Nigerian Journal of Technology (NIJOTECH)
Vol. 43, No. 1, March, 2024, pp.131 - 138
www.nijotech.com

Print ISSN: 0331-8443
Electronic ISSN: 2467-8821
https://doi.org/10.4314/njt.v43i1.15

OPTIMIZATION AND MODELING OF SOLAR ENERGY WITH ARTIFICIAL NEURAL NETWORKS

AUTHORS:
B. E. Ntekim[1], and C. Uppin[2,*]

AFFILIATIONS:
[1,2]Department of Computer Science, Baze University, Abuja, Nigeria.

***CORRESPONDING AUTHOR:**
Email: cv.uppin@bazeuniversity.edu.ng

ARTICLE HISTORY:
Received: 13 September, 2023.
Revised: 13 January, 2024.
Accepted: 15 January, 2024.
Published: 31 March, 2024.

KEYWORDS:
Renewable energy, Solar energy, Artificial Neural Networks, Optimal Prediction, Radiation, Area of Panel, Hourly prediction, Average Monthly prediction, August Break, Mean Absolute Error, Mean Squared Error and Root Mean Squared Error

ARTICLE INCLUDES:
Peer review

DATA AVAILABILITY:
On request from author(s)

EDITORS:
Ozoemena Anthony Ani

FUNDING:
None

Abstract

Solar energy represents one of the emerging frontiers in renewable energy, offering significant potential to address the issues of energy unavailability and instability in Uyo (Nigeria). A crucial step in overcoming these challenges is accurately predicting the amount of solar energy that can be harnessed at a specific location. This research focused on achieving optimal solar power prediction, with the following objectives; identifying and investigating the mathematical relationships between relevant variables and parameters. To ensure precise predictions, artificial neural networks (ANN) were employed, utilizing both forward and backward propagation techniques. The input data for the ANN comprised radiation data obtained from a secondary source, the solar panel's size or area from the manufacturer, the panel's efficiency, and its performance ratio – all of which determined the electricity produced in kilowatts. The ANN was trained and tested using meteorological data, enabling accurate predictions of optimal electricity generation for the location. Notably, the hourly predictions reached their peak by 1 PM at the geographic location (5.2N and 7.5E), indicating that the highest levels of solar power were attainable during this daily period. Moreover, the pattern of monthly average solar power exhibited optimal predictions in January. Influenced by meteorological factors, a significant rise and fall in August, commonly referred to as the 'August Break' featured. The results demonstrated exceptional accuracy with minimal error margins (mean absolute error (MAE) of 0.03, mean squared error (MSE) of 0.0, and root mean squared error (RMSE) of 0.03). This high level of accuracy rendered the predictions reliable, making them suitable for consultancy services. Additionally, the potential for future work and expansion was evident, as the ANN could incorporate five or more years of radiation data for further improvements and insights.

1.0 INTRODUCTION

The demand for renewable energy is on the increase. Renewable energy is energy free of pollution and clean. It is environmentally friendly, unlike fossil fuel. Harnessing this form of energy from nature freely needs some technological skills. This report focused on solar energy. The sun is available everywhere, but not all the time just as the rain is not available all the time. How do we harvest this sun energy and get the optimal performance for usage? What roles do computer technologies play in making this energy system reliable, suitable, affordable, and efficient? Why is non-prediction a problem? These and other research questions were the focus of the study.

The keywords in this research were: optimal prediction, artificial neural networks (ANN) and solar

HOW TO CITE:
Ntekim, B. E., and Uppin, C. "Optimization and Modeling of Solar Energy with Artificial Neural Networks", *Nigerian Journal of Technology*, 2024; 43(1), pp. 131 – 138; *https://doi.org/10.4314/njt.v43i1.15*

energy. Prediction tells the installer/investor the amount of energy he can harness and the maximum or optimal amount in the geographical location of Uyo (5.2N and 7.5E). The ANN determined the high prediction accuracy [1] while the solar energy is the output which is the electricity [2]. The role of computer technology is the application of artificial neural networks in the modeling, optimization and accuracy of predicting optimal solar energy obtained from the solar panel [3]. The artificial neural networks technique is one of the deep learning algorithms that is compared to the workings of neurons in the human brain. In Section 1, related works from other authors are examined as relating to the methods used, their results and findings, and then their conclusion.

Artificial Neural Networks was defined by Zahraa, as a computerized intelligent system used in solving complicated problems in prediction, optimization, modeling, clustering, pattern recognition, simulation and others [4]. It is a data processing technique that uses mathematical functions connecting the input layer to the hidden layer and output layer to predict accurate optimal solar energy [5]. The inspiration for artificial neural networks comes from the knowledge of the human nervous system and how information is processed using a composed system of neurons connected with synapses that fire up other neurons. Signals like electrical pulses come from dendrites through axons which split into thousands of branches. At the end of each branch, synapses convert activity into exciting or inhibiting another dendrite of another neuron. Neurons fire when in excitation. The effectiveness of the synapses is the learning process [3].

Some types of neural networks are: - Artificial Neural Networks (ANN), Recurrent Neural Networks (RNN) and Convolutional Neural Networks (CNN). ANN works with structured data only, whereas RNN and CNN work with unstructured data. This study used ANN to perform complex regression analysis. Complex regression involves solving problems with many independent variables (inputs) to produce a single dependent variable (output) [3].

Zahraa, (2019) [4] used ANN to predict and validate daily solar radiation in three Egyptian cities, using basic backpropagation and backpropagation with momentum and learning rate algorithms methods. According to this method, backpropagation in ANN with momentum and learning rate algorithm was better than the basic backpropagation algorithm because it produced a higher precision in predicting solar radiation. Also, the model produced the lowest values of root mean square error (RMSE), mean absolute bias error (MABE), mean absolute percentage error (MAPE) and highest value of correlation (R^2) for the 3 cities where the study was conducted. Although expensive, further study can be done through using measuring equipment and maintenance for solar radiation [4].

Agbulut et al. (2021) [6] worked on the prediction of daily global solar radiation using different machine-learning algorithms to enhance investment decision-making. Data from various locations in Turkey were studied, evaluated and compared using support vector machine (SVM), artificial neural networks (ANN), deep learning (DL) and kernel nearest neighbor (KNN) techniques. In the training, maximum and minimum temperatures, cloud cover, daily extraterrestrial solar radiation, and length of day were inputs used in training the four algorithms by applying them through several metrics (correlation coefficient - R^2, mean bias error - MBE, mean absolute percentage error - MAPE, relative root mean square error - rRMSE, root mean square error - RMSE, mean absolute error – MAE, mean absolute bias error - MABE and t-stat with the following conclusions drawn:

1. Metric result of correlation (R^2) ranges between 85.5% and 93.6% for all the algorithms depending on location.
2. Metric results of MBE were negative in only two locations (Kirklareli and Nevsehir) using SVM but positive in all other locations for all the algorithms.
3. Metric result of MAPE varied between 15.9% and 30.24%, which is a good prediction.
4. rRMSE results varied between 14.10% and 25.19% for the four algorithms in all locations. These percentages were evaluated as good and fair predictions.
5. In comparing all the metric results for 4 locations, the best among them was Nevsehir province.
6. In general evaluation, DL and ANN techniques results were very close in correlation (R^2) and RMSE metrics.
7. Only metric results for t-stat. parameter showed the difference between DL and ANN in Tokat and Nevsehir provinces. This is because the DL algorithm has high MBE value making ANN very successful in these locations.
8. Error magnitude of ANN is very low in comparison with other algorithms in the study [7], [8].

Sriyignesh, (2021) [3] article applied the regression technique. The equation (1) explains the relationship involving these variables and parameters. To determine the slope coefficient, the covariance of the univariate data was evaluated using the equations. Artificial neural networks could produce power output with linear regression and non-linear regression. Vanilla neural networks worked with structured data while others worked well with unstructured data, like recurrent neural networks and convolutional neural networks. ANN technique processes data through three layers: input layer, hidden layer and output layer. Forward propagation was the process of multiplying weight with each input and then summation. Backpropagation was the process of updating the weights, which requires optimization and loss functions in the model [3]. This is the mathematical model shown in equation (1).

$$Y_i = \beta_0 + \beta_1 X_i + \varepsilon_i \tag{1}$$

Where; $\beta_1 = \frac{\sum(X_i - \bar{X})(Y_i - \bar{Y})}{\sum(X_i - \bar{X})^2}$; Y_i = Dependent variable $\beta_0 = Y$ intercept; β_1 = slope coefficient; X_i = independent variable; ε_i = random error term; $\bar{X}$ = mean of X; $\bar{Y}$ = mean of Y.

Equations for Regression Technique. (Source: Sriyinesh, 2021) [3]

The linear activation function and non-linear activation function as the case may be worked in the hidden layers to produce the output. The artificial neural network activation function works by selecting the most important data and suppressing less important data in the neural network [9]. The activation function plays the role of transforming the summed weighted input into output through the hidden layer and saved as output value which in turn is fed back into the hidden layer which is backpropagation. The regularizer in each layer controlled and prevented overfitting.

Sourcing for data is very important for this work, making cloud computing applications part of the review work. Uppin, (2018) [10] examined the wide usage of cloud computing technology in 134 organizations in India. From this research questionnaire, primary data were gathered, descriptively analyzed, cross-tabulated and using SSPS software to arrive at the following: various organizations like IT organizations used this technology because it saved cost and complexity. Secondly, it saved time used for developing new applications. Thirdly, 50% and 22% of respondents agreed and strongly agreed respectively of greater patronage of this technology soon. Limitations in the area of fewer providers of the technology and less awareness observed. Suggestion of more awareness and advantages of the technology needed. Mandatory training of the staff for effective usage and upgrading this technology is very necessary. This research work applied and benefitted from cloud computing technology [10] as the European Commission website provided databases and other useful services. Through the JRC-PVGIS application [11] weather datasets were obtained, saving time, cost and complexity in logistics.

George and Uppin, (2021) [12] employed a hybrid of Deep learning algorithms (CNN), traditional intrusion and Network monitoring systems to proactively detect malicious packets from genuine and log them as they occur in real-time. The method was summarized in a few steps: (1) Alert received from intrusion detection system (2) Wireshark technique captured or screenshot the packet and transmitted to CNN model (3) Classification of the images by CNN model if malicious packets, further actions are taken as – (i) Storing of packets concurrently in the folder (ii) Read-Only icon functional for preservation, (iii) logs report in text file, (iv) Admin decision after interpretation [12]. The expected results were a minimization of the cost and time of the investigation process. Also, it improves the quality of decision-making using Deep Learning and Artificial Intelligence. The research gap is in proactive detection, delay makes it reactive affecting resolution.

Photovoltaic Software, (2022) [13], examined the production of solar energy by applying a general global formula that identified the variables and the parameters used as area, radiation, efficiency and performance ratio. Investigation of the mathematical relationship of these variables and parameters produced the equation (2) [13]:

$$E = A * r * H * PR \tag{2}$$

Where 'E' is the energy, that is Power in KWh (kilowatt x time), 'A' was the size (area) of the radiation flux in the case study, the size is often stated on the panel by the manufacturer. Similarly, 'r' is the efficiency or yield of the panel. It is the amount of sunlight radiation (irradiance), intensity of sunlight incident on a solar panel that the panel can convert into usable electricity. Available literatures on the efficiency of solar panels with the best efficiency range between 15-23% and can last for 40 years [14]. The target power output was set at 22.80% efficiency yield, being the most efficient panels of 2022 designed by manufacturers [15] and the measured power output (Pmax) was set at an efficiency below 22.80%. 'H' was either hourly, daily or monthly averages of the radiation on the surface area in consideration. In this work, the solar radiation of the case study which is Uyo was obtained by locating the geographic coordinates (latitude 5.2N and Longitude 7.5E) of Uyo, from the European Commission website [11]. The determination of the radiation '*H*' also depends on the geographic location – determined by latitude and independent of longitude because different longitude on the same latitude has the same insolation [16]. "*PR*" was the performance ratio (PR), and coefficient of losses (ranges between 0.5 – 0.9, default value of 0.75).

2.0 MATERIALS AND METHODS

Identification and investigation of the mathematical relationship between the variables and parameters recalled from Section 1 (Equation 2) and used in the determination of solar power. The deployment of ANN for accurate prediction in the location was examined using the research design chart (Figure 1). From the chart, methods of data gathering, data processing, data splitting, model choosing/training and predicting were discussed.

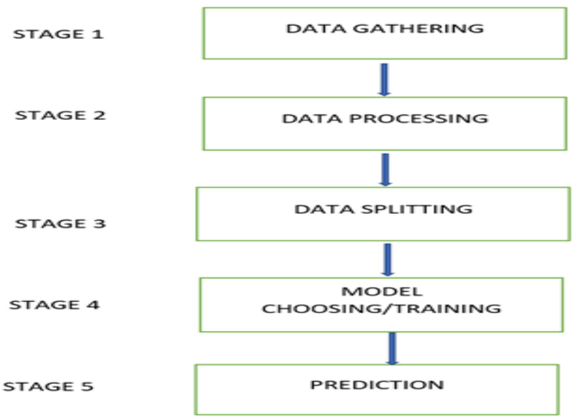
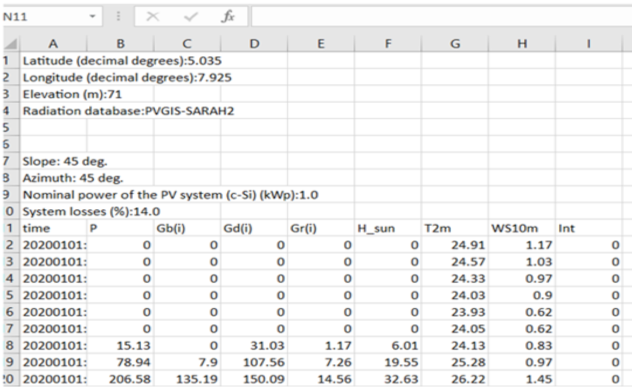


Figure 1: Flow chart of the research design


N11

	A	B	C	D	E	F	G	H	I
1	Latitude (decimal degrees):5.035								
2	Longitude (decimal degrees):7.925								
3	Elevation (m):71								
4	Radiation database:PVGIS-SARAH2								
5									
6									
7	Slope: 45 deg.								
8	Azimuth: 45 deg.								
9	Nominal power of the PV system (c-Si) (kWp):1.0								
0	System losses (%):14.0								
1	time	P	Gb(i)	Gd(i)	Gr(i)	H_sun	T2m	WS10m	Int
2	20200101:	0	0	0	0	0	24.91	1.17	0
3	20200101:	0	0	0	0	0	24.57	1.03	0
4	20200101:	0	0	0	0	0	24.33	0.97	0
5	20200101:	0	0	0	0	0	24.03	0.9	0
6	20200101:	0	0	0	0	0	23.93	0.62	0
7	20200101:	0	0	0	0	0	24.05	0.62	0
8	20200101:	15.13	0	31.03	1.17	6.01	24.13	0.83	0
9	20200101:	78.94	7.9	107.56	7.26	19.55	25.28	0.97	0
0	20200101:	206.58	135.19	150.09	14.56	32.63	26.22	1.45	0

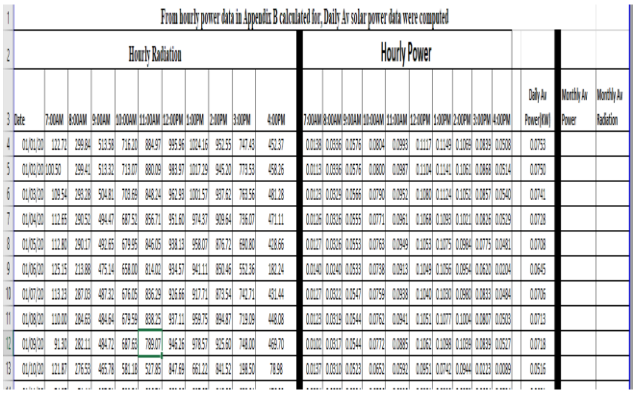
Figure 2: Sample of the original online source dataset

2.1 Data Gathering

European Commission website provided the needed datasets for this work, which is the application of cloud computing [10]. From this online source, the dataset contained 79056 data by 9 columns of the year 2020 being a leap year of 366 days. Column 1 contained the hours and dates data, column 3 contained direct radiation data, column 4 contained diffused radiation data and column 5 contained reflected radiation data. Other columns contained various weather variables not important in this work but are recommended for further research work. Figure 2 was a sample of the original online source dataset while Figure 3 was the Python code of cleansed dataset.

```
df = pd.read_excel(io.BytesIO(uploaded['DATA4D.xlsx']))
```

Figure 3: Python code


From hourly power data in Appendix B calculated for, Daily Av solar power data were computed

	Hourly Radiation											Hourly Power												
3	Date	7:00AM	8:00AM	9:00AM	10:00AM	11:00AM	12:00PM	1:00PM	2:00PM	3:00PM	4:00PM	7:00AM	8:00AM	9:00AM	10:00AM	11:00AM	12:00PM	1:00PM	2:00PM	3:00PM	4:00PM	Daily Av Power(KW)	Monthly Av Power	Monthly Av Radiation
4	01/01/20	122.72	299.84	513.58	716.20	884.97	995.96	1024.16	952.55	747.43	452.37	0.0138	0.0336	0.0576	0.0804	0.0993	0.1117	0.1149	0.1069	0.0839	0.0508	0.0753		
5	01/02/20	100.50	299.41	513.32	713.07	880.09	983.97	1017.29	945.20	773.53	458.26	0.0113	0.0336	0.0576	0.0800	0.0987	0.1104	0.1141	0.1061	0.0868	0.0514	0.0750		
6	01/03/20	109.54	293.28	504.81	703.69	848.24	962.93	1001.57	937.62	763.56	481.28	0.0123	0.0329	0.0566	0.0790	0.0952	0.1080	0.1124	0.1052	0.0857	0.0540	0.0741		
7	01/04/20	112.65	290.52	494.47	687.52	856.71	951.60	974.37	909.64	736.07	471.11	0.0126	0.0326	0.0555	0.0771	0.0961	0.1068	0.1093	0.1021	0.0826	0.0529	0.0728		
8	01/05/20	112.80	290.17	492.65	679.95	846.05	938.13	958.07	876.72	690.80	428.66	0.0127	0.0326	0.0553	0.0763	0.0949	0.1053	0.1075	0.0984	0.0775	0.0481	0.0708		
9	01/06/20	125.15	213.88	475.14	658.00	814.02	934.57	941.11	850.46	552.36	182.24	0.0140	0.0240	0.0533	0.0738	0.0913	0.1049	0.1056	0.0954	0.0620	0.0204	0.0645		
10	01/07/20	113.23	287.03	487.32	676.05	836.29	926.66	917.71	873.54	742.71	431.44	0.0127	0.0322	0.0547	0.0759	0.0938	0.1040	0.1030	0.0980	0.0833	0.0484	0.0706		
11	01/08/20	110.00	284.63	484.64	679.59	838.25	937.11	959.75	894.87	719.09	448.08	0.0123	0.0319	0.0544	0.0762	0.0941	0.1051	0.1077	0.1004	0.0807	0.0503	0.0713		
12	01/09/20	91.30	282.11	484.72	687.63	789.07	946.26	978.57	925.60	748.00	469.70	0.0102	0.0317	0.0544	0.0772	0.0885	0.1062	0.1098	0.1039	0.0839	0.0527	0.0718		
13	01/10/20	121.87	276.53	465.78	581.18	527.85	847.69	661.22	841.52	198.50	78.98	0.0137	0.0310	0.0523	0.0652	0.0592	0.0951	0.0742	0.0944	0.0223	0.0089	0.0516		

Figure 4: Hourly Radiation / Power Data

2.2 Data Processing

This online dataset must be cleansed from noise, null values, duplicated data, typos, missing data, improper formatting, and outliers before it is fed into machine learning. This preprocessing of the dataset before feeding into machine learning is called DATA CLEANING. It is necessary to enhance the accuracy and efficiency of Machine learning. Column 1 was formatted by separating dates from the hours. Column 1 still represents the dates while hours of the dates were transposed to columns 2, 3, 4, . . . 10.

The 3 columns of different hourly radiations (direct, diffused, reflected) were summed up to form one column for radiation. Null values and low insignificant values of radiations were dropped, using radiation data from 9 AM to 4 PM (8 columns). Also, each hourly radiation data was converted to hourly energy/power (in kilowatt) data using the Global Solar Energy formula (Equation 2), [13]. Daily average radiation data and Daily average power were calculated as shown in Figure 4 and the Monthly average data column was also determined. Figure 5 explains these.

```
[ ] X = df.iloc[: , :11]
    X.drop(X.columns[0:3], axis=1, inplace=True)
    X
```

	9:00AM	10:00AM	11:00AM	12:00PM	1:00PM	2:00PM	3:00PM	4:00PM
0	513.58	716.20	884.97	995.96	1024.16	952.55	747.43	452.37
1	513.32	713.07	880.09	983.97	1017.29	945.20	773.53	458.26
2	504.81	703.69	848.24	962.93	1001.57	937.62	763.56	481.28

Figure 5: Downsized coded input radiation dataset

```
[ ] y = df['Daily Av Power(KW)']
    y

0     0.075284
1     0.075002
2     0.074125
```

Figure 6: Part of downsized coding for output power dataset.

In Figure 5, these data formed the input being independent variables (radiation data or x-values) that were fed into machine learning. The output being dependent variables (calculated power data or y-values) were also fed into machine learning shown in Figure 6.

Daily AV power data were obtained by determining each hourly radiation dataset using the Global Solar Energy formula to calculate [13]. From hourly power data in Figure 4 calculated for, Daily Av solar power data were computed, forming a single column data in the Figures. The last column in the same Figure contained the Monthly Av power data.

2.3 Data Splitting

It is a technique employed in evaluating machine learning algorithms [17]. Put simply, the mathematical concept taught with example is the algorithm or model, the students taught formed the training set, then problem or class work formed the testing set, usage of model for more problems is validating set. This is done for better accuracy and unbiased prediction. The dataset was split into the ratio of 80:20, 80% for the training set and 20% for the testing set as shown in Figure 7 Python coding.

```
from sklearn.model_selection import train_test_split
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2)
X_train
```

Figure 7: Split dataset code.

2.4 Creating and Training Model

Artificial Neural Network was the model used in training, which was exhaustively explained in the literature review in Section 1. Figure 8, is the coding for the Splitting, Creating and Training Model.

```
from keras.layers import Dense, Dropout
from keras.models import Sequential
model = Sequential()
model.add(Dense(1000, input_shape=(X_train.shape[1],), activation='relu')) # (features,)
model.add(Dropout(0.5)) # specify a percentage between 0 and 0.5, or larger
model.add(Dense(500, activation='relu'))
model.add(Dropout(0.5)) # specify a percentage between 0 and 0.5, or larger
model.add(Dense(250, activation='relu'))
model.add(Dropout(0.5)) # specify a percentage between 0 and 0.5, or larger
model.add(Dense(1, activation='linear')) # output node
model.compile(optimizer='adam', loss='mse', metrics=['mae','mse'])
model.fit(X_train, y_train , epochs = 250, batch_size=20, verbose=1)
```

Figure 8: Training code for ANN model

In Figure 8 training Model Code was where the forward and backward propagation algorithms were used. From the input layer through activation functions using linear and relu functions as it went through the hidden layer then the output layer and back from the output layer, through the hidden layer to the input layer.

3.0 RESULT AND DISCUSSION

In this section the radiation data collected from the European Commission website for 366 days in a year captured all the solar power fluctuations or flux especially when the solar power intensity was optimal [15]. Low energy/power hours were ignored, not being the focus of the research. This quantitative study made use of descriptive statistical analysis and techniques. When to get the variables were the hours of each day in a year between sunrise and sunset.

3.1 Overview of Data Collection

Figure 4 shows the relevant hourly radiation data and their corresponding hourly power data for each day. From 7 AM to 4 PM every day for 366 days, making it 7686 data. This amount of data was necessary for efficient training and testing of ANN while maintaining the same area ($0.68m^2$). The hourly radiation data formed the x–input independent variable and the y–output dependent variable which is solar energy. The parameters of performance ratio and efficiency, calculated with the area of the panel and radiation produced the output.

3.2 Analysis

Figure 4 was broken into 2 Figures, namely Figure 9a and Figure 9b for clearer analysis. Critically looking at the hourly radiation from Figure 9a between 7 AM and 4 PM daily, described a pattern. The length of bars increased from 7 AM to 1 PM and decreased from 1

https://doi.org/10.4314/njt.v43i1.15

PM to 4 PM daily. The radiation peaked at or maximum around 1 PM most of the days. This visualization explained the hour of optimal radiation, implying that maximum power was generated around the peak hour of the day on a normal seasonal condition in Uyo. A similar pattern was observed for Figure 9b using hourly solar power between the same time range of 7 AM to 4 PM each day. The implication was that any prediction of hourly radiation and solar power in the location must have 1 PM data to form an accurate optimal prediction.

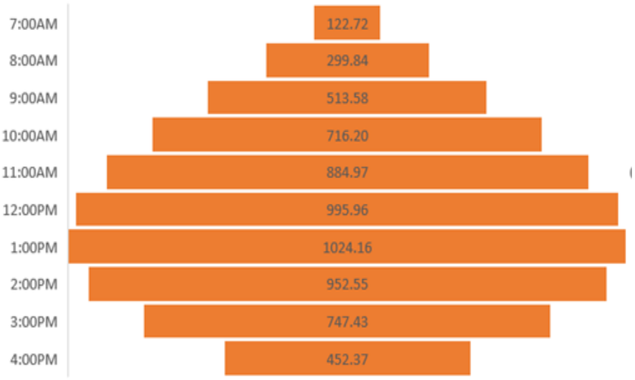


Figure 9a: Pattern of radiation from 7 AM to 4 PM using a PV area of $0.68m^2$

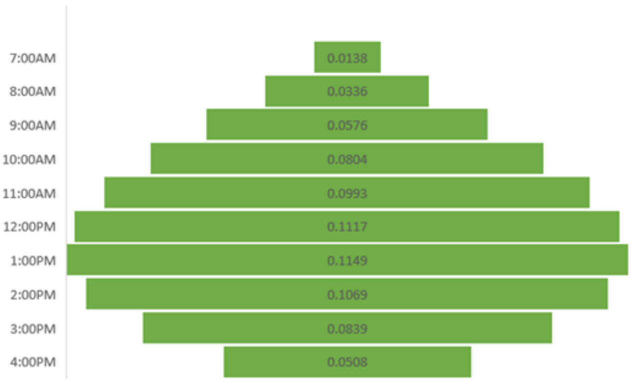


Figure 9b: Pattern of solar energy from 7 AM to 4 PM using a PV area of $0.68m^2$

Figure 10: Monthly Av radiation and energy pattern for the 12 months

The Table 1 shows the calculated monthly average power in kilowatts from January to December (12 months), as well as the monthly average radiation from January to December (12 months). In the description of Figure 10 in the Uyo location, radiation was highest in January, then decreased consecutively till July, a quick crest was formed in August before rising again from September to the highest level in December. The highest point of each month for 12 months formed a curve of a trough in the middle months of the year and a crest at the beginning and ending months of the year, except August. The month of August is known in weather conditions in Nigeria as 'August Break', especially in Southern Nigeria where Uyo is located. Sharp rise and fall in radiation during August Break came from less cloud cover to cause rainfall for about 3 weeks. Implying more radiation reached the panel when the sky was cloudless. The solar power graph (Figure 10) also followed a similar pattern. The monthly radiation and power patterns can be explained as: the beginning and ending months of the year in the location were dry season, the sky was clear of clouds, no blockage of the panel surface from the radiation. So, the panel received more radiation forming the maximum or crest of the curve. On the other hand, in the middle months of the year were months of rainy season in the location, the sky was covered often with clouds, reducing the amount of radiation reaching the panel forming the minimum or trough of the curve. Hence optimal monthly average prediction was in the month of January.

Table 1: Monthly Summary of Radiation & Power

Month	Data4d Power	Data4d Radiation
Jan	0.0706	629.25
Feb	0.065	579.54
Mar	0.0499	444.95
Apr	0.0443	394.95
May	0.0407	362.87
Jun	0.0343	306.01
Jul	0.031	276.15
Aug	0.0379	337.98
Sep	0.0344	306.59
Oct	0.0454	404.93
Nov	0.0508	517.01
Dec	0.0635	565.68

3.3 Summary

The metric results in Table 2 (Figure 11 deductions) from the Python coding of 80% for training and 20% for testing was one of the splitting ratios for machine learning [18]. The Mean Absolute Error (MAE) was 0.03, the Mean Squared Error (MSE) was 0.0 while the Root Mean Squared Error was 0.03 using panel size (area) of $0.68m^2$ and the Mean Absolute Error was 0.04, the Mean Square Error was 0.0 while the Root Mean Square Error was 0.05 using panel area of

$1.0068m^2$. These were very good metrics in prediction. For example, optimal solar power prediction of 500KW with these error margins from this research findings could be of high accuracy and highly reliable. Investors or developers do not need to spend millions for this feasibility study, having the understanding that by consulting a solar energy service, this information will be obtained at a reduced cost.

```
from sklearn import metrics
import numpy as np
from math import ceil
ann_prediction = model.predict(X_test)
mae = metrics.mean_absolute_error(y_test,ann_prediction)
mse = metrics.mean_squared_error(y_test,ann_prediction)
rmse = np.sqrt(metrics.mean_squared_error(y_test,ann_prediction))
print(f'The mean absolute error is {round(mae, 2)}')  #mean absolute error
print(f'The mean squared error is {round(mse,2)}') #mean squared error
print(f'The root mean squared error is {round(rmse,2)}') #root mean squared error

3/3 [==============================] - 0s 9ms/step
The mean absolute error is 0.03
The mean squared error is 0.0
The root mean squared error is 0.03
```

Figure 11: Prediction coding

Table 2: Metric results from 80:20 ratio of training and testing

PV area	MAE	MSE	RMSE
$0.68m^2$	0.03	0.0	0.03
$1.0068m^2$	0.04	0.0	0.05

4.0 CONCLUSION

The hourly analysis revealed that the peak hour for optimal solar power prediction was at 1 PM daily. Furthermore, when considering the monthly averages, January emerged as the peak month for optimal solar power throughout the year. These findings were made possible through the application of the ANN algorithm, which exhibited superior accuracy in predicting solar power in Uyo.

During the research, the relevant variables and parameters were successfully identified and established. The mathematical relationships among these identified factors were demonstrated, providing valuable insights into the prediction process. To implement the ANN, a Python coding model within the Colab software environment was employed. This approach enabled the definition of essential metrics and ultimately led to the highly accurate prediction of solar energy in the specific location.

5.0 RECOMMENDATION

The report utilized data from a one-year timeframe, but there is potential for further studies to extend this period and incorporate data from five or more years to enhance the accuracy of predictions. Additionally, future research could delve into assessing the number of households currently benefiting from solar power systems in the area and the daily consumption of solar power measured in kilowatt-hours (KWH).

Exploring the impact of other weather variables, such as temperature and humidity, on solar energy production could also be a valuable avenue of investigation. Although these variables were not the primary focus of the present research, understanding their influence could offer valuable insights into optimizing solar power generation in the location.

REFERENCES

[1] Alfredo, N., Sonia, L., Marco, M., and Emanuele, G. C. O. "A Selective Ensemble Approach for Accuracy Improvement and Computational Load Reduction in ANN-Based PV Power Forecasting", 2022; IEEE Access, Milan Italy, DOI: 10.1109/ACCESS.2022.3158364

[2] ChérifaKara, M. K., Badia, A., Kamel, K., and Aissa, C. "The impact of the ANN's choice on PV systems diagnosis quality", Volume 240, 15 July 2021, Energy Conversion and Management, https://doi.org/10.1016/j.enconman.2021.114278

[3] Srivignesh, R. "A Walk-through of Regression Analysis Using Artificial Neural Networks in Tensorflow", (2021), Data Science Blogathon. Date accessed 6/1/24

[4] Zahraa, E. M. "Using the artificial neural networks for prediction and validating solar radiation", Journal of the Egyptian Mathematical Society, 2019; Egypt., Pages 1-13. https://doi.org/10.1186/s42787-019-0043-8, Date accessed 6/1/24

[5] Valerio, L. B., Giuseppina, C., and Mariavittoria, D. F. "Artificial Neural Networks to Predict the Power Output of a PV Panel", International Journal of Photoenergy, Volume 2014, Pages 1-12. https://doi.org/10.1155/2014/193083. Date accessed 6/1/24

[6] Agbulut, U., Gurel, A., and Yunus, B. "Predicti-on of daily global solar radiation using different machine learning algorithms: Evaluation and comparison", Semantic Scholar, 2021; https://doi.org/10.1016/j.rser.2020.110114. Accessed: 21/07/23

[7] Umit, A., Ali, E. G., and Yunus B. "Prediction of daily global solar radiation using different machine learning algorithms: Evaluation and comparison", Renewable and Sustainable Energy Reviews, 2020, Elsevier Ltd. Turkey. https://doi.org/10.1016/j.rser.2020.110114

[8] Gurel, A. E., Agbulut, U., and Biçen Y. "Assessment of machine learning, time series, response surface methodology and empirical models in prediction of global solar radiation", Journal of Cleaner Production on Science Direct, 2020; Volume 277, 122353 https://doi.org/10.1016/j.jclepro.2020.122353

[9] Pragati, B. "Activation Functions in Neural Networks [12 Types & Use Cases]", 2021; Microsoft V7 Ltd, London HQ. https://www.v7labs.com/. Date accessed 6/1/24

[10] Uppin, C. "Study of Cloud Computing: Significance & Impact", The International Journal Research Publications, 2018; ISSN: 2251 1563; Research Journal of Science & IT Management (RJSIM) UGC Approved Index Journal. www.theinternationaljournal.org RJS ITM: Volume: 07, Number: 04. Date accessed 9/7/23

[11] European Commission. Joint Research Centre, Energy Efficiency and Renewables Unit, (2023) via E. Fermi 2749, TP 450, I-21027 Ispra (VA), Italy. JRC-PVGIS@ec.europa.eu Date accessed 6/1/24

[12] George, G., and Uppin C. "A Proactive Approach to Network Forensics Intrusion (Denial of Service Flood Attack) Using Dynamic Features, Selection and Convolution Neural Network", Open Journal of Physical Science (OJPS) Nigeria, 2021; ISSN: 2734-2123, Volume: 2; Issue: 2, Pages: 01 - 09 (2021), DOI: 10.52417/ojps. v2i2.237, www.openjournalsnigeria.org.ng. Date accessed 9/7/23

[13] Photovoltaic-Software. "How to calculate the annual solar energy output of a photovoltaic system?", Photovoltaic-software.com. (2023) https://photovoltaic-software.com/principle-ressources/howcalculate-solar-energy-power-pv-systems. Date accessed 6/1/24

[14] Solar Solar. "Solar Power Efficiency", Solar by nature, (2022) Page 1, North Carolina. https://solarbynatureinc.com/category/solar-energy/. Date: accessed 6/1/24

[15] Henderson, M. "Solar Panel Efficiency", Solar Calculator, (2023) Balnarring VIC, Australia. https://solarcalculator.com.au/about-us/ Date accessed 6/1/24

[16] International Geophysics. "An Introduction to Atmospheric Radiation", ScienceDirect.com, (2024), Volume 84; Pages 1-583. Date accessed 6/1/24

[17] Sharma, U. "Why Do We Need Data Splitting?", Faculty (Big Data Analytical), (2021), Technical… , Date accessed 11/4/23 (Utkarsh Sharm… in.linkedin.com)

[18] Jaiswal, S. "Train and Test datasets in Machine Learning", Javatpoint. Noida, (2021), India. https://www.javatpoint.com/train-and-test-datasets-in-machine-learning. Date accessed 6/1/24