American Journal of Mechanical and Industrial Engineering
2024, Vol. 9, No. 4, pp. 63-74
https://doi.org/10.11648/j.ajmie.20240904.11

Research Article

Modelling the Solar Intensity of Asaba Town in Nigeria Using Response Surface Methodology and Machine Learning Techniques

Akpenyi-Aboh Onyekachukwu Nicklette[1, *], Okafor Obiora Clement[2]

[1]Mechanical Engineering Department, Delta State University of Science and Technology, Ozoro, Nigeria
[2]Mechanical Engineering Department, Grundtvig Polytechnic, Oba, Nigeria

Abstract

The use of solar energy devices and machines such as solar pumps, heaters, solar cars, solar air conditioners, solar refrigerators, etc. has been the trend in most developed countries in the world owing to the green nature of solar technology. Based on the vast areas of application of solar technology, the knowledge of the solar intensity of a geographical location and having an estimation model to reduce the frequent measurement of solar-dependent factors are very imperative and helpful to the designers of solar devices and machines. This concern necessitated this research that was centered on determining the optimal value of the solar intensity of Asaba town, Delta state, and developing a mathematical model for estimating the solar intensity of the region. Meteorological data of the region was collected from the Nigerian Meteorological Agency (NIMET) for ten years (2011-2020). The data obtained from NIMET were: rainfall amount, relative humidity, mean temperature, and solar intensity. A multiple linear regression (MLR) model was developed using the collected data and the artificial neural network (ANN) model was also used to estimate solar intensity of the region. Response surface methodology (RSM) tool was employed to perform numerical optimization using the collected data of the region in quest of getting an optimal value of the solar intensity of Asaba town and the combinatorial best factor levels- rainfall, relative humidity, and mean temperature that would yield the optimum value of the response variable- solar intensity. Statistical tools such as mean bias error (MBE), mean percentage error (MPE), root mean square error (RMSE), Nash-Sutcliffe equation (NSE), correlation coefficient, test- statistics, and coefficient of determination (R^2) were used to evaluate the estimation performance of the Quartic polynomial model developed from the optimization process, multiple linear regression, and the artificial neural network models. From the results obtained, the optimal value of solar intensity was 759.687w/m^2 at the factor levels of: rainfall-194.58 mm, relative humidity-28.7989 mmHg and mean temperature of 25.7288 ℃. Also, the statistical validation tools applied revealed that the Quartic polynomial model had a much better performance characteristic than the other estimation models. This study would find application in the field of heating, ventilation, and air-conditioning (HVAC) solar systems by aiding the designers of such systems in knowing and estimating the value of the solar intensity of Asaba and other geographical regions that have similar climatic conditions.

Keywords

Solar Intensity, Response Surface Methodology, Artificial Neural Network, Least-square Regression Analysis, HVAC Systems

[*]Corresponding author: akpenyichimdi@gmail.com (Akpenyi-Aboh Onyekachukwu Nicklette)

Received: 1 August 2024; **Accepted:** 20 September 2024; **Published:** 31 October 2024

1. Introduction

The application of renewable energy systems in various fields has been a global trend owing to the green nature of such technology [1]. Countries, most especially developed ones, have been very much interested in exploiting to the fullest this energy type owing to its non-depletable state and low maintenance culture requirement. In addition, the harmful effect posed by fossil fuel burning on the environment is fast depleting the ozone layer, causing global warming and health-related issues to man. Renewable energy is an energy type that replenishes itself without depleting it through natural means. Among the non-conventional energy resources, solar energy, wind energy, and biomass have emerged as the most prospective options for the future [2]. Other forms of renewable energy are geothermal, hydro-energy, ocean tides, and waves, etc.

Solar energy is derived from the sun. Most recent technologies are hinged principally on solar energy applications. These technologies are seen in the invention of solar cars, solar heaters, solar pumps, solar refrigerators, solar air conditioners, etc. However, some designed solar cars are hybrid systems that still employ the aid of internal combustion (IC) engines in their operations. The intensity of solar energy is influenced by seasonal changes, geographical location, and position of the collector [3]. The knowledge of solar intensity for a particular geographical location is very paramount. This is because, as solar radiation passes through the atmosphere, it undergoes absorption, reflection, refraction, and scattering by various constituents of the atmosphere. The amount of solar radiation finally reaching the surface of the earth depends quite significantly on the concentration of airborne particulate matter, gaseous pollutants, and water (vapor, liquid, or solid) in the sky which can further attenuate the solar energy and change the diffuse and direct radiation ratio [4]. Global solar radiation is divided into two components: diffuse solar radiation and direct solar radiation [4]. The former results from scattering of the solar radiation by gases in the earth's atmosphere, dispersed water droplets, and particulates; while the latter refers to non-scattered solar radiation [4]. Hence, the idea of the amount of solar radiation intensity available for a region is necessary. Such knowledge would help the designers and installers of solar energy devices to improve their designs, enhance performance, and reduce running and maintenance costs by installing the required number of solar panels optimally needed to drive the energy needs of the region, and maximize the available solar radiation of the region.

Sequel to this concern, to help mitigate the frequent measurement of global solar radiation or solar intensity available for use, various estimation models using some meteorological factors like sunshine duration, extraterrestrial solar radiation/intensity, relative humidity, minimum temperature, maximum temperature, amount of precipitation or rainfall, etc. have been developed for various geographical locations by several researchers [5-8]. However, most of the research efforts were geared towards the development of estimation models without an insight into the optimal value of solar intensity or global solar radiation peculiar to the region and the factor combinations that would yield the optimal value. Trabea and Shaltout [9] carried out a study on the correlative assessment of global solar radiation measurements and the meteorological parameters using data from different parts of Egypt. The data employed were: solar radiation, mean daily maximum temperature, mean daily relative humidity, mean daily sea level pressure, mean daily vapor pressure, and hours of sunshine. The result obtained confirmed a good correlation coefficient between solar radiation and the other predictors. Nnabuenyi et al [4] developed an estimation model for Oko town in Anambra State using sunshine and temperature formulation models. Both models were validated using statistical tools and thus, were confirmed suitable for estimating global solar radiation of Oko. In addition, Awachie and Okeke [10] developed a prediction model that uses only the maximum ambient temperature for Nsukka town, Enugu State.

Therefore, this study was based on the determination of the optimal value of solar intensity of Asaba town, Delta State using meteorological data such as rainfall amount, relative humidity, mean temperature, and solar intensity collected from the Nigerian Meteorological Agency (NIMET) for ten (10) years. Also, solar intensity estimation models were developed using multiple linear regression (MLR), artificial neural network (ANN), and quartic polynomial formulation tools.

2. Materials and Method

2.1. Materials

The materials used in this study were the meteorological data of Asaba town obtained from NIMET for a period of ten years (2011-2020). The data include: monthly mean values of rainfall amount (mm), monthly mean values of relative humidity (mmHg), monthly mean values of temperature (℃), and monthly mean values of solar intensity (w/m^2). These data were collected in daily values and were converted to monthly mean values. The study employed these three variables because they are critical in the determination of the solar intensity of a region. Though, other meteorological variables can also be added to improve the estimation performance of the models developed in this study.

Also, the software tools employed were Design Expert 11.0, Minitab 2022, MATLAB 2017, Microsoft Excel 2020, and Python. The design expert and Python software were used for the response surface methodology (RSM) and machine learning optimization respectively, while MATLAB, Excel, and Minitab were employed in the model development and evaluation.

2.2. Method

The data collected from NIMET was studied and the RSM tool in design expert software 11.0 was used to model the optimum value of solar irradiance at a specific factor combination for Asaba town. Machine learning (ML) tool was also used to validate the RSM optimization output. Statistical analysis such as analysis of variance (ANOVA), sequential model sum of squares, etc. was also employed to determine the individual effect contributions of the independent variables (rainfall amount, relative humidity, and temperature) and their respective interactive effect on the response variable-solar intensity. A multiple linear regression (MLR) model was developed for estimating the solar intensity of Asaba town using the meteorological data obtained from the NIMET. An artificial neural network (ANN) tool was equally used to predict the solar intensity of the test site. The MLR and ANN were chosen for this study because they are known to have good estimation performances.

The results obtained from ANN, the developed MLR model, and the quartic polynomial regression model obtained from the RMS application were compared with the measured values of solar intensity (obtained from NIMET) to assess the efficiency and correctness of the models. The developed mathematical models were validated using some statistical measures like mean bias error (MBE), mean percentage error (MPE), root mean square error (RMSE), Nash-Sutcliffe equation, coefficient of correlation (COC), t-statistics test and coefficient of determination (R^2) to get the best-suited model for Asaba town. Graphical study was also employed to visualize, compare, and contrast the degree of association or variance between the measured values of solar intensity and estimated/predicted values obtained from the developed models- MLR, Quartic polynomial (QP) model, and ANN for Asaba town.

2.3. Development of a Multiple Linear Regression (MLR) Model

MLR model was developed for the estimation of solar intensity of Asaba town, Delta State. Unlike simple linear regression, MLR is a useful extension of simple linear regression (SLR) because it accommodates more than one independent variable or regressor. The MLR model is expressed in Equation (1).

$$S = a + b_1x_1 + b_2x_2 + b_3x_3 \tag{1}$$

Where: S= solar intensity (w/m^2), a = the intercept, b_1, b_2, b_3= regression coefficients, x_1 = rainfall amount (mm), x_2 = relative humidity (mmHg), x_3= average temperature (°C).

The empirical simultaneous equations used in computing the values of the regression coefficients are presented thus,

$$\sum x_1^2 b_1 + \sum x_1x_2b_2 + \sum x_1x_3b_3 = \sum x_1S \tag{2}$$

$$\sum x_1x_2b_1 + \sum x_2^2b_2 + \sum x_2x_3b_3 = \sum x_2S \tag{3}$$

$$\sum x_1x_3\, b_1 + \sum x_2x_3b_2 + \sum x_3^2b_3 = \sum x_3S \tag{4}$$

Where:

$$\sum x_i^2 = \sum x_i^2 - \frac{(\sum x_i)^2}{n} \tag{5}$$

$$\sum x_iS = \sum x_iS - \frac{(\sum x_i)(\sum S)}{n} \tag{6}$$

$$\sum x_ix_j = \sum x_ix_j - \frac{(\sum x_i)(\sum x_j)}{n} \tag{7}$$

n = number of data points = 120

The value of a was computed using equation (8) given thus,

$$a = \bar{S} - b_1\overline{x_1} - b_2\overline{x_2} - b_3\overline{x_3} \tag{8}$$

Where: $\bar{S}, \overline{x_1}, \overline{x_2}, \overline{x_3}$ = the mean values of solar intensity, rainfall amount, relative humidity, and temperature.

2.4. ANN Application

ANN model was used to predict the solar intensity of Asaba town and its result was compared to that obtained using the developed MLR and QP models. The NN used three data samples for its training, validation, and testing. The training samples were presented to the network during the training phase, and the network was adjusted according to its error. Validation samples were used to measure the network generalization and to halt training when generalization stopped improving. The testing samples do not affect training and so provide an independent measure of network performance during and after training. ANN uses three training algorithms: Levenberg-Marquardt, Bayesian regularization, and scaled conjugate gradient. Levenberg-Marquardt was used as the training algorithm in this study because it best suits the dataset. From the 120 data samples used in this study, 70%, 15%, and 15% of the samples were used for training, validation, and testing respectively. Figure 1 shows the ANN model used in the prediction operation comprising three (3) inputs, ten (10) hidden neurons, one output layer, and one output.

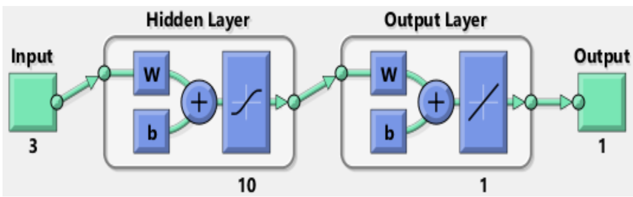

Figure 1. *Applied ANN model.*

The ANN applied in predicting the solar intensity of Asaba used a two-layer feed-forward network with sigmoid hidden neurons and linear output neurons capable of fitting multi-dimensional mapping problems.

2.5. Solar Intensity Optimization Using RSM and ML

To find the optimum factor combination that would maximize the response variable- solar intensity of Asaba town, RSM and ML tools were applied. The independent variables used were: rainfall amount, relative humidity, and temperature, while the response variable was solar intensity. The model output of RSM was used in the ML algorithm to validate the RSM results of the process. Table 1 shows the optimization criteria used to obtain the optimal value of solar intensity.

Table 1. *Optimization criteria used to obtain the optimal solution.*

Variables	Lower value	Upper value	Goal
Rainfall (mm)	5.4	367.7	In range
Relative humidity (mmHg)	27.10	341.40	In range
Temperature (℃)	16.45	37.765	In range
Solar intensity (w/m^2)	237.45	546.8	Maximize

2.6. Statistical Analysis of the Developed Models

2.6.1. Analysis of Variance (ANOVA)

Analysis of variance (ANOVA) is an analysis tool used in statistics that splits an observed aggregate variability found inside a data set into two parts: systematic factors and random factors. The systematic factors have a statistical influence on the given data set, while the random factors do not. Analysts use the ANOVA test to determine the influence that independent variables have on the dependent variable in a regression study. ANOVA analysis was used to study the individual and combined/interactive significant effects of the independent variables on the response variable- solar intensity.

2.6.2. Sequential Model Sum of Squares (SS)

The sum of squares is a statistical technique used in regression analysis to determine the dispersion of data points. In a, the goal is to determine how well a data series can be fitted to a function that might help to explain how the data series was generated. The sum of squares is used as a mathematical way to find the function that best fits (varies least) from the data. The sequential model sum of squares has three types- type 1, type 11, and type 111 depending on the model that outputs the optimal response value at any factor level considered. The sequential model sum of squares is summarily a measure of deviation from the mean value of the variables. It is mathematically computed using equation (9).

$$SS = \sum_{i=0}^{n}(S_{imea} - \bar{S}_{imea})^2 \qquad (9)$$

Where:
S_{imea} = ith value of measured solar intensity,
$\bar{S}_{imea}$ =Mean of measured values of solar intensity,
$S_{imea} - \bar{S}_{imea}$= deviation of each item from the mean value.

2.7. Validation Techniques for the Models

The developed model was properly assessed for correctness using the following statistical measures: mean bias error (MBE), mean percentage error (MPE), root mean square error (RMSE), Nash-Sutcliffe equation, coefficient of correlation (COC), t-statistics test and R-squared value. These statistical tools are lucidly discussed as follows.

2.7.1. The Mean Bias Error (MBE)

A low value of MBE attests or validates a good suitability for response prediction at any factor level considered. A negative value of MBE suggests an underestimation of the predicted value. The relation used for its computation is given by equation (10) [11].

$$MBE = \frac{1}{n}\sum_{i=1}^{n}(S_{ical} - S_{imea}) \qquad (10)$$

Where: n = number of data points, S_{ical} = calculated value of solar intensity, S_{imea} = measured value of solar intensity.

2.7.2. The Mean Percentage Error (MPE)

The correctness of a model is verified if its MPE value falls between -10% and 10%. MPE value is computed using equation (11) [12].

$$MPE(\%) = \frac{1}{n}\sum_{i=1}^{n}\left(\frac{S_{ical}-S_{imeas}}{S_{imeas}}\right) \times 100 \qquad (11)$$

2.7.3. The Root Mean Square Error (RMSE)

The smaller the value of RMSE the better the model's estimation strength and accuracy. Its computational formula is given by equation (12) [13].

$$RMSE = \left[\frac{1}{n}\sum_{i=1}^{n}(s_{ical} - s_{imeas})^2\right]^{\frac{1}{2}} \quad (12)$$

2.7.4. The Nash-Sutcliffe Equation (NSE)

The model's efficiency and correctness are assured if and only if NSE is very close to one. NSE is used for evaluating the predictive accuracy of models, particularly in assessing how well the predicted values from a model match observed data. The formula used to compute it as given by Chen et al (2004) is given by equation (13). [14]

$$NSE = 1 - \frac{\sum_{i=1}^{n}(S_{imea} - S_{ical})^2}{\sum_{i=1}^{n}(S_{imea} - \bar{S}_{mea})^2} \quad (13)$$

Where: $\bar{S}_{mea}$ = the mean measured value of solar intensity.

2.7.5. The Coefficient of Correlation (COC)

The coefficient of correlation, r measures the degree of association or relationship between the measured value of solar intensity and the calculated/estimated value. Models with values of coefficient of correlation closer to one (1) or even one is regarded as efficient models well suited for response prediction at any factor level considered. Karl Pearson's method was used to assess the developed model's association with the measured values of solar intensity. It was computed using the equation given thus.

$$r = \frac{\sum XY}{\sqrt{(\sum X^2)(\sum Y^2)}} \quad (14)$$

Where: X = the difference between the measured values of solar intensity and the mean of the measured solar intensity, and Y = the difference between the estimated solar intensity and the mean of the estimated solar intensity.

2.7.6. T-statistics Test

At a confidence interval of 95% and a significance level of 5%, the t-statistics test was carried out to determine how small its value was based on the fact that the smaller the value the better the performance of the model. In essence, the t-test evaluates whether the prediction model's error patterns are significant enough to suggest a need for adjustment or further improvement. The formula used to compute it is given by equation (15).

$$t = \left[\frac{(n-1)(MBE)^2}{(RMSE)^2 - (MBE)^2}\right]^{\frac{1}{2}} \quad (15)$$

2.7.7. The R-squared Value (R^2)

The coefficient of determination, R^2, is used to analyze how a difference in a second variable can explain differences in one variable. Models having a value of R-squared closer to one are considered the best model. It is also called the coefficient of determination. The formula used to compute its value is given as

$$R^2 = \frac{\text{sum of squares (ss)}}{\text{sum of squares of residuals}} = \frac{\sum_{i=0}^{n}(S_{imea} - \bar{S}_{imea})^2}{\sum_{i=0}^{n}(S_{imea} - \bar{S}_{ical})^2} \quad (16)$$

3. Results And Discussion

3.1. Multiple Linear Regression Model

The multiple linear regression model developed for estimating the solar intensity of Asaba town is given in equation (17)

$$S = 503 - 0.5353x_1 - 0.1655x_2 - 0.62x_3 \quad (17)$$

Where: S = solar intensity; x_1 = rainfall amount; x_2 = relative humidity; x_3 = average temperature. The significance of the coefficient of each factor in the model is shown in Table 2.

Table 2. *Significance of the model coefficients.*

Term	Coef	SE Coef	T-Value	P-Value	VIF
Constant	503	49.5	10.15	0.000	
Rainfall amount (mm)	-0.54	0.0575	-9.32	0.000	2.81
Relative humidity (mmHg)	-0.17	0.0855	-1.94	0.055	2.89
Temperature (℃)	-0.62	1.65	-0.38	0.707	1.22

From Table 2, the P-values of the constant term and rainfall amount were less than the chosen significance level (∝) of 5% for a confidence interval of 95%. Also, the other factors: relative humidity and temperature have their P-values greater than 0.05. This implies that the coefficients of rainfall amount and the constant term had a more significant effect in estimating the response values- solar intensity than the last two factors- temperature and relative humidity. To further decipher the degree of effect contributed by these predictors on the response, an analysis of variance (ANOVA) in Table 3 is shown.

Table 3 shows that the P-value of the MLR model is 0 which is less than the significance level of 5%, hence the developed MLR model is suitable for estimating solar intensity of Asaba town, Delta State. In addition, the P-value of rainfall was less than 5% but was greater than 5% (significance level) for other factors- relative humidity and temperature. This thus implies that rainfall amount had a more significant effect on solar intensity than the other two factors. Figure 2 shows the normal probability graph of residuals against their percentages. Residual is the difference between the measured and model-estimated values of solar intensity of Asaba town, Delta State. The degree of their fittings along the regression line suggests how well a model worked.

Table 3. *Analysis of variance (ANOVA) for the MLR model.*

Source	DF	Adj SS	Adj MS	F-Value	P-Value
Regression	3	508341	169447	110.48	0.000
Rainfall amount (mm)	1	133121	133121	86.80	0.000
Relative humidity (mmHg)	1	5753	5753	3.75	0.055
Temperature (℃)	1	217	217	0.14	0.707
Error	116	177914	1534		
Total	119	686255			

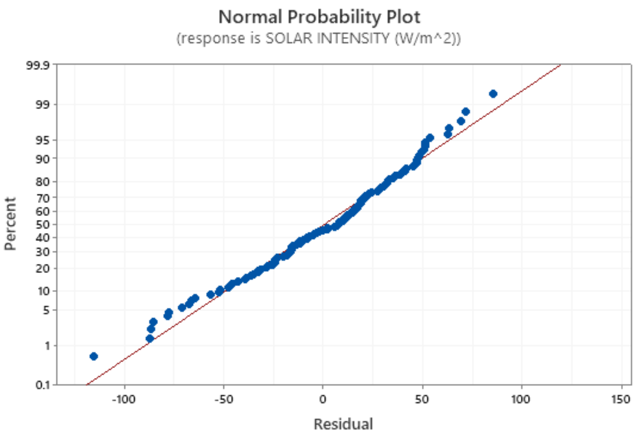

Figure 2. *Normal probability plot of residuals.*

From Figure 2, the residuals did not deviate much from the regression fit but can be seen to almost fit into the line. It further attests to the MLR model's suitability for response-solar intensity prediction/estimation.

3.2. Artificial Neural Network (ANN) Prediction

From the data training and retraining operation geared towards the quest of getting a better solution with a high coefficient of determination, eleven (11) iterations were done using ten (10) neurons and one (1) hidden layer. The performance sequence of the data training operation is given in Figure 3.

From Figure 3, for an eleven (11) epoch used during the iteration process, the best validation performance obtained was 1141.7841 at epoch 5. The detailed breakdown of the iteration process detailing the simulation points and the prediction behavior of the ANN model is shown in Figure 4.

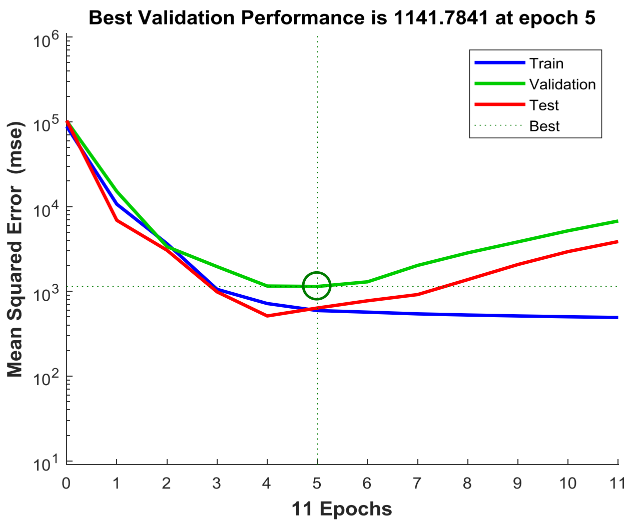

Figure 3. *Performance validation of the ANN model.*

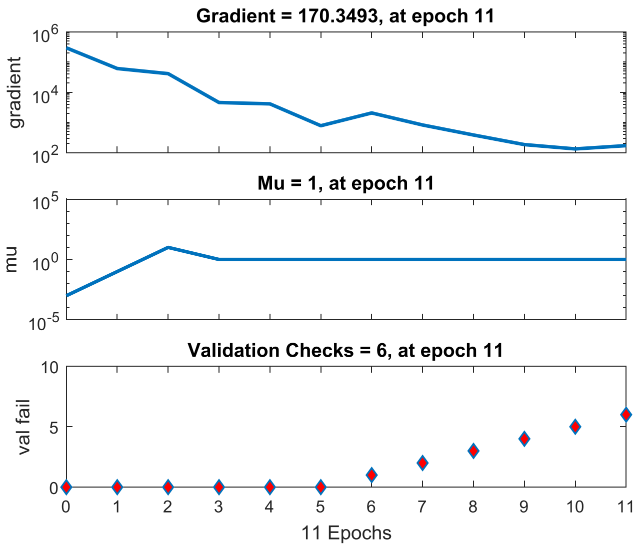

Figure 4*. Graphical plot of the iteration operation.*

Figure 4 also explains the sequence check for an optimal validation performance of the ANN model that improved its prediction/estimation efficiency. The gradient value of 170.3493 points to the degree of decline from epoch 1 to epoch 11 during the validation check in search of an optimal solution. A clearer insight into the ANN model's performance characteristics is revealed through the graphical plots of the training, validation, test, and overall output of the ANN model prediction analysis shown in Figure 5.

From Figure 5, the predicted value of solar intensity did not well align with the regression line. It could be seen vividly that some of the predicted values of solar intensity deviated to a certain degree from the regression fit. The training, validation, test, and overall output all had high values of correlation coefficient (R) of 0.94347, 0.92765, 0.94256, and 0.9385. This thus implied that the predicted values of solar intensity gotten from the ANN model and its measured value from NIMET are not too far beyond range, they are closely related to a certain degree.

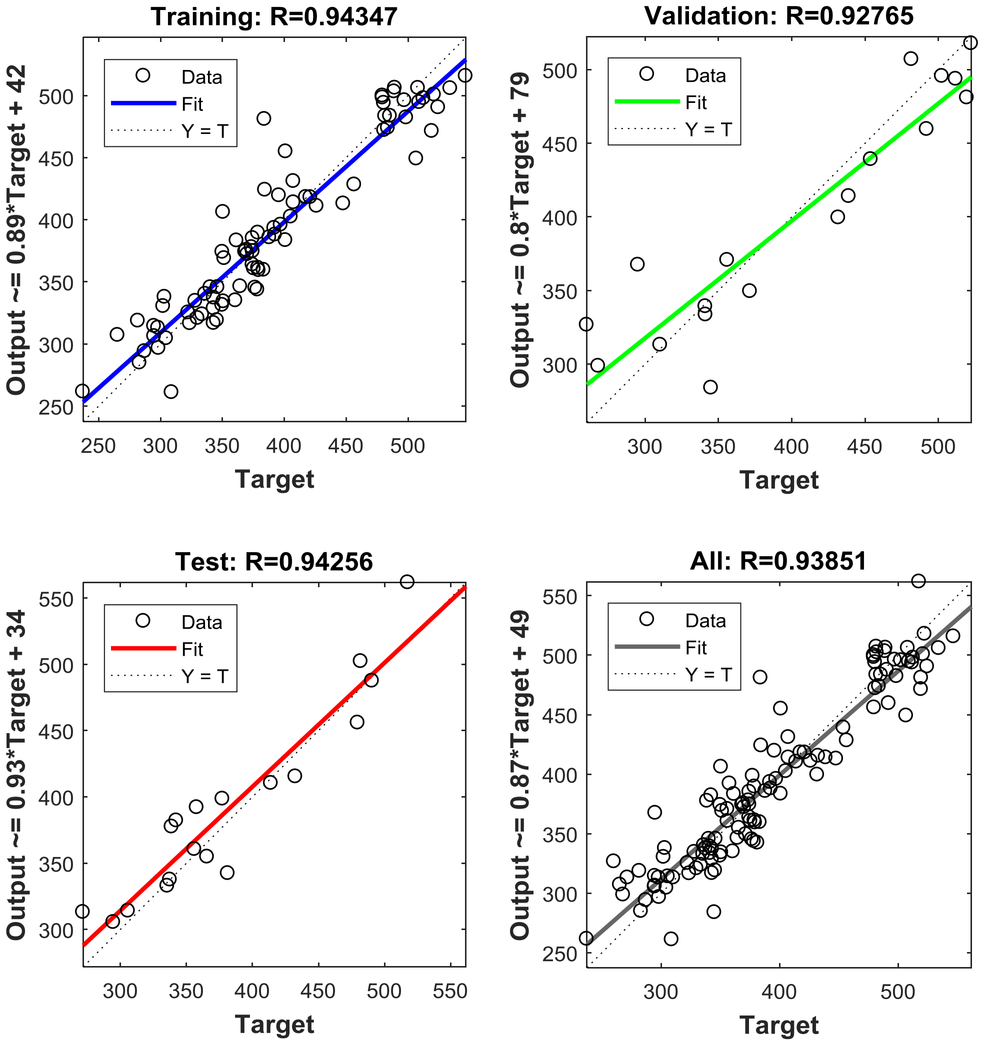

Figure 5*. Graphical plots of the training, validation, test, and overall output of the ANN model.*

3.3. Statistical Analysis Result

The result of the statistical analysis- sequential model sum of squares (SS) performed in quest of finding the best empirical model to be used for the optimization process is shown in Table 4.

From Table 4, quadratic-2 factor interaction (2F1) and quartic-cubic polynomial equation forms were suggested to be employed for the optimization process and their respective.

Table 4. Sequential model sum of squares (type 1).

Source	Sum of Squares	df	Mean Square	F-value	p-value	
Mean vs Total	1.839E+07	1	1.839E+07			
Linear vs Mean	5.083E+05	3	1.694E+05	110.48	< 0.0001	
2FI vs Linear	25270.58	3	8423.53	6.24	0.0006	
Quadratic vs 2FI	33195.13	3	11065.04	10.19	< 0.0001	Suggested
Cubic vs Quadratic	22939.18	10	2293.92	2.38	0.0143	
Quartic vs Cubic	34148.62	15	2276.57	3.10	0.0005	Suggested
Fifth vs Quartic	14859.93	21	707.62	0.9534	0.5289	
Sixth vs Fifth	17111.48	28	611.12	0.7240	0.8097	
Residual	30389.15	36	844.14			
Total	1.908E+07	120	1.590E+05			

Table 5. Model summary statistics.

Source	Std. Dev.	R^2	Adjusted R^2	Predicted R^2	PRESS	
Linear	39.16	0.7407	0.7340	0.7139	1.963E+05	
2FI	36.75	0.7776	0.7658	0.6696	2.268E+05	
Quadratic	32.95	0.8259	0.8117	0.7759	1.538E+05	Suggested
Cubic	31.07	0.8594	0.8326	-5.4840	4.450E+06	
Quartic	27.09	0.9091	0.8728	-164.0088	1.132E+08	Suggested
Fifth	27.24	0.9308	0.8713		*	
Sixth	29.05	0.9557	0.8536		*	

P-values were less than the significance level of 5% for a confidence interval of 95%. The exact polynomial equation to be used in the optimization was clearly defined in the model summary statistics presented in Table 5. From Table 5, both the quadratic and quartic (4th power) polynomial equations were suggested but the quartic polynomial equation was used for the numerical optimization because it had the highest value of R^2 and adjusted R^2 Values. The quartic model developed for response estimation is given in Equation (18). The model is in coded factors, not actual factors.

$$\begin{gathered}\text{Solar intensity} = 346.84-30.99A-52.60B+423.87C+125.76AB-523.49AC+790.79BC-27.97A^2-345.25B^2-2862.44C^2 \\ -3225.89ABC-62.90A^2B+1314.68A^2C+991.82AB^2+2005.53AC^2+3919.52B^2C-324.96BC^2-56.08A^3-821.64B^3 \\ -614.98C^3+127.29A^2B^2+6175.55A^2BC+798.67A^2C^2-7100.08AB^2C+3731.32ABC^2+1218.26B^2C^2+67.38A^3B \\ -1538.75A^3C+554.33AB^3+4069.85AC^3+52094B^3C-6631.39BC^3+21.44A^4-314.29B^4+738.63C^4 \end{gathered} \tag{18}$$

Where: A = rainfall amount, B = relative humidity, C = average temperature,

The equation in terms of coded factors can be used to make predictions about the response for given levels of each factor.

By default, the high levels of the factors are coded as +1 and the low levels are coded as -1. The coded equation is useful for identifying the relative impact of the factors by comparing the factor coefficients.

3.4. Numerical Optimization (RSM)

The selection of the optimum solution was based on the factor combination having the highest desirability value. Figure 6 shows the optimum factor levels that would yield the desired best value of solar intensity for Asaba town, Delta State.

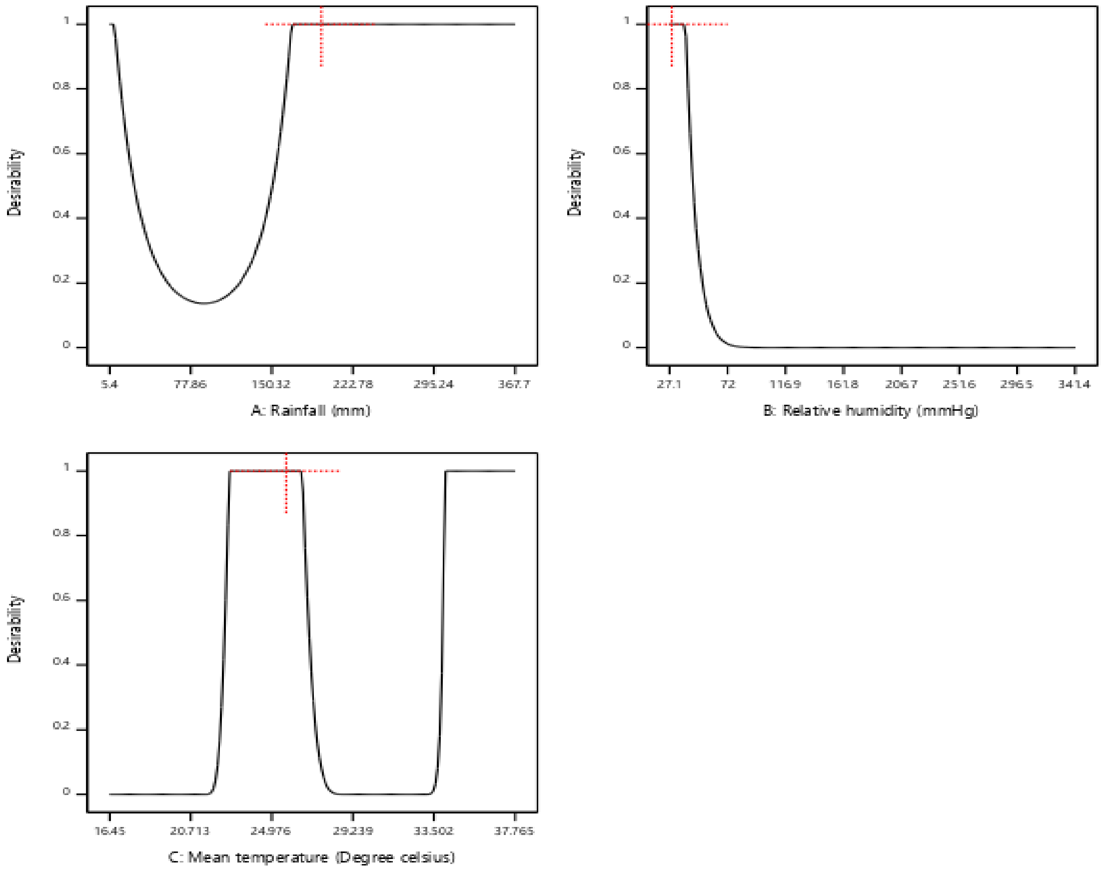

Figure 6. *Plot of desirability values against rainfall, relative humidity and mean temperature.*

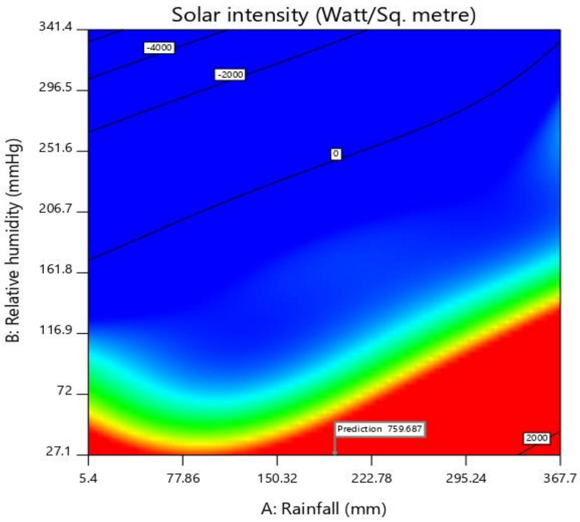

Figure 7. *Contour plot of solar intensity with the predictor variable.*

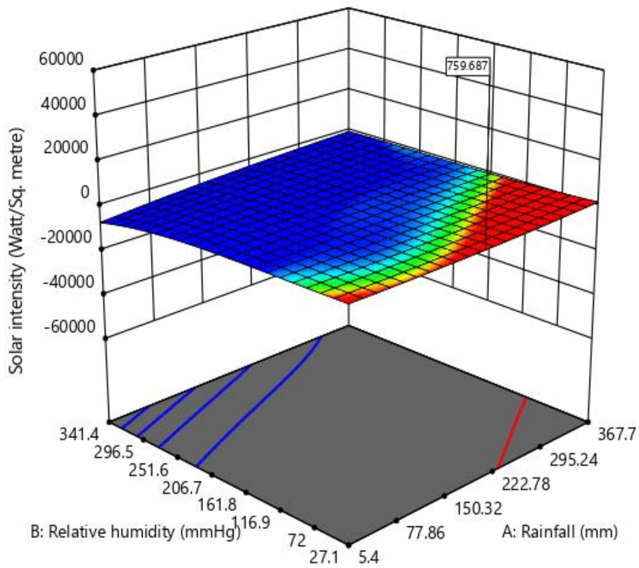

Figure 8. *3D plot of solar intensity and the predictors.*

From Figure 6, it is evident that the highest desirability value of one (1) was obtained at factor points of rainfall- 194.58 mm, relative humidity- 28.7989 mmHg and mean temperature- 25.7288℃. At these factor points, the optimal value of solar intensity would be estimated using the quartic polynomial model that was developed. Further on the response optimization outcome, the contour and 3D plots of the response and predictors are shown in Figures 7 and 8. Figure 7 clearly shows the estimated/predicted optimal value of the solar intensity of the considered test site at the obtained predictors' best values to be 759.687w/m^2. Figure 8 further attests that at the optimum factor combination obtained, the best value of the solar intensity of Asaba town is 759.687w/m^2.

3.5. Graphical Comparison of the Models

To visualize, compare, and contrast the estimated values of the three models developed- MLR, ANN, and quartic polynomial (QP) with that of measured values of solar intensity obtained from NIMET, their graphical relationships would be of immense aid. Figures 9, 10, and 11 show the plots of the MLR model estimates against the measured values of solar intensity; ANN model estimates against the measured values, and QP estimates against the measured values of solar intensity respectively.

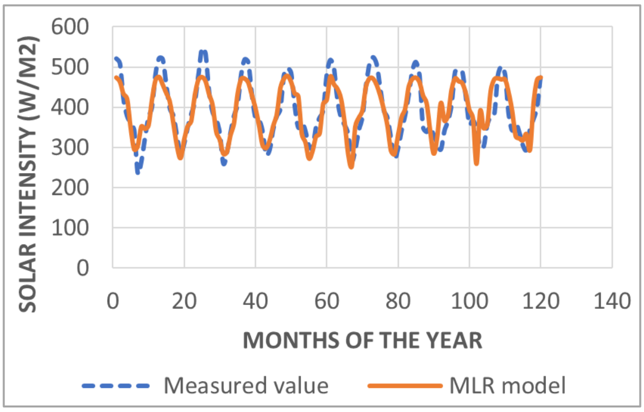

Figure 9. Plot of months of the year against solar intensity for both measured values of solar intensity and MLR model estimates.

From Figure 9, it is evident that the MLR model estimates did not well align with the measured values of solar intensity. The deviations seemed much but the model's performance characteristics would be ascertained better through the application of statistical validation tools employed for model error analysis. Figure 10 depicts a much better performance of the ANN model predictions/estimates than the MLR model. The fittings of the estimated points/values of the ANN model on the curve compared to that of the measured values appeared to have a good correlation coefficient.

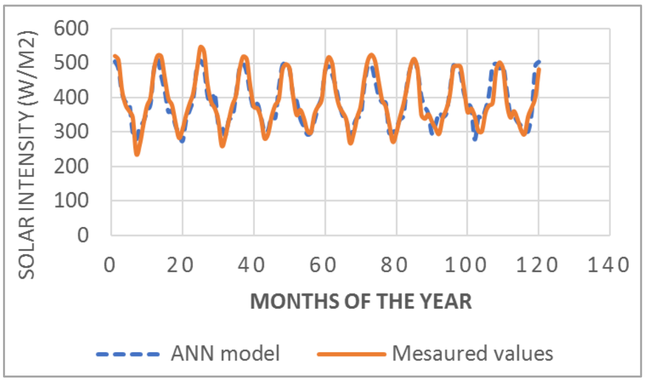

Figure 10. Plot of months of the year against solar intensity for both measured values of solar intensity and ANN model estimates.

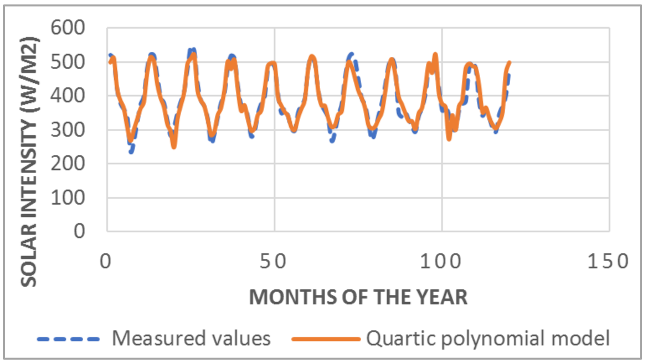

Figure 11. Plot of months of the year against solar intensity for both measured values of solar intensity and QP model estimates.

The QP model estimates compared to the measured values of solar intensity of Asaba town as shown in Figure 11 seemed to have a better performance characteristic than the MLR and ANN model estimations. This assertion was based on the good fittings of the estimated values of solar intensity (QP model) to the measured values. The best model was obtained through the application of statistical validation tools.

3.6. Model Validation Output

To help decipher lucidly the best model for estimating the solar intensity of Asaba town, statistical analysis was performed and the result is presented in table 6.

Table 6. *Model validation.*

Models	MBE	MPE	RMSE	NSE	COC (R)	T-STAT	R^2
MLR	-1.9326E-13	0.9421	38.5048	0.9905	0.861	5.4752E-14	0.7407
ANN	-1.4948	0.1719	28.2549	0.9949	0.9281	0.5166	0.8991
QPM	-1.6437E-13	0.37477	22.7963	0.9967	0.9530	7.8657E-14	0.9091

From Table 6, the Quartic polynomial (QP) model is the best model for estimating the solar incidence of Asaba town based on the following reasons:

1. It has a very low MBE value,
2. Its MPE value is between -10 to 10,
3. Its RMSE value is the smallest compared to the other models,
4. Its NSE value is closer to 1 than the other models,
5. It has the highest correlation coefficient,
6. It has the lowest t-stat value,
7. It has the highest value of the coefficient of determination, R^2.

4. Conclusions

This study focused on determining the optimum value of the solar intensity of Asaba town at specific values of the independent factors (rainfall amount, relative humidity and mean temperature) considered using the meteorological data of the region collected over ten years from the Nigerian Meteorological Agency (NIMET). The numerical optimization process using the response surface methodology approach (RSM) employed a Quartic polynomial (QP) model which had the highest value of adjusted. R^2compared to the quadratic model that was equally suggested. The analysis of variance (ANOVA) performed proved that each of the Quartic polynomial model's terms had an individual and combined interactive effect on the response variable. Other models were equally developed. The highest desirability of one (1) was achieved at the optimal factor levels of: rainfall-194.58 mm, relative humidity-28.7989 mmHg, and mean temperature of 25.7288°C. At this optimal factor level combination, the best value of solar intensity estimated by the Quartic polynomial model was 759.687 w/m^2. To further assess the Quartic model's performance, a multiple linear regression model (MLR) and a prediction model of artificial neural network (ANN) were developed and employed to compare and validate the performance characteristics of the Quartic polynomial model. The multiple linear regression model and artificial neural network models used rainfall amount, relative humidity, and mean temperature as the independent factors, also called predictors, and solar intensity as the response or target variable. The model terms of the developed multiple linear regression model all had a significant effect on the response variable with the variables- rainfall amount and the constant term having a greater effect on the response. The quartic polynomial model had a much better performance characteristic than the other models based on the statistical tools applied in validating the models. This was attested to by its: low value of MBE, MPE value falling between -10 to 10, lowest RMSE value, highest value of NSE, highest value of coefficient of correlation (R), lowest t-statistics value, and highest value of coefficient of determination (R^2). Hence it is more suitable for estimating the solar intensity of Asaba town than the multiple linear regression model and artificial neural network (ANN) applications.

Abbreviations

MBE	Mean Bias Error
MPE`	Mean Percentage Error
RMSE	Root Mean Square Error
NSE	Nash-Sutcliffe Equation
ANN	Artificial Neural Network
RSM	Response Surface Methodology
MLP	Multiple Linear Regression
QP	Quartic Polynomial Model
NN	Neural Network

Acknowledgments

The authors appreciate the management of the Nigerian Meteorological Agency (NIMET), Asaba, Delta State for approving the data used in this study.

Author Contributions

Akpenyi-Aboh, Onyekachukwu Nicklette: Conceptualization, Data curation, Investigation, Writing- original draft, and Writing- review and editing

Okafor, Obiora Clement: Formal analysis, Methodology, Software, Validation, Visualization

Funding

This work is not supported by any external funding.

Data Availability Statement

The data is available from the corresponding author upon reasonable request.

Conflicts of Interest

The authors declare no conflicts of interest.

References

[1] Achebe, C. H., Okafor, O. C., & Obika, E. N. (2020). Design and implementation of a crossflow turbine for Pico hydropower electricity generation. *Heliyon*, *6*(7), 1–13. https://doi.org/10.1016/j.heliyon.2020.e04523

[2] Banjade D., Poudyal K. N., Daponte P., Rapuano S. and Vito L. di, (2010), Estimation of Global Solar Radiation using Empirical Models at Benevento, Italy, National Seminar on Power and Communication Sectors Development (PCSD), Kathmandu, Nepal, 41-44.

[3] Abdulrahim, A. T., Diso, I. S., El-Jummah, A. M. (2011). Solar concentrators' developments in Nigeria: A review. *Continental Journal of Engineering Sciences*, Vol. 6(3), pp. 30-37.

[4] Nnabuenyi, H. O., Okoli, L. N., Nwosu, F. C., and Ibe, G., (2017). Estimation of global solar radiation using sunshine and temperature-based models for Oko town in Anambra State, Nigeria. *American Journal of Renewable and Sustainable Energy*, 3(2): 8-14.

[5] Okundamiya, M. S., Emagbetere, J. O., & Ogujor, E. A. (2015). Evaluation of various global solar radiation models for Nigeria. *International Journal of Green Energy*, *13*(5), 505–512. https://doi.org/10.1080/15435075.2014.968921

[6] Ezemuo, D. T., Ogunoh, C. and AguhP. S., (2017): Photovoltaic Solar Radiation Systems Analysis. *Journal of Scientific and Engineering Research*, 2017, vol.4(9), pp 93-106.

[7] Onyeka V. O., Nwobi-Okoye C. C., Okafor O. C., Madu K. E., Mbah O. M. (2021). Estimation of global solar radiation using empirical models. *Journal of Engineering Sciences*, 8(2), pp. 11-19.

[8] Mbah, O. M., Madueke, C. I., Umunakwe, R., & Okafor, C. O. (2022). Machine Learning Approach for Solar Irradiance Estimation on Tilted Surfaces in Comparison with Sky Models Prediction. *Journal of Engineering Sciences*, *9*(2), G1–G6. https://doi.org/10.21272/jes.2022.9(2).g1

[9] Trabea, A. A., & Shaltout, M. A. M. (2000). Correlation of global solar radiation with meteorological parameters over Egypt. *Renewable Energy*, *21*(2), 297–308. https://doi.org/10.1016/s0960-1481(99)00127-5

[10] Awachie, I. R. N and Okeke, C. E. (1990). New empirical solar model and its use in predicting global solar irradiation. *Nigerian Journal of Solar Energy*, vol. 9, pp. 143-156.

[11] Duffie, J. A., Beckman, W. A. (2013). Solar Engineering of Thermal Processes. John Wiley and Sons, New York, USA.

[12] Robaa, S. M. (2008). Evaluation of sunshine duration from cloud data in Egypt. *Energy*, Vol. 33, pp. 785-795.

[13] Glover, J., McCulloch, J. S. (1958). The empirical relation between solar radiation and hours of sunshine. *Journal of Royal Meteorological Society*, Vol. 84, pp. 172-175.

[14] Chen, R., Ersi, K., Yang, J., Lu, S., Zhao, W. (2004). Validation of five global radiation models with measured daily data in China. *Energy Conversion and Management*, Vol. 45, pp. 1759-1769.

Biography

Akpenyi-Aboh Onyekachukwu Nicklette, obtained her M. Sc. in Energy engineering from Chukwuemeka Odumegwu Ojukwu University. Currently, she is pursuing a PhD on energy systems engineering at the same University. She has contributed several outstanding journal and conference papers on the fields of renewable energy and HVAC systems. She lectures at Delta State University of Science and Technology, Ozoro.

Okafor Obiora Clement is a Ph.D. student at Nnamdi Azikiwe University, Awka, Nigeria. He is an academic staff of Grundtvig Polytechnic, Oba. He bagged first class honours in both his BSc. and MSc. degrees. He has contributed sixteen journal articles and four conference papers on the fields of thermo-fluid systems, renewable energy, HVAC systems, and system design and modelling.

Research Fields

Akpenyi-Aboh Onyekachukwu Nicklette: Renewable energy, Heat systems, Refrigeration systems Air-conditioning systems, Power plant technology

Okafor Obiora Obiora: Renewable energy, Machine design, Modelling and simulation, Machine learning, HVAC systems, Power plant technology