Improving Forecasting Accuracy for Cooling Load of Air-Conditioning System Using NARX Recurrent Neural Network

Fatima Shittu[1], Mustapha Lawal Abdulrahman[2] Shehu Mukhtar [3]

[1&2]Department of Computer Science, Federal Polytechnic, Damaturu
[3]Department of Mathematical Science, Abubakar Tafawa Balewa University, Bauchi.
Correspondence e-mail: fatimashittu@fedpodam.edu.ng

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. A number of technique and computational approaches have been used recently in order to improve the prediction accuracy. These techniques as presented in the literature would likely have decreased accuracy in application if the occupant of the building in the future differs significantly those that was concurrent with the training data. Recurrent 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 proposed a nonlinear autoregressive recurrent neural network to forecast cooling load data for air conditioning system while mitigating the future impact of changes in occupant. The proposed model was evaluated against state-of-the-art prediction techniques use in CLF forecasting using MSE and R on Matlab 2018a. The experimental result shows that the proposed model not only curtail future impact of occupant uncertainty but also outperforms the existing models in terms of accuracy and model fitting achieving a lower value in terms of MSE.

Keywords: Forecasting Accuracy, Cooling Load, NARX, HVAC, CLF Forecasting, MSE

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 (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(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 is aimed at designing and implementing a model that predict electricity consumption in residential buildings by enhancing the prediction reliability using NARX network while including occupancy information as an external factor to the prediction and evaluate the results against the state of the art. The key contribution of this study is implementing the NARX base Network that addresses the high error associated to medium and long term forecast by curtailing future impact of occupancy changes to the prediction results, thus enhancing the accuracy and reliability of the proposed model against state-of-the-art prediction techniques use in energy forecasting. The result from this model will greatly assist in planning and adjusting electricity demand and supply respectively.

The rest of this paper is divided into sub-sections. Section ii details the theoretical background of Deep recurrent network, Section iii presents the formulation and description of the proposed NARX model and the scheme for data normalization and further describes the evaluation setup, Section iv analyzed the simulation results and section v concludes by summarizing the key results

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. Computational Techniques

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).

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 1. Figure (1) presents the configuration of this model.

$$y(n+1)=f[y(n),\ldots,y(n-d_y+1);u(n),u(n-1),\ldots,u(n-d_u+1) \quad (1)$$

Where:

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

$\boldsymbol{y(n)}$ 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).

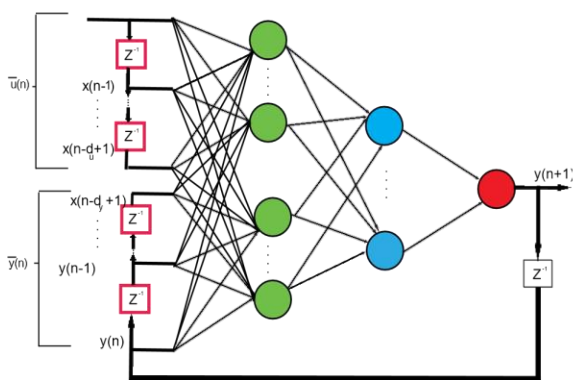

Figure. 1. NARX recurrent neural network architecture

IV. Related Work

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.

Similarly, (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.

Also, (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) analyse 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. Methodology

Neural network is very effective for pattern classifying for unstructured static data which is not related by time constraint. In temporal pattern recognition the pattern evolves in time. Conventional feed forward neural networks are not very effective with temporal data. Hence neural network has to adapt for these changes and has to consider time dependence in the training process. Recurrent neural network uses feedback loop in addition to the feed forward neural network architecture. NARX neural network model is a combination of Multilayer feed forward neural network, recurrent neural network and time delay neural network. In the literature a comparison is made through case study for implementing forecasting on a time series, they concluded the superior performance of NARX over other.

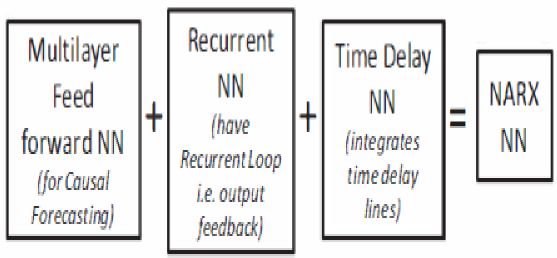

Figure 2: NARX network

NARX model is defined by the following equation

$Y(t) = f\,(x(t-1), x(t-2), \ldots x(t-n_x), y(t-1), y(t-2) \ldots y(t-n_y)$(2)

Where y is the predicted or output variable and x is the input vector.

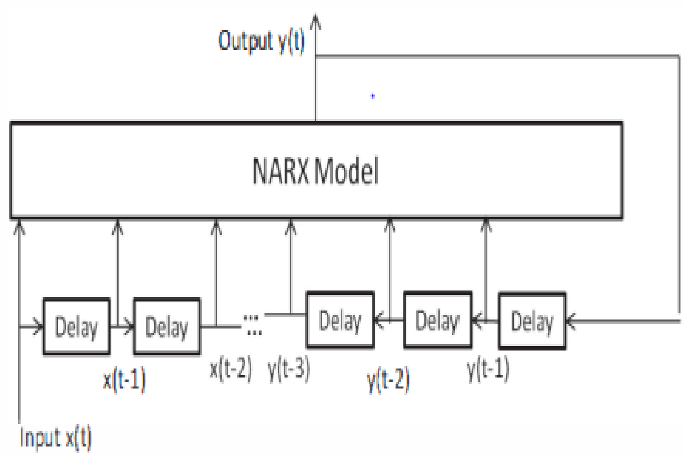

Figure 3: Proposed NARX Model

NARX network uses exogenous inputs along with feedback information enclosing layers of the network. The predicted quantity is regressed on past values of the output parameter and exogenous input parameters. NARX based neural network is formed for the hourly load forecast using MATLAB's neural network toolbox. The NARX network model consists of two layers feed forward network. It has sigmoid transfer function in the hidden layer and a linear transfer function in the output layer. The NARX network used for electric load forecasting application. The hidden layer has a nonlinear function. It is sigmoid for our case.

In this research, the training function used is Levenberg-Marquardth back propagation. The training function updates weights and bias values according to Levenberg-Marqardth optimization. It requires more memory but it is the fastest and good for supervised learning. This training method, like quasi Newton methods, approach second order training speed without computing Hessian matrix. Optimal performance not only depends on selection of the right neural network but also on the tuning process. One must select right number of neurons and delays and divide the data for training, validation and testing in proper ratio, to get best prediction. Selecting large number of neurons and delays may lead to over-fitted network, which may provide a low mean square error (MSE) on training and validation but high MSE on testing and out of sample tests.

VI. Data Collection and Model Input

To train the model for forecasting, the developed NARX network relies heavily on input vectors. Choosing an input vector comprising of external factor produces a better, robust and reliable performance. The proposed model was tested and evaluated on electricity consumption data obtained from the private residential buildings in using Smart Software Tools from residential smart meters. The dataset contains CLD data for 34 apartments for the period 1st January, 2014 to 31st December, 2016. The features of the dataset consisted of a combination of weather, date and CLD load consumption related variables respectively. We also obtained the occupancy data which was used to curtail the impact of occupancy behaviours on the forecasting accuracy.

Evaluation setup

To evaluate the performance of the models in this study, the data at one-hour intervals was partitioned into training, testing and a validation set. The features of the dataset consisted of weather, occupancy behaviour and electric CL consumption related variables respectively. The dataset is Partition into training, test and validation. The first 70% is for training, 15% for test and last 15% for validate respectively. The results will be evaluated using MSE and Error Correlation R.

VII. Results and Discussion

The accuracy of the estimated forecasts of the proposed model was compared with the other models to determine which model gives a more accurate forecast through the use of MSE, Regression and correlation error (R). The result for the forecasting using all the different algorithms by the performance standard achieved by each of the model is displayed and the findings is discussed.

Figure 4 display the error autocorrelation function. The figure shows how the prediction errors are related in time. Ideally, for perfect prediction model, there should be only one nonzero value of the autocorrelation function. These nonzero values should occur at zero lag; this is also called mean square error. Such an autocorrelation function would imply complete un-correlation of predicted errors with each other. Base on Figure 4, the correlations, excluding the zero lag, fall almost around the 95% confidence limits within zero, thus the model appears to be suitable. If there were significant correlation in the prediction errors, then it's possible to improve and enhance the prediction accuracy maybe by changing the neural network structure or increasing the number of delays in the network.

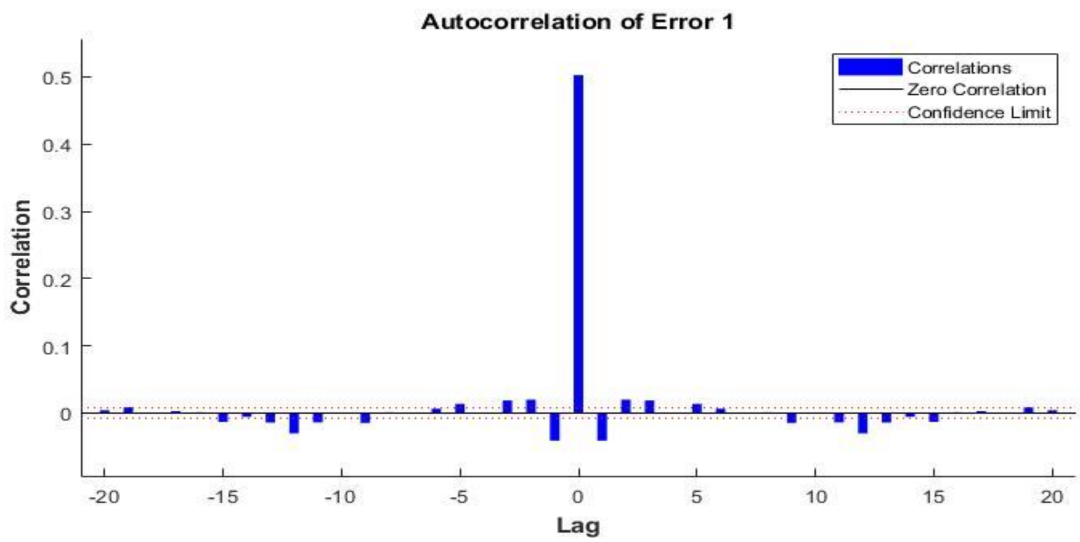

Figure 4. Auto correlation error for the proposed model

Similarly, Figure 5 Shows the Error histogram of the fit set for the proposed model.

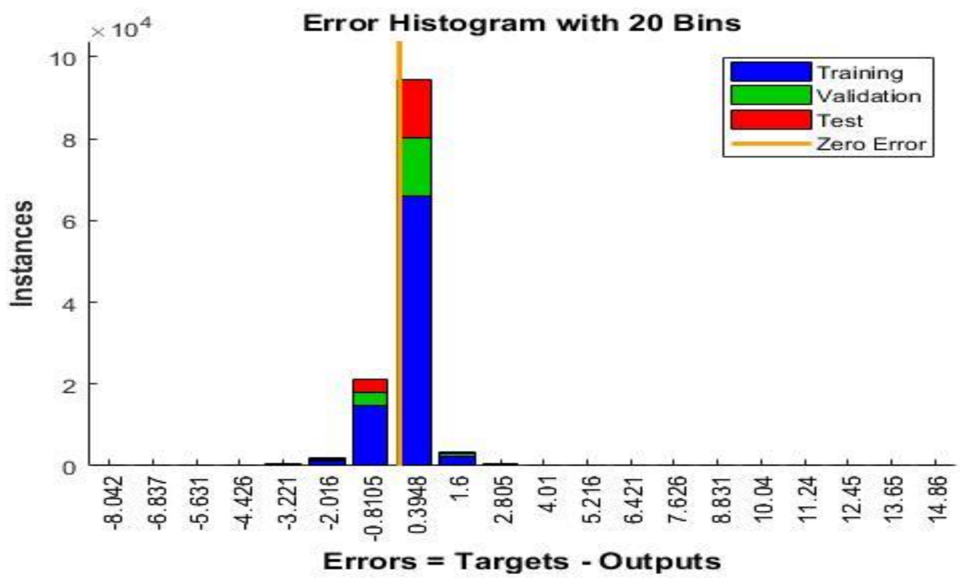

Figure 5 Error Histogram for fit set of the proposed model

Figure 6 shows the regression graph for the proposed model. It shows how accurately our trained model fits the dataset. It was observed that the value of R is 1 which imply good prediction, hence, the model prediction is very close to the test dataset. If it was close to zero, it implies bad prediction which shows that the model completely fails in making a correct prediction. In our own case, the model prediction was close to the actual test which demonstrates good prediction by the proposed model as shown in table 1.

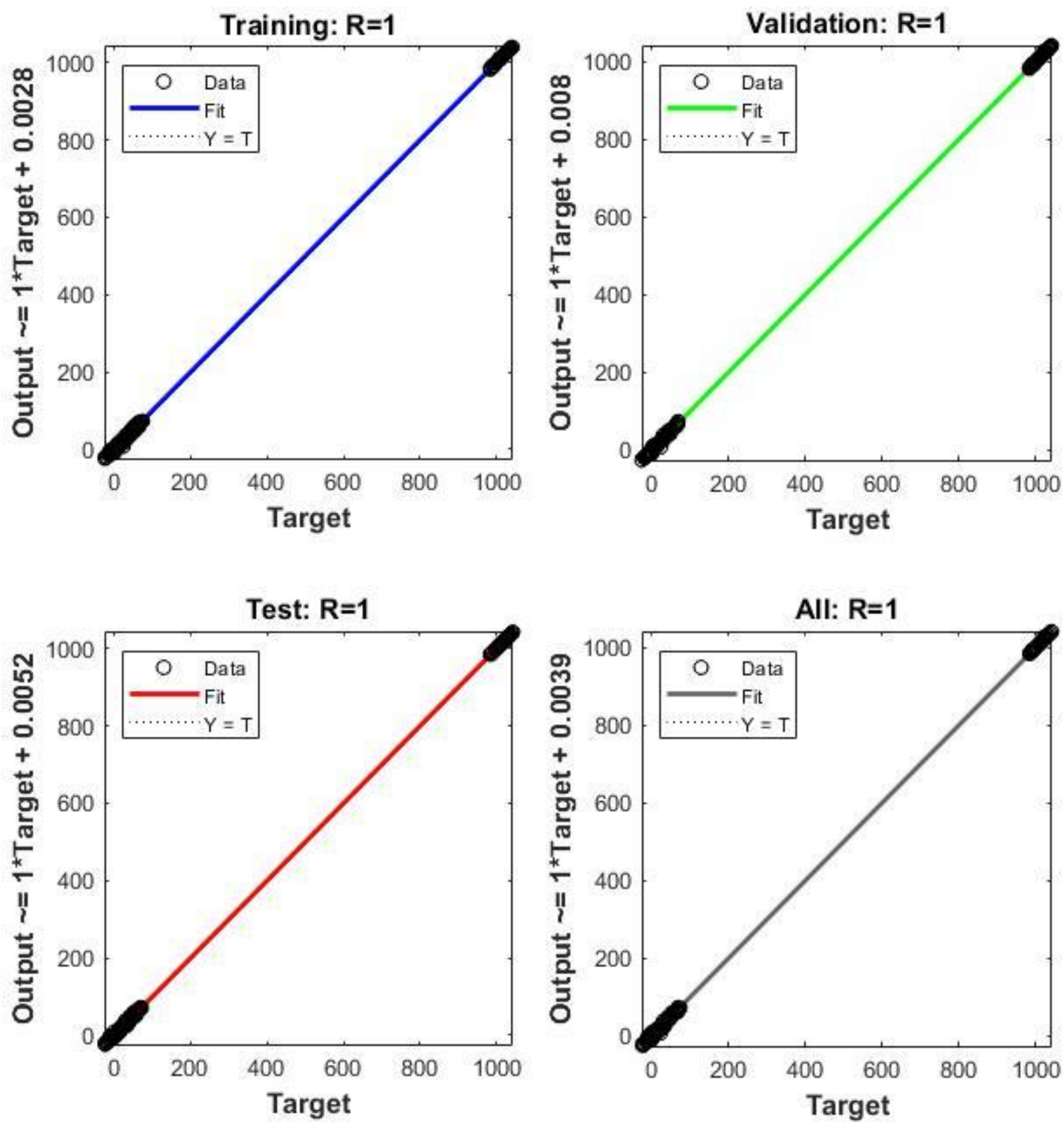

Figure 6. Error regression on forecast vs. actual for the proposed model

Figure 8 shows the validation MSE for the proposed model

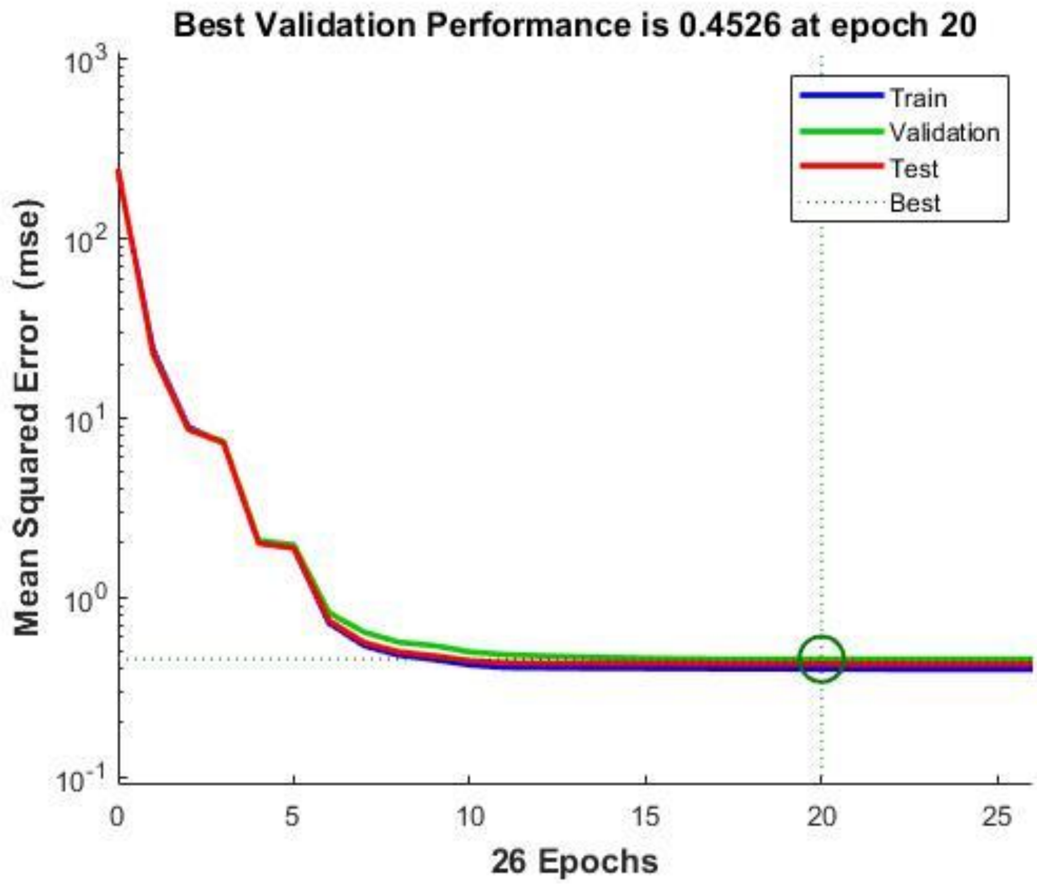

Figure 7. Validation MSE

Table 1 illustrates how the performances of the proposed NARX models compare with those of the MLP model

	Proposed Model	NARX	MLP
MSE	0.44105	0.51094	1.65471
R	1	0.984321	0765838

Clearly, from Figures 6 and 7. The propose model in general perform better than all the existing models achieving a lower value in terms of MSE and R. Moreover, the proposed model combines the occupancy information and predict concurrently with the CL data. Thus, the technique has significantly increased the prediction accuracy and reliability by curtailing future alterations of external factors such as the occupancy uncertainties to the predictions results.

VIII. Conclusion

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. Cooling load forecasting plays a crucial role in the planning, control and operation of heating, ventilating and air-conditioning systems. new paradigms have to be employed that will use automated methods to dynamically forecast the buildings energy consumption to curtail the gap between demand and supply, A number of technique and computational approaches have been used recently in order to improve the prediction accuracy. These techniques as presented in the literature would likely have decreased accuracy in application if the occupant of the building in the future differs significantly to those that was concurrent with the training data.

Recurrent 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 research work put forward a recurrent neural network to improve the forecast accuracy of cooling load data for air conditioning system while mitigating the future impact of changes in occupant. The proposed model was evaluated against NAR and MLP using MSE and R. The experimental result shows that the proposed model not only curtail future impact of occupant uncertainty but also outperforms the existing models in terms of accuracy and model fitting achieving a lower value in terms of MSE.

References

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.

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.

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*.

Zhao, H.-x., & Magoulès, F. (2012). A review on the prediction of building energy consumption. *Renewable and Sustainable Energy Reviews, 16*(6), 3586-3592.