburgh Napier University – [d] School of Engineering – University of Edinburgh, Edinburgh (UK) - [e] Expedition Engineering / Useful Projects Ltd., London (UK) - [f] Cambridge Architectural Research (CAR), Cambridge (UK) – [g] World Green Building Council, London (UK)

## Abstract

Buildings are the world's largest contributors to energy demand, greenhouse gases (GHG) emissions, resource consumption and waste generation. An unmissable opportunity exists to tackle climate change, global warming, and resource scarcity by rethinking how we approach building design. Structural materials often dominate the total mass of a building; therefore, a significant potential for material efficiency and GHG emissions mitigation is to be found in efficient structural design and use of structural materials.

To this end, environmental impact assessment methods, such as life cycle assessment (LCA), are increasingly used. However, they risk failing to deliver the expected benefits due to the high number of parameters and uncertainty factors that characterise impacts of buildings along their lifespans. Additionally, effort and cost required for a reliable assessment seem to be major barriers to a more widespread adoption of LCA. More rapid progress towards reducing building impacts seems therefore possible by combining established environmental impact assessment methods with artificial intelligence approaches such as machine learning and neural networks.

This short communication will briefly present previous attempts to employ such techniques in civil and structural engineering. It will present likely outcomes of machine learning and neural network applications in the field of structural engineering and – most importantly – it calls for data from professionals across the globe to form a fundamental basis which will enable quicker transition to a more sustainable built environment.

*Keywords:* sustainable; structural; materials; embodied carbon; life cycle assessment LCA; machine learning; neural networks.

## 1. Current situation

The built environment is the sector which puts the most pressure on the natural environment [1], and there is urgency in reducing and mitigating the environmental impacts caused by buildings [2]. This is particularly true for the non-operational life cycle stages (e.g. manufacturing of building materials and components, construction, dismantling, waste processing), whose impacts are currently unregulated and seldom calculated [3].

The structure, together with internal fitouts and their replacement cycles, accounts for often the largest mass in a building and has been found to contribute to the majority of its embodied carbon[1] emissions across the whole life cycle [4, 5]. In order to mitigate carbon emissions it is imperative to assess environmental impacts accurately and reliably. Attempts do exist where embodied carbondatabases have been created to facilitate the calculation and benchmarking of building structures and structural materials [6-9], but there remains considerable variations in the application of methodological approaches, data used, and transparency [1, 3]. This variation has led to substantial limitations in drawing conclusions and supporting decision making [3, 10].

The complexity of assessing the life cycle environmental impacts of buildings and building materials is justified by, and probably originates from, the inherent diversity of the construction sector. Construction projects involve multiple stakeholders over various life cycle stages and deal with many products and services – each with their own specific life cycles – that interact dynamically in both time and space [11, 12]. Therefore, even in detailed analyses of structural materials [3] or entire buildings [6], results regularly differ by two orders of magnitude, leaving decision makers unsure of which reliable value to pick.

In complex issues like the one just described, machine and deep learning, and neural networks can represent a further viable approach to problem solving and to support decision-making. This short communication does not aim to be an introduction to machine and deep learning, and neural networks but the interested reader can find thorough reviews and up to date information in the works of Schmidhuber [13] and Nielsen [14].

Significant advances in machine learning and neural networks, and increased applications of these methods, have been made in recent years [13]. In the following sections, we briefly review the current status of machine learning and neural network applications in structural and civil engineering and highlight their potential use in research concerning sustainability related decisions concerned with building structures

[1] We use embodied carbon as a shorthand to indicate the sum of $CO_2$eq emissions occurred due to all activities and components other than the operational energy consumption related to a building's life. More generally, embodied costs or impacts may refer to different units such as energy, carbon, water, natural resource depletion, etc. Carbon dioxide equivalent emissions are also the measuring unit of the Global Warming Indicator (GWI).

and structural materials. The short communication concludes with a call for data to access this yet untapped potential.

## 2. Machine and deep learning in structural and civil engineering

In 1997, Reich [15] reviewed the applications of machine learning in civil engineering at the time, suggesting that the use of machine learning in civil engineering was still in its infancy. However, the tone of his work was very optimistic seemingly suggesting that a broader uptake of machine learning in engineering was imminent. Twenty years have passed and we are yet to see such broader uptake but there is a more realistic hope that the time is now right due to pervasive computing and powerful computers, and the whole big data and internet of things revolution. Recent applications of machine learning and neural networks in structural engineering are mostly related to prediction and modelling of elastic properties of materials [inter alia 16, 17], compressive and bond strength of concrete [e.g. 18, 19], buckling load [20-22], development of cementitious composites [23], and the refinement finite element models [e.g. 24].

The reasons for these specific fields of applications is related to the nature of machine learning and neural networks, that is a large amount of initial data based on which the learning algorithm (or network) can be trained (Figure 1). All the above applications allow for a relatively easy access to large datasets through, for instance, repeated laboratory testing or computer modelling. However, the same cannot be said for embodied and life cycle carbon assessments of building structures and structural materials. Specifically, multiple specimens of concrete could be realised and tested in a controlled laboratory environment to gather the initial data (Figure 1) on, say, compressive strength to start the process. However, if we were to assess the embodied carbon of concrete in different building projects we need to know exact quantities, where the different materials are coming from, the distances they have travelled and the means of transportation, the exact construction activities taking place on site, and so on. This would then allow for the initial data that starts the process.

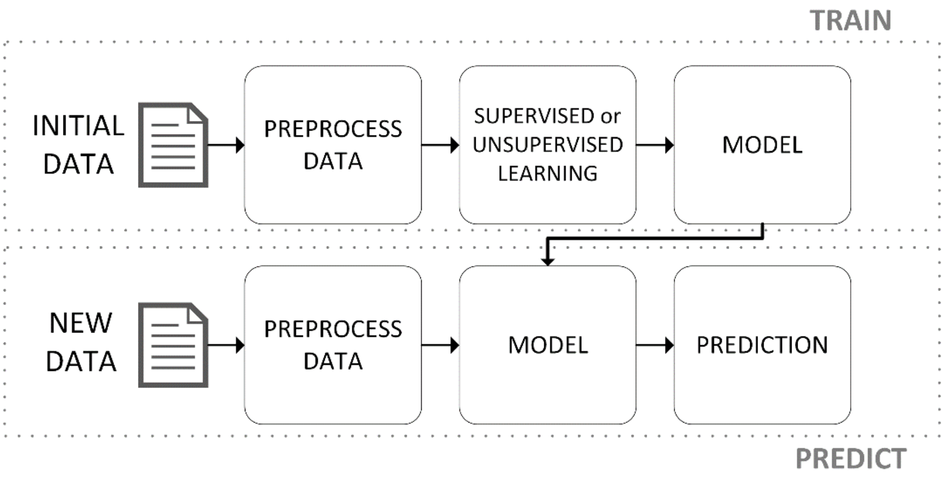


*Figure 1 - Typical workflow for machine learning – adapted from [25]*

It is evident from Figure 1 that having an initial, large-enough data population represents the first step in the machine learning process. Such an opportunity should not be missed, for the predictive power that machine learning and neural networks hold can be truly unexpected and surprising. For instance Gebru et al. [26] used deep learning to estimate the demographic makeup of US neighbourhoods. They found that the "resulting associations are surprisingly simple and powerful; […] for instance if a number of sedans encountered during a drive through a city is higher than the number of pickup trucks, the city is likely to vote for a Democrat during the next presidential election (88% chance); otherwise, it is likely to vote Republican (82%)" [26][2]. Such impressive conclusion however is solidly built on a large amount of initial data.

AI might help solve problems in showing how buildings can be reasonably divided up into sets. The benefit of categorising would be to produce guidance that starting from a building of, say, type A is able to place the likely impacts in a X-Y range, and suggests the best mitigation measures to reduce those impacts. This would be a considerable improvement, and certainly more efficient, than doing an LCA every single time to work our the impact range, the 'hotspots', and the mitigation measures. Automated and intelligent classification does therefore constitute the first barrier to overcome as shown in the example given by Pesto [27] who applied convolutional neaural networks to the problem of classifying US houses by architectural style.

Ultimately, data availability now represents the single and greatest barrier to utilise machine learning and neural networks in understanding the environmental impacts of building structures. Even in the most comprehensive embodied carbon database [6] there are only 144 data points for building structures. Unfortunately, a far greater data population is needed, which leads us to the next section.

## 3. Potential future applications: call for data

Given the current state of machine learning and neural networks, it is possible to foresee what possibilities a successful implementation of machine learning for sustainable structural engineering would unlock. We have envisaged two such opportunities that are within reach; they are shown in Figure 2.

[2] It is worth mentioning, that AI can establish causual links that are not yet understood and explained, and therefore we will have to be careful of correlation without causation.

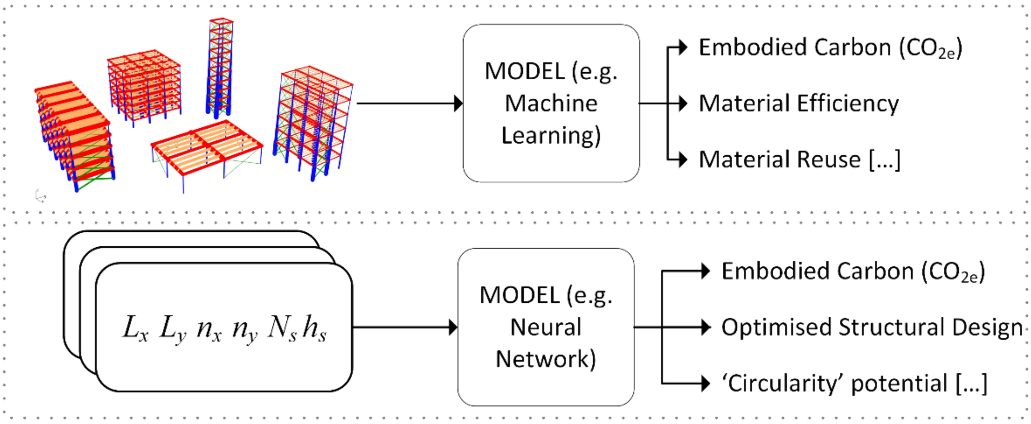


*Figure 2 - Examples of potential successful implementation of machine learning and neural networks in sustainable structural engineering*

For example, Figure 2 shows a case where a trained and validated machine learning algorithm is able to look at visual representation of building structures to accurately and reliably predict their embodied carbon estimates. Alternatively, it could produce a material efficiency index by looking at the difference between the structure fed to the algorithm and equivalent optimised structures (e.g. one with the same structural performance but optimised, i.e. not overdimensioned, structural members) on which the algorithm was trained.

Similarly, a neural network could take very few numerical inputs[3] related to the building structure to again estimate embodied or life cycle carbon emissions, produce optimised designs, or assess the circular economy potential by looking at material reuse at the end of the building's life [28]. The possibilities are endless but they all start with a large and reliable dataset of real and accurate designs of building structures.

For this reason, this short communication aims to launch a broad and open call for data to support such research direction. A submission platform has been created on the website of the Resource Efficient Built Environment Lab (REBEL) [29] where data on geometry, material, building type and other key parameters can be recorded and submitted. REBEL is facilitating this data gathering with an aim to advance sustainable structural engineering and material efficiency.  All data will be treated anonymously and in an aggregated manner and ethics approval has been obtained by the host organisation of the research project [30]. A Data Management Plan (DMP) has been submitted to, and approved by, the Engineering and Physical Sciences Research Council (EPSRC) and is available to any contributor upon request. The dataset that will eventually be put together will be made available freely to anyone for research and non-commercial purposes in agreement with the EPSRC policy framework on research data [31].

Specifically, we aim to collect as broad a set as possible within the following categories:

---

[3] In the example of Figure 2, $L_x$, $L_y$, $n_x$, $n_y$, $N_s$, and $h_s$ refer respectively to the length of the primary span, length of the secondary span, number of primary and secondary spans, number of storeys, and inter-storey height.

1. Design of structural frames with steel as the primary structural material,
2. Design of structural frames with reinforced concrete as the primary structural material,
3. Design of structural frames with structural timber as the primary structural material.

There is no limitation as to parameters such as footprint area, net usable area, number of storeys, or the shape of the buildings. In fact, the more varied the sample population the more accurate the final model will be in dealing with diverse inputs. Based on collected data, which will reflect real built assets, the research team will carry out a detailed and rigorous life cycle assessment (LCA) for each of them. To ensure that these LCAs will not fall short of transparency and reliability, a structured approach will be followed based on the newly published RICS guidance [32]. Data input will be consistently reported, data quality assessed, and assumptions justified and recorded, to ensure that the training population will be as reliable and as transparent as possible. This in turn will produce transparent, verifiable, and reliable predictions. Both the raw data on structural designs (in anonymised and aggregate form) and the LCAs will be made publicly available [31]. Parametric modelling will be used to increase the size of the sample population if necessary. This will form the Initial Data of Figure 1 to build and train the model, i.e. a population of structural designs and their related life cycle carbon emissions. Not to limit any perspective user from submitting a structure, we have decided not specify any file format. Regardless of the file (e.g. dxf, dwg, pdf, Revit model, Rhino model, SketchUp model etc.) a guided file upload system will ask the users relevant questions such as the building type (e.g. school, office, residential development, etc.), the main structural material, and a few others to ensure the dataset is harmonised across the key relevant parameters. There is no deadline for data submission as at any point new inputs can be used either to further train the model or to test its predictive power.

## 4. Concluding remarks

In this short communication, we have highlighted the potential to significantly increase the sustainability of building structures by adopting artificial intelligence approaches such as machine learning algorithm and neural networks.

However, the single and biggest impediment to doing so lies with the lack of large datasets of real-world structural designs across the main structural materials. The availability of such data will allow for more robust and accurate assessments of structural materials and construction details, which consequently create an opportunity for more reliable estimates of the life cycle GHG emissions. Additionally, such robust and reliable dataset will also enable analyses on material efficiency, circularity, wider environmental impact assessments—to name but a few. Not least, the combination of reliable data and AI will produce an unprecedented shift in terms of speed and ease when it comes to LCAs of building structures. In turn, this will widen the LCA user base and will increase the number of decisions which are based also on sustainability aspects and indicators.

We hope that as many stakeholders as possible will contribute their structural designs. Data will be anonymised and treated in an aggregate manner and therefore this sharing effort will truly contribute to the advancement of science. We will also feed back all research findings and outcomes to all organisations who contributed. If willing, organisations will also be acknowledged in white papers and publications arising from this research. This will show their genuine commitment to environmental sustainability, and the role they are playing towards a cleaner, low-carbon future. Mitigating the carbon emissions and environmental impacts caused by buildings is imperative and urgent – this is one opportunity to wholeheartedly contribute to worldwide efforts to tackle climate change and global warming.

## Acknowledgements

The authors gratefully acknowledge the funding received from the UK's Engineering and Physical Sciences Research Council (EPSRC) [Grant No. EP/R01468X/1] for this research.

## References

[1] F. Pomponi, A.M. Moncaster, Embodied carbon mitigation and reduction in the built environment – What does the evidence say?, J Environ Manage, 181 (2016) 687-700.

[2] C. De Wolf, F. Pomponi, A. Moncaster, Measuring embodied carbon dioxide equivalent of buildings: A review and critique of current industry practice, Energy and Buildings, 140 (2017) 68-80.

[3] F. Pomponi, A.M. Moncaster, Scrutinising embodied carbon in buildings: the next performance gap made manifest, Renewable & Sustainable Energy Reviews, 81 (2) (2018) 2431-2442.

[4] S. Kaethner, J. Burridge, Embodied CO2 of structural frames, The Structural Engineer, 90 (5) (2012) 33-40.

[5] M. Webster, H. Meryman, A. Slivers, T. Rodriguez-Nikl, L. Lemay, K. Simonen, Structure and Carbon-How Materials Affect the Climate, SEI Sustainability Committee. Reston, VA: American Society of Civil Engineers, (2012).

[6] K. Simonen, B.X. Rodriguez, C. De Wolf, Benchmarking the Embodied Carbon of Buildings, Technology| Architecture+ Design, 1 (2) (2017) 208-218.

[7] C. De Wolf, F. Yang, D. Cox, A. Charlson, A.S. Hattan, J. Ochsendorf, Material quantities and embodied carbon dioxide in structures, Proceedings of the ICE - Engineering Sustainability, (2015).

[8] C.C.E.L. De Wolf, Material quantities in building structures and their environmental impact, Massachusetts Institute of Technology, 2014.

[9] G.P. Hammond, C.I. Jones, Embodied energy and carbon in construction materials, Proceedings of the ICE - Energy, 161 (2008) 87-98.

[10] A. Säynäjoki, H. Jukka, J. Seppo, H. Arpad, Can life-cycle assessment produce reliable policy guidelines in the building sector?, Environmental Research Letters, 12 (1) (2017) 013001.

[11] M. Erlandsson, M. Borg, Generic LCA-methodology applicable for buildings, constructions and operation services—today practice and development needs, Building and Environment, 38 (7) (2003) 919-938.

[12] W. Collinge, A. Landis, A. Jones, L. Schaefer, M. Bilec, Dynamic life cycle assessment: framework and application to an institutional building, The International Journal of Life Cycle Assessment, 18 (3) (2013) 538-552.

[13] J. Schmidhuber, Deep learning in neural networks: An overview, Neural Networks, 61 (Supplement C) (2015) 85-117.

[14] M.A. Nielsen, Neural Networks and Deep Learning. [Also available at: http://neuralnetworksanddeeplearning.com/], Determination Press, 2015.

[15] Y. Reich, Machine learning techniques for civil engineering problems, Computer-Aided Civil and Infrastructure Engineering, 12 (4) (1997) 295-310.

[16] C. Zopf, M. Kaliske, Numerical characterisation of uncured elastomers by a neural network based approach, Computers & Structures, 182 (Supplement C) (2017) 504-525.

[17] G. Balokas, S. Czichon, R. Rolfes, Neural network assisted multiscale analysis for the elastic properties prediction of 3D braided composites under uncertainty, Composite Structures, 183 (Supplement C) (2018) 550-562.

[18] A. Cascardi, F. Micelli, M.A. Aiello, An Artificial Neural Networks model for the prediction of the compressive strength of FRP-confined concrete circular columns, Engineering Structures, 140 (Supplement C) (2017) 199-208.

[19] F. Yan, Z. Lin, X. Wang, F. Azarmi, K. Sobolev, Evaluation and prediction of bond strength of GFRP-bar reinforced concrete using artificial neural network optimized with genetic algorithm, Composite Structures, 161 (Supplement C) (2017) 441-452.

[20] U.K. Mallela, A. Upadhyay, Buckling load prediction of laminated composite stiffened panels subjected to in-plane shear using artificial neural networks, Thin-Walled Structures, 102 (Supplement C) (2016) 158-164.

[21] Z.u.R. Tahir, P. Mandal, Artificial neural network prediction of buckling load of thin cylindrical shells under axial compression, Engineering Structures, 152 (Supplement C) (2017) 843-855.

[22] S. Tohidi, Y. Sharifi, Neural networks for inelastic distortional buckling capacity assessment of steel I-beams, Thin-Walled Structures, 94 (Supplement C) (2015) 359-371.

[23] J.J. Biernacki, J.W. Bullard, G. Sant, K. Brown, Fredrik P. Glasser, S. Jones, T. Ley, R. Livingston, L. Nicoleau, J. Olek, F. Sanchez, R. Shahsavari, P.E. Stutzman, K. Sobolev, T. Prater, Cements in the 21st century: Challenges, perspectives, and opportunities, Journal of the American Ceramic Society, 100 (7) (2017) 2746-2773.

[24] Y.-S. Park, S. Kim, N. Kim, J.-J. Lee, Finite element model updating considering boundary conditions using neural networks, Engineering Structures, 150 (Supplement C) (2017) 511-519.

[25] R. Adams, Machine Learning for Predictive Modelling, in: MATLAB Expo 2015 UK, 2015.

[26] T. Gebru, J. Krause, Y. Wang, D. Chen, J. Deng, E.L. Aiden, L. Fei-Fei, Using deep learning and Google Street View to estimate the demographic makeup of neighborhoods across the United States, Proceedings of the National Academy of Sciences, (2017) 201700035.

[27] C. Pesto, Classifying U.S. Houses by Architectural Style Using Convolutional Neural Networks, in, Stanford University, 2017.

[28] F. Pomponi, A. Moncaster, Circular economy for the built environment: A research framework, Journal of Cleaner Production, 143 (2017) 710-718.

[29] Resource Efficient Built Environment Lab (REBEL), Call for Data: Machine Learning and Neural Networks for Sustainable Structural Engineering. Available at: http://rebelonline.blog/2017/12/12/call-for-data, 2017.

[30] EPSRC, EP/R01468X/1. Getting the numbers right and getting the right numbers: quantifying the embodied carbon of building structures [PI: Dr F Pomponi], 2017.

[31] EPSRC, Engineering and Physical Sciences Research Council policy framework on research data. [Available at: https://www.epsrc.ac.uk/about/standards/researchdata/ Last Accessed 11th Dec 2017].

[32] RICS, Whole life carbon assessment for the built environment, in, Royal Institution of Chartered Surveyors, London, 2017.