Original Research Article

Solar Radiation Forecasting Models and Its Thermodynamics Analysis in Asaba: Least Square Regression and Machine Learning Approach

Commented [DI1]: Their Thermodynamic

Abstract

Least square regression and machine learning tools were used for the development of global solar radiation forecasting models for Asaba region. 2013-2022 data from Nigerian Meteorological Agency, Asaba was used for this study. The least square regression method was used to develop four global solar radiation -based models, tagged H1, H2, H3 and H4 with characteristic day length, solar declination angle, rainfall amount, etc as its model terms while the machine learning models produced multilayer perceptron, coarse Gaussian model (SVM-based model) and XGBoost model. The prediction factors like mean bias error, mean percentage error, root mean square error, Nash-Sutcliffe equation, coefficient of correlation (R), t-test, and R^2 were considered using the model terms. The results indicates that H4 model outperformed H1, H2, H3, machine learning models (SVM-based model, multilayer perceptron and XGBoost) and other existing models (MA-MME and MLR) with a mean percentage error value of 0.740, RMSE value of 46.588, Nash-Sutcliffe equation value of 0.739, higher R^2 value of 0.7391, t-test value of 2.595E-24 and mean bias error value of -6.88E-12. Thus, H4 model results fell within accepted range. Additionally, the exergy of the global solar radiation of Asaba varied from 20-185 W/m^2 which is good. And this shows that a more efficient and ideal global solar radiation prediction model (H4) has been developed for Asaba and other regions that share similar climatic conditions.

Commented [DI2]: Data from the year 2013-2022

Commented [DI3]: H1

Commented [DI4]: etc.

Commented [DI5]: coefficient of determination (R^2)

Commented [DI6]: was

Commented [DI7]: This

Keywords*:* Global Solar Radiation, Asaba, Least Square Regression, Exergy, Machine Learning Tools.

1.0 Introduction

Energy drives industrialization and it remains the bedrock for national growth and development [[1-2]. Non-renewable energy and renewable energy are the two subset of energy [[1], [3]] Non-renewable form of energy e.g Fossil fuel combustion depletes the ozone layer, thereby, emitting harmful ultraviolet rays to the entire environment via the release of greenhouse gases leading to global warming. And this negative impact on earth

Commented [DI8]: This

has redirected attention to green energy referred to as renewable energy; a form of energy that is cheap, readily available, replenishes itself, doesn't cause economic and political imbalance unlike fossil fuel [1]. Thus, the reason why nations, engineers and scientists consider it as a source that can drive the economy of any country towards industrialization [4]. The various forms of renewable energy are: wind, solar, water, geothermal, biomass etc [3-5]. Among all the various forms of renewable energy, solar energy is more desirable due to its availability and radiant form [[3], [5], [6]].

Commented [DI9]: etc.

Solar energy is gotten from the sun. Duffie and Beckham [7] explained the sun to be a sphere of intensely hot gaseous matter with a diameter of $1.39 \times 10^9 m$ and is, on the average, $1.5 \times 10^{11} m$ from the earth. A view from the earth shows that the sun rotates on its axis about 27 days and the Polar Regions take about 30 days for each rotation [7]. The sun has an effective blackbody temperature of 5777 K. The temperature in the central interior regions is estimated to range from $8 \times 10^6 K$ to $40 \times 10^6 K$ and the density is estimated to be about 100 times that of water [8]. The energy produced in the interior of the solar sphere at temperatures of many millions of degrees must be transferred out to the surface and then be radiated into space as depicted in Figure 1.

Commented [DI10]: obtained

Several technologies such as: solar pumps, solar heaters, photovoltaic systems, solar cars, solar powered heating, ventilation and air-conditioning systems (HVACs) were all designed and built to be driven by solar energy. Solar energy has thus left a huge mark in the world of science and technology to be very promising and an efficient source of energy. In addition, for efficient utilization of these solar driven technologies, sufficient data/information on the amount of solar radiation or irradiation incident on the earth surface for a particular region of interest must be considered and factored into the design of the solar system in order to actualize its optimal performance characteristics. The total amount of solar radiation incident on the earth surface is called global solar radiation (GSR) [9]. Global solar radiation data are very imperative at various steps of the design, engineering, simulation and performance evaluation of any solar energy project [9]. The knowledge of the solar irradiation on the earth's surface is an essential requirement not only in the studies of climate change, environmental pollution, but also in agriculture, hydrology, food industry and non-conventional energy development programs [10]. It should be noted that not all the irradiation actually reaches the earth's surface. The amount of irradiation that would reach the earth's surface depends solely on: the concentration of airborne particulate matter, gaseous pollutants and water (vapour, liquid, or solid) in the sky. These factors contribute to the attenuation of solar radiation and also affect the diffuse and direct radiation ratio [9]. There are two components of global solar radiation, namely: diffuse and direct solar radiation. The former results from scattering caused by gases in the earth's atmosphere, dispersed water droplets and particulates, while the later refers to radiations which were not scattered. Global solar radiation is the algebraic sum of the two components. The amount of irradiation of a particular geographical location in the visible and near-infrared wavelength

Commented [DI11]: as

ranges from $0.3\mu m\ to\ 2\mu m$ and it is measured with a pyranometer. Global solar radiation varies from latitude to latitude [9].

Due to seasonal variations and climatic changes, it has become necessary to frequently measure the amount of solar radiation incident on a particular geographical setting and other climate parameters that are of importance in the design and manufacturing of solar powered appliances and devices. Sequel to the fact that solar radiation measurements are not readily available in many developing countries due to non-availability or inadequate provision of measurement instruments [9]. There is need to research towards the development of mathematical models that could be employed in estimating the amount of solar radiation for a specific geographical site like Asaba, Nigeria, a State Capital that houses an airport and other reputable organizations, with a view to reduce the frequency at which solar irradiation and other climate parameters are being measured with their relevant tools. Njoku et al. [11] performed a study on the analysis of solar radiation measurement in south-west geopolitical zone of Nigeria. They used twenty (20) years data (from 2000-2021) of daily global solar radiation and monthly average hours of bright sunshine sourced from National Aeronautics and Space Administration (NASA) and Nigeria Meteorological Agency (NIMET) in carrying out their studies. From their studies, aided by global solar radiation, clearness index and sunshine hour, sky condition of South West Geopolitical zone in Nigeria was characterized. The worst month for harvest of solar radiation in the zone was August as it recorded the lowest level of clearness index and hours of bright sunshine. The city of Ikeja was considered the worst location for harvest of solar radiation in the geopolitical zone. Coefficients for the Angstrom-Page equation was estimated for the cities and the associated coefficient of determination exhibited low variation for each study location.

Additionally, some researchers have also employed artificial intelligence tools like artificial neutral networks, least square regression model, and machine learning models in predicting the amount of solar radiation of a particular geographical location [10]. Machine learning (ML) model according to Junliang et al. [12] have an intrinsic excellent capability of dealing with non-linear functions approximation, data sorting, pattern recognition, optimization, clustering, and trend detection. Also, artificial neural network (ANN) is the model widely used for predicting operations and it's the machine learning utilized here. The ANN imitates the mechanism of the brain. Ertekin and Evrendiliek [13] estimated daily global solar radiation in Al-Madinah of Saudi Arabia using multi-layer perceptron (MLP) model from sunshine duration, air temperature and relative humidity data. They discovered that the MLP model performed better than the linear model but not the quadratic model. Behrang et al. [14] studied the performance of two ANN models- MLP and radial basis function (RBF) models in predicting daily values of global solar radiation based on sunshine duration and other meteorological data for the Dezful city in Iran. The MLP model performed better than the other models. Also,

Huang et al. [10] compared the prediction performances of random forest (RF) model, multiple linear regression (MLR) and simple persistence methods in modeling daily global solar radiation using the meteorological data of four stations across Australia. They discovered that the RF model had better estimation performance than others. In addition, Huang et al. [10] worked on "solar radiation prediction using different machine learning algorithms and implications for extreme climate events and performed data preprocessing and factors selection based on meteorological factors and solar radiation data from 1980 to 2016 for Ganzhou, China by developing twelve machine learning models to estimate and contrast each day and monthly values of solar radiation and a stacking model using Sklearn and the Xgb library. And concluded that the stacking model and the XGBoost model gave the ideal model in forecasting solar radiation for Ganzhou, China". But as much as we are aware, no one has worked on global solar radiation in Asaba and its region, let alone utilizing different least square regression models (H1, H2, H3 and H4) and machine learning models (SVM-based model, multilayer perceptron and XGBoost) to forecast weather conditions and its thermodynamics analysis for the benefit of the society at large and to meet UN`s millennium development target of clean energy for everyone [2]. The input factors of 10 years collected from NIMET are; monthly average daily hours of bright sunshine, monthly mean minimum and maximum temperatures, monthly mean values of relative humidity, monthly mean values of rainfall and monthly mean global solar radiation while the monthly mean daily sunshine duration known as characteristic day length, sunset hour angle, solar declination angle and other parameters of interest were calculated using the appropriate equations.

Thus, the need to research on solar radiation forecasting models and its thermodynamic analysis in Asaba using least square regression and machine learning approach by collecting ten (10) years (2013-2022) meteorological data of Asaba region from Nigeria Meteorological Agency (NIMET). Then, compare and validate with exist findings globally.

Commented [DI12]: They concluded

Commented [DI13]: Therefore, this research was conducted

Commented [DI14]: their

Commented [DI15]: Then these findings would be compared and validated globally.

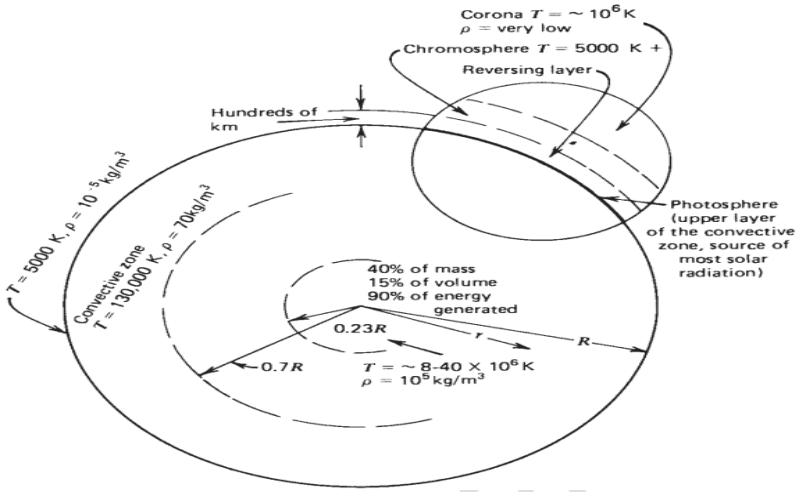

Figure 1: The structure of the sun [7]

2.0 Materials and Methods

2.1. *Material collection*

The materials used in this study includes: the climatic data of Asaba town, Delta State, latitude 6.2059°N and longitude 6.6959°E, for a period of ten years (2013-2022), measuring instruments and software packages. The following climatic data of the study site were collected from the Nigeria Meteorological Agency (NIMET), Asaba- monthly average daily hours of bright sunshine, monthly mean minimum and maximum temperatures, monthly mean values of relative humidity, month amount of rainfall and monthly mean global solar radiation.

2.1.1 *A brief of the Climatic Data from NIMET Asaba region Utilized in this Study with Input Factors*

Table 1- 6 highlights the input factors considered and its data collected from NIMET, Asaba region for10years.

Table 1: Monthly mean values of maximum temperature (℃)

	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022
JAN	34.6	35.4	36.6	36.6	36.4	35.3	35.6	35.7	35.4	34.8
FEB	35.2	36.1	36.9	38.4	37.9	35.8	36	36.2	35.8	35.4
MAR	33.6	34.3	35.1	35.4	37.3	34.9	35.4	35.3	34.3	33.9
APR	33.8	33.4	34.2	35.5	34.9	33.7	34.9	34.6	33.8	34.1
MAY	32.4	32.5	33.8	34.4	34.4	32.6	34	33.2	33.1	32.8
JUNE	31.8	31.8	32.8	32.3	33.3	31.9	31.6	34.8	32.9	32.4
JULY	32.2	32.4	31.5	31.8	31.3	30.3	30	30.3	32.1	31.4
AUG	32.5	33.2	32.8	30.9	29.8	30.8	30.2	30.7	32.8	32.1
SEPT	30.3	31.8	31.8	32	30.3	31	31.3	29.9	33.1	31.2
OCT	31.1	32.8	32.8	33.5	32.4	30.2	31.1	31.4	32.8	31.9
NOV	32.6	34.1	33.8	35.7	34.4	33.1	33.1	34.4	34.8	32.8
DEC	33.8	34.7	34.8	35.5	34.6	34.8	35.2	35.2	34.1	33.6

Table 2: Monthly mean values of minimum temperature (℃)

	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022
JAN	22.3	21.3	22.5	22.4	22.8	22.1	22.1	21.7	22.1	21.8
FEB	23.3	22.1	22.1	23.4	22.4	21.6	23.4	22.3	21.6	21.8
MAR	23.5	24.1	23.8	23.8	23.3	23.4	24.3	22.8	22.8	22.8
APR	24.4	24.3	23.6	24.1	24.1	24.2	24.1	23.5	23.6	23.1
MAY	23.8	23.3	24.2	24.8	24.4	24.8	23.8	24.2	24.4	23.4
JUNE	24.2	23.8	24.6	24.1	24.8	25.2	24.4	24.8	24.2	23.8
JULY	25.2	24.3	24.8	23.5	24	24.8	24.7	24.4	24.1	24.6
AUG	25.4	25.4	24.6	23.1	23.6	24.4	25.1	24.3	24.3	24.6
SEPT	24.8	24.2	23.4	24.4	23.6	24.6	24.6	24.1	24.6	24.5
OCT	24.3	24.2	24.2	23.4	24.4	24.4	24.3	24.3	25.1	24.1
NOV	24.2	24.1	24.6	24.1	24.1	23.8	23.8	24.8	23.5	23.1
DEC	23.9	23.5	24.9	23.7	24.4	23.2	22.5	24.1	23.4	22.2

Table 3: Monthly mean values of relative humidity (%)

	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022
JAN	48	42	64	48	47	56	58	58	48	46
FEB	54	56	68	35	37	62	43	54	52	48
MAR	58	62	70	52	35	68	49	65	64	62
APR	62	74	75	55	52	74	58	72	74	74
MAY	74	78	82	69	62	76	65	78	78	76
JUNE	86	82	88	76	78	79	72	76	82	80
JULY	84	86	86	85	84	85	78	84	80	88
AUG	78	84	87	87	89	86	84	87	88	86
SEPT	87	88	85	83	83	81	86	89	86	88
OCT	82	86	84	76	76	79	84	78	84	82
NOV	78	76	78	65	64	72	78	72	78	75
DEC	71	72	72	58	62	64	73	68	70	65

Table 4: Monthly mean values of bright sunshine (hours)

	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022
JAN	6.42	7.24	8.23	8.23	7.12	8.25	9.23	8.23	7.54	8.34
FEB	7.34	6.58	8.12	7.51	6.23	7.48	8.17	9.1	7.32	10.24
MAR	6.4	5.43	7.45	6.54	5.23	6.48	8.42	7.23	6.12	7.23
APR	5.1	5.52	7.23	6.23	5.12	6.14	7.21	6.43	6.34	6.45
MAY	5.43	6.45	6.45	5.54	4.48	5.51	5.23	5.23	5.23	6.32
JUNE	6.32	7.11	6.12	5.1	4.23	4.52	4.45	5.43	6.23	5.45
JULY	5.43	6.23	5.5	4.23	4.48	4.32	5.23	4.56	4.56	4.46
AUG	4.32	4.13	4.48	4.31	5.14	4.35	5.12	6.23	4.45	5.4
SEPT	5.12	4.54	4.54	4.58	5.45	4.52	4.23	5.21	5.13	6.23
OCT	4.53	5.12	5.58	4.46	6.13	4.51	4.43	4.45	5.42	5.12
NOV	6.24	6.12	6.48	5.42	6.42	5.34	6.32	6.21	6.12	4.38
DEC	5.43	5.41	7.35	6.54	5.13	5.23	6.45	5.41	6.34	5.31

Table 5: Monthly mean values of global solar radiation (w/m^2)

	2013	**2014**	**2015**	**2016**	**2017**	**2018**	**2019**	**2020**	**2021**	**2022**
JAN	620	680	710	720	740	720	680	680	700	640
FEB	680	640	640	780	680	740	720	640	710	580
MAR	520	570	560	730	620	680	640	580	680	540
APR	540	540	520	680	580	640	600	560	640	480
MAY	480	460	470	640	600	580	540	480	580	460
JUNE	510	470	380	580	580	510	490	480	460	420
JULY	380	510	420	540	540	480	520	420	440	470
AUG	420	420	450	510	480	460	480	460	480	460
SEPT	460	480	480	480	460	480	440	470	440	520
OCT	470	450	540	460	520	480	460	430	460	510
NOV	570	480	580	540	560	580	530	520	480	540
DEC	620	540	620	620	540	610	580	600	460	580

Table 6: Monthly mean values of rainfall amount (mm)

	2013	**2014**	**2015**	**2016**	**2017**	**2018**	**2019**	**2020**	**2021**	**2022**
JAN	25.9	41.9	0	0	3.9	0	0	0	0	0
FEB	5.1	25	35.7	18.2	0	233.2	32	0	0	0
MAR	68	115.3	174.9	184.6	78.1	17.4	6.9	151.8	25.9	23.4
APR	204	154.6	47.8	91.6	268.9	219.5	32.5	136.5	232.5	371.5
MAY	293	299.3	270.8	186.2	354.7	211.6	389.8	222.6	278	433.6
JUNE	311.6	381.2	277.8	622.5	373.5	141.3	334.9	291.6	579.5	259.5
JULY	255.7	334.9	371.2	482.4	681.5	846.6	616.4	277.7	304.7	489.4
AUG	171.6	299.2	385.2	325.8	499.4	343.1	782.1	123.9	530.3	190.9
SEPT	235.9	214.2	727	353.1	263.3	745.3	811.1	443	324.4	381.5
OCT	224.3	218.3	263.3	164.4	404.3	133.3	311.7	348.2	349.6	405.3
NOV	19.9	70.5	53.3	15	14.2	69.3	102.2	10.2	276.5	0
DEC	0	27.8	34	0	0	0	0	83.9	64.8	0

2.1.2 Instruments Used in Measuring the Climatic Parameters

The instrument are Campbell-Stokes Recorder, Solarimeter, Minimum Thermometer, Maximum Thermometer, Todler Radar, Hygrometer, Rain gauge, MATLAB Software 2017, Microsoft Excel 2016 and Minitab 2021

Commented [DI16]: Used in this study were

2.2 Research Method

The multivariate formulation models developed by Cohen et al. [15] which has greatly proven to be efficient in the estimation of global solar radiation were used to estimate the global solar radiation values of the Asaba region, Delta State. The least square method was employed to develop multiple linear regression (MLR) model and multivariate models for the estimation of global solar radiation of Asaba town, Delta State. In addition, machine learning models such as multilayer perceptron network (MLP), support vector machine (SVM) and extreme gradient boosting (XGBoost) which are the classified types of artificial neural network (ANN), Kernel-based algorithms (KBA) and tree-based assembly (TBA) respectively, were also used to estimate the values of global solar radiation of the study site. The estimation performances of the machine learning models, empirical multivariate models (Cohen's models), multiple linear regression model and the newly developed multivariate regression models were statistically compared using the indices of mean bias error (MBE), root mean square error (RMSE), mean percentage error (MPE), coefficient of correlation (R), t-statistic test, Nash-Sutcliffe equation (NSE) and coefficient of determinant (R^2) in order to ascertain the correctness and significance of the new models and as well depict the best model. Graphical visualizations were also used in the comparative study of these models.

2.2.1 Determination of Model Development Parameters

Solar declination angle (δ)

Cooper [16], gave an approximate equation for determining the declination angle (δ) to be:

$$\delta = 23.45 \sin\left(360 \frac{284+n}{365}\right) \quad (1)$$

Sunset hour angle (ω_s)

According to Duffie and Beckman [7], the sunset hour angle is computed using the relation:

$$\omega_s = \cos^{-1}(-\tan\emptyset\tan\delta) \quad (2)$$

Monthly mean daily sunshine duration (N)

This was computed using Duffie and Beckham [7] model,

$$N = \frac{2}{15}\cos^{-1}(-\tan\emptyset\tan\delta) = \frac{2}{15}\omega_s \quad (3)$$

Monthly mean daily extraterrestrial global solar radiation (Z_O)

This was determined using the relation expressed by Duffie and Beckham [7],

$$Z_O = \frac{24\times3600G_{sc}}{10^6\pi}\left(1 + 0.033\cos\frac{360n}{365}\right)\left(\cos\emptyset\cos\delta\sin\omega_s + \frac{\pi\omega_s}{180}\sin\emptyset\sin\delta\right) \quad (4)$$

Where: G_{sc} = The solar constant = 1367 w/m^2, $\emptyset$ = the latitude of the area considered (degrees)

δ = Solar declination angle (degrees), ω_s = Sunset hour angle (degrees)

The computed monthly mean values of solar declination angle, sunset hour angle and characteristic day length using equations 1, 2, and 3 respectively are delineated in tables 7 to 9.

Table 7: Monthly mean values of solar declination angle (degrees)

	2013	**2014**	**2015**	**2016**	**2017**	**2018**	**2019**	**2020**	**2021**	**2022**
JAN	-20.90	-21.27	-21.44	-21.75	-21.60	-20.34	-21.10	-20.73	-20.14	-20.73
FEB	-13.00	-13.29	-13.62	-14.27	-13.95	-11.58	-12.95	-12.27	-11.23	-12.27
MAR	-2.40	-2.82	-3.22	-4.02	-3.62	-0.81	-2.42	-1.61	-0.40	-1.61
APR	9.40	9.41	9.04	8.29	8.67	11.23	9.78	10.51	11.58	10.51
MAY	18.80	18.79	18.55	18.04	18.30	19.93	19.03	19.49	20.14	19.49
JUNE	23.10	23.31	23.27	23.15	23.21	23.44	23.35	23.41	23.45	23.41
JULY	21.20	21.52	21.67	21.97	21.83	20.64	21.35	21.01	20.44	21.01
AUG	13.50	13.78	14.11	14.74	14.43	12.10	13.45	12.79	11.75	12.79
SEPT	2.20	2.22	2.62	3.42	3.02	0.20	1.81	1.01	-0.20	1.01
OCT	-9.60	-9.60	-9.23	-8.48	-8.86	-11.40	-9.97	-10.69	-11.75	-10.69
NOV	-18.90	-19.15	-18.91	-18.42	-18.67	-20.24	-19.38	-19.82	-20.44	-19.82
DEC	-23.00	-23.34	-23.29	-23.18	-23.24	-23.45	-23.37	-23.42	-23.45	-23.42

Table 8: Monthly mean values of sunset hour angle (degrees)

	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022
JAN	87.63	87.58	87.56	87.52	87.54	87.70	87.60	87.65	87.72	87.65
FEB	88.57	88.53	88.50	88.42	88.46	88.73	88.57	88.65	88.77	88.65
MAR	89.74	89.69	89.65	89.56	89.61	89.91	89.74	89.83	89.96	89.83
APR	91.03	91.03	90.99	90.90	90.95	91.23	91.07	91.15	91.27	91.15
MAY	92.11	92.11	92.08	92.02	92.05	92.25	92.14	92.20	92.28	92.20
JUNE	92.65	92.68	92.67	92.66	92.66	92.69	92.68	92.69	92.69	92.69
JULY	92.41	92.45	92.47	92.51	92.49	92.34	92.43	92.39	92.31	92.39
AUG	91.49	91.52	91.56	91.63	91.60	91.33	91.49	91.41	91.29	91.41
SEPT	90.24	90.24	90.28	90.37	90.33	90.02	90.20	90.11	89.98	90.11
OCT	88.95	88.95	88.99	89.07	89.03	88.75	88.91	88.83	88.71	88.83
NOV	87.87	87.84	87.87	87.93	87.90	87.71	87.82	87.76	87.69	87.76
DEC	87.36	87.32	87.33	87.34	87.33	87.31	87.32	87.31	87.31	87.31

Table 9: Monthly mean values of characteristic day length (hours)

	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022
JAN	11.68	11.68	11.67	11.67	11.67	11.69	11.68	11.69	11.70	11.69
FEB	11.81	11.80	11.80	11.79	11.79	11.83	11.81	11.82	11.84	11.82
MAR	11.97	11.96	11.95	11.94	11.95	11.99	11.97	11.98	11.99	11.98
APR	12.14	12.14	12.13	12.12	12.13	12.16	12.14	12.15	12.17	12.15
MAY	12.28	12.28	12.28	12.27	12.27	12.30	12.29	12.29	12.30	12.29
JUNE	12.35	12.36	12.36	12.35	12.36	12.36	12.36	12.36	12.36	12.36
JULY	12.32	12.33	12.33	12.33	12.33	12.31	12.32	12.32	12.31	12.32
AUG	12.20	12.20	12.21	12.22	12.21	12.18	12.20	12.19	12.17	12.19
SEPT	12.03	12.03	12.04	12.05	12.04	12.00	12.03	12.01	12.00	12.01
OCT	11.86	11.86	11.87	11.88	11.87	11.83	11.85	11.84	11.83	11.84
NOV	11.72	11.71	11.72	11.72	11.72	11.69	11.71	11.70	11.69	11.70
DEC	11.65	11.64	11.64	11.65	11.64	11.64	11.64	11.64	11.64	11.64

Table 10 shows the day number in the year used in computing the solar declination angle and other parameters.

Table 10: Day number in the years

	2013	**2014**	**2015**	**2016**	**2017**	**2018**	**2019**	**2020**	**2021**	**2022**
JAN	17	15	14	12	13	20	16	18	21	18
FEB	47	46	45	43	44	51	47	49	52	49
MAR	75	74	73	71	72	79	75	77	80	77
APR	105	105	104	102	103	110	106	108	111	108
MAY	135	135	134	132	133	140	136	138	141	138
JUNE	162	166	165	163	164	171	167	169	172	169
JULY	198	196	195	193	194	201	197	199	202	199
AUG	228	227	226	224	225	232	228	230	233	230
SEPT	258	258	257	255	256	263	259	261	264	261
OCT	288	288	287	285	286	293	289	291	294	291
NOV	318	319	318	316	317	324	320	322	325	322
DEC	344	349	348	346	347	354	350	352	355	352

2.2.3 The Proposed New Models

The proposed new models for estimating global solar radiation were developed using least square multivariate regression method and it employed rainfall amount and solar declination angle, in addition to some of the other variables used in the multivariate models of Cohen et al. [15] as its independent or predictive variables. The proposed new models are presented in equations 5-8.

$$H = \cos\emptyset + \cos n + \cos^2 n + \cos\delta + \cos^2\delta + T_{max} + \frac{\bar{n}}{N} + \frac{T_{max}}{R.H} + \left(\frac{T_{max}}{R.H}\right)^2 + \left(\frac{T_{max}}{R.H}\right)^3 + \left(\frac{T_{max}}{R.H}\right)^4 + \left(\frac{\bar{n}}{N}\right)^2 + \left(\frac{\bar{n}}{N}\right)^3 + \frac{R.F}{T_{max}} \quad (5)$$

$$H = \cos\emptyset + \cos n + \cos^2 n + T_{max} + \frac{\bar{n}}{N} + \frac{T_{max}}{R.H} + \left(\frac{T_{max}}{R.H}\right)^2 + \left(\frac{T_{max}}{R.H}\right)^3 + \left(\frac{T_{max}}{R.H}\right)^4 + \left(\frac{\bar{n}}{N}\right)^2 + \left(\frac{\bar{n}}{N}\right)^3 + \frac{R.F}{T_{max}} \quad (6)$$

$$H = \cos\emptyset + \cos n + \cos^2 n + \cos\delta + \cos^2\delta + T_{max} + \frac{\bar{n}}{N} + \frac{T_{max}}{R.H} + \left(\frac{T_{max}}{R.H}\right)^2 + \left(\frac{T_{max}}{R.H}\right)^3 + \left(\frac{T_{max}}{R.H}\right)^4 + \left(\frac{\bar{n}}{N}\right)^2 + \left(\frac{\bar{n}}{N}\right)^3 + \frac{R.F}{T_{max}} + \left(\frac{R.F}{T_{max}}\right)^2 + \left(\frac{R.F}{T_{max}}\right)^4 \quad (7)$$

$$H = \cos\emptyset + \cos n + \cos^2 n + \cos\delta + \cos^2\delta + T_{max} + \frac{\bar{n}}{N} + \frac{T_{max}}{R.H} + \left(\frac{T_{max}}{R.H}\right)^2 + \left(\frac{T_{max}}{R.H}\right)^3 + \left(\frac{T_{max}}{R.H}\right)^4 + \left(\frac{\bar{n}}{N}\right)^2 + \left(\frac{\bar{n}}{N}\right)^3 + \frac{R.F}{T_{max}} + \left(\frac{R.F}{T_{max}}\right)^2 + \left(\frac{R.F}{T_{max}}\right)^4 + \frac{R.F}{R.H} + \left(\frac{R.F}{R.H}\right)^2 + \left(\frac{R.F}{R.H}\right)^4 \quad (8)$$

Where: H = global solar radiation, δ = solar declination angle, T_{max} = maximum temperature, R.F = rainfall amount, R.H. = relative humidity, $\emptyset$ = latitude of the place, n = day number the year, N = characteristic day length, $\bar{n}$ = hours of bright sunshine, $\frac{\bar{n}}{N}$ = sunshine index.

2.2.3.1 Equations utilized for output analysis

The Mean Bias Error (MBE)

The relation used for its computation is given by equation 9.

$$MBE = \frac{1}{n}\sum_{i=1}^{n}(Z_{ical} - Z_{imea}) \quad (9)$$

Where: n = number of data points, Z_{ical} = calculated value of global solar radiation,

Z_{imea} = measured value of global solar radiation

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 10.

$$\text{MPE}(\%) = \frac{1}{n}\sum_{i=1}^{n}\left(\frac{Z_{ical} - Z_{imeas}}{Z_{imeas}}\right) \times 100 \quad (10)$$

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 11

$$RMSE = \left[\frac{1}{n}\sum_{i=1}^{n}(z_{ical} - z_{imeas})^2\right]^{\frac{1}{2}} \quad (11)$$

The Nash-Sutcliffe Equation (NSE)

Model's efficiency and correctness is assured if and only if NSE is very close to one. The formula used to compute it as given by Chen et al. [17] is given by equation 12

$$NSE = 1 - \frac{\sum_{i=1}^{n}(Z_{imea} - Z_{ical})^2}{\sum_{i=1}^{n}(Zi_{mea} - Z_{mea})^2} \quad (12)$$

Where: $\bar{Z}_{mea}$ = the mean measured value of global solar radiation

The Coefficient of Correlation (R)

The coefficient of correlation measures the degree of association or relationship between the measured value of global solar radiation 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. The Karl Pearson's method was used to assess the developed model's association with the measured values of global solar radiation. It was computed using the equation 13.

$$R = \frac{\sum XY}{\sqrt{(\sum X^2)(\sum Y^2)}} \quad (13)$$

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

T-statistics Test

The formula used to compute it is given by equation 14

$$t = \left[\frac{(n-1)(MBE)^2}{(RMSE)^2-(MBE)^2}\right]^{\frac{1}{2}} \quad (14)$$

Coefficient of Determination (R^2)

A value of R^2 close to 1 suggests that the model is best suited for global solar radiation estimation. It is mathematically given as equation 15,

$$R^2 = \frac{\text{sum of squares (ss)}}{\text{sum of squares of residuals}} = \frac{\sum_{i=1}^{n}(Z_{imea}-\bar{Z}_{imea})^2}{\sum_{i=1}^{n}(Z_{imea}-\bar{Z}_{ical})^2} \quad (15)$$

2.2.4 Thermodynamic Analysis: Energy and Exergy Analysis of Global Solar Radiation of Asaba Town

The exergy analysis of Asaba town, Delta state was based on the second law of thermodynamics. It can be used to determine the theoretical limits for the performance of most commonly used engineering systems, such as: heat engine, refrigerators, solar panels, as well as predicting the degree of completion of chemical reactions. Therefore, to estimate the exergy solar radiation on the earth, consider a machine placed on the earth surface as shown in Figure 2.

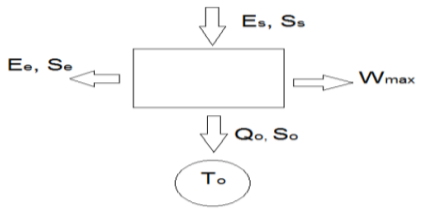

Figure 2: Energy and exergy interactions in a cyclic machine

The machine extracts the maximum work W_{max} obtainable from the energy sourceE_s, and it delivers heat Q_o to the environment at temperatureT_o. The machine transmits the energy E_e by radiation.

According to the energy and entropy balance of the cyclic machine, the exergy solar radiation expressed by Nayak and Tiwari [18] is given as,

$$E_{xs} = H\left[1 - \frac{4}{3}\left(\frac{T_1}{T_2}\right) + \frac{1}{3}\left(\frac{T_1}{T_2}\right)^4\right] \quad (16)$$

Where: E_{xs} = the exergy of the solar energy, H = global solar radiation (w/m^2), T_1 = ambient temperature of the region, which is equivalent to the minimum temperature (K), T_2 = sun's temperature of the region, equivalently equal to the maximum temperature (K), T_o = sink temperature.

Thus, the exergy global solar radiation for Asaba region was computed using equation 16 and was equally shown graphically to clearly delineate the lost energies of solar irradiance in the study site.

3. **Results and Discussions**

3.1 Summary of the Results of the Statistical Evaluation of the Models

Table 11 shows the result of the statistical evaluation of the models- MMA, MMB, MMC, MMD, MME, MLR, H1, H2, H3, H4, MLP, SVM and XGBoost.

It is worthy of note that the best performed model was obtained via a statistical evaluations considering the following statistical tools: mean bias error (MBE), mean percentage error (MPE), root mean square error (RMSE), Nash-Sutcliffe equation (NSE), coefficient of correlation (R), t-statistics test, and coefficient of Determination (R^2).

Table 11: Statistical evaluation of the models

	PERFORMANCE INDICES						
Models	**MBE**	**MPE**	**RMSE**	**NSE**	**R**	**t-stat.**	**R^2(%)**
MMA	-3.78E-12	-0.833	50.520	0.693	0.833	8.162E-13	0.6932
MMB	-3.75E-12	-0.833	50.625	0.692	0.832	6.529E-25	0.6919
MMC	-6.68E-12	-0.833	50.314	0.696	0.834	2.096E-24	0.6957
MMD	-1.03E-11	-0.833	49.325	0.708	0.841	5.189E-24	0.7075
MME	-3.43E-12	0.847	50.248	0.696	0.835	5.545E-25	0.6965
MLR	1.14E-14	0.958	53.480	0.656	0.810	5.407E-30	0.6562
H1	-9.96E-12	0.777	47.442	0.729	0.854	5.245E-24	0.7294
H2	-1.09E-11	0.780	47.555	0.728	0.853	6.252E-24	0.7281
H3	-8.99E-12	0.750	46.846	0.736	0.858	4.382E-24	0.7362
H4	-6.88E-12	0.740	46.588	0.739	0.860	2.595E-24	0.7391
MLP	0.5315	0.604	38.183	0.825	0.908	0.0231	0.9999
SVM	-6.3685	-0.368	47.583	0.728	0.856	2.1705	0.6900
XGBoost	-5.9040	-0.477	39.372	0.814	0.906	2.7374	0.6200

From Table 11, the new models (H1-H4) especially H4 model performed better than the multivariate (MMA-MME) and the MLR models because its: MPE value fall within -10 to 10%, lower RMSE value, higher NSE value (that is much closer to 1), higher value of r value, higher R^2 value and lower t-stat. and MBE value. The H4 model outperformed the models developed by Cohen et al. [15] and the MLR model. In addition, the H4 model also outperformed the machine learning model- SVM, and it had a closer performance characteristic of the MLP and XGBoost models.

3.2 *Result of the Exergy Analysis of Asaba Town*

The result measured GSR and exergy GSR against the 120 months of the collected dataset respectively.

It was observed that the lost GSR energy could be observed to vary between 20-185W/m^2. This therefore suggests that the GSR of Asaba town is not fully exploited as the lost GSR energy could be utilized for other meaningful purposes. Thus, this demonstrates the effect of the atmospheric bodies which absorb some of the energy before it gets to the earth surface. This lost GSR energy need to be harnessed so as to improve the

operational characteristics of solar energy consuming devices such as the heating, ventilation and air-conditioning system (HVAC), solar cars, solar pumps, solar heaters, etc. See appendix A for exergy GSR values.

Commented [DI17]: Appendix A can be referred for exergy GSR values.

4.0 Conclusion

The study presents the thermodynamic analysis and GSR prediction model development for Asaba town, Delta state. The multivariate models of Cohen et al. [15] and the MLR model were outperformed by the four new models developed in this study based on the statistical indices (mean bias error (MBE), mean percentage error (MPE), root mean square error (RMSE), Nash-Sutcliffe equation (NSE), coefficient of correlation (R), t-statistics test, and coefficient of Determination R^2) that were used. The four new models (H1-H4) were modifications of the multivariate models (MMA, MMB, MMC, MMD, and MME) that were achieved by introducing some other parameters like: solar declination angle and rainfall amount, in addition to other polynomic model terms of various degrees that were built using least square regression methodology. The H4 model had better performance characteristics when compared to the other models based on these satisfied conditions: MPE value of 0.740 fall within -10 to 10%, lower RMSE value of 46.588, higher NSE value of 0.739 (that is much closer to 1), higher R value of 0.860, higher R^2 value of 0.7391 and lower t-stat. value of 2.595E-24 and MBE value of -6.88E-12. Furthermore, the H4 model had better prediction performance than the SVM model and a much closer performance characteristics of the MLP and XGBoost machine learning models. Therefore, a GSR prediction models that have better performance characteristics than the existing multivariate and MLR models have been developed in this study. These models are presented thus as equation 17.

$$H_1 = \cos\emptyset + \cos n + \cos^2 n + \cos\delta + \cos^2\delta + T_{max} + \frac{\bar{n}}{N} + \frac{T_{max}}{R.H} + \left(\frac{T_{max}}{R.H}\right)^2 + \left(\frac{T_{max}}{R.H}\right)^3 + \left(\frac{T_{max}}{R.H}\right)^4 + \left(\frac{\bar{n}}{N}\right)^2 + \left(\frac{\bar{n}}{N}\right)^3 + \frac{R.F}{T_{max}} \quad (17a)$$

$$H_2 = \cos\emptyset + \cos n + \cos^2 n + T_{max} + \frac{\bar{n}}{N} + \frac{T_{max}}{R.H} + \left(\frac{T_{max}}{R.H}\right)^2 + \left(\frac{T_{max}}{R.H}\right)^3 + \left(\frac{T_{max}}{R.H}\right)^4 + \left(\frac{\bar{n}}{N}\right)^2 + \left(\frac{\bar{n}}{N}\right)^3 + \frac{R.F}{T_{max}} \quad (17b)$$

$$H_3 = \cos\emptyset + \cos n + \cos^2 n + \cos\delta + \cos^2\delta + T_{max} + \frac{\bar{n}}{N} + \frac{T_{max}}{R.H} + \left(\frac{T_{max}}{R.H}\right)^2 + \left(\frac{T_{max}}{R.H}\right)^3 + \left(\frac{T_{max}}{R.H}\right)^4 + \left(\frac{\bar{n}}{N}\right)^2 + \left(\frac{\bar{n}}{N}\right)^3 + \frac{R.F}{T_{max}} + \left(\frac{R.F}{T_{max}}\right)^2 + \left(\frac{R.F}{T_{max}}\right)^4 \quad (17c)$$

$$H_4 = \cos\emptyset + \cos n + \cos^2 n + \cos\delta + \cos^2\delta + T_{max} + \frac{\bar{n}}{N} + \frac{T_{max}}{R.H} + \left(\frac{T_{max}}{R.H}\right)^2 + \left(\frac{T_{max}}{R.H}\right)^3 + \left(\frac{T_{max}}{R.H}\right)^4 + \left(\frac{\bar{n}}{N}\right)^2 + \left(\frac{\bar{n}}{N}\right)^3 + \frac{R.F}{T_{max}} + \left(\frac{R.F}{T_{max}}\right)^2 + \left(\frac{R.F}{T_{max}}\right)^4 + \frac{R.F}{R.H} + \left(\frac{R.F}{R.H}\right)^2 + \left(\frac{R.F}{R.H}\right)^4 \quad (17d)$$

In addition, the thermodynamic analysis performed in this study for the test region- Asaba town, Delta state, showed that the exergy GSR of the region varied between 20-185 W/m^2.

Declarations

Research grants

This study has not received any funding.

Additional material

Consider the appendix to be supplemental information.

Data Availability

Consider data availability to be made available on request.

References

[1] S. C. Iweka and K. C. Owuama. "Biogas Yielding Potential of Maize Chaff Inoculated with Cow Rumen and Its Characterization," *Journal of Energy Research and Reviews*, pp. 34–50, Nov. 2020, doi: 10.9734/jenrr/2020/v6i330171.

[2] S. C. Iweka, F. C. Ozioko, E. D. Edafiadhe, and T. F. Adepoju, "Bio-oil production from ripe pawpaw seeds and its optimal output: Box-Behnken Design and Machine Learning approach," *Scientific African*, vol. 21, p. e01826, Sep. 2023, doi: 10.1016/j.sciaf.2023.e01826.

[3] S. C. Iweka, K. C. Owuama, J. L. Chukwuneke, and O. A. Falowo. "Optimization of biogas yield from anaerobic co-digestion of corn-chaff and cow dung digestate: RSM and python approach," *Heliyon*, vol. 7, no. 11, p. e08255, Nov. 2021, doi: 10.1016/j.heliyon.2021.e08255.

[4] A. A. Imam, A. Abusorrah, and M. Marzband, "Potentials and opportunities of solar PV and wind energy sources in Saudi Arabia: Land suitability, techno-socio-economic feasibility, and future variability," *Results in Engineering*, vol. 21, p. 101785, Mar. 2024, doi: 10.1016/j.rineng.2024.101785.

[5] S. C. Iweka, O. A. Falowo, A. A. Amosun, and E. Betiku. "Optimization of microwave-assisted biodiesel production from watermelon seeds oil using thermally modified kwale anthill mud as base catalyst," *Heliyon*, vol. 9, no. 7, p. e17762, Jul. 2023, doi: 10.1016/j.heliyon.2023.e17762.

[6] A. Choudhary, D. Pandey, and S. Bhardwaj, "A Review for the Development of ANN Based Solar Radiation Estimation Models," in *Smart Innovation, Systems and Technologies*, Singapore: Springer

Commented [DI18]: References should be in the Journal's format or standard format.

Singapore, 2020, pp. 59–66. Accessed: Jan. 21, 2024. [Online]. Available: http://dx.doi.org/10.1007/978-981-15-5971-6_7

[7] J. A. Duffie and W. A. Beckman, *Solar Engineering of Thermal Processes*. John Wiley & Sons, 2013.

[8] A. B. Da'ie, "Developing mathematical models for global solar radiation intensity estimation at Shakardara, Kabul," *International Journal of Innovative Research and Scientific Studies*, vol. 4, no. 2, pp. 133–138, Mar. 2021, doi: 10.53894/ijirss.v4i2.68.

[9] H. O. Nnabuenyi, L. N. Okoli, F. C. Nwosu, and G. Ibe, "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*, Vol. 3, No. 2, 2017, pp. 8-14.

[10] J. Huang, A. Troccoli, and P. Coppin, "An analytical comparison of four approaches to modelling the daily variability of solar irradiance using meteorological records," *Renewable Energy*, vol. 72, pp. 195–202, Dec. 2014, doi: 10.1016/j.renene.2014.07.015.

[11] M. C. Njoku, D. U. Egwuagu, D. Ncharam, And C. C. Ndefo, "Analysis of Solar Radiation Measurement in South-West Geopolitical Zone of Nigeria", *Global Scientific Journals*, vol. 10, Issue 2, Feb. 2022. ISSN 2320-9186

[12] F. Junliang, W. Lifeng, Z. Fucang, C. Huanjie, Z. Wenzhi, W. Xiukang, and Z. Haiyang, "Empirical and machine learning models for predicting daily global solar radiation from sunshine duration: A review and case study in China," *Renewable and Sustainable Energy Reviews*, vol. 100, pp. 186–212, Feb. 2019, doi: 10.1016/j.rser.2018.10.018.

[13] C. Ertekin and F. Evrendilek, "Spatio-temporal modeling of global solar radiation dynamics as a function of sunshine duration for Turkey," *Agricultural and Forest Meteorology*, vol. 145, no. 1–2, pp. 36–47, Jul. 2007, doi: 10.1016/j.agrformet.2007.04.004.

[14] M. A. Behrang, E. Assareh, A. Ghanbarzadeh, and A. R. Noghrehabadi, "The potential of different artificial neural network (ANN) techniques in daily global solar radiation modeling based on meteorological data," *Solar Energy*, vol. 84, no. 8, pp. 1468–1480, Aug. 2010, doi: 10.1016/j.solener.2010.05.009.

[15] J. Cohen, P. Cohen, S. G. West, and L. S. Aiken, *Applied Multiple Regression/Correlation Analysis for the Behavioral Sciences*. Routledge, 2013.

[16] P. I. Cooper, "The absorption of radiation in solar stills," *Solar Energy*, vol. 12, no. 3, pp. 333–346, 1969, doi: 10.1016/0038-092x(69)90047-4.

[17] R. Chen, K. Ersi, J. Yang, S. Lu, and W. Zhao, "Validation of five global radiation models with measured daily data in China," *Energy Conversion and Management*, vol. 45, no. 11–12, pp. 1759–1769, Jul. 2004, doi: 10.1016/j.enconman.2003.09.019.

[18] M. Neri, D. Luscietti, and M. Pilotelli, “Computing the Exergy of Solar Radiation From Real Radiation Data,” *Journal of Energy Resources Technology*, vol. 139, no. 6, Jun. 2017, doi: 10.1115/1.4036772.