Accepted manuscript doi: 10.1680/jener.23.00017

Submitted: 02 April 2023

Published online in 'accepted manuscript' format: 10 November 2023

Manuscript title: Prediction of energy performance of residential buildings using regularized neural models

Authors: Komal Siwach[1], Harsh Kumar[2], Nekram Rawal[3], Kuldeep Singh[4] and Anubhav Rawat[5]

Affiliations: [1]Maharshi Dayanand University, Rohtak, India; [2]Indian Institute of Technology Delhi, India; [3]Civil Engineering Department, Motilal Nehru National Institute of Technology-Allahabad, Prayagraj, India; [4]Faculty of Engineering, University of Nottingham, Nottingham, UK and [5]Applied Mechanics Department, Motilal Nehru National Institute of Technology-Allahabad, Prayagraj, India

Corresponding author: Anubhav Rawat, Applied Mechanics Department, Motilal Nehru National Institute of Technology-Allahabad, Prayagraj – 211004, India.

E-mail: anubhav-r@mnnit.ac.in

Accepted manuscript doi: 10.1680/jener.23.00017

Abstract

Human habitats are one of the major consumers of energy. Therefore, in the current age of increasing carbon footprints, analyzing energy efficiency of a building is imminent, which has been taken up in the current work. Machine learning based Artificial Neural Network-ANN approach is used in the current work to study building-energy-performance. Total eight parameters; relative compactness, surface area, wall area and roof area of the building, overall height, and orientation of the building, glazing area and its distribution are selected as the input parameters and heating and cooling loads as the output parameters. The network prediction capability was checked by comparing the predictions of the ANN architecture with the benchmark test case. A well trained and validated ANN is used to predict 96 conditions by varying glazing area and glazing area distribution. ANN is found to capture the physics efficiently. This study revealed that there is a significant potential to improve the energy efficiency of the building and the maximum saving in the cooling load can be as high as 20.67% for a fraction of the glazing areas equal to 0.15 if glazing area distribution is kept 32.5% in North, and 22.5% each in the East, South and West.

Keywords: Energy; Building Energy Performance; Machine Learning; Neural Network; Glazing Area; Heating and Cooling Load; Built Environment

Accepted manuscript doi: 10.1680/jener.23.00017

INTRODUCTION

In today's era of globalization, industrialization and capitalization where economies are open to each other, there has been an influx of the masses to move to cities and this caused the number of cities to increase substantially from 5000 to approximately 10000 in last 50 years (Zlotnik 2017). During these years the major concern of the policy makers for governments have been to provide an abode to this population (Ozarisoy & Altan 2022). The houses so constructed require many amenities which are dependent on Energy supply from various sources. In haste, many a time the energy resources have been overburdened to construct the required buildings and in many of the cases the buildings were not given a due care for their energy efficiency resulting in more carbon foot prints across the globe. These urban buildings currently accounts for around 40% of the carbon emissions worldwide (Ma et al. 2012, Virote & Neves-Silva 2012, Abergel et al. 2018). It becomes much more important considering the fact that the life span of these buildings in much long. Changes in climate due to greenhouse gasses emissions and heat energy from buildings have increasingly affected the distribution and composition of ecosystems. This causes a very harmful impact on air quality, global biodiversity, agriculture, natural resources etc. in addition, it is responsible for high concern respiratory, circulatory, and skin diseases. Therefore, every year united nation's organization like UNFCCC emphasizing to reduce carbon emissions in particularly in Urban areas (Kelly 2020).
The Urbanization is required because cities generate more than 80% of global GDP, new ideas and innovations, industrialization, improved education and health facilities, availability of internet etc. Urban areas are responsible for two- thirds of primary energy consumption and 70% of carbon dioxide emissions only from the energy sector worldwide whereas the land surface covered by the urban areas is only 3%. Moreover, the ratio of the urban population to the global population will increase from 55% to 68% between now and 2050. As a result, more buildings will be needed which will be devastating for environmental health, however to attain sustainability is the only key to survive (Zlotnik 2017).

In these circumstances urbanization and building creation shall only expand rather than being remained curtailed. It is thus imperative to design these buildings in such a manner so that they consume and waste minimal energy to save carbon footprints. This proportion of energy consumption depends on many parameters which need to be taken care of during the architectural design of a building. The parameters include weather conditions around the building, the age of the building, geometry design of the building, services offered, activities being carried out in the building, and last but not the least the glazing.

Accepted manuscript doi: 10.1680/jener.23.00017

It is clear from above literatures that zero carbon foot in buildings require an attentive design of every parameter involved in energy performance of a building. For designers of such systems prior energy data of buildings is very vital. This data can be used by variety of software available to predict the energy need thereby suitable models for performance predictions. But the main problem lies in the dynamics of the data. Due to which the analysis of different data is very cumbersome and time-consuming exercise. Therefore, a quick tool is vital which will help in concept design of urban buildings efficiently under these dynamic conditions. AI based artificial neural networks are answers to them.

Conducting experiments and doing detailed energy analysis are expensive cumbersome exercises. Therefore some alternative approaches are the need of the hour. In a recent past a remarkable study was performed by Forrester (1994) to highlight the importance of System dynamics approach which paved way for future Machine learning based works.

Artificial Neural Networks is one such modelling approach of machine learning based modelling. Many studies are there in literature which highlights the importance of ANN based modelling (Gupta et al. 2017, Azadeh et al. 2011, Lippmann 1987, and Kang 2017) has also proven its efficacy in plethora of problems of core engineering related to mechanical, civil, electrical, computer science etc. This depicts that ANN is one of the most flexible tools having tremendous learning ability under dynamic systems and may prove to be vital for energy performance prediction of buildings. In comparison to other tools of artificial intelligence, ANN has already been proved to be a better option in solving the dynamical non-linear behaviors of engineering systems involving bigger dimension data sets (Haykin 2010). Zhang et al. 2021 reviewed the applications of machine learning for modelling and analysing the building energy system. They summarized 128 different machine learning algorithms namely algorithms of Regression, Support Vector Machine (SVM), Neural Network (NN), Deep Learning, Tree-based Algorithms, Hybrid Algorithms, Autoregressive methods, Extreme Learning, Bayesian Networks, Case-based Reasoning (CBR), Meta Learning, k-Nearest Neighbors (kNN), Gaussian Process and Mixture Models, Ensemble Methods, Fuzzy Timeseries Algorithm etc can be applied to predict building energy load.

Banihashemi et al. 2017 used the ANN approach analyzing building energy performance by postulating a unique approach in which they introduced the energy hybrid function both in ANN and the decision tree. This enabled the network to deal with the continuous data. Deviating from traditional physics methods, Paterson et al. 2017 gave a machine learning based approach for building performance prediction. For performance analysis they used thermal comfort and energy consumption as the input layer parameters and could give the results having and errors within 22-23% in terms of mean absolute percentage error (MAPE). Considering the same factors of thermal comfort and energy consumption Zhai and Soh 2017 tried to optimize heating, ventilation and air conditioning systems. With application of two algorithms sparse firefly algorithm and sparse augmented firefly algorithm (AFA) they could gave a program to save energy by upto 30% at no compromise to thermal comfort. Ahn and Cho 2017 gave another ANN base model to control the room parameters for energy efficiency through the air supply and its temperature together. Using this model, they could reduce the energy consumption by 2% by avoiding the control errors upto 88%.

For Zero Energy Building (nZEB) in cold climate conditions, Xiaolong et al. 2018 did an energy performance analysis using ANN to minimize heating and cooling needs of these buildings in particularly for China. Using their approach, they could predict the energy needs upto an accuracy of approximately 95%. Turhan et al. 2014 studied many parameters using ANN approach to analysis the effect of these parameters, like geometry, layout climate etc., over building energy performance in terms of heat load. They did the whole exercise to compare the results of ANN with a software called KEP-IYTE-ESS (Turhan et al. 2014 and KEP-SDM 2008) which was a building energy performance simulation tool. They found that the ANN predictions were as good as the physical modelling using software tool with prediction rate of 0.977.

Razmi et al. 2022 recently conducted a study to optimize and improve the energy performance during the architectural design of a building. Their PCA-ANN integrated NSGA-III model was specially focused for buildings like dormitories have multi-type buildings. The model had the R^2 score of 0.99 with a prediction of thermal performance improvement in the range of 15-42% with respect to the existing base models. Elbeltagi and Wefki 2021 used ANN to predict the building energy performances in the early stages of its design with a reasonable accuracy. Chujie et al. 2021 did a detailed literature survey work and reviewed various works in recent past (Deb et al. 2017, Wang and Srinivasan 2017, Wei et al. 2018, Amasyali and El-Gohary 2018, Bourdeau 2019, Mohandes et al. 2019, Wang et al. 2019, Roman et al. 2020, Fathi et al. 2020, and Sun et al. 2020) on energy performance analysis using ANN and concluded that ANN can be used for predicting building performances effectively. Their work basically focused on various ANN architecture for e.g., MLP, RBF, WNN, NARX, ENN, RNN, LSTM and CNN etc.

Bui et al. 2020 further gave a new ANN model using Electromagnetism dependent Firefly model to evaluate the heating performance and cooling performance of buildings. D'Amico and Ciulla 2022 worked on the similar aspects to check the economics of ANN based model in predicting the building energy efficacy. Xiang et al. 2022 used a three-step model approach using GeneticAlgorithm-Neural Network (GA-ANN) for building energy performance analysis to make the designers of such buildings predicting the real time, easy to use and trust worthy assessment in decision making. Olu-Ajayi et al. 2022a adopted a machine learning (MI) based approach and compared nine algorithms for better suitability to building energy performance. Along with that probably first time they used hyper parameter tuning for best results in building performance analysis through MI approach. Olu-Ajayi et al. 2022b further used various ML approaches like utilized various machine learning techniques such as Artificial Neural Network (ANN), Gradient Boosting (GB), Deep Neural Network (DNN), Random Forest (RF), Stacking, K Nearest Neighbour (KNN), Support Vector Machine (SVM), Decision tree (DT) and Linear Regression (LR) for the prediction of annual energy consumption in residential buildings. They concluded that the DNN (Deep Neural Network) is the most efficient model for energy prediction. The unique aspect of their work was to include the building design features as the part of the model. Zhang et al. 2019 developed and evaluated cooling load models using K-means clustering method and k-nearest neighbor (kNN) classification method. It is suggested that consideration of different load modes can effectively improve the accuracy of prediction of the day-ahead model. Further Yang, et al 2023), López-Ochoa et al 2023, Ma, et al 2023, Krarti 2023 and Wang, et al 2023 recently did some remarkable researches to highlight the importance of machine learning approaches in the field of energy optimization.

The above detailed literature survey reveals that, the Machine learning approach in particular ANN can address the design of building services in new buildings by optimizing the efficiency of different energy management, heating, ventilation, air conditioning and water systems as it can simultaneously handle multiple parameters efficiently to explicitly manifest their importance by avoiding costly experimentation for each parameter individually. It may provide a more detailed alternative to traditional building energy benchmarking and rating schemes. AI can either enable or inhibit the delivery of all 17 goals and 169 targets recognized in the 2030 Agenda for Sustainable Development (Vinuesa 2020). AI may act as an enabler on 134 targets (79%) across all SDGs, generally through a technological improvement, which may allow it to overcome certain present limitations. The major research contributions of the present work are summarized as follows:

- In the present work one of the widely accepted tools of AI, artificial neural networks (ANN) have been chosen to establish the energy efficiency of a building situated in Greece.
- Furthermore, the current work involves comparing the prediction performance of three neural variants models namely Levenberg-Marquardt (LM), Bayesian Regularization (BR) and Scaled Conjugate Gradient (SCG) for the target prediction task. The aim here is to find the best model for achieving reliable and accurate forecasting performance.
- The impact of several building parameters; relative compactness, surface area of the building, wall area of the building, roof area of the building, overall height, and orientation of the building, glazing area and glazing area distribution on the heating and cooling loads of the buildings is well investigated and discussed in the present study.
- The current research study reveals critical information about several factors to improve the energy efficiency of the buildings.

ARTIFICIAL NEURAL NETWORK (ANN) BASED MODELLING

Based on literature survey, it is manifested that there exist alternative modeling and analysis approaches for the design of energy efficient buildings. Artificial Neural Networks are one of those most efficient machine learning based approaches which have gained popularity based on its potential to effectively simulate a physical systems of real world and very low cost. Details of this methodology which has been adopted in the current work are explained in subsequent sections.

Structure of ANN

ANN is a processing system, design which is inspired by the structure and functioning of the human brain, that can be trained to analyse and model nonlinear, complex processes that are not well explored or understood. A schematic diagram showing signal processing by an artificial neuron is shown inFigure 1.The similarity between a human neuron and artificial neuron is shown in Figure 1 (a), where dendrite is taken as input, and synaptic terminal is taken as output.

An artificial neural receives signal as an input and multiply these inputs with the weights corresponding to each input signal. The cumulative signal is processed by an activation function to get an output as shown in Figure 1(b). An artificial neural network is combination of several such neurons. The general-purpose ANN structure consists of various hidden layers and an input and an output layer. Each hidden layer may comprise of multiple

neurons. The connection from input layer to hidden layers and hidden layers to output layer remains through weight activation function and optimization functions only to update the value of weights in each iteration called as learning of the ANN network (Zell et a. 1994). Flow diagram of neural network is shown in Figure 1(c) which explains the network as the matrix of neurons and its topology or connections between them which is governed by its learning rule. Each output of a neuron becomes the input to the next layer neuron. Complex relationships can be modelled with the help of multiple layers of neurons with different neurons in each layer training is done by sequentially applying inputs and adjusting their corresponding weights according to a specified algorithm until the desired target value is obtained. While training, the weights for each input will end up to a specific value, till its output approximately reaches the desired target.

Mathematical Modelling of a Neuron

A neuron, which is known to mimic the functionality of biological neuron, is obtained by formulating a simplified function into a mathematical formula, which will further be the basis of the mathematical model of neural network. In this section the mathematical model of Feed Forward Back Propagation (FF-BP) and Bayesian Regularization (BR) is explained.

Accepted manuscript doi: 10.1680/jener.23.00017

Feed Forward Back Propagation

Feed Forward Back Propagation (FF-BP) model can be categorized in two phases. During first phase, the input is presented and propagated forward through the network to compute the output value for each neuron. This neuron is then compared to the target output, resulting in an error signal for each output unit. The second phase involves a backward pass through the network during which the error signal is passed to each unit in the network and the appropriate weight change is made. Mathematical modelling of FF-BP algorithm is presented as below.

The input vector x_i is defined as

$$x_i = [x_1, x_2 \cdots\ x_n] \tag{1}$$

and the weight vector w_{ji} is defined as

$$w_{ji} = \left[w_{j1}, w_{j2} \cdots\ w_{jn}\right] \tag{2}$$

The induced local field v_j is obtained by

$$v_j = \sum_{i=1}^{i=n} w_{ji} x_i \tag{3}$$

where, n is the number of inputs and the bias b_j is given as

$$b_j = w_{j0} x_0 \tag{4}$$

Finally, the output y_j is obtained as.

$$y_j = \varphi(v_j + b_j) \tag{5}$$

where, φ is the activation function which determines the behavior of a node. In current work sigmoid function is taken as activation function following Angel et al. 2017 as shown in Eq. (6), and its value varies from 0 to 1.

$$\varphi(v_j(n)) = \frac{1}{1+exp(-av_j(n))} \tag{6}$$

where, $a > 0$ and $-\infty < v_j(n) < \infty$.

The error signal at the output of neuron j at n^{th} iteration is calculated as

$$e_j(n) = t_j(n) - y_j(n) \tag{7}$$

The total error energy or cost function for neuron j at n^{th} iteration is given by

$$\xi(n) = \frac{1}{2} \sum_{j \in m} e_j^2\ (n) \tag{8}$$

where, m is the number of neurons in output layer

Now if N is the total number of the training data set, then the average cost function will be,

$$\xi_{avg} = \frac{1}{N} \sum_1^N \xi(n) \tag{9}$$

The goal of the training of the network is to adjust the weights to get closer to the target value. Thus, the objective of iterative training is to minimize the averaged cost function. As average cost function is a function given as $\xi(n) = f\ (e, y, v, w)$, its partial derivative is evaluated by the chain ruleas given in Eq. (10).

$$\frac{\partial \xi\ (n)}{\partial w_{ji}} = \frac{\partial \xi\ (n)}{\partial e_j(n)} \frac{\partial e_j(n)}{\partial y_j(n)} \frac{\partial y_j(n)}{\partial v_j(n)} \frac{\partial v_j(n)}{\partial w_{ji}(n)} \tag{10}$$

Individual constituents of Eq. (10) can be derived by differentiating Eq. (8), Eq. (7), Eq. (5) and Eq. (3). These constituents are given in Eq. (11) - Eq. (14).

$$\frac{\partial \xi\ (n)}{\partial e_j(n)} = e_j(n) \tag{11}$$

$$\frac{\partial e_j(n)}{\partial y_j(n)} = -1 \tag{12}$$

$$\frac{\partial y_j(n)}{\partial v_j(n)} = \varphi_j' \left(v_j(n) \right) \tag{13}$$

$$\frac{\partial v_j(n)}{\partial w_{ji}(n)} = y_i(n) \tag{14}$$

After finding all the derivatives and substituting them back in Eq. (10) we may obtain,

$$\frac{\partial \xi(n)}{\partial w_{ji}(n)} = -e_j(n)\varphi_j' \left(v_j(n) \right) y_i(n) \tag{15}$$

The updated weights can be written as given in Eq. (16),

$$\Delta w_{ji}\ (n) = -\eta \frac{\partial \xi(n)}{\partial w_{ji}(n)} = \eta\ e_j(n)\varphi_j' \left(v_j(n) \right) y_i(n) \tag{16}$$

where, η is the learning rate parameter.

These weights are calculated and updated iteratively for all the neurons used in ANN.

Bayesian Regularization

The intricacy of the back propagation neural network (BPNN), monitored by neurons in the hidden layer and their associated weights may leads to overfitting, i.e., the network tries to make the error as small as possible for the training set but performs poorly when new data is presented. However, a robust network model should be able to generalize accurately, i.e., it should predict accurately even when presented with new data. Therefore, Bayesian regularization-based learning of BPNN models is utilized to achieve better generalization and minimal over-fitting for the trained networks following Gouravarajua et al. 2020. To generalize the neural network, the cost function is modified using a regularization method. A penalty term is added to the performance index and performance index has been defined as

$$F(\overline{w}) = \mu \xi(n) + \nu E_D \tag{17}$$

where, μ, ν are regularisation parameter, E_D is the squared network weights and $\xi(n)$ is the total error energy or cost function. If $\mu \ll \nu$, smaller errors are generated, while if$\mu \gg \nu$, there is a reduced weight size, hence BR is used to find the optimum regularization parameters between μ and ν. Target is to choose the weights which maximize the posterior probability distribution of the weights $P\left(\overline{w} | D, \mu, \nu, M_N\right)$ given a certain training data D, where weights $(\overline{w})$are considered as random variables and $M_N$as neural network architecture.

According to Bayes' rule, posterior probability distribution depends on the likelihood function, the prior density, and the normalization factor for a neural network model M_N and can be evaluated from

$$P(\overline{w}\mid D, \mu, \nu, M_N) = \frac{P(D \mid \overline{w}, \nu, M_N) P(\overline{w}\mid \mu, M_N)}{P(D \mid \mu, \nu, M_N)} \tag{18}$$

The likelihood function is given as

$$P(D \mid \overline{w}, \nu, M_N) = \frac{e^{(-\nu E_D)}}{Z_D(\nu)} \tag{19}$$

where, $Z_D(\nu) = (\frac{\pi}{\nu})^{Q/2}$ and $Q = n \times m$,

And Prior density is given as,

$$P(\overline{w}\mid \mu, M_N) = \frac{e^{(-\mu E_W)}}{Z_W(\mu)} \tag{20}$$

The Posterior probability distribution is given as,

$$P(\overline{w}\mid D, \mu, \nu, M_N) = \frac{e^{(-\mu E_W - \upsilon E_D)}}{Z_D(\mu, \nu)} = \frac{e^{(-F(\overline{w}))}}{Z_D(\mu, \nu)} \tag{21}$$

where, $Z_D(\mu, \nu)$ is the normalization factor.

The complexity of the model M_N is governed by regularization parameters(μ, ν), which need to be estimated from the data. Therefore, Bayes' rule is again applied to optimize them as follows

$$P(\mu, \nu \mid D, M_N) = \frac{P(D \mid \mu, \nu, M_N) P(\mu, \nu \mid M_N)}{P(D \mid M_N)} \tag{22}$$

$$\mu^* = \frac{\gamma}{2E_W(\overline{w}^*)} \tag{23}$$

where $P(\mu, \nu \mid D, M_N)$ denotes the assumed uniform prior density for the parameters μ and ν. It was seen that by maximising the likelihood function, the posterior probability maximises automatically. However, it can be noted that the likelihood function is the normalisation factor. For solving the likelihood function, the objective function around the minimal point $\overline{w}^*$is expanded with a Taylor series expansion, and the optimum value of the regularization parameters can be evaluated using,

$$\nu^* = \frac{Q-\gamma}{2E_D(\overline{w}^*)} \tag{23}$$

where, γ represent the number of effective parameters exhausted in minimising the error function given as

$$\gamma = K - \mu^* tr(H^*)^{-1} \tag{24}$$

For $0 \leq \gamma \leq K$, the Hessian matrix of the objective function is evaluated at $\overline{w}^*$ and is calculated using the Gauss-Newton approximation as

$$H^* \approx J^T J \tag{25}$$

where,J is the Jacobian matrix formed by the first derivatives of the network error e with respect to network weights w_{ij}.

The normalization factor $Z_D(\mu, \nu)$ can then be approximated as

$$Z_D(\mu, \nu) \approx (2\pi)^{\frac{K}{2}} (det(H^*))^{-\frac{1}{2}} exp(-F(\overline{w}^*)) \tag{26}$$

Bayesian Regularisation (BR) ANNs are extra vigorous than usual back-propagation networks and can lessen the necessity for long cross-validation. BR algorithm is a process that changes a nonlinear regression into a “well-modelled” statistical problem in the means of a ridge regression. Hence BR algorithm generally takes more time but is accurate with smaller data sets.

Complex relationships can be modelled with the help of the multiple layers of the neurons with varying number of neurons in each layer. The weights for each input are the most important parameter to analyse the impact that input variable produces over the target output. The input variables having a stronger influence on the output parameter will have larger weights on specific input variables.

Training Network

Selection of the training network is the most important step while performing ANN analysis. The network consists of several layers termed as the input layer, the hidden layer, and the output layer. The input node is the left most node, while the output node is the right most one, the layers between the input and output layers are called as the hidden layers as depicted in Figure 1 (c). The network is basically trained by manipulating the values of hidden layers. The signals travel from the input layer to the output layer through hidden layers. Each neuron of the neural network is connected to the previous and next layer by an imaginary line with different weights and after each iteration (epoch), network updates these weights according to the desired output.

The output of every node is the weighted sum of the input multiplied with its weight for every connection. To produce the output vector, the weighted sum of all of the neurons of the input passes through an activation function. An activation function may be sigmoid, RELU, Tanh, Linear etc. Sigmoid has been used as the activation function in the current work for all the hidden layers following Angel et al. 2017. A sigmoid function is a real function defined for all the real values and has a non-negative derivative. Mathematically, sigmoid function is defined as,

$$y = \frac{1}{1+(e)^{-x}} \tag{28}$$

where, x is the input parameter and y is the output parameter.

Network Development

MATLAB is used for the development of the required ANN network for the study. MATLAB is a programming language and numeric computing environment developed by MathWorks which provides the platform to create an artificial neural network (ANN), in which the data set can be trained using nnstart tool as shown in Figure 2. The saved network with the best performance (having least MSE) is used to predict the heating and cooling loads with the testing data or new data. The iteration continued till optimum number of neurons in hidden layer is reached, which is governed by value of the mean square error (MSE).One of the representative networks used in this study is shown in Figure 3. The details of this network shall be explained in section 3.3.

Accepted manuscript doi: 10.1680/jener.23.00017

DEVELOPMENT OF NEURAL NETWORK FOR MODELING HEATING AND COOLING LOAD OF RESIDENTIAL BUILDING

To develop the required neural network for the optimization of the heating and cooling loads of the residential buildings, the data set of Tsanas and Xifara (2012) is used in the present study. Figure 4 shows the schematic diagram depicting parameters investigated in the present study and range of investigated parameters are given in Table 1.

Benchmark Data Set

The data used in the present study is taken from Tsanas and Xifara 2012, which is openly available from the following link https://archive.ics.uci.edu/dataset/242/energy+efficiency. Statistical analysis of this data has already been done by the authors of this article Tsanas and Xifara 2012. Additionally, Chou and Bui 2014 used the same data set and carried out analysis using support vector regression (SVR), artificial neural network (ANN), classification and regression tree (CART), chi-squared automatic interaction detector (CHAID) and general linear regression (GLR).

Likewise, Cheng and Cao 2014, proposed evolutionary multivariate adaptive regression splines (EMARS) which is a hybrid model of multivariate adaptive regression splines (MARS) and artificial bee colony (ABC). They utilised the same data set which is used in the present investigation. In this study, EMARS performance was compared against five other AI models, including MARS, back-propagation neural network (BPNN), radial basis function neural network (RBFNN), classification and regression tree (CART), and support vector machine (SVM). They further compared the performance of EMARS with the random forest (RF). For hypothesis test, the comparison was done based on mean absolute error (MAE) values using one-tailed t-test with unequal size and unknown variances. Other studies such as Boukabour and Masmoudi 2021, Abdelkader et al. 2020, Irfan and Ramlie 2021, Yang et al. 2017, Gkioulekas et al. 2021 has shown statistically representativeness of the data considered in the present study. Moreover, In the present study 10-fold cross-validation is conducted to generalize statistical analysis for an independent data set.

In the benchmark study of Tsanas and Xifara (2012), the buildings considered were situated in Greece, with seven persons residing in it and idle in activity (which correlates to 70 W of energy).For the buildings, the humidity was considered as 60%, the air speed as 0.30 m/s, the illumination level as 300 Lux, the sensible heat exchange as 5 W/m^2 and the latent heat exchange as 2 W/m^2. The infiltration rate of air into the building was considered as 0.5 with sensitivity of wind as0.25 air change/hour. The air change/hour was set as the internal design condition and taken as an assumption for the study. Further, it was assumed that the operation hours are 15 to 20 hours on the weekdays and 10 to 20 hours are over the weekends for the heat transfer calculations.These buildings were composed of 18 elementary cubes and the total volume 771.75 m^3 of the buildings were kept same whereas parameters such as surface area, roof area, glazing area etc. listed in Table 1were varied.

These geometric and operating parameters were modelled and simulated using Autodesk Ecotect analysis simulation software.

A total of 768 building samples were obtained from different combinations of

(a) Twelve building forms for the four of the orientations with three glazing areas, having five different glazing area distribution scenarios [Table 2]

(b) Twelve building forms for the four of the orientations without glazing. All the 768 simulated buildings are characterized by the eight building parameters (X1-X8 in Table 1) taken as input parameters and two output parameters (Y1and Y2 in Table 1) as heating load and cooling load for the development of the optimized neural network. Table 1 summarizes the details of the input and the output variable with their mathematical representation and the number of possible values of each of the variable. The functional dependence of input and output parameters used in this study is shown by parallel plot in Figure 5. It can be observed from this figure that both heating load (HL) and cooling load (CL) exhibit complex relation for all the input parameters (X1 to X8).

Performance Analysis of Developed Network

MATLAB developed by Mathworks is used to create an artificial neural network (ANN) with nntool (MATLAB 2018).The process of network optimization shall be explained under section 3.3.The performance of this network has been analyzed with the help of regression value and mean square error (MSE) value.

Regression plot is the performance parameters to analyse the network performance as shown in Figure 6. Regression plot is shown for a representative It creates a regression line between 2 parameters and helps to visualize their linear relationships. It is the plot between output and target value, which shows how closer the output is to the target value. The value of regression coefficient is 0.99865 for the training data as shown inFigure 6, for the testing data, it is 0.99826 and overall, the regression coefficient is 0.9986which indicates that network can take output closer to the target value.

Mean Square Error is the average of squares of difference between target and output parameter values for all the training data as given in Eq. (29). It shows how close the network can reach to the target value. Lesser the value of MSE, the better is training of the network. The variation of MSE of current work with epochs is shown in Figure 7 for one representative among all the investigated networks. The least value of MSE in the training data i.e.,1.59e-04for the shown case is obtained at 258 epoch while 2.9e-04for the testing data.

$$MSE = \sum_{i=1}^{n} \frac{((Heating\ Load)_{Actual,i} - (Heating\ Load)_{Predicted,i})^2}{n} \quad (29)$$

Accepted manuscript doi:
10.1680/jener.23.00017

Network Sensitivities Analysis

The sensitivity of any network depends on factors like training of the selected algorithm, number of hidden layers, number of neurons in different hidden layers and segregation of data into training, validation, and testing.

Sensitivity on the Training Algorithm

In the current work three different algorithms are used to train the network. These networks are: Levenberg-Marquardt (LM), Bayesian Regularization (BR) and Scaled Conjugate Gradient (SCG). The data used for training is divided as 70% for training, 15% for validation and 15% for testing with a single hidden layer. Number of neurons are varied starting from 8 and then increased in the multiples of 8. The MSE of trained network is evaluated on the normalized data. This value may not be a true performance parameter and hence it is decided that the network predictions on the denormalised data for the untrained data will be used to evaluate the MSE performance of the investigated networks. Results obtained from this study are presented in Table 3 and Figure 8. Following are the observations drawn from the obtained results:

- It is observed that the MSE decreases up to2.46e-01 with the increase in number of the neurons for LM training algorithm. It starts further decreasing for BR training algorithm when different number of neurons are taken from 8 to 80. But it starts increasing up to7.16e+00 with the training algorithm SCG, hence increase in number of neurons is terminated as soon as the MSE starts increasing with the number of neurons.
- Further, it is observed that the MSE is least (1.15e-01) for the BR algorithm as compared to the LM and SCG algorithms. The MSE for the SCG algorithm is higher(5.30e+00)as compared to the LM (2.11e-01) algorithm, thus, for the data set used in this study, the BR is the better algorithm, which gives best results with 16 neurons.

Sensitivity analysis of the ANN network-Hidden Layers

Performance of an artificial neural network is dependent on the number of hidden layers and the number of neurons. In order to select optimal number of the hidden layers and the number of neurons in each of the hidden layer, a systematic study is conducted by varying the hidden layers and the number of neurons in each hidden layer. The investigated hidden layers and combination of neurons in each layer is shown in Table 4. Artificial neural networks are developed for each case shown in Table 4. As mentioned in section 3.3.1the data used for training of these networks is divided as 70% for the training, 15% for the validation and 15% for the testing and BR can be selected as a training algorithm to train the network. It is seen that with additional neurons in the network the MSE value reduces and the minimum value of MSE is observed when the number of neurons in each of the two hidden layers are 24 and 12 respectively as mentioned in test case 4 of Table 4. However, it should be

Accepted manuscript doi:
10.1680/jener.23.00017

noted that this minimum mean square error is more than the one observed for single hidden layer network. Hence, it is decided to use network BR-16 (single hidden layer) and BR-24-12 (two hidden layer) for further analysis.

K-Fold Cross Validation

K-Fold Cross Validation is popularly known for evaluation of performance of algorithms. For k-fold cross validation, firstly, data is divided into k disjoints folds (training and testing random sets) having approximately similar number of instances. Afterwards, every fold plays the role to test the model induced from the other k-1 folds when it is repeatedly performed as depicted in Figure 9.The purpose of k fold cross validation is to obtain reliable accuracy estimates. In the present study, 10 folds are used to evaluate the performance of selected networks i.e., BR-16 and BR-24-12. The average MSE of 10 folds for the trained network (BR-16) is given in Table 5. In each fold, MSE of the trained network on the normalized data used for validation is evaluated. Likewise in each fold, trained network is used to predict heating load and cooling load on the preserved 33 data points. The MSE is evaluated based on the prediction of these 33 conditions and presented in Table 5. The average MSE of the network on the trained normalized data is 2.28E-04 and MSE on the untrained data set of HL and CL is 0.158 and 1.26 respectively. K-fold analysis is also done on the network BR-24-12 i.e., two hidden layer network and results are presented in Table 6. The average MSE of the network on the trained normalized data is 7.14E-04 and MSE on the untrained data set of HL and CL is 1.5 and 4.41 respectively. These values indicates that the networks can be utilized for further analysis.

Network and prediction

After the sensitivity analysis it is observed that the best possible network consists of single hidden layer with 16 neurons in them, while the Bayesian Regularization algorithm is the best possible algorithm to be used for the data set considered in the present study. However, performance prediction of network BR-24-12 is also comparable. Hence, predictions of these two networks is discussed throughout the paper. Further, a comparison between the predicted values of heating and cooling loads using both the selected networks and the values obtained from Tsanas and Xifara (2012) is shown in the Figure 10 and Figure 11 for the entire data set of 768 samples. It is seen that the prediction values from the network traces the non-linear trend of the benchmark data set accurately despite of the fact that only 70% data was used for training of the network. The absolute error and its occurrence can be understood from the histogram presented in Figure 12 for both the heating and cooling loads. It can be observed from Figure 12 that absolute error is skewed towards the zero error for both the selected networks. The important observations drawn are discussed as follows:

Accepted manuscript doi: 10.1680/jener.23.00017

- For the heating load predictions, 331& 336 data sets are within the range of 0 < 0.5 absolute error and 291& 320are within the range of -0.5 < 0 for the network BR-16 and BR-24-12 respectively. Thus, for 84.62%& 89.25% of the heating load data set, absolute error is within the range of -0.5 to 0.5 for the network BR-16 and BR-24-12 respectively. A maximum absolute error of 1.89&1.71 is noticed for a single data point which is corresponding to 12.56 %&11.34 % error for the heating load data set for the network BR-16 and BR-24-12 respectively.
- Similar to the heating load data, cooling load predictions are in good agreement with the benchmark data set of Tsanas and Xifara (2012). For the cooling load predictions, 263&278 data sets are within the range of 0 < 0.5 absolute error and 222 &281 are within the range of -0.5 < 0. Thus, for 65.99% &76.05% of the cooling load data set, absolute error is within the range of -0.5 to 0.5 for the network BR-16 and BR-24-12 respectively. A maximum absolute error of 3.54& 2.03is noticed for a single data point which is corresponding to 10.34% &6.08% error for the cooling load data set for the network BR-16 and BR-24-12 respectively. The mean square error in predicting heating and cooling load is found to be 0.115 and 0.749 for the network BR-16 and0.21 and 0.876 for the network BR-24-12 respectively.

VERIFICATION OF NETWORK PREDICTION CAPABILITY

For the verification of network prediction capability out of the eight input parameters as mentioned in Table 1 four most important parameters are selected namely Glazing area distribution, Wall area, Surface Area, and Relative Compactness. Untrained data set was reserved for these parameters for verification of network prediction capability before training the networks investigated in this study. Then with the help of the trained networks (BR-16 & BR-24-12), predictions for heating (HL) and cooling (CL) loads of the building are done for different values of the four parameters mentioned above. The results so obtained for HL and CL are drawn in Figure 13 and Figure 14 respectively and compared with the actual values obtained for the same parameters by Tsanas and Xifara (2012). The details discussion regarding varying four parameters vs HL and CL is given as follows:

- The variation of heating load with glazing area distribution is shown in Figure 13 (a) along with the benchmark test case. It can be observed that the variation in the heating load with glazing area distribution is well captured by the trained ANN. A maximum deviation of 2.35% and 2.11% is observed for the normalized glazing area distribution of 1 for the network BR-16 and BR-24-12 respectively.
- The variation of heating load with the wall area follows a non-linear trend in the benchmark study as shown in Figure 13 (b). Heating load decreases when the normalized wall area increases from 294 to 318.5 after that a sudden peak in the heating load was observed for wall area of 343; further increment in the wall area leads to decrease in the heating load. This non-

linear behavior is well captured by the trained neural network. A maximum deviation of 1.4% is observed for the normalized wall area of 294 for the network BR-16 and 0.92% is observed for the normalized wall area of416.5 for the network BR-24-12.

- The variation of heating load with the normalized surface area is shown in Figure 13 (c). It can be observed that heating load increases with the increase in the normalized surface area and attains a maximum value for the normalized surface area of 784. Heating load decreases in the benchmark test case with further increment in the normalized surface area. This trend is also captured by the trained neural networks with a reasonable accuracy. The maximum deviation of 1.25 %&0.71% is observed in the predicted results for the normalized surface area of 735 for the network BR-16 and BR-24-12 respectively.
- Heating load variation with the relative compactness follows the similar trend as was followed by the normalized surface area. Heating load increases when relative compactness increases from 0.62 to 0.64 and thereafter it decreases as shown in Figure 13 (d). It can be observed from this figure that the trained network can predict this trend. The maximum deviation of 7.13 % & 2.23% is observed in the predicted results fora relative compactness of 0.74 for the network BR-16 and BR-24-12 respectively.

In a similar context, a comparison of predicted cooling load with the benchmark results for the above mentioned four parameters are shown in Figure 14 and the corresponding critical observations are discussed as follows:

- The variation of cooling load with the glazing area distribution is almost uniform in the benchmark test case as shown in Figure 14 (a). The trained ANN network under predicted cooling load by 2.44% for the normalized glazing area distribution of 4 for the network BR-16 and over predicted by 5.44% for the normalized glazing area distribution of 1 for the network BR-24-12.
- A non-linear variation of the cooling load with the normalized surface area, wall area and relative compactness can be observed from Figure 14 (b-d) respectively. These trends are well predicted by the trained neural networks. The maximum deviation in the predictions of the normalized surface area, wall area and relative compactness is 6.58%, 4.5% and 2.87% respectively for the network BR-16. The maximum deviation in the predictions of the normalized surface area, wall area and relative compactness is 3.36%, 4.29% and 1.18% respectively for the network BR-24-12. Thus, it can be concluded that the predictions of the trained neural network are well within 10% deviation for the untrained data set and trained neural network can be employed to predict energy performance of the building within the reasonable accuracy.

Accepted manuscript doi:
10.1680/jener.23.00017

OPTIMIZATION OF ENERGY PERFORMANCE USING DEVELOPED NETWORK

In this section the optimization of building glazing area distribution for four different fractions of the glazing areas (0.15, 0.2, 0.3 and 0.35) is discussed. Twelve building glazing area distribution investigated are summarized in Table 7. The trained network with a structure of 8-16-2 is used to predict in total 96 different conditions of glazing area distribution between North, East, South, and West for four fractions of glazing areas (0.15, 0.2, 0.3 and 0.35).The predicted heating and cooling load are depicted in Figure 15 and the important conclusion drawn are discussed as follows:

- It can be observed that with the decrease in the glazing area fraction, the requirement of heating and cooling load decreases. This is because of the higher rate of the heat transfer from the glazing area as compared to the walls. When the glazing area is decreased, the heat loss/gain to/from the surrounding is also decreased. Hence, heating, and cooling load decrease with decreases in the glazing area.
- It can also be observed from Figure 15 (a) that there is a potential to decrease the heating load if glazing area distribution is carefully designed. For glazing area (GA) = 0.35, a minimum heating load is required for Test Number 8 which is listed in Table 7. For this test case, 15% of the glazing area is distributed in the North and West and 35% in the East and in the South. The saving in normalized heating load for Test-8 as compared to Test-1 is around 4 % for all the glazing area fractions investigated in this study.
- On the contrary to the heating load, the potential to save in cooling load is significantly higher as it can be seen in Figure 15 (b). The minimum and the maximum predicted cooling loads are corresponding to Test cases of 1 and 9 respectively for all the glazing area fractions investigated in this study. There is a potential of 18.97% saving in the cooling load for a glazing area of 0.35 if glazing area distribution is kept 32.5% in North, 22.5% in the East I, in the South and West respectively (Test Case -1) as compared to the glazing area distribution of 15% in the North and West, 25% in the East and 45% in the South (Test Case-9). It is also interesting to note that there is a potential to save 20.67% of the cooling load for a glazing area of 0.15. The minimum and the maximum predicted cooling load for a glazing area of 0.15 is corresponding to Test Case-1 and Test Case-9 respectively.
- Based on the combined minimum heating and cooling load, Test Case-1 is recommended i.e., glazing area distribution of 32.5% in the North, 22.5% each in the East, in the South and West. A saving of 6.9% and 7.87 % can be achieved in the combined heating and cooling load as compared to Test Case-9 for the glazing area of 0.35 and 0.15 respectively. For the other two investigated glazing areas, the net saving is within the range of saving achieved from the glazing area of 0.35 and 0.15.

Thus, a significant improvement in the energy efficiency of the residential buildings can be achieved if due care is given while planning or constructing residential buildings. Artificial neural network-based predictions can give a reliable and quick input to the town planners.

CONCLUSIONS

The current study is conducted to obtain an AI based optimized ANN network to find the minimum heating and cooling loads of a building and there by predict the building energy performance with respect to a number of parameters viz. surface area, wall area, glazing area, glazing area distribution etc. An ANN structure 8-16-2 with Bayesian Regularization algorithm is found to be the optimum ANN network structure for the prediction of the building energy performance with the MSE as 1.59e-04 for the network validation and 0.115 and 0.749 for prediction on the untrained data. The optimality of the network is also confirmed by the comparison of the HL and the CL values predicted by the network and the actual values available in literature, where the error is found to be less than 10% only. A well trained and validated artificial network is used to predict 96 conditions by varying glazing area and glazing area distribution.

This study revealed that there is a significant potential to improve the energy efficiency of the building. The maximum saving in the cooling load can be as high as 20.67% for a glazing area of 0.15 if glazing area distribution is kept 32.5% in the North and 22.5% each in the East, South and West. The net saving of the combined cooling and heating load can be as high as 6.9% and 7.87 % for glazing area of 0.35 and 0.15 respectively.

Nomenclature

b_j	Bias
H^*	Hessian matrix
J	Jacobian Matrix
v_j	Induced local field
w_{ji}	Weight vectors
x_i	Input vector
y_j	Output
Z_D	normalization factor
Greek Symbols	
φ	Activation function
ξ	Cost function
η	learning rate

Accepted manuscript doi: 10.1680/jener.23.00017

Abbreviations

ANN Artificial Neural Network
BR Bayesian Regularization
CL Cooling Load
CNN Convolutional Neural Network
DNN Deep Neural Network
ENN Ensemble Neural Network
GA Glazing Area
GDP Gross Domestic Product
HL Heating Load
KNN K-Nearest Neighbor
LM Levenberg-Marquardt
LR Logistic Regression
LSTM Long short-term memory
MAPE Mean Absolute Percentage Error
ML Machine Learning
MLP Multi-Layer Perceptron
MSE Mean Square Error
NARX Nonlinear Auto-Regressive Exogenous Model
RBF Radial Basis Function
RF Random Forest
RNN Recurrent Neural Networks
SCG Scaled Conjugate Gradient
SVM Support Vector Machine
UNFCCC United Nations Framework Convention on Climate Change
WNN Wavelet Neural Networks

References

[1] Zlotnik, H., 2017, “World Urbanization: Trends and Prospects." In New Forms of Urbanization,” *Routledge*.

[2] Ma, Z., Cooper, P., Daly, D., and Ledo, L., 2012, “Existing Building Retrofits: Methodology and State-of-the-Art,” Energy Build., 55, pp. 889–902.

[3] Virote, J., and Neves-Silva, R., 2012, “Stochastic Models for Building Energy Prediction Based on Occupant Behavior Assessment,” Energy Build., 53, pp. 183–193.

[4] Abergel, T., Dean, B., Dulac, J., & Hamilton, I., 2018, “Global Alliance for Buildings and Construction, 2018 Global Status Report,” United Nations Environ. Int. Energy Agency, p. 325.

[5] Kelly, L., “UNFCCC in Bonn: Bringing Transparency, Understanding, and Clarity to Countries’ Post-2020 Climate Contributions.”

[6] Gupta, A. K., Kumar, P., Sahoo, R. K., Sahu, A. K., and Sarangi, S. K., 2017, “Performance Measurement of Plate Fin Heat Exchanger by Exploration: ANN, ANFIS, GA, and SA,” J. Comput. Des. Eng., 4(1), pp. 60–68.

[7] Azadeh, A., Saberi, M., Anvari, M., and Mohamadi, M., 2011, “An Integrated Artificial Neural Network-Genetic Algorithm Clustering Ensemble for Performance Assessment of Decision Making Units,” J. Intell. Manuf., 22(2), pp. 229–245.

[8] Lippmann, R. P., 1987, “An Introduction to Computing with Neural Nets,” IEEE ASSP Mag., 4(2), pp. 4–22.

[9] Kang, H. J., 2017, “Development of an Nearly Zero Emission Building (NZEB) Life Cycle Cost Assessment Tool for Fast Decision Making in the Early Design Phase,” Energies, 10(1).

[10] Haykin, S., 2010, *Neural Networks and Micromechanics*.

[11] Banihashemi, S., Ding, G., and Wang, J., 2017, “Developing a Hybrid Model of Prediction and Classification Algorithms for Building Energy Consumption,” Energy Procedia, 110(December 2016), pp. 371–376.

[12] Paterson, G., Mumovic, D., Das, P., and Kimpian, J., 2017, “Energy Use Predictions with Machine Learning during Architectural Concept Design,” Sci. Technol. Built Environ., 23(6), pp. 1036–1048.

[13] Zhai, D., and Soh, Y. C., 2017, “Balancing Indoor Thermal Comfort and Energy Consumption of ACMV Systems via Sparse Swarm Algorithms in Optimizations,” Energy Build., 149, pp. 1–15.

[14] Ahn, J., and Cho, S., 2017, “Dead-Band vs. Machine-Learning Control Systems: Analysis of Control Benefits and Energy Efficiency,” J. Build. Eng., 12(April), pp. 17–25.

[15] Xu, X., Feng, G., Chi, D., Liu, M., and Dou, B., 2018, “Optimization of Performance Parameter Design and Energy Use Prediction for Nearly Zero Energy Buildings,” Energies, 11(12).

[16] Turhan, C., Kazanasmaz, T., Uygun, I. E., Ekmen, K. E., and Akkurt, G. G., 2014, “Comparative Study of a Building Energy Performance Software (KEP-IYTE-ESS) and ANN-Based Building Heat Load Estimation,” Energy Build., 85, pp. 115–125.

[17] Razmi, A., Rahbar, M., and Bemanian, M., 2022, “PCA-ANN Integrated NSGA-III Framework for Dormitory Building Design Optimization: Energy Efficiency, Daylight, and Thermal Comfort,” Appl. Energy, 305(October 2021), p. 117828.

[18] Elbeltagi, E., and Wefki, H., 2021, “Predicting Energy Consumption for Residential Buildings Using ANN through Parametric Modeling,” Energy Reports, 7, pp. 2534–2545.

[19] Lu, C., Li, S., and Lu, Z., 2021, “Building Energy Prediction Using Artificial Neural Networks: A Literature Survey,” Energy Build., p. 111718.

[20] Deb, C., Zhang, F., Yang, J., Lee, S. E., and Shah, K. W., 2017, “A Review on Time Series Forecasting Techniques for Building Energy Consumption,” Renew. Sustain. Energy Rev., 74(January), pp. 902–924.

[21] Wang, Z., and Srinivasan, R. S., 2017, “A Review of Artificial Intelligence Based Building Energy Use Prediction: Contrasting the Capabilities of Single and Ensemble Prediction Models,” Renew. Sustain. Energy Rev., 75(September 2015), pp. 796–808.

[22] Wei, Y., Zhang, X., Shi, Y., Xia, L., Pan, S., Wu, J., Han, M., and Zhao, X., 2018, “A Review of Data-Driven Approaches for Prediction and Classification of Building Energy Consumption,” Renew. Sustain. Energy Rev., 82(August 2017), pp. 1027–1047.

[23] Amasyali, K., and El-Gohary, N. M., 2018, “A Review of Data-Driven Building Energy Consumption Prediction Studies,” Renew. Sustain. Energy Rev., 81(January 2016), pp. 1192–1205.

[24] Bourdeau, M., Zhai, X. qiang, Nefzaoui, E., Guo, X., and Chatellier, P., 2019, “Modeling and Forecasting Building Energy Consumption: A Review of Data-Driven Techniques,” Sustain. Cities Soc., 48(February), p. 101533.

[25] Mohandes, S. R., Zhang, X., and Mahdiyar, A., 2019, “A Comprehensive Review on the Application of Artificial Neural Networks in Building Energy Analysis,” Neurocomputing, 340, pp. 55–75.

[26] Wang, Y., Chen, Q., Hong, T., and Kang, C., 2019, “Review of Smart Meter Data Analytics: Applications, Methodologies, and Challenges,” IEEE Trans. Smart Grid, 10(3), pp. 3125–3148.

[27] Roman, N. D., Bre, F., Fachinotti, V. D., and Lamberts, R., 2020, “Application and Characterization of Metamodels Based on Artificial Neural Networks for Building Performance Simulation: A Systematic Review,” Energy Build., 217.

[28] Fathi, S., Srinivasan, R., Fenner, A., and Fathi, S., 2020, “Machine Learning Applications in Urban Building Energy Performance Forecasting: A Systematic Review,” Renew. Sustain. Energy Rev., 133(September), p. 110287.

[29] Sun, Y., Haghighat, F., and Fung, B. C. M., 2020, “A Review of The-State-of-the-Art in Data-Driven Approaches for Building Energy Prediction,” Energy Build., 221, p. 110022.

[30] Bui, D. K., Nguyen, T. N., Ngo, T. D., and Nguyen-Xuan, H., 2020, “An Artificial Neural Network (ANN) Expert System Enhanced with the Electromagnetism-Based Firefly Algorithm (EFA) for Predicting the Energy Consumption in Buildings,” Energy, 190.

[31] D’Amico, A., and Ciulla, G., 2022, “An Intelligent Way to Predict the Building Thermal Needs: ANNs and Optimization,” Expert Syst. Appl., 191(December 2021), p. 116293.

[32] Li, X., Liu, S., Zhao, L., Meng, X., and Fang, Y., 2022, “An Integrated Building Energy Performance Evaluation Method: From Parametric Modeling to GA-NN Based Energy Consumption Prediction Modeling,” J. Build. Eng., 45(2), p. 103571.

[33] Olu-Ajayi, R., Alaka, H., Sulaimon, I., Sunmola, F., and Ajayi, S., 2022, “Machine Learning for Energy Performance Prediction at the Design Stage of Buildings,” Energy Sustain. Dev., 66, pp. 12–25.

[34] Olu-Ajayi, R., Alaka, H., Sulaimon, I., Sunmola, F., and Ajayi, S., 2022, “Building Energy Consumption Prediction for Residential Buildings Using Deep Learning and Other Machine Learning Techniques,” J. Build. Eng., 45(July 2021), p. 103406.

[35] Zhang, Q., Tian, Z., Ding, Y., Lu, Y., and Niu, J., 2019, “Development and Evaluation of Cooling Load Prediction Models for a Factory Workshop,” J. Clean. Prod., 230, pp. 622–633.

[36] Zhang, L., Wen, J., Li, Y., Chen, J., Ye, Y., Fu, Y., and Livingood, W., 2021, “A Review of Machine Learning in Building Load Prediction,” Appl. Energy, 285(December 2020), p. 116452.

[37] Vinuesa, R., Azizpour, H., Leite, I., Balaam, M., Dignum, V., Domisch, S., Felländer, A., Langhans, S. D., Tegmark, M., and Fuso Nerini, F., 2020, “The Role of Artificial Intelligence in Achieving the Sustainable Development Goals,” Nat. Commun., 11(1), pp. 1–10.

[38] Zell, A., Mache, N., Hübner, R., Mamier, G., Vogt, M., Schmalzl, M., and Herrmann, K.-U., 1994, “SNNS (Stuttgart Neural Network Simulator) In: Skrzypek J. (Eds),” *Neural Network Simulation Environments, The Kluwer International Series in Engineering and Computer Science Vol 254. Springer, Boston, MA*, pp. 165–186.

[39] Angel, M., Mojena, R., Sanchez, A., Sagaro, R., Sanchez, M., Carvajal, H., Roberto, C., and Lima, C., 2017, “Neural Network Analysis for Erosive Wear of Hard Coatings Deposited by Thermal Spray : Influence of Microstructure and Mechanical Properties,” Wear, 376–377, pp. 557–565.

[40] Gouravaraju, S., Narayan, J., Sauer, R. A., and Gautam, S. S., 2020, “A Bayesian Regularization-Backpropagation Neural Network Model for Peeling Computations,” 8295, pp. 0–1.

[41] MATLAB, 2018, *9.7.0.1190202 (R2019b)*, The MathWorks Inc., Natick, Massachusetts.

[42] Tsanas, A., and Xifara, A., 2012, “Accurate Quantitative Estimation of Energy Performance of Residential Buildings Using Statistical Machine Learning Tools,” Energy Build., 49, pp. 560–567.

[43] Yang, Xiaolin, Zhuoxi Chen, Yukai Zou, and Fengdeng Wan. "Improving the Energy Performance and Economic Benefits of Aged Residential Buildings by Retrofitting in Hot–Humid Regions of China." *Energies* 16, no. 13 (2023): 4981.

[44] López-Ochoa, Luis M., Jesús Las-Heras-Casas, Juan M. González-Caballín, and Manuel Carpio. "Towards nearly zero-energy residential buildings in Mediterranean countries: The implementation of the Energy Performance of Buildings Directive 2018 in Spain." *Energy* 276 (2023): 127539.

[45] Ma, Lingyong, Dan Luo, Huiling Hu, Qing Li, Ruitong Yang, Shu Zhang, and Dong Li. "Energy performance of a rural residential building with PCM-silica aerogel sunspace in severe cold regions." *Energy and Buildings* 280 (2023): 112719.

[46] Krarti, Moncef. "Optimal energy performance of dynamic sliding and insulated shades for residential buildings." *Energy* 263 (2023): 125699.

[47] Wang, Xia, Jiachen Yuan, Kairui You, Xianrui Ma, and Zhaoji Li. "Using Real Building Energy Use Data to Explain the Energy Performance Gap of Energy-Efficient Residential Buildings: A Case Study from the Hot Summer and Cold Winter Zone in China." *Sustainability* 15, no. 2 (2023): 1575.

[48] Ozarisoy, Bertug, and Hasim Altan. "Significance of occupancy patterns and habitual household adaptive behaviour on home-energy performance of post-war social-housing estate in the South-eastern Mediterranean climate: Energy policy design." *Energy* 244 (2022): 122904.

[49] Jay W. "System dynamics, systems thinking, and soft OR." *System dynamics review* 10, no. 2-3 (1994): 245-256.

[50] J.S. Chou, D.K. Bui, "Modeling heating and cooling loads by artificial intelligence for energy-efficient building design", Energy Build. 82 (2014) 437–446. https://doi.org/10.1016/j.enbuild.2014.07.036.

[51] M.Y. Cheng, M.T. Cao, Accurately predicting building energy performance using evolutionary multivariate adaptive regression splines, Appl. Soft Comput. J. 22 (2014) 178–188. https://doi.org/10.1016/j.asoc.2014.05.015.

[52] S. Boukabour, A. Masmoudi, Semiparametric Bayesian networks for continuous data, Commun. Stat. - Theory Methods. 50 (2021) 5974–5996. https://doi.org/10.1080/03610926.2020.1738486.

[53] E.M. Abdelkader, A. Al-Sakkaf, R. Ahmed, A comprehensive comparative analysis of machine learning models for predicting heating and cooling loads, Decis. Sci. Lett. 9 (2020) 409–420. https://doi.org/10.5267/j.dsl.2020.3.004.

[54] M. Irfan, F. Ramlie, Analysis of Parameters which Affects Prediction of Energy Consumption in Buildings using Partial Least Square (PLS) Approach, J. Adv. Res. Appl. Sci. Eng. Technol. 25 (2021) 61–68. https://doi.org/10.37934/araset.25.1.6168.

[55] L. Yang, S. Liu, S. Tsoka, L.G. Papageorgiou, A regression tree approach using mathematical programming, Expert Syst. Appl. 78 (2017) 347–357. https://doi.org/10.1016/j.eswa.2017.02.013.

[56] I. Gkioulekas, L.G. Papageorgiou, Tree regression models using statistical testing and mixed integer programming, Comput. Ind. Eng. 153 (2021) 107059. https://doi.org/10.1016/j.cie.2020.107059.

Accepted manuscript doi:
10.1680/jener.23.00017

[57] Turhan, Cihan, Tugce Kazanasmaz, Ilknur Erlalelitepe Uygun, Kenan Evren Ekmen, and Gulden Gokcen Akkurt. "Comparative study of a building energy performance software (KEP-IYTE-ESS) and ANN-based building heat load estimation." *Energy and Buildings* 85 (2014): 115-125.

[58] KEP-SDM, Dwelling Energy Performance-Standart Assessment Procedure, Chambers of Mechanical Engineers, Izmir, Turkey (2008)

Table 1. Input and output parameters, following Tsanas and Xifara (2012).

Input/Output Parameter	Symbol Used	Number of Level Studied
Relative Compactness	X1	12
Surface Area	X2	12
Wall Area	X3	7
Roof Area	X4	4
Overall Height	X5	2
Orientation	X6	4
Glazing Area	X7	4
Glazing Area Distribution	X8	6
Heating Load	Y1	586
Cooling Load	Y2	636

Table 2. Glazing area distribution scenarios Tsanas and Xifara (2012).

Distribution Scenario simulated for all three Glazing Areas	% of Glazing Area
Uniform	Equal 25% on every of the four sides
North	55% in north and rest remaining three having equal 15%
East	55% in north and rest remaining three having equal 15%
South	55% in north and rest remaining three having equal 15%
West	55% in north and rest remaining three having equal 15%

Table 3. Performance of the single hidden layer networks investigated in the present study.

			Trained Data	Untrained Data	
Test Case	Training Algorithm	No. of Neurons	Network MSE (Normalized Data)	MSE_HL (De-normalized Data)	MSE_CL (De-normalized Data)
1	Levenberg-Marquardt	8	5.38E-04	2.53E-01	1.80E+00
2		16	2.30E-04	2.11E-01	1.46E+00
3		24	1.78E-04	1.53E-01	9.66E-01
4		32	1.48E-04	2.46E-01	1.45E+00
5	Bayesian Regularization	8	5.67E-04	2.35E-01	2.21E+00
6		16	1.59E-04	1.15E-01	7.49E-01
7		24	9.55E-05	8.11E-02	8.93E-01
8		32	6.30E-05	1.73E-01	1.19E+00
9		64	2.73E-05	5.86E-01	9.43E-01
10		80	3.36E-05	1.17E-01	8.86E-01
11	Scaled Conjugate Gradient	8	5.22E-03	1.48E+01	1.59E+01
12		16	4.48E-03	5.30E+00	1.24E+01
13		24	4.02E-03	7.16E+00	1.74E+01
14		32	3.27E-03	6.73E+00	9.07E+00

Table 4. Performance of the deep learning networks investigated in the present study.

			Trained Data	Untrained Data	
Test Case	Hidden Layer-1 No. of Neurons	Hidden Layer-2 No. of Neurons	Network MSE (Normalized Data)	MSE_HL (De-normalized Data)	MSE_CL (De-normalized Data)
1	16	8	1.65E-04	1.04E-01	1.30E+00
2	16	12	1.52E-04	1.87E-01	1.92E+00
3	24	8	8.44E-05	2.05E-01	1.33E+00
4	24	12	9.03E-05	2.10E-01	8.76E-01
5	24	16	9.53E-05	2.58E-01	2.53E+00
6	32	12	6.58E-05	4.08E-01	1.94E+00
7	32	16	7.34E-05	1.26E-01	1.72E+00
8	32	20	7.63E-05	5.23E-01	1.46E+00
9	40	20	6.56E-05	4.87E-01	3.10E+00

Table 5. k-fold cross validation for network BR_16

	Trained Data	Untrained Data	
Fold Number	Network MSE (Normalized Data)	MSE_HL (De-normalized Data)	MSE_CL (De-normalized Data)
1	3.25E-04	1.07E-01	9.09E-01
2	2.46E-04	1.36E-01	8.59E-01
3	1.97E-04	1.42E-01	1.22E+00
4	1.55E-04	1.33E-01	9.22E-01
5	3.17E-04	1.29E-01	9.09E-01
6	2.36E-04	1.46E-01	6.94E-01
7	2.10E-04	2.30E-01	1.26E+00
8	1.45E-04	1.76E-01	1.36E+00
9	2.07E-04	1.48E-01	1.18E+00
10	2.45E-04	2.27E-01	3.27E+00
Mean	2.28E-04	1.58E-01	1.26E+00

Table 6. k-fold cross validation for network BR_24_12

	Trained Data	Untrained Data	
Fold Number	Network MSE (Normalized Data)	MSE_HL (De-normalized Data)	MSE_CL (De-normalized Data)
1	8.90E-05	2.27E+00	5.64E+00
2	6.87E-05	2.83E+00	6.17E+00
3	8.21E-05	8.65E-01	2.60E+00
4	7.10E-05	2.08E-01	8.45E-01
5	5.53E-05	3.92E-01	1.16E+00
6	8.22E-05	6.51E-01	2.85E+00
7	6.65E-05	8.59E-01	2.59E+00
8	7.85E-05	1.70E+00	5.51E+00
9	6.06E-05	2.80E+00	8.10E+00
10	5.97E-05	2.40E+00	8.70E+00
Mean	7.14E-05	1.50E+00	4.41E+00

Table 7. Percentage of Glazing Area Distribution

Test Case	North	East	South	West
1	32.5	22.5	22.5	22.5
2	40	20	20	20
3	47.5	17.5	17.5	17.5
4	45	25	15	15
5	35	35	15	15
6	25	45	15	15
7	15	45	25	15
8	15	35	35	15
9	15	25	45	15
10	15	15	45	25
11	15	15	35	35
12	15	15	25	45

Figure captions

Figure 1. a) Similarity between human neuron and artificial neuron b) flow diagram of single neuron c) a schematic diagram showing signal processing by an artificial neuron.

Figure 2. Flow diagram of the procedure adopted to train the data set in MATLAB.

Figure 3. A schematic diagram of the deep learning network

Figure 4. A schematic diagram showing parameters investigated in the benchmark study of Tsanas and Xifara (2012).

Figure 5. Parallel plot of data set (Tsanas and Xifara (2012)) showing functional dependency of (a) heating load and (b) cooling load on the input parameters.

Figure 6. Performance analysis of the network (a) regression of training data set (b) regression of test data set (c) regression of Target.

Figure 7. Variation of Mean Square Error (MSE) with Epochs for one representative network.

Figure 8. Comparison of prediction performance of BR-16 (best performing) and SCG-8 (worst performing) networks among the tested networks with single hidden layer.

Figure 9. Schematic showing K-Fold cross validation procedure.

Figure 10. Comparison of actual (Tsanas and Xifara (2012)) and predicted heating load for the dataset used in network development and training.

Figure 11. Comparison of actual (Tsanas and Xifara (2012)) and predicted cooling load for the dataset used in network development and training.

Figure 12. Histogram showing absolute error variation for (a) Heating Load (b) Cooling Load

Figure 13. Comparison of experimental and predicted results of heating load for the untrained data (a) Glazing area distribution (b) Wall area (c) Surface Area (d) Relative Compactness

Figure 14. Comparison of experimental and predicted results of cooling load for the untrained data (a) Glazing area distribution (b) Wall area (c) Surface Area (d) Relative Compactness.

Figure 15. Influence of glazing area distribution and glazing area on the heating and cooling load.

Accepted manuscript doi: 10.1680/jener.23.00017

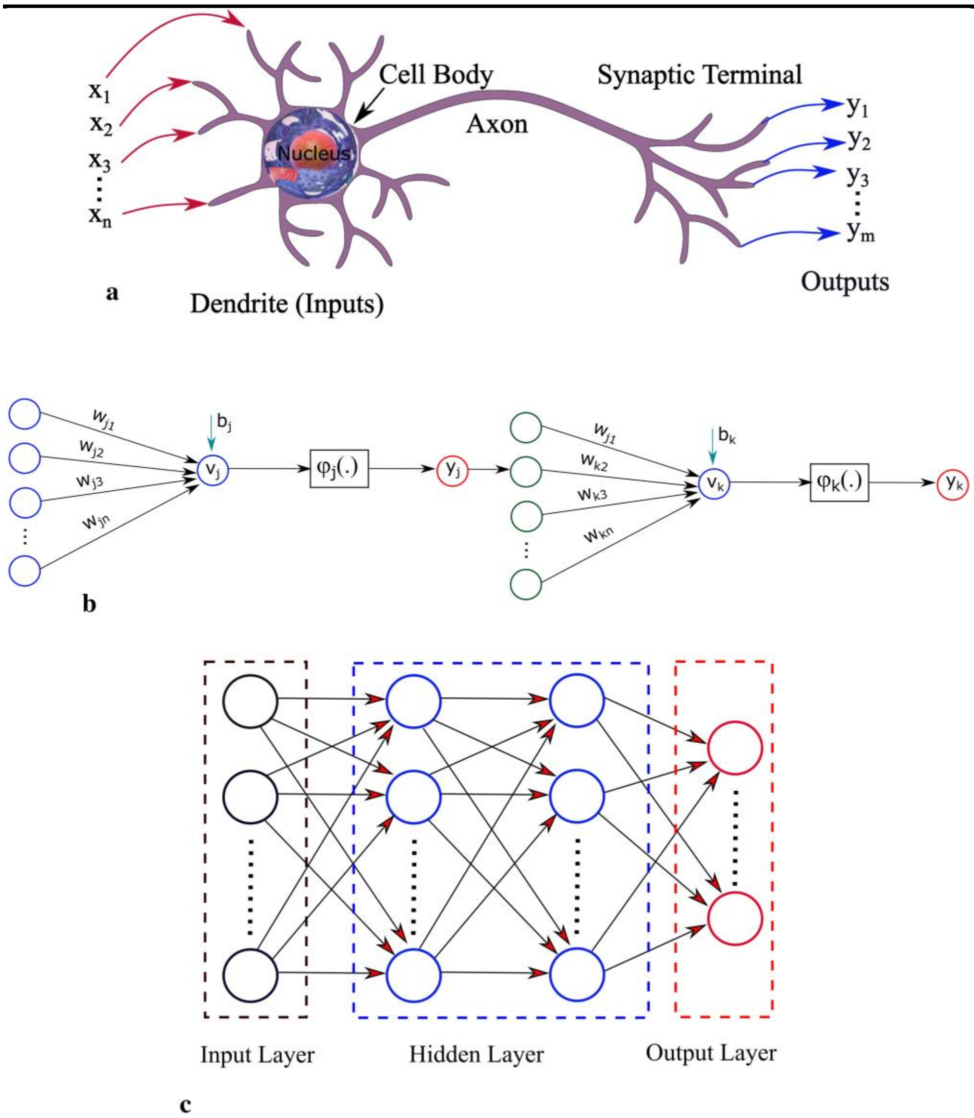

Figure 1

Accepted manuscript doi:
10.1680/jener.23.00017

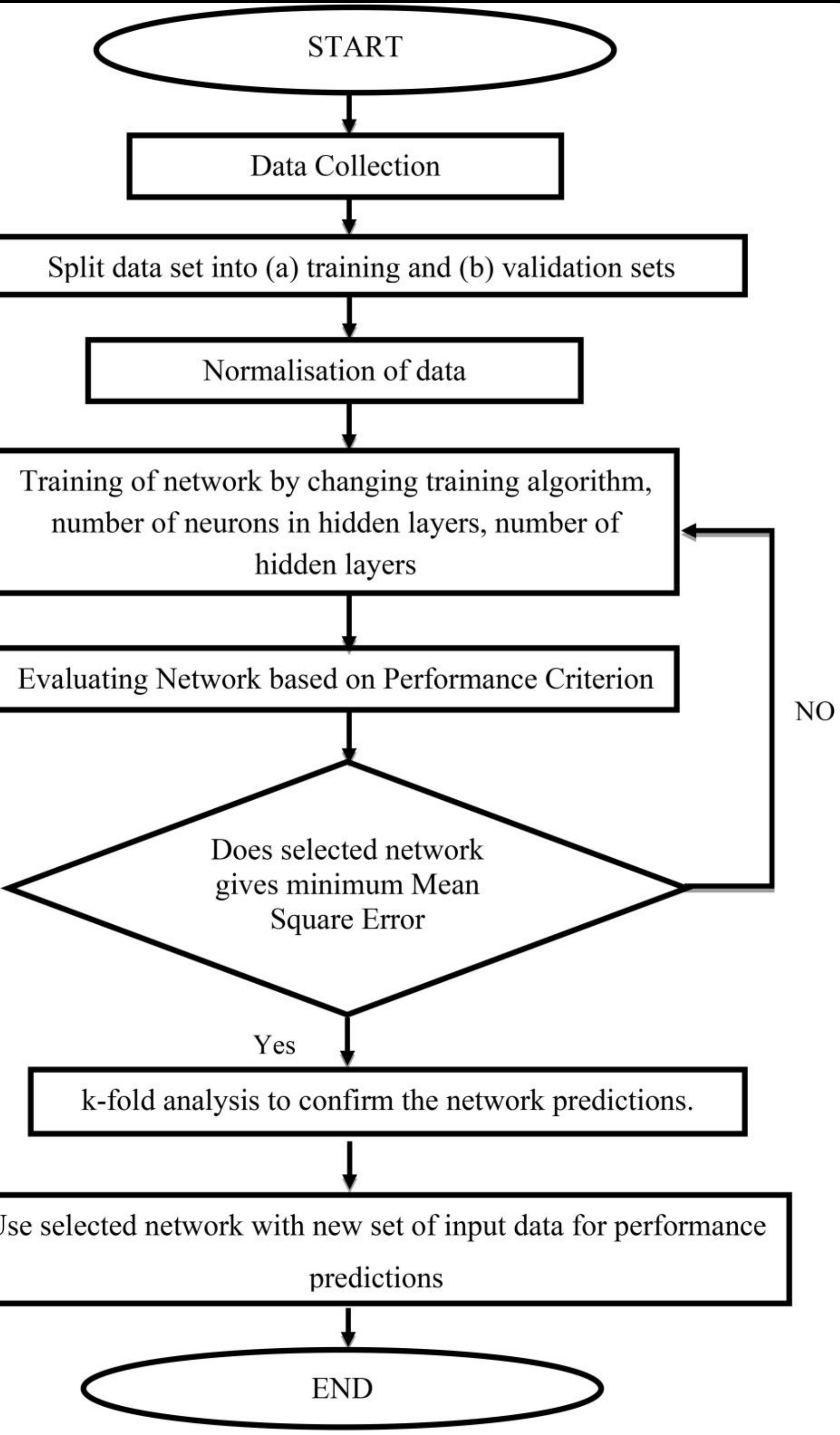

Figure 2

Accepted manuscript doi:
10.1680/jener.23.00017

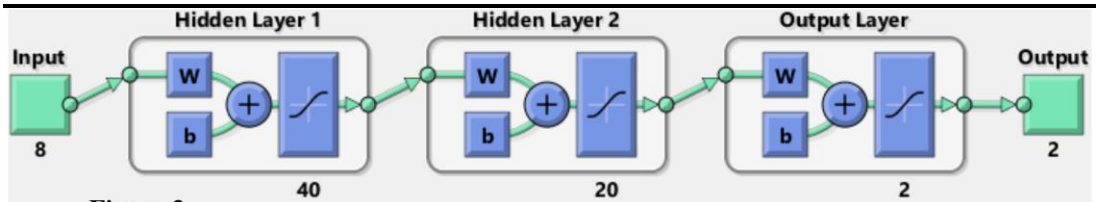

Figure 3

Accepted manuscript doi:
10.1680/jener.23.00017

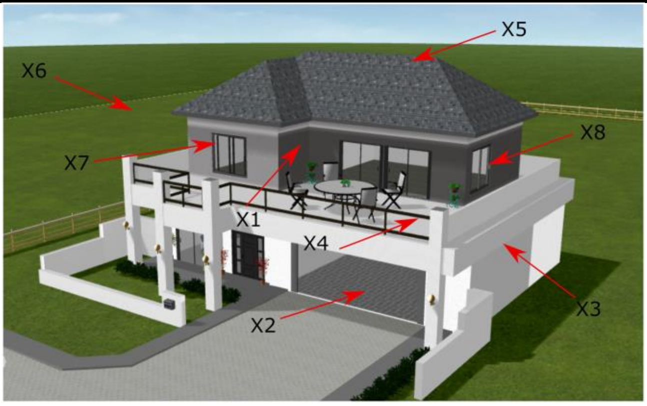

Figure 4

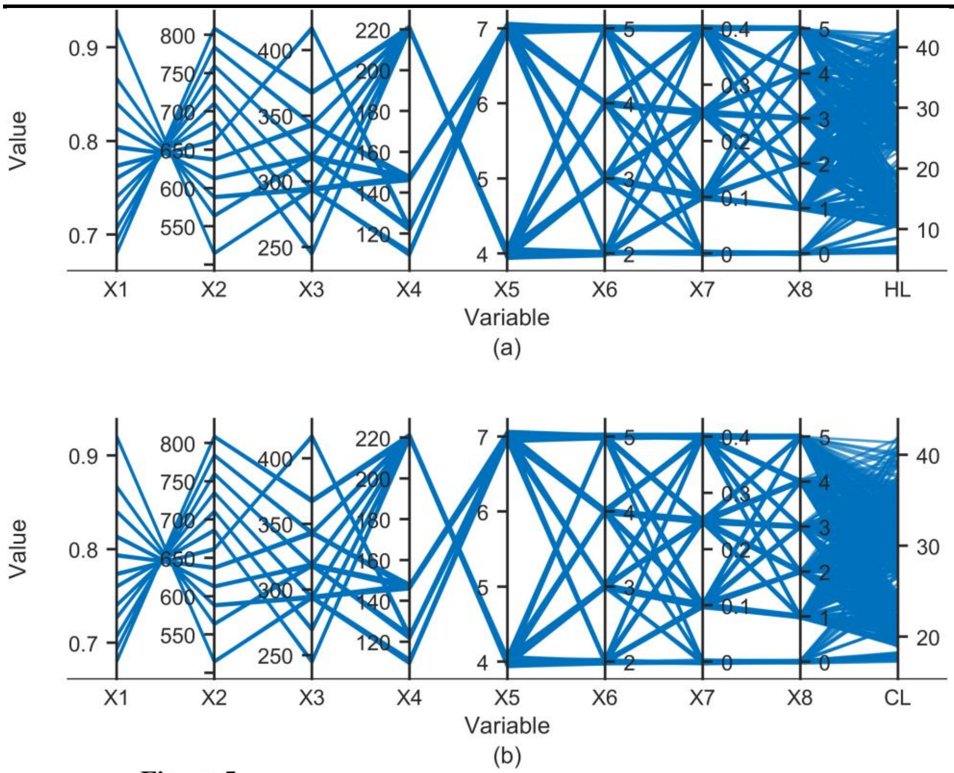

Figure 5

Accepted manuscript doi:
10.1680/jener.23.00017

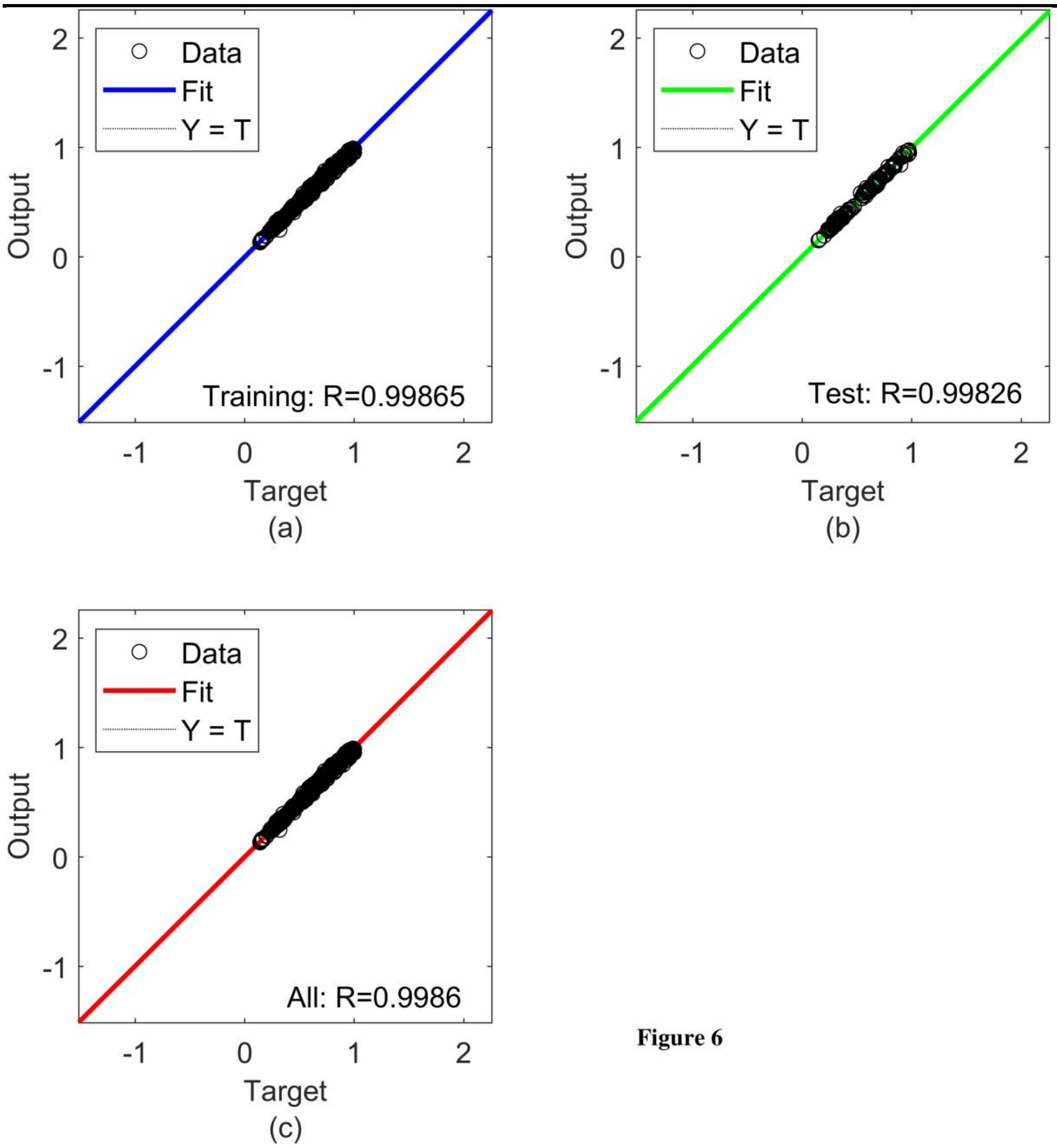

Figure 6

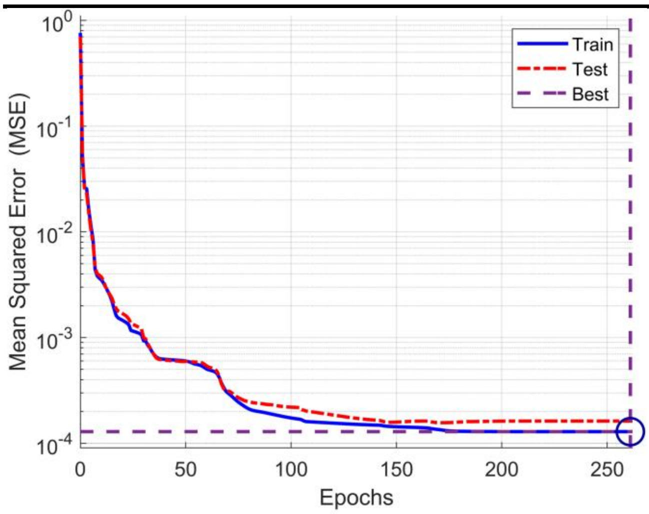
Mean Squared Error (MSE)
Epochs
Train
Test
Best
10^{0}
10^{-1}
10^{-2}
10^{-3}
10^{-4}
0
50
100
150
200
250

Figure 7

Accepted manuscript doi:
10.1680/jener.23.00017

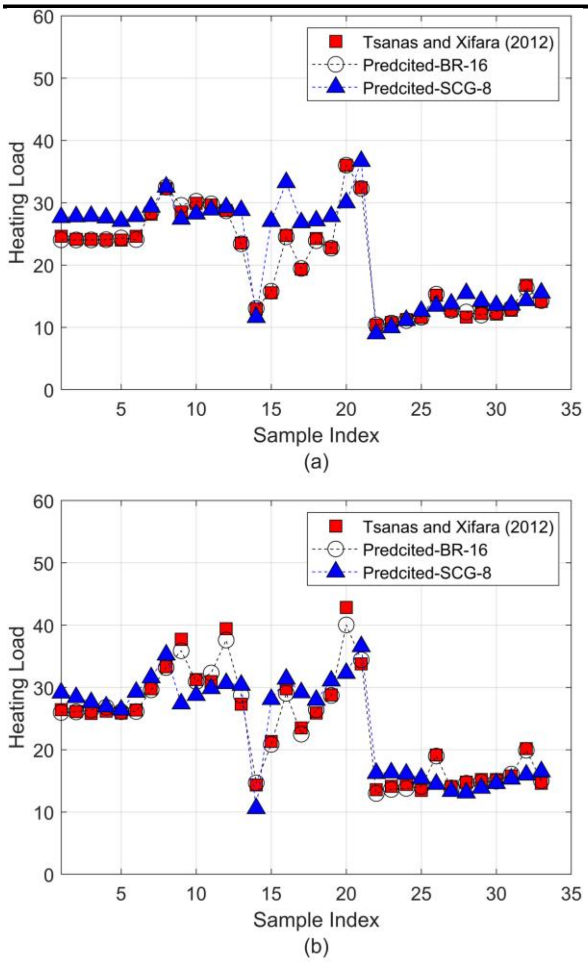

Figure 8

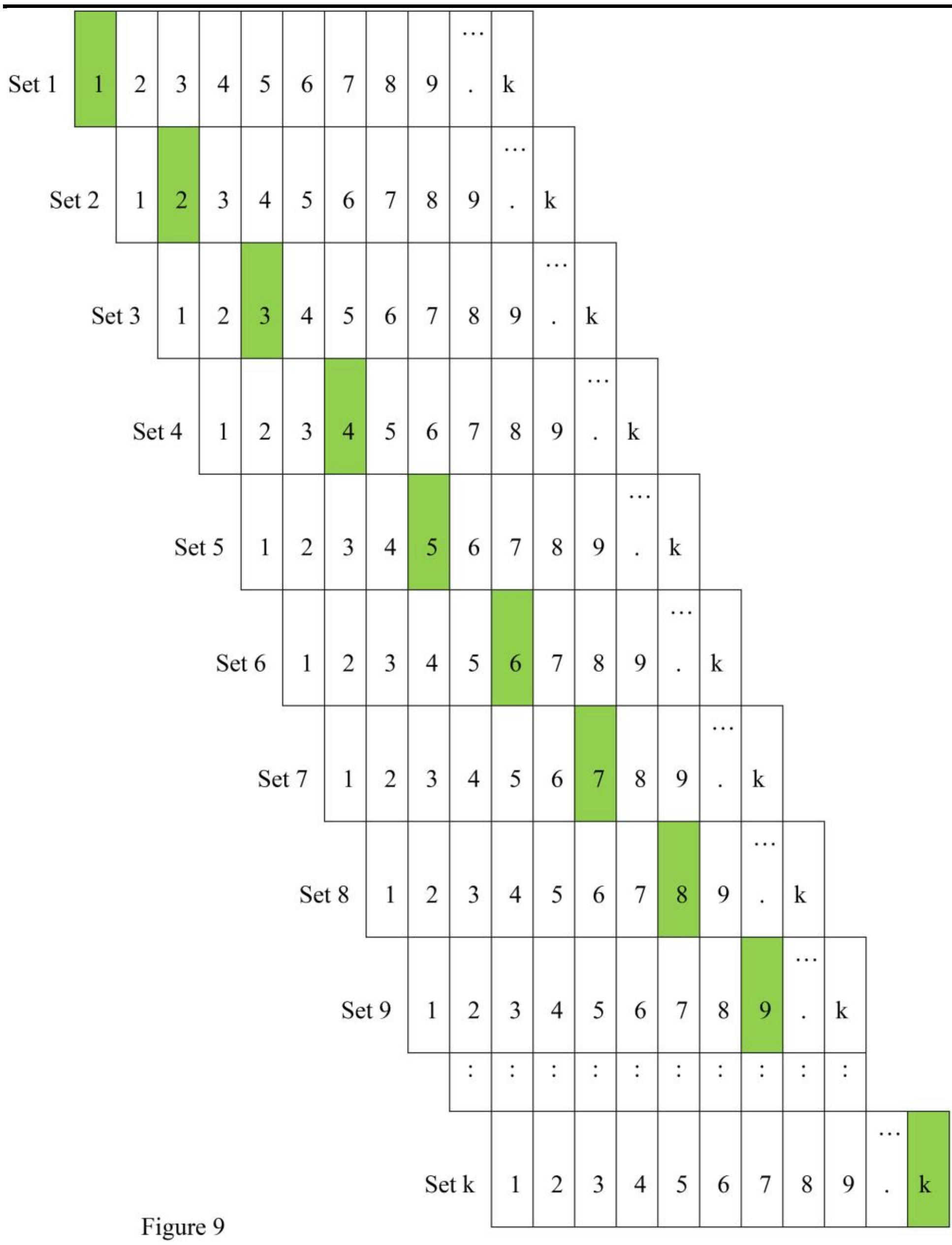
Set 1 1 2 3 4 5 6 7 8 9 … k
Set 2 1 2 3 4 5 6 7 8 9 … k
Set 3 1 2 3 4 5 6 7 8 9 … k
Set 4 1 2 3 4 5 6 7 8 9 … k
Set 5 1 2 3 4 5 6 7 8 9 … k
Set 6 1 2 3 4 5 6 7 8 9 … k
Set 7 1 2 3 4 5 6 7 8 9 … k
Set 8 1 2 3 4 5 6 7 8 9 … k
Set 9 1 2 3 4 5 6 7 8 9 … k
Set k 1 2 3 4 5 6 7 8 9 … k

Figure 9

Accepted manuscript doi:
10.1680/jener.23.00017

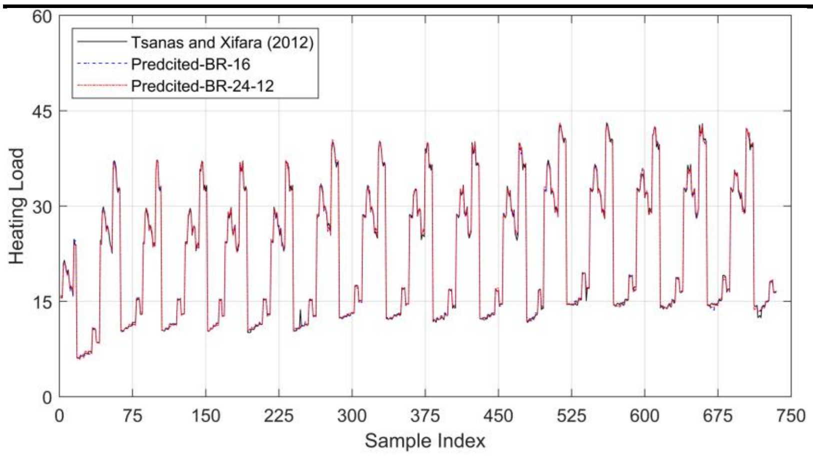

Figure 10

Accepted manuscript doi:
10.1680/jener.23.00017

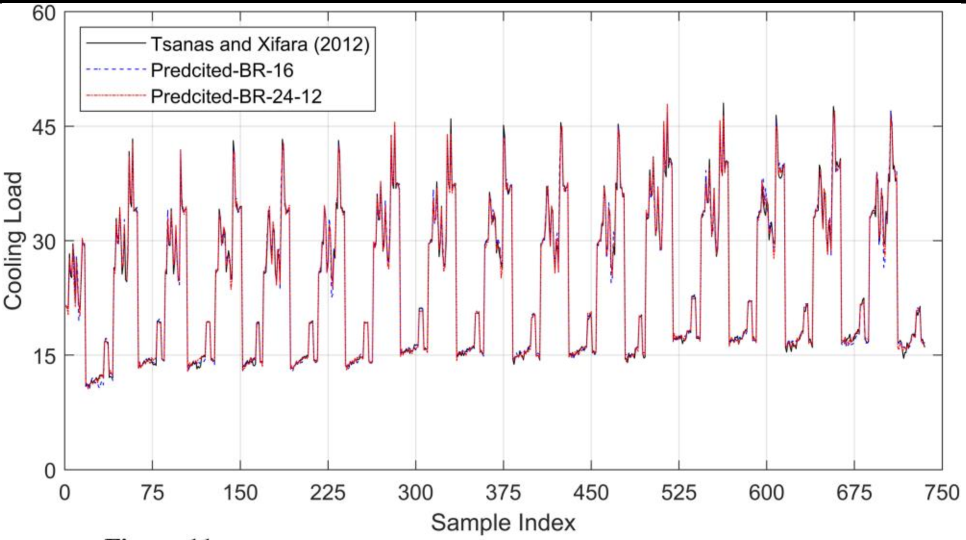

Figure 11

Accepted manuscript doi:
10.1680/jener.23.00017

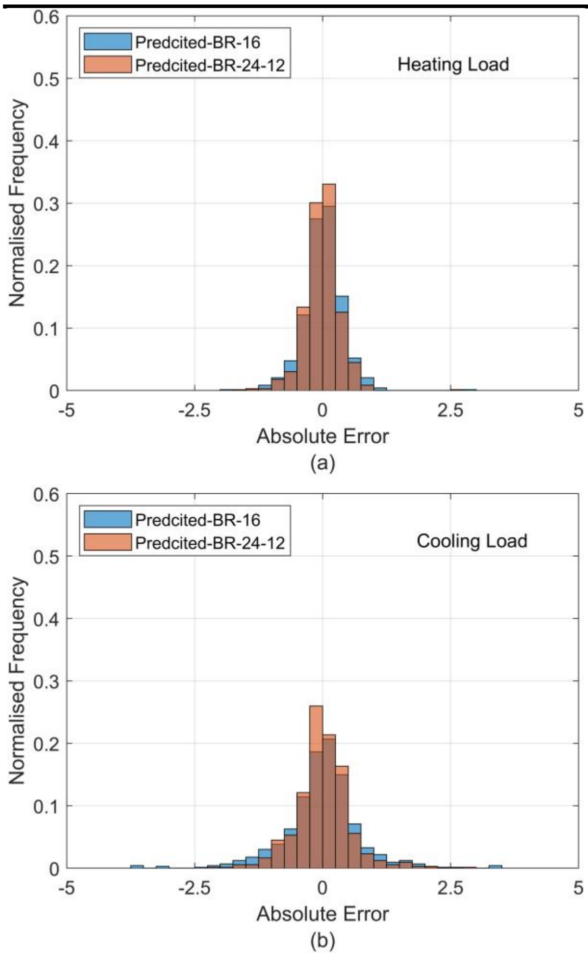

Figure 12

Accepted manuscript doi: 10.1680/jener.23.00017

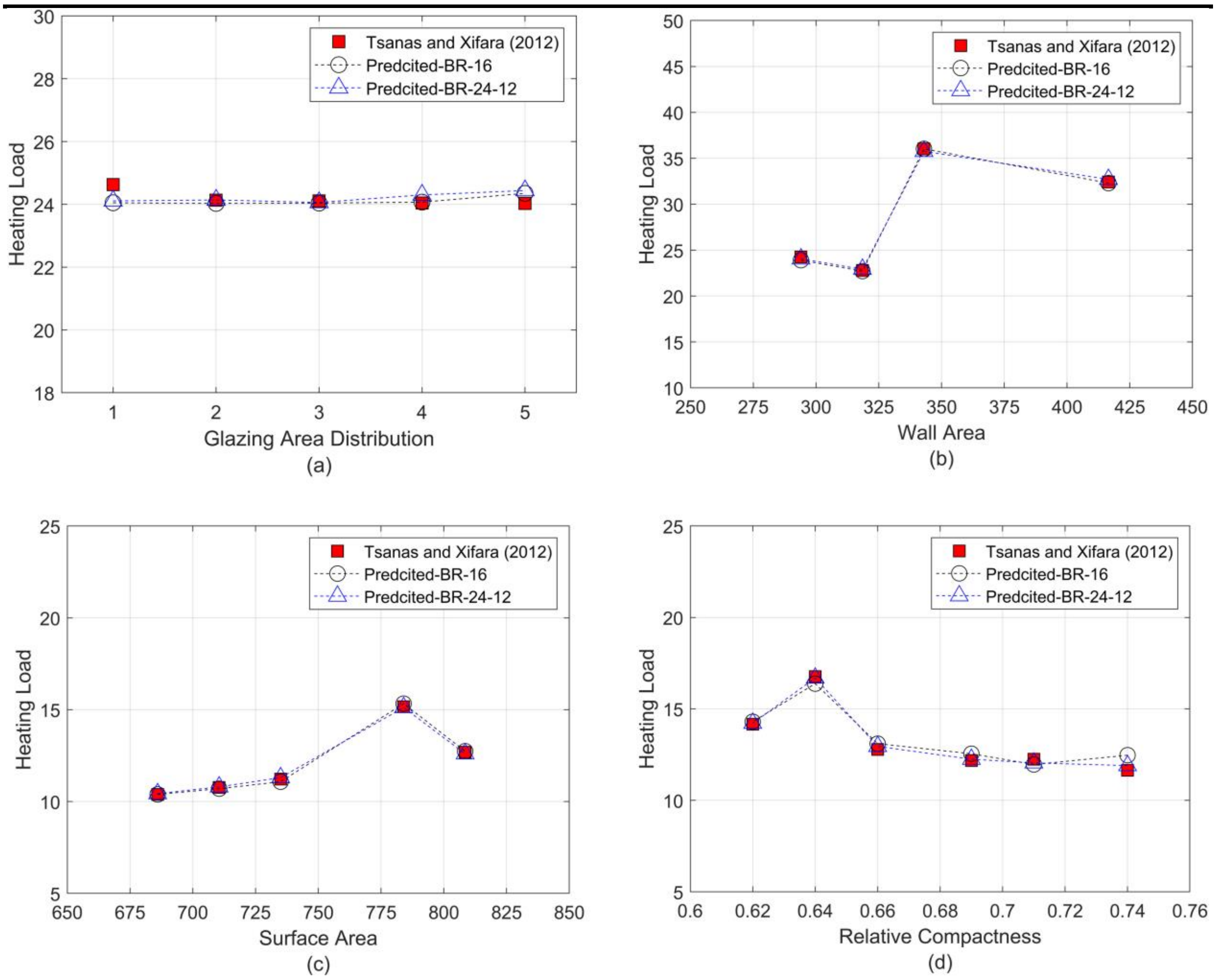

Figure 13

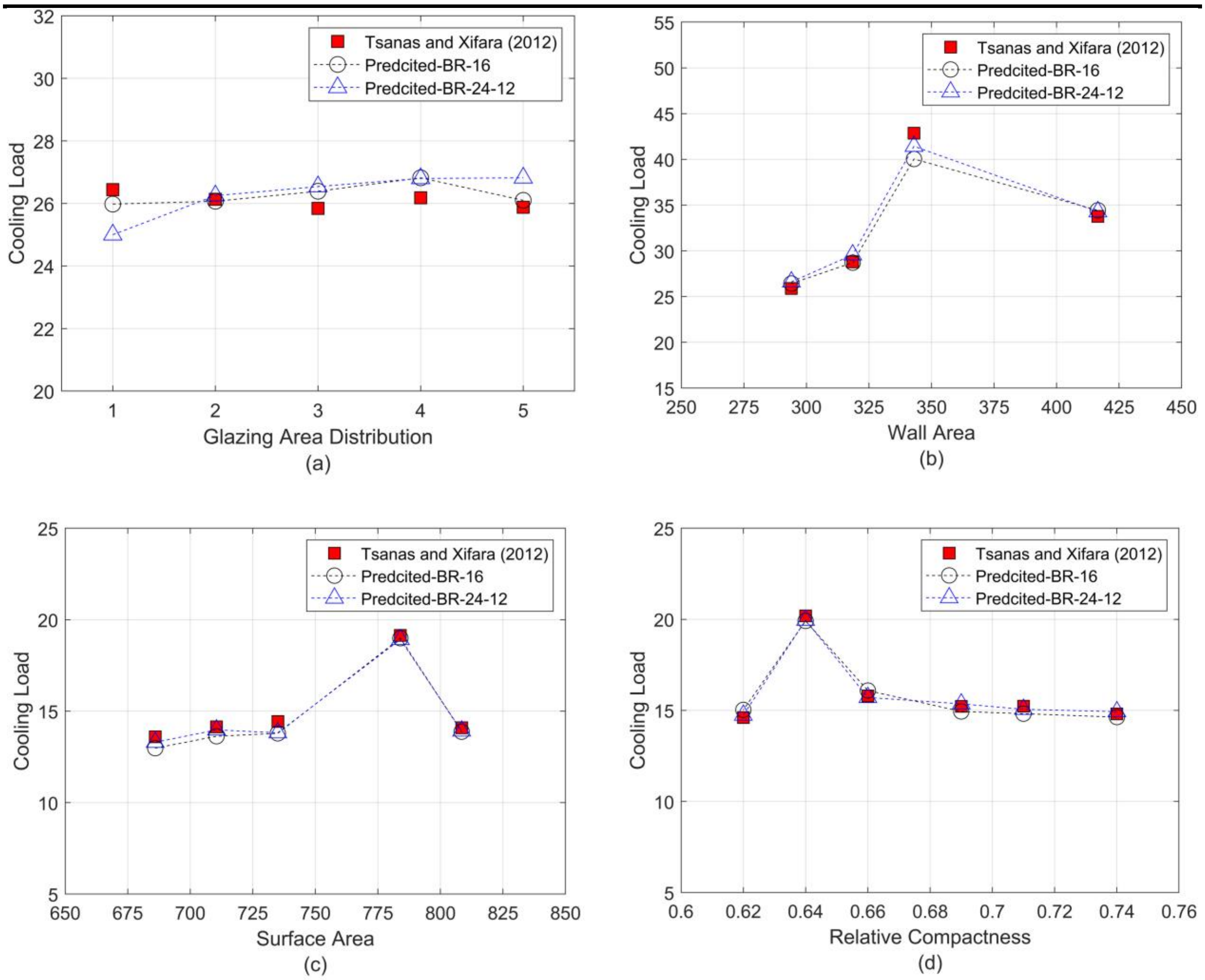
Tsanas and Xifara (2012)
Predcited-BR-16
Predcited-BR-24-12
Cooling Load
Glazing Area Distribution
(a)
Wall Area
(b)
Surface Area
(c)
Relative Compactness
(d)

Figure 14

Accepted manuscript doi:
10.1680/jener.23.00017

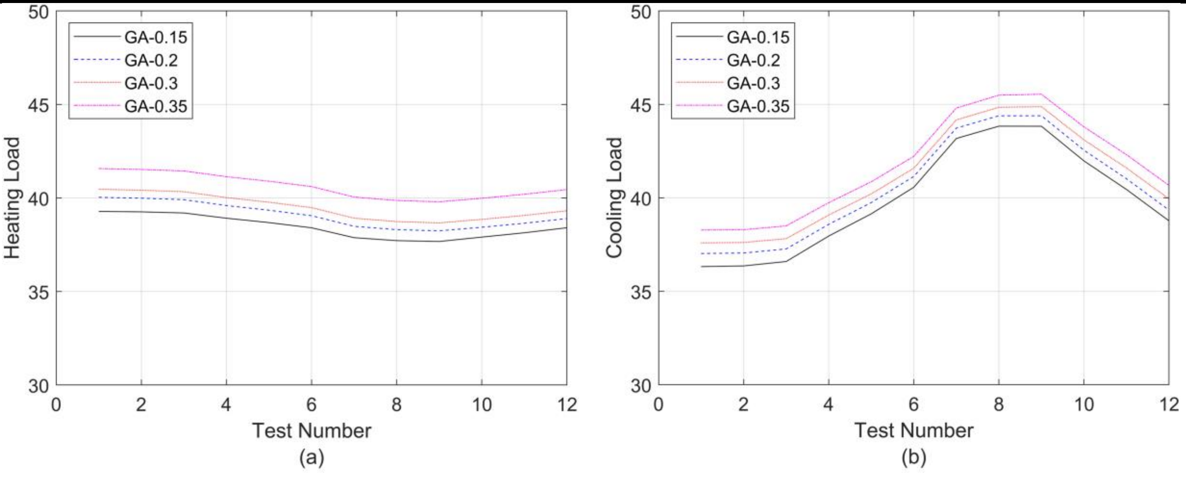

Figure 15