Abstract. Modeling solar radiation is a necessity for the utilization of the benefits it brings to mankind. Time series analysis has proved to stand out amidst other statistical tools when estimating and forecasting solar radiations and their variations. In this paper, a mixture of the Autoregressive Moving Average (ARMA) and Generalized Autoregressive Conditional Heteroscedasticity (GARCH) time series models were implemented on the solar radiation series for Port Harcourt meteorological station, located at the south-southern part of Nigeria, to capture and model the conditional mean and volatility that may exist in the series. After subjecting the models to some evaluation metrics for model adequacy, the results gave appropriate ARMA model for the station, indicated the presence of volatility in the radiation series and these volatilities were modeled using the combination of ARMA-GARCH models, which produced a better estimate than the ARMA models alone.

Keywords. Models; Solar radiation; ARMA; GARCH; Volatility

MSC. 78Axx

Received: May 30, 2017 Revised: June 29, 2017 Accepted: July 11, 2017

1. Introduction

In most developed countries, the use of solar energy as an alternative source for generating power is gaining an edge over other sources, despite its maintenance expenses. Solar radiation
is the radiant energy emanating from the sun towards the earth surface. Solar energy warms our planet and gives us our everyday wind and weather. The solar energy keeps the earth in a state of warmth and produces the wind and weather, experienced on a daily basis. In the absence of this radiant energy from the sun, the earth becomes too cool for living creatures to survive (Nwankwo and Nnabuchi [13]).

The sun is an unending source of natural energy that when compared with other forms of renewable energy, has the potential for a broad range of applications due to its accessibility. The closer the earth is to the sun, the more the intensity of solar energy it receives. Some factors that affect the amount of solar radiation the earth’s surface receives are the geographic region, time of day, time of year, local landscape and local climate condition (Innovateus [?]).

Instruments for measuring solar radiation are generally called Solarimeters. Pyranometers are special kinds of a simple whole sky Solarimeter which is used to measure global solar radiation meanwhile a Pyrheliometer measures beam or direct solar radiation Nwankwo and Nnabuchi [13].

Nigeria is located in the western region of Africa at latitude between 4°N and 13°N and longitude 3°E and 15°E made up of 36 states. A little grasp as to how a particular geographical location in Nigeria encountered variation in solar energy distribution across it might lead to one discovering that variation in solar energy distribution across it might lead to one discovering that the various states that make up Nigeria possesses different meteorological data which accounts for variation in solar energy distribution (Osueke et al. [15]).

Though the measurement of radiant energy transmitted by the sun is not being totally covered in developing countries like Nigeria, meteorological indexes like temperature, rainfall, sunshine hours, etc. are used to extrapolate the solar energy reaching the earth’s surface (Togrul [22]).

Knowing that for various states in Nigeria, there are varying solar radiation intensities. For instance, there is a higher intensity of sun-rays in the Northern part compared to the southern part of Nigeria, so these differences would be considered to improve the accuracy of the models.

In the research community, ARMA methods are widely used and popular time series models compared to other models like Artificial Neural Network Models, Markov Chains, Fuzzy networks, etc. McKenzie [12]. The ARMA models are very flexible; therefore, they can be used in various types of time series with different orders. Finally, it offers a regular pervasiveness at individual phases (identification, estimation and diagnostic checks) for a suitable model. In these models, one of the greatest difficulties is the need for enormous amount of data (Ji and Chee [9]).

Forecasts are critically important for use in monitoring solar systems, energy systems sizing, and optimization and utility applications. Utilities and independent system operators use forecasting information to manage generation and distribution. Extra-terrestrially, there is no stochasticity in solar irradiance; hence, deterministic models are often used to model this data. At ground level, the success of Seasonal ARIMA models are attributed to their abilities to capture the stochastic component of the irradiance series due to the effects of the ever-changing
atmospheric conditions (Ranganai and Nzuza [18]).

The electrical energy produced by solar power systems are not as unchanging as those generated by other sources, which implies that the integration of solar energy into traditional electrical systems are quite demanding (Lewis [19]). In making practical and creative use of solar radiation, modelling is one of the critical difficulties (Yang et al. [26]).

In spite of the fact that accurate mean estimates of the solar radiation are brought to light through various models as proposed by professionals, the heteroscedastic nature of solar radiation is often lost, that is, the variance responsible for non-linearity is mostly left out when modelling solar radiation (Sun et al. [20]).

Through this paper, we are interested in applying a time series statistical tool that have been extensively utilized in finance and financial decisions to solar energy and achieve a better estimate of the mean and volatility (variations) in solar radiation received in Nigeria. Although, countless researchers in Nigeria who are more of physicists and engineers have developed some good models for estimating global radiation, there is little or no attention on modeling and forecasting solar radiation using time series tools especially S/ARMA, GARCH models for mean and volatility of solar radiation series.

Time is an important factor in virtually every aspect of life and human endeavours, which have made researchers from various works of life, explore all areas ranging from economy, business, archaeology, engineering, academia etc. As a result of these, time series analysis has grown to be relevant in all of these fields.

Among the most effective approaches for analysing time series data is the model introduced by Box and Jenkins, Autoregressive Integrated Moving Average (ARIMA). For instance, in a study by, Chowdhury and Biswas [3] an ARIMA model was developed in MATLAB environment for simulating and forecasting the rainfall data for the study area Krisnanaga, India using the Box-Jenkins methodology. The rainfall data covered the period of 1971 to 2010, where the first thirty (30) years i.e. from 1971 to 2000 of the data was used for the model development and the remaining ten (10) years i.e. from 2001 to 2010 of the data was used to verify the developed model. From the study, it was found that the ARIMA model (1,0,0)(1,1,1)_{12} is suitable for forecasting monthly rainfall over the study area and further suggested that the model could be used for forecasting the monthly rainfall for up-coming years.

Appropriate solar data modeling and reliable forecasting of solar radiation is vital for designing, performance prediction and monitoring solar energy conversion systems. One class of models used successfully in the literature to achieve this are the short-memory Box-Jenkins seasonal/non-seasonal Autoregressive Integrated Moving Average (S/ARIMA) stochastic models (Craggs et al. [5]; Zaharim et al. [27]; Voyant et al. [24]).

Generally, it is well-known in time series analysis that the ARMA-GARCH models are used for modeling the mean and volatility in finance (Engle [6], Bollerslev [1]), yet these models have not received much attention in the energy community except for wind-speed estimation (Ewing et al. [8]; Payne and Carroll [17]; Liu et al. [11]).
Recently, Sun et al. [20] conducted an empirical investigation on solar radiation series using ARMA-GARCH models. Representative dataset from two China stations were incorporated into six different ARMA-GARCH models to estimate the mean and volatility of monthly solar radiation time series which out-performed the traditional point forecasting models like the simple Artificial Neural Network (ANN), because ANN was a poorer model in dealing with volatility of solar radiation data. The results reported showed that the ARMA-GARCH(-M) models are efficient in estimating radiation series.

The remaining part of this paper is sectioned as follows: Section 2 reviews the general ARMA and GARCH methodologies. Section 3 describes in details the representative meteorological site under investigation. Section 4 uses the daily solar radiation time series from the site to describe the appropriate ARMA-GARCH model for estimating the mean and volatility that exist in the series with the summaries of the results from the study were given. Finally, in Section 5 a brief remark was made to conclude the paper.

2. Method

2.1 Foundations for ARMA Models

A stationary time series \( \{S_t\} \) is said to be an autoregressive moving average process of order \( p \) and \( q \) written as ARMA\((p,q)\), if it satisfies the difference,

\[
S_t - \phi_1 S_{t-1} - \cdots - \phi_p S_{t-p} = \epsilon_t + \theta_1 \epsilon_{t-1} + \cdots + \theta_q \epsilon_{t-q} .
\]  
(2.1)

\( \{S_t\} \) are the solar radiation series, \( \{\epsilon_t\} \) are white noise (shocks) for the solar radiation process and the coefficients \( \phi \)'s and \( \theta \)'s are such that the model is stationary and invertible. For stationarity, the roots of \( \Phi(B) \) must lie outside the unit circle i.e. \(|B| > 1 \) while the invertibility condition is that the roots of \( \theta(B) \) must lie outside the unit circle.

A general non-seasonal ARIMA\((p,d,q)\) model is \( \Phi(B) \nabla d S_t = \theta(B) \epsilon_t \), where \( \nabla(B) = I - B \).

\[
\Phi(B) = 1 - \phi_1 B - \phi_2 B^2 - \cdots - \phi_p B^p
\]
and

\[
\theta(B) = 1 + \theta_1 B + \theta_2 B^2 + \cdots + \theta_q B^q ,
\]

\[
\Phi(B) \nabla d S_t = \theta(B) \epsilon_t .
\]  
(2.2)

For non-stationary series \( \{S_t\} \), Box and Jenkins [2] proposed that differencing of sufficient order \( d \) could make the series stationary. If the \( d \)th difference denoted by \( \{\nabla^d S_t\} \) satisfies \( \Phi(B) \nabla d S_t = \theta(B) \epsilon_t \), then \( \{S_t\} \) is said to follow an autoregressive integrated moving average model of order \( p, d \) and \( q \), denoted by ARIMA\((p,d,q)\).

The Box-Jenkins procedure is focuses on fitting an ARIMA model to a data, which are of three parts: Identification, Estimation and Verification.

A popular way to choose \( p \) is by minimizing Akaike’s AIC (an information criterion), introduced by Akaike ([?], [?]) defined as,

\[
AIC = -2 \log L + 2k ,
\]  
(2.3)
where $k$ is the number of parameters estimated, (in the above case $p$). The optimal model order is determined by the value of $k$ for which $\text{AIC}_{(k)}$ is minimum.

### 2.2 Foundations for GARCH Models

To model volatility in the series if it exists, we either use Autoregressive Conditional Heteroscedasticity (ARCH) or Generalized ARCH models suggested by Engle [6], and Bollerslev [7] for univariate volatility, having the following properties:

**ARCH Model**

$$s_t = \mu + \epsilon_t,$$  \hspace{1cm} (2.4)

where $s_t$ is the return series (transformed solar radiation series), $\mu$ is a constant and $\epsilon_t$ is the random shock (error term) which is distributed as $\epsilon_t = \sigma_t \Omega_t$ and $(\Omega_t)$ is a sequence of identically and independently distributed random variable with mean zero and variance unity. Then for $a_0 > 0$ and $a_i \geq 0 \ (i > 0)$, we have the innovation,

$$\sigma^2_t = a_0 + \sum_{i=1}^{p} a_i \epsilon^2_{t-i}. \hspace{1cm} (2.5)$$

The model in (2.5) is called ARCH($p$) model. Note, the distribution of $\Omega_t$ can be standard normal, standardized student-$t$, generalized error distribution (GED) or skewed student-$t$ distribution.

**GARCH Model**

We know that for the ARCH model of Engle [6], conditional variance $\sigma^2_t$ is determined based on the dependencies among lags of the return series alone. In the GARCH model, lags of the conditional variance, $\sigma^2_{t-j} \ (j > 0)$ are introduced to further remove the linear dependencies in the return series.

**GARCH Specification**

GARCH($p, q$) Model is then specified as

$$\sigma^2_t = a_0 + \sum_{i=1}^{p} a_i \epsilon^2_{t-i} + \sum_{j=1}^{q} \beta_j \sigma^2_{t-j}. \hspace{1cm} (2.6)$$

Then for $a_0 > 0$, $a_i, \beta_j \geq 0 \ (i, j > 0)$, and $\sum_{i,j=1}^{\max(p,q)}(a_i + \beta_j) < 1$, the GARCH($p, q$) model in (2.6) can be parameterized by applying $\sigma_t = \epsilon^2_t - \sigma^2_t$. Then, we have

$$\epsilon^2_t = a_0 + \sum_{i=1}^{\max(p,q)}(a_i + \beta_j)\epsilon^2_{t-i} + \alpha_t - \sum_{j=1}^{q} \beta_j \alpha_{t-j} \hspace{1cm} (2.7)$$

which is an ARMA representation of the squared residuals, $\epsilon^2_t$.

### 3. Solar Radiation Data from the Sites

The solar radiation dataset for this study were obtained from the Nigerian Meteorological Agency (NIMET), Oshodi, Lagos State office, Nigeria. The parameter made available was the solar radiation series measured in millilitres (ml) using the Gunn-Bellani Radiation Integrator.
as the instrument for reading the radiation in those stations. The representative site under investigation was Port Harcourt with longitude (06.57°E) and latitude (05.01°N) as seen in Table 1 and Figure 1. Furthermore, on average the solar irradiance per square meter is 153.23 W/m$^2$ annually and away from normality in terms of the Kurtosis and Skewness values which are greater than zero, as shown in Table 1.

The investigation periods were from the 1st of January 2011 to 31st of December, 2015 which covered daily observations within those periods. In order to understand the data, some basic statistical summaries like means, variances, etc. were conducted on the data.

Before using the dataset from the stations, a standard conversion was made from ml to watts per sq. meters (1 ml to 13.153 W/m$^2$). And the reason for use of Gunn-Bellani Radiation Integrator relative to a Solarimeter for taking solar radiation readings was because the former was inexpensive and easy to use compared to the later. Computations and graphic displays were done using R statistical programming applications.

Figure 1. Map showing the location of Port Harcourt (Viewsphotos 23)

4. Results and Discussion

Figure 2 shows the time series plot of the solar radiations measured in watts per squared meters (Watts/Sq.m) for Port Harcourt station to the years of observations. From the plot, there seems to be trends and seasonality patterns in the solar radiations throughout these years and also that there appears to be some kind of stationarity in the daily solar radiations for the site. Furthermore, the plot shows that in 2012 the radiation received on a particular day was so extreme (about 400 W/m$^2$) and the 4th quarter of every year, there is a drastic drop in the intensity of solar radiation received in Port Harcourt. It is also worth noting, that the time plot is not adequate enough to give us all the information that is needed. After the solar radiation time plot was constructed, the next step was to perform a test for autocorrelation for the lags in the series using autocorrelation function (ACF) and partial autocorrelation (PACF) plots to have a visual display of its behaviour, which indicates a slow decay as lags increases.
and a cut off at lag 10 for the ACF and PACF residuals of the series respectively. Also, a test for stationarity in the series was further conducted via the ADF and KPSS tests. All of these tests are the basic time series conditions that must be satisfied in order for a particular model to be appropriate for estimation and forecasting purposes. 

Augmented Dickey Fuller (ADF) Test of Table 2 reports a \( p \)-value that is less than 0.05 for Port Harcourt, therefore we reject the null hypothesis of the presence of a unit root. This implies that the solar radiation for the site is stationary and need no differencing, before it can be modeled using Box-Jenkins methodology. Kwiatowski-Phillips-Schmidt-Shin (KPSS) test, the long memory test of Table 2 reports a \( p \)-value that is greater than 0.05, therefore we do not reject the null hypothesis of stationarity for Port Harcourt which agrees with their respective ADF. The next step is to fit the ARMA model to the radiation series.

The ACF plots for Port Harcourt (Figure 3) shows some significant lags. Further confirmation was carried out via Box-Ljung diagnostic test. The Null Hypothesis is that the autocorrelation is not different from 0. The Box-Ljung test in Table 2 with a reported \( p \)-value greater than 0.05 for Port Harcourt implies that we do not reject the null hypothesis of insignificant autocorrelations. Also, the model must follow Normal distribution with mean zero and a constant variance. Squared residuals plot for the site shows cluster of volatility at some points in time. Furthermore, since the ACF and PACF of the squared residuals (Figure 3) for the site displays some significant lags, it implies that we can model volatility for average solar radiation in this site because, there exists a strict white noise (disturbance) which is independent with zero mean and normally distributed. The residuals therefore show some patterns that might be modeled. To implement this, we used the GARCH method to model the conditional variance of the series. The \( p \)-values of Jacque-Bera diagnostic test for volatility is less than 0.05, indicating statistical significance and that of Box-Ljung test for residuals is greater than 0.05, and so we cannot reject the null hypothesis that the autocorrelation of the residuals is different from 0 as observed in Table 3. The model therefore, adequately represents the residuals.

The ARMA(1,2) or ARIMA(1,0,2) model for solar radiation from Port Harcourt site is:

\[
S_{PHt} = 0.9867S_{PHt-1} + \varepsilon_{PHt} - 0.9389\varepsilon_{PHt-1} + 0.0562\varepsilon_{PHt-2},
\]

(4.1)

where \( \{S_{PHt}\} \) are the stationary time series for Port Harcourt Solar radiations, \( \{\varepsilon_{PHt}\} \) are the white noise (or shocks) existing in the series.

The GARCH(1,1) model for Port Harcourt is;

\[
\sigma^2_{PHt} = 74.249 + 0.0845\varepsilon^2_{PHt-1} + 0.8798\sigma^2_{PHt-1}.
\]

(4.2)

The Mixed ARMA-GARCH Model: \(-\)ARMA(1,2)+GARCH(1,1)

\[
S_{PHt} + \sigma^2_{PHt} = 0.9867S_{PHt-1} + \varepsilon_{PHt} - 0.9389\varepsilon_{PHt-1} + 0.0562\varepsilon_{PHt-2} + 74.249 + 0.0845\varepsilon^2_{PHt-1} + 0.8798\sigma^2_{PHt-1},
\]

(4.3)

where \( \{S_{PHt}\} \) are the stationary time series for Port Harcourt Solar radiations, \( \{\varepsilon_{PHt}\} \), \( \{\varepsilon^2_{PHt}\} \) and \( \{\sigma^2_{PHt}\} \) are the white noise (or shocks) and their squares the residual of the ARMA model, and the conditional variance of the GARCH model existing in the series.
Figure 2. Time Plot, ACF and PACF plots for Port Harcourt

Table 1. Summary statistics of the solar radiation series and Location for the Port Harcourt Site

<table>
<thead>
<tr>
<th>Site</th>
<th>Mean</th>
<th>Std. dev.</th>
<th>Min.</th>
<th>Max.</th>
<th>Skewness</th>
<th>Kurtosis</th>
<th>Longitude</th>
<th>Latitude</th>
<th>Height (m)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Port. Harcourt</td>
<td>153.23</td>
<td>52.58</td>
<td>3.95</td>
<td>394.58</td>
<td>0.09</td>
<td>0.27</td>
<td>05.01' N</td>
<td>06.57'E</td>
<td>247.0</td>
</tr>
</tbody>
</table>

Table 2. Shows the test for stationarity and residual analysis for the site

<table>
<thead>
<tr>
<th>Site</th>
<th>Augmented Dickey Fuller Test</th>
<th>KPSS Test</th>
<th>Residual (Box-Ljung) Test</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Lag Value</td>
<td>P-Value</td>
<td>Lag Value</td>
</tr>
<tr>
<td>Port Harcourt</td>
<td>12</td>
<td>-4.5374</td>
<td>0.01</td>
</tr>
</tbody>
</table>

Table 3. Normality test of residuals for ARMA-GARCH Results for the Port Harcourt site

<table>
<thead>
<tr>
<th>Diagnostic tests</th>
<th>Jacque-Bera</th>
<th>Box-Ljung Test</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>X-Squared</td>
<td>d.f</td>
</tr>
<tr>
<td></td>
<td>78.197</td>
<td>2</td>
</tr>
</tbody>
</table>

Table 4 shows the estimates of both models, their standard errors, their AICs and log-likelihoods, with the mixed model having a lower log-likelihood and AIC when compared to the
single model which making the former a better model for prediction. Table 5 is the forecast for Solar Radiation in Port Harcourt for first week of the new year 2016 using only the ARMA(1,2) model which neither considers volatility or reflect changes as new information are available but focuses only on analysing time series data linearly. In other words, the mixed model will consider modeling the noise existing in the ARMA based on the conditional variances as seen in last column of the table. In comparison to the actual value, it is observed that both the ARMA(1,2) and its combination with GARCH(1,1) closely converges to the actual solar radiation readings for Port Harcourt. Looking at Table 5, it is no doubt that the combined model has a significant accuracy to the single ARMA model to address the volatility that exists in the radiation series observed at Port Harcourt meteorological site.

Table 4. Statistical estimates and diagnostic tests for Port Harcourt meteorological site

<table>
<thead>
<tr>
<th>Parameters</th>
<th>ARMA</th>
<th>ARMA-GARCH</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Estimate</td>
<td>Std. Error</td>
</tr>
<tr>
<td>µ</td>
<td>156.1666</td>
<td>9.0270</td>
</tr>
<tr>
<td>ar1</td>
<td>0.9867</td>
<td>0.0049</td>
</tr>
<tr>
<td>ma1</td>
<td>0.9389</td>
<td>0.0245</td>
</tr>
<tr>
<td>ma2</td>
<td>0.0562</td>
<td>0.0243</td>
</tr>
<tr>
<td>α₀</td>
<td>74.2488</td>
<td>19.04922</td>
</tr>
<tr>
<td>α₁</td>
<td>0.0845</td>
<td>0.01406</td>
</tr>
<tr>
<td>β₁</td>
<td>0.8799</td>
<td>0.02053</td>
</tr>
</tbody>
</table>

Diagnostic Tests

| Log-Likelihood | -9539.07 | -9472.043 |
| AIC            | 19088.14 | 18950.09  |
| MAPE           | 14.606   | 14.3271   |
| RMSE           | 17.118   | 18.662    |
| R²             | 0.013    | 7.13E-05  |

Table 5. One-Week ahead forecast for 2016

<table>
<thead>
<tr>
<th>Day</th>
<th>Point Forecast</th>
<th>95% Lower</th>
<th>95% Upper</th>
<th>Actual Radiation</th>
<th>ARMA(1,2)-GARCH(1,1)</th>
</tr>
</thead>
<tbody>
<tr>
<td>01-Jan-16</td>
<td>223.788</td>
<td>135.658</td>
<td>311.918</td>
<td>193.35</td>
<td>227.2</td>
</tr>
<tr>
<td>02-Jan-16</td>
<td>223.554</td>
<td>135.324</td>
<td>311.784</td>
<td>228.36</td>
<td>226.6</td>
</tr>
<tr>
<td>03-Jan-16</td>
<td>222.661</td>
<td>133.961</td>
<td>311.361</td>
<td>238.07</td>
<td>226.0</td>
</tr>
<tr>
<td>04-Jan-16</td>
<td>222.779</td>
<td>132.625</td>
<td>310.934</td>
<td>227.54</td>
<td>225.3</td>
</tr>
<tr>
<td>05-Jan-16</td>
<td>220.909</td>
<td>131.314</td>
<td>310.505</td>
<td>201.24</td>
<td>224.7</td>
</tr>
<tr>
<td>06-Jan-16</td>
<td>220.051</td>
<td>130.029</td>
<td>310.073</td>
<td>199.92</td>
<td>224.0</td>
</tr>
<tr>
<td>07-Jan-16</td>
<td>219.204</td>
<td>128.768</td>
<td>309.64</td>
<td>226.23</td>
<td>223.4</td>
</tr>
</tbody>
</table>
5. Conclusion

We observed that, the proposed model which closely mimics the solar radiations received in Port Harcourt is the combined ARMA(1,2)-GARCH(1,1) model. The GARCH model makes provision for variations (or volatility) that exist in the surface radiation compared to the single ARMA(1,2) model which only focuses on the linearity of the radiation time series. From the one-week ahead forecast, it was observed that as the day increases, both models follow a consistent decreasing pattern relative to the actual values. It is important to recall the mathematical expression of both model combinations as follows:

The Mixed ARMA-GARCH Model: –ARMA(1,2)+GARCH(1,1)

\[
S_{PH_t} + \sigma^2_{PH_t} = 0.9867S_{PH_{t-1}} + \varepsilon_{PH_t} - 0.9389\varepsilon_{PH_{t-1}} + 0.0562\varepsilon_{PH_{t-2}} + 74.249 + 0.0845\varepsilon^2_{PH_{t-1}} + 0.8798\sigma^2_{PH_{t-1}}.
\]
It can therefore, be safely recommended that, the ARMA(1,2)-GARCH(1,1) is an adequate model to forecast the solar radiation for Port Harcourt, Rivers State, South-Southern Nigeria, which is an integral part in the application of solar energy and systems in the energy sector of the economy.

**Acknowledgment**

We cease this opportunity to appreciate God Almighty for this work, for the strength and insight given to make this paper a reality. Also, the Nigerian Meteorological Agency (NIMET) for making the data available for this research. The partial support received from Covenant University, Nigeria is also acknowledged.

**Competing Interests**

The authors declare that they have no competing interests.

**Authors’ Contributions**

All the authors contributed significantly in writing this article. The authors read and approved the final manuscript.

**References**


