

## Article

# Development of an Electrical Energy Consumption Model for Malaysian Households, Based on Techno-Socioeconomic Determinant Factors

Boni Sena <sup>1,2</sup>, Sheikh Ahmad Zaki <sup>1,\*</sup> , Hom Bahadur Rijal <sup>3</sup> , Jorge Alfredo Ardila-Rey <sup>4</sup> , Nelidya Md Yusoff <sup>5</sup>, Fitri Yakub <sup>1</sup> , Farah Liana <sup>1</sup> and Mohamad Zaki Hassan <sup>5</sup> 

<sup>1</sup> Malaysia-Japan International Institute of Technology, Universiti Teknologi Malaysia, Kuala Lumpur 54100, Malaysia; boni.sena@ft.unsika.ac.id (B.S.); mfitri.kl@utm.my (F.Y.); farahlina@utm.my (F.L.)

<sup>2</sup> Faculty of Engineering, University of Singaperbangsa Karawang, Karawang 41361, Indonesia

<sup>3</sup> Faculty of Environmental Studies, Tokyo City University, Yokohama 224-8551, Japan; rijal@tcu.ac.jp

<sup>4</sup> Department of Electrical Engineering, Universidad Técnica Federico Santa María, Santiago de Chile 8940000, Chile; jorge.ardila@usm.cl

<sup>5</sup> Razak Faculty of Technology and Informatics, Universiti Teknologi Malaysia, Kuala Lumpur 54100, Malaysia; nelidya.kl@utm.my (N.M.Y.); mzaki.kl@utm.my (M.Z.H.)

\* Correspondence: sheikh.kl@utm.my

**Abstract:** Energy-saving strategies are required to address the increasing global CO<sub>2</sub> and electrical energy consumption problems. Therefore, the determinant factors of electrical energy consumption consist of socio-demographic changes, occupant behavior, house and appliance characteristics, or so-called techno-socioeconomic factors, which all need to be assessed. Statistics models, such as the artificial neural network (ANN), can investigate the relationship among those factors. However, the previous ANN model only used limited factors and was conducted in the developed countries of subtropical regions with different determinant factors than those in the developing countries of tropical regions. Furthermore, the previous studies did not investigate the various impacts of techno-socioeconomic factors concerning the performance of the ANN model in estimating monthly electrical energy consumption. The current study develops a model with a more-in depth architecture by examining the effect of additional factors such as socio-demographics, house characteristics, occupant behavior, and appliance characteristics that have not been investigated concerning the model performance. Thus, a questionnaire survey was conducted from November 2017 to January 2018 with 214 university students. The best combination factors in explaining the monthly electrical energy consumption were developed from occupant behavior, with 81% of the variance and a mean absolute percentage error (MAPE) of 20.6%, which can be classified as a reasonably accurate model. The current study's findings could be used as additional information for occupants or for companies who want to install photovoltaic or wind energy systems.

**Keywords:** electrical energy consumption model; artificial neural network; socio-demographic; house characteristics; occupant behavior; appliance characteristic



**Citation:** Sena, B.; Zaki, S.A.; Rijal, H.B.; Ardila-Rey, J.A.; Yusoff, N.M.; Yakub, F.; Liana, F.; Hassan, M.Z. Development of an Electrical Energy Consumption Model for Malaysian Households, Based on Techno-Socioeconomic Determinant Factors. *Sustainability* **2021**, *13*, 13258.

<https://doi.org/10.3390/su132313258>

Academic Editor: Farzad Jalaei

Received: 18 October 2021

Accepted: 26 November 2021

Published: 30 November 2021

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

The increasing CO<sub>2</sub> level worldwide, due to electrical energy consumption, is the fundamental factor of environmental damage such as the depletion of the ozone layer, global warming, and climate change. The emission of CO<sub>2</sub> has increased at a rate of 1.3% every year [1]. The emission of CO<sub>2</sub> can be classified as being from sources such as coal, natural gas, and oil in the industry and building sectors. The building sector consists of commercial and residential sectors that contribute 28% of the global energy-related CO<sub>2</sub> emissions [2]. Malaysia had a global commitment to reduce 45% CO<sub>2</sub> emission intensity by 2030, as stated in the Kuala Lumpur Low Carbon Society 2030 Blueprint [3]. Therefore,

optimization and efficiency improvement in the residential and commercial sections are necessary to achieve low carbon emissions in Malaysia.

The consumed energy levels in Malaysia escalated due to the contribution of residential sectors, being responsible for 20.7% of the total consumed energy [4]. The average electrical energy consumption was 345 kWh per month, from a survey of 348 samples in Malaysia [5]. Zaki et al. [6] performed a survey on occupant behavior regarding air conditioner consumption for a low-cost apartment in Kuala Lumpur. They confirmed that occupants tended to use AC without being affected by outdoor temperature in terms of the usage of AC in the bedrooms. Furthermore, most occupants used AC at night-time, a maximum of 50% of them being used in bedrooms. Ranjbar et al. [7] analyzed the time series datasets of electrical energy consumption during the El Nino climate situation. They found that the relationship between electrical energy usage and the outdoor air temperature was more notable in the night-time because the occupant used AC in the night-time rather than in the daytime. Hisham et al. [8] analyzed the air conditioning and total load diversity in typical residential buildings in Malaysia, and revealed that 75% of the occupants utilized AC almost every day for about five to six hours per day, at the rate of 0.93 kWh/day during the daytime and 3.43 kWh/day during the night-time. The correlation coefficient between total daily AC consumption and the mean daily outdoor temperature ranged from 0.1 to 0.48. Furthermore, Aqilah et al. [9] investigated the electrical energy consumption and indoor thermal environment for typical air-conditioned residential dwellings in Malaysia. The results showed that the total and the daily averages are 14.5 kWh/day and 3.9 kWh/day, respectively. The consumed electrical energy is predicted to increase because of economic improvement, occupant behavior and the availability of appliances at home. Therefore, a strategy for reducing electrical energy needs to be implemented to address the problem in Malaysia.

The strategy implementation is quite complex, since many determinant factors could affect the consumed electrical energy levels in a residential household's consumed electrical energy, such as sociodemographic issues, house characteristics, appliance characteristics, and occupant behavior factors. Moreover, those factors could have interconnected interactions with each other, as established by Jones et al. [10]. For example, Kim [11] studied determinant factors for a household in Korea, based on 2250 dwellings; Kim concluded that the factors showed a direct and indirect correlation with electrical energy consumption. It also stated that the use of a cooling appliance had a significant correlation with age, the education level of the head of the household, income, family composition, total floor area and the number of people living in the building. Furthermore, Chen et al. [12] examined the factors affecting household electrical energy consumption in China, based on 642 households during winter and 838 households during the summer. They ascertained that the number of rooms, number of people living in it, the age of the occupants, the number of AC units owned, education level, and AC consumption were interconnected. Similar findings from Shahi et al. [13] verified the energy usage in rural, semi-urban and urban areas in Nepal, based on 442 households, which concluded that the energy usage was linked to income, job, number of residents, education level and the use of LED lamps.

Mathematical models are required to evaluate the determinant factors affecting electrical energy consumption, such as an artificial neural network (ANN). Previous studies using the ANN model, such as that by Roque [14], developed the model based on 136 houses in Canada, and concluded that the combination of sociodemographic factors, the building, and the number of hours spent at home could explain 91% of the variance. Buratti et al. [15] used the ANN model based on house characteristics and additional power systems in Italy, and found a variance of 98%. Khayatian and Sarto [16] applied the ANN model to evaluate the energy performance certificate of residential houses, based on 220,000 houses in Italy, which concluded the variance as being between 81 and 91%. Contrary to the findings of previous studies, Lee et al. [17] utilized the ANN model to predict electrical energy consumption based on socio-demographic factors and occupant behavior, and stated that the model only could explain 37% to 42% of the variance. Hawkins et al. [18] established an

ANN model to investigate the determinant factors affecting electrical energy usage in the UK for university buildings, based on 1872 rooms. The results showed that the model had a mean absolute percentage error (MAPE) of 34% and a variance (CV) coefficient of 33.6%. Pombeiro et al. [19] developed an ANN model for one commercial building in Portugal, based on the real-time dataset of one month with 3381 datasets, and concluded that the model could explain 68% of the variance, with a mean absolute error (MAE) of 8.5 and mean square error (MSE) of 17.2%. Beccali et al. [20] studied the development of the ANN model to find the energy performance of 151 buildings in Italy, and concluded that the MAE of the model was 1.99. Most of the previous studies were conducted in a temperate region with limited determinant factors, as performed by Roque [14], Buratti et al. [15], Lee et al. [17], and Hawkins et al. [18], while other studies did not consider determinant factors in their research, such as Khayatian and Sarto [16], Pombeiro et al. [19], and Beccali et al. [20].

There are several studies of residential energy consumption in a tropical climate, especially in Malaysia. Zaki et al. [21] developed a model for generating air-conditioner operation schedules, based on 38 dwellings in Malaysia. The results showed that the probability density model offered good performance in predicting the observation data. Ranjbar et al. [22] analyzed the short-term measurement of real-time electrical energy consumption in Kuala Lumpur during hot weather, and concluded that the AC contributed 19.4% to 52.3% of the total consumed electricity. Swarno et al. [23] observed the diurnal variation of the urban microclimate in Kuala Lumpur, and indicated that urban microclimatic parameters were influenced by monsoon seasons and the urban topography. Sena et al. [24] constructed a conceptual framework of electrical energy consumption modeling for Malaysian households, using an artificial neural network. They suggested that the combination of sociodemographics, house characteristics, occupant behavior, and appliance characteristics were important in developing an ANN model for estimating electrical energy consumption. Aqilah et al. [9] observed real-time electrical energy consumption for selected Malaysian households at middle-income levels. They classified three types of appliances: cold, active and standby appliances. The results showed that cold appliances such as refrigerators consumed 118.8 kWh/month, followed by cold and hot appliances, such as water dispensers, which consumed 52.5 kWh/month, and standby appliances, such as televisions, which consumed 0.9 kWh/month. The consumed electricity was higher in the night-time, due to the use of the air conditioner. Wijaya and Tezuka [25] assessed the determinant factors using a multiple linear regression model; this confirmed that sociodemographic, occupant behavior and appliance characteristics showed a significant contribution to the electrical energy consumption for residents in Indonesia. Azlina et al. [26] studied influencing factors such as sociodemographics, house characteristics, occupant behavior and appliance characteristics, and concluded that those factors had a significant impact on the variability of household electricity consumption. Some studies were conducted in tropical countries; however, the comprehensive impact of the determinant factors of techno-socioeconomics with respect to electrical energy consumption, based on the ANN model approach, was not investigated.

Most of the previous electrical energy consumption models based on determinant factors were developed in temperate regions, such as those of Hawkins et al. [18], Beccali et al. [20], and only a few studies were conducted in tropical regions, as were those performed by Wijaya and Tezuka [25] and Azlina et al. [26]. The determinant factors of techno-socioeconomic decisions in the temperate region might differ from those in the tropical region due to the difference in socio-economic level, cultures, and climates condition, which could affect the model's performance. Previous studies did not consider the various impacts of determinant factors in a tropical region with respect to the performance of the ANN model. The current study aims to investigate the performance of the ANN model, based on different determinant factors. This study also aims to determine the best combination of determinant factors regarding Malaysian households, when estimating monthly electrical energy consumption.

## 2. Materials and Methods

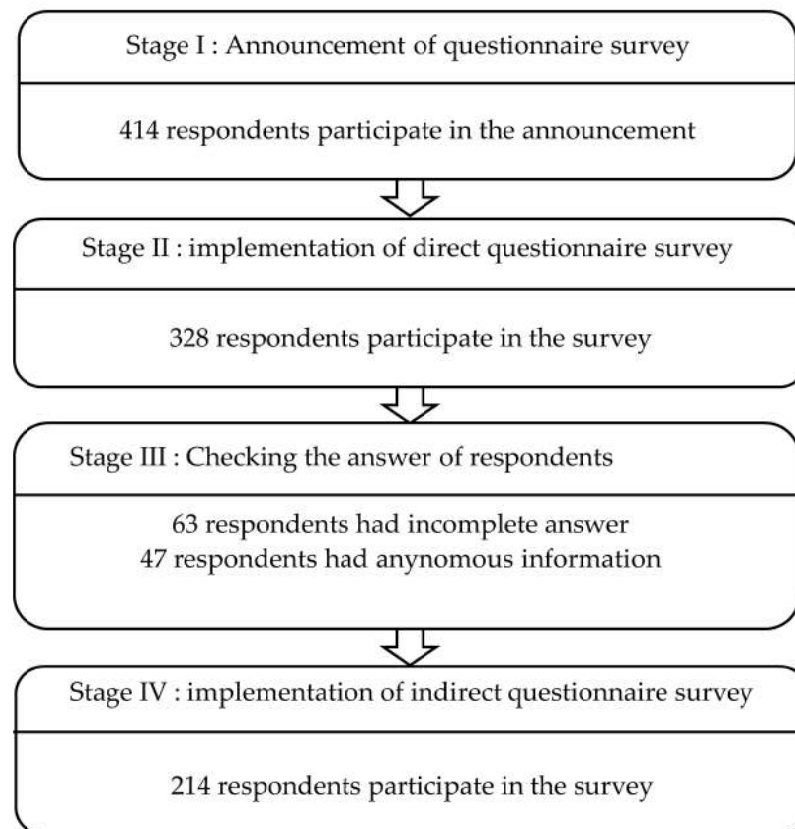
### 2.1. Questionnaire and Survey

Datasets in this research were collected from a field survey that was performed on respondents from students of the Malaysia-Japan International Institute of Technology (MJIT) and the Universiti Teknologi Malaysia (UTM), Kuala Lumpur, from November 2017 to January 2018. The process of finding potential respondents was implemented based on random sampling, which means that all respondents had a similar possibility of being selected for the field survey. The respondents willing to participate in the survey would share their scheduled time and place with the main author. The information about the survey had been announced to 414 students by visiting their classes, with the permission of the lecturers. However, only 328 students were involved in the survey, which was divided into two parts, a direct and an indirect survey.

The direct questionnaire survey aimed to find general information regarding determinant factors from the respondents, and was conducted from the middle of November to the middle of December 2017. At this stage, the author would implement a direct survey by distributing the questionnaire to all respondents. First, they would fill in the questionnaire after receiving direction from the main author. Then, the respondents would fill in the questionnaire based on information from their family or parents' house, since most of them were staying in the university dormitory at the time of the survey. The general information regarding their households would be acquired at this stage. Finally, the main author would check the answers of the respondents, which can be classified into two types: verified and unverified questionnaires.

A verified questionnaire implies that the questionnaire's responses were plausible as common answers from people who were given similar questions. The respondents respond to all questions, except for detailed questions that needed to be answered by their parents or family members. On the other hand, an unverified questionnaire implies that the responses to the questionnaire were dubious, or that the responses were not completed. For example, an anonymous questionnaire was classified as unverified because the main author could not confirm the questionnaire respondent's answers. After checking all responses, the main author found 63 unfinished or dubious responses and 47 anonymous questionnaires. Therefore, the unverified questionnaires would be removed from further processing to avoid invalid results. The number of verified questionnaires was 214, and these would be carried forward to an indirect questionnaire survey. Figure 1 shows the process of data collection based on the questionnaire and survey, from the beginning stage of the announcement to the last stage of the indirect questionnaire survey.

An indirect questionnaire survey aimed to extract detailed information regarding determinant factors in the households, and was performed from January to April 2018. The indirect questionnaire survey would be implemented by obtaining detailed information using social media apps, because the respondents' parents or other household members must determine some information. Some questions at this stage regarded monthly electrical energy bills and consumption figures, total family income, floor space area, detailed house characteristics (number of windows, number of specific room types), details regarding appliance types (period of purchasing appliances, power rating of an appliance, and detailed occupant behaviors (occupancy schedule for each household member, usage profile of household appliances). The respondents obtained their monthly electrical energy consumption from their electricity bill, consisting of detailed information on how much electricity was consumed in a month. The verified and completed questionnaires from direct and indirect questionnaire surveys will be analyzed further. Details of the process and the results of the questionnaire survey were provided by Sena et al. [27]. Summary of questionnaire survey was provided in Appendix A. Figure 2 shows the flowchart of a direct and indirect questionnaire survey.



**Figure 1.** Summary of participants from the beginning to the end of the questionnaire survey process.

The respondents were divided into 113 males (52.8%) and 101 females (47.2%). The average age was 22 years old, with a standard deviation of 6 years. The total number of respondents who provided their house location was 179 students (83.6%). Most of the surveyed houses were in Selangor (24.3%), Kuala Lumpur (14.5%), and Johor (11.2%). Putrajaya and Labuan were not included in the survey because the respondents' houses were not located in those places. In addition, 35 respondents did not provide the location of their houses for reasons of privacy. The type of house for respondents comprised a detached house (82.2%), apartment (10.3%), townhouse (3.3%), shop house (0.9%), and others (6%). Figure 3 shows the types of surveyed houses. The distribution of the sample size for houses in each state, based on the collected datasets, is shown in Table 1.

**Table 1.** Number of buildings per state.

States	Building Numbers
Selangor	52
Kuala Lumpur	31
Johor	24
Pahang	14
Sarawak	11
Perak	10
Kelantan	9
Negeri Sembilan	8
Kedah	7
Penang	6
Melaka	3
Sabah	2
Terengganu	2

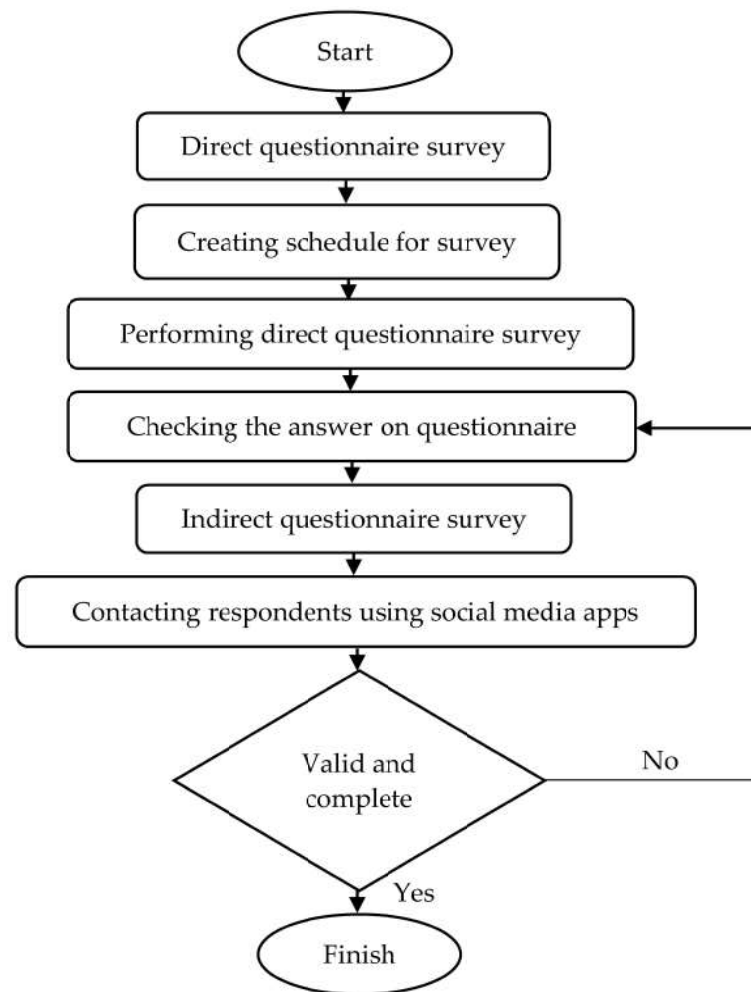


Figure 2. Flowchart of a direct and indirect questionnaire survey [27].



Figure 3. Typical types of houses investigated. (a) Typical single-storey house [28]. (b) Typical double-storey house [29]. (c) Typical three-storey house [30].

Figure 4 shows the climograph for Malaysia, based on secondary data from the weather station in Sultan Abdul Aziz Shah Airport, Selangor Malaysia [31]. Malaysia has a hot and humid climate throughout the year; therefore, the monthly average outdoor temperature and humidity is almost constant in most towns, as confirmed by Kubota et al. [32], Swarno et al. [23], and Zaki et al. [33]. The average monthly outdoor temperature was 27.8 °C, and the average monthly relative humidity was 82%. The maximum average monthly outdoor temperature was 29 °C, and the maximum average monthly relative humidity was 88%, while the minimum average monthly outdoor temperature was 25.2 °C and the minimum average monthly relative humidity was 78%.

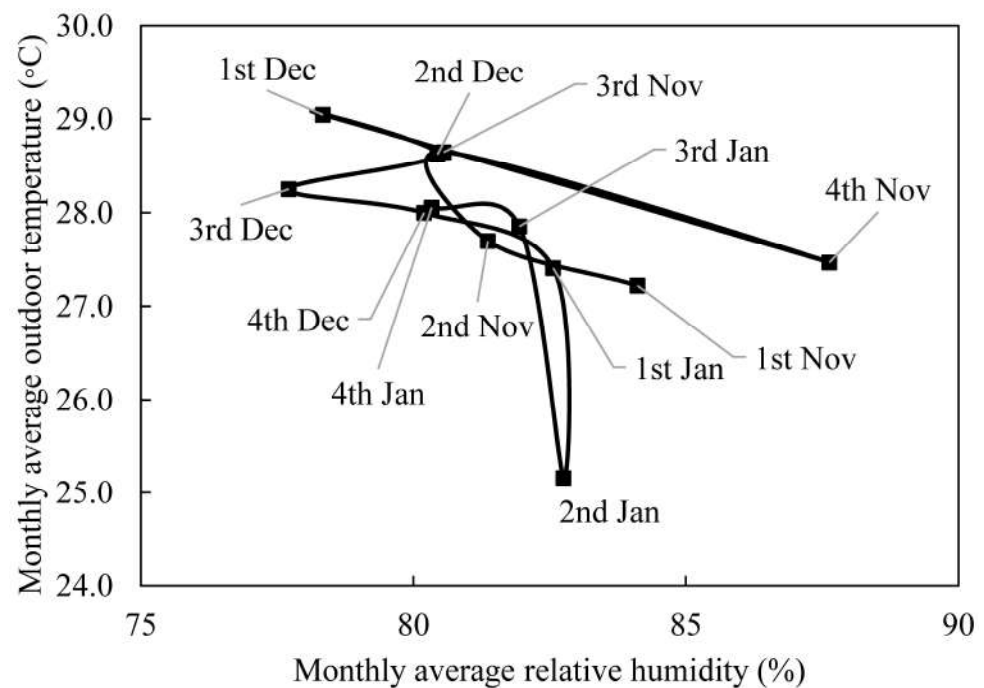


Figure 4. Climograph of Malaysia.

Each point in the climograph represents the average weekly outdoor temperature and average weekly relative humidity. For example, “1st Dec” means the first week of December, with an average relative humidity of 78% and an average outdoor temperature of 29.1 °C. The weekly average relative humidity and outdoor temperature sequentially decreased from 1 December to 2 December, 3 November, and 4 November.

## 2.2. Statistical Methods

The artificial neural network (ANN) model was utilized as a statistical method to estimate monthly electrical energy consumption, based on a comprehensive set of determinant factors. The feed-forward of multi-layer perceptron neural networks was selected as a type of ANN model in this paper. This is because Akbari et al. [34] confirmed that the network had been commonly used with a non-linear activation function, allowing the network to learn non-linear and linear relationships between input and output models. The development of the ANN model started by setting the type and parameter. Equation (1) shows the basic ANN model, as established by Biswas et al. [35], where input ( $x$ ), weight function ( $w$ ), bias function ( $\theta$ ), output model ( $y$ );  $f(\cdot)$  is an activation function that determines the output of the neuron.

$$y = f(wx + \theta) \quad (1)$$

The 214 verified questionnaire survey responses would be used as the source of the ANN model. Data on techno-socioeconomic factors, which consists of socio-demographics, house characteristics, occupant behavior and appliance characteristics as acquired from the questionnaire, would be used as an input of the ANN model. A detailed explanation of those factors has been provided by Sena et al. [27]. Weight function ( $w$ ) and bias function ( $\theta$ ) would be randomly determined while developing the ANN model. Data regarding monthly electrical energy consumption, as taken from electric bills obtained from the surveyed house from July 2017 to January 2018, would be utilized as an output for the ANN model.

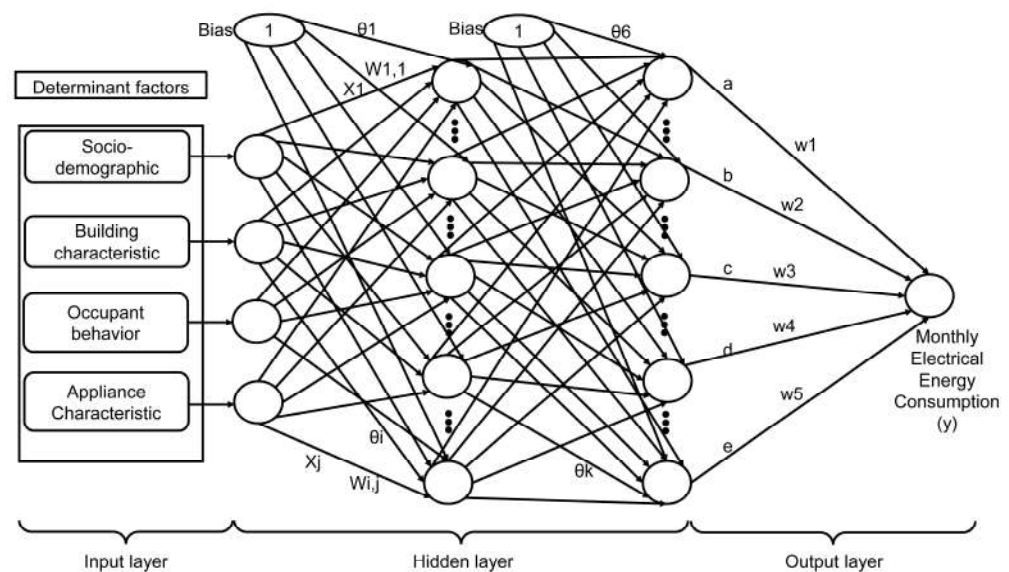
The datasets were divided up for two different purposes: data for developing an ANN for a 184-dataset model and data from the testing model for 29 datasets. One piece of data for the testing model was removed from the analysis because the value of monthly electrical energy consumption was too high compared to other datasets, reducing the

model's performance. Table 2 shows the statistical descriptive data for developing and testing the model.

**Table 2.** Statistic descriptive data of developing and testing model.

Parameters	Data of Developing Model	Data of Testing Model
Average	444.8	368.3
Max.	873.7	837.3
Min.	64.7	100
Std. Dev.	176.0	170.9
Variance	420.0	406.0
Range	809.0	737.3
N	184	29

The data sets used for developing the ANN model were divided into three types: training, validation, and testing for 70%, 15%, and 15% of 184 datasets, respectively. The ANN model would randomly select the classifications of those types of datasets. In order to investigate the model performance, the data of the testing model for 29 datasets would be used, which were different from the data used for developing the ANN model. The hidden layers were determined randomly and changed until the ANN model achieved a minimum mean square error (MSE). Four training functions were used in this paper: the Levenberg–Marquadt (LM), scaled conjugate gradient (SCG), Broyden–Fletcher–Goldfarb–Shanno (BFGS), and the one-step secant (OSS).



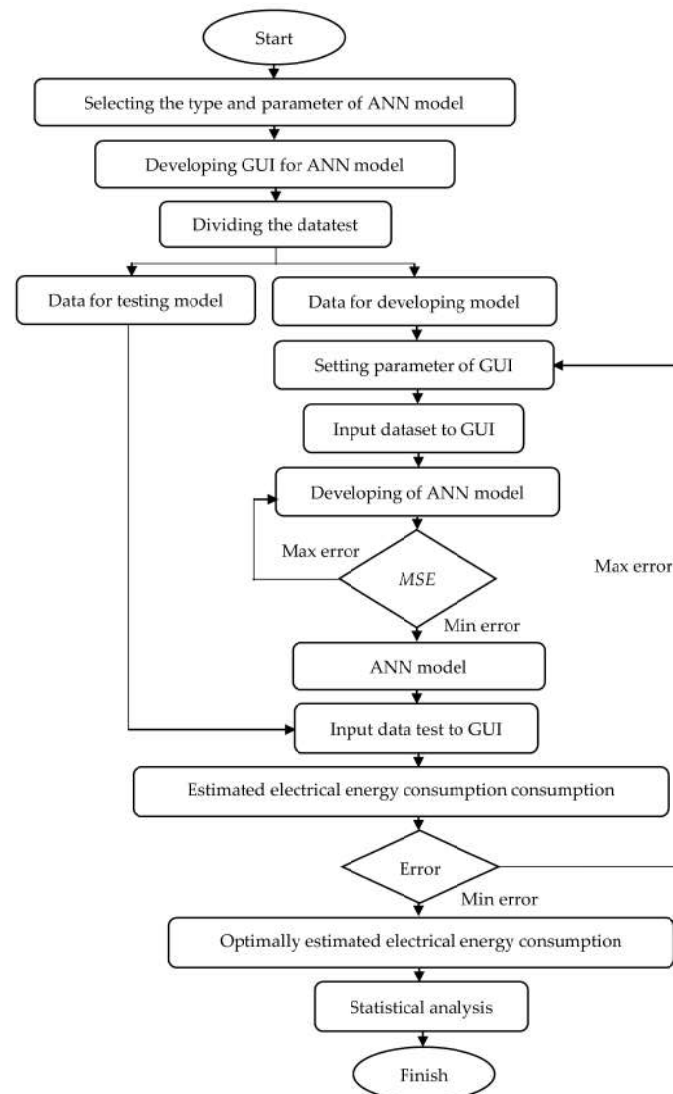
**Figure 5.** A framework of an ANN model, based on determinant factors [24].

A graphical user interface (GUI) was developed after selecting the type and parameter of the ANN model. The GUI was used as the interface for the user, who employed the ANN model to estimate monthly electrical energy consumption, based on different techno-socioeconomic factors. The data of the developing model would initially be put into GUI before developing the model. After finishing the development of the model, the data for the testing model would be put into it, and the estimated monthly electrical energy consumption would emerge via the GUI. The estimated monthly electrical energy consumption would be manually compared to those of data of the testing model, due to the limitations of computer specification. If the errors between the estimated and observed data were still high, the development process of the ANN model would be reconstructed again by re-setting the parameters of the ANN model and re-inputting the dataset into the GUI. The process would be repeated until the error between the estimated and observed monthly electrical energy consumptions was considered to be at



the minimum. Finally, model validation would be performed by comparing the observed and estimated monthly electrical energy consumption and by using a statistical method, as suggested by Buratti et al. [15]. The basic ANN model was developed based on the objectives of the current study, consisting of socio-demographic determinant factors, house characteristics, occupant behavior and appliance characteristics as the input model ( $x_i$ ) and monthly electrical energy consumption as the output model ( $y$ ), as shown in Figure 5.

Statistical methods were utilized to measure how much error in the model was based on absolute error (AE), mean absolute error (MAE), mean squared error (MSE), root mean square error (RMSE), mean absolute percentage error (MAPE), a sum of square error (SSE), the sum of squared total (SST), coefficient of determination ( $R^2$ ), coefficient of variance (CV), and mean bias error (MBE). These methods can be used to determine which model has the best performance in estimating the monthly electrical energy consumption, based on the combination of determinant factors. Figure 6 shows the flowchart of development and validation of the ANN model, based on determinant factors affecting monthly electrical energy consumption.

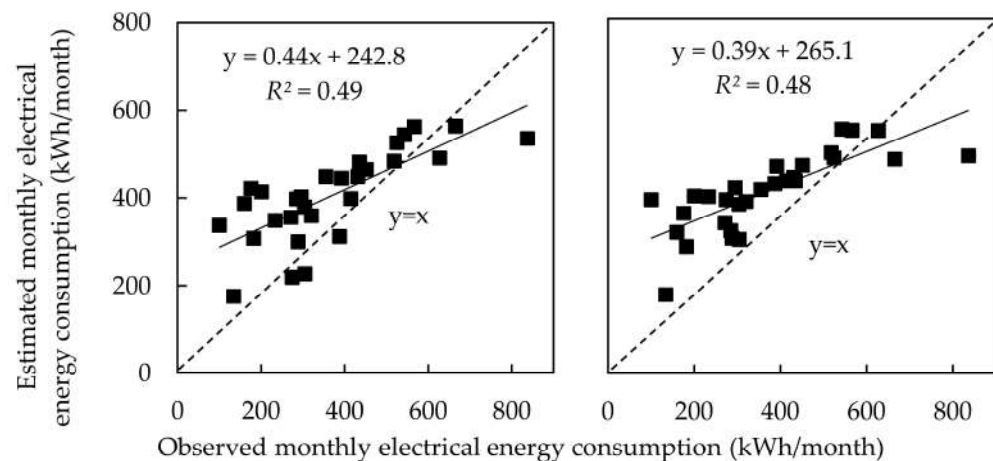


**Figure 6.** Flowchart of the validating non-linear model, based on determinant factors affecting monthly electrical energy consumption, using an ANN.

### 3. Results and Discussion

#### 3.1. Performance of the Model Based on House Characteristic Factors

The performance of the model based on house characteristic factors was assessed in this section to investigate how much impact the factors have, to explain the variance of observed and estimated electrical energy consumption. ANN models for electrical energy consumption were constructed, taking the number of rooms as the input of the model and monthly electrical energy consumption as the model output. The  $x$ -axis shows the observed monthly electrical energy consumption, which means the data comes from a field survey on Malaysian households. The  $y$ -axis shows the estimated monthly electrical energy consumption using the data results derived from the ANN model in the current study. The ANN model based on the Broyden–Fletcher–Goldfarb–Shanno (BFGS) training function showed slightly better performance than that of the model based on the one-step secant (OSS) training function, explaining 49% and 48% of the variance of observed and estimated monthly electrical energy consumption, respectively, as shown in Figure 7.



**Figure 7.** Estimated and observed monthly electrical energy consumption for the house characteristics model, based on BFGS and OSS.

Table 3 shows the electrical energy consumption performance for the ANN model, based on the house characteristic factors. The ANN model based on the house characteristics model factors, using the BFGS training function, showed a slightly higher error than the ANN model using the OSS training function for MAE, RSME, and MBE. However, both models showed a MAPE of less than 25%, which means that those models were classified as a reasonable estimation. The CV value for both models was more than 30%, which means that both models had high variance. However, the  $R^2$  value was less than 50%, which means that both models still had a low correlation between observed and estimated monthly electrical energy consumption.

**Table 3.** Performance of house characteristics model using BFGS and OSS.

Parameter	Training Function	
	BFGS	OSS
Mean absolute error (MAE) (kWh)	91.9	88.8
Mean square error (MSE) (kWh)	14,929.3	15,293.2
Root mean square error (RMSE) (kWh)	30.9	30.6
Mean absolute percentage error (MAPE) (%)	24.0	22.5
Coefficient of determination ( $R^2$ )	0.49	0.48
Coefficient of variance (CV) (%)	33.8	34.2
Mean bias error (MBE) (kWh)	25.8	25.0

Note: BFGS: Broyden–Fletcher–Goldfarb–Shanno, OSS: one step secant.

The current ANN model based on house characteristics had an almost similar performance to that reported by Hawkins et al. [18], as shown in Table 4. Hawkins et al. [18] stated that the MAPE (mean absolute percentage error) and CV (coefficient of variance) were 34% and 33.6%, respectively. On the other hand, the current model found that MAPE and CV were 22.5% and 34.2%, respectively. Hawkins et al. [18] constructed an ANN model based on house and weather factors for a university building in the United Kingdom. The datasets of their model consisted of two types: data for developing a model with 148 datasets and data for testing the model with 20 datasets, while the current study used 184 datasets for the developing model and 29 datasets for the testing model, and one dataset of the testing model was removed because the value was too high. Therefore, the current study had an almost similar performance to that of the model used by Hawkins et al. [18]. Both studies had a similar coefficient of variance (CV) due to the discrepancy in the input model. The current study only utilized a limited input model regarding the number of rooms, while Hawkins et al. [18] used many numerical and categorical data types without any data normalization.

**Table 4.** First Comparison between the current study and that of Hawkins et al. [18].

ANN Type	Current Study		Hawkins et al. [18]	
	MAPE	CV	MAPE	CV
OSS	22.5%	34.2%	34%	33.6%
BFGS	24%	33.8%	34.7%	34.3%

Note: BFGS: Broyden–Fletcher–Goldfarb–Shanno, OSS: one step secant, MAPE: mean absolute percentage error, CV: coefficient of variance.

Compared to other studies, Lee et al. [17] established the ANN model based on socio-demographic factors, such as age, income, gender, education level, job, and time indoors for a household in South Korea. It uses genuine real-time and categorical data with a normalization process, which was better than that of the current study, as shown in Table 5. Nevertheless, the current study used correlation analysis to separate the datasets that had low and insignificant correlations from the analysis, as explained by Sena et al. [27], while Lee et al. [17] did not use the same method. The current study showed better performance than that of Lee et al. [17] by explaining 49% and 40% of the observed electrical energy consumption variance, respectively. In contrast, Lee et al. [17] reported less error because they had larger datasets for the training ANN model than the current study, with MSEs of  $1.04 \times 10^4$  and  $1.5 \times 10^4$ , respectively.

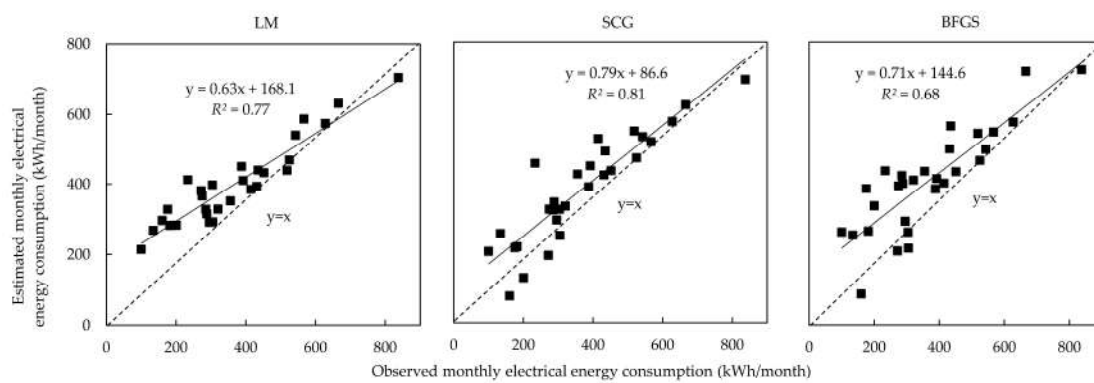
**Table 5.** First Comparison between the current study and Lee et al. [17].

ANN Type	Current Study		Lee et al. [17]	
	MSE	$R^2$	MSE	$R^2$
OSS	$1.53 \times 10^4$	0.48	$1.05 \times 10^4$	0.42
BFGS	$1.5 \times 10^4$	0.49	$1.04 \times 10^4$	0.40

Note: BFGS: Broyden–Fletcher–Goldfarb–Shanno, OSS: one step secant, MSE: mean square error,  $R^2$ : coefficient of determination.

### 3.2. Performance of Model Based on Occupant Behavior Factors

The performance of the model based on occupant behavior factors was carried out in this section to find out how the impact of certain factors explained the variance in observed and estimated electrical energy consumption. ANN models for electrical energy consumption were constructed based on the set temperature of the AC, usage of the router, the water dispenser and hairdryer on weekdays as the input of the model and monthly electrical energy consumption as the model's output. The performance of the ANN model in estimating monthly electrical energy consumption, using the Levenberg–Marquardt (LM), scaled conjugate gradient (SCG), and Broyden–Fletcher–Goldfarb–Shanno (BFGS) training functions, are shown in Figure 8.



**Figure 8.** Estimated and observed monthly electrical energy consumption for the occupant behavior model, based on LM, SCG and BFGS. Note: x: observed monthly electrical energy consumption, y: estimated monthly electrical energy consumption,  $R^2$ : coefficient of determination.

Table 6 shows the examination of electrical energy consumption performance for the ANN model, based on occupant behavior factors. The ANN model based on the SCG training function had better variance, explaining 81% of the variance, than those based on the LM and BFGS training functions, which only had a variance of 77% and 68% of the observed and estimated monthly electrical energy consumption. The  $R^2$  value showed more than 65%, which means that the models had a moderate correlation between the observed and estimated monthly electrical energy consumption. The value of CV from the models showed less than 25%, which implies a high variance. On the other hand, the value of MAPE for the model, based on LM and SCG, showed a value of less than or equal to 20%, which means that the models had a good estimation for monthly electrical energy usage. In contrast, the model based on BFGS showed 24% of MAPE, which means the model had a reasonable estimation.

**Table 6.** Performance of the occupant behavior model using LM, SCG and BFGS.

Parameter	Training Function		
	LM	SCG	BFGS
Mean absolute error (MAE) (kWh)	63.6	58.1	80.6
Mean square error (MSE) (kWh)	6650.7	5534.7	9479.5
Root mean square error (RMSE) (kWh)	24.4	31.0	31.4
Mean absolute percentage error (MAPE) (%)	18.0	20.6	24.0
Coefficient of determination ( $R^2$ )	0.77	0.81	0.68
Coefficient of variance (CV) (%)	22.5	20.6	26.9
Mean bias error (MBE) (kWh)	17.9	16.3	22.7

Note: LM: Levenberg–Marquadt, SCG: scaled conjugate gradient, BFGS: Broyden–Fletcher–Goldfarb–Shanno.

Compared to the previous research, the current ANN model had a smaller error than that of Hawkins et al. [18], with 18% and 34% MAPE, respectively, as shown in Table 7. In addition, the current ANN model showed less coefficient of variance than that of Hawkins et al. [18], with a CV of 22.5% and 33.6%, respectively. The current ANN model showed better performance in estimating observed monthly electrical energy consumption than that of Hawkins et al. [18] due to the different factors of the input model. The current model utilized occupant behavior as an input of the model; in contrast, Hawkins et al. [18] took house characteristic factors as inputs of the model. Occupant behavior tends to be a non-linear characteristic, and house characteristic factors showed a linear habit regarding electrical energy consumption, as Tam et al. [36] confirmed. Therefore, the current model based on occupant behavior showed better performance than the previous model based on house characteristics.

**Table 7.** Second Comparison between the current study and that of Hawkins et al. [18].

ANN Type	Current Study		Hawkins et al. [18]	
	MAPE	CV	MAPE	CV
LM	18.0%	22.5%	34%	33.6%
SCG	20.6%	20.6%	34.7%	34.3%
BFGS	24.0%	26.9%	34.7%	33.7%

Note: LM: Levenberg–Marquadt, SCG: scaled conjugate gradient, BFGS: Broyden–Fletcher–Goldfarb–Shanno, MAPE: mean absolute percentage error, CV: coefficient of variance.

Kim et al. [37] developed an electrical energy consumption model based on the ANN model for South Korea: socio-demographics, house characteristics, occupant behavior, and appliance characteristics. They utilized a large population sample from the secondary database from government institutions, combining numeric- and categorical-type datasets. On the other side, the current model also utilized combinations based on those types of datasets. However, it only covered a small number of samples. The current study showed better performance than Kim et al. [37] by explaining 11% and 70% of the variance, as shown in Table 8. The current model utilized correlation analysis to assess the linear relationship among the determinant factors of monthly electrical energy consumption and multiple linear regression, to determine the best combinations of those factors, as explained by Sena et al. [27]. In contrast, Kim et al. [37] utilized multiple linear regression to assess the determinant factors without correlation analysis. The effect of a non-linear characteristic within the datasets of the determinant factors would affect the performance of the ANN model, as in the model used by Kim et al. [37].

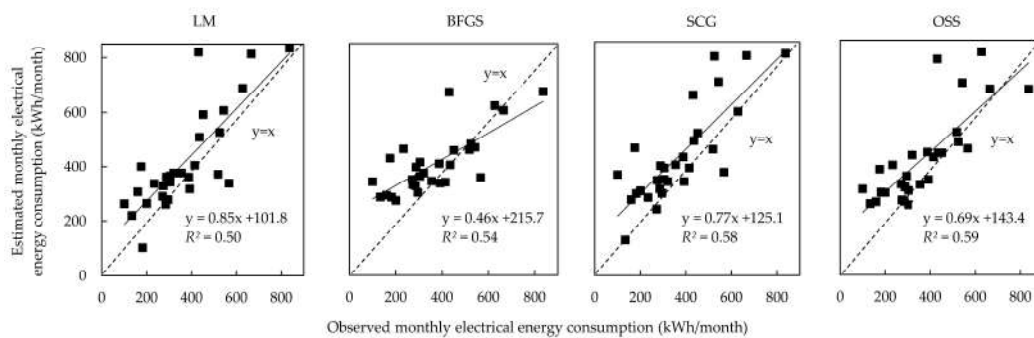
**Table 8.** First Comparison between the current study and that of Kim et al. [37].

ANN Type	Current Study		Kim et al. [37]	
	MSE	$R^2$	MSE	$R^2$
SCG	$5.5 \times 10^3$	0.81	$7.69 \times 10^5$	0.11
LM	$6.6 \times 10^3$	0.77	$5.11 \times 10^5$	0.07
BFGS	$9.5 \times 10^3$	0.68	$2.59 \times 10^6$	0.08

Note: LM: Levenberg–Marquadt, SCG: scaled conjugate gradient, BFGS: Broyden–Fletcher–Goldfarb–Shanno, MSE: mean square error,  $R^2$ : coefficient of determination.

### 3.3. Performance of Model Based on Socio-Demographic and Appliance Characteristic Factors

The performance of the model based on socio-demographic and appliance characteristic factors was conducted in this section, in order to investigate the effect of those factors to represent the variance of observed and estimated electrical energy consumption. As a result, ANN models for electrical energy consumption were developed based on the number of wall-mounted ACs, the number of chargers of smartphones, the number of AC energy stars, the education level of the household head, the number of standalone freezers, and the number of fluorescent lamps as the input of the model, and monthly electrical energy consumption as the output of the model. As a result, an ANN model based on the OSS training function had better performance in explaining monthly electrical energy consumption, with a variance of 59%, than those of models based on LM, BFGS, and SCG, which explained 50%, 54%, and 58%, respectively, as shown in Figure 9.



**Figure 9.** Estimated and observed monthly electrical energy consumption for the socio-demographic and appliance characteristic model, based on LM, BFGS, SCG and OSS. Note:  $x$ : observed monthly electrical energy consumption,  $y$ : estimated monthly electrical energy consumption,  $R^2$ : coefficient of determination.

Table 9 shows the examination of electrical energy consumption performance for the ANN model, based on socio-appliance factors. The ANN model that was based on the OSS training function showed better variance, explaining 59% of the variance, than those based on LM, SCG and BFGS training functions, which only had variances of 49%, 58% and 54% of the observed, estimated monthly electrical energy consumption. All models had a MAPE of less than 25%, which means that those models gave a reasonably accurate estimation of monthly electrical energy consumption. The CV values showed as equal to and more than 30%, which means that the models showed high variance. On the other hand, the  $R^2$  values in these models were less than 60%, which means that the correlation between the observed and estimated monthly electrical energy consumption was relatively low.

**Table 9.** Performance of socio-appliance model using LM, SCG, BFGS and OSS.

Parameter	Training Function			
	LM	SCG	BFGS	OSS
Mean absolute error (MAE) (kWh)	89.7	83.5	93.7	81.3
Mean square error (MSE) (kWh)	14,914.0	12,266.4	13,560.9	11,444.8
Root mean square error (RMSE) (kWh)	32.3	27.2	30.8	26.4
Mean absolute percentage error (MAPE) (%)	24.5	20.7	24.8	21.0
Coefficient of determination ( $R^2$ )	0.49	0.58	0.54	0.59
Coefficient of variance (CV) (%)	33.7	30.6	32.2	29.6
Mean bias error (MBE) (kWh)	25.2	23.5	26.3	22.8

Note: LM: Levenberg–Marquadt, SCG: scaled conjugate gradient, BFGS: Broyden–Fletcher–Goldfarb–Shanno, OSS: one-step secant.

The ANN model based on socio-demographic and appliance factors showed better variance than that of Lee et al. [17], explaining 42% and 60%, respectively. In contrast, the current study had a higher mean square error (MSE) due to the lack of training dataset for the ANN model, with an MSE of  $1.14 \times 10^4$  and  $1.05 \times 10^4$ , respectively. Table 10 compares the results of Lee et al. [17] and the current study, based on  $R^2$  and MSE. The difference in performance between the current model and Lee et al. [17] is because of the input of the ANN model. The current model used a combination of socio-demographic and appliance characteristics. However, Lee et al. [17] only considered limited socio-demographic and occupant behavior factors. Furthermore, the current model utilized continuous and categorical-type datasets, while Lee et al. [17] utilized categorical, continuous, and binary-type datasets. In the model used by Lee et al. [17], only monthly income offered the continuous type of data, while occupancy time used binary datasets. The binary datasets could reduce the model's performance because the datasets used limited numbers, such as 0 and 1. On the other hand, the continuous type used a wide range of numbers from 0 to 9. Nevertheless, the current model had a higher MSE than that of Lee et al. [17] due to the limitation of the number of samples.

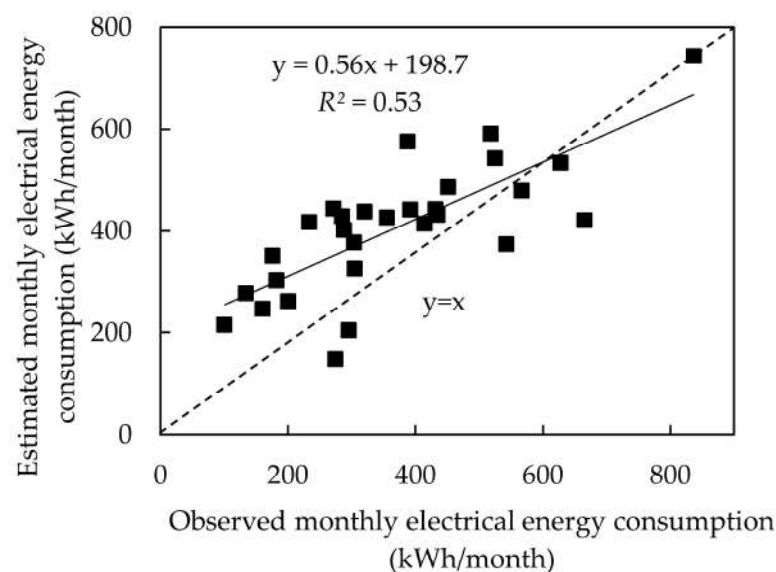
**Table 10.** Second Comparison between the current study and that of Lee et al. [17].

ANN Type	Current Study		Lee et al. [17]	
	MSE	$R^2$	MSE	$R^2$
OSS	$1.14 \times 10^4$	0.60	$1.05 \times 10^4$	0.42
LM	$1.49 \times 10^4$	0.49	$1.04 \times 10^4$	0.40
BFGS	$1.36 \times 10^4$	0.54	$1.09 \times 10^4$	0.38
SCG	$1.23 \times 10^4$	0.58	$1.04 \times 10^4$	0.38

Note: LM: Levenberg–Marquadt, SCG: scaled conjugate gradient, BFGS: Broyden–Fletcher–Goldfarb–Shanno, OSS: one-step secant, MSE: mean square error,  $R^2$ : coefficient of determination.

### 3.4. Performance of Model Based on Socio-Demographic and Occupant Behavior Factors

The performance of the model based on socio-demographic and occupant behavior factors was investigated in this section to determine how much impact those factors had on the variance of observed and estimated monthly electrical energy consumption. The ANN model for electrical energy consumption was constructed based on the number of wall-mounted ACs, the number of chargers for smartphones, the number of AC energy stars, the education level of the household head, the number of standalone freezers, and the number of fluorescent lamps as inputs of the model and monthly electrical energy consumption as the output of the model. As a result, the ANN model based on the Levenberg–Marquadt (LM) training function could explain 53% of the monthly electrical energy consumption variance for socio-demographic and occupant behavior factors, as shown in Figure 10.



**Figure 10.** Observed and estimated monthly electrical energy consumption for socio-demographic and occupant behavior, based on LM. Note: x: observed monthly electrical energy consumption, y: estimated monthly electrical energy consumption,  $R^2$ : coefficient of determination.

Table 11 shows the performance of the ANN model based on socio-occupant factors using LM training functions. The model showed a MAPE of 28.2%, which can be classified as a reasonably accurate estimation. However, the model showed a higher MAPE than the previous model based on socio-appliance factors, which means that the current model had poorer performance in estimating monthly electrical energy consumption than the previous model. The value of CV in the model was more than 30%, which means that the model had a high variance. On the other hand, the  $R^2$  value in this model was less than 60%, which means that the correlation between the observed and estimated monthly electrical energy consumption, based on socio-demographic and occupant behavior factors, was quite low. Therefore, the current model showed poorer performance than models based on socio-demographic and appliance characteristic factors.

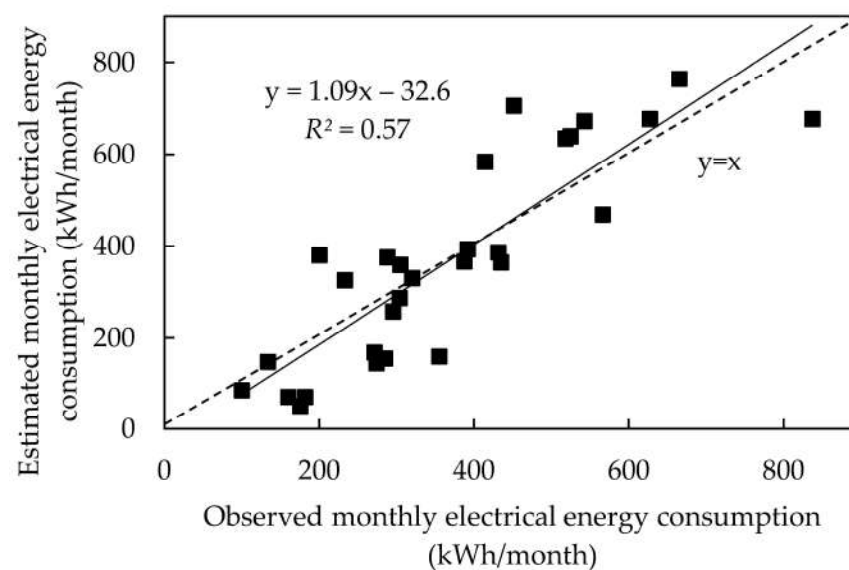
**Table 11.** Performance of socio-occupant model using the Levenberg–Marquardt model.

Parameter	Training Function
	Levenberg–Marquardt
Mean absolute error (MAE) (kWh)	99.5
Mean square error (MSE) (kWh)	13,620.9
Root mean square error (RMSE) (kWh)	34.7
Mean absolute percentage error (MAPE) (%)	28.2
Coefficient of determination ( $R^2$ )	0.53
Coefficient of variance (CV) (%)	32.3
Mean bias error (MBE) (kWh)	28.0

The ANN model based on socio-demographic and occupant behavior showed better variance than that of Kim et al. [37], with 53% and 11%, respectively. The different types of model input also contributed to the poorer performance of that of Kim et al. [37]. The current study showed less mean square error (MSE) than showed by Kim et al. [37], with an MSE of  $1.36 \times 10^4$  and  $7.69 \times 10^5$ . The impact of multicollinearity and non-linear relationships among the determinant factors could reduce the model performance as reported by Kim et al. [37]. Furthermore, the current model removed factors that had low significance and multicollinearity, as explained by Sena et al. [27]. However, the current model had fewer data training and testing sets than that of Kim et al. [37].

### 3.5. Performance of Model Based on Appliance Characteristics and Occupant Behavior Factors

The performance of the model based on appliance characteristics and occupant behavior factors was carried out in this section to examine the impact of those factors on the variance of monthly electrical energy consumption. An ANN model for electrical energy consumption was constructed based on the number of wall-mounted ACs, the usage of AC on weekdays, the number of standalone freezers, the number of flat-screen TVs, and the number of AC energy stars, the number of chargers for smartphones and the number of fluorescent lamps. The model of the one-step secant (OSS) training function could explain 57% of the variance, as shown in Figure 11.



**Figure 11.** Observed and estimated monthly electrical energy consumption for appliance characteristic and occupant behavior factors, based on the OSS training function. Note: x: observed monthly electrical energy consumption, y: estimated monthly electrical energy consumption,  $R^2$ : coefficient of determination.



Table 12 shows the performance of the ANN model based on appliance-occupant factors, using OSS training functions. The model showed a MAPE of 44.0%, which can be classified as a reasonably accurate estimation. However, the model showed a higher MAPE than the previous model based on socio-appliance and socio-occupant factors. The  $R^2$  value was 57%, which means that the correlation between the observed and estimated monthly electrical energy consumption was relatively low. However, the current model also showed higher variance than the previous model of socio-occupant factors. Thus, the results established that the current model based on appliance characteristics and occupant behavior showed a better monthly electrical energy consumption variance. Still, the model had higher errors than those of the model based on socio-appliance and socio-occupant factors. Furthermore, the CV value was more than 30%, which means that the model had a high variance.

**Table 12.** Performance of appliance-occupant model using the one-step secant training function.

Parameter	Training Function
	One Step Secant
Mean absolute error (MAE) (kWh)	93.8
Mean square error (MSE) (kWh)	12560.9
Root mean square error (RMSE) (kWh)	70.8
Mean absolute percentage error (MAPE) (%)	44.0
Coefficient of determination ( $R^2$ )	0.57
Coefficient of variance (CV) (%)	31.0
Mean bias error (MBE) (kWh)	26.4

Based on the OSS training function, the current model had a variance of 57% regarding monthly electrical energy consumption, which was better than that of Lee et al. [17] and Kim et al. [37], which only explained 42% and 11% of the variance, respectively. However, Lee et al. [17] showed less error than the current study and Kim et al. [37], with an MSE of  $1.05 \times 10^4$ ,  $1.26 \times 10^4$ , and  $7.69 \times 10^5$ , respectively. Similar to the previous models, the effect of multicollinearity and the non-linear relationship could reduce the model performance of Lee et al. [17] and Kim et al. [37].

### 3.6. Performance of Model Based on Socio-Demographic, Appliance Characteristic and Occupant Behavior Factors

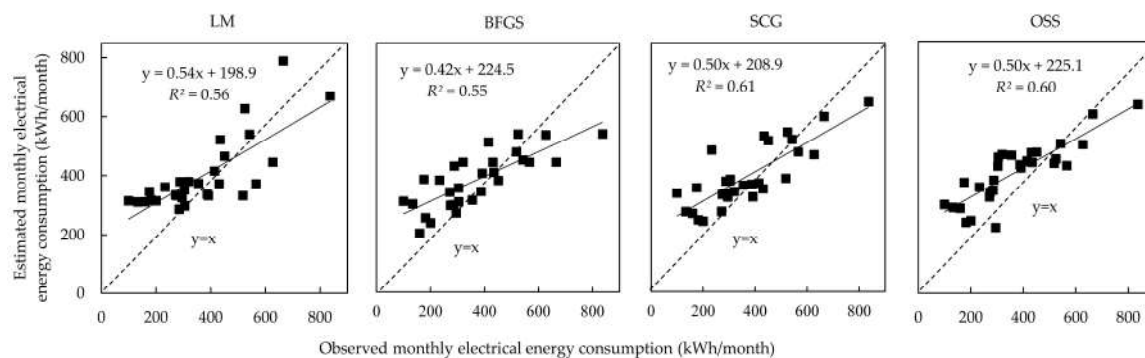
The performance of the model based on socio-demographic, appliance characteristic and occupant behavior factors were implemented in this section, to investigate the effect of those factors on the variance in monthly electrical energy consumption. ANN models for electrical energy consumption were developed based on the number of wall-mounted ACs, usage of AC on weekdays, the number of standalone freezers, the number of flat-screen TVs, and the number of AC energy stars, the number of chargers for smartphones, the number of fluorescent lamps and the level of education of the household head.

Table 13 shows the performance of the ANN model based on socio-appliance-occupant factors using LM, BFGS, SCG, and OSS training functions. The model based on the SCG training function showed a slightly better  $R^2$  than that of the model based on the OSS training function, by explaining 61% and 60% of the variance in monthly electricity electrical energy consumption, as shown in Figure 12. Similarly, the LM and BFGS training function model had an almost similar variance, with 56% and 55%, respectively. The value of  $R^2$  could be classified as a moderate performance for the model based on SCG and OSS, while the model based on LM and BFGS had low performance. Those models had a MAPE of less than 25%, which could be classified as a reasonably accurate estimation. However, those models' performances were still less good than those of the ANN model based on occupant behavior. However, these models still had better performance than those of the ANN model based on socio-appliance, socio-occupant and appliance-occupant factors. The value for CV was higher than 28%, which means that the models had high variance.

**Table 13.** Performance of the socio-appliance-occupant model, based on statistic analysis.

Parameter	Training Function			
	LM	BFGS	SCG	OSS
Mean absolute error (MAE) (kWh)	92.7	86.7	85.7	92.9
Mean square error (MSE) (kWh)	12,917.8	13,087.8	11,530.7	11,638.1
Root mean square error (RMSE) (kWh)	31.4	29.0	28.0	28.8
Mean absolute percentage error (MAPE) (%)	24.8	22.6	22.4	24.3
Coefficient of determination ( $R^2$ )	0.56	0.55	0.61	0.60
Coefficient of variance (CV) (%)	31.4	31.6	29.7	29.8
Mean bias error (MBE) (kWh)	26.1	24.4	24.1	26.11

Note: LM: Levenberg–Marquadt, SCG: scaled conjugate gradient, BFGS: Broyden–Fletcher–Goldfarb–Shanno, OSS: one-step secant.



**Figure 12.** Observed and estimated monthly electrical energy consumption for socio-demographic, appliance characteristic and occupant behavior models, based on LM, BFGS, SCG and OSS. Note: x: observed monthly electrical energy consumption, y: estimated monthly electrical energy consumption,  $R^2$ : coefficient of determination.

The ANN model based on socio-demographics, appliance characteristic and occupant behavior using the SCG training function had better variance than those of Lee et al. [17] and Kim et al. [37], with variances of 61%, 42% and 11%, respectively, as shown in Table 14. Due to the limitations of data training, the current model showed higher error than that of Lee et al. [17]. Still, the current model had less error than that of Kim et al. [37] because of the impact of multicollinearity in the determinant factors, with MSEs of  $1.05 \times 10^4$ ,  $1.19 \times 10^4$ ,  $7.69 \times 10^5$ , respectively. The current study removed multicollinearity and non-linear relationships based on Pearson correlation analysis, as explained by Sena et al. [27], while Lee et al. [17] and Kim et al. [37] did not remove these factors from the models.

**Table 14.** Performance of current study and those of Lee et al. [17] and Kim et al. [37].

ANN Type	Current Study		Lee et al. [17]		Kim et al. [37]	
	MSE	$R^2$	MSE	$R^2$	MSE	$R^2$
SCG	$1.15 \times 10^4$	0.61	$1.05 \times 10^4$	0.42	$7.69 \times 10^5$	0.11
OSS	$1.16 \times 10^4$	0.60	$1.04 \times 10^4$	0.40	$5.11 \times 10^5$	0.07
LM	$1.29 \times 10^4$	0.56	$1.09 \times 10^4$	0.38	$2.59 \times 10^6$	0.08
BFGS	$1.31 \times 10^4$	0.55	$1.04 \times 10^4$	0.38	$1.51 \times 10^6$	0.08

Note: LM: Levenberg–Marquadt, SCG: scaled conjugate gradient, BFGS: Broyden–Fletcher–Goldfarb–Shanno, OSS: one-step secant. MSE: mean square error,  $R^2$ : coefficient of determination.

#### 4. Discussion

The ANN model of electrical energy consumption, based on various determinant factors of techno-socio economics, has been developed using different types of training functions, such as the Levenberg–Marquadt (LM), scaled conjugate gradient (SCG), Broyden–Fletcher–Goldfarb–Shanno (BFGS), and one-step secant (OSS). However, the ANN model based on house characteristic factors showed poorer performance than the model based on occupant behavior, by explaining 49% and 81% of the variance, respectively.

Furthermore, the model based on occupant behavior showed less error than the model based on house characteristics for mean absolute error (MAE), mean square error (MSE), root mean square error (RMSE), mean absolute percentage error (MAPE), coefficient of determination ( $R^2$ ), and mean bias error (MBE). Thus, the results confirmed that the occupant behavior factors could moderately explain the monthly electrical energy consumption variance, rather than house characteristic factors.

Additional factors in the ANN model reduced the model's performance in estimating monthly electrical energy consumption. The combination of socio-demographic and appliance characteristics on the ANN model only achieved 59%. This was less than the ANN model, which only achieved 59%; this was less than that of the model based on occupant behavior factors, while the combined socio-occupant and appliance-occupant factors could explain 53% and 57% of the variance, respectively. On the other hand, a combination of socio-demographics, appliance characteristics and occupant behavior could explain 61% of the variance. Therefore, a combination of socio-demographics, appliance characteristics, and occupant behavior could explain 61% of the variance better than doubled-up combined factors on the model. Therefore, those ANN models using combined determinant factors could be classified as reasonably accurate estimates, based on less than 50% MAPE.

An ANN model based on socio-appliance-occupant factors showed less error than others based on the combined factors in MSE, RSME, CV and MBE. In contrast, the lowest MAE was demonstrated by the ANN model based on socio-appliance factors. The ANN model based on combined factors showed higher error than those of the models based on single factors such as house characteristics and occupant behavior. The results confirmed that the increase of the explanatory variable on the model tended to reduce the performance of the ANN model when estimating monthly electrical energy consumption.

The current ANN model showed better performance in estimating monthly electrical energy consumption compared to those of Lee et al. [17], Hawkins et al. [18], and Kim et al. [37]. Nevertheless, the current model showed a higher MSE than that of Lee et al. [17] and Kim et al. [37], due to the limitations of data testing. On the other hand, the current model showed a similar CV compared to Hawkins et al. [18], due to a similar capacity of data testing.

#### *Limitations and Future Study*

The current study had a limitation in the number of samples. The total datasets used in this paper were 214, divided into two sets: 184 for developing the ANN model and 29 for testing model performance. The current study only covered a population sample of students in the Universiti Teknologi Malaysia for 21,997 students, with a confidence level of 95% and a margin of error of 7% based on a sample calculation from Cochran [38].

The performance of the current model was also restricted due to the small number of samples. The model showed reasonable accuracy based on MAPE and explained 81% of the monthly electrical energy consumption variance. Therefore, the model only achieved moderate accuracy in estimating monthly electrical energy consumption. The ANN model was limited to a feed-forward multi-layer perceptron neural network with the activation function of a hyperbolic tangent. The model optimization was manually performed by randomly selecting several hidden layers, and the number was not used to classify the model performance. The calculation of model performance was performed based on the work of Buratti et al. [15], due to the limitations of computer specification. The classifications of the ANN model were based on five training functions: the Levenberg–Marquardt, Broyden–Fletcher–Goldfarb–Shanno, the scaled conjugate gradient, and one-step secant Bayesian regularization. Some results showed overestimated monthly electrical energy consumption, which was caused by the limitations of data training.

Future studies should increase the number of samples in Malaysia. Proportionate and random samples were stratified based on certain parameter levels such as income, occupation or educational background. Therefore, the model performance could be improved with a more specific classification. Future studies need to consider other approaches, such

as a top-down approach, a bottom-up approach based on engineering methods, and a hybrid model based on the grey box approach. The comparison between estimated and observed monthly electrical energy consumption could be developed into the baseline of Malaysian households as an essential energy management system. The developed model in this paper could be improved by using other scopes such as automation control or mechatronics to create intelligent appliances.

## 5. Conclusions

The research was intended to investigate the performance of electrical energy consumption based on the determinant factors of techno-socioeconomics from 184 items of data for developing the model and 29 items for the testing model. Therefore, the ANN model was selected to investigate the non-linear characteristic between determinant factors and monthly electrical energy consumption.

The results of the ANN model based on single factors showed the following:

- The model based on occupant behavior using scaled conjugate gradient explained 81% of the variance with a MAPE of 18%.
- The model based on house characteristics by Broyden–Fletcher–Goldfarb–Shanno explained 49% of the variance with a MAPE of 22.5%.

The ANN model based on double and triple combined factors was described by the following:

- The model based on socio-demographic and appliance characteristics factors using the one-step secant showed the best performance with a MAPE of 21% and an  $R^2$  of 0.59.
- The model based on socio-demographics and occupant behavior according to Levenberg–Marquardt explained 53%, with a MAPE of 28.2%.
- The model based on appliance characteristics and occupant behavior using the one-step secant presented at 57% with a MAPE of 44%.
- The model based on socio-demographics, appliance characteristics and occupant behavior using scaled conjugate gradient, explained 61% with a MAPE of 22.4%.

The ANN model based on occupant behavior factors showed the best performance compared to other models, which means that the factors could represent the monthly electrical energy consumption for Malaysian households. At the same time, the model based on house characteristics showed poorer performance than that of the model based on occupant behavior. The impact of socio-demographic and appliance characteristics was almost similar; however, in some models, appliance characteristics tended to have more effect on the monthly electrical energy consumption variance.

The findings of the current study recommended that the occupant behavior, socio-demographic and appliance characteristic factors should be considered for the additional information of determinant factors when designing an energy-efficiency strategy in the household. This is because those factors showed a moderately significant impact on the variance of monthly electrical energy consumption. In comparison, house characteristics did not significantly impact the variance of monthly electrical energy consumption. However, the house characteristic factors might have an indirect impact on other factors. Therefore, house characteristic factors could be moderately considered determinant factors, although the impact might not be too significant compared to other factors.

**Author Contributions:** Data curation, B.S.; Formal analysis, B.S.; Funding acquisition, S.A.Z., J.A.A.-R., F.L. and M.Z.H.; Investigation, B.S.; Methodology, B.S.; Project administration, S.A.Z.; Supervision, S.A.Z., N.M.Y. and F.Y.; Validation, B.S.; Writing—original draft, B.S.; Writing—review & editing, H.B.R. and N.M.Y. All authors have read and agreed to the published version of the manuscript.

**Funding:** This study is financially supported by the Takasago Thermal Engineering Co. Ltd. (Vot 4B424) and Matching Grant (Vot 01M89) projects of Universiti Teknologi Malaysia.

**Data Availability Statement:** The data presented in this study are available on request from the corresponding author.

**Conflicts of Interest:** The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

## Appendix A

**Table A1.** Techno-socio economic factors in the ANN model.

Variables	Questions	Descriptions
Education level	Please indicate the education level of the head of your household or the person who take responsibility for your family	(1) Secondary education (SPM) (2) Pre-university (Matriculation) (3) University
Number of energy stars	Please indicate the energy star for owned selected appliances (refrigerators, fans, televisions, and air conditioners)	Units
Number of rooms	Please indicate the number of rooms according to each type of room	Room
Total number of appliances	Please indicate how many appliances are in your house	Units
Set temperature of AC	Please indicate how you set the temperature of your AC appliance	Celsius

## References

- Olivier, J.G.J.; Peters, J.A.H.W. *Trends in Global CO<sub>2</sub> and Total Green House Gas Emissions: 2018 Report*; PBL Netherlands Environmental Assessment Agency: The Hague, The Netherlands, 2018.
- UN Environment and International Energy Agency. *Towards a Zero-Emission, Efficient, and Resilient Buildings and Construction Sector: Global Status Report 2017*; World Green Building Council: London, UK, 2017.
- Ho, C.S.; Chau, L.W.; Ling, H.T.; Rahim, M.N.A.; Zulibrahim, M.Z.; Hishammuddin, M.A.H.; Abdullah, R.; Saari, N.S.; Jausus, N. *A Greener Better Kuala Lumpur. A Summary for Policymakers*; Kuala Lumpur Low Carbon Society Blueprint 2030: Kuala Lumpur, Malaysia, 2018.
- Suruhanjaya Tenaga. *National Energy Balance*; Suruhanjaya Tenaga (Malaysian Energy Commission): Putrajaya, Malaysia, 2017.
- Ahmed, M.S.; Mohamed, A.; Homod, R.Z.; Shareef, H.; Khalid, K. Awareness on energy management in residential buildings: A case study in kajang and putrajaya. *J. Eng. Sci. Technol.* **2017**, *12*, 1280–1294.
- Zaki, A.S.; Hanip, M.N.F.; Hagishima, A.; Yakub, F.; Ali, M.M.S. Survey of resident behaviour related to air conditioner operation in low-cost apartments of Kuala Lumpur. *Chem. Eng. Trans.* **2018**, *63*, 259–264.
- Ranjbar, N.; Zaki, S.A.; Yusoff, N.M.; Hagishima, A. Time series data analysis of household electricity usage during El-Nino in Malaysia. In *Chemical Engineering Transactions*; Italian Association of Chemical Engineering-AIDIC: Milano, Italy, 2017; pp. 379–384. [\[CrossRef\]](#)
- Hisham, N.A.; Salim, S.A.Z.S.; Hagishima, A.; Yakub, F.; Saipol, H.F.S. Statistical analysis of air-conditioning and total load diversity in typical residential buildings. *Bull. Electr. Eng. Inform.* **2021**, *10*, 2299. [\[CrossRef\]](#)
- Aqilah, N.; Zaki, S.A.; Hagishima, A.; Rijal, H.B.; Yakub, F. Analysis on electricity use and indoor thermal environment for typical air-conditioning residential buildings in Malaysia. *Urban Clim.* **2021**, *37*, 100830. [\[CrossRef\]](#)
- Jones, R.V.; Fuertes, A.; Lomas, K.J. The socio-economic, dwelling and appliance related factors affecting electrical energy consumption in domestic buildings. *Renew. Sustain. Energy Rev.* **2015**, *43*, 901–917. [\[CrossRef\]](#)
- Kim, M.J. Characteristics and determinants by electrical energy consumption level of households in Korea. *Energy Rep.* **2018**, *4*, 70–76. [\[CrossRef\]](#)
- Chen, J.; Wang, X.; Steemers, K. A statistical analysis of a residential energy consumption survey study in Hangzhou, China. *Energy Build.* **2013**, *66*, 193–202. [\[CrossRef\]](#)
- Shahi, D.K.; Rijal, H.B.; Shukuya, M. A study on household energy-use patterns in rural, semi-urban and urban areas of Nepal Based on Field Survey. *Energy Build.* **2020**, *223*, 110095. [\[CrossRef\]](#)
- Roque, M. Survey and Artificial Neural Network Analysis on Occupant's Household Energy Use in a High-Rise Multi-Unit Residential Building in Toronto, Canada. Master's Thesis, Ryerson University, Toronto, ON, Canada, 2013.
- Buratti, C.; Orestano, F.C.; Palladino, D. Comparison of the energy performance of existing buildings by means of dynamic simulations and artificial neural networks. *Energy Procedia* **2016**, *101*, 176–183. [\[CrossRef\]](#)
- Khayatian, F.; Sarto, L. Application of neural networks for evaluating energy performance certificates of residential buildings. *Energy Build.* **2016**, *125*, 45–54. [\[CrossRef\]](#)
- Lee, S.; Jung, S.; Lee, J. Prediction model based on artificial neural network for user-based building energy consumption in South Korea. *Energies* **2019**, *12*, 608. [\[CrossRef\]](#)
- Hawkins, D.; Hong, S.M.; Raslan, R.; Mumovic, D.; Hanna, S. Determinants of energy use in UK higher education buildings using statistical and artificial neural network methods. *Int. J. Sustain. Built Environ.* **2012**, *1*, 50–63. [\[CrossRef\]](#)
- Pombeiro, H.; Santos, R.; Carreira, P.; Silva, C.; Sousa, J.M. Comparative assessment of low-complexity models to predict electrical energy consumption in an institutional building: Linear regression vs. fuzzy modeling vs. neural networks. *Energy Build.* **2017**, *146*, 141–151. [\[CrossRef\]](#)

20. Beccali, M.; Ciulla, G.; Brano, V.L.; Galatioto, A.; Bonomolo, M. Artificial neural network decision support tool for assessment of the energy performance and the refurbishment actions for the non-residential building stock in Southern Italy. *Energy* **2017**, *137*, 1201–1218. [[CrossRef](#)]
21. Zaki, S.A.; Hagishima, A.; Fukami, R.; Fadhilah, N. Development of a model for generating air-conditioner operation schedules in Malaysia. *Build. Environ.* **2017**, *122*, 354–362. [[CrossRef](#)]
22. Ranjbar, N.; Sheikh, A.Z.; Yusoff, N.M.; Yakub, F.; Hagishima, A. Short-term measurements of household electricity demand during hot weather in Kuala Lumpur. *Int. J. Electr. Comput. Eng.* **2017**, *7*, 1436. [[CrossRef](#)]
23. Swarno, H.A.; Zaki, S.A.; Hagishima, A.; Yusup, Y. Characteristics of wind speed during rainfall event in the tropical urban city. *Urban Clim.* **2020**, *32*, 100620. [[CrossRef](#)]
24. Sena, B.; Zaki, S.A.; Yakub, F.; Yusoff, N.M.; Ridwan, M.K. Conceptual framework of modelling for Malaysian household electrical energy consumption using artificial neural network based on techno-socio economic approach. *Int. J. Electr. Comput. Eng. (IJECE)* **2018**, *8*, 1844–1853. [[CrossRef](#)]
25. Wijaya, M.E.; Tezuka, T. A comparative study of households' electrical energy consumption characteristics in Indonesia: A techno-socioeconomic analysis. *Energy Sustain. Dev.* **2013**, *17*, 596–604. [[CrossRef](#)]
26. Azlina, A.A.; Kamaludin, M.; Abdullah, E.S.Z.E.; Radam, A. Factors influencing household end-use electricity demand in Malaysia. *Adv. Sci. Lett.* **2016**, *22*, 4120–4123. [[CrossRef](#)]
27. Sena, B.; Zaki, S.A.; Rijal, H.B.; Ardila-Rey, J.A.; Yusoff, N.M.; Yakub, F.; Ridwan, M.K.; Muhammad-Sukki, F. Determinant factors of electricity consumption for a Malaysian household based on a field survey. *Sustainability* **2021**, *13*, 818. [[CrossRef](#)]
28. 24 Units Single Storey Cluster Semi Detached Houses at Taman Tawas Jaya. Available online: <http://mansion-builders.com/taman-tawas-jaya/24-units-single-storey-cluster-semi-detached-houses-at-taman-tawas-jaya/> (accessed on 27 October 2019).
29. House for Sale in Miri Sarawak Malaysia. Available online: <http://emrealagency.blogspot.com/2011/04/house-for-sale-in-miri-sarawak-malaysia> (accessed on 27 October 2019).
30. Categories of Home in Malaysia. Available online: <https://www.propsocial.my/topic/226/categories-of-homes-in-malaysia-posted-by-prosocial-editor> (accessed on 27 October 2019).
31. Weather Company. Subang Airport, Selangor, Malaysia Weather History. Available online: <https://www.wunderground.com/history/daily/my/subang-airport/WMSA> (accessed on 31 January 2019).
32. Kubota, T.; Jeong, S.; Toe, D.H.C.; Ossen, D.R. Energy consumption and air-conditioning usage in residential buildings of Malaysia. *J. Int. Dev. Coop.* **2011**, *17*, 61–69.
33. Zaki, S.A.; Syahidah, S.W.; Shahidan, M.F.; Ahmad, M.I.; Yakub, F.; Hassan, M.Z.; Md Daud, M.Y. Assessment of outdoor air temperature with different shaded area within an urban university campus in hot-humid climate. *Sustainability* **2020**, *12*, 5741. [[CrossRef](#)]
34. Akbari, M.; Asadi, P.; Besharati-Givi, M.K.; Khodabandehlouie, G. Artificial neural network and optimization. In *Advances in Friction-Stir Welding and Processing*; Woodhead Publishing: Sawston, UK, 2014; pp. 543–599. [[CrossRef](#)]
35. Biswas, M.R.; Robinson, M.D.; Fumo, N. Prediction of residential building energy consumption: A neural network approach. *Energy* **2016**, *117*, 84–92. [[CrossRef](#)]
36. Tam, V.W.; Almeida, L.; Le, K. Energy-related occupant behaviour and its implications in energy use: A chronological review. *Sustainability* **2018**, *10*, 2635. [[CrossRef](#)]
37. Kim, S.; Jung, S.; Baek, S.M. A model for predicting energy usage pattern types with energy consumption information according to the behaviors of single-person households in South Korea. *Sustainability* **2019**, *11*, 245. [[CrossRef](#)]
38. Cochran, W.G. *Sampling Techniques*, 3rd ed.; John Wiley & Sons: New York, NY, USA, 1977.