SEMARAK ILMU PUBLISHING

CFD Letters

Journal homepage:
https://semarakilmu.com.my/journals/index.php/CFD_Letters/index
ISSN: 2180-1363

Long Short-Term Memory Neural Network Model for the Control of Temperature in A Multi-Circuit Air Conditioning System

Ibrahim Oleolo[1,*], Hayati Abdullah[1], Ismail Mustapha[2], Maziah Mohamad[1], Mohammad Nazri Mohd Jaafar[1], Hasan Alqaraghuli[4], Emmanuel Abiodun Abioye[4], Akeem Olowolayemo[3], Balogun Wasiu Adebayo[4], Sapiah Sulaiman[5],

[1] School of Mechanical Engineering, Faculty of Engineering, Universiti Teknologi Malaysia, 81310 UTM Johor Bahru, Johor, Malaysia
[2] School of Computing, Computer Science Department, Universiti Teknologi Malaysia, 81310 UTM Johor Bahru, Johor, Malaysia
[3] Department of Computer Science, KICT, International Islamic University Malaysia
[4] Division of Control and Mechatronics Engineering, School of Electrical Engineering Universiti Teknologi Malaysia, 81310 UTM Johor Bahru, Johor, Malaysia
[5] Smart Digital Community Research Alliance Universiti Teknologi Malaysia, 81310 UTM Johor Bahru, Johor, Malaysia

ARTICLE INFO

Article history:
Received 8 September 2022
Received in revised form 30 September 2022
Accepted 22 October 2022
Available online 31 December 2022

Keywords:
Temperature control; Neural Network; Multi-Circuit Air Conditioning System

ABSTRACT

Temperature control is important in energy management of buildings. Air conditioning system contributes a high percentage of the total energy consumption, the compressor, which is a major component of the Air conditioning system, utilizes up to 90% of the energy. This can drastically be reduced by varying the frequency of the compressor with respect to the required indoor temperature, as such, reducing the overall energy usage of the air conditioning system. The combination of a well-tuned controller and variable frequency drive can be used to achieve this. It is important to develop a good model which can be used to design the controller. Although there are published research works in the development of models for the control of air conditioning systems, there seems to be a lack of study in the area of multi-circuit centralized air conditioning system. In this study, two models were developed using Long Short Term Memory Neural Network and Recurrent Neural Network, utilizing compressor speed and indoor air temperature of a multi-circuit water cooled packaged unit as input and output respectively. Comparing the two models, results shows that the Long Short-Term Memory Neural Network model performed better across evaluation metrics such as R-squared, Mean Squared Error and Mean Absolute Error, with the value of 0.9638, 0.0049, and 0.0190 respectively.

1. Introduction

The control mechanism and optimization settings play a major role in how well an air conditioning system performs in terms of energy efficiency and good thermal comfort. The air conditioning system's dependability and stability are so critical to the extent that even a slight system inconsistency might result in serious consequences, like energy wastage and too much deviation from set point. The air conditioning system is a major energy consumer in the building sector, accounting

* *Corresponding author.*
E-mail address: engrkunlephd@gmail.com (Ibrahim Oleolo)

https://doi.org/10.37934/cfdl.14.12.8498

for over half of all the electrical energy used in commercial buildings [1]. Thermal comfort on the other hand is an essential requirement in building because it affects the productivity and health of the occupants. Considering these factors, the appropriate control operations in the HVAC system can go a long way in saving a lot of energy without compromising the indoor thermal comfort [2]. Therefore, proper energy management in the HVAC system largely depends on the control system and optimization parameters [3]. This has also been confirmed by several study on efficient energy management in HVAC, that energy efficiency in HVAC can be improved through advance control algorithm. A good control algorithm which makes energy consumption reduction possible is dependent on accurate dynamic modeling and accurate optimization techniques [4]. It is a very challenging task to develop an accurate and effective model for HVAC systems, even with its similarities with other process control, there are some features that makes it more challenging like non-linearity of the system, time-varying disturbances and set-points. This points to the reality why modeling techniques continue to become more advanced in HVAC system control [3]. It has been demonstrated that it is more viable and cost-effective to improve the control algorithm in order to gain more efficiency [5]. Because of its non-linearity, external disturbances, set points value, and time-varying dynamics, air conditioning system control is often considered to be more demanding, sophisticated, and distinctive than other forms of control systems [6]. As a result, there is a need for a good model as systems must be well modelled in order to evaluate and regulate energy usage and indoor air quality. There are several diverse model applications in controlling energy consumption in air conditioning systems. In general, there are three types of modeling methods used for air conditioning systems. The first type of modeling is the data driven approach, in which the input and output system data are collected and a connection is established between the input and output variables using mathematical formulation [3, 7]. The second technique is physics-based, in which system models are built using controlling physics rules and in-depth knowledge of the basic procedure. Mass balance, heat transfer, momentum, and flow balance are the fundamental energy laws upon which physics-based models are constructed. From these laws, a number of mathematical equations can be derived and solved. These models can be lumped parameter or distributed. A study [8] suggest a physics-based linear parametric Autoregressive-moving average with exogenous terms (ARMAX) room temperature model in an office building where thermodynamic equations are used to determine a linear regression model's structure and order. The third type is the grey box technique, in which the basic framework of the model is built using physics-based approach, and the model parameters are calculated using parameter estimated algorithms utilizing the system's measured data and requiring extensive understanding of the system. Because they employ the physics-based approach to generate the model structure and data to estimate the model parameters, grey box models combine the benefits of both physics-based and data-driven models. The combination of both physics-based and data-driven models makes it a more complex model. Physics-based methods form the basic structure of the model. A work that focuses on modeling thermal activity in buildings was done by previous study [9], to develop an efficient HVAC control algorithm. This is based on an approximate temperature model defined by the grey box approach for three different rooms in the same house.

The data driven approach is being utilized in this study. The aim of the data-driven approach is to establish a model for existing systems and to estimate system parameters using calculated data input and output. The data-driven method has been widely used in many fields and recent trends have shown that in HVAC applications it has also gained attention [10]. A study [6] categorized data-driven modelling into nine (9) major techniques. Such techniques include models of the frequency domain, algorithms of data mining, statistical models, geometric models, fuzzy logic models, state-space models case-based reasoning models, stochastic models and instant models.

Data-driven models are trying to find a real system approximation. To achieve an approximate model, the device identification process for data-driven modelling is useful for specific applications based on measured data and previous system awareness [11]. In this study data were collected for the parameters of indoor temperature and humidity, outdoor temperature and humidity, supply air temperature and humidity as well as mixed air temperature and humidity. The flow of the constant air volume air conditioning system was measured by an air flow velocity transducer. The power measurement of the two compressor was carried out using a power transducer. The data was analysed using variants of the Artificial Neural Network models namely the Long Short-Term Memory (LSTM) Neural Network model and the Recurrent Neural Network model [12]. The LSTM has been widely used in time series prediction and analysis, since it was first published by previous study [13]. The LSTM has been implemented across areas such as image analysis and processes, language translation and speech recognition. However, only a few research works have used the LSTM model to analyse and predict energy consumption. It was demonstrated in a study that while LSTM has a tough time predicting minute electric energy usage, it does a good job in predicting hourly electric energy consumption [14]. These findings demonstrate that the LSTM model has a lot of potential for predicting energy use. The capacity of the LSTM to solve non-linear problems and retain past data is superior to some other variant of neural network when compared to the time series model [15]. RNN on the hand can learn patterns from previous records, as well as generalise and forecast future load patterns for unknown data. During the previous few years, RNN had gained attention among researchers especially in the area of dynamic systems which can be described using networks with feedback connections. In this study, the results from both LSTM and RNN models were compared using evaluation metrics such as the Mean Squared Error, Mean Absolute Error, and R-Squared. Due to many elements such as the climatic environment and occupancy, air-conditioning consumption of energy is subject to quick variations but obeys periodic rules [16]. If these periodic criteria for precisely measuring the energy consumption of an air-conditioning system are properly mastered, a good energy-saving operation mode can be obtained [17].

1.1 Recurrent Neural Network

A RNN is a special case of ANN, the aim is to predict next steps in sequence of observation by comparing previous steps. The fundamental objective of RNNs is to foresee future trends by using sequential observations and learning from previous stages [18]. The knowledge obtained during the previous phases of reading sequential data is stored in the hidden layers of RNNs. Using earlier information, RNN do same task for each element in the sequence, predicting future unseen sequential data, hence they are referred to as a recurrent. The main drawback of a generic RNN is that it only recollects a few previous stages in the sequence, thus, unsuitable for retaining longer data sequences. Long Short-Term Memory (LSTM) recurrent network has been proposed to address this problem using the “memory cells” [19]. In Figure 1, the RNN architecture is displayed. The vector w(t), which represents the current word while employing 1 of N coding (thus, its size is equal to the size of the vocabulary), and the vector s(t-1), which represents output values in the hidden layer from the previous time step, are concatenated to create the vector x(t) [20]. The input, hidden, and output layers of the network, which was trained using conventional backpropagation, are all present. The following is how values in these layers are calculated.

$$x(t) = [w(t)^T s(t-1)^T]^T \tag{1}$$

$$s_j(t) = f\left(\sum_i x_i(t)u_{ji}\right) \tag{2}$$

$$y_k(t) = g\left(\sum_j s_j(t)v_{kj}\right) \tag{3}$$

where $f(z)$ and $g(z_m)$ are sigmoid and softmax activation functions

$$f(z) = \frac{1}{1+e^{-z}}, \quad g(z_m) = \frac{e^{z_m}}{\sum_k e^{z_k}}$$

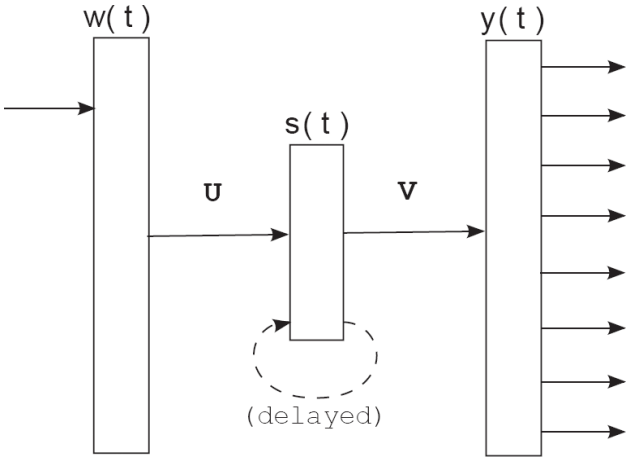


Fig. 1. Simple RNN Structure [18]

1.2 LSTM Model

The LSTM resulted from further development and overcoming the error in RNN [13] This issue is the long time lag which were inaccessible to the existing RNN architecture, considering the fact that backpropagated error either blows up or decays exponentially. LSTM is a special kind of RNN with additional features to memorize the sequence of data. It is composed of memory blocks, which are similar to differentiable versions of the memory chips used in digital computers. Each one has one or more recurrently connected memory cells as well as three multiplicative units (input, output, and forget gates) that provide continuous analogues for the write, read, and reset operations for the cells [21]. The input to the cells is multiplied by the input gate's activation, the output to the net is multiplied by the output gate's activation, and the prior cell values are multiplied by the forget gate's activation. Only the gates allow the net to connect with the cells. The structure of the LSTM is shown in Figure 2 below.

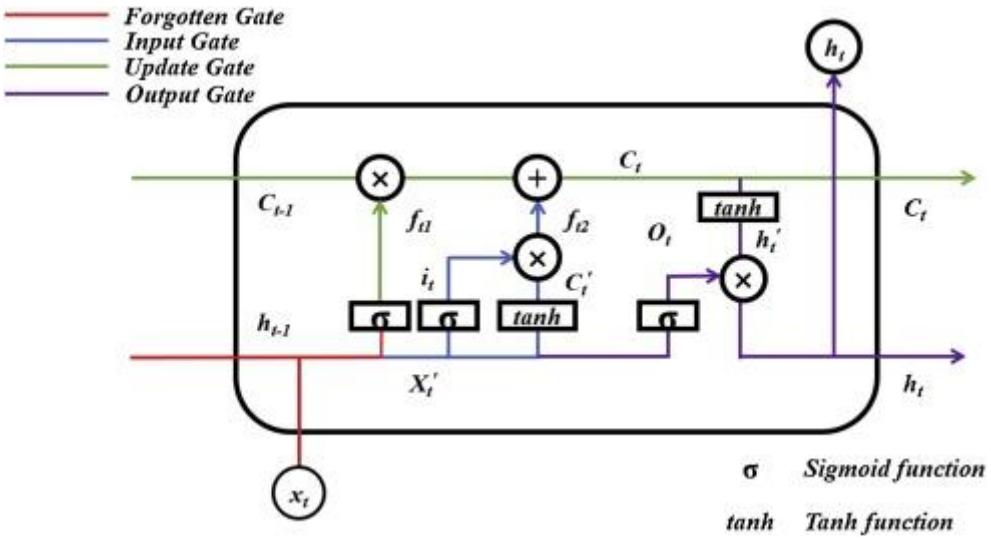


Fig. 2. LSTM Structure [15]

The preceding data and historic data, which is analyzed by the cells, are used to anticipate the data for the next moment. As indicated in Figure 3, the LSTM is made up of cells with similar structure. Each cell contains three input parameters, C_{t-1} (historically stored information), X_t (historical data), and h_{t-1} (prediction of prior cell outcomes and cell input parameter) [15, 22].

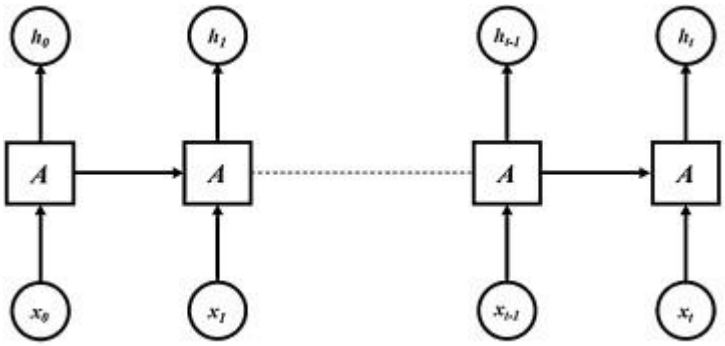


Fig. 3. LSTM Cell [15]

The data h_{t-1}, is processed by the preceding cell, and the input data of the present time X_t are combined via a matrix to form Eq. (1):

$$X'_t = [h_{t-1}, X_t] \tag{4}$$

In the forgotten gate, the LSTM obtains the capacity to filter data. X't is processed by the sigmoid function to obtain f_{t1} which is calculated thus:

$$f_{t1}=\sigma(W_f \cdot X'_t + b_f) \tag{5}$$

In calculating C'_t, X'_t is processed by the tanh function and sigmoid function to obtain i_t:

$$i_t =\sigma(W_i \cdot X'_t + b_t) \tag{6}$$

$$C'_t= \tanh(W_c \cdot X'_t + b_c) \tag{7}$$

$$f_{t2} = i_t \times C'_t \quad (8)$$

C_t is restructured in the update gate. The historical data is then derived by multiplying C_t and ft_1 with the matrix. Combining with ft_2, the neutral unit output C_t can be expressed as given below:

$$C_t = f_{t1} \times C_{t-1} + f_{t2} \quad (9)$$

The LSTM outputs the result in the output gate. To convert X'_t to O_t, the sigmoid function is employed. As a result, O_t decides which C_t should be kept. In addition, the tanh function transforms Ct into h'_t. The final data h_t is obtained by multiplying h'_t and O_t which is expressed below:

$$O_t = \sigma (W_o \cdot X'_t + b_o) \quad (10)$$

$$h_t = O_t * \tanh \quad (11)$$

2. Methodology

This procedure followed in this study to design the model for temperature control in multi-circuit air conditioning system are summarised in Figure 4.

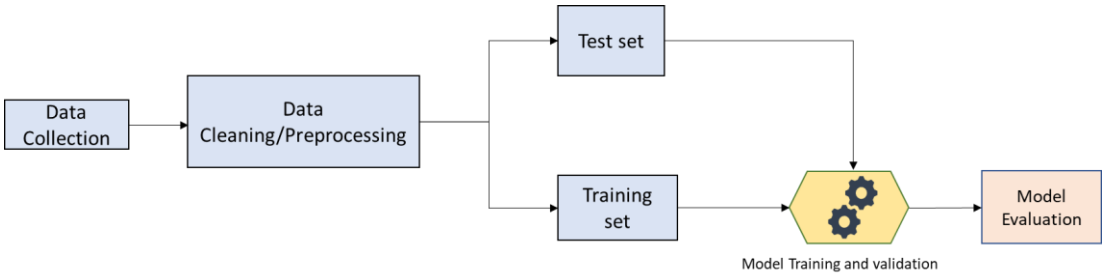


Fig. 4. Data analysis flowchart

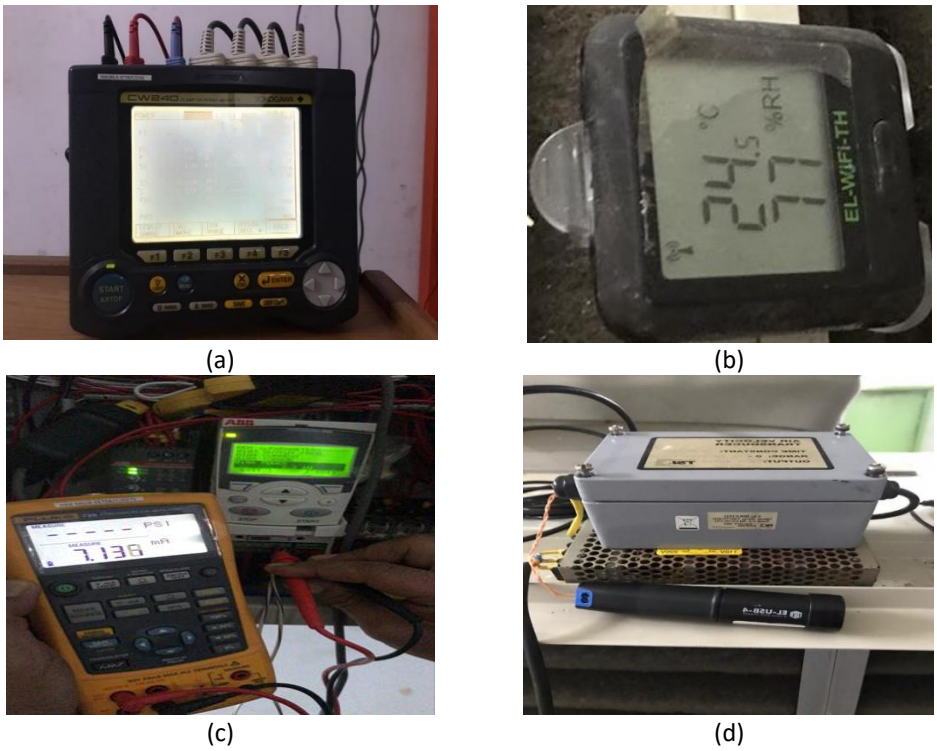
2.1 Experimental Setup and Data Collection

The multi-circuit air conditioning system used in this study is located in an office building (Block F54) in Universiti Teknologi Malaysia, Johor Bahru, Malaysia. The air conditioning system is a two-refrigeration circuit water-cooled packaged unit (WCPU). An evaporator unit with expansion valve, a water-cooled condenser, and a scroll type compressor makes up each circuit. In the event that one of the compressors fails, the two refrigeration circuits are designed to supply partial capability. The on/off thermostat and power connectors are situated in the AHU (Air Handling Unit) room in which the WCPU is located. The duct network that travels from the AHU to the indoor space constitutes the air distribution portion. The indoor space serves as an office environment, with an average of 40 occupants. The working hours are from 8:00 a.m. to 5:00 p.m. on Sunday to Wednesday, and 8:00 a.m. to 3.30 p.m. on Thursday. The set-point temperature for the air conditioning system was 24°C. Data gathering is a key step in creating an accurate model that accurately captures the system's physics as closely as possible. The power of the compressor is the input data, while the indoor temperature is the output data. Table 1 lists the measuring instruments with the parameters being measured.

Table 1
Parameters measured and instruments

No	Parameter	Instrument	Accuracy
1	Compressor Power	Clamp-on Power Analyzer	0.5% of full scale
2	Indoor Temperature	EasyLog wifi Temperature Data Logger	±0.3°C / ±0.6°F
3	Outdoor Temperature	EasyLog wifi Temperature Data Logger	±0.3°C / ±0.6°F
4	Supply Air	Air flow meter	±2.0% of reading
5	Indoor Relative Humidity	EasyLog wifi Humidity Data Logger	±2% RH
6	Outdoor Relative Humidity	EasyLog wifi Humidity Data Logger	±2% RH

In Figure 5 below the instruments that were used for data collection were shown during the two-week period of collecting data.

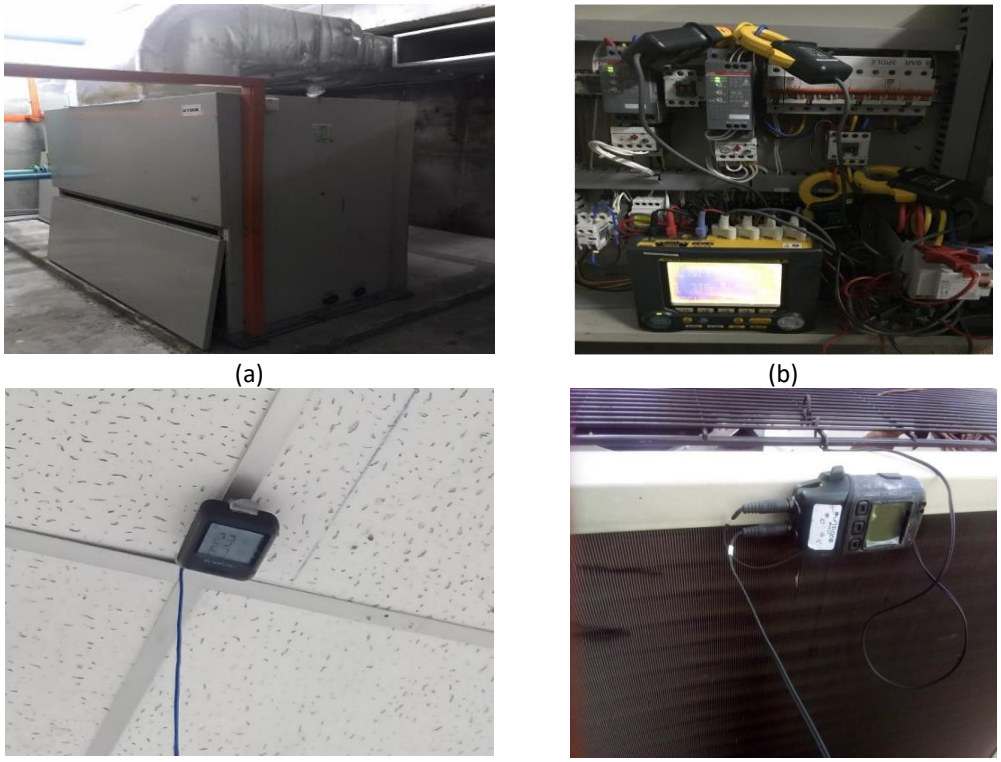


(a) (b)

(c) (d)

Fig. 5. Data Collection Instruments (a) Yokogawa IM CW240E Clamp-on Power Analyzer (b) EasyLog wifi Temperature and Humidity Data Logger (c) Power meter (d) Air flow meter

Figure 6 depicts the installation of the measuring equipment utilized in the data collection of crucial parameters. The locations were properly monitored for the duration of data collection to ensure seamless data collection.



(a) (b)

(c) (d)

Fig. 6. Data Collection instrument location (a) Air handling Unit (b) Installation of power analyzer at the power and control panel (c) The indoor data logger installation (d) The outdoor data logger installation

2.2 Data Description/Preprocessing

The experimental data is a time series data of per minute measurements of two compressor power inputs and a corresponding indoor temperature output for a duration of 4486 minutes. The summary statistics of each of the input and output variables are presented in Table 2. As shown in the table, the range of input and output variables vary which has been reported in pertinent works to affect the performance of neural networks [23]. Hence, each of the input and output variables are normalized such that each falls between 1 and -1 to avoid the associated shortcomings and speedup model convergence.

Table 2
Summary Statistics of Data

Statistics	Temperature	Compressor input 1	Compressor input 2
Count of samples	4486.0	4486.0	4486.0
mean	21.75	7716.16	8363.42
std	1.13	1092.04	471.96
min	19.35	0.0	1200.0
25%	20.91	7600.0	8100.0
50%	21.58	7800.0	8300.0
75%	22.23	8100.0	8700.0
max	26.61	8700.0	9200.0

2.3 Model Architectural Setup and Configuration

For this work, the proposed LSTM model has a single hidden layer with 128 hidden LSTM cells. This setup was enough to model the given data over a maximum training period of 200 epochs. Similarly, a baseline RNN model with a single hidden of 128 neurons was also built as a comparative method to the baseline. To prevent overfitting, two regularization methods, early stopping and dropout rate of 20%, were employed as part of the training phase. The early stopping simply monitors the validation error and stops the model training if there are improvement in performance after 5 consecutive epochs to ensure both the LSTM and RNN models do not overfit the training data. The RNN model used for comparison has similar architectural setup as the proposed LSTM. Adam optimizer with a learning rate of 0.0001 was used for model optimization. We trained each model using stochastic gradient decent with batch size of 32. All the models were implemented within pytorch deep learning framework environment [24].

For the predictive task of indoor temperature prediction, 50% of the data was used as training set while the remaining 50% served as the validation set. The sequence length represents the duration or length of input a time series prediction model is trained on for to predict the output for the next timestamp. For instance, we can choose to train our model for the 5 consecutive sequence (minutes) of data to predict the temperature for the 6th minute. Smaller sequences have been reported in the literature to lack relevant historical information necessary for good model performance while very large sequences appear to overwhelm the model with redundant information and consequently the deteriorate overall performance. Hence, we trained the models with varying sequence lengths, ranging from 60 to 10 minutes.

2.4 Evaluation Metrics

Three popular yet complementary evaluation metrics that were used to measure the performance of the proposed regression LSTM and RNN models in this research are as follows:

2.4.1 Coefficient of determination (R^2 or R-squared)

Given a set of expected outputs Y and model predicted outputs $\mathrm{f(X)}$, where $\mathrm{Y_j}$ and $\mathrm{f(X)_j}$ are respectively the actual and predicted for the $\mathrm{X_j th}$ sequence of data, we define the mean of the actual outputs as:

$$\mathrm{Mean_{out}} = \frac{1}{\mathrm{n}}\sum_{\mathrm{j}}^{\mathrm{n}} \mathrm{Y_j} \tag{12}$$

R^2 or R-squared [25] is given as follows:

$$\mathrm{R^2} = 1 - \frac{\sum_{\mathrm{j}}\left(\mathrm{Y_j} - \mathrm{f(X)_j}\right)^2}{\sum_{\mathrm{j}}\left(\mathrm{Y_j} - \mathrm{Mean_{out}}\right)^2} \tag{13}$$

where the numerator and denominator are the residual sum of squares and total sum of squares respectively. While the value ranges from -∞ to +1, the closer the value is to +1, the better is the result.

2.4.2 Mean square error (MSE)

Mean Square Error is given by Eq. (12). It has been found to be especially useful in detecting outlier within a predicted range.

$$\mathrm{MSE} = \frac{1}{\mathrm{n}}\sum_{\mathrm{j}}^{\mathrm{n}}(\mathrm{Y_j} - \mathrm{f(X)_j})^2 \tag{14}$$

The range of values for MSE is between 0 to +∞ with the best value being 0.

2.4.3 Mean absolute error (MAE)

Similarly, the range of values for MAE is between 0 to +∞ with the best value being 0. However, unlike the MSE, MAE is the sum of the absolute difference of the actual and the predicted outputs.

$$\mathrm{MAE} = \frac{1}{\mathrm{n}}\sum_{\mathrm{j}}^{\mathrm{n}}|\mathrm{Y_j} - \mathrm{f(X)_j}| \tag{15}$$

3. Results

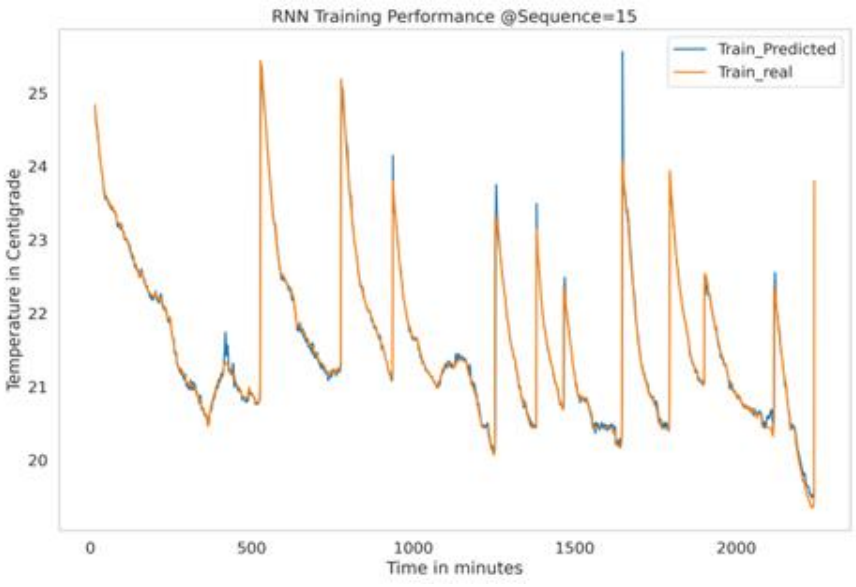
The results of the proposed LSTM model as well as the RNN model for varying sequence lengths across the different evaluation metrics is presented in Table 3. Varying the sequence lengths provided some useful insight on the performance of the models. Smaller sequence length generally tends to yield better performance compared to longer ones.

Although smaller sequence lengths are susceptible to model overfitting, the regularization methods (early stopping and dropout) employed in this studied sufficiently prevented the models from it. It can be observed from the table that although LSTM model trained with sequence length of 25 minutes yielded the best training performance but such impressive performance could not be replicated on the validation set which the true measure of the model performance. Hence, the best LSTM model is the one trained using sequence length of 30 minutes. Similar trend is observed on the RNN model which showed the best performance with data of sequence length 15 minutes.

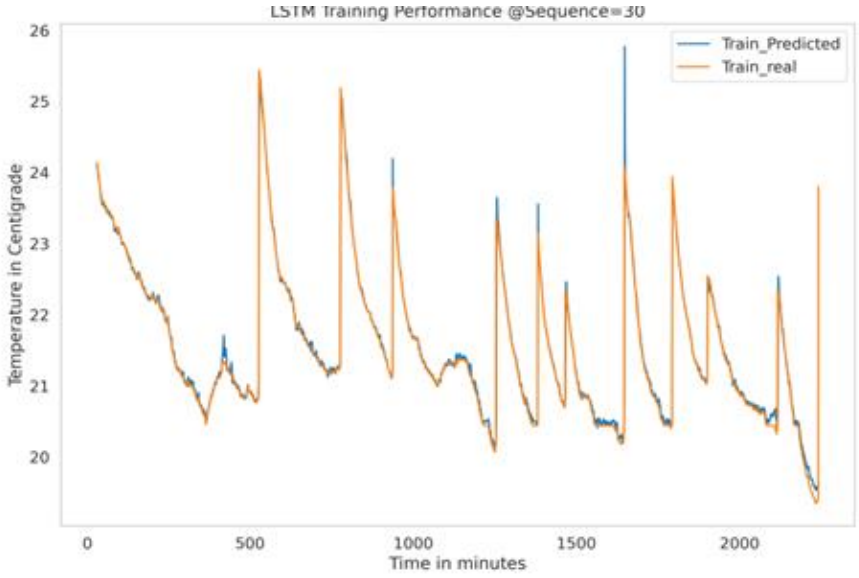
Table 3
Training and validation performance of LSTM and RNN models across varying sequence lengths

Sequence Length in minutes	Phase	LSTM			RNN		
		R-squared	MSE	MAE	R-squared	MSE	MAE
60	Training	0.9712	0.0030	0.0165	0.9725	0.0028	0.0159
	Testing	0.9514	0.0066	0.0226	0.9521	0.0065	0.0226
55	Training	0.9748	0.0026	0.0151	0.9735	**0.0027**	0.0143
	Testing	0.9585	0.0056	0.0205	0.9562	0.0060	0.0180
50	Training	0.9751	0.0026	0.0148	0.9727	0.0029	0.0155
	Testing	0.9590	0.0056	0.0200	0.9537	0.0063	0.0207
45	Training	0.9739	0.0027	0.0163	0.9705	0.0031	0.0166
	Testing	0.9524	0.0065	0.0228	0.9295	0.0095	0.0280
40	Training	0.9740	0.0026	0.0153	0.9736	0.0028	0.0146
	Testing	0.9536	0.0055	0.0231	0.9543	0.0062	0.0194
35	Training	0.9746	0.0027	0.0165	0.9718	0.0030	0.0185
	Testing	0.9547	0.0061	0.0242	0.9516	0.0065	0.0223
30	Training	0.9766	**0.0025**	0.0154	0.9745	0.0028	0.0166
	Testing	**0.9638**	**0.0049**	**0.0190**	0.9588	0.0055	0.0197
25	Training	**0.9775**	**0.0025**	**0.0142**	0.9751	**0.0027**	**0.0141**
	Testing	0.9555	0.0060	0.0215	0.9573	0.0057	0.0183
20	Training	0.9759	0.0027	0.0166	0.9750	0.0028	0.0176
	Testing	0.9592	0.0055	0.0208	0.9595	0.0054	0.0199
15	Training	0.9767	0.0026	0.0146	0.9766	**0.0027**	0.0150
	Testing	0.9580	0.0056	0.0194	**0.9612**	**0.0052**	**0.0171**
10	Training	0.9767	0.0027	0.0156	**0.9770**	**0.0027**	0.0151
	Testing	0.9581	0.0056	0.0206	0.9604	0.0053	0.0185

The training and validation performance of both the best LSTM and RNN models are also illustrated in Figure 7 and 8 respectively. It shows the time series graph of the training and validation performance of both the LSTM and RNN models.



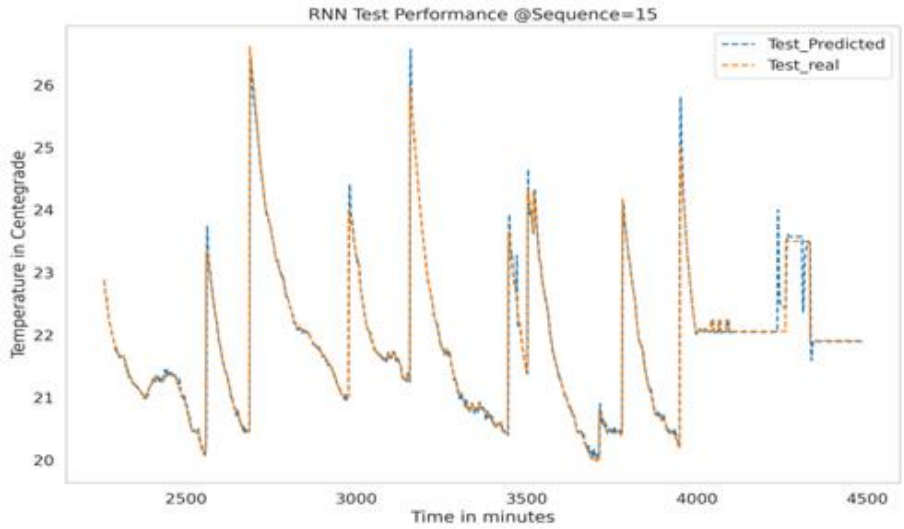
(a)



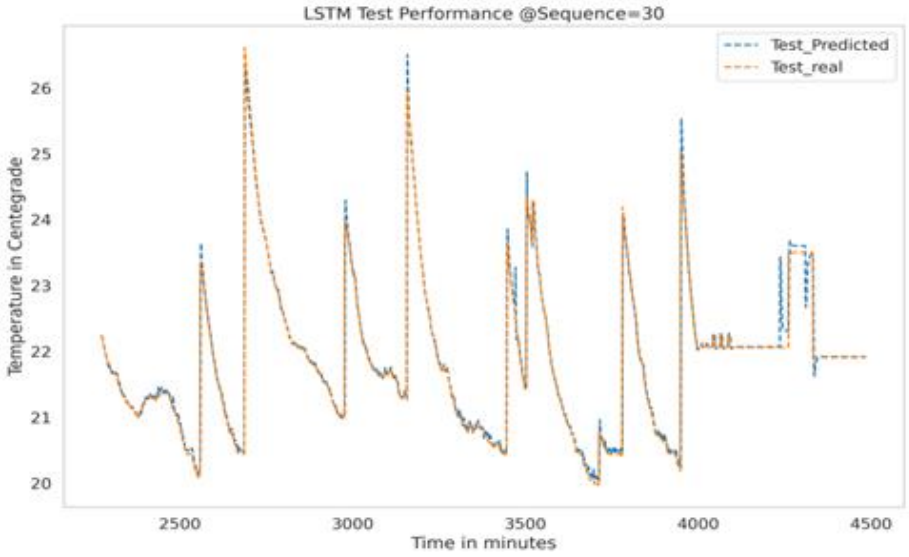
(b)

Fig. 7. Training Performance of RNN and LSTM models trained on 15- and 30-minutes sequence lengths respectively (a) RNN (b) LSTM

It can be observed from Figure 7 that both models fit the training data (*i.e.,* the real output) adequately with LSTM doing so better. This is also evident from the results obtained for the evaluation metrics in terms of the Mean Squared Error (MSE), Mean Absolute Per Error (MAE) and R-squared as shown in Table 3. Figure 8 similarly shows the graph of the validation predictions of both LSTM and RNN models with the real output. While both models were able to capture the distribution of the real output in the early part of the time series, the RNN model notably deviates from the true values as it approaches the end.



(a)



(b)

Fig. 8. Validation Performance of RNN and LSTM models trained on 15 and 30 minutes sequence lengths respectively (a) RNN (b) LSTM

The true performance of any machine learning model can only be shown by it performance on a validation/test dataset [26]. In contrast to the RNN model, the LSTM model shows minimal deviation from the real output as it approaches the end of the time as in Figure 8(b). This is largely possible due to the presence of memory cells which allows LSTM to remember longer sequence of data. The superior validation performance of the best LSTM model over the corresponding best RNN model can also be seen in Table 3 where it shows better R-squared and MSE of 0.9638 and 0.0049 compared to 0.9612 and 0.0052 for the RNN model respectively. However, in terms of MAE, the best RNN model shows better performance of 0.0171 over the LSTM model with 0.0190. Moreover, results for the LSTM and RNN models as illustrated in Table 3 shows good training performance for both models and replicating similar performance on the validation data indicates that both models did not overfit the the training data.

4. Conclusions

Selecting an accurate model that best describes the dynamics of the system is very important to develop a control process which will reduce energy usage without compromising the desired indoor condition [27]. This research contributes to the body of knowledge in the field of building and energy management, particularly in the design of controllers for air conditioning systems. It is important to note that the results of the proposed model in this study are applicable to the experimental set-up for the case study reported in this paper. However, a similar approach for obtaining dynamic models can be used for other centralized multi-circuit air conditioning systems in buildings.

Acknowledgement

This work was supported by the Ministry of Higher Education Malaysia and Universiti Teknologi Malaysia through the Translational Research Grant (UTM-TRG) Vot 4J337.

References

[1] Abdullah, Hayati, Oleolo Ibrahim, Mohammad Nazri Mohd Jaafar, Maziah Mohamad, Akmal Baharain, and Sapiah Sulaiman. "Energy Savings In a Multi-Circuit Water-Cooled Packaged Unit Air-Conditioning System." *Journal of*

Advanced Research in Fluid Mechanics and Thermal Sciences 76, no. 1 (2020): 39-53. https://doi.org/10.37934/arfmts.76.1.3953

[2] Fasiuddin, M., and I. Budaiwi. "HVAC system strategies for energy conservation in commercial buildings in Saudi Arabia." *Energy and Buildings* 43, no. 12 (2011): 3457-3466. https://doi.org/10.1016/j.enbuild.2011.09.004

[3] Afroz, Zakia, G. M. Shafiullah, Tania Urmee, and Gary Higgins. "Modeling techniques used in building HVAC control systems: A review." *Renewable and sustainable energy reviews* 83 (2018): 64-84. https://doi.org/10.1016/j.rser.2017.10.044

[4] Wei, Xiupeng, Andrew Kusiak, Mingyang Li, Fan Tang, and Yaohui Zeng. "Multi-objective optimization of the HVAC (heating, ventilation, and air conditioning) system performance." *Energy* 83 (2015): 294-306. https://doi.org/10.1016/j.energy.2015.02.024

[5] Oleolo, Ibrahim, Hayati Abdullah, Maziah Mohamad, Mohammad Nazri Mohd Jaafar, Akmal Baharain, and Sapiah Sulaiman. "Multi-Circuit Air-Conditioning System Modelling for Temperature Control." *Journal of Advanced Research in Fluid Mechanics and Thermal Sciences* 83, no. 2 (2021): 14-24. https://doi.org/10.37934/arfmts.83.2.1424

[6] Afram, Abdul, and Farrokh Janabi-Sharifi. "Theory and applications of HVAC control systems–A review of model predictive control (MPC)." *Building and Environment* 72 (2014): 343-355. https://doi.org/10.1016/j.buildenv.2013.11.016

[7] Oleolo, Ibrahim, Hayati Abdullah, Maziah Mohamad, Mohammad Nazri Mohd Jaafar, and Akmal Baharain. "Model Selection for the Control of Temperature in a Centralized Air Conditioning System." *Journal of Advanced Research in Applied Mechanics* 74, no. 1 (2020): 10-20.

[8] Wu, Siyu, and Jian-Qiao Sun. "Two-stage regression model of thermal comfort in office buildings." *Building and Environment* 57 (2012): 88-96. https://doi.org/10.1016/j.buildenv.2012.04.014

[9] Maasoumy, Mehdi, Alessandro Pinto, and Alberto Sangiovanni-Vincentelli. "Model-based hierarchical optimal control design for HVAC systems." In *Dynamic Systems and Control Conference*, vol. 54754, pp. 271-278. 2011. https://doi.org/10.1115/DSCC2011-6078

[10] Kusiak, Andrew, and Guanglin Xu. "Modeling and optimization of HVAC systems using a dynamic neural network." *Energy* 42, no. 1 (2012): 241-250. https://doi.org/10.1016/j.energy.2012.03.063

[11] Oleolo, Ibrahim, Hayati Abdullah, Maziah Mohamad, Mohammad Nazri Mohd Jaafar, Akmal Baharain, and Sapiah Sulaiman. "MODEL ANALYSIS FOR THE CONTROL OF HUMIDITY IN A WATER-COOLED CENTRALIZED AIR CONDITIONING SYSTEM." In *THE 9th INTERNATIONAL GRADUATE CONFERENCE ON ENGINEERING, SCIENCE AND HUMANITIES*, p. 59. UNIVERSITI TEKNOLOGI MALAYSIA, 2022.

[12] Koury, R. N. N., L. Machado, and K. A. R. Ismail. "Numerical simulation of a variable speed refrigeration system." *International journal of refrigeration* 24, no. 2 (2001): 192-200. https://doi.org/10.1016/S0140-7007(00)00014-1

[13] Hochreiter, Sepp, and Jürgen Schmidhuber. "Long short-term memory." *Neural computation* 9, no. 8 (1997): 1735-1780. https://doi.org/10.1162/neco.1997.9.8.1735

[14] Marino, Daniel L., Kasun Amarasinghe, and Milos Manic. "Building energy load forecasting using deep neural networks." In *IECON 2016-42nd Annual Conference of the IEEE Industrial Electronics Society*, pp. 7046-7051. IEEE, 2016. https://doi.org/10.1109/IECON.2016.7793413

[15] Zhou, Chonggang, Zhaosong Fang, Xiaoning Xu, Xuelin Zhang, Yunfei Ding, and Xiangyang Jiang. "Using long short-term memory networks to predict energy consumption of air-conditioning systems." *Sustainable Cities and Society* 55 (2020): 102000. https://doi.org/10.1016/j.scs.2019.102000

[16] Wang, Jian Qi, Yu Du, and Jing Wang. "LSTM based long-term energy consumption prediction with periodicity." *Energy* 197 (2020): 117197. https://doi.org/10.1016/j.energy.2020.117197

[17] Kim, Tae-Young, and Sung-Bae Cho. "Predicting residential energy consumption using CNN-LSTM neural networks." *Energy* 182 (2019): 72-81. https://doi.org/10.1016/j.energy.2019.05.230

[18] Siami-Namini, Sima, and Akbar Siami Namin. "Forecasting economics and financial time series: ARIMA vs. LSTM." *arXiv preprint arXiv:1803.06386* (2018).

[19] Pacella, Massimo, and Gabriele Papadia. "Evaluation of deep learning with long short-term memory networks for time series forecasting in supply chain management." *Procedia CIRP* 99 (2021): 604-609. https://doi.org/10.1016/j.procir.2021.03.081

[20] Mikolov, T., S. Kombrink, and L. Burget. "Cernocky, ˇ J., Khudanpur, S.: Extensions of recurrent neural network language model." In *2011 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 5528-5531. https://doi.org/10.1109/ICASSP.2011.5947611

[21] Graves, Alex, and Jürgen Schmidhuber. "Framewise phoneme classification with bidirectional LSTM networks." In *Proceedings. 2005 IEEE International Joint Conference on Neural Networks, 2005.*, vol. 4, pp. 2047-2052. IEEE, 2005. https://doi.org/10.1016/j.neunet.2005.06.042

[22] Liu, Pengfei, Xipeng Qiu, and Xuanjing Huang. "Recurrent neural network for text classification with multi-task learning." *arXiv preprint arXiv:1605.05101* (2016).
[23] Bishop, Christopher M. *Neural networks for pattern recognition*. Oxford university press, 1995. https://doi.org/10.1201/9781420050646.ptb6
[24] Paszke, Pytorch. "An imperative style, high-performance deep learning library." *Adv. Neural Inf. Process. Syst* 32: 8026.
[25] Wright, Sewall. "Correlation and causation." (1921).
[26] Ying, Xue. "An overview of overfitting and its solutions." In *Journal of physics: Conference series*, vol. 1168, no. 2, p. 022022. IOP Publishing, (2019). https://doi.org/10.1088/1742-6596/1168/2/022022
[27] Aprea, C., R. Mastrullo, and C. Renno. "Experimental analysis of the scroll compressor performances varying its speed." *Applied thermal engineering* 26, no. 10 (2006): 983-992. https://doi.org/10.1016/j.applthermaleng.2005.10.023