Machine Learning Baseline Energy Model (MLBEM) to Evaluate Prediction Performances in Building Energy Consumption

Rijalul Fahmi Mustapa
College of Engineering, Universiti Teknologi MARA, Johor Branch, 81750 Masai, Johor, Malaysia
rijalulfahmi@uitm.edu.my

Muhammad Asraf Hairuddin
College of Engineering, Universiti Teknologi MARA, Johor Branch, 81750 Masai, Johor, Malaysia | Institute for Big Data Analytics and Artificial Intelligence (IBDAAI), Universiti Teknologi MARA, Shah Alam 40450, Malaysia
masraf@uitm.edu.my (corresponding author)

Atiqah Hamizah Mohd Nordin
College of Engineering, Universiti Teknologi MARA Johor Branch, 81750 Masai, Johor, Malaysia
atiqah26@uitm.edu.my

Nofri Yenita Dahlan
Solar Research Institute, Universiti Teknologi MARA, 40450 Shah Alam, Selangor, Malaysia
nofriyenita012@uitm.edu.my

Ihsan Mohd Yassin
Microwave Research Institute, Universiti Teknologi MARA, 40450 Shah Alam, Selangor, Malaysia
ihsan_yassin@uitm.edu.my

Nur Dalila Khirul Ashar
Department of Computer and Communication Systems, Faculty of Engineering, Universiti Putra Malaysia, 43400 Serdang, Selangor, Malaysia | College of Engineering, Universiti Teknologi MARA, Johor Branch, 81750 Masai, Johor, Malaysia
nurdalila306@uitm.edu.my

Received: 1 May 2024 | Revised: 17 June 2024 | Accepted: 20 June 2024

 DOI: https://doi.org/10.48084/etasr.7683

ABSTRACT

Electric Energy Consumption (EEC) prediction for building operations can be performed using a Baseline Energy Model (BEM), which is vital to ensure the efficiency of the EEC estimates with its respective independent variables. However, developing the BEM to represent the relationship between independent variables can be a complex task due to the EEC variability in an educational building that differs during its operation period. The best-suited BEM must be continuously improvised to achieve good modeling with accurate and reliable predictions that capture the building operations' current dynamics. This study aims to conduct a comparative performance assessment between deep learning, machine learning, and statistical models to develop the BEM and, therefore, predict the EEC of the building for 24, 48, 72, and 96 hours, while considering the operation of the lecture weeks and the associated number of students and staff. The hours and temperature are considered as independent variables to be tested with residual error evaluations, whilst the correlation coefficient, coefficient of determination, and training time are also taken into account. Three models with different categories involving Long Short-Term Memory (LSTM),

Support Vector Regression (SVR), and AutoRegressive Integrated Moving Average with Exogenous inputs (ARIMAX) were compared, concluding that SVR was the best and can be used as a universal model in the Machine Learning Baseline Energy Model (MLBEM) studies. Accurate EEC prediction will offer a huge advantage for building operators to properly monitor, plan, and manage the EEC, hence avoiding excessive utility bills.

Keywords-energy efficiency forecast; machine learning; deep learning; baseline model; buildings

I. INTRODUCTION

The surge in the expansion of new infrastructure buildings is expected to lead to an increase in Electric Energy Consumption (EEC), and the energy demand from buildings is expected to increase by 50% by 2050. This accounts for factors, such as population growth, urbanization, and higher living standards contributing to higher greenhouse gas emissions. The trend of increasing EEC will greatly impact greenhouse gas production. The International Energy Agency (IEA) reported that global electricity consumption increased by 1.7%, reaching 22,848 TWh at the end of 2019 [1]. Furthermore, global electricity consumption increased by 2% in 2022, and Asia experienced the highest increase of 3.6% [2]. This rising trend in electricity consumption may have adverse environmental effects. Notably, the increase in EEC is primarily attributed to energy use in buildings [3-6]. At a global level, several plans tried to promote energy efficiency measures to reduce electricity consumption in buildings, mitigating the environmental impact of increased energy use. Such plans involve Malaysia's National Energy Efficiency Action Plan (2016-2035) [7], China's 14th Five-Year Plan for Renewable Energy Development (2021-2025) [8], India's National Mission for Enhanced Energy Efficiency (NMEEE) [9], and Germany's Energy Efficiency Strategy 2050 [10].

Prior to any energy efficiency steps being implemented to quantify savings, it is crucial for building operators to establish a good Baseline Energy Model (BEM). This BEM model represents the EEC with its respective independent variables. It is important to establish a model driven by the data, such as the Machine Learning Baseline Energy Model (MLBEM), in order to evaluate the electric energy savings due to its capability to recognize the pattern underlying the examined dataset. Therefore, MLBEM can be considered a powerful tool that can be used to observe the pattern of the EEC, enabling operators to predict the future EEC in the building. The energy savings were calculated based on the difference between the predicted and the measured EEC after the implementation of the energy efficiency steps. A bad BEM will provide inaccurate predictions, which deviate, thus underestimating energy savings. Conversely, a good BEM will provide accurate prediction and will therefore quantify good energy savings. Further understanding of energy savings quantification is depicted in Figure 1 [11]. The figure illustrates two important sections: the baseline period and the post-Energy Conservation Measures (ECM) period. The baseline period is where the measurement of EEC is conducted. During this period, the EEC is known as the normal consumption without any ECM intervention. After any ECMs are implemented, the period is known as the post-ECM period, where EEC is measured after ECMs have been conducted. During the post-ECM period, an adjustment of the EEC is necessary to estimate the EEC as if no ECMs were conducted during this period. This adjustment is established by predicting the EEC using the BEM. The adjustment of EEC through prediction is made since energy savings from ECM activities are not simply the subtraction of the post-ECM period EEC from the baseline period EEC. This is due to the fact that the independent variables that govern the EEC during the baseline period and the post-ECM period may change. Hence, predicting the EEC in order to adjust it through a BEM is a vital action.

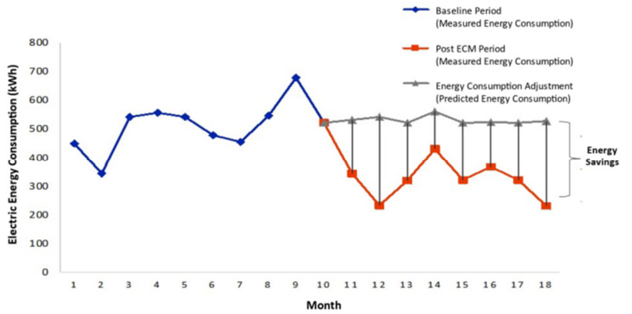

Fig. 1. International Performance Measurement and Verification Protocol (IPMVP) energy savings quantification framework.

The prediction of EEC in Malaysia is not new, as certain studies have been conducted on this subject [12-14]. Authors in [13] developed a BEM using a hybrid artificial neural network to perform EEC predictions of a commercial building. EEC was predicted during the post-energy efficiency period and then subtracted from the measured energy consumption for energy savings quantification. Authors in [12] predicted the EEC utilizing models, such as Support Vector Machine (SVM), Artificial Neural Networks (ANN), and K-Nearest Neighbor (KNN), aiming to improve the accuracy of the prediction made by the building energy management system. A prediction of energy demand in Malaysia [14] was made deploying a Non-linear Autoregressive Exogenous Artificial Neural Network (NARX-ANN) and a Multiple Non-linear Regression (MNLR) model. It is important to note that the prediction model from NARX-ANN has a higher Coefficient of Determination (R^2) compared to the MNLR model. Prediction serves as an important aspect in planning activities. Hence, models that provide the best accuracies are of utmost importance.

Challenges arise where the BEM development requires an appropriate modelling technique to accurately model the EEC with its related independent variables. Furthermore, no known specific investigations have used a certain analysis of the selection of independent variables. However, an in-depth review of the independent variables selected, such as occupations in the educational building, is available [15-18]. The modeling technique is essential to model the EEC with its related independent variables that have a strong correlation in the building. Different operation periods will lead to different EECs based on the occupied spaces in the building. Discrimination of the operation period has to be executed. Thus, this work will model two different BEMs representing two different operation periods in one educational building, i.e. lecture week and non-lecture week. The method utilizes Machine Learning (ML), where a learning algorithm was employed to study the historical data without prior knowledge to become a model representing the EEC [14, 19, 20]. The choice of the modeling technique for the EEC prediction significantly influences the accuracy of predictions. The use of ML techniques for EEC prediction purposes is rapidly increasing [21, 22]. Furthermore, it is apparent that the ML models in [23, 24] perform well, with the EEC data typically recorded over the course of a year. Notably, ML has gained prominence due to its ability to provide accurate predictions after the modeling stage, making it a robust choice. In [12], ML techniques (SVM and ANNs) are proposed to predict EEC in Malaysia's smart buildings, addressing the issue of low precision in building energy management system predictions. Authors in [25] involved large-scale energy prediction in the residential and commercial sectors. Independent factors such as the Gross Domestic Product (GDP) were used in the ML models, resulting in highly accurate predictions. Energy predictions in various countries were performed using ANNs with the NARX model for heating loads in [26]. Authors in [27] focussed on predicting cooling loads employing various Deep Learning (DL) and ML methods. In [28], weather data served as independent variables for the EEC prediction in Malaysian educational buildings. Notably, the predictions of the Long Short-Term Memory (LSTM) model were compared with (SVR) and Gaussian process regression, with LSTM demonstrating superior performance using four years of EEC data.

However, in practice, regular measurement and recording of EEC may not be feasible for certain building operators due to the high cost of installing dedicated building energy management systems. This limitation results in a scarcity of data for building owners to make accurate predictions about their buildings' EEC. Furthermore, as far as is known, a limited number of studies involve the prediction of EEC using data collected for less than a year and comparative modelling between DL, ML, and statistical model analysis in MLBEM studies.

II. DATASET DESCRIPTION

The building selected for the case study is the building of the Electrical Engineering Studies College of Engineering. It is a 6-story building, and the spaces contained within the building are summarized in Table I. The building's main activity is to conduct teaching and learning activities. The case study will consider the building's operation during lecture weeks. Mainly, the occupants in the building use the lecturers' offices, classrooms, and laboratories. These spaces are primarily utilized weekly from Sunday to Thursday, which are the working days. The main operating hours are from 8:00 a.m. to 5:00 p.m. The measurement of the EEC and the collection of independent variables will be discussed below.

A. *Electrical Energy Consumption*

The building receives a three-phase 0.415 kV source from the secondary voltage of an 11 kV/0.415 kV transformer. The 11 kV voltage is supplied by the local utility. This work will measure the EEC in the building from Sunday to Thursday. The EEC used for this study is the hourly EEC from 12:00 a.m. until 11:00 p.m. The measured EEC will be the input to the MLBEM. A data logger was connected at the 0.415 kV voltage level in the main switch room, which supplied the voltage to the building. The EEC interval used in this work is the hourly EEC interval. The EEC during weekends and public holidays will not be included in this study since the EEC in the building is relatively low. Additionally, modeling the EEC during weekends and public holidays may require a special model, as the factors governing the EEC during these periods may not be the same as on weekdays. The measured hourly EEC is displayed in Figure 2.

TABLE I. LIST OF SPACE AND NUMBER OF UNITS IN THE BUILDING

Space	Number of units
Lecturer's office room	80
Classroom	4
Laboratory	24
Meeting room	5
Toilets	16

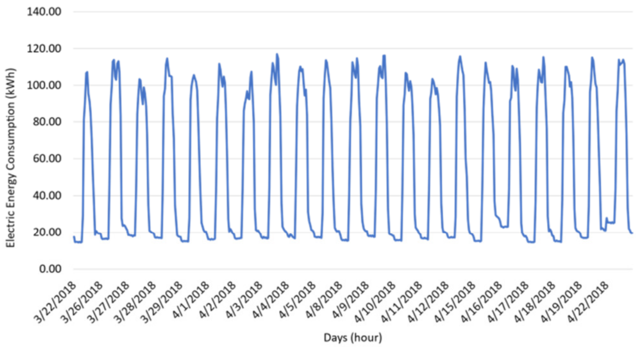

Fig. 2. Measured hourly EEC.

B. *Independent Variables*

The independent variables that may impact the EEC are lecturers' occupancy in office rooms, lecturers' occupancy in classrooms, lecturers' occupancy in laboratories, students' occupancy in classrooms, students' occupancy in laboratories, and outdoor temperature. In total, six different independent variables will be used as factors governing the EEC in the building. Details of the independent variable data collection will be further explained below.

1) *Lecturer's Occupancy in the Building*

Lecturer occupancy in the building is segmented into three categories. The first category is lecturers in office rooms. Occupancy in lecturers' office rooms is counted based on the attendance system recorded in a database. This is due to the fact that most lecturers spend their time in their office rooms except for classes and laboratory sessions. Lecturers need to scan a biometric fingerprint when arriving at and leaving the building. The attendance record will reveal the timestamp for the arrival and departure of each lecturer. If one of the timestamps is not available in the record, it will be counted for the day. It is assumed that the loads in the office rooms are fully utilized during office hours. Meanwhile, lecturers in classrooms and laboratories will be counted based on the lecture and laboratory session timetable. The presence of lecturers in classrooms can have a significant impact due to the use of plug loads such as LCD projectors and LED TVs. The presence of lecturers in these spaces will be recorded during lecture weeks from 8:00 a.m. to 5:00 p.m. From 12:00 a.m. until 7:00 a.m. and from 6:00 p.m. until 11:00 p.m., there will be no occupants. Hence, occupancy is zero during these hours.

2) *Student's Occupancy and Outdoor Temperature*

The occupancy of students in the lecture rooms and laboratories was counted based on the lectures and laboratory sessions conducted in these spaces. First, the students' subjects and laboratory sessions carried out in these spaces were identified from the timetable. Next, the number of students registered for the subjects and laboratory sessions in these spaces was retrieved from the student information system. Then, the total number of students occupying the spaces was counted. It is assumed that the lectures and laboratory sessions took place during the measurement of EEC. Similar to lecturers' occupancy, students' occupancy will be zero from 12:00 a.m. to 7:00 a.m. and from 6:00 p.m. to 11:00 p.m. The outdoor temperature will be based on hourly historical data retrieved from www.weatherunderground.com using nearby satellite data. A sample of 24-hour tabulated data of the EEC and respective independent variables is illustrated in Figure 3 for clearer understanding.

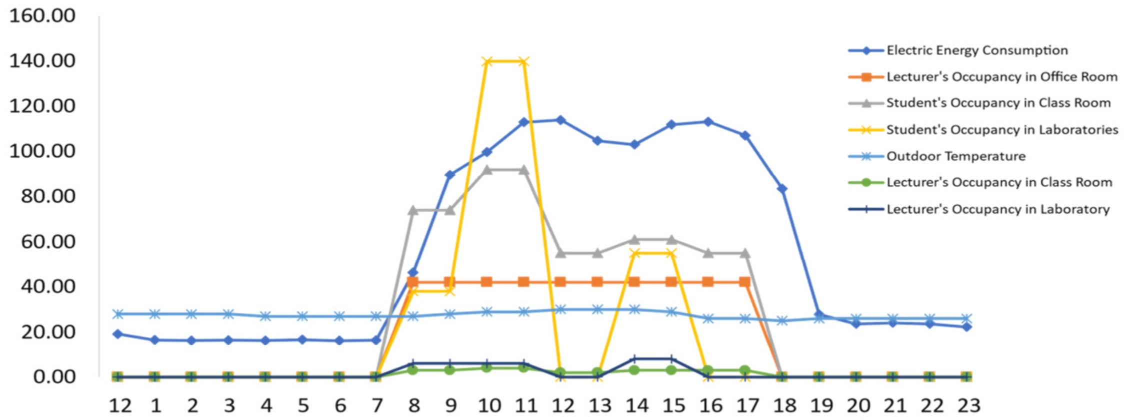

Fig. 3. Hourly EEC plot with the independent variables.

III. METHODOLOGY

The flowchart displayed in Figure 4 depicts the comprehensive methodology of this study, which aims to compare the performance of three ML models, i.e. LSTM, SVR, and AutoRegressive Integrated Moving Average with Exogenous inputs (ARIMAX) in order to compare the state-of-the-art of BEM for building EEC.

A. *Dataset Acquisition and Pre-Processing*

The independent input variables represent the most influential aspects of academic operations that could impact the EEC of the building that was acquired to develop the MLBEM. The input variables comprise time and date, staff occupancy, and the number of students in classrooms or laboratories. These reflect the occupancy levels and activities occurring within the building and weather conditions determined using temperature data. A total of 496 hours of data points were collected during the lecture weeks. The data were partitioned into training and testing sets, respectively, comprised 400 and 96-hours of data points. The testing set aimed to assess the accuracy and predictive capability of the trained models on unseen data.

Following the data collection process, the data were pre-processed prior to being incorporated into the ML models. During this step, the *mapminmax* method was used to normalize the independent data between -1 and 1, placing the data on the same scale and avoiding data domination during the learning process. In the subsequent process, the pre-processed data went on to model training, which was divided into three distinct paths for each ML model under consideration. The hyperparameters of each model were tweaked to achieve the greatest possible performance of BEM, and the models were initially evaluated to estimate the performance of the MLBEM model in the step ahead of 12 hours. Further analysis was conducted to examine the performance of the MLBEM to be tested for the next step ahead in 24, 48, 72, and 96 hours. The developed models were evaluated on how they learned from the training data and how they could generalize to data they had not previously encountered. In the final phase, a comparison of the models' overall performance with respect to the prediction outcome and actual data was conducted using a number of metrics. This includes residual error, correlation coefficient, and coefficient of determination.

B. *Model Configurations*

1) *Long Short-Term Memory (LSTM)*

An LSTM network is a Recurrent Neural Network (RNN) that is designed to handle sequential data and is suited for time series analysis and can be categorised as a DL model. It is a member of the class of ANNs that include a memory cell and gating mechanisms. It is efficient in identifying long-term dependencies in the data and learning complex temporal patterns.

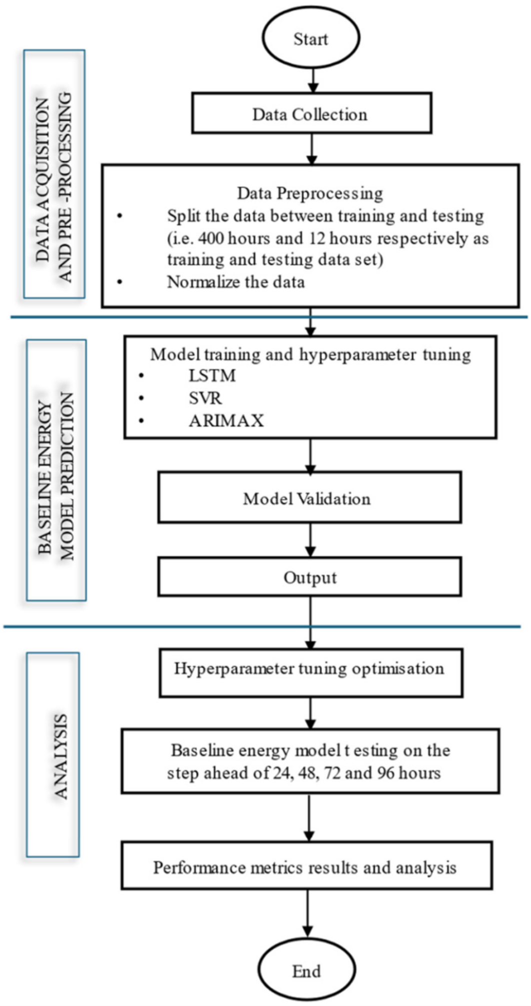

Fig. 4. Workflow diagram of the overall methodology proposed for MLBEM developed using the ML approach.

2) *Support Vector Regression (SVR)*

SVR is normally used in regression analysis, which extends the original method based on an SVM when developing the prediction model. One advantage of SVR is that it can capture linear and non-linear data relationships. It finds the most accurate estimation from the observed target values for all training data by identifying a hyperplane in a feature space with a high number of dimensions. The SVR algorithm seeks to minimize an epsilon-insensitive loss function and simultaneously maximize the margin surrounding the hyperplane to generate precise regression predictions.

3) *AutoRegressive Integrated Moving Average with Exogenous Inputs (ARIMAX)*

The ARIMAX model is one of the statistical time series models that extends the capabilities of the Autoregressive Integrated Moving Average (ARIMA) model. The additional component of independent variable X is incorporated into the ARIMA model. Accordingly, it could improve prediction model accuracy due to its ability to capture relationships between past observations and future values.

C. Evaluating Model Performance

A comprehensive evaluation of the prediction models related to building EEC is essential for quantifying the developed MLBEM in terms of model's accuracy, efficiency, and reliability. In this study, the performance is assessed through residual error metrics, namely Root Mean Square Error (RMSE), Mean Square Error (MSE), and Mean Absolute Error (MAE). Other metrics were also considered, including correlation coefficient (R), coefficient of determination (R^2), and training time. Parameters to formulate performance metrics involve yi as the actual data instance, $\hat{y}_i$ as the predicted data instance, $\overline{yi}$ as the mean of the actual data instance, $\bar{y}$ as the mean of the predicted data instance, and n as the number of data.

RMSE quantifies the model's accuracy by measuring the differences between the predicted values and the actual observations. It is particularly effective for assessing predictive models in EEC analysis, where a lower RMSE signifies a closer alignment between predictions and real-world data:

$$\text{RMSE} = \sqrt{\frac{1}{n}\sum\nolimits_{i=1}^{n}(y_i - \hat{y}_i)} \quad (1)$$

MSE represents the average squared difference between the estimated values and the actual data. It serves as a critical indicator of variance in the predicted EEC for the building, providing insights into the model's tendency to underpredict or overpredict:

$$\text{MSE} = \frac{1}{n}\sum\nolimits_{i=1}^{n}(y_i - \hat{y}_i)^2 \quad (2)$$

MAE quantifies the average magnitude of errors in MLBEM, thereby providing a straightforward measure of accuracy and predicted model reliability:

$$\text{MAE} = \frac{1}{n}\sum\nolimits_{i=1}^{n}\left|y_i - \hat{y}_i\right| \quad (3)$$

R evaluates the strength and direction of the linear relationship between predicted and actual values. In developing the MLBEM, a high correlation coefficient indicates the insights of a strong predictive relationship related to how well (or good fit) the prediction model aligns with the observed data:

$$\text{R} = \frac{n((\sum\nolimits_{i=1}^{n}(\hat{y}_i - \bar{y})(y_i - \bar{y}_i)}{\sqrt{(\hat{y}_i - \bar{y})^2} * \sqrt{(y_i - \bar{y}_i)^2}} \quad (4)$$

R-squared measures the variability between the predicted and the actual values, demonstrating how well the MLBEM fits the observed data points. This metric provides an intuitive understanding of the proportion of variance explained by the model in which high R-squared values indicate a better fit:

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y}_i)^2} \quad (5)$$

Training time refers to the duration required to train the model. This metric is crucial for evaluating the computational efficiency of different modelling approaches in the context of EEC prediction.

D. Hyperparameter Tuning

Each model's hyperparameters were tuned for optimal performance [29-31]. Table II presents a summary of various tuning hyperparameters.

TABLE II. HYPERPARAMETERS USED TO OPTIMIZE THE MLBEM

Model	Hyperparameter	Tuning range
LSTM	No. of hidden units	10,20,30,40,50,60
	Learning rate	0.0001,0.001,0.01,0.1
	Activation function	sigmoid
	Optimization	Adam
	Drop out	0.5
SVR	Epsilon	0.1,1,5,10,15,20
	Regularization parameter, C	0.1,1,2,3,4,5,6,7,8,9,10
	Kernel	Radial Basis Function (RBF)
ARIMAX	Autoregressive lags (ARLags)	1,2,3,4,5,6,7,8,9,10
	Differencing order, D	1,2,3,4,5
	Moving average lags (MALags)	1,2,3,4,5,6,7,8,9,10

To assess the best performance of the LSTM network, the number of hidden units and learning rate varied within a specified tuning range. These hyperparameters represent the learning capacity of the LSTM network and the rate at which the model updates its weights during training. Main hyperparameter tuning for the SVR model was conducted on two parameters, namely epsilon and C, which respectively signify the width of the margin of tolerance for errors and the trade-off between a training error and the margin within the defined range. It varies from 0.1 to 20 to understand its impact on the model's performance. The tuning process for the ARIMAX model focuses on AutoRegressive Lags (ARlags), D (order of differencing), and Moving Average Lags (MALags). The ARlags and MALags parameters indicate the dependency on past values and errors, respectively, and vary from 1 to 10. Meanwhile, the D parameter signifies the number of times the observations are differenced, varying from 1 to 5.

IV. RESULTS AND DISCUSSION

The most optimal hyperparameter tuning is initially examined on the 12-hour step ahead. The result is tabulated in Table III. The LSTM model performed best when configured with 40 hidden units and a learning rate of 0.01. The SVR model worked best with an epsilon of 0.1 and a C value of 10. The ideal parameters for the ARIMAX model were determined to be ARlags = 2, D = 1, and MALags = 5. These tuning parameters demonstrate that each model has unique setups that can improve its forecasting ability. Following the selection of hyperparameters, the resulting models for 12-hour forecasting were evaluated using the performance metrics and data provided in Table IV. The LSTM model had a higher MAE of 2.790 than the other two models. Nevertheless, the highest R^2 value of 0.700, R-value of 0.837, and training time of 56.32 s were recorded. This suggests that the LSTM model revealed a significant ability to explain the variability of the data set, even though the model has drawbacks in terms of computation model training.

TABLE III. BEST-TUNED PARAMETERS FOR 12-HOUR EEC PREDICTION

Model	Hyperparameter	Selected values
LSTM	No. of hidden units	40
	Learning rate	0.01
	Activation function	sigmoid
	Optimization	Adam
	Drop out	0.5
SVR	Epsilon	0.1
	Regularization parameter, C	10
	Kernel	RBF
ARIMAX	Autoregressive lags (ARLags)	2
	Differencing order, D	1
	Moving average lags (MALags)	5

Meanwhile, the SVR model demonstrated a promising compromise between error metrics and training time. It had the lowest RMSE, MSE, MAE, and training time values among the three models, with 1.914, 3.664, 0.866, and 0.686, respectively, indicating the best-fit and efficient training model. However, lower R^2 and R values, of 0.159 and 0.443, respectively, can be observed in SVR as compared to the LSTM model.

TABLE IV. PERFORMANCE METRICS FOR 12-HOUR EEC PREDICTION

Metrics	12-hours		
	LSTM	**SVR**	**ARIMAX**
RMSE	2.985	1.914	3.220
MSE	8.909	3.664	10.371
MAE	2.790	0.866	2.540
R	0.837	0.443	0.066
R^2	0.700	0.159	2.280
Training time	56.323	0.686	3.390

The ARIMAX model produced the highest error metrics, with RMSE and MSE values of 3.220 and 10.371. Although it reported the highest R^2 value of 2.280, its R of 0.066 was the lowest among the three models, exhibiting a poor correlation between the predicted and actual data instances.

The result performance of each model was further experimented with and tested using various metrics at different future step hours, i.e. 24, 48, 72, and 96, as displayed in Table V. The efficacy of the models was determined by their ability to correctly predict steps ahead of EEC. Each step hour demonstrates that different metrics recorded were inconsistently performed to predict the outcome of the MLBEM. Analysis of the residual error metrics was highly contributed by ARIMAX compared to LSTM and SVR, disclosing the weakest predictive relationship in terms of correlation analysis with the dataset. It can be observed that the ARIMAX model is unable to perform well in BEM. On the contrary, SVR's residual error metrics, specifically RMSE, MSE, and MAE, manifest the least error when compared between actual and predicted values of building EEC.

Similarly, SVR also demonstrates the shortest training time, indicating its superior efficiency. Therefore, it showcased a promising balance of prediction accuracy and computational efficiency throughout the testing of BEM for the 24- to 96-hours steps ahead. The R and R^2 values between LSTM and SVR are comparable most of the time. Although the LSTM model exhibited a strong capability to consistently explain the variability in the dataset, the longer training time introduced during the model training has presented a computational challenge, which may be due to the complex network layer.

TABLE V. PERFORMANCE METRICS OF EEC PREDICTION AT VARIOUS STEP-AHEAD HOURS

Metrics	24-hours		
	LSTM	**SVR**	**ARIMAX**
RMSE	7.516	3.411	39.491
MSE	56.491	11.638	1559.527
MAE	5.448	1.855	31.531
R	0.994	0.997	0.737
R^2	0.980	0.993	0.075
Training time	54.867	1.215	3.804
Metrics	**48-hours**		
	LSTM	**SVR**	**ARIMAX**
RMSE	8.231	3.400	34.241
MSE	67.745	11.562	1172.435
MAE	5.787	1.541	26.639
R	0.993	0.997	0.705
R^2	0.966	0.993	0.278
Training time	55.702	0.739	3.889
Metrics	**72-hours**		
	LSTM	**SVR**	**ARIMAX**
RMSE	8.434	3.621	40.348
MSE	71.140	13.108	1627.932
MAE	5.489	1.494	32.004
R	0.992	0.996	0.653
R^2	0.957	0.992	-0.030
Training time	56.479	0.742	3.566
Metrics	**96-hours**		
	LSTM	**SVR**	**ARIMAX**
RMSE	8.275	3.625	38.819
MSE	68.481	13.137	1506.890
MAE	5.671	1.545	30.106
R	0.990	0.996	0.611
R^2	0.959	0.992	0.039
Training time	56.277	0.718	3.851

Besides observing the BEM performances through residual error metrics and good fit model accuracy, visual insights through comparative graphical illustration can demonstrate the pattern of the BEM test performances, as depicted in Figure 5. When analysing the ARIMAX plotting of the predicted output of the BEM, the most apparent picture displays the large deviation or discrepancy to inaccurately produce the BEM, resulting in a weak fitted model. SVR plotting consistently tracks the actual building EEC throughout all hour steps. In addition, as the hour step increased, BEM with LSTM and SVM could exactly capture the pattern of peak values in building EC up to 72- and 96-hour steps, respectively. Visualisation of the training and testing dataset, along with its R and R^2 analysis for the SVR model, is presented in Figures 6 and 7. Considering the maximum step ahead of 96-hours, the relationship between the predicted and actual data instances of building EEC suggests a strong linear correlation, with a correlation coefficient of 0.996 ($R\approx1$) and a coefficient of determination of 0.992 ($R\approx0.99$). Based on this, it is evident that the SVR can efficiently model the EEC of the building for MLBEM studies.

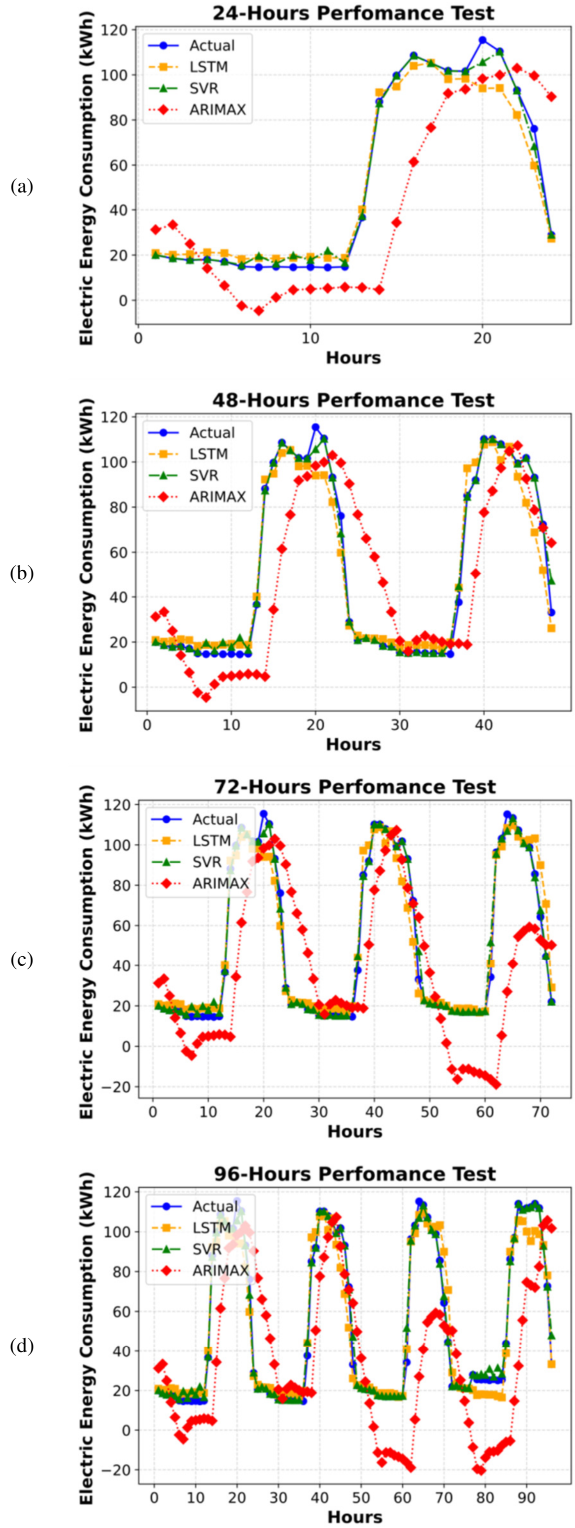

Fig. 5. MLBEM tested on varying step ahead hours performance at (a) 24, (b) 48, (c) 72, and (d) 96 hours.

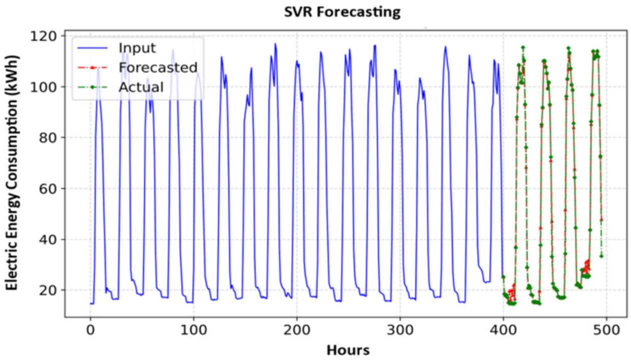

Fig. 6. Illustration of the best performance between actual and predicted EEC plotting.

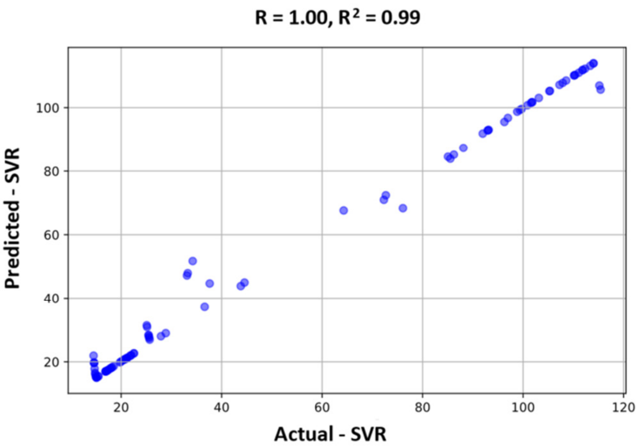

Fig. 7. The best R and R^2 metrics of the SVR model for MLBEM.

V. CONCLUSION AND FUTURE SCOPE

Comparative evaluations between deep learning, machine learning, and statistical models provide a comprehensive elucidation of how the model configurations influence the predictive performances and their variability in developing the predictive models for estimating EEC in an educational building based on the outdoor temperature, and the number of occupants in the building within the space of offices room, classrooms, and laboratories. Among the investigated models, SVR was discovered to be the most appropriate and effective for up to 96-hour predictions to develop as MLBEM that involves EEC prediction for building applications. This is due to the main attributes and well-balanced trade-off in terms of its high R and R^2 of at least 99% performances and the least computational training time. Overall, SVR consistently performed well as the model capability in capturing the patterns underlying the data instance in EEC.

These findings are promising for the advancement of the system in the future and continuous improvement in incorporating more variables. This includes building attributes and external influences, yielding a more comprehensive understanding of energy usage patterns and enhancing precision and dependability in EEC prediction. Thereby, the developed MLBEM is reliable. Moreover, the real-time integration of the MLBEM is achieved by continuously monitoring the building EEC, facilitating prompt identification of irregularities or excessive EEC. This real-time feedback enables the implementation of more proactive and targeted energy management strategies.

ACKNOWLEDGEMENT

The authors acknowledge the Research Management Centre (RMC) of Universiti Teknologi MARA (UiTM) for awarded grant 600-RMC 5/3/GPM (049/2022)) which supports our research work and provides financial assistance throughout the study.

REFERENCES

[1] "Electricity Information - Data product," *IEA*. https://www.iea.org/data-and-statistics/data-product/electricity-information.

[2] "World Power consumption | Electricity consumption | Enerdata." https://yearbook.enerdata.net/electricity/electricity-domestic-consumption-data.html.

[3] T. Hong, L. Yang, D. Hill, and W. Feng, "Data and analytics to inform energy retrofit of high performance buildings," *Applied Energy*, vol. 126, pp. 90–106, Aug. 2014, https://doi.org/10.1016/j.apenergy.2014.03.052.

[4] D. Kadric, R. Blazevic, H. Bajric, and E. Kadric, "Evaluation of Energy Renovation Measures for Hospital Buildings using the PSI Method," *Engineering, Technology & Applied Science Research*, vol. 14, no. 1, pp. 12753–12758, Feb. 2024, https://doi.org/10.48084/etasr.6558.

[5] F. A. AlFaraidy and S. Azzam, "Residential Buildings Thermal Performance to Comply With the Energy Conservation Code of Saudi Arabia," *Engineering, Technology & Applied Science Research*, vol. 9, no. 2, pp. 3949–3954, Apr. 2019, https://doi.org/10.48084/etasr.2536.

[6] A. Zerroug and E. Dzelzitis, "A Study of Modeling Techniques of Building Energy Consumption," *Engineering, Technology & Applied Science Research*, vol. 10, no. 1, pp. 5191–5194, Feb. 2020, https://doi.org/10.48084/etasr.3257.

[7] "National Energy Efficiency Action Plan (NEEAP) | ESCAP Policy Documents Managment." https://policy.asiapacificenergy.org/node/1269.

[8] "China released its 14th Five-Year Plan for Renewable Energy with quantitative targets for 2025 - Sino-German Cooperation on Climate Change, Environment, and Natural Resources." https://climatecooperation.cn/climate/china-released-its-14th-five-year-plan-for-renewable-energy-with-quantitative-targets-for-2025/.

[9] "NMEEE | BUREAU OF ENERGY EFFICIENCY, Government of India, Ministry of Power." https://beeindia.gov.in/en/programmes/nmeee.

[10] "Germany's Energy Efficiency Strategy 2050 - SUSTAINABLE TRANSITION CHINA." https://transition-china.org/energyposts/germanys-energy-efficiency-strategy-2050/.

[11] *International Performance Measurement & Verification Protocol*, Revised ed. Oak Ridge, TN, USA: U.S. Department of Energy, 2002.

[12] M. K. M. Shapi, N. A. Ramli, and L. J. Awalin, "Energy consumption prediction by using machine learning for smart building: Case study in Malaysia," *Developments in the Built Environment*, vol. 5, Mar. 2021, Art. no. 100037, https://doi.org/10.1016/j.dibe.2020.100037.

[13] W. N. W. Md Adnan, N. Y. Dahlan, and I. Musirin, "Development of Hybrid Artificial Neural Network for Quantifying Energy Saving using Measurement and Verification," *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 8, no. 1, pp. 137–145, Oct. 2017, https://doi.org/10.11591/ijeecs.v8.i1.pp137-145.

[14] B. V. Ayodele, S. I. Mustapa, N. Mohammad, and M. Shakeri, "Long-term energy demand in Malaysia as a function of energy supply: A comparative analysis of Non-Linear Autoregressive Exogenous Neural Networks and Multiple Non-Linear Regression Models," *Energy Strategy Reviews*, vol. 38, Nov. 2021, Art. no. 100750, https://doi.org/10.1016/j.esr.2021.100750.

[15] Y. Huang, Y. Yuan, H. Chen, J. Wang, Y. Guo, and T. Ahmad, "A novel energy demand prediction strategy for residential buildings based on ensemble learning," *Energy Procedia*, vol. 158, pp. 3411–3416, Feb. 2019, https://doi.org/10.1016/j.egypro.2019.01.935.

[16] Z. Dong, J. Liu, B. Liu, K. Li, and X. Li, "Hourly energy consumption prediction of an office building based on ensemble learning and energy consumption pattern classification," *Energy and Buildings*, vol. 241, Jun. 2021, Art. no. 110929, https://doi.org/10.1016/j.enbuild.2021.110929.

[17] T. Zhao, J. Xu, C. Zhang, and P. Wang, "A monitoring data based bottom-up modeling method and its application for energy consumption prediction of campus building," *Journal of Building Engineering*, vol. 35, Mar. 2021, Art. no. 101962, https://doi.org/10.1016/j.jobe.2020.101962.

[18] Y. Ahn and B. S. Kim, "Prediction of building power consumption using transfer learning-based reference building and simulation dataset," *Energy and Buildings*, vol. 258, Mar. 2022, Art. no. 111717, https://doi.org/10.1016/j.enbuild.2021.111717.

[19] R. Mena, F. Rodríguez, M. Castilla, and M. R. Arahal, "A prediction model based on neural networks for the energy consumption of a bioclimatic building," *Energy and Buildings*, vol. 82, pp. 142–155, Oct. 2014, https://doi.org/10.1016/j.enbuild.2014.06.052.

[20] H. R. Khosravani, M. D. M. Castilla, M. Berenguel, A. E. Ruano, and P. M. Ferreira, "A Comparison of Energy Consumption Prediction Models Based on Neural Networks of a Bioclimatic Building," *Energies*, vol. 9, no. 1, Jan. 2016, Art. no. 57, https://doi.org/10.3390/en9010057.

[21] B. Grillone, S. Danov, A. Sumper, J. Cipriano, and G. Mor, "A review of deterministic and data-driven methods to quantify energy efficiency savings and to predict retrofitting scenarios in buildings," *Renewable and Sustainable Energy Reviews*, vol. 131, Oct. 2020, Art. no. 110027, https://doi.org/10.1016/j.rser.2020.110027.

[22] L. Zhang *et al.*, "A review of machine learning in building load prediction," *Applied Energy*, vol. 285, Mar. 2021, Art. no. 116452, https://doi.org/10.1016/j.apenergy.2021.116452.

[23] D. Koschwitz, J. Frisch, and C. van Treeck, "Data-driven heating and cooling load predictions for non-residential buildings based on support vector machine regression and NARX Recurrent Neural Network: A comparative study on district scale," *Energy*, vol. 165, pp. 134–142, Dec. 2018, https://doi.org/10.1016/j.energy.2018.09.068.

[24] R. Lei and J. Yin, "Prediction method of energy consumption for high building based on LMBP neural network," *Energy Reports*, vol. 8, pp. 1236–1248, Jul. 2022, https://doi.org/10.1016/j.egyr.2022.02.071.

[25] S. A. Nabavi, A. Aslani, M. A. Zaidan, M. Zandi, S. Mohammadi, and N. Hossein Motlagh, "Machine Learning Modeling for Energy Consumption of Residential and Commercial Sectors," *Energies*, vol. 13, no. 19, Jan. 2020, Art. no. 5171, https://doi.org/10.3390/en13195171.

[26] G. Aruta, F. Ascione, O. Boettcher, R. F. D. Masi, G. M. Mauro, and G. P. Vanoli, "Machine learning to predict building energy performance in different climates," *IOP Conference Series: Earth and Environmental Science*, vol. 1078, no. 1, Jun. 2022, Art. no. 012137, https://doi.org/10.1088/1755-1315/1078/1/012137.

[27] P. A. Schirmer, I. Mporas, and I. Potamitis, "Evaluation of Regression Algorithms in Residential Energy Consumption Prediction," in *3rd European Conference on Electrical Engineering and Computer Science*, Athens, Greece, Dec. 2019, pp. 22–25, https://doi.org/10.1109/EECS49779.2019.00018.

[28] M. Faiq *et al.*, "Prediction of energy consumption in campus buildings using long short-term memory," *Alexandria Engineering Journal*, vol. 67, pp. 65–76, Mar. 2023, https://doi.org/10.1016/j.aej.2022.12.015.

[29] E. Beard, J. Brown, and L. Shahab, "Association of quarterly prevalence of e-cigarette use with ever regular smoking among young adults in England: a time–series analysis between 2007 and 2018," *Addiction*, vol. 117, no. 8, pp. 2283–2293, 2022, https://doi.org/10.1111/add.15838.

[30] T. Fu and X. Li, "Hybrid the long short-term memory with whale optimization algorithm and variational mode decomposition for monthly evapotranspiration estimation," *Scientific Reports*, vol. 12, no. 1, Dec. 2022, Art. no. 20717, https://doi.org/10.1038/s41598-022-25208-z.

[31] T. N. Tran, B. M. Lam, A. T. Nguyen, and Q. B. Le, "Load forecasting with support vector regression: influence of data normalization on grid search algorithm," *International Journal of Electrical and Computer Engineering*, vol. 12, no. 4, pp. 3410–3420, Aug. 2022, https://doi.org/10.11591/ijece.v12i4.pp3410-3420.