IJIE

Penerbit UTHM

INTERNATIONAL JOURNAL OF INTEGRATED ENGINEERING
ISSN: 2229-838X e-ISSN: 2600-7916

Vol. 18 No. 3 (2026) 510-522
https://publisher.uthm.edu.my/ojs/index.php/ijie

Investigating the Energy Consumption in Malaysian Educational Buildings: A Linear Regression Analysis for Baseline Energy Modelling

Norfakhira Mohd Nor[1], Mohd Azahari Razali[1]*, Nur Syafiqah Adha Narrudin[1], Masataro Suzuki[2], Nur Erma Nardiah Nazami[3]

[1] *Faculty of Mechanical and Manufacturing Engineering, Universiti Tun Hussein Onn Malaysia, Parit Raja, Batu Pahat, Johor, 86400, MALAYSIA*

[2] *Department of System Safety Engineering, Institute of GIGAKU, Nagaoka University of Technology, Nagaoka, Niigata, JAPAN*

[3] *Sanaiden Group Berhad, Level 8, Uptown 1, No. 1 Jalan SS21/58, Damansara Utama, Petaling Jaya, Selangor, 47400, MALAYSIA*

*Corresponding Author: azahari@uthm.edu.my
DOI: https://doi.org/10.30880/ijie.2026.18.03.033

Article Info

Received: 15 November 2025
Accepted: 13 March 2026
Available online: 30 April 2026

Keywords

Energy consumption, linear regression, independent variables, educational buildings, baseline energy model

Abstract

The baseline energy models currently developed for Universiti Tun Hussein Onn Malaysia (UTHM) using conventional linear regression fail to meet the International Performance Measurement and Verification Protocol (IPMVP) threshold of R^2>0.75, indicating insufficient explanatory power for reliable energy performance verification. The consumption of energy within the Malaysian universities is increasingly significant due to extensive facilities and growing student populations. This study analysed energy consumption patterns at UTHM and evaluated energy models using linear regression. Three independent variables included in the analysis were the number of students, the number of working days, and cooling degree days (CDD), with monthly consumption from 2022 to 2024 analysed. Simple Linear Regression (SLR) results indicated weak relationships, with correlation coefficients (R) ranging from 0.05 to 0.40. The strongest explanatory power was provided by the number of working days. The Multiple Linear Regression (MLR) enhanced the model performance, yielding coefficients of determination (R^2) between 0.25 and 0.43, explaining 25%–43% of the variance. However, these values remain below the IPMVP requirement. The findings highlight the limitations of traditional linear regression and suggest the need to incorporate additional operational variables and advanced predictive approaches, such as Artificial Neural Networks (ANN) and Long Short-Term Memory (LSTM) models, to improve prediction accuracy and institutional energy planning. These findings provide useful guidance for improving energy management and conservation practices in educational institutions.

1. Introduction

Malaysian universities consume a considerable amount of energy, particularly electricity. This is primarily because of the extensive facilities and the number of students and staff. The utility of energy in university construction is emerging as a big problem in the country. It contributes a lot to the overall energy requirement and the running expenses of these institutions. Academic structures in Malaysia tend to utilise approximately two-thirds of their energy in their mechanical ventilation system, water heating, and cooling system [1]. The remaining third is spent on lighting and other electrical appliances. In university buildings, lighting by itself constitutes 10 to 30 percent of electricity consumption [1]. Malaysian public universities typically contain a lot of structures that may arise in their usage, such as administration buildings, library buildings, student centres, cafeterias, and residential buildings (dorms). An example is the UiTM Perak campus, which has forty-eight buildings, and its total floor area is approximately 82,848 m^2 with the capacity to house approximately 5,000 people, comprising students and staff [1]. Energy consumption in older buildings will be higher due to ageing equipment and infrastructure, as is the case in Universiti Utara Malaysia (UUM) [2]. Every year, the government invests billions of Malaysian Ringgit in the electricity of the public universities. There is one report that the country spends RM 2.7 billion a year on electricity use across universities [1]. Individual buildings such as the Chancellery consumed 1.6 million kWh at UUM in 2018, and this amounts to approximately RM 840,000 [2]. There is also heavy power consumption in student residential halls. Major halls, such as Inasis Bank Rakyat, consume more than 820,000 kWh each year [2].

Studies at the UiTM Tapah campus demonstrated that the university could reduce its Building Energy Index through ECM to 135 kWh/m^2/year. This implies greater energy efficiency [1]. In this regard, it is possible to classify Energy Conservation Measures (ECM) into no-cost, medium-cost, and high-cost strategies that focus on electrical as well as mechanical systems [1]. The energy-saving measurement and verification can be implemented according to the scheme presented in the International Performance Measurement and Verification Protocol (IPMVP) [3]. Also, one more policy was introduced to facilitate energy efficiency measures as a part of overall ECMs to cut energy use, namely the National Energy Efficiency Action Plan (NEEAP), which was proposed in 2015-2025 [4]. Some of the universities, such as Universiti Teknologi Malaysia (UTM), have taken campaigns to create awareness on energy management. They have also installed energy-efficient lights and installed systems that monitor energy consumption. Such initiatives have seen a fall in the energy consumption and expenditure by 21.6 percent in the last two years between 2009 and 2011 [2]. The conduct of the students and the staff significantly influence the use of energy. Research results reveal a moderate degree of energy awareness, and energy-saving practices are detectable among the Malaysian university students [5]. The greater the awareness, the more energy savings that will be achieved. Making things simple, such as encouraging people to switch off lights and computers, among others, when they are not in use, will help save on costs and energy consumption [5]. The artificial neural networks and other data-driven models are being designed that will be used to forecast energy consumption in university buildings to enhance it. These models assist the facilitation managers to plan and implement energy-saving measures in a more effective manner [6].

Universiti Tun Hussein Onn Malaysia (UTHM) uses electricity as the main energy source since it serves as the source of lighting and air-conditioning systems. Such systems consume a large fraction of the overall energy usage in university buildings [7]. This study discusses the details of measuring and analysing energy consumption in various buildings of UTHM. This includes lecturer rooms, classrooms, and residential colleges. As an example, the G3 Building recorded the most energy expenditure per month, amounting to RM44,979.22. That was followed by residential colleges and faculty blocks [7]. The data on the consumption of energy in 2018-2020 were selected to evaluate UTHM activities in comparison to faculties and facilities. It was revealed that buildings equipped with engineering laboratories consume more energy than those with computer laboratories since the former need more ventilation and air conditioning [8]. The importance of the Building Energy Index (BEI) as an important measure to determine the energy performance of a building has also been mentioned by other recent studies. As an example, the residential colleges like Kolej Kediaman Tun Dr Ismail (KKTDI) and Kolej Kediaman Tun Fatimah (KKTF) have registered lower BEI values than the national standard of 135 kWh/m^2/year, which shows that they perform better in terms of energy performance. This can be partly blamed on less occupancy as a result of the Movement Control Orders caused by Covid-19 [9]. It is important to note that the period from 2022 to 2024 represents a post-pandemic operational transition phase. The year 2022 reflects a partial recovery following Malaysia's Movement Control Orders (MCO) with hybrid learning and reduced campus occupancy. In 2023, campus activities were largely normalised, while 2024 shows full operational recovery with increased laboratory usage and research activities. These operational shifts may significantly influence energy consumption behaviour and model stability.

A regression model analysis of energy consumption in a government university usually involves using statistical methods to understand and predict how different factors affect the university's energy use. Linear regression (LR) models are also a popular choice for predicting energy consumption [10]. Because of their simplicity, these regression-based models are preferred for predicting energy use. Linear regression (LR) was popular in the past literature [11-14], especially following its inclusion in the IPMVP standard as an accepted approach in the determination of baseline energy models and energy use projections. These models are constructed using independent variables which are normally well comprehended and measurable. Nonetheless, when using LR to forecast the energy consumption of buildings, it is necessary to note that the findings may vary due to the differences in characteristics of buildings, geographical location [15], and load characteristics [16]. Selecting appropriate independent variables is a critical step in developing a reliable prediction model. This is because a building's energy consumption is closely linked to the activities and conditions represented by those variables. The process of identifying and choosing suitable independent variables, therefore, requires careful consideration. Different building types, climatic regions, and load profiles may involve different variables that directly influence energy use. If inappropriate variables are selected, the resulting predictions are likely to be inaccurate. Although many studies report highly accurate energy prediction results, relatively few provide an in-depth discussion on how the chosen independent variables actually influence energy consumption. One recent study addressing independent variables was reported in [6], focusing on an institutional building at Multimedia University, Malacca Campus. However, its scope was limited to that specific case study. The paper reviewed the issue of predicting energy consumption in campus buildings at Multimedia University (MMU) based on Long Short-Term Memory (LSTM). Besides the prediction model, the relative importance of the various independent variables, including weather-related factors, was also investigated in the study. The results showed that both environmental and schedule-related variables play a significant role in influencing energy use.

Furthermore, several environmental parameters were found to be significant in enhancing the accuracy in prediction, such as air pressure, temperature, humidity, wind velocity, rainfall, duration of rainfall, and the amount of rainfall. Based on the Shapley Additive explanation (SHAP) to assess the feature analysis, the paper discovered that temperature, wind speed, rain duration and quantity of rainfall had a positive impact on the performance of the model. Among these, temperature and rainfall emerged as especially important variables during the model training process. Previous research also indicates that rainfall, particularly when accompanied by wind, can reduce cooling energy demand in hot climates. In addition to environmental factors, operational conditions were found to strongly influence model accuracy. The most significant parameters were found to be the type of lockdown (no Movement Control Order (MCO) and Recovery MCO) and the type of day (weekday, weekend, or holiday). Energy consumption typically decreases significantly during weekends and public holidays, often by 15–20%, as buildings are either closed or operating with much lower occupancy levels.

Under the present scenario in UTHM, the value of the coefficient of determination (R^2) reported in the literature has been found to be less than 0.75. Given that an R^2 near 1.0 is typically necessary to guarantee high explanatory power, this small value indicates that the existing models are performing an analysis on common factors that have been previously done by other researchers, such as the number of students [17], the number of working days [18] and cooling degree day (CDD) [19]. Although other variables such as floor area, equipment density, building age or laboratory types are recognised in literature as significant contributors to energy use, this study intentionally evaluates commonly used operational variables first. This approach allows assessment of whether traditional predictors alone are statistically sufficient before expanding toward more complex structural variables. Besides that, this low value shows that the current models are not explaining sufficient variation in the data. In the specific context of UTHM, previously developed regression-based baseline models consistently reported coefficients of determination below 0.75, which is the minimum acceptable value recommended by IPMVP for reliable performance verification. This statistical underperformance provides the main motivation of the present study, which aims to evaluate whether commonly used operational variables are sufficient to construct an acceptable baseline energy model for UTHM.

The primary aim of the research is to examine the energy consumption pattern within UTHM's main campus and to consider the variation in energy consumption based on the academic calendar. Besides this primary objective, the research also aims to identify and assess the independent variables that could affect energy consumption in educational buildings and to develop a baseline model of energy use. A baseline energy model is an important tool for energy management staff, as it allows them to monitor and discuss changes in a building's energy consumption when its operation or structure is modified [20]. A model like this will assist the owners of buildings and energy managers, as well as utility providers, in their planning and implementation of effective energy management strategies. A well-designed baseline model is very useful when it is created based on the appropriate structure and analysis of the consumed energy data and other related independent variables to make informed decisions. When making the baseline model, it is necessary to define which independent variables are the most suitable that may influence energy consumption. This study employs both single linear regression (SLR) and multiple linear regression (MLR) models. The performance of the MLR model will be assessed using the correlation coefficient (R) and the coefficient of determination (R^2). Data on monthly energy usage will be gathered from a university building, along with selected independent variables that are likely to influence energy consumption. The data will be synchronised with the same time periods as the recorded energy data. The impact of the independent variables will be examined through SLR and MLR techniques, followed by a correlation analysis. The final step involves comparing results from both models to gain insights into how the chosen independent variables affect energy consumption in UTHM buildings.

2. Methodology

To achieve the aims of this study, several methodological procedures have been implemented. The first step in the research process is to choose an educational facility to be the case study. The building on the main campus of the Parit Raja branch of Universiti Hussein Onn Malaysia has been selected as the research case. The independent factors influencing the building's energy use were selected based on commonly used UTHM operational variables that affect energy consumption. The energy usage dataset was obtained from verified Tenaga Nasional Berhad (TNB) utility billing records provided by the Sustainable Campus Office (SCO). After developing the baseline energy model using a multiple linear regression (MLR) approach, the resulting model was applied to estimate energy consumption. All statistical analysis for the SLR and MLR models was carried out using the Excel Data Analysis ToolPak. Data preprocessing and visualisation were undertaken with Microsoft Excel to guarantee reproducibility. Based on the IPMVP standard, the coefficient correlation (R^2) value's minimal threshold will be determined. In the next subsections, several of the methodology's explanations will be expanded upon (Fig. 1).

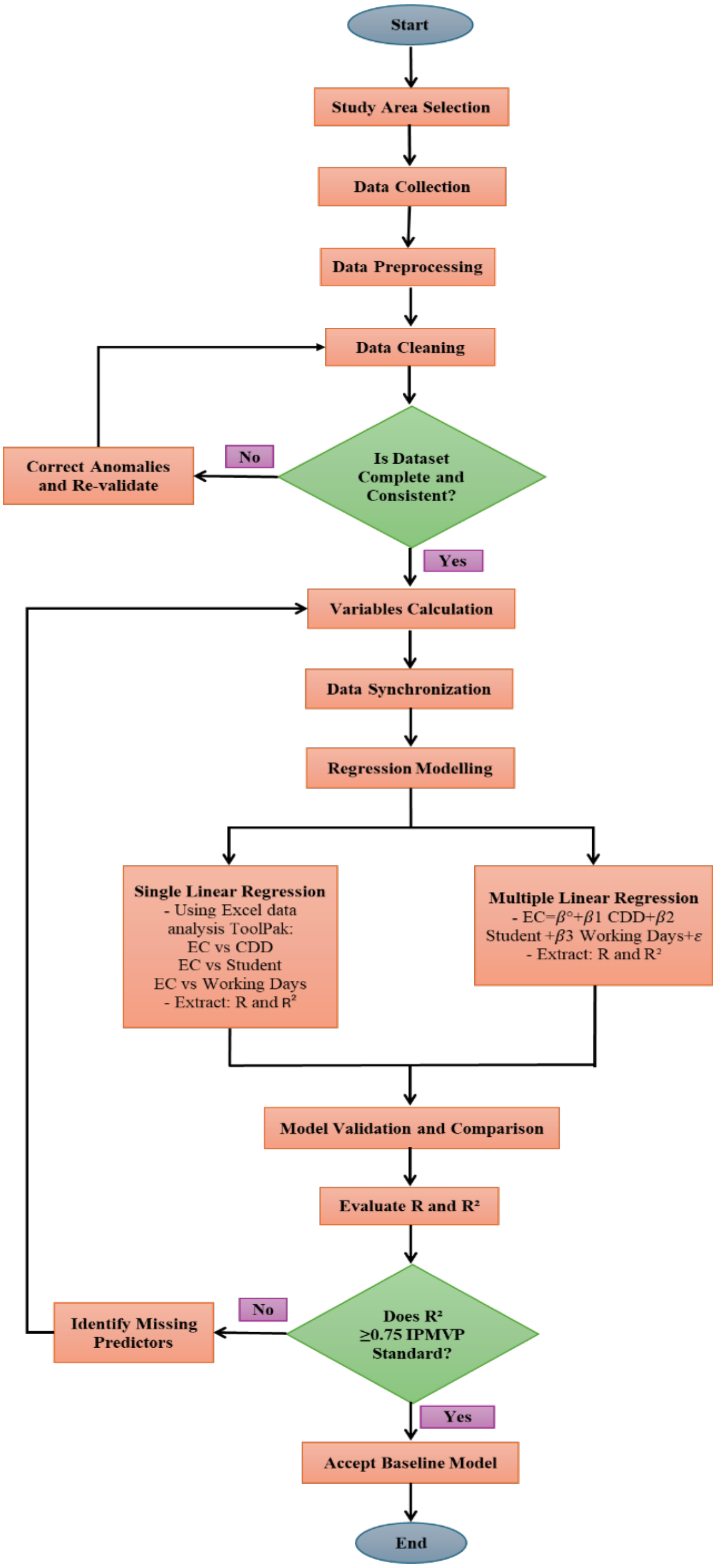

Fig. 1 *Methodology of the present study*

2.1 University Building

To achieve the objectives of this study, several methodological stages were carried out. The first step involved selecting an appropriate educational building as the case study. This research focuses on a main campus building at Universiti Tun Hussein Onn Malaysia (UTHM). Fig. 2 illustrates the layout of UTHM's main campus. It is situated at approximately 2°N latitude and 103°E longitude in Parit Raja, Batu Pahat, Johor, Malaysia. The UTHM main campus has 330,139.31 m^2 of areas of university buildings. Its main campus includes several important buildings, such as the Tunku Mahkota Ismail Hall, the Sports Complex, the Sultan Ibrahim Mosque, the Sultan Ibrahim Hall, and the Tunku Tun Aminah Library. To make management and identification easier, the facilities are divided into three main categories, such as academic, administration and residential college. Fig. 3 below presents a pie chart of the categories of facility buildings in UTHM. The largest fraction (58%) is comprised of academic buildings, which indicates the significance of teaching and learning infrastructure to the institution. The rate of administrative buildings is 37%, and the buildings have the required support and management services. In the meantime, residential colleges make up just 5% and serve as a type of student accommodation.

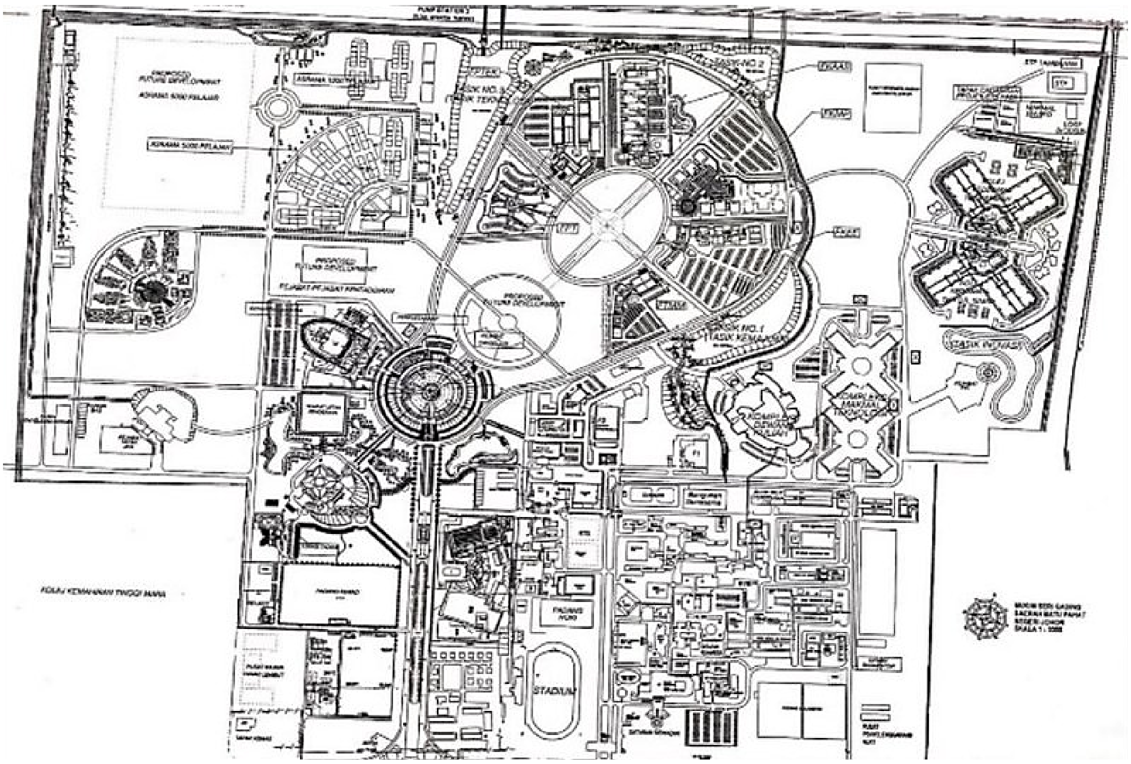

Fig. 2 *Layout of the main campus of UTHM*

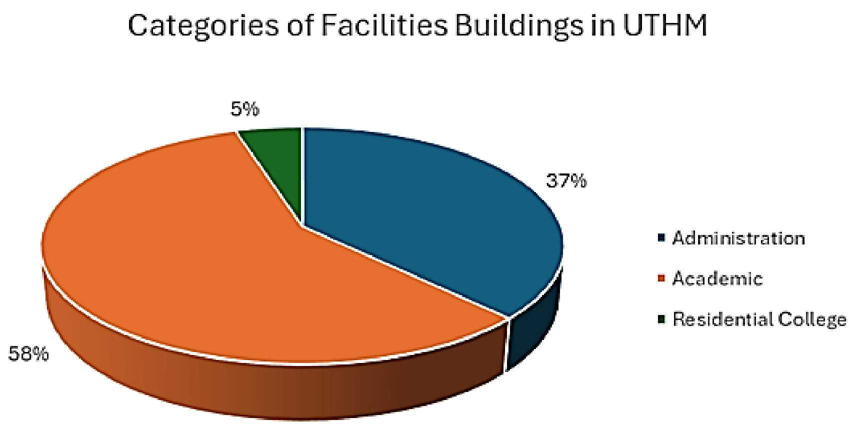

Fig. 3 *Distribution of categories of facilities buildings in UTHM*

2.2 Energy Consumption Measurement

Historical energy consumption data (preferably electricity in kWh) will be collected ideally at monthly intervals, spanning at least three years from 2022 until 2024 to capture seasonal and operational variations. The overall campus electricity consumption data was obtained from verified Tenaga Nasional Berhad (TNB) utility billing records provided by the Sustainable Campus Office (SCO) at Universiti Tun Hussein Onn Malaysia (UTHM). These records represent the official energy consumption data for the study period (2022-2024).

2.3 Independent Variables

This study was conducted to identify the factors influencing energy consumption in university buildings to develop a predictive model of energy use. The independent variables were selected based on the commonly used UTHM operational variables that initially impact energy consumption in the building. The factors identified as substantially affecting the energy consumption by university buildings on the main campus of UTHM are as follows:

- Number of students. The number of students represents the total registered on-campus enrolment per semester. It does not directly measure daily attendance; therefore, variations in actual daily occupancy may contribute to unexplained variance in the regression models.
- Number of working days
- Cooling degree day (CDD). The cooling degree day is an independent variable that measures the daily average temperature that deviates from the base temperature, which is fixed at 24°C. The base temperature of 24°C was selected because it corresponds to typical HVAC indoor setpoints used in Malaysian institutional buildings, and it aligns with tropical thermal comfort standards. The Malaysia Meteorological Department (MetMalaysia) provided the daily average temperatures, which are used to calculate cooling degree days (CDD).

2.4 Linear Regression Model

A linear regression model describes the relationship between a system's input and output using a straightforward mathematical approach. It fits a regression line to the data to represent and quantify the connection between the input and output variables. A single linear regression (SLR) model is shown in Eq. (1). In Eq. (1), y indicates the dependent variable, or the output of the model [17]. The term β_0 signifies the intercept of the regression line, while β_1 denotes the coefficient linked to the independent variable (x_1). The term $\mathcal{E}$ refers to the model's error component, which accounts for unexplained variations.

$$y = \beta_\circ + \beta_1 x_1 + \varepsilon \tag{1}$$

Multiple Linear Regression (MLR) is a statistical technique employed to investigate the relationships between one dependent variable and two or more independent variables. This method formulates a linear equation based on observed data and assumes that the relationship between the independent variables and the dependent variable is linear. By utilising this approach, researchers can evaluate the effect of each independent variable on the dependent variable while controlling for the other variables. Generally, a multiple linear regression model can be mathematically expressed, as demonstrated in Eq. (2) [17].

$$y = \beta_\circ + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \varepsilon \tag{2}$$

The performance of both the Simple Linear Regression (SLR) and Multiple Linear Regression (MLR) models will be evaluated using the correlation coefficient (R) and the coefficient of determination (R^2). The values of these metrics will be compared to assess the effectiveness of each model. Additionally, the R^2 values for both models will be discussed in relation to the standards set by the International Performance Measurement and Verification Protocol (IPMVP) [3].

3. Results and Discussions

3.1 Energy Consumption Pattern

Energy consumption is considered one of the most significant indicators of campus functioning because it not only shows the number of activities that occurred but also the intensity of resource use. The 2022-2024 data on the monthly energy consumption of UTHM's main campus in Fig. 4 show clear trends of growth, variation across different months, and semester factors affecting overall demand. The most noticeable feature about the data is that energy consumption is steadily increasing between 2022 and 2024. Almost every month, 2022 has the lowest levels of consumption, then 2023 shows moderate increases, and the highest value is in 2024. This positive pattern is a strong indication of increased energy demand in UTHM over the long term. This trend may be attributed to various factors such as campus growth, a rise in the number of students, an increase in academic and research activities and the use of more energy-consuming technologies and facilities. This phenomenon is a strong pointer to the fact that UTHM is witnessing increased growth in its operating capacity, and consequently, this growth implies increased energy consumption.

The total trend is positive, but energy consumption is unevenly spread across the year. Clear variation can be seen during the months, showing the effect of semester and institutional variability. In the early months (January-March), the level of energy consumption starts moderately in January every year, and in 2024, the increment is impressive when compared to the previous years. February always displays a decline, possibly owing to less academic activity in this month, perhaps because of semester holidays or fewer operations. In March, demand goes down, and 2024 is marked with very high demand compared to the past years. Between April and August, there is a clear peak seen in all three years around June and July. This is a time when energy consumption values are always at their highest, especially in 2023 and 2024, when the values are almost at the reach of three million kWh. This may be due to a confluence of high academic activity, warmer weather resulting in greater use of air conditioning and perhaps academic or research courses. In August, though, it begins to decline, a reaction to a peak of activity in the middle of the year.

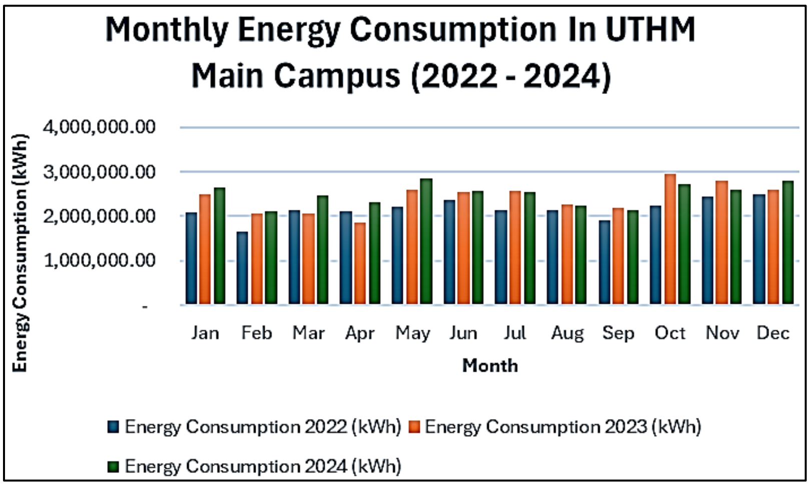

Fig. 4 *Monthly energy consumption from 2022 until 2024 (kWh)*

Late in the year (September to December), September represents a minimum energy use in the three years, which suggests a downturn in campus activities, potentially because of the semester transitions. But October and November are the months with sharp increases, especially in 2023, when consumption reaches almost three million kWh. December is followed by steadily high use in 2024, which contrasts with more moderate figures in previous years. This trend shows that the months at the end of the year are becoming more energy demanding, probably because of the increased academic programmes, academic exams, or institutional event preparations.

In conclusion, the energy consumption pattern of the UTHM main campus between 2022 and 2024 presents the picture of an evolving energy consumption pattern, semester fluctuations, and the rising reliance on electricity to support the campus operations. The steady increase over the years points to an expanding and increasingly active campus, and the frequent peaks in the middle and late months of a year point to the significance of institutional and climatic reasons. This increase in energy demand will be instrumental in ensuring the sustainability, efficiency and cost-effectiveness of the UTHM as the institution expands in the coming years.

3.2 Single Linear Regression Model

The findings have been summarised regarding three input variables: Cooling Degree Days (CDD), the number of students, and the number of working days. These variables are examined in connection with energy consumption during the period from 2022 to 2024, as presented in Table 1. The simple linear regression (SLR) analysis yields the coefficients of correlation (R) and determination (R^2) for each variable, as detailed in the table. The weakest performance is shown by the variable EC_CDD among the three independent variables. Over the three years, the correlation coefficients vary between 0.05 and 0.32, while the coefficients of determination are negligible (0.00 to 0.10). In 2022, R=0.17 and R^2=0.03 show a very weak positive relationship; CDD only explains 3% of the variance in energy consumption. This means that cooling demand due to weather had little influence on the energy use.

While in 2023, R=0.32 and R^2=0.10 show that the correlation became stronger, and CDD explained 10% of the variance. This implies a stronger correlation between weather conditions and energy consumption in 2023, which may be attributed to warmer weather or more frequent use of air conditioning.

Subsequently, for 2024, R=0.05 and R^2=0.00 show the correlation is nearly zero with no explanatory power. That means that CDD had almost no effect on the annual energy consumption in the year 2024, which may be explained by certain factors about operations or behaviour that overcame the effects of climatic factors. Although Malaysia is located in a tropical climate with consistently high ambient temperatures and solar irradiance, Cooling Degree Days (CDD) exhibited negligible explanatory power in 2024. The negligible explanatory power of Cooling Degree Days (R^2=0.00) in 2024 does not imply the absence of cooling demand but rather reflects the limited variability of tropical temperature conditions, saturation of HVAC base load operation, and the dominance of internal gains over climatic influences. Additionally, monthly data aggregation may have weakened short-term temperature-energy relationships. These findings suggest that in tropical institutional buildings, operational and occupancy-related variables may exert greater influence on energy consumption than external climatic indicators. This means that EC_CDD does not have much meaningful linear association with the output variable. Its explanatory ability is highly insignificant, showing that EC_CDD is either irrelevant as an explanatory variable in this model or that its effect is non-linear and cannot be easily expressed in a non-complicated regression model.

Table 1 *Correlation coefficient (R) and the coefficient of determination (R^2) for each variable within a Simple Linear regression (SLR) model*

Output_input	2022		2023		2024	
	(R)	(R^2)	(R)	(R^2)	(R)	(R^2)
EC_CDD	0.17	0.03	0.32	0.10	0.05	0.00
EC_no of students	0.27	0.07	0.19	0.04	0.23	0.05
EC_number of working days	0.38	0.15	0.40	0.16	0.29	0.09

The variable EC, the number of students, illustrates an even better trend over time. In the case of 2022, it can be said that the correlation coefficient is 0.27, and the determination coefficient is 0.07, showing that there is a weak positive correlation with the explained variance of 7%. Energy demand was slightly affected by the student population. In 2023, the correlation is 0.19, and the R^2 is 0.04. Although the population increased, the number of students did not correlate closely with the energy consumption changes, possibly because of efficiency gains or changes in the use of the facilities. In 2024, the correlation is virtually at 0.23, and the R^2 is literally 0.05. The correlation increased a bit, yet the explanatory power was low at 5 percent. This implies that although students are also a source of energy consumption, their contribution on its own is not high. Students are correlated with energy consumption in a consistent but weak relationship, suggesting that population growth is not the exclusive factor leading to an increase in energy demand; instead, other drivers can be more important.

Using the three identified variables, EC, and the number of working days, simply shows the best relationship with the dependent variable. In 2022, R=0.38 and R^2=0.15 show the highest correlation among all predictors, explaining 15% of the variance. This highlights the evident contribution of operational schedules toward energy utilisation. In 2023, it is also important to note that the correlation is 0.40 and R^2 is 0.16, which shows that working days are the only accounting factors that explain 16% of the variance in the result. As can be seen in this trend, the working days have a consistent effect on the outcome variable. While in 2024, R=0.29 and R^2=0.09 show the correlation falls greatly below the first, which explains only 9% of the variance, meaning that working days continue to have an impact, but other variables might have been more influential in 2024. The number of working days has the strongest correlation and explanatory power within every three years, which again demonstrates that it is the most relevant single predictor of the three variables with which it is considered.

The results of the SLR model show weak to moderate correlations, with an average R value of less than 0.40 and an average R^2 value of less than 0.20. This shows that none of the predictors (CDD, student numbers, and working days) can be fully attributed to annual energy consumption variations within the UTHM campus. The number of working days is the strongest predictor of energy demand, highlighting the operational calendar as a crucial factor. The effect of Cooling Degree Days was not consistent, indicating that weather may have an effect on the daily energy changes, but does not have a strong effect on annual consumption. The students' population has a low contribution, meaning that population growth does not directly correlate to similar proportional growth in the use of energy.

According to the International Performance Measurement and Verification Protocol (IPMVP), a reliable regression model for energy performance verification should attain an R^2 value greater than 0.75 or at least approach this benchmark. These benchmarks are crucial to ensure that the calculated energy savings are both credible and defensible. However, when each independent variable is modelled separately, none of them provides sufficient explanatory power to meet the IPMVP standard. This indicates that energy consumption cannot be adequately explained by a single factor alone. Instead, more advanced modelling approaches, such as multivariable regression, interaction effects, or non-linear models, are necessary to capture the complexity of influencing factors. The relatively low R^2 values further suggest that important explanatory variables are missing and should be identified in future research.

3.3 Multiple Linear Regression Model

The results presented in Table 2 offer an overview of the effectiveness of the MLR model in predicting annual energy consumption for the period from 2022 to 2024. However, as opposed to the SLR, the MLR method assesses many independent variables at once. This approach is more proper because, in the context of energy consumption, it is difficult to expect one factor to prove and convince its influence, thus affecting energy consumption. The table presents annual consumption figures as well as two important model evaluation statistics, the coefficient of correlation (R) and the coefficient of determination (R^2). The records show that energy has continued to be used on a yearly basis, resulting in an annual rise of energy consumption to 25,958,962 kWh in 2022, 29,006,176 kWh in 2023, and 30,020,457 kWh in 2024. This trend of an increase in energy demand points to an absolute growth in energy demand over time. These rises can be attributed to institutional growth, more occupancy, high use of technology or external reasons like climatic conditions. The consistent increase illustrates the importance of effective predictive models because energy planning and resource allocation will have to be synchronised with the increase in demand.

The values of the coefficients of correlation in 2022, 2023 and 2024, which appear in Table 2, are 0.50, 0.66 and 0.51, respectively. Overall, the reported values indicate a moderate positive relationship between the selected predictors and energy consumption. For instance, in 2022, the correlation coefficient (R) of 0.50 suggests a moderate positive association between the independent variables and energy consumption. While the predictors show some relationship to consumption, the correlation is not particularly strong. R=0.66 in 2023 reflects a higher positive relationship than in 2022, indicating that the predictors were closer to the real consumption trends in this year. In 2024, R=0.51 indicates a moderate positive correlation, not as strong as in 2023 but slightly stronger than in 2022. This shows there is a weak predictive relationship, but weaker than it was in the previous year. The MLR model can be considered a reasonable success in aligning correspondences between the predictions and the actual energy consumption, the correlation being far from perfect. The consistency of these values over successive years hints that there is a more or less constant correlation between the selected independent variables and energy consumption, implying that the model used is capturing some sort of persistent drivers of energy use. Nevertheless, the finding that R is not greater than 0.75 also shows that the model has a clear limitation in the prediction strength.

Table 2 *Correlation coefficient (R) and the coefficient of determination (R^2) for 2022-2024 within a Multiple Linear Regression (MLR) model*

	Year		
	2022	2023	2024
Annual energy consumption (kWh)	25,958,962.00	29,006,176.00	30,020,457.00
Coefficient of correlation (R)	0.50	0.66	0.51
Coefficient of determination (R^2)	0.25	0.43	0.27

The coefficient of determination (R^2) shows the percentage of the variance in the dependent variable which is due to the regression model. The increased values of R^2 indicate that the predictors are more useful to explain the fluctuations in energy consumption. In this analysis, the R^2 in 2022, 2023 and 2024 were 0.25, 0.43 and 0.27, respectively. It implies that the MLR model can be used to explain around 25-43% of the annual energy consumption variation. In particular, the R^2 value of 0.25 in 2022 means that the independent variables included in the model accounted for only 25 percent of the changes in energy use, which implies that other kinds of factors that were not controlled had a stronger influence on the yearly energy demand in 2022. In contrast, the R^2 value of 0.43 in 2023 shows that the model explained 43% of the variance, representing a noticeable improvement in predictive performance compared to 2022. This shows that the independent variables used were more effective in explaining the drivers of energy demand, perhaps because of more consistent operational patterns or better data quality. According to the R^2 value of 0.27 in 2024, it captures 27% of the variance, which is less than in 2023. Despite an improvement compared to 2022, the comparatively low explanatory power shows that the influence of external or unmeasured factors on energy consumption in 2024 was more significant. This can indicate fluctuation in climate, changes in campus activities, or unusual changes in the use of facilities.

The correlation coefficient (R) and coefficient of determination (R^2) both improved from 2022 to 2023, suggesting stronger predictive power and better alignment of the regression model during this period. However, both values declined in 2024, which may indicate increased unpredictability in consumption or the influence of external variables not captured in the model. Overall, the model demonstrates that multiple variables together offer a better explanation of energy demand than single-variable regressions (SLR), but its predictive ability remains moderate, with less than half of the variance explained even in the strongest year (2023).

The model needs to satisfy certain conditions to be considered a valid linear regression model that can be used to perform a minimum amount of energy modelling, as well as to determine the energy savings. The guidelines to this process are provided by the International Performance Measurement and Verification Protocol (IPMVP), which is a globally recognised framework. The determination of the coefficient, R^2 shows how close the correlation is between energy use and its dependent variables. The IPMVP specifies that regression models used for energy savings verification should typically have an R^2 value greater than 0.75. R^2=1.0 shows a perfect correlation, while R^2>0.75 is an acceptable value for correlation. This standard level is what makes the model capable of capturing most of the variation in energy use and, therefore, serves as a good baseline on which to measure energy savings.

Table 2 highlights that the MLR model of UTHM's main campus energy consumption indicates a changing predictive performance over the years 2022–2024. The model was most effective in the year 2023, with an explanatory power of 43% and a strong correlation (R=0.66); however, the model exhibited a declining explanatory power in the year 2024, even though the absolute consumption was high. The decline in explanatory power in 2024 may be attributed to operational factors not captured in the model, such as increased engineering laboratory utilisation, expansion of research instrumentation, and extended operational hours toward the end of the year. These high-load activities introduce variability beyond occupancy and calendar-based predictors. The existing model can provide valuable baselines but cannot be relied on to provide exact forecasts or precise policy planning. This weakness points to the importance of more detailed methods of energy modelling within UTHM. More factors (temperature, occupancy, technological efficiency, etc.) will need to be included in future research, and more complex methods like non-linear models or machine learning will be investigated. These would allow making decisions more accurately and aid in making better decisions. Further research should consider incorporating wider and more dynamic variables to better account for the unexplained variance to make more exact forecasts and achieve better energy management.

4. Conclusions

In conclusion, the energy consumption trend within Universiti Tun Hussein Onn Malaysia (UTHM) was analysed, and the baseline energy models were investigated through the application of linear regression. The analysis of monthly consumption data in 2022-2024 showed that the energy demand is growing continuously, affected by academic activity, climatic conditions, and institutional growth. Simple Linear Regression (SLR) models were found to have poor explanatory power, and the number of working days was found to be the best independent variable. It is still not sufficient to qualify under the International Performance Measurement and Verification Protocol (IPMVP). Additional Linear Regression (MLR) gave better results; coefficients of determination (R^2) between 0.25 and 0.43 showed that including more than one variable increases predictive performance. The Multiple Linear Regression (MLR) model explained between 25% and 43% of the variance in annual energy consumption, indicating moderate but insufficient predictive capability relative to the IPMVP requirement of R^2>0.75. Although model performance improved during the stabilised operational year of 2023, explanatory strength declined again in 2024, suggesting the presence of additional unmodelled drivers.

However, the models are still limited and leave significant variations unexplained. This makes it essential to expand the scope of independent variables to include additional operational, technological, and environmental factors. In future studies, it is important to use a wider set of predictors to enhance the intensity of the energy consumption model. These can be building age, type of laboratory (engineering laboratories and computer laboratories), equipment density, relative humidity, efficiency of the HVAC system, and building orientation. Further studies are also needed to investigate more advanced approaches to enhance the accuracy and strength of energy consumption forecasts, including machine learning strategies and non-linear regression models. Moreover, to ensure the prediction of nonlinear and temporal interactions in campus energy consumption, advanced predictive techniques like Artificial Neural Networks (ANN) and Long Short-Term Memory (LSTM) models [6] are suggested. Nonetheless, these limitations do not mean the study is not a useful basis on which baseline energy models can be built to enable more effective energy planning, conservation strategies and policymaking in Malaysian educational institutions.

Acknowledgement

Communication of this research is made possible through monetary assistance by Universiti Tun Hussein Onn Malaysia, and this research was also supported by the Universiti Tun Hussein Onn Malaysia (UTHM) Research Management Centre (RMC), provided through the GPPS (vot J053).

Conflict of Interest

Authors declare that there is no conflict of interests regarding the publication of the paper.

Author Contribution

The authors confirm contribution to the paper as follows: ***study conception and design:*** *Norfakhira Mohd Nor, Mohd Azahari Razali;* ***data collection:*** *Norfakhira Mohd Nor;* ***analysis and interpretation of results:*** *Norfakhira Mohd Nor;* ***draft manuscript preparation:*** *Norfakhira Mohd Nor, Nur Syafiqah Adha Narrudin, Masataro Suzuki, Nur Erma Nardiah Nazami. All authors reviewed the results and approved the final version of the manuscript.*

References

[1] Nur Azfahani Ahmad, Muhammad Hilmi Dzulkefli & Fiza Abdul Rahim (2023) Assessing the energy performance of public university in Malaysia by using energy conservation measure (ECM): A case study of UiTM Tapah, Malaysia, *International Journal of Sustainable Construction Engineering and Technology*, *14(3)*, 309-319, https://doi.org/10.30880/ijscet.2023.14.03.027.

[2] Azra Nugraha Chandrakusuma & Shafini Mohd Shafie (2019) Energy efficiency: A study of energy consumption pattern in UUM Campus. In *Symposium on Technology Management and Logistics (STML Go Green),* 4, 28-37. https://repo.uum.edu.my/id/eprint/28487

[3] John Cowan (2002, October). *International performance measurement and verification protocol: Concepts and options for determining energy and water savings-Vol. I*, International Performance Measurement & Verification Protocol, 1. https://escholarship.org/uc/item/68b7v8rd

[4] Kementerian Tenaga, Teknologi Hijau dan Air (2015). National Energy Efficiency Action Plan, Malaysia. https://www.pmo.gov.my/wp-content/uploads/2019/07/National-Energy-Efficiency-Action-Plan.pdf

[5] Sock Yen Ng (2012). Energy awareness and energy-use behaviour of students in Malaysian university [Doctoral dissertation, Universiti Teknologi Malaysia].

[6] Muhammad Faiq, Kim Geok Tan, Chia Pao Liew, Ferdous Hossain, Chih-Ping Tso, Li Li Lim, Adam Yoon Khang Wong 7 Zulhilmi Mohd Shah (2023) Prediction of energy consumption in campus buildings using long short-term memory, *Alexandria Engineering Journal*, *67*, 65-76, https://doi.org/10.1016/j.aej.2022.12.015

[7] Ade Asmi, Autino Djamaris & Nur Sufini Hanizen Sulaiman (2023) A Study on Energy Efficiency in UTHM Buildings, *International Journal of Innovation Science, Engineering & Technology*, 10(01), 58-63. http://www.ijiset.com/

[8] Nur Adila Mohd Nizam, & Alvin John Lim (2023) Data analytic on energy consumption in UTHM building campus, *Recent Trends in Civil Engineering and Built Environment*, *4*(3), 541-551. https://penerbit.uthm.edu.my/periodicals/index.php/rtcebe/article/view/5842

[9] Nur Fatihin Najila Mohd Jaffri 7 Junaidah Jailani (2023) Study of building energy index and energy efficiency measures for Residential Colleges at UTHM main campus, *Recent Trends in Civil Engineering and Built Environment*, *4(1)*, 367-372. https://publisher.uthm.edu.my/periodicals/index.php/rtcebe/article/view/5824

[10] Jiwon Kim, Younghoon Kwak, Sun-Hye Mun & Jung-Ho Huh (2022) Electric energy consumption predictions for residential buildings: Impact of data-driven model and temporal resolution on prediction accuracy, *Journal of Building Engineering*, *62*, 105361, https://doi.org/10.1016/j.jobe.2022.105361

[11] Nor Shahida Razali & Nofri Yenita Dahlan (2015) Whole facility measurement for quantifying energy saving in an office building, Malaysia, *Applied Mechanics and Materials*, *785*, 676-681. https://doi.org/10.4028/www.scientific.net/AMM.785.676

[12] F. M. A. Rahman, Nofri Yenita Dahalan & Nor Shahida Razali (2015) Modelling adjusted baseline energy in an office building using artificial neural network, *Applied Mechanics and Materials*, *785*, 655-660, https://doi.org/10.4028/www.scientific.net/AMM.785.655

[13] Suhaidi Mohd Aris, Nofri Yenita Dahlan, Mohd Nasrun Mohd Nawi, Tengku Ahmad Nizam & Mohamad Zamhari Tahir (2015) Quantifying energy savings for retrofit centralized hvac systems at Selangor state secretary complex, *Jurnal Teknologi (Sciences & Engineering)*, *77(5)*, 93-100, https://doi.org/10.11113/jt.v77.6125

Int. Journal of Integrated Engineering Vol. 18 No. 3 (2026) p. 510-522

[14] F. C. Barnard & L. J. Grobler (2012, August 15-16) *Baseline service level adjustment methodologies for energy efficiency projects on compressed air systems in the mining industry* [In *2012 Proceedings of the 9th Industrial and Commercial Use of Energy Conference*]. IEEE, Town, South Africa, 1-8. https://ieeexplore.ieee.org/abstract/document/6330146

[15] Haizhou Fang, Hongwei Tan, Ningfang Dai & Xiaolei Yuan (2021, November 7-8). *Day-ahead prediction method of hourly building energy consumption in transition season* [In 2021 IEEE 7th International Conference on Cloud Computing and Intelligent Systems (CCIS)]. IEEE, Xi'an, China, pp. 376-380), https://doi.org/10.1109/CCIS53392.2021.9754671

[16] Guorong Sha & Qing Qian (2018, October 17-20). *Prediction of commercial building lighting energy consumption based on EPSO-BP* [In 2018 18th International Conference on Control, Automation and Systems (ICCAS)], IEEE, PyeongChang, South Korea, pp. 1035-1040. https://ieeexplore.ieee.org/abstract/document/8571584

[17] Rijalul Fahmi Mustapa, Atiqah Hamizah Mohd Nordin, Muhammad Asraf Hairuddin, Wan Suhaifiza W. Ibrahim, Siti Aliyah Mohd Saleh, N. Y. Dahlan, & Ihsan Mohd Yassin (2023, March 6-7). *Educational building's energy consumption independent variables analysis using linear regression model: a comparative study* [In 2023 IEEE 3rd International Conference in Power Engineering Applications (ICPEA)], IEEE, Putrajaya, Malaysia, pp. 202-207, https://doi.org/10.1109/ICPEA56918.2023.10093222

[18] Muhammad Faiq, Kim Geok Tan, Chia Pao Liew, Ferdous Hossain, Chih-Ping Tso, Li Li Lim, Adam Yoon Khang Wong & Zulhilmi Mohd Shah (2023). Prediction of energy consumption in campus buildings using long short-term memory. *Alexandria Engineering Journal*, *67*, 65-76, https://doi.org/10.1016/j.aej.2022.12.015

[19] Syahrunizam Buyamin, Nor Azura Osman & Azhar Ramli (2024) Optimising energy consumption in educational facilities: A case study on energy conservation practices in Malaysia, *Politeknik & Kolej Komuniti Journal of Engineering and Technology*, *9(2)*, 136-150. https://app.mypolycc.edu.my/journal/PMJET/article/view/713/396

[20] Rijalul Fahmi Mustapa, N. Y. Dahlan, Ihsan Mohd Yassin, AH Mohd Nordin & Azlee Zabidi (2019) Energy consumption prediction through linear and non-linear baseline energy model, *Indonesian Journal of Electrical Engineering and Computer Science*, *17(1)*, 102-109, https://doi.org/10.11591/ijeecs.v17.i1.pp102- 109