A STUDY ON THE APPLICATION OF ARTIFICIAL INTELLIGENCE TECHNIQUES FOR PREDICTING THE HEATING AND COOLING LOADS OF BUILDINGS

Sushmita Das,[1] Aleena Swetapadma,[2*] and Chinmoy Panigrahi[3]

ABSTRACT

The prediction of the heating and cooling loads of a building is an essential aspect in studies involving the analysis of energy consumption in buildings. An accurate estimation of heating and cooling load leads to better management of energy related tasks and progressing towards an energy efficient building. With increasing global energy demands and buildings being major energy consuming entities, there is renewed interest in studying the energy performance of buildings. Alternative technologies like Artificial Intelligence (AI) techniques are being widely used in energy studies involving buildings. This paper presents a review of research in the area of forecasting the heating and cooling load of buildings using AI techniques. The results discussed in this paper demonstrate the use of AI techniques in the estimation of the thermal loads of buildings. An accurate prediction of the heating and cooling loads of buildings is necessary for forecasting the energy expenditure in buildings. It can also help in the design and construction of energy efficient buildings.

KEYWORDS

building energy performance, heating and cooling load, Artificial Neural Network, Support Vector Machine, Iteratively Reweighted Least Squares, Random Forest

I. INTRODUCTION

Rapid advances in building technology and growing urbanization has led to a vast increase in the number of buildings worldwide. This increase has led the building sector to become a major energy consuming entity. Energy consumption by buildings is as high as 32% of the overall energy consumption worldwide [1]. Thus, it has become a key area for researchers to devise methods to make buildings more energy efficient and environmentally friendly. The energy consumption in buildings can be attributed to their lighting system, heating, ventilation and air conditioning (HVAC) system, pumping system and various other electrical equipment used by building occupants. Among these subsystems in a building, the HVAC systems consume the largest amount of energy. Almost 30% of the energy consumed by buildings can be attributed to

1. Department of Electrical Engineering, Trident Academy of Technology, Bhubaneswar, India
2. School of Computer Engineering, KIIT University, Bhubaneswar, India;
*Corresponding author: aleena.swetapadma@kiit.ac.in
3. School of Electrical Engineering, KIIT University, Bhubaneswar, India

their HVAC systems [2, 3, 5, 6]. Thus, there is ample scope to make buildings energy efficient by making HVAC systems more efficient. In order to design better HVAC systems for buildings with regard to reduced energy consumption, it is important to forecast the heating and cooling loads of buildings as accurately as possible. Conventional methods of calculating the heating and cooling loads of buildings is very time consuming. An alternative approach to the prediction of heating and cooling loads of buildings is by using AI techniques. There may be two approaches, namely blackbox modelling and gray box modelling. In blackbox modelling, usually the sensor data from buildings is used for modelling and a prediction of the heating and cooling loads; whereas, in the gray box modelling approach, building characteristics such as surface area, glazing, wall area, relative compactness, orientation etc. are also taken into account in addition to data collected through sensors and weather data. Several AI techniques like Artificial Neural Networks (ANNs), Support Vector Regression (SVR), Classification and Regression Tree (CART), Iteratively Reweighted Least Squares (IRLS), General Regression Neural Networks (GRNN), Random Forest (RF), Chi-Squared automatic interaction detector, ensemble models, deep learning algorithm and linear regression models have been used for the purpose of predicting the heating and cooling loads of buildings. Table 1 displays the key traits of each AI technique mentioned. This paper presents a review of research that has used AI techniques for the estimation of the heating and cooling loads in buildings.

2. DISCUSSIONS OF MODELS PROPOSED FOR HL AND CL ESTIMATION OF BUILDINGS

This section presents an overview of the AI techniques used for the estimation of HL and CL of buildings, the building types considered, the software used to calculate the HL and CL and set the reference values and the input variables considered for each model. Table 2 presents the overview of the models discussed in the next section:

3. DISCUSSION OF THE MODELS

A. Tsanas and A. Xifara in [12] have proposed statistical machine learning models for accurately predicting the heating load and cooling load of residential buildings. The authors have selected eight input variables for prediction of the heating load and cooling load which are the two output variables. The input variables are selected based on previous literature which suggests these variables to be the factors primarily affecting a building's energy requirements. The variables chosen as inputs are relative compactness, surface area, wall area, roof area, overall height, orientation, glazing area and glazing area distribution. The building dataset was generated using Ecotect. The volume of all the buildings was the same (771.75 m^3). However, the surface areas and dimensions of the buildings were different. Low U-Value building materials were considered. A total of 768 buildings were studied. These building samples were generated using Ecotect considering 12 building forms, 3 glazing area variations with 5 glazing area distributions each for 4 orientations and 12 building forms for 4 orientations without glazing ((12x3x5x4=720) + (12x4=48) = 768). The machine learning techniques used to analyse the data were IRLS (Iteratively Reweighted least Squares) and RF (Random Forests). The values of HL and CL obtained from Ecotect were assumed to be the actual values and the results obtained using the proposed models were compared against these actual values. The results showed that RF greatly outperformed IRLS. The performance of the two models was evaluated in terms of MAE, MSE and MRE. A limitation of the study is that climate and occupancy were assumed

TABLE 1. Various methods used.

AI Technique	Advantages	Disadvantages
ANN	Capable of learning and modelling nonlinearities and complexities in data. Ability to generalize and predict new data presented to model. Does not impose any limitations on the input variables and their distribution.	Computational time taken for deep learning networks is long. Huge data is required for ANN architectures with multiple layers. Are not probabilistic.
SVR	Has powerful regularization characteristics which are helpful in generalizing the model to new data. Ability to perform satisfactorily on datasets with several attributes even if number of cases available for training the model are less.	Satisfactory performance largely depends on the proper selection of kernel. Computation time taken is longer. Complex algorithms are used and large memory required for programming in big tasks.
CART	Ability to perform feature selection. Data preparation tasks like tackling differences in parameter scales and accounting for missing values in the data are not required.	Even a small variation in data can result in a large change in the design of the optimal decision tree. Often gives less accurate results in comparison to other techniques.
IRLS	Can be used with Gauss-Newton and Levenberg-Marquart algorithms.	Computational time taken is longer.
GRNN	Applicable to regression, prediction and classification problems. Uses single pass learning approach so no back propagation is required. High accuracy of prediction.	Huge data involved can make computation costlier. Not possible to improve network by optimisation.
RF	Highly flexible and accurate. Does not require data preparation. Helps overcome the problem of over fitting faced while using a single decision tree. Performs satisfactorily even if a large amount of data is missing.	Highly complex and takes a lot of time to construct as compared to decision trees. Less intuitive. Computational time is very large.

to be constant. Including these as input variables will improve the accuracy of prediction for HL and CL.

The authors in [13] have used different data analytic methods for the prediction of heating and cooling loads of buildings. The methods used are Support vector regression (SVR), ANN, Classification and Regression Tree (CART), General Linear Regression, Chi-Squared Automatic Interaction Detector and Ensemble/Combined models of two, three, four and five single models. The building data was the same as the data used in [12]. Ecotect was used to simulate twelve building types with the same volume and same materials used for construction but their

TABLE 2. Overview of models described.

Authors and Reference no.	No. of input variables	Input variables	Software used for calculation of HL and CL	Outputs	AI Techniques used	Building Typology
A. Tsanas and A. Xifara [12]	8	1. relative compactness 2. surface area 3. wall area 4. roof area 5. overall height 6. orientation 7. glazing area 8. glazing area distribution	Ecotect	HL and CL	RF and IRLS	Residential
J. Chou and D. Bui [13]	8	Same as in [12]	Same as in [12]	HL and CL	SVR, ANN, CART, Ensemble Models	Residential
M. Cheng and M. Cao [14]	8	Same as in [12]	Same as in [12]	HL and Cl	EMARS, MARS, BPNN, RBFNN, CART, SVM	Residential
D. Kapetanakis *et al* [15]	9	1. atmo-spheric temperature 2. ambient rela-tive humidity 3. wind speed 4. solar radiation 5. clearness of the sky 6. indoor air temperature 7. indoor rela-tive humidity 8. occupancy 9. CO_2 levels	Energy Plus	HL	Multilayer Perceptron ANN, RBFNN	Commercial
Y. Sonmez *et al* [16]	8	Same as in [12]	Same as in [12]	HL and CL	k-NN, abc-kNN, ga-kNN, classical ANN, ga-ANN, abc-ANN	Residential
C. Deb and *et al* [17]	5	1. air conditioned load 2. non-aircondi-tioned load 3. air temperature 4. relative humidity 5. solar radiation	----	CL	Feed forward ANN	Commercial (Educational)

TABLE 2. *(Continued)*

Authors and Reference no.	No. of input variables	Input variables	Software used for calculation of HL and CL	Outputs	AI Techniques used	Building Typology
C. Fan *et al* [18]	7	1. Month 2. Day 3. Hour 4. Minute 5. Day type 6. Outdoor temperature 7. Outdoor relative humidity	------	CL	Deep Learning	Commercial (Educational)
B. Gunay *et al* [19]	6	1. Outdoor air temperature 2. Horizontal solar irradiance 3. Wind speed 4. Moisture content of outdoor air 5. Electrical loads 6. Work hour	Energy Plus	HL and CL	ANN	Commercial (Office)
H. Pombeiro *et al* [20]	10	1. Weekday 2. Hour 3. Minute 4. Occupancy 5. Temperature 6. Relative humidity 7. Wind speed 8. Atmospheric pressure 9. Precipitation 10. Solar radiation	------	HL and CL	Fuzzy networks, Neural networks	Commercial (Institutional)
A.Jihad and M.Tahiri [21]	6	1. Height 2. Relative Compactness 3. Wall surface 4. Building surface 5. Orientation 6. Window ratio	DesignBuilder	HL and CL	ANN	Residential

dimensions and surface areas were different. Building activities were assumed to be sedentary in nature. Eight input variables, namely Relative compactness, Surface area, Wall area, Roof area, Overall height, Orientation, Glazing area, Glazing area distribution were considered and the two output variables were the heating load and cooling load. The performances of the various models used for estimation of heating and cooling load were compared using evaluation indicators like Root Mean Square Error (RMSE), Mean Absolute Error (MAE), Mean Absolute Percentage Error (MAPE), Linear Correlation Coefficient R and a Synthesis Index (SI) which is a combination of RMSE, MAE, MAPE AND 1-R. The performance outputs of the different models showed that the cooling load was predicted with the highest accuracy using the ensemble model of SVR and ANN with a SI of 0.11; whereas, the model best suited for estimation of the heating load was the SVR model with a SI of 0.00. It was also seen that the computation time taken by the suggested models was significantly less and within a few seconds. The limitation of this work is the use of default settings in the models, both for the single and ensemble models. The models proposed in the paper are only applicable to the twelve building types which are specified. The authors suggest that higher accuracy in prediction of HL and CL of the buildings can be achieved by optimization of parameters in the models (single and ensemble) by using evolutionary computing and swarm intelligence techniques.

The authors M. Cheng and M. Cao in [14] have proposed an AI model EMARS (Evolutionary Multivariate Adaptive Regression Splines) to accurately forecast the energy performance of buildings. The EMARS model is a hybrid of MARS (Multiple Adaptive Regression Splines) and Artificial Bee Colony. The proposed model was developed using data from [12]. The performance of the proposed EMARS model was compared with that of five other AI models namely, MARS, Back Propagation Neural Network (BPNN), Radial Basis Function Neural Network (RBFNN), CART and Support Vector Machine (SVM). The performance of the EMARS model was compared with that of other models on the basis of RMSE (Root Mean Squared Error), MAPE (Mean Absolute Percentage Error), MAE (Mean Absolute Error) and R^2 (Coefficient of Determination). The results showed that among all models, the proposed EMARS model had the lowest RMSE values of 1.00 for cooling load and 0.47 for heating load. The EMARS model gave superior results even in terms of MAPE, MAE and R^2. The accuracy of predicting the heating load was better than that of the cooling load. The parameter analysis showed that the two factors primarily affecting the heating load are surface area and roof area whereas the cooling load is equally affected by six out of the seven factors used for mapping. The limitation of the work is the use of simulation data. The model should be tested taking actual datasets.

Data driven or blackbox models have been developed for the estimation of heating loads in commercial buildings in the work proposed in [15]. The data is collected from Building Energy Management Systems (BEMS) installed in the buildings and in addition weather data is also included. To resolve the problem of lack of some data in the data collected from BEMS, a reference building developed using Energy Plus was set as a benchmark. Sensors are located in the buildings and these sensor data are the BEMS variables. Statistical techniques were used for analyzing the correlation between the input variables and output variable. The output variable is the heating load of the building which is to be predicted and the input variables chosen for analysis were atmospheric temperature, ambient relative humidity, wind speed, solar radiation, clearness of the sky, indoor air temperature, indoor relative humidity, occupancy and CO_2 levels. The predictive models developed were Regression models- Multiple Linear, Multiple Nonlinear and Generalized Linear and ANN models- Multilayer Perceptron and Radial Basis Function. The performance accuracy of each model was evaluated using an equation involving Mean Absolute Error (MAE) and maximum and minimum predicted values. Correlation studies done by calculating Pearson coefficient and Spearman analysis showed that there is a very strong correlation- between *a.* Sky clearness and Solar Radiation and *b.* Heating Load and Gas Consumption. A combination of different situations led to the development of 90 regression models and 60 ANN models. The regression models had an average accuracy of 74.52 % whereas the ANN models had an average accuracy of 83.31 %. The most accurate regression model was obtained using generalized linear regression technique and the accuracy was 92.1%. The most accurate ANN model was generated using the multilayer perceptron approach and the accuracy was 97%. The common inputs to both these models were atmospheric temperature, ambient relative humidity, solar radiation and indoor air temperature. It was concluded that models based on ANN predicted the heating load of a building better than models based on regression techniques.

The authors in [16] have used hybrid Artificial Intelligence (AI) techniques to predict the heating load and cooling load of buildings. The data required for analysis is obtained from [12]. In this work six machine learning algorithms have been used for predicting the heating load and cooling load. These are classical estimator knn (k-nearest neighbor), artificial-bee-colony based k- nearest neighbor (abc-knn), genetic algorithm based knn (ga-knn), classical ANN, ga-ANN and abc-ANN. Out of the 768 datasets, 576 were used for training and 192 for testing and verification. Results showed that among the knn models, the abc-knn gave the best performance with a MAE of 1.10. It was observed that the input variables most affecting the cooling load are surface area, overall height and glazing area whereas the input least affecting the cooling load of a building is glazing area distribution. Among the ANN models, the abc-ANN model performed best with a MAE of 0.52. When compared with the abc-knn, both the adaptive ANN models performed better. While MAE for abc-knn is 1.10, it is 0.61 for ga-ANN and 0.52 for abc-ANN. It has been concluded that the accuracy of predicting the cooling load and heating load of buildings can be significantly improved with the use of heuristic based hybrid knn and ANN methods.

A model using ANN was developed to forecast the daily cooling load energy consumption for three institutional buildings in a university campus in Singapore [17]. The three buildings were part of the same school but the purpose and use of each building varies. Energy consumption data of the three buildings was collected over two years. The forecasting model developed was made more relevant by the inclusion of weather data namely air temperature, relative humidity and solar radiation. The air conditioning load was correlated with these three climatic factors and the results of correlation showed that none of the climatic factors considered had any significant impact on the air conditioning load of the buildings. Rather it was concluded that the key factors affecting cooling load energy used in these buildings are building occupancy and indoor conditions. A feedforward ANN network consisting of an input layer, hidden layer and output layer was used. Using a trial and error method, it was seen that when the number of hidden neurons is 20, the model gave accurate results in less time. Thus, the number of neurons in the hidden layer was fixed to 20. Different ANN architectures were obtained by adding more hidden layers and tested for prediction accuracy; it was observed that the computing time for the 5-20 network was the least. A 5-20 network implies five input neurons and one hidden layer of 20 neurons. A Bayesian regularization algorithm was used as the training algorithm and the training was done in MATLAB. The energy consumption data was equally divided into five classes ranging from very low, low, medium, high and very high. This was done to reduce the high variation in data. The data was also distributed into 25 class numbers. This helps in differentiating between the different week days, and there is no overlapping of data between class levels of different days. The inputs to the ANN were the previous five days' measured energy consumption data in the form of energy classes. With this approach of energy classes and class numbers, the model predicted results with an accuracy as high as 0.9794. The model is able to predict the daily cooling load for the next 20 days with good prediction accuracy by taking only the previous five days measured data values as inputs. The study suggests that in the future a forecasting model can be developed to make it applicable for a wider range of institutional buildings.

In [18], the authors have proposed a method to forecast the cooling load for the next 24 hours by use of deep learning algorithms. Deep learning can be defined as a method which involves the combination of various machine learning algorithms. Deep learning algorithms are capable of displaying nonlinear and complicated formations in big data. The building whose data has been used in this study is an educational building in Hong Kong. The building mainly consists of classrooms, offices and a data center. The total floor area is 11,000 m^2 out of which the conditioned area is 8500 m^2. Annual data from 2015 was collected and the collected data had a time interval of 30 minutes. The set of collected data had the mentioned variables: time variables such as month, day, hour, minute and day type (i.e. weekday or weekend), external temperature, external relative humidity, supply chilled water temperature, returned chilled water temperature and flow rate of chilled water temperature. The total number of observations in the dataset is 15,792. Deep learning technique can be utilized in a supervised way to form a forecasting model or in an unsupervised way for retrieving important attributes from unanalyzed data. In this work, first feature extraction was done to retrieve relevant attributes from the building data which were then given as inputs to the model. Four types of feature extraction methods were used to extract important engineering, statistical, structural and deep learning features from the building data. This step of feature extraction shows the unsupervised aspect of deep learning. Supervised deep learning is highlighted by the seven forecasting techniques used to form prediction models using various feature sets. The prediction techniques used are Multiple Linear Regression, Elastic Net, Random Forests, Gradient Boosting Trees, Support Vector Regression, Extreme Gradient Boosting and Deep Learning. Seventy percent of the dataset was used for training, 15% for validation and the remaining 15% for testing. The performance of the prediction models is evaluated on the basis of indicators like Mean Absolute Error (MAE), Root Mean Square Error (RMSE) and Coefficient of Variation of the Root Mean Square Error (CV-RMSE). Results show that the models developed using Multiple Linear Regression and Elastic Net give the poorest performances among all models proposed. Among the remaining five techniques, the model developed using Extreme Gradient Boosting gives the best performance. The prediction outputs using linear techniques is found to be very poor as compared to the nonlinear prediction techniques. In other words, it is the nonlinear prediction techniques that are preferable over linear techniques in such situations. The model using Extreme Gradient Boosting gave the best performance in the prediction of cooling load for the next 24 hours. It can been concluded that using the features of the extraction step can greatly enhance the prediction capability of the models using nonlinear forecasting techniques.

The authors in [19] have proposed inverse blackbox modelling for estimation of heating and cooling load of buildings. The buildings chosen for the study are five office buildings in Ottawa, Canada. The data collected for the study were from five buildings during three years' hourly data of heating and cooling load, hourly electricity consumption data for plug-in equipment, lighting, fans and pumps for the same period for each building, simultaneous weather data such as external temperature, solar irradiance, relative humidity and wind speed recorded at 15 minute intervals at a nearby weather station. Two types of inverse blackbox models were developed with one based on Linear Regression and the other based on ANN. The ANN models were tested for two configurations- *a.* one hidden layer and one output layer and *b.* two hidden layers and one output layer. Sigmoid activation functions were used in hidden layers whereas linear activation functions were used in the output layer. The Levenberg- Marquardt back propagation algorithm was used for parameter estimation. In the Linear Regression based model, the method of least squares was adopted for estimating unknown parameters. In both the models, the inputs were added one at a time to study their effect on the performance of the model. Before training the models, a correlation study was done between the inputs and the heating and cooling loads by calculating the average of the Pearson correlation coefficients for all the five buildings under study. The evaluation criteria used for assessing the performance of the developed models are Root Mean Square Error (RMSE) and R^2 values. The results showed lower RMSE values for ANN based models as compared to the linear regression-based model. This can be attributed to ANN's ability to handle nonlinearities in data better. There was no significant difference in performance of the two ANN models which implies that increasing the complexity of the model will not necessarily lead to better performance. An analysis of performance results of the three models led to the selection of a single layer ANN model with six inputs and one hour input record as the final inverse blackbox model.

Pombeiro and et al. in [20] have done a comparative study of models for predicting the energy usage in an institutional building. The models used are based on linear regression, fuzzy networks and neural networks. The work shows that fuzzy and neural network models have much better prediction accuracy than linear regression models. The name of the building chosen for the case study is IST-T located in Oeiras, Portugal, with a Mediterranean climate, which can be classified as mild in nature. The datasets used for the models were electricity consumption data having energy readings for the month of May, from 01/05/2014 to 31/05/2014, with data being recorded at an interval of every 15 minutes, weather data constituting temperature, relative humidity, wind speed, precipitation, atmospheric pressure and solar radiation, occupancy data and time data pertaining to type of day i.e. weekday or weekend and time of day. The energy usage data was collected every 15 minutes by a smart meter installed in the electricity board of the building which was then converted into kWh. Since the building does not have any sensors, an indirect method using local area connected WiFi users was adopted. The weather data was collected from a weather station located on the roof of the building. The correlation between the input variables was studied, and it was observed that temperature has a high correlation with relative humidity (54.3%) followed by temperature with solar radiation (29.9 %). Another significant correlation was between relative humidity and solar radiation being 29.6%. A correlation of 21.5% was observed between occupancy and time of day. For each of the three models, the datasets were divided into input and output data with a division of 60% training data and 40% test data. The MATLAB Statistics Toolbox was used for developing the linear regression model. The function LinearModel.fit was used for the model. The fuzzy model proposed in the work was formed using the function genfis2 from the MATLAB Fuzzy Logic Toolbox. The genfis2 function applies the Sugeno type FIS with Subtractive Clustering algorithms. A feed-forward back propagation network with one hidden layer using fitnet function from the MATLAB Neural Network Toolbox was used for the neural network model developed. In the fuzzy and neural models, a low number of parameters were varied to train the models, the aim being to develop low complexity models. In the fuzzy model, the radii parameter was varied whereas in the neural model, the number of neurons in the hidden layer was varied. The results show that the linear regression model was not effective enough to justify the load profile. The VAF was 35.9 % and the MAE was 11.8 kWh. The fuzzy model showed considerable improvement with a VAF of 70.3% and MAE of 6.0 kWh. The neural network modelled also showed better performance when compared to the linear regression model with a VAF of 61.6% and

MAE of 8.5 kWh. Results showed that occupancy data when used improves the performance of models. Weather data did not provide any significant improvement in performance of the models, the reason being the chiller which is the main energy consuming load in the building was operated everyday at the same period, regardless of the temperature variation.

Authors A. Jihad and M. Tahiri in [21] have predicted the heating load and cooling load of residential buildings in the climatic zone of Agadir, Morocco, by using ANN as a learning Algorithm. The database of buildings required for the study was generated using Design Builder. The heating load and cooling load requirements for different building scenarios was then simulated with Energy Plus. The methodology used for dataset generation in the study is similar to the work done by Tsanas and Xifara in [13]. The learning algorithm invovled a gradient descent optimization. The ReLU (Rectified Linear Unit) was used as an activation function at the output of each neuron to eliminate all negative input values and increase the convergence speed. The ANN model trains the data of the 5625 buildings generated using Design Builder for prediction of heating and cooling loads. The highest prediction accuracy of 98.6% was reached after 25 million iterations. The validity of the proposed model was tested using data from three residential buildings in Morocco. The heating and cooling loads for the three buildings predicted using the proposed neural network model were compared with the simulated HL and CL data obtained using Design Builder. The proposed model showed an accuracy of 97.6% for the test data and an accuracy of 98.7% for the training data. The ANN model when compared to the model developed using Design Builder showed an improved prediction accuracy of 94.8% and 98.5% for training and test data respectively. It was concluded that the model can easily predict the energy requirements of a new or an existing building without having to calculate thermodynamic balance or use the simulation software. The limitation of the work is the use of only one climatic zone; therefore, the model cannot be used for situations involving other climatic zones.

4. CONCLUSION

In most of the buildings, whether residential or commercial, the important objectives of thermal comfort and reduced energy consumption are not achieved. Recently developed models for predicting energy requirements of buildings are using AI techniques and results show that these strategies have improved the situation. A difficult and time consuming aspect in this area of study is the collection of actual building energy datasets. However, this limitation has been greatly reduced recently with many studies using real data collected from buildings through smart energy meters, sensors and other means. Another limitation observed in most of the research is that the models developed will not be applicable if the climatic zone in which the building is situated changes. Accurate prediction of a building's energy requirements needs to consider building occupancy as an important factor. However, in some of the studies mentioned, the occupancy has been assumed to be constant or considered less significant. It is important to keep in mind that in most buildings the cooling equipment is turned on everyday for a fixed duration irrespective of building occupancy. The building energy manager has an important role to play in this aspect.

Various techniques like ANN, SVM, Chi Squared Automatic Detection Error, linear regression, ensemble models, IRLS, RF etc. have been used for the purpose of prediction of heating and cooling loads of buildings. But results show that models using ANN are able to address the issue of accurate forecasting of building energy requirements more effectively. It can be concluded that ANN is indeed a powerful tool used in the study of energy performance of buildings. The literature study in this area also signifies the usefulness of ANN in the energy analysis studies of buildings. Table 3 summarises the results obtained in each paper, their merits and demerits, and a scope for improvement in future work. This work does not intend to undermine the importance of statistical learning techniques, but based on results, suggests ANN to be a superior method to address nonlinearities present in building energy data and thus be able to provide a better platform for evaluation, estimation and prediction of building energy requirements.

TABLE 3. Results of methods.

Research Reference no.	Outputs	Merits	Demerits	Future improvement
[12]	HL and CL	The input variables chosen are same as in previous literature and helps in comparing results presented in other papers considering the same variables. Benefits of RF in energy studies is demonstrated	Glazing area, although an important variable is not the most correlated with neither HL nor CL	Climate and occupancy are assumed to be constant. Use of simulated data
[13]	HL and CL	Computation time taken by the suggested models is very less	1. Use of default settings in the models, both for the single and ensemble models 2. Models proposed in the paper are only applicable to the twelve building types which are simulated	1. Use of evolutionary computing and swarm intelligence techniques for optimization of parameters in the models (single and ensemble) for achieving 1. 1. Higher accuracy in prediction of HL and CL of the buildings 2. Testing of model using actual datasets
[14]	HL and CL	Accuracy of proposed EMARS model higher than BPNN, RBFNN, CART and SVM	Use of simulation data	Testing of model using actual datasets

(conitnues)

TABLE 3. *(Continued)*

Research Reference no.	Outputs	Merits	Demerits	Future improvement
[15]	CL	Inclusion of zone air temperature for prediction of HL	Effect of occupancy on HL prediction not significant	Application of model to various types of commercial buildings in different climatic regions
[16]	HL and CL	Performance of classical ANN is enhanced by using adaptive ANN	Use of simulated data	Model can be tested using actual datasets
[17]	CL	High variability in data is tackled by division of data into classes and class numbers	Only three institutional buildings considered	Generalize the model for a wider range of institutional buildings
[18]	CL	Computation time is reduced	Computation time increases with feature numbers and is maximum when the RAW feature set is used	Control strategies to regulate operations in response to demand can be developed
[19]	HL and CL	1.Use of actual datasets 2.Weather conditions taken into consideration 3.Useful where no or limited sensor data available for greybox modelling	Only data from five large office buildings used	Can be extended for analysing energy usage of different building typologies
[20]	HL and CL	Significance of occupancy in prediction of HL and CL	1. Other occupancy data like lecture schedules not considered 2. Energy tariff data not included	1. Comparison of the models for a different test dataset 2. Use of higher range of historical data
[21]	HL and CL	Actual data used as test data for the neural network model	Only one climatic region is considered	Can be extended to other climatic regions

ABBREVIATIONS

AI	Artificial Intelligence
ANN	Artificial Neural Network
CART	Classification and Regression Tree
CL	Cooling Load
GRNN	General Regression Neural Networks
HL	Heating Load
HVAC	Heating Ventilation and Air Conditioning
IRLF	Iteratively Reweighted Least Squares
MAE	Mean Absolute Error
MAPE	Mean Absolute Percentage Error
MRE	Mean Relative Error
MSE	Mean Square Error
RBFNN	Radial Basis Function Neural Network
RF	Random Forest
RMSE	Root Mean Square Error
SVM	Support Vector Machine
SVR	Support Vector Regression

5. REFERENCES

[1] Urge-Vorsatz D, Cabeza LF, Serrano S, Barreneche C, Petrichenko K. Heating and cooling energy trends and drivers in buildings. Renew Sust Energy Rev 2015;41:85–98.

[2] Perez-Lombard L, Ortiz J, Pout C. A review on buildings energy consumption information. Energy Build 2008;40:394–8.

[3] Li XW, Wen J. Review of building energy modeling for control and operation, Renew Sustain Energy Rev 2014;37:517–37.

[4] W.G. Cai, Y. Wu, Y. Zhong, H. Ren, China building energy consumption: situation, challenges and corresponding measures, Energy Policy 37 (6) (2009) 2054–2059.

[5] G. Platt, J. Li, R. Li, G. Poulton, G. James, J. Wall, Adaptive HVAC zone modelling for sustainable buildings, Energy and Buildings 42 (2010) 412–421.

[6] R. Yao, B. Li, K. Steemers, Energy policy and standard for built environment in China, Renewable Energy 30 (2005) 1973–1988.

[7] A. Yezioro, B. Dong, F. Leite, An applied artificial intelligence approach towards assessing building performance simulation tools, Energy and Buildings 40 (2008) 612–620.

[8] S. Sholahudin, Hwataik Han, Heating load predictions using the static neural networks method, Int. J. Technol. 6 (2015) 946–953.

[9] Hamid R. Khosravani, María Del Mar Castilla, Manuel Berenguel, Antonio E. Ruano, Pedro M. Ferreira, A comparison of energy consumption prediction models based on neural networks of a bioclimatic building, Energies 9 (2016) 57.

[10] Chengdong Li, Zixiang Ding, Dongbin Zhao, Jianqiang Yi, Guiqing Zhang, Building energy consumption prediction: an extreme deep learning approach, Energies 10 (2017) 1525.
[11] S. Aman, Improving Energy Use Forecast for Campus Micro-grids using Indirect Indicators, in: 2011 IEEE 11th International Conference Data MiningWorkshops (ICDMW), Vancouver, BC, Canada, 2010, pp. 389–397.
[12] A. Tsanas, A. Xifara "Accurate quantitative estimation of energy performance of residential buildings using statistical machine learning tools. Energy and Buildings. Volume 49, 2012, pp. 560–567.
[13] Jui-Sheng Chou, Dac-Khuong Bui, Modeling heating and cooling loads by artificial intelligence for energy-efficient building design, Energy and Buildings 82 (2014) 437–446.
[14] M. Cheng, M. Cao, "Accurately predicting building energy performance using evolutionary multivariate adaptive regression splines," Applied Soft Computing 22 (2014) 178–188.
[15] Dimitrios-Stavros Kapetanakis, Eleni Mangina, El Hassan Ridouane, Konstantinos Kouramas, Donal Finn, COMPARISON OF PREDICTIVE MODELS FOR FORECASTING BUILDING HEATING LOADS, 14th Conference of International Building Performance Simulation Association, Hyderabad, India, Dec. 7–9, 2015.
[16] Yusuf Sonmez; Uğur Guvenc; H. Tolga Kahraman; Cemal Yilmaz, "A comperative study on novel machine learning algorithms for estimation of energy performance of residential buildings," Smart Grid Congress and Fair (ICSG), 2015 3rd International Istanbul.
[17] Chirag Deb, Lee Siew Eang, Junjing Yang, Mattheos Santamouris, Forecasting diurnal cooling energy load for institutional buildings using Artificial Neural Networks, Energy and Buildings 121 (2016) 284–297.
[18] Cheng Fan, Fu Xiao, Yang Zhao, A short-term building cooling load prediction method using deep learning algorithms, Applied Energy 195 (2017) 222–233.
[19] Burak Gunay, Weiming Shen, Guy Newsham, Inverse black box modeling of the heating and cooling load in office buildings, Energy and Buildings 142 (2017) 200–210.
[20] H. Pombeiro, R. Santos, P. Careirra, C. Silva, J. Soussa," Comparative assessment of low-complexity models to predict electricity consumption in an institutional building: Linear regression vs. fuzzy modeling vs. neural networks," Energy and Buildings 146 (2017) 141–151.
[21] Alaoui Sosse Jihad, Mohamed Tahiri, "Forecasting the heating and cooling load of residential buildings by using a learning algorithm "gradient descent," Morocco," Case studies in Thermal Engineering 12 (2018) 85–93.
[22] S. Sholahudin, Hwataik Han, Heating load predictions using the static neural networks method, Int. J. Technol. 6 (2015) 946–953.
[23] Hamid R. Khosravani, María Del Mar Castilla, Manuel Berenguel, Antonio E. Ruano, Pedro M. Ferreira, A comparison of energy consumption prediction models based on neural networks of a bioclimatic building, Energies 9 (2016) 57.
[23] Chengdong Li, Zixiang Ding, Dongbin Zhao, Jianqiang Yi, Guiqing Zhang, Building energy consumption prediction: an extreme deep learning approach, Energies 10 (2017) 1525.
[24] J.W. Taylor, Short-term load forecasting with exponentially weighted methods, IEEE Transactions on Power Systems 27 (1) (2012) 458–464.
[25] H.B. Gunay, W. O'Brien, I. Beausoleil-Morrison, Control-oriented inverse modeling of the thermal characteristics in an office, Science and Technology for the Built Environment (2016) 1–20.
[26] P. Bacher, H. Madsen, Identifying suitable models for the heat dynamics of buildings, Energy and Buildings 43 (7) (2011) 1511–1522.
[27] D. Lazos, A.B. Sproul, M. Kay, Optimisation of energy management in commercial buildings with weather forecasting inputs: a review, Renewable and Sustainable Energy Reviews 39 (2014) 587–603.
[28] C. Martani, D. Lee, P. Robinson, R. Britter, C. Ratti, ENERNET: studying the dynamic relationship between building occupancy and energy consumption, Energy Build. 47 (2012) 584L 591.
[29] I. Berry, J. Conner, A. Braun, I. Berry, T. Roman, L. Adams, J.P. Pasnak, J. Conner, R. Scheck, A. Braun, The Cacti Manual, The Cacti Group, 2012.