Real-time Environmental Feedback at the Early Design Stages

Creating a conceptual analysis tool by teaching artificial neural networks with design inputs and monitored energy consumption data

Greig Paterson[1], Sung Min Hong[2], Dejan Mumovic[3], Judit Kimpian[4]
[1,3]Bartlett School of Graduate Studies, UCL, UK, [2]UCL Energy Institute, UK, [4]Aedas R&D, UK
[1,3]http://www.bartlett.ucl.ac.uk/graduate, [2]http://www.bartlett.ucl.ac.uk/energy, [4]http://www.aedas.com/
[1]greig.paterson.10@ucl.ac.uk, [2]s.hong@ucl.ac.uk, [3]d.mumovic@ucl.ac.uk, [4]judit.kimpian@aedas.com

Abstract. *It has been argued that traditional building simulation methods can be a slow process, which often fails to integrate into the decision making process of non-technical designers, such as architects, at the early design stages. Furthermore, studies have shown that predicted energy consumption of buildings during design is often lower than monitored energy consumption during operation.*
In view of this, this paper outlines research to create a user friendly design tool that predicts energy consumption in real-time as early design and briefing parameters are altered interactively. As a test case, the research focuses on school design in England. Artificial neural networks (ANNs) were trained to predict the energy consumption of school designs by linking actual heating and electrical energy consumption data from the existing building stock to a range of design and briefing parameters.
Keywords. *Environmental design tool; energy prediction; artificial neural networks; building operational performance; schools.*

INTRODUCTION

There are many environmental 'design aids' available, with the objective of helping designers make sustainable design decisions. These design aids can largely be grouped into the following categories (Morbitzer, 2003):

- Design guidelines / rules of thumb
- Steady state calculation methods
- Correlation based methods
- Physical modelling
- Building simulation

Given that environmental design problems tend to be 'wicked' (Rittel and Webber, 1973), and thus distinctly novel and unique, rules of thumb, basic calculations and correlation methods are often inadequate techniques (Morbitzer, 2003; Pratt and Bosworth, 2011) and physical modelling has the disadvantage of being very costly (Morbitzer, 2003). When used correctly, the most powerful design aid available for the analysis of environmental performance is building simulation (Morbitzer, 2003). Building simulation is, however, rarely used by architects at the early design stages (Pratt and Bosworth, 2011).

Architect and psychologist Lawson (2004; 2006) states that simulation tools are not 'design' tools but 'evaluation' tools which are used to assess designs after they have been designed. A major barrier is the time taken to input all the required information, such that the designer can only afford to do it after the major design decisions have been made (Lawson, 2004). Also, the design space is constrained by the fact that commonly used building simulation tools produce static design proposals - it is therefore difficult, given time and economic constraints, to produce a wide range of design options (Pratt and Bosworth, 2011). In this way, the design space is sparingly populated because the models are discrete rather than continuous, thus omitting 'in- between' solutions (Pratt and Bosworth, 2011).

Furthermore, research, such as that carried out by CarbonBuzz [1], highlight the fact that the actual energy consumption of buildings regularly exceeds the design estimates, often by more than double.

Real world problems have complex and non-linear interactions, therefore system behavior is often best learned through observations rather than modelling (Samarasinghe, 2007). In view of this, an alternative approach at predicting energy consumption in buildings is to collect large amounts of actual energy and design data and analyse the patterns between the two. One such source of actual 'observed' energy data in the UK are Display Energy Certificates (DECs) (CIBSE, 2009). One method of learning the relationships between energy consumption and design inputs are artificial neural networks (ANNs). ANNs are machine learning techniques inspired by the structure and processes of biological neural networks that take place within the brain (Haykin, 1999). ANNs were found to be suitable for assessing determinants of energy use in higher education buildings in London, UK (Hawkins et al., 2012).

In light of the above, two questions emerge:

1. Can an ANN based method for a design tool be developed that offers non-technical users the ability to predict energy consumption in real-time as they explore the design space?
2. Can such a tool be based on actual energy consumption, rather than simulated data, in an accurate manner?

As a test case, the research focuses on school design in England. The purpose of this paper is firstly to summarise the data collection process and describe the ANN method. Finally, the tool user interface development and preliminary results will be presented.

DATA COLLECTION

The data collection process was a desktop study with the aim of collecting as much design and briefing data as is freely available on hundreds of schools across England. Table 1 and Table 2 outline the input and output parameters for the ANN models. The energy data used to train the ANNs were sourced from the Display Energy Certificate (DEC) database, which are stored in the non-domestic energy performance register maintained by Landmark [2]. The annual electricity and heating fuel use (kWh/m^2/annum) figures were used as the output in this study. The following criteria were used to select the school buildings for analysis, ensuring the buildings are comparable with each other:

- The school has a valid DEC
- The school has one main building
- Age of construction and material use are consistent

Data on 465 schools have thus far been collected.

In addition to energy consumption, other data collected from the DEC database were total useful floor area (m^2) and building environmental conditioning type. The number of pupils in each school was gathered from the Department for Education's (UK) EduBase public portal [3] and heating and cooling degree days were acquired from the Central Information Point [4]. The geometric and site data were gathered by measurement or visual inspec-

Table 1
ANN inputs.

Input Parameter	Input Neuron Type	Data Range / Activation Criteria	Description
Construction Year	Continuous	1860-2010	Year the school was built
Phase of Education	Binary	(-1) Primary/elementary, (1) secondary/high school	Primary schools or secondary schools/sixth form colleges
Number of Pupils	Continuous	44-2013	Part-time pupils divided by 2, plus the number of full-time pupils
Internal Environmental Conditioning	Categorical	(-1) Nat. vent, (0) mixed mode, (1) mech. vent	Primary internal environmental conditioning strategy
Site Exposure	Categorical	(-1) Exposed, (0) semi-sheltered, (1) sheltered	'Exposed': no obstructions present (4 x the height of the school away); 'semi-exposed': obsts. lower than the school; 'sheltered': obsts. taller than the school.
Orientation	Continuous	-45° - +45°	Angle at which the external walls differ from absolute north, south, east and west. Positive angle for clockwise orientations.
North Façade Adjacency	Binary	(-1) Open, (1) obstructed	Obstructed if a building or tree is within 1 x the height of the building from the majority of the façade orientation
South Façade Adjacency	Binary	(-1) Open, (1) obstructed	See North Façade Adjacency
East Façade Adjacency	Binary	(-1) Open, (1) obstructed	See North Façade Adjacency
West Façade Adjacency	Binary	(-1) Open, (1) obstructed	See North Façade Adjacency
Floor Area	Continuous	$861m^2$-$15396m^2$	Total usable floor area
Building Depth Ratio	Continuous	2.50-16.60	Building volume / exposed external wall area
Compactness Ratio	Continuous	1.01-4.59	Perimeter of the building footprint / perimeter of a circle with the same area as the building footprint
Surface Exposure Ratio	Continuous	1.71-5.67	Building volume / exposed surface area
North Glazing Ratio	Continuous	0.00-0.13	Glazed area on the north façade / total floor area
South Glazing Ratio	Continuous	0.00-0.15	Glazed area on the south façade / total floor area
East Glazing Ratio	Continuous	0.00-0.11	Glazed area on the east façade / total floor area
West Glazing Ratio	Continuous	0.00-0.14	Glazed area on the west façade / total floor area

Input Parameter	Input Neuron Type	Data Range / Activation Criteria	Description
Glazing Type	Binary	(-1) Single, (1) double	Single or double/secondary glazing
Roof Shape	Binary	(-1) Pitched, (1) flat	Pitched or flat roof
Roof Glazing	Binary	(-1) None, (1) glazing	Existence of any roof glazing
Heating Degree Days	Continuous	1635.6-2843.3	Heating degree days during the DEC monitoring period
Cooling Degree Days	Continuous	73.9.7-425.2	Cooling degree days during the DEC monitoring period

Table 1 continued ANN inputs.

Output	Output Neuron Type	Data Range	Description
Heating Energy Consumption	Continuous	7-272kWh/m^2/annum	Annual heating fuel use
Electricity Energy Consumption	Continuous	7-95kWh/m^2/annum	Annual electricity fuel use

Table 2 ANN outputs.

tion from the online map software Digimap [5], Bing Maps [6] and Google Earth [7].

The building height was derived by multiplying the average number of storeys by 3.62m - the average floor-floor height of schools in the UK (Steadman et al., 2000). The building volume was then derived by multiplying the building height with the building footprint area, measured from Digimap [5]. Glazing percentages were measured from Bing Map [6] images using bespoke code developed in the Processing programming environment [8].

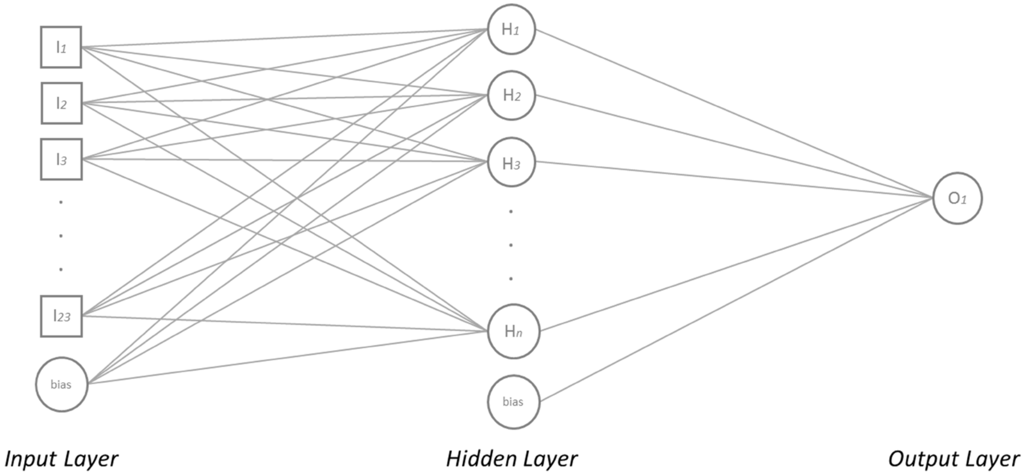
The construction year of the buildings were collected from each school's website where available otherwise they were derived from historical digital map software [5]. Data on schools of varying ages were collected to increase the size of the database, giving the neural network more data to learn from. A proportion of the differences in, for example, fabric quality and building systems between newer schools and older schools are likely to be picked up in the construction year neuron. Therefore, this neuron will exist within the trained network in the final design tool but fixed to the most recent date.

ANN ARCHITECTURE

All ANNs were constructed in Matlab [9]. The aim of the ANN method is to predict the energy consumption outputs (Table 2) based on a set of inputs (Table 1). A multilayer perceptron network was used for the study - Figure 1 shows the conceptual structure of this ANN. The hidden layer enables the system to generate nonlinear and complex relationships by intervening between the input and output neurons (Haykin, 1999). Each neuron in the input and output layer took continuous, categorical or binary values as outlined in Table 1 and Table 2. Prior to the training of the network, all continuous inputs were normalised to values between -1 and 1 to generalise the calculation process. Two ANN models were constructed, one with heating energy consumption as an output and one with electrical energy consumption as an output - both ANN models included all of the input parameters (Table 1).

A Levenberg-Marquardt backpropagation supervised training technique was used to train the feedforward network to recognise the patterns that exist in the dataset. The prediction performance of the ANN was assessed by validating the ANN with 10% of the gathered database on which the ANN had not been trained - the testing dataset. 10% of the gathered database was used to stop the training process before overlearning occurred (Demuth et al., 2008) and the remaining 80% of the database was used to train the network. The number of neu-

Figure 1
Conceptual Structure of the ANN.



rons in the hidden layer were altered between 2, 4, 8, 16 and 32 neurons. Each network configuration was trained five hundred times and the ANN with the lowest mean squared error (1) was selected for further analysis. Further analysis consisted of calculating the coefficient of determination (R^2) and the below performance indicators, (2) and (3):

Mean squared error (MSE) = $\sum_i^n \frac{(\hat{Y}_i - Y_i)^2}{n}$ (same units as output) (1)

Root-mean squared error (RMSE) = $\sqrt{\frac{\sum_i^n (\hat{Y}_i - Y_i)^2}{n}}$ (same unit as output) (2)

Mean absolute percentage error (MAPE) = $\frac{\sum_i^n \frac{|\hat{Y}_i - Y_i|}{Y_i}}{n}$ (%) (3)

Where Y_i and $\hat{Y}_i$ are the target and predicted outputs respectively for the training, testing or stopping configuration *i* and *n* is the total number of configurations in the training, testing or stopping datasets.

USER INTERFACE

Figure 2 shows a representation of the tool user interface. The tool is currently being developed in the Processing programming environment [8]. The ANN algorithms are integrated into this environment with MATLAB Builder JA [10].

The tool allows the user to sketch the footprint of the building by clicking and dropping vertices in an input window - these vertices can later be dragged or deleted. All other inputs are entered via sliders (continuous inputs) and tick boxes (categorical/binary inputs) thereby encouraging the user to 'play' and test different options, encouraging exploration of 'in-between' solutions in the design space. The ability to gain feedback in real-time results in the user being able to 'animate' the results and learn the relationships between the design inputs and energy outputs by the acceleration of change in the results as the design space is explored.

RESULTS AND DISCUSSION

ANN configurations with two and eight neurons in the hidden layer were found to produce the least prediction errors for heating and electricity energy consumption respectively. Table 3 summarise the results of the errors for the best performing ANN configurations. The electricity output was predicted with a mean absolute percentage error (MAPE) of 19.3%, while the heating output was predicted with a MAPE of 20.5%. These errors are an improvement

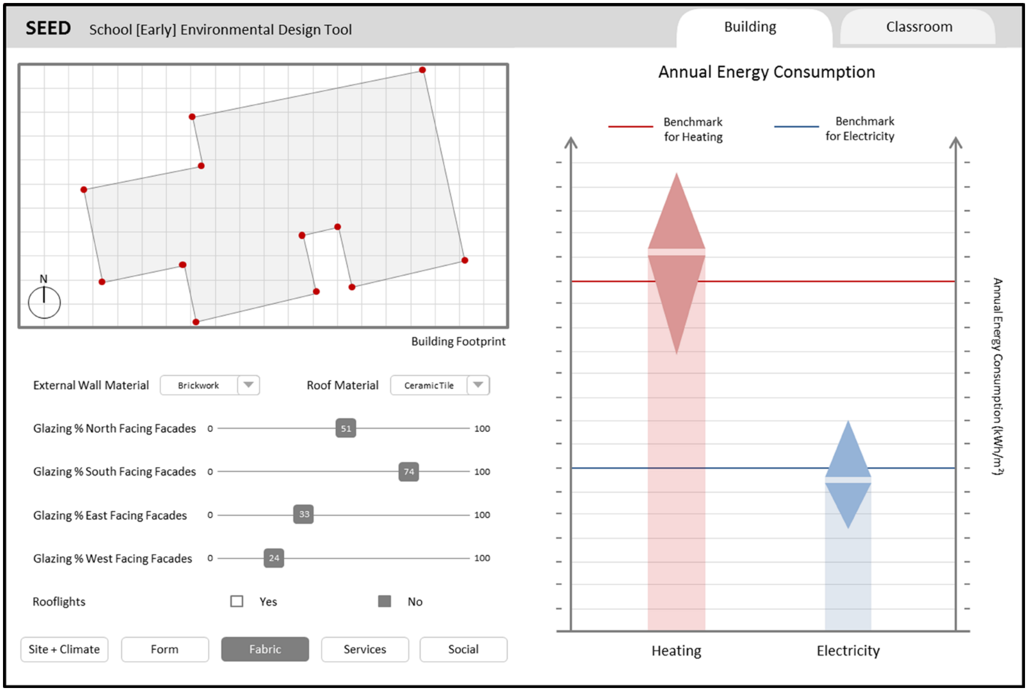


Figure 2
Representation of the user interface.

of 10.0% and 6.7% for heating and electricity energy consumption respectively, when compared to using the Chartered Institution of Building Services Engineers (CIBSE) Technical Memorandum 46 (TM46) Energy Benchmarks as energy performance indicators (Table 4). As mentioned in the introduction, Hawkins et al. (2012) used an ANN method to assess the energy determinants in higher education buildings in London, UK. The ANN method by Hawkins et al. produced MAPEs of 25.1% and 34.8% for heating and electricity fuel use respectively - the results from the research in this paper better these errors by 4.6% and 15.5% for heating and electricity respectively.

Figure 3 show scatter plots of the ANN predictions vs actual annual heating and electricity energy consumption from the testing dataset. The coefficient of determination (R^2) shows that the 23 design and briefing parameters (ANN inputs) explain 39% and 41% of the variation in annual heating and electricity energy consumption of the schools respectively.

From this initial study it appears that the ANN

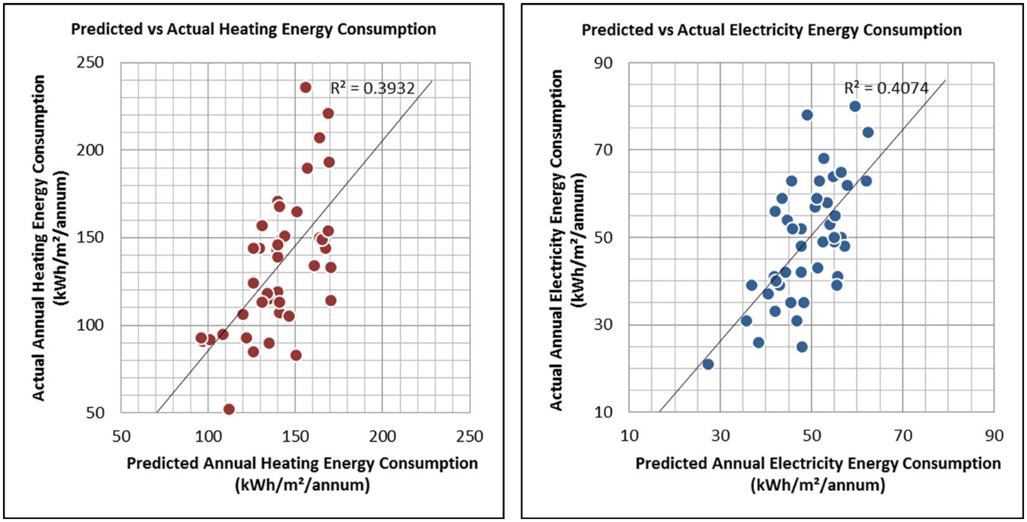
ANN Output	RMSE (kWh/m²/annum)	MAPE (%)
Heating Energy Consumption	30.5	20.5
Electricity Energy Consumption	10.8	19.3

Table 3
Prediction errors of the ANNs - calculated from the ANN testing dataset.

TM46 Benchmark	RMSE (kWh/m²/annum)	MAPE (%)
Heating Energy Consumption	41.3	30.5
Electricity Energy Consumption	16.7	26.0

Table 4
Prediction errors of the CIBSE TM46 Benchmarks - calculated from the ANN testing dataset.

Figure 3
Scatter plots of predicted ANN vs actual heating (left) and electricity (right) energy consumption.



method is viable for predicting energy consumption in existing school buildings. Nevertheless, further research is planned to improve the performance of this method and ensure it is viable for new school designs as outlined in the further work section.

CONCLUSION

This paper outlines research to create a user friendly design tool that predicts energy consumption in real-time as early design and briefing parameters are altered interactively. As a test case, the research focused on school design in England. Artificial neural networks (ANNs) were trained to predict the energy consumption of school designs by linking actual heating and electrical energy consumption data from the existing building stock to a range of design and briefing parameters. The initial design of the user interface was introduced in this paper.

For the energy consumption predictions, the ANN mean absolute percentage error (MAPE) was 20.5% for heating and 19.3% for electricity. The coefficient of determination (R^2) was 39% and 41% for heating and electricity energy consumption respectively. The aforementioned errors were compared with another method and study and produced lower errors, as outlined in the previous section. Nevertheless, it is desirable to reduce these errors further and improve the R^2 values. In order to improve both the performance of the ANN method and increase the relevance of the tool, further design inputs are likely to be required. The nature of this desktop study was to collect as many design and briefing inputs as are freely available. Acquiring further inputs, such as building services and fabric data, may require direct communication with individual schools or local authorities. This process is likely to be time consuming however is being pursued. Further actions to improve the ANN performance, as well as ensuring the tool is relevant to the design process and applicable to new school designs, are outlined in the following section.

It should be noted that the development of this tool does not have the objective of replacing traditional building simulation - instead it aims to act as a user friendly sanity check for non-technical designers, such as architects, at the early design stages.

FURTHER WORK

There are a number of developments underway in order to make the method of prediction in this research more accurate and the use of the design tool more relevant. These developments include collecting data on more schools across England; the pursuit of additional input parameters, as mentioned in the previous section; refining the input parameters for heating and electricity energy predictions separately; and the exploration of alternative ANN architectures. The final tool will go through a validation process using a number of new schools as case studies to ensure the method is applicable to new school designs. Finally, as the user interface develops, it will be tested by focus groups within industry.

ACKNOWLEDGEMENTS

This research was undertaken as part of an Engineering Doctorate (EngD) project at the UCL EngD Centre in Virtual Environments, Imaging and Visualisation (VEIV). The work is sponsored by the Engineering and Physical Sciences Research Council (EPSRC) and Aedas, with the R&D team within Aedas providing much support and collaboration.

REFERENCES

CIBSE, 2008, *TM46 Energy Benchmarks*, Chartered Institution of Building Services Engineers Publications, London, UK.

CIBSE, 2009, *CIBSE TM47 Operational Ratings and Display Energy Certificates*, Chartered Institution of Building Services Engineers Publications, London, UK.

Demuth, H, Beale, M, Hagan, M, 2009, *Neural Network Toolbox: User's Guide,* Mathworks, MA, USA.

Hawkins, D, Hong, SM, Raslan, R, Mumovic, D, Hanna, S, 2012, 'Determinants of Energy Use in UK Higher Education Buildings Using Statistical and Artificial Neural Network Methods', *International Journal of Sustainable Built Environment*, 1(1), pp. 50-63.

Haykin, S, 1999, *Neural Networks: A Comprehensive Foundation*, Prentice Hall, NJ, USA.

Lawson, B, 2004, *What Designers Know*, Architectural Press, Oxford, UK.

Lawson, B, 2006, *How Designers Think: The Design Process Demystified*, 4th Edition, Architectural Press, Oxford, UK.

Morbitzer, C, 2003, *Towards the Integration of Simulation into the Design Process*, PhD Thesis, University of Strathclyde, Glasgow, UK.

Pratt, KB, Bosworth, DE, 2011, 'A Method for the Design and Analysis of Parametric Building Energy Models', *Proceedings of Building Simulation 2011: 12th Conference of International Building Performance Simulation Association*, Sydney, Australia, pp. 2499-2506.

Rittel, HW, Webber, MM, 1973, 'Dilemmas in a General Theory of Planning', *Policy Sciences*, 4, pp. 155-69.

Samarasinghe, S, 2007, *Neural Networks for Applied Sciences and Engineering: From Fundamentals to Complex Pattern Recognition*, Auerbach Publications, NY, USA.

Steadman, P, Bruhns, HR, Rickaby, PA, 2000, 'An Introduction to the national Non-Domestic Building Stock Database', *Environment and Planning B: Planning and Design*, 27(1), pp. 3-10.

[1] http://www.carbonbuzz.org/
[2] https://www.ndepcregister.com/
[3] http://www.education.gov.uk/edubase/home.xhtml
[4] http://www.landmark.co.uk/solutions/registers/nondomestic/cip
[5] http://digimap.edina.ac.uk/
[6] http://www.bing.com/maps/
[7] http://www.google.com/earth/index.html/
[8] http://processing.org/
[9] http://www.mathworks.co.uk/
[10] http://www.mathworks.co.uk/products/javabuilder/