

Article



Operational Energy and Carbon Cost Assessment Model for Family Houses in Saudi Arabia

Othman Alshamrani¹, Adel Alshibani^{2,3} and Awsan Mohammed^{2,4,*}

- ¹ College of Architecture and Planning Building Engineering, Imam Abdulrahman Bin Faisal University, Dammam 31451, Saudi Arabia; osalshamrani@iau.edu.sa
- ² Construction Engineering and Management Department, King Fahd University of Petroleum and Minerals, Dhahran 31261, Saudi Arabia; alshibani@kfupm.edu.sa
- ³ Interdisciplinary Research Center of Construction and Building Materials, King Fahd University of Petroleum and Minerals, Dhahran 31261, Saudi Arabia
- ⁴ Interdisciplinary Research Center of Smart Mobility and Logistics, King Fahd University of Petroleum and Minerals, Dhahran 31261, Saudi Arabia
- * Correspondence: awsan.mohammed@kfupm.edu.sa; Tel.: +966-13-8607575

Abstract: In Saudi Arabia, housing projects account for almost half of the total electricity consumed by the construction industry because of the large number of housing projects compared to other types of buildings. This paper proposes a quantitative approach using a multiple linear regression assessment model to predict the energy cost and environmental cost of housing building in Saudi Arabia. It was developed to assist house owners in Saudi Arabia in estimating the monthly energy cost and associated operational carbon cost according to several predictor parameters. Based on related literature, these parameters were reviewed and discussed by experts from the Ministry of Housing. They included building location, wall type, number of occupants, window type, envelope insulation, building age, building area, number of air-conditioning units and their systems, and lighting system. The model development process included five main stages: collecting the energy and carbon cost data from completed operating housing units, categorizing the collected data based on parameters, diagnosing the quality of gathered data and filtering outlier data if any, building and generating a model, and lastly, testing and validating the model. More than 77 datasets were collected across the country during different times of the year. The findings of this study reveal that the relationship between the number of users and the building area with the energy cost is significant and that the number of users is more correlated to the energy cost than the age of the building or the number of central air conditioners installed. Moreover, the results show that the developed model has the ability to predict energy and carbon costs with high accuracy. The developed model serves as a decision support tool for householders and decision makers in the Ministry of Housing to control the predicted parameters. This would be beneficial for the housing unit owners for allocating constrained budgets.

Keywords: energy cost; assessment model; regression analysis; energy parameters; carbon cost

1. Introduction

Buildings in Saudi Arabia are progressively demanding significant energy needs, particularly during the summer season, due to rising air-conditioning demands associated with extremely high outside temperatures across the country. Saudi Arabia's electricity expenditure accounts for over one-third of its total regular oil production [1]. The building sector uses about 80% of the overall electricity produced [2].

Moreover, Saudi Arabia's population is continuously increasing, and the economy is rapidly expanding, resulting in a surge in suburban construction. In 2016, the Saudi housing minister announced a new construction program that sought to build about 1.5 million new homes over the next seven to eight years. Presently, housing projects use approximately



Citation: Alshamrani, O.; Alshibani, A.; Mohammed, A. Operational Energy and Carbon Cost Assessment Model for Family Houses in Saudi Arabia. *Sustainability* 2022, *14*, 1278. https://doi.org/10.3390/su14031278

Academic Editor: Alejandro Rescia

Received: 13 November 2021 Accepted: 18 January 2022 Published: 24 January 2022

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



Copyright: © 2022 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/). half of the overall power use of the construction stock [3]. This is because of flaws in construction code, layout procedures, metropolitan layout, and building products [4–7]. In comparison to the worldwide construction industry, which uses around 40% of overall energy consumption, Saudi Arabia has the highest residential energy consumption. It is expected to rise annually by 5–8% [8]. By 2035, domestic oil production will equal the consumption of domestic oil. Thus, the Saudi government and major corporations developed the new program of Saudi Energy Efficiency in an effort to cut energy costs by 30%. The Saudi government announced the establishment of the Saudi Energy Efficiency Center (SEEC) in November 2010.

Furthermore, in April 2016, Saudi 2030 Vision was launched by the government focusing on emerging towns and on how to attain ecological sustainability [9]. However, while the government's leadership has been trying to reduce the existing energy intake, the energy consumption of buildings keeps increasing, leading to an increase in the energy cost of buildings.

Realizing the impact of buildings' energy consumption on overall energy costs, the government issued the Saudi Building Code in 2018. With such a policy, it hopes to address the perennial problem of high energy consumption and its associated costs.

A recent study has identified some of the issues that Saudi residential buildings face [1]. It noted that a large portion of their electricity consumption had gone toward AC units and around 70% of the constructed houses were not thermally sealed. In Saudi Arabia, 1, 811.5 million kWh of electricity is generated daily using four million barrels of oil, which represents half of the daily oil exports and one-third of daily oil production [1]. Furthermore, the building sector consumes 80% (649 million kWh) of the total energy produced daily which is the highest rate in the region.

A construction project entails several phases, including incorporation of intangible design, comprehensive design, development, building, and process and preservation as shown in Figure 1. Despite the availability of many types of commercial software, it is difficult to predict a project's power, construction, and maintenance costs precisely at the theoretical stage due to a lack and/or inaccuracy of information [1]. Data become available as the development progresses, allowing more precise cost assessments. As the design moves from the intangible to the tangible stage, adjusting it becomes a challenge [10–13].



Figure 1. Typical stages in a construction project.

The energy consumption in Saudi buildings has been studied extensively [14–16]. Researchers Taleb and Sharples [14] aimed to create strategies to provide sustainable housing in Saudi Arabia by researching the country's current energy and water consumption. This study endeavored to identify inherent design shortcomings in the existing Saudi buildings. According to Taleb and Sharples, there are several methods by which Saudi buildings can be made more energy-efficient. The insulation of exterior walls and roofs, the implementation of superior glazing systems, and smart lighting are among the methods needed. Ashraf and Almaziad [15] used an energy simulation program to assess the effects of a multilevel educational building's facades on its energy consumption. When it came to residential facilities, Alrashed and Asif [16] studied the impact of factors such as weather conditions, residence types, envelopes, cooking devices, and air-conditioning systems. Their study used a 2012 survey of 115 residences, including 25 traditional houses, 28 apartments, and 62 villas. They suggested using double glazing and a mini-split air-conditioning system to reduce energy consumption.

The previous studies indicated that there is scarcity of providing an accurate measurement tool for evaluating architectural designs that were energy-efficient at the conceptual design phase. In addition, there was a lack of computer-based applications, which could guide owners and architectural engineers in Saudi Arabia to select the most energy-efficient designs early in a project's conceptual phase to minimize energy costs except that of Alshibani and Alshamrani [1]. The literature shows that software available for modeling buildings had several limitations, including their lack of accuracy, the requirement for a substantial amount of data for modeling, which could be unavailable in the conceptual design phase, and the required time it took to finish the modeling process. The time taken depended on the buildings' and spaces' dimensions [17]. Information technology advancements have contributed to the development of many computer-based tools. A prominent example is artificial neural networks, one of the greatest commonly used techniques for developing prediction models across various fields [18,19]. There is a growing number of applications for predicting energy consumption that utilize Artificial Neural Networks (ANNs) [14,20–27].

ANNs are more effective and accurate at predicting the energy consumption of buildings than ordinary simulation models or regression techniques [23]. For instance, the Abductory Induction Mechanism (AIM), developed by Abdel-Aal et al. [20], estimates the domestic consumption of energy in Saudi Arabia's eastern region on a monthly basis. The model takes into account only demographic measures, economic indices, and weather parameters [21]. Nasr et al. created an artificial-neural-network-based model for electricity consumption based on weather-dependent variables and time series to predict electricity use in Lebanon. A hybrid growth model devised by Meng et al. [22] was used to predict China's monthly electrical energy consumption by extracting the trend of cumulative consumption. Although these models attempted to predict monthly electric energy consumption, they neglected to consider the architectural design of the building and its integration with the architectural elements. Using an ANN, Kumar et al. [24] assessed the heat load and total carbon emissions for a multi-story building. Karattasou et al. [25] suggested integrating statistical processes such as cross-validation, information criterion, and hypothesis testing into ANN models to improve performance.

Using a three-layered back-propagation ANN, Ekici and Aksoy [26] predicted the thermal energy requirement of different building types with a forecasting percentage at 94.8–98.5%. Mena-Yedra et al. [17] introduced a short-term ANN-based model to estimate the electricity demand at a solar energy research center and found admissible results. Along with ANNs, CAD tools have become advanced enough to estimate the building design parameters with ease, accuracy, and speed. Alshiabani and Alshamrani [1] developed an ANN/BIM model to predict the total electric energy costs of residential buildings in Saudi Arabia using a neural network. With the Kingdom's climate in mind, the system aimed to help architects design residential buildings to reduce energy costs. Many models and systems have been developed for different applications and jurisdictions for energy consumption in Saudi Arabia's building sector. Furthermore, regression [28] and neural-network-based models [29] were developed to predict the energy consumption in school facilities in Saudi Arabia. The findings indicate the capability of the proposed model to

estimate the energy consumption. Gao et al. [30] also developed a model based on the regression for estimating energy exchange efficiency. The results of the study indicate the suitability of the developed model to predict energy-exchange efficiency under different conditions. Mastrucci et al. [31] developed a statistical model based on a geographical information system to predict the consumption of energy of residential buildings. The important feature of the proposed model was the ability to predict consumption of energy on a large scale using an accurate, fast, and simplified tool without using a huge input dataset or assumptions. Gardezi et al. [32] developed a model based on multi-variant regression and a life cycle assessment method for embodied carbon footprint prediction in a residential building. The model was validated and made predictions against observed values of five. More housing units different from case studies also ensured efficiency and consistency.

Based on the aforementioned literature and to the best of the authors' knowledge, the previous developed models individually and collectively lack accurate prediction of the monthly energy footage and associated carbon emission of Saudi buildings. For residential structures, predicting the energy cost is a vital consideration, which can help the architect select the most energy-efficient alternative that meets practical objectives. During the conceptual design phase, engineers often have difficulty predicting power costs without adequate data.

This research is intricate due to the complexities in the curve of energy expenditure [33,34], along with incomplete data for the design of Heating, Ventilation, and Air Conditioning (HVAC) and illumination systems and envelope system. Consequently, this paper introduces a quantitative approach using a multiple-regression-based model to predict the monthly energy footage and associated carbon costs of Saudi housing buildings. The carbon cost is embodied in the model in order to meet the Saudi government's goal that is reflected in the decarbonation strategy which, consequently, will help in reducing the environmental impact, climate change, and global warming. In this paper, the factors influencing operational energy and carbon costs are identified. In addition, the correlation between the predictor parameters and the response factor is investigated. Moreover, the average monthly energy and carbon cost of housing projects in Saudi Arabia are predicted. The best regression model is then developed and validated using actual data from real-time projects. The main findings show that building area, central air conditioning, building age, and the number of users are the main factors that influence the model's prediction of energy and operational carbon costs.

The model can be used in other countries with climates similar to that of the Middle East and Gulf region to estimate the energy costs of residential buildings. However, the electricity tariff rate for a particular country needs to be taken into account. Since the factors that influence energy costs are identified and the correlation between these factors is investigated, the developed model can be used as a framework. The model could be modified based on the residential building's requirements for it to be adoptable for similar applications in other countries. Furthermore, the developed model could be retrained using new datasets rather than starting from scratch. This study is expected to serve as a milestone, assisting professionals and researchers in making quick, effective, and long-term energy cost assessments, decisions, and solutions.

2. Methodology

This study employed a real-time energy and carbon cost dataset of Saudi family houses to develop an energy and carbon cost assessment model. The average cost of energy for each month was determined by identifying and describing various input parameters. These parameters included wall, window type, number of occupants, envelope insulation, building age, building location, building area, air-conditioning type, the number of air conditioners, and lighting system. In this study, data gathered to build a model were as follows:

- Wall type: concrete block, clay, Siporex, and precast;
- Window type: single, double, and triple glazing;
- Envelope insulation: applied or not applied;
- Building area: 150 m^2 to 1000 m^2 ;
- Building age: 1 to 40 years;
- Building location: Dammam, Khobar, Qatif, Dhahran, Thoqbah;
- Number of users: 2 to 15 occupants;
- Air-conditioning type: window, split, and central units;
- Number of AC units: 1 to 16 units;
- Lighting systems: fluorescent, condescend, and LED.

To predict the monthly energy and carbon costs, different scenarios with various parameter combinations were run. The tested scenarios were categorized based on some changes to the alternatives for a single parameter while the other parameters remained constant. Figure 2a,b show the predicted costs for various window types and wall systems, respectively, whereas other parameters are held constant.



Clay block



Siporex

Wall system

Precast Panel

The developed model sought to diagnose any possible correlation between response variables (monthly energy and carbon cost) and predictor variables (energy and building factors). The model used multiple regression techniques to analyze the relationship between the monthly energy and carbon cost of Saudi family houses to develop a prediction model. Figure 3 depicts the development of the multiple linear regression model in five stages: the gathering of actual average monthly energy and carbon costs of existing houses,

the model.

6 of 18



Figure 3. Linear regression model development process.

This study used Grubbs' test, an outlier testing program, to identify the presence of outliers, as they have the potential to influence the model's predictive performance. According to the null hypothesis, the sample belongs to the same normal population, but the alternative hypothesis states that the extreme values are outliers. Results are shown in Figure 4 and Table 1. Based on the results, there are no outliers at a 5% significance level, and all data pertain to the same population.



Figure 4. Outlier plot of energy consumption cost in Saudi riyal.

Table 1. Grubbs' test results for outliers.

Parameter	Ν	Mean	Standard Deviation	Lowest Value	Highest Value	G	Р
Energy and Carbon Cost (SAR)	76	1643	1116	201	5124	3.12	0.098

A total of 77 house-building datasets were collected in Saudi Arabia, throughout the eastern region, to gather energy cost data. An assessment model was made using 85% of the 77 datasets (65 points) to predict the energy consumption of Saudi family homes, with the remaining 15% (12 points) chosen randomly to validate the developed regression model. Table 2 shows a sample of the data gathered in this study.

Project	Location	Area	Age	Insulation	Wall Type	Users No.	Window Type	WAC	SAC	CAC	Lighting System	Monthly Energy Cost
1	1	335	35	0	1	13	1	6	6	0	2	1075
2	1	500	3	0	1	6	1	0	11	0	2	348
3	1	700	3	0	2	7	2	0	9	0	2	942
4	2	550	20	0	3	4	1	4	2	0	2	588
5	1	237	6	1	4	5	2	0	8	0	1	298
6	1	250	3	1	1	5	1	0	7	0	1	225
7	5	250	2	0	1	5	2	0	9	0	3	400
8	3	550	5	1	3	6	2	0	5	0	1	1000
9	1	370	5	0	1	6	2	0	9	0	1	320
10	2	300	3	1	4	7	2	0	10	0	3	395
11	2	200	30	1	1	5	2	6	0	0	3	150
12	3	500	25	0	1	7	2	0	7	4	1	600
13	1	350	5	1	1	8	2	0	7	0	1	112
14	1	500	9	1	4	7	2	0	0	13	3	2000
15	3	300	3	0	4	4	2	0	10	0	1	253
16	1	340	4	0	1	5	2	0	8	0	1	220
17	1	400	6	1	1	8	2	3	15	0	3	1160

 Table 2. A sample of the collected energy costs data of projects.

Preliminary data analysis consisted of two major tests: calculating the best subset regression and identifying any data interaction and correlation. In order to develop a model, four steps were taken: generating the regression model, reviewing elementary factors, conducting the residual study, and selecting the validation model.

A correlation analysis was performed before the regression analysis, and the correlation coefficient "R" expressed the relationship between variables. Table 3 shows the correlation between the input factors. As can be seen from the table, the R-value between any two variables is less than 0.7. This indicates that all variables were part of the regression analysis.

	Α	В	С	D	Ε	F	G	Н	J	K	L
А	1.00										
В	0.09	1.00									
С	0.33	-0.07	1.00								
D	-0.08	0.08	-0.46	1.00							
E	-0.25	-0.09	-0.05	-0.18	1.00						
F	-0.13	0.06	0.14	-0.03	-0.05	1.00					
G	0.05	-0.03	-0.45	0.46	-0.03	-0.09	1.00				
Н	-0.09	-0.02	0.58	-0.39	-0.01	0.24	-0.56	1.00			
J	0.03	-0.03	-0.19	0.12	-0.02	0.14	0.21	-0.36	1.00		
K	-0.07	0.33	-0.26	0.30	0.10	-0.10	0.26	-0.30	-0.53	1.00	
L	-0.35	0.13	-0.16	-0.15	0.03	-0.01	-0.12	-0.14	0.01	-0.07	1.00

Table 3. Correlation coefficients among the input predictors.

Furthermore, scatter plots were used to visualize and analyze the data and determine the relationship between the input variables and their responses (i.e., energy cost). In addition to their simplicity, scatter plots can accurately show the linear as well as the non-linear relationship between two given factors. For instance, the relationship between the number of users and the building area with the energy cost was significant, as shown in Figure 5a,b. Moreover, Figure 5 reveals that the number of users was more correlated to the energy cost than the age of the building or the number of central air conditioners installed.



Figure 5. Cont.



Figure 5. Scatter Plot of Energy and carbon cost versus (**a**) area of building, (**b**) number of occupants, (**c**) age of building, and (**d**) number of central AC units.

3. Results and Discussion

The monthly energy and carbon costs of Saudi family houses were assessed using multiple linear regression modeling techniques. The developed model can assist building

owners in selecting building systems, materials, and architectural design that can help reduce the operating energy and carbon cost of their houses. The developed model can predict energy and carbon costs based on predictor variables. These variables include wall type, number of occupants, window type, envelope insulation, building age, building location, building area, air conditioner type, number of air conditioners, and lighting systems.

3.1. Best Subset Result

Table 4 displays the generated regression-based models in each line by analyzing the subset regression output. The lowest standard deviation (S) and Cp values are 12.0 and 239.4, respectively. *S* is the standard deviation of residuals, and Cp is expressed in the following equation:

$$C_{p} = \frac{SSE_{P}}{MSE(X_{1}...X_{p-1})} - (n-2p)$$
(1)

where C_p is used to evaluate the fit of a regression model estimated using ordinary least squares, SSE_P is the error sum of squares for the fitted subset regression model with p parameters (p-1 predictors), $MSE(X_1...X_{p-1})$ is the unbiased estimate of variance, and n is the number of observations.

Table 4. Best subset regression analysis.

Vars	R-Sq	R-Sq (Adj)	R-Sq (Pred)	Mallows Cp	S	Α	В	C	D	Ε	F	G	Н	J	К	L
1	31.6	30.6	27	177.5	460.52		Х									
1	29	27.9	25	186.8	469.32										Х	
2	49.3	47.8	45	117.2	399.53								Х		Х	
2	46.7	45	41.3	126.6	409.94						Х				Х	
3	64.9	63.2	60.8	64.5	335.18								Х	Х	Х	
3	62.9	61.1	58.3	71.7	344.79		Х						Х		Х	
4	74.5	72.9	69.6	32.7	288.01		Х				Х		Х		Х	
4	71.6	69.7	67.4	43	304.05		Х						Х	Х	Х	
5	80.3	78.7	76.4	14.1	254.9		Х				Х		Х		Х	Х
5	77.4	75.5	72.2	24.5	273.39		Х				Х		Х	Х	Х	
6	81.3	79.4	76.7	12.8	250.77		Х				Х		Х	Х	Х	Х
6	81	79.1	76.6	13.8	252.71		Х				Х	Х	Х		Х	Х
7	82.1	80	76.1	12	247.43		Х	Х			Х		Х	Х	Х	Х
7	82	79.8	76.3	12.5	248.39		Х	Х			Х	Х	Х		Х	Х
8	83.1	80.8	76.7	10.3	242.07		Х	Х	Х		Х		Х	Х	Х	Х
8	83	80.6	76.8	10.9	243.33		Х	Х			Х	Х	Х	Х	Х	Х
9	83.7	81.1	77.2	10.3	240.01		Х	Х	Х		Х	Х	Х	Х	Х	Х
9	83.4	80.8	76.4	11.4	242.29	Х	Х	Х	Х		Х		Х	Х	Х	Х
10	84.1	81.3	77.2	10.8	238.96	Х	Х	Х	Х		Х	Х	Х	Х	Х	Х
10	83.8	80.9	76.7	11.9	241.32		Х	Х	Х	Х	Х	Х	Х	Х	Х	Х
11	84.4	81.2	76.9	12	239.41	Х	Х	Х	Х	Х	Х	Х	Х	Х	Х	Х

The best subset also indicates that the peak values for R^2 and R^2 , (adjusted) in the developed energy and carbon cost assessment regression model, are 84.4% and 81.2%, respectively. According to the best subset regression analysis, all of the predictors in this model are significant and need to be in the regression model development.

3.2. Developed Regression Model

The R^2 value shows that the predictors account for 84.4% of the variance in "energy and carbon cost" (response variable) for the developed regression model. The R^2 (adjusted) accounts for the number of predictors in the model. Based on these two values, the model is most appropriate for the data. The developed linear regression model is:

$$AECC = (-698 - 28.1A + 1.074B + 11.8C + 120.9D - 18.5E + 76.6F + 139.7G + 80.7H + 22.4J + 88.5K - 167.1L) \times 1.83$$
(2)

where *AECC* is the average of monthly energy and carbon cost in Saudi riyals (SAR), *A* is the building location (Dammam 1, Khobar 2, Qatif 3, Dhahran 4, Thoqbah 5), *B* is the gross area varied (150 m² to 1000 m²), *C* is the building age (1 year–40 years), *D* is the insulation (applied 1, not applied 2), *E* is the wall system (concrete block 1, clay 2, Siporex 3, and precast 4), *F* is the number of users (2 persons to 15 persons), *G* is the window type (single 3, double 2, and triple glazing 1), *H* is the number of window air-conditioning units (1–16 units), *J* is the number of split air conditioners (1–16 units), *K* is the number of central air conditioners (1–16 units), and *L* is the lighting systems (condescend 3, fluorescent 2, and LED 1).

3.3. Assessment Tests of Model Adequacy

The analysis of the response factor (monthly energy and carbon cost) versus the predictors indicates that some predictor variables have a positive correlation (direct proportion) with the resulting cost, while others have a negative relationship (direct reflex). For example, a positive relationship exists among building age, building area, insulation, number of users, and air-conditioning types. A direct reflex correlation exists among wall type, building location, and lighting system versus the corresponding energy and carbon cost factor.

Based on the preliminary tests, the coefficient of determination, R^2 , and R^2 (adjusted) values are 84.4% and 81.2%, respectively. In the developed model, R^2 indicates that predictor variables explain 84.4% of the variance in the response variable (energy and carbon cost). R^2 (adjusted) is a modified value of R^2 associated with several terms of description in the developed model. The standard deviation value (S) is 12, and the R^2 value shows that the model is suitable for the data.

Furthermore, the T-test determines whether the predictor variables significantly correlate with the response variable or not. Most of the *p*-values for the estimated coefficients for each predictor in Table 5 are less than 0.05. Thus, in this case, the null hypothesis was rejected, whereas the alternative hypothesis was accepted. The majority of the predictor variables significantly correlate with the response variable "monthly energy and carbon costs" at 0.1 α -level. There is, however, a different pattern when it comes to other predictors, including location, wall type, and insulation, which have an insignificant correlation with the response factor. In addition, Table 5 shows *p*-values for the input variables. The analysis shows the significance of the developed model at 0.05 α -level. As a result, at least one coefficient should not be zero in the selected regression model. Some of the predictors have a high *p*-value which might not reflect the significance of these variables. However, as proven by the best subset analysis, they should still be included in the model to achieve the best predictions of energy and carbon cost.

The Pareto charts in Figure 6 with the absolute values of the standardized effects represent all of the standardized effects on the response; they are t-statistics. However, statistical significance was only found for effects that exceeded the dashed line. In Figure 6, the dashed line marks the 2.00 abscissa for a 0.05 significance level. The statistically significant variables and interaction factors in this test were the building area, AC systems, users, and building age. The strongest predictor was the building area (*B*) with *t*-value = 5.73, followed by central air conditioning (*K*) with *t*-value = 5.52, followed by window air conditioning (*H*) and the number of users (*F*) with *t*-values = 4.38 and 4.03, respectively. The factors of building age (*C*) and lighting system (*L*) had close t-values of \cong 2.6. Based on the *t*-test, factors such as wall type (*E*) predictor recorded the lowest significant correlation due to the expected similar u-values of a different wall system. The second-lowest significant correlated factor was location (*A*) since these distinct locations were within the same region and climate zone.

Term	Coef	<i>t</i> -Value	<i>p</i> -Value	VIF
Constant	-698	-2.89	0.006	
Loc.	-28.1	-1.38	0.174	2.04
Area	1.074	5.73	0.000	1.54
Age	11.80	2.79	0.007	2.42
Ins.	120.9	1.43	0.159	2.02
Wall	-18.5	-0.89	0.379	1.21
User	76.6	4.03	0.000	1.44
Win.	139.7	1.60	0.115	1.80
WAC	80.7	4.38	0.000	3.70
SAC	22.4	1.79	0.078	3.90
CAC	88.5	5.52	0.000	4.52
Light	-167.1	-2.57	0.013	1.76

Table 5. Statistical diagnostics of the developed regression model.

Figure 6. Pareto chart of the standardized effects.

3.4. Residual Analysis Result

The normal probability graph shown in Figure 7 indicates that all error terms follow an almost normal distribution. There are a few deviations from normality that can be observed in the residuals' normal probability chart. An analysis of these observations reveals that they are possible abnormalities. By excluding these abnormalities, the values of regression coefficients, such as R², can be improved. However, the model would not represent the real-world data available. The residual analysis results are accepted because minor deviations from normality are not a cause for concern [25].

Figure 8 shows how outliers appear in the histogram of the residual plot. Small bars on either side of the histogram indicate a high standard deviation from the mean value, which means the data for these values do not agree perfectly with the model. To estimate the probability of errors and outliers in the normal probability plots, the output of Minitab for the outliers and unusual observations was analyzed in this study. Table 6 presents standardized residual observations that have a considerable impact on the characteristics of the developed model. These observations influence the normal probability plot of residuals. Some of these observations are excluded from the model to improve the

regression model without affecting the process for predicting the operational energy and carbon cost (response factor).

Figure 7. Normal probability plot from residual analysis.

Figure 8. Histogram of the residual plots.

Table 6. Program output for unusual observations.

OBS	Energy Cost	Fit	Residual	Std Residual
29	2000	308	1692	2.93 R
41	3000	1769	1231	2.31 R
56	1590	3176	-1586	-3.23 R
67	1800	483	1317	2.26 R
68	1975	2633	-658	-1.81 X
77	4800	4470	330	0.70 X
78	5200	2822	2378	4.14 R

R Large residual, X Unusual X.

3.5. Economic Analysis

The economic analysis of the predictor parameters in the developed regression model shows that the highest energy and carbon cost is caused by applying different lighting types at 310 SAR followed by window types at 256 SAR. The triple glazing can reduce the cost by 512 SAR compared to the single one. The second-highest energy and carbon cost is recorded at 220 SAR when the thermal insulation is not applied in the external walls. One hundred and sixty-two rivals is the energy and carbon cost value that will be added for

each central air-conditioning unit while the window air-conditioning unit will increase the cost by 148 SAR. The developed regression model shows that every family member can increase the energy and carbon cost by 140 SAR. The location of the building will affect the cost by 50 SAR while each split air-conditioning unit will increase the cost by 41 SAR. Building a different wall system can help in reducing the cost by 34 SAR while the age of the building (each year) can increase the energy and carbon cost by 22 SAR. Finally, the outcome of the developed regression model shows that the energy and carbon cost will increase by 2.0 SAR/m², depending on the area of the house.

3.6. Model Validation

Model validation is the process of evaluating the proposed model's performance and accuracy with new data by applying it to new datasets. In this paper, twelve actual points of the energy cost and associated operational carbon cost from real projects were used to validate the developed model. These points represented 15% data points which were plotted to compare the in-hand gathered data and the predictive model. The validation relied on a mathematical model proposed by Zayed et al. [24]. It calculated the average invalidity and valid percentage to test the validity of data using the following formula:

$$AIP = \frac{\sum_{i=1}^{n} \left| 1 - \left(\frac{E_i}{C_i} \right) \right|}{n}$$
(3)

and
$$AVP = 1 - AIP$$
 (4)

where *AVP*, *AIP*, C_i , E_i , and *n* denote the rate of average validity, average invalidity rate, actual value, predicted value, and observation numbers, respectively. *AIP* value varies between 0 and 1.

0 0 10

$$AIP = \frac{0.249}{12} = 0.0208$$
$$AVP = 1 - AIP = AVP = 0.9791$$

As seen from the AVP value, the predicted model's accuracy is 97.9% which is adequate.

The validation plot chart in Figure 9 illustrates the comparison of predicted and actual operational energy and carbon costs. Based on the graph, predicted values are close to the real response values. The resulting outputs of the second validation test are, therefore, deemed acceptable.

Figure 9. The developed regression model's plot validation.

3.7. Model Testing and Training

A hypothetical case study was assumed in this study to test and train the developed multiple linear regression model for assessing the energy and associated carbon cost. A new family house in the city of Dammam with a medium area size of (500 m^2) was selected for this case study. The total number of users was seven persons, and the number of air conditioners was 10 units. The test was performed by applying different possible scenarios for window types, air-conditioning systems, lighting systems, and various wall systems while the insulation is becoming mandatory in the city. The various test scenarios were conducted using the developed regression model in this study as follows:

$$AECC = (-698 - 28.1A + 1.074B + 11.8C + 120.9D - 18.5E + 76.6F + 139.7G + 80.7H + 22.4J + 88.5K - 167.1L) \times 1.83$$
(5)

Various tested scenarios with different parameter combinations were conducted on the selected case study to measure the economic impact of different parameter alternatives. The tested scenarios were grouped according to some changes of the alternatives for one parameter while the other parameters were kept fixed. The comparisons were conducted based on the monthly predicted energy and carbon cost as well as the total assessed annual energy and carbon cost as shown in Table 7. The lowest annual energy and carbon cost for the different window scenarios is predicted for the triple-glazing window (10,780 SAR), while the highest cost is predicted for the single-glazing window (16,900 SAR). Hence, the triple-glazing window can reduce the cost by 36.2% compared to the single-glazing window. The lowest annual energy and carbon cost for the different wall-system scenarios is predicted for the precast concrete panel (18,700 SAR), while the highest cost is predicted for the clay block system (19,500 SAR). Hence, the precast concrete panel can reduce the cost by 4.1% compared to the clay block.

Wall System	Window Type	Air-Conditioning System Lighting Syster		Predicted Monthly Cost (SAR)	Annual Energy and Carbon Cost (SAR)				
Different window-type scenarios									
Concrete block	single	split	fluorescent	1410	16,900				
Concrete block	double	split	fluorescent	1155	13,850				
Concrete block	triple	split	fluorescent	900	10,780				
Different wall-system scenarios									
Clay block	triple	window	LED	1625	19,500				
Siporex	triple	window	LED	1590	19,100				
Precast Panel	triple	window	LED	1560	18,700				
		Different AC-s	ystem scenarios						
Clay block	double	split	condensing	1425	17,100				
Clay block	double	window	condensing	2490	29,900				
Clay block	double	central	condensing	2635	31,625				
Different lighting-system scenarios									
Precast panel	single	central	condensing	2825	33,880				
Precast panel	double	central	LED	1950	23,475				
Precast panel	triple	central	fluorescent	2000	24,075				

Table 7. Predicted energy and carbon cost for various tested scenarios.

The lowest annual energy and carbon cost for the different air-conditioning-system scenarios is predicted for the split AC unit (17,100 SAR), while the highest cost is predicted

for the central AC system (31,625 SAR). Hence, the split AC system can reduce the cost by 46.0% compared to the central AC system. The lowest annual energy and carbon cost for the different lighting-system scenarios is predicted for the LED system (23,475 SAR), while the highest cost is predicted for the condensing lighting system (33,880 SAR). Hence, the LED lighting system can reduce the cost by 30.7% compared to the central AC system. For the overall tested scenarios applying the developed regression model, it was found that the lowest annual predicted energy and carbon cost is recorded at (10,780 SAR), while the highest is recorded at (33,880 SAR). The total annual energy and carbon cost can be reduced by 68.2% when different systems are combined as presented in Table 7.

Hence, the developed multiple regression model can be a significant tool that will help homeowners in predicting and reducing their annual energy and carbon cost. It can be used as a supportive decision-making tool that will enable them in selecting the most economical and environmentally friendly system options as proven in the tested case study.

The limitations of this study include applying multiple linear regression techniques to create the operational energy and carbon cost. The assessment model was developed based on a limited number of collected data and could be used only in areas in eastern Saudi Arabia with a similar climate zone. The model is limited for application on family houses with the following features: gross area (150–1000 m²), age of the building (1–40 years), number of occupants (2 to 16 persons), and specific wall and window types. This study is limited to investigating the correlation between parameters (predictors) and predicted monthly energy and carbon costs.

4. Conclusions

This research proposes a quantitative regression-based model for predicting the monthly energy footage and associated carbon costs of Saudi buildings. The proposed model was adequately developed and verified using actual datasets. The results reveal that factors affecting the energy consumption and carbon costs of family-houses the most are large building area, central air-conditioning system, number of window air-conditioning units, number of users, age of the building, and lighting system. The model was tested and validated with a high accuracy of 97.9%, using mathematical and graphical validation. The developed regression model can serve as a decision support tool to enable building owners and decision makers in the Ministry of Housing to control the energy consumption, energy cost, and emitted carbon cost of Saudi family houses based on some predictor variables. Future work could extend the current research on energy and carbon cost to cover other types of buildings at different locations with different climate zones. Other modeling techniques, such as fuzzy logic, genetic algorithms, and hybrid methods, can be included. Research can incorporate more data to ascertain the correlation with other parameters, such as the wall system, thermal insulation type, and thickness. Moreover, the model's applicability would be enhanced further if it were integrated with other factors, especially if the model were used in other countries. Furthermore, energy prices are influenced by a variety of factors, including inflation. These elements could be incorporated into the model in the future.

It is also possible that the developed model can be extended to incorporate the Life Cycle Analysis (LCCA) model for economic analysis for the selection of architectural design with the least cost.

Author Contributions: Conceptualization, O.A.; methodology, O.A., A.A. and A.M.; software, O.A.; validation, O.A. and A.A.; formal analysis, O.A. and A.M.; investigation, A.A.; data curation, A.A. and O.A.; writing—original draft preparation, O.A.; writing—review and editing, A.M. and A.A.; supervision, O.A.; project administration, A.A. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: It can be made available upon request to the corresponding author.

Acknowledgments: The authors would like to thank King Fahd University of Petroleum and Minerals and Imam Abdulrahman bin Faisal University for the support and facilities provided to carry out this research.

Conflicts of Interest: The authors declare no conflict of interest.

References

- Alshibani, A.; Alshamrani, O.S. ANN/BIM-based model for predicting the energy cost of residential buildings in Saudi Arabia. J. Taibah Univ. Sci. 2017, 11, 1317–1329. [CrossRef]
- Abd-ur-Rehman, H.M.; Al-Sulaiman, F.A.; Mehmood, A.; Shakir, S.; Umer, M. The potential of energy savings and the prospects of cleaner energy production by solar energy integration in the residential buildings of Saudi Arabia. *J. Clean. Prod.* 2018, 183, 1122–1130. [CrossRef]
- KAPSARC [Internet]. Electricity Consumption by Sectors. Available online: https://datasource.kapsarc.org/explore/dataset/ electricity-consumption-by-sectors/SearchinGoogleScholar (accessed on 13 February 2020).
- 4. Lin, J.S.C.; Hsieh, P.L. Refinement of the technology readiness index scale: A replication and cross-validation in the self-service technology context. *J. Serv. Manag.* 2012, 23, 34–53. [CrossRef]
- Malik, A. Saudi Arabia Seeks Partners to Build 1.5 Million Homes. 2016. Available online: https://gulfbusiness.com/saudiarabia-seeks-partners-build-1-5-million-homes/ (accessed on 19 October 2021).
- Hovde, P.J. The factor method for service life prediction from theoretical evaluation to practical implementation. In Proceedings
 of the 9th International Conference on Durability of Building Materials and Components, Brisbane, Australia, 17–20 March 2002.
- Ginestet, S.; Marchio, D.; Morisot, O. Improvement of buildings energy efficiency: Comparison, operability and results of commissioning tools. *Energy Convers. Manag.* 2013, 76, 368–376. [CrossRef]
- Saudi Arabia, Public Utilities Report. February 2016. Available online: http://www.jeg.org.sa/data/modules/contents/uploads/ infopdf/2871.pdf (accessed on 19 June 2021).
- 9. Vision 2030, Saudi Arabia's Vision for 2030. 2018. Available online: https://vision2030.gov.sa/en/media-center (accessed on 19 June 2021).
- Pezzotta, G.; Sassanelli, C.; Pirola, F.; Sala, R.; Rossi, M.; Fotia, S.; Koutoupes, A.; Terzi, S.; Mourtzis, D. The Product Service System Lean Design Methodology (PSSLDM): Integrating product and service components along the whole PSS lifecycle. *J. Manuf. Technol. Manag.* 2018, 29, 1270–1295. [CrossRef]
- 11. Taylor, H.F.W. 1 Portland cement and its major constituent phases. In Cement Chemistry; Thomas Telford Ltd.: London, UK, 1997.
- 12. Halmetoja, E. The conditions data model supporting building information models in facility management. *Facilities* **2019**, *37*, 484–501. [CrossRef]
- 13. Kashkooli, A.M.S.; Vargas, G.A.; Altan, H. A semi-quantitative framework of building lifecycle analysis: Demonstrated through a case study of a typical office building block in Mexico in warm and humid climate. *Sustain. Cities Soc.* 2014, 12, 16–24. [CrossRef]
- 14. Taleb, M.H.; Sharples, S. Developing sustainable residential buildings in Saudi Arabia: A case study. *Appl. Energy* **2009**, *88*, 383–391. [CrossRef]
- 15. Ashraf, N.; Almaziad, F. *Effects of Facade on the Energy Performance of Education Building in Saudi Arabia*; World Sustainable Building: Barcelona, Spain, 2014; ISBN 978-84-697-1815-5.
- 16. Alrashed, F.; Asif, M. Trends in residential energy consumption in Saudi Arabia with particular reference to the Eastern Province. *J. Sustain. Dev. Energy Water Environ. Syst.* **2014**, *4*, 376–387. [CrossRef]
- 17. Mena-Yedra, R.; Rodriguez, F.; Castilla, M.M.; Arahal, M.R. A Neural Network Model for Energy Consumption Prediction of CIESOL Bioclimatic Building. In *International Joint Conference SOCO'13-CISIS'13-ICEUTE'13 Salamanca*; Springer: Cham, Switzerland, 2014.
- 18. Kehily, D.; McAuley, B.; Horem, A. Leveraging Whole Life Cycle Costs When Utilising Building Information Modelling Technologies. *Int. J. 3-D Inf. Model.* 2002, 1, 40–49. [CrossRef]
- 19. Kraus, W.E.; Watt, S.; Larson, P.D. *Challenges in Estimating Costs Using Building Information Modelling*; AACE International Transactions: Morgantown, WV, USA, 2007; pp. IT11–IT13.
- Abdel-Aal, R.; Al-Garni, A.Z.; Al-Nassar, Y.N. Modelling and forecasting monthly electric energy consumption in eastern Saudi Arabia using Abductive Networks. *Energy* 1997, 22, 911–921. [CrossRef]
- Nasr, G.E.; Badr, E.A.; Younes, M.R. Neural networks in forecasting electrical energy consumption: Univariate and Multivariate Approaches. In Proceedings of the Fourteenth International Florida Artificial Intelligence Research Society Conference, Key West, FL, USA, 21–23 May 2001; pp. 489–492.
- 22. Meng, M.; Shang, W.; Niu, D. Monthly electric energy consumption forecasting using multi-window moving average and hybrid growth models. *J. Appl. Math.* 2014, 243171. [CrossRef]
- 23. Al-Hamed, S.; Wahby, M.F. Prediction of potato yield based on energy inputs using artificial neural networks and C-sharp under Saudi Arabia conditions. *Biosci. Biotechnol. Res. Asia* 2016, 2, 631–644. [CrossRef]

- Kumar, R.; Aggarwal, R.K.; Sharma, J.D. Energy analysis of a building using artificial neural network: A review. *Energy Build*. 2013, 65, 352–358. [CrossRef]
- Karatasou, S.; Santamouris, M.; Geros, V. Modeling and predicting building's energy use with artificial neural networks: Methods and results. *Energy Build.* 2006, *8*, 949–958. [CrossRef]
- 26. Ekici, B.B.; Aksoy, U.T. Prediction of building energy consumption by using artificial neural networks. *Adv. Eng. Softw.* 2009, *5*, 356–362. [CrossRef]
- 27. Kumar, R.; Aggarwal, R.K.; Sharma, J.D. Estimation of total energy load of building using artificial neural network. *Energy Environ. Eng.* **2008**, *2*, 25–35. [CrossRef]
- Mohammed, A.; Alshibani, A.; Alshamrani, O.; Hassanain, M. A regression-based model for estimating the energy consumption of school facilities in Saudi Arabia. *Energy Build.* 2021, 237, 110809. [CrossRef]
- 29. Alshibani, A. Prediction of the energy consumption of school buildings. Appl. Sci. 2020, 10, 5885. [CrossRef]
- Gao, H.; Li, Z.; Qiu, S.; Yang, B.; Li, S.; Wen, Y. Energy exchange efficiency prediction from non-linear regression for membranebased energy-recovery ventilator cores. *Appl. Therm. Eng.* 2021, 197, 117353. [CrossRef]
- Mastrucci, A.; Baume, O.; Stazi, F.; Leopold, U. Estimating energy savings for the residential building stock of an entire city: A GIS-based statistical downscaling approach applied to Rotterdam. *Energy Build*. 2014, 75, 358–367. [CrossRef]
- 32. Gardezi, S.S.S.; Shafiq, N.; Zawawi, N.A.W.A.; Khamidi, M.F.; Farhan, S.A. A multivariable regression tool for embodied carbon footprint prediction in housing habitat. *Habitat Int.* **2016**, *53*, 292–300. [CrossRef]
- Abdel-Aal, R.E.; Al-Garni, A.Z. Forecasting monthly electric energy consumption in eastern Saudi Arabia using univariate time-series analysis. *Energy* 1997, 22, 1059–1069. [CrossRef]
- Alshamrani, O.S.; Alshibani, A. Automated decision support system for selecting the envelope and structural systems for educational facilities. *Build. Environ.* 2020, 181, 106993. [CrossRef]