Journal of Ambient Intelligence and Humanized Computing
https://doi.org/10.1007/s12652-019-01317-y

ORIGINAL RESEARCH

Forecasting heating and cooling loads of buildings: a comparative performance analysis

Sanjiban Sekhar Roy[1] · **Pijush Samui**[2] · **Ishan Nagtode**[1] · **Hemant Jain**[1] · **Vishal Shivaramakrishnan**[3] · **Behnam Mohammadi-ivatloo**[4,5]

Received: 4 November 2018 / Accepted: 6 May 2019

Abstract

Heating load and cooling load forecasting are crucial for estimating energy consumption and improvement of energy performance during the design phase of buildings. Since the capacity of cooling ventilation and air-conditioning system of the building contributes to the operation cost, it is ideal to develop accurate models for heating and cooling load forecasting of buildings. This paper proposes a machine-learning technique for prediction of heating load and cooling load of residential buildings. The proposed model is deep neural network (DNN), which presents a category of learning algorithms that adopt nonlinear extraction of information in several steps within a hierarchical framework, primarily applied for learning and pattern classification. The output of DNN has been compared with other proposed methods such as gradient boosted machine (GBM), Gaussian process regression (GPR) and minimax probability machine regression (MPMR). To develop DNN model, the energy data set has been divided into training (70%) and testing (30%) sets. The performance of proposed model was benchmarked by statistical performance metrics such as variance accounted for (VAF), relative average absolute error (RAAE), root means absolute error (RMAE), coefficient of determination (R^2), standard deviation ratio (RSR), mean absolute percentage error (MAPE), Nash–Sutcliffe coefficient (NS), root means squared error (RMSE), weighted mean absolute percent error (WMAPE) and mean absolute percentage Error (MAPE). DNN and GPR have produced best-predicted VAF for cooling load and heating load of 99.76% and 99.84% respectively.

Keywords Heating load · Cooling load · Building · Deep neural network · Gradient boosted machine · Gaussian process regression · Minimax probability machine regression

Abbreviations

DNN	Deep neural network
GBM	Gradient boosted machine
GPR	Gaussian process regression
MPMR	Minimax probability machine regression
VAF	Variance accounted for (VAF)
RAAE	Relative average absolute error
RMAE	Root means absolute error
R^2	Coefficient of determination
RSR	Standard deviation ratio
MAPE	Mean absolute percentage error
NS	Nash–Sutcliffe coefficient

✉ Behnam Mohammadi-ivatloo
mohammadi@ieee.org

Sanjiban Sekhar Roy
s.roy@vit.ac.in

Pijush Samui
pijushsamui@gmail.com

Ishan Nagtode
naktodeishan.premanand2015@vit.ac.in

Hemant Jain
hemanham@gmail.com

Vishal Shivaramakrishnan
vishaalshivaram@gmail.com

[1] School of Computer Science and Engineering, Vellore Institute of Technology, Vellore 632014, India

[2] Department of Civil Engineering, NIT Patna, Patna 800005, India

[3] School of Business, University of Connecticut, 2100 Hillside Rd, Storrs, CT 06268, USA

[4] Faculty of Electrical and Computer Engineering, University of Tabriz, Tabriz, Iran

[5] Institute of Research and Development, Duy Tan University, Da Nang 550000, Vietnam

RMSE	Root means squared error
WMAPE	Weighted mean absolute percent error
SVDD	Support vector data description
HVAC	Heating, ventilation, and air conditioning
MARS	Multivariate adaptive regression splines
ELM	Extreme learning machine
AHSRAE	American Society of Heating, Refrigerating and Air-Conditioning Engineers
EBP	Energy performance of buildings
GHG	Greenhouse gas
UNFCCC	United Nations Framework Convention on Climate Change
BMS	Building management system
i.i.d	Independent and identically distributed

1 Introduction

Modern world has seen how developing nations have a ceaseless urge to procure a status of a developed country. In acquiring the same, these countries however forget that sustainable energy and reducing carbon dioxide emissions is paramount (British Petroleum 2013). To meet the demand of the people for necessities like accommodation, it has resorted to make major renovations and replacements in buildings. Recent studies have shown that improper building design and structure, which happens due to hasty planning, contribute to a whopping 40% of the carbon dioxide emissions (Xu et al. 2012). Insufficient time and space constraints to complete the given building project has led to improper building design making India grab a fourth position to emit in amount of CO_2 gas emission. This is detrimental to the people living on earth [GHG data from UNFCCC (United Nations Framework Convention on Climate Change)] (Xu et al. 2012). Thus, it has become even more important that measures be taken to reduce this amount. This paper attempts to mitigate this problem by using out the energy performance of buildings (EBP) by reduction of energy consumption of the buildings, which in recent years has seen a huge amount of interest taken by scholars (Martínez-Molina et al. 2016). An attempt to increase the EPB could ameliorate this gore situation. Thus, forecasting heating and cooling load, estimated from simple features of buildings like surface area, wall area, overall height etc. could help us find out EPB (Kreider et al. 1995). These methods are widely used even in HVAC projects (U.S. Department of Energy 2009). Another important thing to notice is the fact that predicting heating and cooling load could help us reducing the electricity consumption in buildings, which in turn reduces CO_2 emissions. Studies have also found a direct correlation between electricity consumption and CO_2 emissions in Nigeria (Akpan and Akpan 2012), which goes on to show that the hypothesis suggested at first is applicable elsewhere too and not only to Nigeria. Authors (Chung and Rhee 2014) investigated on the opportunities for energy conservation and found out that the building occupancy and its design are the two major factors, which are directly related to EPB. Another study by Gul and Patidar (2015) on energy consumption also confirms the study analysis made by Chung and Rhee (2014) stating that these rules are applicable not only in Korea but also in elsewhere of the world, underlining the statement that building design plays a vital role for energy conservation. Hence, we can conclude that predicting heating load and cooling load will not only help us determine the renovated building structure but also will help us reduce the electricity consumption depending on occupancy pattern and also it will help us making smart buildings (Amiribesheli et al. 2015; Amiribesheli and Bouchachia 2018; Andreu and Angelov 2013). Importance of prediction of heating load and cooling load in buildings is highlighted in Fig. 1.

The need for the prediction of energy consumption in building constructions is a growing interest among researchers such as Yang et al. (2014), Li and Li (2015), Deb et al. (2016), and Malkawi et al. (2016). Moreover, building areas impact the natural resources in a significant manner as stated by Lechtenböhmer and Schüring (2011). Therefore, the buildings offer amenities for human requirements and their uncountable aid, to the society cannot be overlooked. Besides, structures had negative effects to the atmosphere during the last decades as stated by the various authors (Li and Li 2015). The manufacturing process of the contemporary world is based on a foundation of metals, hydrocarbons, and electric power. These have strong connections; each is

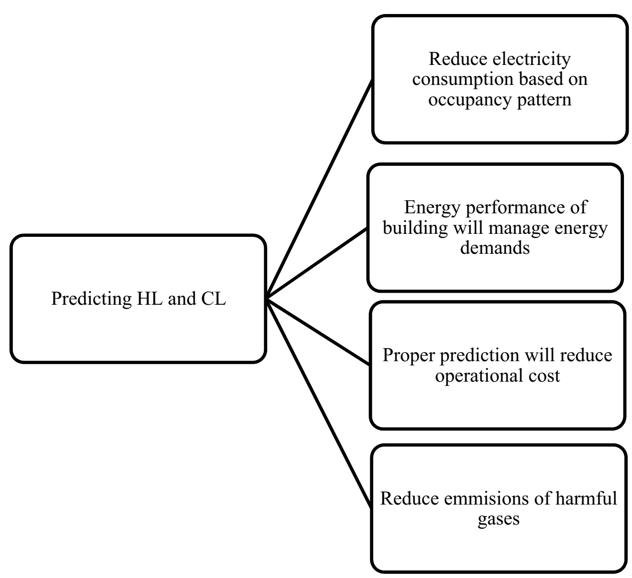

Fig. 1 Importance of prediction of heating load and cooling load in buildings

manageable if only there is enough energy to yield the others. The consumption of energy globally in the year 2013 was about 12,928.4 million tonnes (Energy Consumption by Sector, 2015; European Energy Agency 2015). Global consumption of energy in the year 2008 was about 474 Exajoules (EJ), most of which was supplemented by fossil fuels. Worldwide electricity consumption has also risen by 70% during the period 1990–2008 (Energy Consumption by Sector, 2015; European Energy Agency 2015). The important fact is that buildings constitute roughly 40% of worldwide energy ingesting and thus, have a significant role in the energy arcade with 30% of global CO2 emissions. In addition, Madadnia et al.(2013) and Ahmad et al. (2016) stated that heating and cooling loads are met by the building's HVAC system. Sensors and automation technology are used very often to calculate the heating and cooling load, but more advanced commercial building management system (BMS) sometimes fails to predict the heating and cooling load of buildings with a high degree of accuracy. This forecasting of heating and cooling load and consumption of energy are thus challenging tasks due to the number of interacting factors involved, such as a large variety of appliances and customization of buildings to satisfy the ever increasing demands of the population. The demand for a better approach for load forecasting remains high and more robust forecasting models need to be developed in order to assist engineers and scientists to properly assess sustainability issues during the construction phase of buildings. HVAC regulations have been mentioned in the work of Tsanas and Xifara (2012). Their work has carried out simulations that have generated 12 different building shapes. In HVAC area of research ASHRAE (American Society of Heating, Refrigerating and Air-Conditioning Engineers) is the most famous regulatory body. The fundamental idea of ASHRAE is to advance the arts and sciences of heating, cooling, ventilation, air conditioning, refrigeration and correlated human issues to aid the growing needs of public. HVAC works like an impetus in controlling the atmosphere inside the building. This HVAC framework is in charge of all energy use in a building. Thus, anticipating of heating and cooling load specifically preserves vitality.Researchers have worked a lot on forecasting the heating and cooling loads of buildings. Different types of machine learning algorithms have been employed for prediction of cooling and heating load of buildings, successfully. Deb et al. (2016) have used artificial neural network. ANN is widely used in cases where even the non-linear hypotheses fail to learn. Khayatian et al. (2016) also used the ANN for forecasting the energy performance. Yang et al. (2014) have employed LS-SVM (Least Square-Support Vector Machine) and they have compared SVM with the ANN, which uses a backpropagation algorithm to learn. Yang et al. (2014) have shown that SVM performs better compared to ANN in reducing MAPE and R-squared value. Yu et al. in their work have used hierarchical multi-class SVDD and achieved good precision (Yu et al. 2010). Roy et al. (2018) also have worked on predictive models of multivariate adaptive regression spline, extreme learning machine and a hybrid model of MARS and ELM. Their results were very impressive. The detailed literature survey has been mentioned in Table 1.

This paper aims to provide a solution to the imperative need for energy efficiency by proposing four models that predict the heating load and cooling load of residential buildings. The prediction of heating load and cooling load are beneficial in many ways, for example, it will ensure that an informed decision in regards to sustainability can be made while designing a building. Predictive models for loads can also be used to construct buildings that are more energy efficient. The overall framework of the proposed work is shown in Fig. 2.

In this work, we address the following contributions:

1. People have limited experience of DNN applications for heating load and cooling load predictions of buildings. Single layers give 'shallow' kind of learning. Shallow network requires more layers to become a 'deep' neural network. Therefore, there is a need to understand the perspective of DNN application on heating and cooling load prediction of residential building.
2. In this work, DNN has shown the non-linear manifold performance to the high dimensions of energy efficiency data.
3. In addition, this work, also compares GBM, GPR and MPMR performances with DNN.
4. The aim of this work is also to develop an effective software to estimate energy consumption of buildings during design phase itself.

The remaining part of this paper is organized as follows. Section 2 briefly describes the theoretical background of the proposed models. The details of materials and methods and outcome of the proposed models are described in Sect. 3. Section 4 incorporates the result and discussion of proposed work. In addition, this section also carries out a performance comparison of proposed models. Finally, Sect. 5 presents the conclusion part of this work.

2 Theoretical background

2.1 Deep neural network

This work has checked potential capability of deep neural network (DNN) for prediction of heating and cooling load of buildings depending on various buildings' parameters that have been discussed earlier. The deep learning technique

Table 1 Detailed literature survey of heating and cooling load prediction of buildings

References	Model	Data set	Error index	Conditional attributes	Independent variable
Banihashemi et al. (2017)	ANN and decision tree	Rhino 5	NMSE, RMSE	Wall, insulation, roofing material, windows glazing, floor ground system, Ttype of main space heating, type of main space cooling, building orientation, window to wall ratio ceiling height (m), lighting (Lux) energy load	Heating and cooling energy (kWh)
Samuel et al. (2017)	BIM model	Arboleda case study	NMSE, RMSE, MAE	Climate, building floor area, occupant, equipment, frequency, occupants' behaviour number of heated rooms, number of cooled room	Building energy consumption
Nilashi et al. (2017)	EM clustering, ANFIS	UCI repository	MAE, RMSE, MAPE	Relative compactness, surface area, wall area, roof area, overall height, building orientation, glazing area and glazing distribution	Heating load and cooling load
Naji et al. (2016)	ELM	Energy Plus simulations	RMSE, r, R^2	Building location elevation, window glazing, window frame, HVAC	Energy consumption
Gunay et al. (2017)	One layer artificial neural network	Ottawa, Canada	RMSE, r, R^2	Cooling load, heating load, outdoor temperature, solar irradiance, wind speed, moisture content, electrical load	Heating and cooling load
Fan et al. (2017)	Deep learning	Ottawa, Canada	RMSE, MAE, CV-RMSE	Outdoor temperature, outdoor RH, chilled water supply temperature, chilled water return temperature, chilled water flowrate, cooling load	Cooling load
Zuazua-Ros et al. (2017)	Chilled ceiling element model	Spanish Government	NMSE, RMSE, MAE	Inlet temperatures, fluid flow rate, average heat dissipation power per unit area of the panel	Cooling load
Lindelöf (2017)	Bayesian heating model	ASHRAE handbook	NMSE, RMSE, MAE	The base temperature, building's heat loss coefficient, base load, and daily heating variability	Heating load
Chou and Bui (2014)	SVR and ANN	Ecotect	MAE, RMSE, MAPE	Relative compactness (RC), surface area, wall area, roof area, overall height, orientation, glazing area, glazing distribution, and cooling load/heating load	Heating load and cooling load

Table 1 (continued)

References	Model	Data set	Error index	Conditional attributes	Independent variable
Guo et al. (2017)	ELM	TSCFLP	MAE, RMSE, MAPE	Building's physical characteristics, heating/cooling loads, occupancy transitions	Heating load and cooling load
Sánchez-Oro et al. (2016)	Variable neighbourhood search algorithm	Comunidad Autónoma de Madrid	MAE, RMSE, MAPE	Gross domestic product (GDP), population, exports (in Euros)	Energy demand

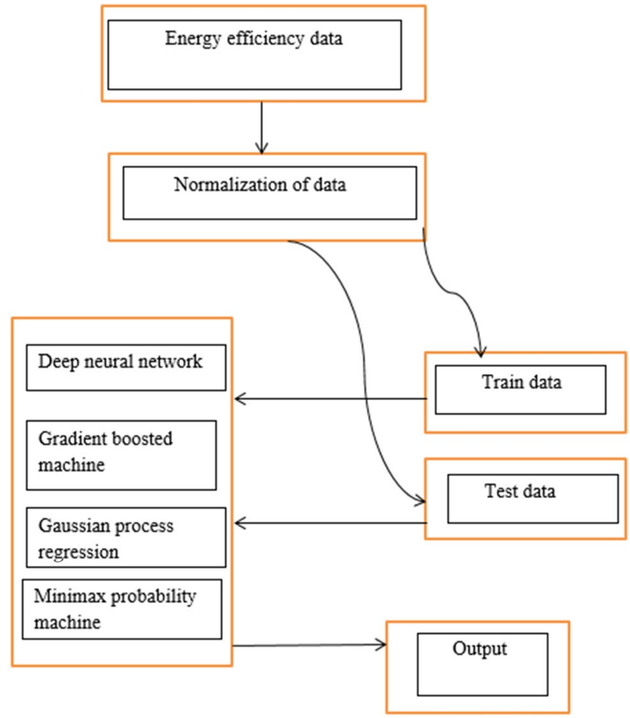

Fig. 2 Overall flow of the proposed work

includes the category of learnings that uses nonlinear information in several steps using hierarchical structures as stated by Deng and Dong (2014). Deep learning primarily is used for learning and pattern classification. Higher level concepts are generated using the lower level features in the hierarchy of deep architecture. Deep learning can be referred to as the intersection of pattern recognition, neural network and graphical modelling. Prediction on massive data set fits well with the deep learning architecture. Artificial neural networks (ANN) frequently are used for deep learning referred by Reg Fellow et al. (2016). Deep neural network has an edge over other machine learning classifiers as stated by Bengio et al. (2013). Proposed deep neural network estimates the intrinsic patterns in the heating and cooling load of buildings. In addition, deep learning logically allows multi-task learning, which considers all features of building shapes using single lyer deep neural network. This network primarily has self-learning units comprising of two or more than two layers. DNN is a feed forward neural network that has multiple hidden layers. In between the input and the output layers, DNN has hidden units. Diagrammatic description of deep neural network is presented as Fig. 3. The hidden unit p is able to adopt a logistic function for mapping below input x_{p} to the scalar state y_{p} of successive layers.

In literature, the deep neural network has captured a lot of research attention recently, since this model is able to extend the distinct advantages compared to its counterparts.

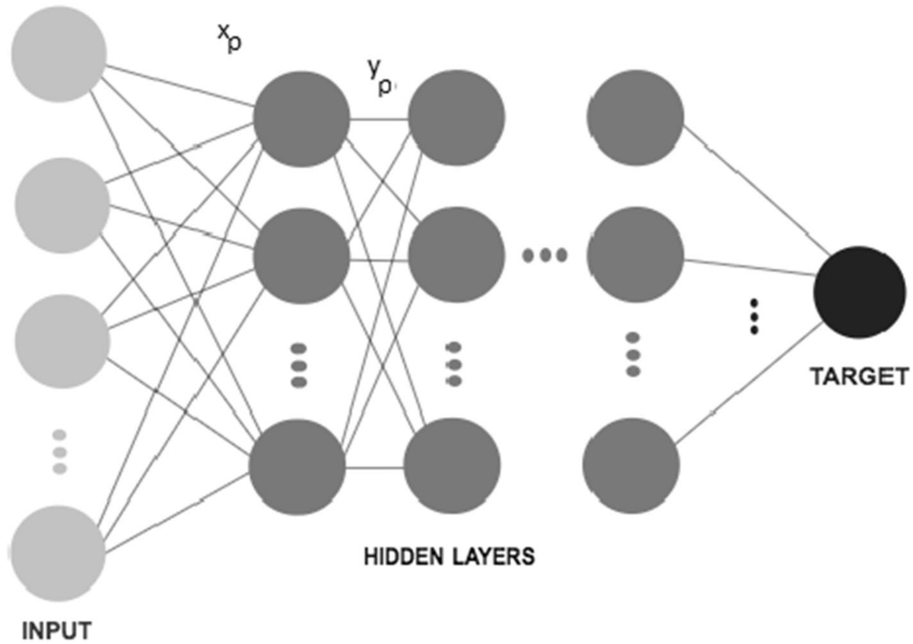

Fig. 3 Diagrammatic description of deep neural network

In DNN network, the output of ith neuron x_i can be determined by (1) and (2):

$$x_i = f(\xi_i), \tag{1}$$

$$f(\xi_i) = \vartheta_i + \sum\nolimits_{j \in \tau_i} -1 W_i X_j, \tag{2}$$

ξ_i is the ith-neuron's potential and $f\left(\xi_i\right)$ is the transfer function. The transfer function can be written as follows:

$$f(\xi_i) = \frac{1}{1 + \exp(-\xi_i)}. \tag{3}$$

The overall objective cost function can be written as sum of squared error:

$$E = \sum 1/2(x_o - \hat{x_o}), \tag{4}$$

where x_o and $\hat{x_o}$ are computed and target values, calculated on output neurons and the gradient decent (GD) algorithm is applied to update the weights incrementally after each epoch. The cost function $J(\cdot)$, the sum of squared errors, $SSE = (J\,(w))$, can be written as:

$$J(w) - \frac{1}{2} \sum_i \left(\text{target}^i - \text{output}^i\right)^2. \tag{5}$$

Equation (5) can also be represented as:

$$\Delta w_j = -\eta \frac{\partial J}{\partial w_j}, \tag{6}$$

while Eq. (6) can be represented as:

$$w : w + \Delta w, \tag{7}$$

where, η is the learning rate, stochastic gradient descent updates the weights after every training data i, and each weight j:

$$\Delta w_j = -\eta \frac{\partial J}{\partial w_j} \left(\text{target}^{(i)} - \text{output}^{(i)}\right)^2 x_j(i). \tag{8}$$

Deep belief network has various qualities, i.e., it's composed of multiple layers with the connection between the layers but not between the neurons of each layer. Eight input variable is used for DNN model. The response variables are y1 and y2, which represents heating load and cooling load, respectively. The cooling load is the output variable.

2.2 Gradient boosted machine

Gradient boosted machine is an ensemble algorithm that combines weak classifiers or weak prediction models. In this work, the weak learners are classification trees (Natekin and Knoll 2013). Predicted results are generated by improving estimations using a differentiable loss function. It fits consecutive trees by taking into account net loss of previous trees. Each tree is partially present in the final solution. Tree boosting uses nonlinear regression algorithm to improve the accuracy of the model. Boosting trees decreases the speed of the algorithm and lacks transparency but has the best accuracy in some cases. The GBM algorithm has a lot of hyper parameters and it is sensitive to noise and extreme values. To avoid GBM overfitting the data a suitable stopping point is required, but it is often the best possible (Natekin and Knoll 2013).

2.3 Introduction to Gaussian Process Regression

Gaussian process is a generalization of multivariate Gaussian distribution to a Gaussian distribution of variable of unbounded dimension. Every subset of dimensions of such a variable follows a multivariate Gaussian distribution. Let's look at what we mean by a multivariate Gaussian distribution (Rasmussen and Williams 2006).

Any random vector $\in \mathbb{R}^n$, with $\mu \in \mathbb{R}^n$ and $\sum \in \mathbb{R}^{n\times n}$ as mean vector and covariance matrix respectively, we have multivariate Gaussian probability distribution for X:

$$P(X;\mu,\Sigma) = \frac{1}{(2\pi)^{n/2}|\Sigma|}\exp\left(-\frac{1}{2}(X-\mu)^T\Sigma^{-1}(X-\mu)\right). \tag{9}$$

Interestingly, it turns out that Gaussian distributions can be used to define a distribution over some mapping *f* of finite set of elements, say $X = \{x_1, x_2 \ldots, x_m\}$ such that:

$$f(x_1) = a_1, f(x_2) = a_2, \ldots f(x_m) = a_m. \tag{10}$$

Now, if we generalize the mapping for an unbounded number of variables and define a certain structure to compute Σ as well as μ and ensure that every finite subset of mappings follow a multivariate Gaussian distribution such that:

$$\begin{pmatrix} f(x_1) \\ f(x_2) \\ \vdots \\ f(x_m) \end{pmatrix} \sim N\left(\begin{pmatrix} m(x_1) \\ m(x_2) \\ \vdots \\ m(x_m) \end{pmatrix}, \begin{bmatrix} k(x_1,x_1) & \cdots & k(x_1,x_m) \\ \vdots & \ddots & \vdots \\ k(x_m,x_1) & \cdots & k(x_m,x_m) \end{bmatrix}\right). \tag{11}$$

We denote this for unbounded number of variables in terms of GP as,

$$f(.) \sim GP(m(.), k(.;.)). \tag{12}$$

The process above is what we call Gaussian Process. To completely specify the GPR we need to define m and Σ explicitly. In our implementation, we have taken $m(x) = 0$ and defined $k(x_1, x_2)$ as:

$$k(x,x') = \exp\left(10^{\Bumpeq}||x-x'||^2\right). \tag{13}$$

This is called as the squared exponential kernel function. It supplies vital information of correlation between two points x and x′. ‖ x − x′ ‖ being the distance between the two points x and x′; β is a parameter which governs the ripples per unit of the functions. In this way, the parameter β governs the smoothness of functions. Therefore, it governs the model.

In the comparison table, we have used a Gaussian process with linear kernel as below:

$$k(x,x') = ||x-x'||. \tag{14}$$

We have arrived at this point by defining the prior distribution of our model that is actually the distribution of functions we were talking about all the way. Now, as per the Bayesian approach we will update our prior probability with each training example to arrive at some posterior distribution of functions. Now, considering we have obtained the posterior distribution of functions, we would predict the value for a new test input X^* *as* Y^* following the Bayesian approach and using already proven results to arrive at the prediction as:

First, we will compute the posterior distribution of the functions h* by using the previous posterior distribution function h as prior:

$$\begin{pmatrix} h \\ h^* \end{pmatrix}\Bigg| X, X^* \sim N\left(\vec{0}, \begin{pmatrix} K(X,X) & K(X,X^*) \\ K(X^*,X) & K(X^*,X^*) \end{pmatrix}\right). \tag{15}$$

We have the posterior distribution from the above. Now we proceed with prediction:

$$\begin{pmatrix} Y \\ Y^* \end{pmatrix}\Bigg| X, X^* = \begin{pmatrix} h \\ h^* \end{pmatrix}, \tag{16}$$

we have:

$$\begin{aligned} &Y^*|Y,X,X^* \sim N(\mu^*,\Sigma^*) \\ &where,\ \mu^* = K(X^*,X)K^{-1}(X,X)Y \\ &and,\ \Sigma^* = K(X^*,X^*) - K(X^*,X)K^{-1}(X,X)K(X,X^*). \end{aligned} \tag{17}$$

The parameter β looks like it is missing and one may ask what is the use of β then. A closer look at the kernel function will answer the question. The parameter is actually governing the formation of distribution of posterior before the test inputs were observed.

2.4 Introduction to minimax probability machine regression

MPMR stands for mini-max probability machine regression and it construction based on 'minimax probability machine' classification by using kernel function.It has got its name from the fact that it frames the regression model in such a way so as to maximize the minimum probability of any prediction, lying in a bounded space around the true function (Strohmann et al. 2004). We assume that there exists an underlying regression function such that:

$$y = f(x) + \epsilon, \tag{18}$$

where, x represents the attribute vector and y represents the target; ϵ represents the noise with $E(\epsilon) = 0$ and $\sigma^2(\epsilon) = 0$.

Considering the underlying principle of MPMR, we now try to come up with an approximation to f. Detailed discussion of MPMR can be found in the work of Strohmann et al. (2004) turns out that the MPM regression model is a kernel based model that iterates the kernel function over each training input vector with some weightage to each training input vector and adds a bias to it to give a prediction. Mathematically, we can write it as:

$$\hat{y} = \hat{f}(x) = \sum_{i=1}^{i=N} \delta_i K(x_i, x) + \alpha. \tag{19}$$

Here, α is the bias and δ_i is the weight for ith training input vector. Moreover, α *and* δ_i are the outputs of MPMR and they provide the complete description of model. The kernel function can be expressed mathematically as:

$$K(x_i, x) = \exp\left\{ -\frac{(x_i - x)(x_i - x)^T}{2\sigma^2} \right\}, \tag{20}$$

where, σ is the width of radial basis function. We have chosen this kernel function as it is far more robust and intuitive compared to other Kernel functions.

3 Materials and methodology

3.1 Data

We have experimented with an energy efficiency dataset that has been created by Tsanas and Xifara (2012). Tsanas and Xifara (2012) have simulated twelve dissimilar building shapes in Ecotect software and have generated this dataset (https://archive.ics.uci.edu/ml/datasets/energy+efficiency#).The dataset comprises of 768 samples and 8 features, aiming to predict two real valued responses. There are eight attributes in the dataset namely $X1, X2, X3$ *up to* $X8$ and two decision variables $y1$ *and* $y2$. According to Tsanas and Xifara (2012), "Simulating building energy aspects is a widely used approach despite the fact that it is impossible to guarantee that the simulation findings will perfectly reflect actual data in the real world (HL and CL)". Nevertheless, the simulated results provide good indication of the likely percentage change and underlying trend of the actual data, enabling energy comparisons of buildings. Even if the data used in this study obtained via the simulations, it might be biased. Moreover, any inconsistency in the simulated data and actual real-world data does not affect the methodology developed in this study. The following are the parameters associated with the attributes and decision variables. Buildings differ in the glazing area, the glazing area distribution, the orientation, and other parameters. With consideration of preliminary cube (3.5×3.5×3.5); onwers of the data set have created twelve buildings forms with each building made up with eighteen cubes of elementary types. Ecotect software was used to create the simulated buildings. The same materials for each eighteen element was used for all buildings forms. The volume of 771.75 m^3 is the same for all buildings with slightly different dominations and surface areas. The lowest U value is used along with the updated common material for selection. The authors have used the building features such as walls (1.78), roofs (0.500), floors (0.860) and windows (2.260). The assumptions were made during simulations that buildings are in Greece, Athens. The clothing was taken as 0.6 clo, 60% humidity, 0.30 m/s as air speed, and with 300 lx as lighting level. The inside gains were fixed to sensible (5) and latent (2 W/m^2), along with the rate of infiltration was set to 0.5 for air change rate with sensitivity of wind as 0.25 air changer per hour. Simulating various settings as functions of the aforementioned characteristics, the data set considers 768 building shapes. The dataset includes 768 samples and each sample contains eight attributes and two responses (y1 and y2) as mentioned earlier. The aim of this modelling is to use the eight features to predict each of the two responses correctly. This work considers y1 as heating load predictor and y2 as cooling load predictor.

3.2 Experimental results obtained by proposed methods

The output of the proposed models DNN, GBM, GPR, and MPMR have been shown in terms of tables. For experimentation purpose, we have adopted R programming and MATLAB programming environment. For DNN and GBM package 'h2o' have been used. The next two adopted algorithms (GPR and MPMR) we have written program in MATLAB environment. The comparative simulations has been accomplished by Weka tool which is a free software that has the collections of many machine learning algorithms. The performances of all developed models have been judged with standard metrics of error measures, such as VAF, RAAE, RMAE, R2, RSR, NMBE, MAPE, NS, RMSE, and WMAPE. The performances of the four different proposed algorithms are shown in Table 2 for heating load and Table 3 for cooling load. DNN really performed average compared to other algorithms. The actual variable is d_t *or* Y^A and the predicted variable is y_t *or* Y. The formulations of these all error measurements have been given as below:

$$VAF = \left(1 - \frac{var(d_t - y_t)}{var d_t}\right) \times 100, \tag{21}$$

$$RAAE = \sum_{1}^{n} \frac{|Y^A - Y|}{n\sigma_Y}, \tag{22}$$

Table 2 Different error indexes of heating load (y1) with different forecasting methods

Y1	DNN	GBM	GPR	MPMR
VAF	98.05	96.54	99.84	88.022
RAAE	0.1632	0.25	0.033	0.167
RMAE	0.893	−0.004	0.153	0.663
R2	0.93	0.92	0.99	0.99
RSR	0.25	0.26	0.042	0.002
NMBE	3.199	−10.40	−0.146	−1.188
MAPE	7.57	11.16	1.442	13.104
NS	0.93	0.92	0.997	0.99
RMSE	2.71	2.81	0.426	0.059
WMAPE	0.0671	0.104	0.014	0.101

Table 3 Different errors on cooling load (y2)

Y2	DNN	GBM	GPR	MPMR
VAF	99.76	98.53	99.13	89.55
RAAE	0.318	0.213	0.063	0.229
RMAE	0.285	0.210	0.324	0.811
R2	0.864	0.94	0.99	0.99
RSR	0.367	0.23	0.169	0.0031
NMBE	−11.04	−7.55	−0.384	−0.408
MAPE	12.33	8.061	2.341	24.18
NS	0.862	0.94	0.99	0.99
RMSE	3.66	2.37	1.67	0.0791
WMAPE	0.115	0.077	0.024	0.157

$$RMAE = \frac{\max\left(Y - Y^{A}\right)}{\sigma_Y}, \tag{23}$$

$$R^2 = \frac{\sum_1^n \left(d_t - d_{mean}\right)^2 - \sum_1^n (d_t - y_t)^2}{\sum_1^n (d_t - y_t)^2}, \tag{24}$$

$$RSR(\%) = \sum_1^n \frac{RMSE}{\sqrt{\frac{\sum_1^n (d_t - \bar{d})^2}{N}}}, \tag{25}$$

$$NMBE = \frac{\frac{1}{n}\sum_i^N (d_t - y_t)}{\frac{1}{n}\sum_i^N d_i} \times 100, \tag{26}$$

$$MAPE(\%) = \frac{1}{n}\sum_1^n \frac{\left|(d_t - y_t)\right|}{d_t} \times 100, \tag{27}$$

$$NS = 1 - \frac{\sum_1^n \left(d_t - y_t\right)}{\sum_1^n \left(d_t - d_{mean}\right)^2}, \tag{28}$$

$$RMSE = \sqrt{\frac{\sum_1^n (d_t - y_t)^2}{n}}, \tag{29}$$

$$MAPE = \frac{\sum_1^n \left|\frac{d_t - y_t}{d_t}\right| \times d_t}{\sum_1^n d_i}. \tag{30}$$

4 Discussion and analysis

From Tables 2 and 3 various errors can be seen, which have been obtained by DNN, GBM, GPR and MPMR for the prediction of heating and cooling load of buildings, respectively. This is evident from Tables 2 and 3, the performance of DNN is poor compared to GPR and MPMR based on the R^2 values. Also, its performance is average when compared to other errors. Table 3 shows that the performance of DNN is the least compared to all the other algorithms while calculating cooling load. Specifically, there is a huge gap in performance between DNN and GBM; GBM in turn performs poorly compared to GPR and MPMR. Fan et al. (2017) paper promotes DNN for the prediction of cooling load by showing the speed of prediction to be fastest compared to other algorithms. They did not document the errors obtained extensively, and not much can be figured out from the only three types of errors they have documented. In this work, we have considered the errors extensively by properly documenting 11 such errors, which are sufficient to give complete idea of DNN's performance. The results obtained after testing the MPMR model has been tabulated in the Table 2 and the r-squared value obtained for heating load on test data 0.99 lies in the range of best prediction models. For heating load prediction, GPR has the highest VAF value of 99.84 and minimal errors in terms of RAAE, WMAPE. GPR and MPMR highest R^2 values i.e. 99%. GBM has the lower RMAE value of −0.004. MPMR has best RSR value and NS value. GBM has NMBE as −10.40. GPR has MAPE value as 1.442. MPMR has obtained best RMSE of 0.059. In case of cooling load prediction, DNN has highest VAF and NMBE value of 99.76 and −11.04 respectively. GBM has best RMAE value of 0.210. GPR has best RAAE of 0.0638, MAPE value of 2.341, NS value of 0.99, and WMAPE value of 0.024. Lastly MPMR has obtained best RSR (0.0031), RMSE (0.0791). Furthermore, errors can be minimized in this case also if we consider i.i.d noise variables. Since the

Table 4 Comparison with other methods for heating load

	Gaussian process (linear Kernel)	Linear regression	ANN	RBF network	SVM
R^2	0.901	0.92449	0.9883	0.8043	0.9137
Correlation Coeff.	0.9664	0.9599	0.9945	0.8838	0.9675
Mean absolute error	2.5005	0.1959	0.0851	0.3535	2.2792
Root mean squared error	3.1737	0.2788	0.1056	0.4585	2.9581
Relative absolute error	25.7917%	21.4315%	9.3057%	40.2363%	23.5084%
Root relative squared error	28.3969%	27.851%	10.55%	46.7624%	26.4673%

noise is not symmetric and also the model is MPMR, it will be difficult to integrate noise variables with MPMR model. In the previous case (Gaussian process model); the assumption of i.i.d. noise variables will not make the theory complicated, which is easy to understand as both of them are derived from Gaussian distribution. One can show that the assumption of i.i.d. noise variables with $N(0, \sigma^2)$ will add an extra term $\sigma^2 I$ where I is an identity matrix of appropriate dimension with $K^{-1}(X, X')$ in μ^* and $\sum^*$ from Eq. (17). We have also carried out comparative studies for other machine learning algorithms viz. Gaussian regression (linear Kernel), linear regression, ANN, RBF network and SVM based on the same data set. The results have been tabulated in Tables 4 and 5. Table 4 tabulates the results obtained for prediction of heating load, whereas, Table 5 tabulates the results obtained for prediction of cooling load. From Table 4, we can observe that ANN outperforms every other algorithm with R^2 equal to 0.988. While, RBF network gives a really poor performance with R^2 equal to 0.804. Linear regression being a very traditional algorithm has still managed to outperform complex algorithms like SVM and Gaussian process (linear Kernel). ANN has also performed very well in optimizing root relative squared error reducing it to only 10% for heating load (Table 4). The predictions obtained by ANN are highly correlated with the actual values, as the correlation coefficient is 99.5%. While, SVM could not perform very well on R^2, the predicted values are highly correlated with actual values when compared to GR, linear regression and RBF. From Table 5, we can observe that linear regression has outperformed all the other algorithms in case of cooling load prediction. We would have expected ANN to outperform the other algorithms like it did for heating load but that is not the case. All the remaining algorithms have produced same results as for heating load with a decrease in R^2 of 0.02–0.03. When it comes to correlation, we can see that ANN has predicted the most highly correlated values with actual data points.

5 Conclusions

In recent times, energy consumptions of buildings represents a considerable percentage of total energy used and that directly contributes to the environmental pollution. In this work, the four proposed machine learning models yielded very encouraging overall performance. The performances of DNN has been compared with GPR, GBM and MPMR. In terms of overall performance GPR and MPMR have surpassed other two methods. The limitations these models face at the moment is the lack of more real time data. Nevertheless, this research work has focussed on buildings, the same approach can also be adopted to various other energy amenities, and to the time series prediction models based on the purpose of usage and obtainability of related data such as energy tariff mechanism, weather data and metrics of energy conservation measures. Finally, it will be interesting to observe that how heating and cooling load predictions of buildings can help constructing smart buildings. In future

Table 5 Comparison with other methods for cooling load

	Gaussian process (linear Kernel)	Linear regression	ANN	RBF network	SVM
R^2	0.88	0.9032	0.8925	0.7884	0.8843
Correlation Coeff.	0.9493	0.948	0.9529	0.8935	0.9461
Mean absolute error	2.3889	2.1469	2.3077	3.1339	2.37
Root mean squared error	3.2811	2.959	3.1184	4.3648	3.2355
Relative absolute error	26.451%	25.342%	25.5522%	35.404%	26.3466%
Root relative squared error	32.0971%	31.718%	30.5057%	44.6714%	31.6504%

work, for heating and cooling load predictions of buildings authors can also explore evolutionary algorithms. The computational time taken by these proposed methods can also be explored further.

References

Ahmad MW, Mourshed M, Yuce B, Rezgui Y (2016) Computational intelligence techniques for HVAC systems: a review. Build Simul 9:359–398

Akpan GE, Akpan UF (2012) Electricity consumption, carbon emissions and economic growth in Nigeria. Int J Energy Econ Policy 2:292–306

Amiribesheli M, Bouchachia H (2018) A tailored smart home for dementia care. J Ambient Intell Hum Comput 9(6):1755–1782

Amiribesheli M, Benmansour A, Bouchachia A (2015) A review of smart homes in healthcare. J Ambient Intell Hum Comput 6(4):495–517

Andreu J, Angelov P (2013) An evolving machine learning method for human activity recognition systems. J Ambient Intell Hum Comput 4(2):195–206

Banihashemi S, Ding G, Wang J (2017) Developing a hybrid model of prediction and classification algorithms for building energy consumption. Energy Procedia 110:371–376

Bengio Y, Courville A, Vincent P (2013) Representation learning: a review and new perspectives. IEEE Trans Pattern Anal Mach Intell 35:1798–1828. https://doi.org/10.1109/TPAMI.2013.50

British Petroleum (2013) BP statistical review of world energy, June 2014. Br Pet www.bp.com/statisticalreview. https://doi.org/10.1016/j.egypro.2013.06.172. Accessed June 2018

Chou JS, Bui DK (2014) Modeling heating and cooling loads by artificial intelligence for energy-efficient building design. Energy Build 82:437–446. https://doi.org/10.1016/j.enbuild.2014.07.036

Chung MH, Rhee EK (2014) Potential opportunities for energy conservation in existing buildings on university campus: a field survey in Korea. Energy Build 78:176–182. https://doi.org/10.1016/j.enbuild.2014.04.018

Deb C, Eang LS, Yang J, Santamouris M (2016) Forecasting diurnal cooling energy load for institutional buildings using artificial neural networks. Energy Build 121:284–297. https://doi.org/10.1016/j.enbuild.2015.12.050

Deng L, Dong Y (2014) Foundations and trends in signal processing. Signal Proces 7:3–4

European Energy Agency (2015) Final energy consumption by sector and fuel. Indic Assess | Data maps 20. CSI 027/ENER 016

Fan C, Xiao F, Zhao Y (2017) A short-term building cooling load prediction method using deep learning algorithms. Appl Energy 195:222–233. https://doi.org/10.1016/j.apenergy.2017.03.064

Goodfellow I, Bengio Y, Courville A (2016) Deep learning. MIT press

Gul MS, Patidar S (2015) Understanding the energy consumption and occupancy of a multi-purpose academic building. Energy Build 87:155–165. https://doi.org/10.1016/j.enbuild.2014.11.027

Gunay B, Shen W, Newsham G (2017) Inverse blackbox modeling of the heating and cooling load in office buildings. Energy Build 142:200–210. https://doi.org/10.1016/j.enbuild.2017.02.064

Guo P, Cheng W, Wang Y (2017) Hybrid evolutionary algorithm with extreme machine learning fitness function evaluation for two-stage capacitated facility location problems. Expert Syst Appl 71:57–68. https://doi.org/10.1016/j.eswa.2016.11.025

Khayatian F, Sarto L, Dall'O' G (2016) Application of neural networks for evaluating energy performance certificates of residential buildings. Energy Build 125:45–54. https://doi.org/10.1016/j.enbuild.2016.04.067

Kreider JF, Claridge DE, Curtiss P et al (1995) Building energy use prediction and system identification using recurrent neural networks. J Sol Energy Eng 117:161. https://doi.org/10.1115/1.2847757

Lechtenböhmer S, Schüring A (2011) The potential for large-scale savings from insulating residential buildings in the EU. Energy Effic 4:257–270. https://doi.org/10.1007/s12053-010-9090-6

Li Y, Li X (2015) Natural ventilation potential of high-rise residential buildings in northern China using coupling thermal and airflow simulations. Build Simul 8:51–64. https://doi.org/10.1007/s12273-014-0188-1

Lindelöf D (2017) Bayesian estimation of a building's base temperature for the calculation of heating degree-days. Energy Build 134:154–161. https://doi.org/10.1016/j.enbuild.2016.10.038

Madadnia J, Vakiloroaya V, Samali B (2013) Modelling and performance prediction of an integrated central cooling plant for HVAC energy efficiency improvement. Build Simul 6:127–138. https://doi.org/10.1007/s12273-013-0104-0

Malkawi A, Yan B, Chen Y, Tong Z (2016) Predicting thermal and energy performance of mixed-mode ventilation using an integrated simulation approach. Build Simul 9:335–346. https://doi.org/10.1007/s12273-016-0271-x

Martínez-Molina A, Tort-Ausina I, Cho S, Vivancos JL (2016) Energy efficiency and thermal comfort in historic buildings: a review. Renew Sustain Energy Rev 61:70–85. https://doi.org/10.1016/j.rser.2016.03.018

Naji S, Keivani A, Shamshirband S et al (2016) Estimating building energy consumption using extreme learning machine method. Energy 97:506–516. https://doi.org/10.1016/j.energy.2015.11.037

Natekin A, Knoll A (2013) Gradient boosting machines, a tutorial. Front Neurorobot. https://doi.org/10.3389/fnbot.2013.00021

Nilashi M, Dalvi-Esfahani M, Ibrahim O et al (2017) A soft computing method for the prediction of energy performance of residential buildings. Measurement. https://doi.org/10.1016/j.measurement.2017.05.048

Rasmussen CE, Williams CK (2006) Gaussian processes for machine learning, vol 2. MIT Press, Cambridge, MA, p 4

Roy SS, Roy R, Balas VE (2018) Estimating heating load in buildings using multivariate adaptive regression splines, extreme learning machine, a hybrid model of MARS and ELM. Renew Sustain Energy Rev 82:4256–4268

Samuel EI, Joseph-Akwara E, Richard A (2017) Assessment of energy utilization and leakages in buildings with building information model energy. Front Archit Res 6:29–41. https://doi.org/10.1016/j.foar.2017.01.002

Sánchez-Oro J, Duarte A, Salcedo-Sanz S (2016) Robust total energy demand estimation with a hybrid variable neighborhood search—extreme learning machine algorithm. Energy Convers Manag 123:445–452. https://doi.org/10.1016/j.enconman.2016.06.050

Strohmann TR, Belitski A, Grudic GZ, DeCoste D (2004) Sparse greedy minimax probability machine classification. Adv Neural Inf Process Syst 16:105–112

Tsanas A, Xifara A (2012) Accurate quantitative estimation of energy performance of residential buildings using statistical machine learning tools. Energy Build 49:560–567. https://doi.org/10.1016/j.enbuild.2012.03.003

U.S. Department of Energy (2009) Buildings energy data book. http://buildingsdatabook.eren.doe.gov/, pp 1–271. Accessed June 2018

Xu X, Taylor JE, Pisello AL, Culligan PJ (2012) The impact of place-based affiliation networks on energy conservation: an holistic model that integrates the influence of buildings, residents and the neighborhood context. Energy Build 55:637–646. https://doi.org/10.1016/j.enbuild.2012.09.013

Yang L, Yan H, Lam JC (2014) Thermal comfort and building energy consumption implications—a review. Appl Energy 115:164–173
Yu J, Lee H, Im Y et al (2010) Real-time classification of internet application traffic using a hierarchical multi-class SVM. KSII Trans Internet Inf Syst 4:859–876. https://doi.org/10.3837/tiis.2010.10.009
Zuazua-Ros A, Martín Gómez C, Ramos JC, Bermejo-Busto J (2017) Towards cooling systems integration in buildings: experimental analysis of a heat dissipation panel. Renew Sustain Energy Rev 72:73–82