# Computational Approach for Cooling Load Forecasting of Air-Conditioning System: A Survey and Open Challenges

Fatima Shittu[1], Mustapha Lawal Abdulrahman[1], Aminu Saleh[2] and Shehu Muktar[1]

[1]Department of Computer Science, Federal Polytechnic, Damaturu

[2]Department of Mathematical Science, Abubakar Tarawa Balewa University Bauchi.
Correspondence e-mail: fsbinta1234@gmail.com

**Abstract**

*Due to the high energy consumption in buildings, cooling load forecasting plays a crucial role in the planning, control and operation of heating, ventilating and air-conditioning systems. The building sector is a major consumer of energy worldwide and a large amount of energy is used for Heating, Ventilation and Air Conditioning (HVAC). One cause of high consumption in HVAC systems lies in their frequent failure to operate as intended after a period of operation, even with correct commissioning. To curtail the gap between demand and supply, new paradigms have to be employed that will use automated methods to dynamically forecast the buildings energy consumption. Many influencing factors make the prediction complex including: occupancy behaviors, weather conditions, parameter and model selection, missing data or incomplete information, and computation complexity of some models. Neural networks are recently applied for solving load forecasting due to their wide application. Deep neural network models provide a practical approach to energy consumption prediction. This paper offers a review of deep neural network for building energy consumption prediction that utilize machine learning algorithm. Based on this Review, existing research gaps are identified and future research directions are highlighted in the context of building energy consumption prediction.*

***Keywords:*** **Load Forecasting, Air-Conditioning, Cooling Load, Energy Consumption**

## I. Introduction

Buildings represent a large portion of the world's energy consumption and associated $CO_2$ emissions. Prediction of building energy consumption is crucial for improved decision making towards reducing energy consumption and $CO_2$ emissions (Amasyali & El-Gohary, 2018). Due to the high energy consumption in buildings, cooling load forecasting plays a crucial role in the planning, control and operation of heating, ventilating and air-conditioning systems. The building sector is a major consumer of energy worldwide and a large amount of energy is used for Heating, Ventilation and Air Conditioning (HVAC). One cause of high consumption in HVAC systems lies in their frequent failure to operate as intended after a period of operation, even with correct commissioning. In many buildings, energy performance is not a concern as long as building comfort can be maintained(Ji, Xu, Duan, & Lu, 2016).

Generally, accurate forecasting of cooling load is significantly important for the planning and design of HVAC systems. In addition, it also can be used to enhance the operating performance of HVAC systems, such as adjusting the starting time of cooling equipment to satisfy start-up demand, minimizing the electric on-peak load and improving the energy utilization efficiency of cooling storage systems. However, the cooling load data series always exhibit nonlinear and dynamic features because of the chaotic nature of the global weather system and life behaviors of human beings, making it very difficult to be forecasted accurately and reliably (Ji et al., 2016).

Traditionally, four types of methodologies have been developed in the literature, including persistence, physical modeling, statistical approaches and soft-computing based methods, to fulfill cooling load forecasting task of HVAC. The persistence method is the simplest way of producing a forecast. It assumes that the predicted value at next time step is equal to the latest measured value because of the high autocorrelation in cooling load time series data. Therefore, this method is generally implemented as a benchmark for comparison to baseline the performance of newly developed forecasting approaches.

Physical modeling methods aim to establish a mathematical model for cooling load forecasting (CLF) by using environmental information, such as building envelopes and wall materials, and meteorological data, such as outdoor dry-bulb temperature, relative humidity and solar irradiation. The environmental and meteorological information are used to characterize actual building thermal behaviors (Fu, 2018). The existing simulation tools based on physical modeling methods include Energy Plus, TRNSYS and ESP-r. However, physical modeling methods generally require intensive computation resources and timely accurate response and thus real time short-term forecasting cannot be guaranteed {Lim, 2017}. In other words, this type of methods is less efficient for real-time implementation. In addition, it is very difficult to collect some of the environmental and meteorological parameters because the measurement of each data comes at a price.

Therefore, physical modeling methods are not suitable for the management of real energy systems. Physical approach mainly depends on the physical principles and building details and its properties to characterized building behaviors, thus refers to as the white box models. The performance of this model is subject to fluctuation because it requires large amount of information thus becoming computationally expensive if the assumption of this information is not completely fulfill since the model capture building response to influential factors including the outdoors and indoors environment (H.-x. Zhao & Magoulès, 2012), But a major drawback of this deterministic model is not being able to account for complex consumption behaviors and difficulty in obtaining the input parameters required, thus making poor predictions sometime up to 100% lost in accuracy(H.-x. Zhao & Magoulès, 2012).

Although an alternative to such physics models which is statistical and machine learning models and have been successfully applied in this context such as linear regression, support vector machine Extreme learning machine, stochastic model (CRBM and FCRBM), Long Short-Term Memory, Recurrent Neural Network, Elman Neural Network, Deep Recurrent Neural Network and Ensemble Technique. Statistical approaches try to develop a relationship between historical data samples and future cooling load though regression analysis. Up to date, multiple linear regression (MLR), nonlinear autoregressive exogenous (NARX) models and autoregressive integrated moving average (ARIMA) models were commonly implemented for CLF (Fu, 2018).

This paper, offers a review of deep neural network and mainly focuses on RNNs for building energy consumption prediction that utilize machine learning algorithm including Artificial neural network, recurrent neural network, Elman Neural Network and Deep recurrent neural network. The remainder of this paper is divided into sub-sections. Section two (2) will discuss the advantage and motivation of the study, section three (3) will provide an overview of some of the deep learning prediction method discussed in the literature, section four (4) will provide the application of deep learning in the context of energy prediction studies and finally, section five (5) will discussed the review and future research gap/direction will be recommended

## II. MOTIVATION

The estimation of the total building energy usage can help architects and engineers to conceive more clearly the building energy efficiency level. This can be carried out early in the design and during the construction of a building resulting in more realistic and accurate energy efficient building design (Naji et al 2015). Also, accurate load forecast can reduce high cost of over- aid under – contracts on balancing markets due to load prediction errors. Moreover, it keeps power market efficient and provide a better understanding of the dynamics of the monitored system on the other hand, a wrong prediction cause either a load overestimation, which leads to the excess supply and consequently more cost and contract curtailments for market participant or a load underestimation resulting in failures in gathering enough provision thereby more costly supplementation services. Base on the prevailing needs in terms of utilities, there is need to adjust electricity generation with respect to the corresponding consumption which presently is a formidable task.

## III. LITERATURE REVIEW

### 2.1 Computational Techniques

The review of fundamental concept pertaining to this proposed researched work is presented in the following sub sections. A brief introduction on computing approaches used in the field of cooling load forecasting as discussed in the literature.

### 2.1.1 Artificial Neural Network

ANN has been applied in different areas because of their human intelligence performance e.g. speech recognition, robotic control, image processing, data mining, and energy forecasting (Singh & Dwivedi, 2018). Trial and error is the most widely method of determining the number of hidden layer to train the network, training algorithm such as Back Propagation is used. The error in the network weight is minimized during the training phase, when one of the following occur the training is terminated i.e. the gradient performance is below a threshold or the maximum iteration is at its peak or the required value is met for the error that is minimized (Cheng & Li, 2008). Despite the computational power of ANN, it comes with a number at draw backs such as model over fitting, random weight optimization sensitivity and probability of local optima convergence (Tieleman, 2008)**.**

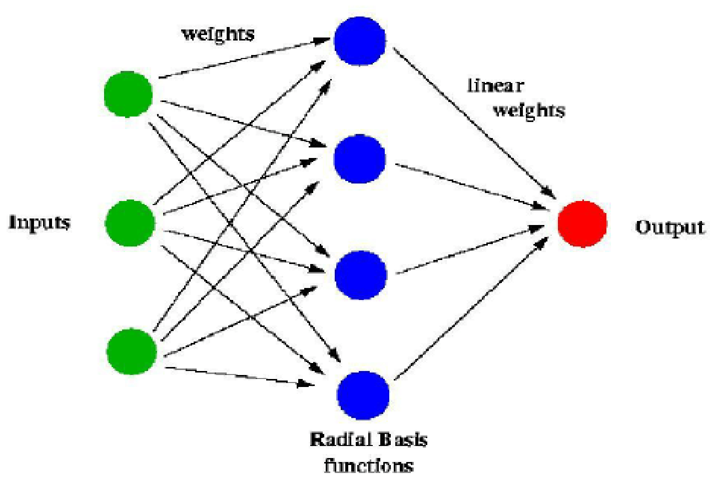


**Figure 1: Block Diagram of Artificial Neural Network {Laxmi, 2014}**

### 2.2.2 Deep Belief Networks

A deep belief network (DBN) for forecasting can be seen as a two-step algorithm:

i. The DBN performs a feature learning to reduce the dimensionality of the input data set.

ii. An additional layer, e.g., a linear layer, is added to carry out the forecasting.

In more detail, each layer of a DBN consists of a Restricted Boltzmann Machine (RBM). An RBM is a two-layer stochastic ANN which learns a probability distribution over the input data set. The layers are organized like a funnel, which helps to learn features within the data. To train a DBN for regression, two training steps have to be performed: In the first training step, the DBN is trained in an unsupervised manner by contrastive divergence. This results in a topology which can abstract the data to a reduced set of features. The second training step starts with appending an ANN, e.g., one layer of fully connected neurons, to the pre-trained topology. The newly attached layers have to be trained with the desired targets to perform a prediction. It is possible to train only the newly attached layer or to train the whole DBN again(Gensler, Henze, Sick, & Raabe, 2016).

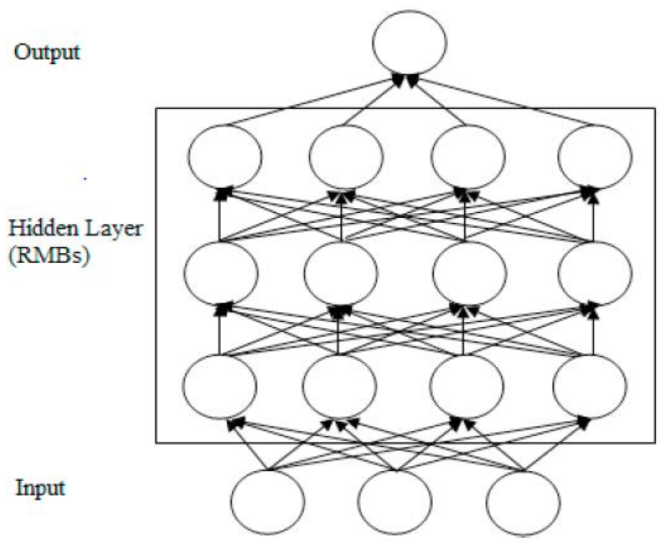


**Figure 2: DBN Structure(Chen, He, & Tso, 2017)**

### 2.2.3 Deep Recurrent Neural Network (GRU and LSTM)

A recurrent neural network can be extended to a deep recurrent neural network via different ways as shown by (Pascanu, Gulcehre, Cho, & Bengio, 2013), the concept of depth in RNN was initially argued, but base on the architecture of RNN, three areas can be make deeper in other to construct a DRNN, this includes; hidden to hidden function, hidden to hidden transition and hidden to output function (Pascanu et al 2014). There is different variant of DRNN but we will focus on the popular DRNN use in the context of electricity forecasting including LSTM and GRU. The idea of $RNN_S$ with gated units was first introduced by Hochreiter and schmidhuber. LSTM is suitable for modeling long range dependencies, LSTM's architecture has memory blocks Compared to RNN which has hidden units (Chniti, Bakir, & Zaher, 2017).

### 2.2.4 Nonlinear Autoregressive Neural Network (NARX)

In the literature, a nonlinear autoregressive network with exogenous inputs (NARX) is a part of discrete-time nonlinear systems. This hybrid design involves the genetic algorithm (GA)-based optimization technique in the optimal brain strategy by determining the optimal networks and involving the external inputs. This sophisticated architecture made it more effective than traditional regression models such as AR, MA, ARIMA. In the energy management field, this algorithm is often used owing to its great abilities in time series dependencies analysis for prediction purposes. The equation of the NARX is defined as follows in equation (5). Figure (1) presents the configuration of this model. $y(n+1)=f[y(n),...,y(n-d_y+1);u(n),u(n-1),...,u(n\ d-+_u 1)$ (5)

Where:

$\boldsymbol{u}(\boldsymbol{n})$ Input of the model at discrete time step n

$\boldsymbol{y}(\boldsymbol{n})$ model output at discrete time step n

$\boldsymbol{d_u} \geq 1$ Input memory order

$\boldsymbol{d_y} \geq 1$ Output memory order

In the standard NARX network, we have a two-layer feed forward network, with a sigmoid transfer function in the hidden layer and a linear transfer function in the output layer. This network has a specific feature by involving a tapped delay lines to store previous values of the x(t) and y(t) sequences. Note that the output of the NARX network, y(t), is fed back to the input of the network (through delays).

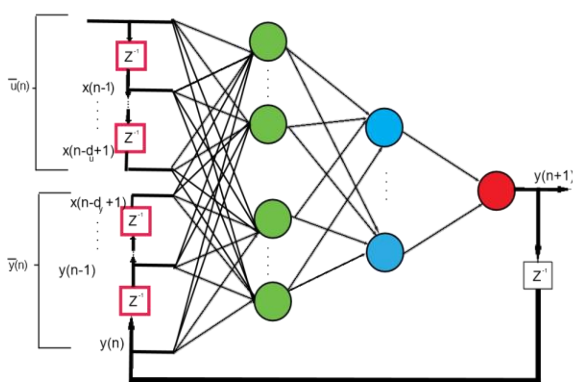


**Figure. 3. NARX recurrent neural network architecture**

## IV. APPLICATION OF COMPUTATIONAL TECHNIQUES FOR COOLING LOAD FORECASTING

This section gives the review of literatures that utilized machine learning techniques for the estimation of energy consumption in both residential and commercial buildings.

There have been limited amount of work with regard to medium to long term forecasting either sub-hourly or hourly intervals with long term prediction being more difficult and complex task to achieve, a relative error often in excess of 40-50 % is associated with medium to long term forecasting (Mocanu, Nguyen, Gibescu, & Kling, 2016). Potential improvement on prediction accuracy of the aforementioned machine learning can be obtained using deep neural network where modeling of more complex functioning is allowed by the use of multiple layers of abstraction. These approaches are being applied recently in the context of energy prediction for example (Fu, 2018) employ deep learning approach for the determination of cooling load prediction which outperforms SVM and BPNN in terms of accuracy and stability.

Similarly, (Fan, Xiao, & Zhao, 2017) Propose a deep learning for building profile load predictions, an enhance performance was obtained especially when the deep neural network is use in an unsupervised manner. It was also shown in the studies of (Mocanu et al., 2016) where the forecasting accuracy was greatly improved by the integration of DBN into a reinforcement algorithm that can target features of the buildings.

Equally important, (Fan et al., 2014) proposed an ensemble model for the prediction of energy consumption load, the proposed model obtain higher accuracy than individual model but the proposed model depends highly on base model selection.

(Guo, Nazarian, Ko, & Rajurkar, 2014) presents a robust hourly cooling-load forecasting method based on time-indexed autoregressive with exogenous inputs (ARX) models and compare it to ANN and ARX, the numerical case studies show the proposed prediction method performs better than some ANN and ARX forecasting models. The work in (Qiang et al., 2015) which Presents cooling load prediction of heating, ventilating and air-conditioning (HVAC) systems in office buildings using PCA of meteorological factors, cumulative effect of high temperature (CEHT) and dynamic two-step correction and The predicted load of the proposed model has acceptable agreement with actual load, where the mean absolute relative error is less than 8%. Although this result has relatively higher error compare to other predictors.

In addition, (Ji et al., 2016) proposed an hourly cooling load prediction model, called the ‘‘RC-S” model optimized with GA and compare it against Existing RCS Model. The proposed ‘‘RC-S” cooling load calculation method is more accurate than the existing RC model and much simpler than whole building simulation models. The study can further be expanded into a prototype toolkit for hourly building cooling load calculation.

(Fan et al., 2017) using deep learning algorithms for short-term building cooling load prediction and compared it with MLR, ELN, RF, GBM, SVR, XGB and DNN. The findings are enlightening and could bring more flexible and effective solutions for building energy predictions.

(Ding, Zhang, Yuan, & Yang, 2018) Proposed short-term and ultra-short-term cooling load prediction models for office buildings and evaluate it against GA-SVR and GA-WD-SVR . The prediction results indicate that the GA-SVR prediction model performs better for short-term cooling load prediction with MRE and R2 of 6.5% and 73.1%, respectively, while the GA-WD-SVR prediction model performs better for ultra-short-term cooling load prediction with MRE and R2 of 4.6% and 88.7%, respectively.

(Ding et al., 2018) analyses the effect of various input variables on prediction accuracy by evaluating it with ANN and SVM. It is concluded that the prediction models with optimized input combinations perform better than those without optimization.

The study in (Fu, 2018) proposed Deep Belief Network based Ensemble Approach for Cooling Load Forecasting of Air-conditioning System and compare it to persistence, BPNN and SVM algorithms. The numerical results demonstrate that the proposed forecasting framework not only exhibits high stability and robustness, but also performs better than the benchmarks compared with. One of the major the limitations of this research is that the DBN based forecasting framework has been implemented for short-term cooling load forecasting problems in real buildings and other exogenous variables, such as solar irradiation, work schedules and wind speeds were not considered in the work.

## V. DISCUSSIONS AND OPEN OPPORTUNITIES

The existing study in (Fu, 2018) proposed Deep Belief Network based Ensemble Approach for Cooling Load Forecasting of Air-conditioning System and compare it to persistence, BPNN and SVM algorithms. The numerical results demonstrate that the proposed forecasting framework not only exhibits high stability and robustness, but also performs better than the benchmarks compared with. One of the major the limitations of this research is that the DBN based forecasting framework has been implemented for short-term cooling load forecasting problems in real buildings and other exogenous variables, such as solar irradiation, work schedules and wind speeds were not considered in the work. (Hashmi, Arora, & Priolkar, 2015), (Buitrago & Asfour, 2017) and (Oprea et al., 2018) opined that NARX technique outperformed FFNN for forecasting multi step ahead electric load most especially dynamic time series system achieving a 30% improvement in forecasting performance. Prediction accuracy and reliability need to be improved upon by exploring the various AI techniques and comparing them against the best algorithms.

Based on our review as presented in table 1, we found that most of the work focuses on forecasting short term, There have been limited amount of work with regard to medium to long term forecasting either sub-hourly or hourly intervals with long term prediction being more difficult and complex task to achieve, a relative error often in excess of 40-50 % is associated with medium to long term forecasting (Mocanu et al., 2016) Also, Ensemble and hybrid technique which have shown to provide reliable and better results still surfers due to their implementation and computational complexity (Wang & Srinivasan, 2017), more research is needed to bridge this gap with simpler but effective models.

The importance of changes in building operation and other exogenous variables, such as solar irradiation, work schedules and wind speeds which are highly uncertain thus affecting the prediction accuracy and reliability (Mocanu et al., 2016) need to be looked into in order to develop a generalized model based on machine learning algorithm

## VI. CONCLUSION

The building sector is a major consumer of energy worldwide and a large amount of energy is used for Heating, Due to the high energy consumption in buildings, cooling load forecasting plays a crucial role in the planning, control and operation of heating, ventilating and air-conditioning systems. Ventilation and Air Conditioning (HVAC). One cause of high consumption in HVAC systems lies in their frequent failure to operate as intended after a period of operation, even with correct commissioning. To curtail the gap between demand and supply, new paradigms have to be employed that will use automated methods to dynamically forecast the buildings energy consumption. This paper offers a review of deep neural network for building energy consumption prediction that utilize machine learning algorithm. Based on this Review, existing research gaps are identified and future research directions are highlighted in the context of building energy consumption prediction.

**Reference**

Amasyali, K., & El-Gohary, N. M. (2018). A review of data-driven building energy consumption prediction studies. *Renewable and Sustainable Energy Reviews, 81*, 1192-1205.

Buitrago, J., & Asfour, S. (2017). Short-term forecasting of electric loads using nonlinear autoregressive artificial neural networks with exogenous vector inputs. *Energies, 10*(1), 40.

Chen, Y., He, K., & Tso, G. K. (2017). Forecasting crude oil prices: a deep learning based model. *Procedia computer science, 122*, 300-307.

Cheng, J., & Li, Q. (2008). Reliability analysis of structures using artificial neural network based genetic algorithms. *Computer methods in applied mechanics and engineering, 197*(45-48), 3742-3750.

Chniti, G., Bakir, H., & Zaher, H. (2017). *E-commerce time series forecasting using LSTM neural network and support vector regression.* Paper presented at the Proceedings of the International Conference on Big Data and Internet of Thing.

Fan, C., Xiao, F., & Zhao, Y. (2017). A short-term building cooling load prediction method using deep learning algorithms. *Applied energy, 195*, 222-233.

Fu, G. (2018). Deep belief network based ensemble approach for cooling load forecasting of air-conditioning system. *Energy, 148*, 269-282.

Gensler, A., Henze, J., Sick, B., & Raabe, N. (2016). *Deep Learning for solar power forecasting—An approach using AutoEncoder and LSTM Neural Networks.* Paper presented at the 2016 IEEE international conference on systems, man, and cybernetics (SMC).

Hashmi, M. U., Arora, V., & Priolkar, J. G. (2015). *Hourly electric load forecasting using nonlinear autoregressive with exogenous (narx) based neural network for the state of goa, india.* Paper presented at the 2015 International Conference on Industrial Instrumentation and Control (ICIC).

Ji, Y., Xu, P., Duan, P., & Lu, X. (2016). Estimating hourly cooling load in commercial buildings using a thermal network model and electricity submetering data. *Applied energy, 169*, 309-323.

Mocanu, E., Nguyen, P. H., Gibescu, M., & Kling, W. L. (2016). Deep learning for estimating building energy consumption. *Sustainable Energy, Grids and Networks, 6*, 91-99.

Pascanu, R., Gulcehre, C., Cho, K., & Bengio, Y. (2013). How to construct deep recurrent neural networks. *arXiv preprint arXiv:1312.6026*.

Singh, P., & Dwivedi, P. (2018). Integration of new evolutionary approach with artificial neural network for solving short term load forecast problem. *Applied energy, 217*, 537-549.

Tieleman, T. (2008). *Training restricted Boltzmann machines using approximations to the likelihood gradient.* Paper presented at the Proceedings of the 25th international conference on Machine learning.

Wang, Z., & Srinivasan, R. S. (2017). A review of artificial intelligence based building energy use prediction: Contrasting the capabilities of single and ensemble prediction models. *Renewable and Sustainable Energy Reviews, 75*, 796-808.