Original Research Article

Time Series and Empirical Orthogonal Transformation using Meteorological Parameters across the Climatic Zones in Nigeria

.

ABSTRACT

Time series and empirical orthogonal transformation analysis was carried out for four (4) selected tropical sites, which are situated across the four different climatic zones, viz. Sahelian, Midland, Guinea savannah and Coastal region in Nigeria using measured monthly average daily global solar radiation, maximum and minimum temperatures, sunshine hours, rainfall, wind speed, cloud cover and relative humidity meteorological data during the period of thirty one years (1980-2010). Seasonal Auto Regressive Integrated Moving Average (ARIMA) models were developed along with their respective statistical indicators of coefficient of determination (R^2), Root Mean Square Error (RMSE), Mean Absolute Percentage Error (MAPE) and Mean Absolute Error (MAE). The results indicated that the models were found suitable for one step ahead global solar radiation forecast for the studied locations. Furthermore, the results of the time series analysis revealed that the model type for all the meteorological parameters show a combination of simple seasonal with one or more of either ARIMA, winter's additive and winter's multiplicative with the level been more significant as compared to the trend and seasonal variations for the exponential smoothing model parameters in all the locations. The results of the correlation matrix revealed that the global solar radiation is more correlated to the mean temperature except for Akure where it is more correlated to the sunshine hours; the mean temperature is more correlated to the global solar radiation; the rainfall is more correlated to the relative humidity and the relative humidity is more correlated to the rainfall in all the locations. The results of the component matrix revealed that three seasons are identified in Nguru located in the Sahelian region namely, the rainy, the cool dry (harmattan) and the hot dry seasons while in Zaria, Makurdi and Akure located in the Midland, Guinea savannah and Coastal zones two distinct seasons are identified namely, the rainy and dry seasons.

Keywords: Time series, empirical orthogonal transformation, climatic zones, ARIMA, correlation matrix, component matrix.

1. INTRODUCTION

A time series is basically defined as a set of statistical data that are usually collected at regular intervals. It is also a sequence of observations on a variable measured at successive points in time or over successive periods of time [1]. The measurements can be taken every hour, day, week, month, or year, or at any other regular interval found suitable by the investigator. Understanding the pattern of the data is an important factor to know how the time series has behaved in the past and would give an idea on how it can behave at that particular time of investigation. If such behaviour can be expected to continue in the future, the past pattern can be used to decide on an appropriate forecasting method.

Time series data occur naturally in many field of application. The methods of time series analysis predate those for general stochastic processes and Markov Chains. The aims of time series analysis are to describe and summarize time series data, fit low-dimensional models, and make forecasts [2].

World climate change has been one of the most important topics in water resources studies. Weather parameters such as precipitation, monthly temperature and relative humidity forecasting could be practically useful in making decisions, risk management and most favourable usage of water resources. These three meteorological parameters have incontestable effects on hydrological cycle, production of crops products cycle, water usage specifically agricultural usage, people activities and the environments [3]. The correctness of the forecasting of the PV plant output power is mainly dictated by the accuracy of the forecasting of the solar energy resources. The forecasting of the solar irradiance is presently done by statistical methods, among which the Autoregressive Integrated Moving Average (ARIMA) model [4] is the most popular.

Empirical orthogonal function (EOF) analysis or principal component analysis (PCA) have become standard statistical techniques in the geophysical sciences of meteorology and oceanography particularly in the field of climate research [5]. This method of statistical analysis is used to reduce any complicated data set into a finite and small number of new variables. It uses correlation between variable to determine a smaller number of new variables called components that can give vital information about the data [5].

Time series analysis of solar radiation data is vital in predicting long-term average performance of solar energy systems. Thus, several researches have been carried out in this area under discussion. In the study of Sulaiman et al. [6], the Box-Jenkins approach was applied to daily solar radiation data from four different locations in Malaysia. The deterministic component is removed using Fourier analysis and the stochastic component subject to analysis by Auto Regressive Moving Average (ARMA). From their study, it was found that the residuals are best described by ARMA (2,0). Zaharim et al. [7] employed ARMA Box-Jerkins method to model global solar radiation data in Malaysia. Hejase and Assi [8] uses time series methods to predict the monthly average daily global solar radiation using mean air temperature, mean wind speed, daily sunshine hours, relative humidity and global solar radiation obtained from the National Centre of Meteorology and Seismology (NCMS) in Abu Dhabi during the period from 1995 – 2007. Several forecasting approaches have been used in literature. Among these, the most effective in producing hour-ahead predictions are based on empirical regression, neural networks [9] and time-series models (e.g., ARMA, ARIMA) [10, 11]. However, day/month/year ahead forecast has been carried out by several researchers.

Analysis of the inter annual variability of solar radiation and sunshine hour for Brazil was investigated by Tiba and Fraidenraich [12] to generate statistical parameter for model checking or to be used as input data of synthetic time series generation, it was reported that the AR-1 is the recommended method for monthly solar radiation synthetic time series generation, with auto-correlation coefficient varying from 0.30 to 0.40 for the localities in the north of Brazil and zero for other zones. Also, Rich et al. [13] developed statistical models, Artificial Neural Network (ANN) models, satellite imaging based models, numerical based models and hybrid methods for solar irradiance forecasting. In their work, they found that regressive methods such as AR, MA, ARMA and ARIMA take advantage of the correlated nature of the irradiance observations and tend to work well in both data-poor and data-rich environments. Guo et. al. [14] used both ARMA and GARCH models and its extension for forecasting wind speed. SalahShour et al. [15] carried out a study on potential survey stations in Khuzestan province in order to generate solar electricity, their studies was conducted, using data from 15 variables surrounding stations Khuzestan using empirical orthogonal transformation (factor analysis) for ranking the province of solar power.

More recently, Tijjani et al. [16] carried out a time series analysis on magnetic indices of Auroral Electrojet (AE), Auroral Upper (AU), Auroral Lower (AL) and Auroral Oval (AO) during the period of six years (2008 – 2013) using SPSS version 16.0 with expert modeler. The year 2014 was also forecasted in their study. In another study Aliyu et al. [5] used Empirical Orthogonal Function (EOF) to assess the MODIS C006 LV2 aerosol AOD and AE products which was compared to AERONET AOD and AE observations data. Their results showed that seasonal variation of AOD peaks during the dry season from December to February and reaches minimum during summer in August 2008. The analysis of EOF revealed a good correlation between MODIS and AERONET AOD were observed on the correlation matrices in all the data.

The aim of this present study is to compare the time series and empirical orthogonal transformation statistical analysis on the meteorological parameters of monthly average daily global solar radiation, sunshine hours, wind speed, mean temperature, rainfall, cloud cover and relative humidity during the period of thirty one years (1980 – 2010) for four (4) selected tropical sites, which are situated across the four different climatic zones, viz. Sahelian, Midland, Guinea savannah and Coastal region in Nigeria with a view of investigating their variation.

2. METHODOLOGY

2.1 ACQUISITION OF DATA

It has been reported according to the World Meteorological Organization [17] and Ojo and Adeyemi [18] that to ensure the optimal climate modeling, data series should extend to at least thirty years long. In this regard, the measured monthly average daily global solar radiation, maximum and minimum temperatures, sunshine hours, rainfall, wind speed, cloud cover and relative humidity utilized in this study covered a period of thirty one years (1980-2010) for all the locations under investigation. The meteorological data for the four (4) selected study areas situated across the four climatic zones in Nigeria were obtained from the Nigerian Meteorological Agency (NIMET), Oshodi, Lagos, Nigeria. The stations that are located within the climatic zones of the study areas for this present research work are shown in **Fig. 1.**

Fig. 1. A map of Nigeria showing the four climatic zones and the studied stations.

2.2 TIME SERIES

Time series has been defined by Prema and Uma Rao [1] as a sequence of observations on a variable measured at successive points in time or over successive periods of time The measurements are usually taken every hour, day, week, month, or year, or at any other regular interval depending on the investigation. The forecasting of the solar irradiance is currently done by statistical methods, among which the Autoregressive Integrated Moving Average (ARIMA) model [4] is the most popular and was adopted using expert modeller with IBM SPSS version 20 software in this present study. The expert modeller decides the best model as to either ARIMA or exponential smoothing based on which model gives the highest coefficient of determination (R^2). In exponential smoothing, the level, trend and seasonal variations for the meteorological parameters are described graphically with their respective estimates and significant level mainly at 95% confidence level.

An ARIMA model expresses the observation Z_t at the time t as a linear function of previous observations, a current error term and a linear combination of previous error terms. A seasonal ARIMA model is usually denoted $ARIMA\,(p,d,q) \times (P,D,Q)_s$ and contains the following terms: (1) $AR(p)$ – the non-seasonal autoregressive term of order p; (2) $I(d)$ – the non-seasonal differencing of order d; (3) $MA(q)$ – the non-seasonal moving average term of order q; (4) $AR_s(P)$ – the seasonal autoregressive term of order P; (5) $I(D)$ – the seasonal differencing of order D; (6) $MA_s(Q)$ – the seasonal moving average term of order Q.

In general the seasonal ARIMA model is expressed as [19]:

$$\phi_p(B)\Phi_P(B^s)\nabla^d\nabla_s^D Z_t = c + \theta_q(B)\Theta_Q(B^s)\varepsilon_t \tag{1}$$

where B is the backshift operator defined by the expression:

$$BZ_t = Z_{t-1} \quad (2)$$

where ε_t is the random error at time t usually assumed with normal distribution, zero mean and standard deviation σ_a (white noise). Sometimes an adjustment constant c is included in equation (1). The ARIMA model was developed using coefficients ϕ, Φ, θ and Θ obtained from the analysis.

$$\phi_p(B) = 1 - \phi_1 B - \phi_2 B^2 \ldots \ldots - \phi_p B^p \text{ , } \theta_q(B) = 1 - \theta_1 B - \theta_2 B^2 \ldots \ldots . -\theta_q B^q$$

$$\nabla^d = (1 - B)^d \text{ , } \nabla_s^D = (1 - B^s)^D$$

$$\Phi_P(B^s) = 1 - \Phi_1 B^s - \Phi_2 B^{2s} \ldots \ldots . \Phi_P B^{ps} \text{ , } \Theta_Q(B^s) = 1 - \Theta_1 B^s - \Theta_2 B^{2s} \ldots \ldots . \Theta_Q B^{Qs}$$

Other formulations using the concept of equation (1) and (2) are

$$B^2 Z_t = Z_{t-2} \text{ , } B^{12} Z_t = Z_{t-12} \text{ , } \nabla Z_t = Z_t - Z_{t-1} = (Z_t - BZ_t) = (1 - B) Z_t \text{ , } \nabla^2 Z_t = \nabla(Z_t - Z_{t-1}) = \nabla Z_t - \nabla Z_{t-1} \text{ , } \nabla_s Z_t = Z_t - Z_{t-s} = (1 - B^s) Z_t$$

ARIMA model ignores the independent variable completely, and uses past and present values of dependent variable to produce accurate short-term forecasting [20].

2.3 EMPIRICAL ORTHOGONAL TRANSFORMATION

The Empirical Orthogonal transformation Analysis or simply the factor analysis attempts to identify underlying variables, or factors, that explain the pattern of correlations within a set of observed variables. It is often used in data reduction to identify a small number of factors that explain most of the variance that is observed in a much larger number of manifest variables.

The factor analysis using IBM SPSS Statistics version 20 was carried out to determine the correlation matrix, Kaiser-Meyer-Olkin (KMO) and Bartlett's test, component matrix and scree plots for the studied locations. The correlation matrix shows how each of the meteorological parameters is correlated to other parameters; the KMO and Bartlett's test measures the sampling adequacy and test of sphericity in which the acceptable value should be ≥ 0.05 (50%); the component matrix obtained for this study was to determine the prevalence of rainy/dry season in each component based on rainfall, relative humidity, solar radiation and mean temperature. High rainfall and relative humidity and low solar radiation and mean temperature indicates prevalence of rainy season while high solar radiation and mean temperature and low rainfall and relative humidity indicates prevalence of dry season.

3. RESULTS AND DISCUSSION

3.1 SAHEL SAVANNAH

Time Series for Nguru

The seasonal $ARIMA\ (0,0,0) \times (0,1,1)_{12}$ model with $seasonal\ MA = 0.789$ developed for Nguru with the global solar radiation as dependent variable is given by the expression

$$Z_t = Z_{t-12} - 0.789\varepsilon_{t-12} + \varepsilon_t \quad (3)$$

The R^2, RMSE, MAPE and MAE found in this study for Nguru are 54.5 %, 1.573 $MJm^{-2}day^{-1}$, 4.628 % and 1.098 % respectively.

Table 1a: Model Description for Nguru

			Model Type
Model ID	Global solar radiation	Model_1	Simple Seasonal
	Sunshine hours	Model_2	Simple Seasonal
	Wind speed	Model_3	Simple Seasonal
	T_{mean}	Model_4	Simple Seasonal
	Rainfall	Model_5	Winters' Additive
	Cloud cover	Model_6	ARIMA(0,0,2)(0,0,0)
	Relative humidity	Model_7	Simple Seasonal

Table 1a shows that the model type for the model description is simple seasonal for all the meteorological parameters for Nguru, except for rainfall and cloud cover with winter's additive and ARIMA models respectively.

Table 1b: Exponential Smoothing Model Parameters for Nguru

Model			Estimate	Sig.
Global solar radiation-Model_1	No Transformation	Alpha (Level)	0.099	0
		Delta (Season)	6.20E-05	0.998
Sunshine hours-Model_2	No Transformation	Alpha (Level)	0.099	0
		Delta (Season)	8.91E-06	1
Wind speed-Model_3	No Transformation	Alpha (Level)	0.300	0
		Delta (Season)	5.76E-06	1
T_{mean}-Model_4	No Transformation	Alpha (Level)	0.200	0
		Delta (Season)	9.24E-06	1
Rainfall-Model_5	No Transformation	Alpha (Level)	0.021	0.072
		Gamma (Trend)	0	0.85
		Delta (Season)	0.001	0.959
Relative humidity-Model_7	No Transformation	Alpha (Level)	0.099	0.001
		Delta (Season)	8.91E-08	1

Table 1c: ARIMA Model Parameters for Nguru

					Estimate	Sig.
Cloud cover-Model_6	Cloud cover	No Transformation	Constant		6.971	0
			MA	Lag 2	-0.393	0

Table 1 (b&c) shows the exponential smoothing model for the meteorological parameters for Nguru. The results showed that all the parameters have only level and seasonal variations, except for rainfall with level, trend and seasonal variations and cloud cover with ARIMA model. It is obvious that the variation of the level is more dominant as compared to the trend and seasonal variations since the significant level for the level is less than 0.05 except for the rainfall with 0.072 while that of the trend and seasonal variations are greater than 0.05 at 95% confidence level. The estimates of the ARIMA model for cloud cover are given as $constant = 6.971;\ and\ MA, Lag\ 2 = -0.393$.

Empirical Orthogonal Transformation for Nguru

Table 1d: Correlation Matrix for Nguru

		GSR	SSH	WS	T_{mean}	RF	CC	RH
Correlation	GSR	1	-0.004	-0.041	0.663	-0.140	-0.064	-0.002
	SSH	-0.004	1	-0.02	-0.007	-0.203	-0.007	-0.074
	WS	-0.041	-0.02	1	-0.174	-0.112	-0.005	-0.231
	T_{mean}	0.663	-0.007	-0.174	1	0.139	-0.035	0.351
	RF	-0.140	-0.203	-0.112	0.139	1	0.058	0.780
	CC	-0.064	-0.007	-0.005	-0.035	0.058	1	0.071
	RH	-0.002	-0.074	-0.231	0.351	0.780	0.071	1

where GSR is the global solar radiation in $MJm^{-2}day^{-1}$, SSH is the sunshine hours in hours, WS is the wind speed in ms^{-2}, T_{mean} is the mean temperature in ^{0}C, RF is rainfall in mm, CC is the cloud cover and RH is the relative humidity in percentage (%).

Table 1e: KMO and Bartlett's Test for Nguru

Kaiser-Meyer-Olkin Measure of Sampling Adequacy.		0.502
Bartlett's Test of Sphericity	Approx. Chi-Square	726.904
	Df	21
	Sig.	0

In Table 1d, the correlation matrix for Nguru shows how each of the meteorological parameters is correlated to other parameters. It was observed that the global solar radiation is more correlated to the mean temperature with 66.3 %. The sunshine hours and the wind speed have negative correlation with other meteorological parameters. The mean temperature is more correlated to the global solar radiation with 66.3 %. The rainfall and cloud cover are more correlated to the relative humidity with 78.0% and 7.1% respectively.

The relative humidity is more correlated to the rainfall with 78.0%. The results showed that a negative correlation (inverse relationship) exists between the global solar radiation and the meteorological parameters of sunshine hours, wind speed, rainfall, cloud cover and relative humidity. Negative correlations (inverse relationship) exist between the sunshine hours and the meteorological parameters of global solar radiation, wind speed, mean temperature, rainfall, cloud cover and relative humidity. Negative correlations (inverse relationship) exist between the wind speed and the meteorological parameters of global solar radiation, sunshine hours, mean temperature, rainfall, cloud cover and relative humidity. Negative correlations (inverse relationship) exist between the mean temperature and the meteorological parameters of sunshine hours, wind speed, and cloud cover. Negative correlations (inverse relationship) exist between the rainfall and the meteorological parameters of global solar radiation, sunshine hours and wind speed. Negative correlations (inverse relationship) exist between the cloud cover and the meteorological parameters of global solar radiation, sunshine hours, wind speed and mean temperature. Negative correlations (inverse relationship) exist between the relative humidity and the meteorological parameters of global solar radiation, sunshine hours and wind speed.

In Table 1e, the Kaiser-Meyer-Olkin (KMO) shows that the sampling adequacy of 50.2% was achieved. The Bartlett's test of sphericity gives degree of freedom of 21 and it's significant at 95% confidence level.

Table 1f Component Matrix for Nguru

	Component		
	1	2	3
Relative humidity	0.876	-0.303	0.084
Rainfall	0.768	-0.509	-0.094
Global solar radiation	0.311	0.858	-0.144
T_{mean}	0.645	0.662	-0.027
Sunshine hours	-0.207	0.196	0.791
Wind speed	-0.385	-0.043	-0.538
Cloud cover	0.064	-0.212	0.277

Extraction Method: Principal Component Analysis.
a. 3 components extracted.

Table 1f shows the component matrix for Nguru. For component 1 the relative humidity and rainfall has correlation of 87.6 % and 76.8 % while the mean temperature and global solar radiation has correlation of 64.5 % and 31.1 %, this is an indication that the rainy season is prevalence. Component 2 shows that the relative humidity and rainfall has negative correlation of 30.3 % and 50.9 % while the mean temperature and global solar radiation has correlation of 66.2 % and 85.8 %, this is an indication that the dry season is prevalence. Component 3 shows that the relative humidity and rainfall has positive and negative correlation of 8.4 % and 9.4 % while the mean temperature and global solar radiation has negative correlation of 2.7 % and 14.4 % this is an indication that the cool dry season is prevalence. The study region revealed that three seasons are identified; the rainy season, the cool dry (harmattan) season and the hot dry season.

Fig. 2a shows the scree plot for Nguru. The eigenvalue decreases from 2.20 corresponding to component number 1 until it gets to eigenvalue 1.00 corresponding to component number 3 and further decreases to eigenvalue 1.80 to 0.25 with a negative slope of about 1.55. The eigenvalue decreases from 0.25 to 0.20 corresponding to component numbers 6 and 7 respectively. It was observed that the eigenvalue of at least 1 for Nguru was found to be component numbers 1, 2 and 3.

3.2 MIDLAND ZONE

Time Series for Zaria

The seasonal $ARIMA\,(0,0,1)\times(0,1,1)_{12}$ model with $MA=-0.245$ and $seasonal\,MA=0.808$ developed for Zaria with the global solar radiation as dependent variable is given by the expression

$$Z_t = Z_{t-12} + 0.245\varepsilon_{t-1} - 0.808\varepsilon_{t-12} - 0.198\varepsilon_{t-13} + \varepsilon_t \quad (4)$$

The R^2, RMSE, MAPE and MAE found in this study for Zaria are 64.0 %, 1.395 $MJm^{-2}day^{-1}$, 5.141 % and 1.101 % respectively.

Table 2a: Model Description for Zaria

			Model Type
Model ID	Global solar radiation	Model_1	Simple Seasonal
	Sunshine hours	Model_2	Simple Seasonal
	Wind speed	Model_3	Simple Seasonal
	T_{mean}	Model_4	Simple Seasonal
	Rainfall	Model_5	Simple Seasonal
	Cloud cover	Model_6	Simple Seasonal
	Relative humidity	Model_7	Winters' Additive

Table 2a shows that the model type for the model description is simple seasonal for all the meteorological parameters for Zaria, except for relative humidity with winter's additive model.

Table 2b: Exponential Smoothing Model Parameters for Zaria

Model			Estimate	Sig.
Global solar radiation-Model_1	No Transformation	Alpha (Level)	0.100	0
		Delta (Season)	3.42E-05	0.998
Sunshine hours-Model_2	No Transformation	Alpha (Level)	0.1	0
		Delta (Season)	2.37E-05	0.999

Wind speed-Model_3	No Transformation	Alpha (Level)	0.300	0
		Delta (Season)	0	0.996
T_{mean}-Model_4	No Transformation	Alpha (Level)	0.100	0
		Delta (Season)	2.86E-06	1
Rainfall-Model_5	No Transformation	Alpha (Level)	0.008	0.263
		Delta (Season)	7.72E-05	0.996
Cloud cover-Model_6	No Transformation	Alpha (Level)	0.824	0
		Delta (Season)	0.075	0.198
Relative humidity-Model_7	No Transformation	Alpha (Level)	0.039	0.008
		Gamma (Trend)	8.38E-05	0.902
		Delta (Season)	0.001	0.959

Table 2b shows the exponential smoothing model for the meteorological parameters for Zaria. The results showed that all the parameters have only level and seasonal variations, except for relative humidity with level, trend and seasonal variations. It is obvious that the variation of the level is more dominant as compared to the trend and seasonal variations since the significant level for the level is less than 0.05 except for the rainfall while that of the trend and seasonal variations are greater than 0.05 at 95% confidence level.

Empirical Orthogonal Transformation for Zaria

Table 2c: Correlation Matrix for Zaria

		GSR	SSH	WS	T_{mean}	RF	CC	RH
Correlation	GSR	1	0.290	0.115	0.355	-0.434	0.003	-0.517
	SSH	0.290	1	0.035	-0.136	-0.267	-0.195	-0.313
	WS	0.115	0.035	1	-0.088	-0.118	-0.113	-0.327
	T_{mean}	0.355	-0.136	-0.088	1	0.062	0.196	0.221
	RF	-0.434	-0.267	-0.118	0.062	1	0.156	0.607
	CC	0.003	-0.195	-0.113	0.196	0.156	1	0.238
	RH	-0.517	-0.313	-0.327	0.221	0.607	0.238	1

Table 2d: KMO and Bartlett's Test for Zaria

Kaiser-Meyer-Olkin Measure of Sampling Adequacy.		0.572
Bartlett's Test of	Approx. Chi-	607.016

Sphericity	Square	
	df	21
	Sig.	0

In Table 2c, the correlation matrix for Zaria shows how each of the meteorological parameters is correlated to other parameters. It was observed that the global solar radiation is more correlated to the mean temperature with 35.5 %. The sunshine hours, wind speed and mean temperature are more correlated to the global solar radiation with 29.0 %, 11.5 % and 35.5 % respectively. The rainfall and cloud cover are more correlated to the relative humidity with 60.7 % and 23.8 % respectively. The relative humidity is more correlated to the rainfall with 60.7 %. The results showed that a negative correlation (inverse relationship) exists between the global solar radiation and the meteorological parameters of rainfall and relative humidity. Negative correlations (inverse relationship) exist between the sunshine hours and the meteorological parameters of mean temperature, rainfall, cloud cover and relative humidity. Negative correlations (inverse relationship) exist between the wind speed and the meteorological parameters of mean temperature, rainfall, cloud cover and relative humidity. Negative correlations (inverse relationship) exist between the mean temperature and the meteorological parameters of sunshine hours and wind speed. Negative correlations (inverse relationship) exist between the rainfall and the meteorological parameters of global solar radiation, sunshine hours and wind speed. Negative correlations (inverse relationship) exist between the cloud cover and the meteorological parameters of sunshine hours and wind speed. Negative correlations (inverse relationship) exist between the relative humidity and the meteorological parameters of global solar radiation, sunshine hours and wind speed. In Table 2d, the Kaiser-Meyer-Olkin (KMO) shows that the sampling adequacy of 57.2 % was achieved. The Bartlett's test of sphericity gives degree of freedom of 21 and it's significant at 95% confidence level.

Table 2e: Component Matrix for Zaria

	Component	
	1	2
Relative humidity	0.874	0.073
Rainfall	0.77	-0.093
Global solar radiation	-0.673	0.611
Sunshine hours	-0.553	-0.103
Wind speed	-0.375	-0.187
T_{mean}	0.129	0.853
Cloud cover	0.358	0.508

Extraction Method: Principal Component Analysis.
a. 2 components extracted.

Table 2e shows the component matrix for Zaria. For component 1 the relative humidity and rainfall has correlation of 87.4 % and 77.0 % while global solar radiation and mean temperature has negative and positive correlation of 67.3 % and 12.9 %, this is an indication that the rainy season is prevalence. Component 2 shows that relative humidity and rainfall has positive and negative correlation of 7.3 % and 9.3 % while global solar radiation and mean temperature has correlation of 61.1 % and 85.3 %, this is an indication that the dry season is prevalence. The study region revealed that two distinct seasons are identified; the rainy season and the dry season.

Fig. 2b shows the scree plot for Zaria. The eigenvalue decreases sharply from 2.4 to 1.4 corresponding to component numbers 1 and 2 with a negative slope of about 1. The eigenvalue further decreases from 1.40 to 0.30 and finally to 0.20 corresponding to component number 7. It was observed that the eigenvalue of at least 1 for Zaria was found to be component numbers 1, 2 and 3.

3.3 GUINEA SAVANNAH

Time Series for Makurdi

The seasonal $ARIMA\,(0,1,2) \times (0,1,1)_{12}$ model with $MA, Lag\ 1 = 0.607$, $MA, Lag\ 2 = 0.230$ and $Seasonal\ MA = 0.886$ developed for Makurdi with the global solar radiation as dependent variable is given by the expression

$$Z_t = Z_{t-1} + Z_{t-12} - Z_{t-13} - 0.607\varepsilon_{t-1} - 0.230\varepsilon_{t-2} - 0.886\varepsilon_{t-12} + 0.538\varepsilon_{t-13} + 0.204\varepsilon_{t-14} + \varepsilon_t \quad (5)$$

The R^2, RMSE, MAPE and MAE found in this study for Makurdi are 68.2 %, 1.282 $MJm^{-2}day^{-1}$, 5.606 % and 1.030 % respectively.

Table 3a: Model Description for Makurdi

			Model Type
Model ID	Global solar radiation	Model_1	Simple Seasonal
	Sunshine hours	Model_2	Simple Seasonal
	Wind speed	Model_3	Winters' Multiplicative
	T_{mean}	Model_4	Simple Seasonal
	Rainfall	Model_5	Winters' Additive
	Cloud cover	Model_6	ARIMA(1,0,11)(0,0,0)
	Relative humidity	Model_7	Simple Seasonal

Table 3a shows that the model type for the model description is simple seasonal for all the meteorological parameters for Makurdi, except for wind speed and rainfall with winter's multiplicative and winter's additive; and cloud cover with ARIMA model.

Table 3b: Exponential Smoothing Model Parameters for Makurdi

Model	Estimate	Sig.

Global solar radiation-Model_1	No Transformation	Alpha (Level)	0.200	0
		Delta (Season)	1.75E-05	0.999
Sunshine hours-Model_2	No Transformation	Alpha (Level)	0.2	0
		Delta (Season)	8.42E-05	0.997
Wind speed-Model_3	No Transformation	Alpha (Level)	0.269	0
		Gamma (Trend)	0.001	0.942
		Delta (Season)	0.012	0.328
T_{mean}-Model_4	No Transformation	Alpha (Level)	0.1	0
		Delta (Season)	1.72E-05	0.999
Rainfall-Model_5	No Transformation	Alpha (Level)	0.057	0.013
		Gamma (Trend)	1.49E-07	1
		Delta (Season)	3.62E-05	0.999
Relative humidity-Model_7	No Transformation	Alpha (Level)	0.300	0
		Delta (Season)	1.52E-05	1

Table 3c: ARIMA Model Parameters for Makurdi

					Estimate	Sig.
Cloud cover-Model_6	Cloud cover	No Transformation	Constant		6.888	0
			AR	Lag 1	0.971	0
			MA	Lag 1	0.405	0
				Lag 2	0.5	0
				Lag 11	-0.125	0

Table 3 (b&c) shows the exponential smoothing model for the meteorological parameters for Makurdi. The results showed that all the parameters have only level and seasonal variations, except for wind speed and rainfall with level, trend and seasonal variations and cloud cover with ARIMA model. It is obvious that the variation of the level is more dominant as compared to the trend and seasonal variations since the significant level for the level is less than 0.05 except for the rainfall while that of the trend and seasonal variations are greater than 0.05 at 95% confidence level. The significant level for the estimates of the ARIMA model is significant at 95% confidence level.

Empirical Orthogonal Transformation for Makurdi

Table 3d: Correlation Matrix for Makurdi

		GSR	SSH	WS	T_{mean}	RF	CC	RH
Correlation	GSR	1	0.462	-0.027	0.428	-0.491	-0.252	-0.416
	SSH	0.462	1	-0.014	0.199	-0.475	-0.151	-0.322
	WS	-0.027	-0.014	1	0.295	-0.009	-0.015	-0.182
	T_{mean}	0.428	0.199	0.295	1	-0.276	-0.018	-0.249
	RF	-0.491	-0.475	-0.009	-0.276	1	0.142	0.608
	CC	-0.252	-0.151	-0.015	-0.018	0.142	1	0.205
	RH	-0.416	-0.322	-0.182	-0.249	0.608	0.205	1

Table 3e: KMO and Bartlett's Test for Makurdi

Kaiser-Meyer-Olkin Measure of Sampling Adequacy.		0.688
Bartlett's Test of Sphericity	Approx. Chi-Square	588.911
	df	21
	Sig.	0

In Table 3d, the correlation matrix for Makurdi shows how each of the meteorological parameters is correlated to other parameters. It was observed that the global solar radiation and wind speed are more correlated to the mean temperature with 42.8 % and 29.5 % respectively. The sunshine hours and mean temperature are more correlated to the global solar radiation with 46.2 % and 42.8 % respectively. The rainfall and cloud cover are more correlated to the relative humidity with 60.8 % and 20.5 % respectively. The relative humidity is more correlated to the rainfall with 60.8 %. The results showed that a negative correlation (inverse relationship) exists between the global solar radiation and the meteorological parameters of wind speed, rainfall, cloud cover and relative humidity. Negative correlations (inverse relationship) exist between the sunshine hours and the meteorological parameters of wind speed, rainfall, cloud cover and relative humidity. Negative correlations (inverse relationship) exist between the wind speed and the meteorological parameters of global solar radiation, sunshine hours, rainfall, cloud cover and relative humidity. Negative correlations (inverse relationship) exist between the mean temperature and the meteorological parameters of rainfall, cloud cover and relative humidity. Negative correlations (inverse relationship) exist between the rainfall and the meteorological parameters of global solar radiation, sunshine hours, wind speed and mean temperature. Negative correlations (inverse relationship) exist between the cloud cover and the meteorological parameters of global solar radiation, sunshine hours, wind speed and mean temperature. Negative correlations (inverse relationship) exist between the relative humidity and the meteorological parameters of global solar radiation, sunshine hours, wind speed and mean temperature.

In Table 3e, the Kaiser-Meyer-Olkin (KMO) shows that the sampling adequacy of 68.8 % was achieved. The Bartlett's test of sphericity gives degree of freedom of 21 and it's significant at 95% confidence level.

Table 3f: Component Matrix for Makurdi

	Component	
	1	2
Rainfall	-0.799	0.146
Global solar radiation	0.782	-0.124
Relative humidity	-0.745	-0.054
Sunshine hours	0.671	-0.258
Cloud cover	-0.34	0.278
Wind speed	0.164	0.847
T_{mean}	0.547	0.552

Extraction Method: Principal Component Analysis.
a. 2 components extracted.

Table 3f shows the component matrix for Makurdi. For component 1 the rainfall and relative humidity has negative correlation of 79.9 % and 74.5 % while global solar radiation and mean temperature has correlation of 78.2 % and 54.7 % this is an indication that the rainy season is prevalence. Component 2 shows that rainfall and relative humidity has positive and negative correlation of 14.6 % and 5.4 % while global solar radiation and mean temperature has negative and positive correlation of 12.4 % and 55.2 % this is an indication that the dry season is prevalence. The study region revealed that two distinct seasons are identified; the rainy season and the dry season.

Fig. 2c shows the scree plot for Makurdi. The eigenvalue decreases sharply from 2.70 to 1.25 corresponding to component numbers 1 and 2 with a negative slope of about 1.45. The eigenvalue further decreases from 1.25 to 0.40. It was observed that the eigenvalue of at least 1 for Makurdi was found to be components numbers 1, 2 and 3.

3.4 COASTAL ZONE

Time Series for Akure

The seasonal $ARIMA\ (0,1,1) \times (0,1,1)_{12}$ model with $MA = 0.742$ and $seasonal\ MA = 0.923$ developed for Akure with the global solar radiation as dependent variable is given by the expression

$$Z_t = Z_{t-1} + Z_{t-12} - Z_{t-13} - 0.742\varepsilon_{t-1} - 0.923\varepsilon_{t-12} + 0.685\varepsilon_{t-13} + \varepsilon_t \qquad (6)$$

The R^2, RMSE, MAPE and MAE found in this study for Akure are 67.1 %, 1.764 $MJm^{-2}day^{-1}$, 7.568 % and 1.309 % respectively.

Table 4a: Model Description for Akure

			Model Type
Model ID	Global solar radiation	Model_1	Simple Seasonal
	Sunshine hours	Model_2	Simple Seasonal
	Wind speed	Model_3	Simple Seasonal
	T_{mean}	Model_4	Winters' Additive
	Rainfall	Model_5	Simple Seasonal
	Cloud cover	Model_6	ARIMA(0,0,10)(1,0,0)
	Relative humidity	Model_7	Simple Seasonal

Table 4a shows that the model type for the model description is simple seasonal for all the meteorological parameters for Akure, except for the mean temperature and cloud cover with winter's additive and ARIMA models respectively.

Table 4b: Exponential Smoothing Model Parameters for Akure

Model			Estimate	Sig.
Global solar radiation-Model_1	No Transformation	Alpha (Level)	0.299	0
		Delta (Season)	4.28E-06	1
Sunshine hours-Model_2	No Transformation	Alpha (Level)	0.2	0
		Delta (Season)	3.34E-05	0.999
Wind speed-Model_3	No Transformation	Alpha (Level)	0.2	0
		Delta (Season)	9.50E-05	0.996
T_{mean}-Model_4	No Transformation	Alpha (Level)	0.008	0.411
		Gamma (Trend)	5.63E-06	0.998
		Delta (Season)	0.001	0.969
Rainfall-Model_5	No Transformation	Alpha (Level)	0.1	0
		Delta (Season)	7.38E-06	1

Relative humidity-Model_7	No Transformation	Alpha (Level)	0.1	0
		Delta (Season)	4.75E-05	0.999

Table 4c: ARIMA Model Parameters for Akure

					Estimate	Sig.
Cloud cover-Model_6	Cloud cover	No Transformation	Constant		6.622	0
			MA	Lag 1	-0.258	0
				Lag 6	-0.257	0
				Lag 10	-0.205	0
			AR, Seasonal	Lag 1	0.376	0

Table 4 (b&c) shows the exponential smoothing model for the meteorological parameters for Akure. The results showed that all the parameters have only level and seasonal variations, except for mean temperature with level, trend and seasonal variations and cloud cover with ARIMA model. It is obvious that the variation of the level is more dominant as compared to the trend and seasonal variations since the significant level for the level is less than 0.05 except for the mean temperature while that of the trend and seasonal variations are greater than 0.05 at 95% confidence level. The estimates of the ARIMA model for cloud cover are significant at 95% confidence level.

Empirical Orthogonal Transformation for Akure

Table 4d: Correlation Matrix for Akure

		GSR	SSH	WS	T_{mean}	RF	CC	RH
Correlation	GSR	1	0.638	-0.016	0.528	-0.330	-0.142	-0.417
	SSH	0.638	1	-0.145	0.569	-0.471	-0.164	-0.497
	WS	-0.016	-0.145	1	0.077	0.089	0.093	0.153
	T_{mean}	0.528	0.569	0.077	1	-0.323	-0.034	-0.361
	RF	-0.330	-0.471	0.089	-0.323	1	0.373	0.642
	CC	-0.142	-0.164	0.093	-0.034	0.373	1	0.512
	RH	-0.417	-0.497	0.153	-0.361	0.642	0.512	1

Table 4e: KMO and Bartlett's Test for Akure

Kaiser-Meyer-Olkin Measure of Sampling Adequacy.		0.763
Bartlett's Test of Sphericity	Approx. Chi-Square	851.817
	df	21
	Sig.	0

In Table 4d, the correlation matrix for Akure shows how each of the meteorological parameters is correlated to other parameters. It was observed that the global solar radiation and mean temperature are more correlated to the sunshine hours with 63.8 % and 56.9 % respectively. The sunshine hours is more correlated to the global solar radiation with 63.8 %. The wind speed, rainfall and cloud cover are more correlated to the relative humidity with 15.3 %, 64.2 % and 51.2 % respectively. The relative humidity is more correlated to the rainfall with 64.2 %. The results showed that a negative correlation (inverse relationship) exists between the global solar radiation and the meteorological parameters of wind speed, rainfall, cloud cover and relative humidity. Negative correlations (inverse relationship) exist between the sunshine hours and the meteorological parameters of wind speed, rainfall, cloud cover and relative humidity. Negative correlations (inverse relationship) exist between the wind speed and the meteorological parameters of global solar radiation and sunshine hours. Negative correlations (inverse relationship) exist between the mean temperature and the meteorological parameters of rainfall, cloud cover and relative humidity. Negative correlations (inverse relationship) exist between the rainfall and the meteorological parameters of global solar radiation, sunshine hours and mean temperature. Negative correlations (inverse relationship) exist between the cloud cover and the meteorological parameters of global solar radiation, sunshine hours and mean temperature. Negative correlations (inverse relationship) exist between the relative humidity and the meteorological parameters of global solar radiation, sunshine hours and mean temperature.

In Table 4e, the Kaiser-Meyer-Olkin (KMO) shows that the sampling adequacy of 76.3 % was achieved. The Bartlett’s test of sphericity gives degree of freedom of 21 and it’s significant at 95% confidence level.

Table 4f: Component Matrix for Akure

	Component	
	1	2
Sunshine hours	0.814	0.261
Relative humidity	-0.81	0.332
Rainfall	-0.742	0.269
Global solar radiation	0.731	0.383
T_{mean}	0.665	0.522
Cloud cover	-0.463	0.658
Wind speed	-0.15	0.442

Extraction Method: Principal Component Analysis.
a. 2 components extracted.

Table 4f shows the component matrix for Akure. For component 1 the rainfall and relative humidity has negative correlation of 74.2 % and 81.0 % while global solar radiation and mean temperature has correlation of 73.1 % and 66.5 % this is an indication that the rainy season is prevalence. Component 2 shows that rainfall and relative humidity has correlation of 26.9 % and 33.2 % while global solar radiation and mean temperature has correlation of 38.3 % and 52.2 % this is an indication that the dry season is prevalence. The study region revealed that two distinct seasons are identified; the rainy season and the dry season.

Fig. 2d shows the scree plot for Akure. The eigenvalue decreases sharply from 3.20 to 1.25 corresponding to component numbers 1 and 2 with a negative slope of about 1.95. The eigenvalue decreases subsequently from 1.25 to 0.40. It was observed that the eigenvalue of at least 1 for Akure was found to be components numbers 1, 2 and 3.

4. CONCLUSION

Time series and empirical orthogonal transformation analysis was investigated for four (4) selected tropical sites, situated across the four different climatic zones, viz. Sahelian, Midland, Guinea savannah and Coastal region in Nigeria using measured monthly average daily global solar radiation, maximum and minimum temperatures, sunshine hours, rainfall, wind speed, cloud cover and relative humidity meteorological data during the period of thirty one years (1980-2010). Seasonal ARIMA models were developed for all the locations under study. The ARIMA models developed are one step forecast as it forecasts the values of global solar radiation in the next interval, this provides a form of relief in case of system failure i.e., the pyranometer that is intended to be replaced before the next interval of measurement after the forecasting. The coefficient of determination (R^2) for the developed ARIMA models are $\geq 54.5\%$ for all the locations; Akure has the highest with $R^2 = 69.7\%$. The statistical indicators of Root Mean Square Error (RMSE), Mean Absolute Percentage Error (MAPE) and Mean Absolute Error (MAE) were also obtained for the study areas. The results of the model type indicated by the meteorological parameters in all the locations shows that the simple seasonal is more dormant as compared to the ARIMA, winter's additive and winter's multiplicative. The results of the correlation matrix revealed that the global solar radiation is more correlated to the mean temperature except for Akure where it is more correlated to the sunshine hours; the mean temperature is more correlated to the global solar radiation; the rainfall is more correlated to the relative humidity and the relative humidity is more correlated to the rainfall in all the locations. The Kaiser-Meyer-Olkin (KMO) showed sampling adequacy greater than 50 % for all the studied locations. The Bartlett's test of sphericity gives degree of freedom of 21 and it's significant at 95% confidence level for all the studied locations. The results of the component matrix revealed that three seasons are identified in Nguru located in the Sahelian region namely, the rainy, the cool dry (harmattan) and the hot dry seasons while in Zaria, Makurdi and Akure located in the Midland, Guinea savannah and Coastal zones two distinct seasons are identified namely, the rainy and dry seasons. The scree plots showed that the eigenvalue of at least 1 for all the studied locations was found to be components numbers 1, 2 and 3.

COMPETING INTERESTS

Authors have declared that no competing interests exist.

REFERENCES

1. Prema V, Uma Rao K. Development of statistical time series models for solar power prediction. Renewable Energy. 2015; 83: 100-109.
2. Brockwell PJ, Davis RA. Time series: theory and methods. 2nd ed. Springer. 2013.
3. Shamsnia SA, Shahidi N, Liaghat A, Sarraf A, Vahdat SF. Modeling of Weather Parameters Using Stochastic Methods (ARIMA Model)(Case Study: Abadeh Region, Iran). 2011 International Conference on Environment and Industrial Innovation IPCBEE vol.12 (2011) © (2011) IACSIT Press, Singapore.
4. Box GEP, Jenkins GM. Time series analysis. Forecasting and Control, Holden-Day, San Francisco. 1970.
5. Aliyu R., Tijjani BI, Sharafa SB. Assesment and Comparison of MODIS AOD and AE Product over Ilorin Aeronet Station using the Statistical Analysis of Empirical Orthogonal Function (EOF). Bayero Journal of Pure and Applied Sciences. 2018; 11(1): 223 – 229.
6. Sulaiman MY, Hlaing OOWM, AbdWahab M, Sulaiman ZA. Analysis of residuals in daily solar radiation time series. Renewable Energy. 1997; **11**: 97-105.
7. Zaharim A, Razali AM, Gim TP, Sopian K. Time Series Analysis of Solar Radiation data in the Tropics. European Journal of Scientific Research. 2009; Vol. 25 No 4. 672 – 678.
8. Hejase HAN, Assi AH. Time-Series Regression Model for Prediction of Monthly and Daily Average Global Solar Radiation in Al Ain City-UAE. Proceedings of the Global Conference on Global Warming held on 11 – 14 July, 2011, Lisbon, Portugal. 2011; Pp 1 – 11.
9. Mellit A, Benghanem M, Kalogirou S. An adaptive wavelet network model for forecasting daily total solar-radiation. Applied Energy. 2006; vol. 83, no. 7, pp. 705-722.

10. Perdomo R, Banguero E, Gordillo G. Statistical modeling for global solar radiation forecasting in Bogota. In Photovoltaic Specialists Conference (PVSC), 2010 35th iEEE, jun 2010; pp. 002 374-002 379.
11. Raji MHTA, Boyo AO. Analysis of global solar radiation data as time series data for some selected cities of western part of Nigeria, International Journal of Advanced Renewable Energy Research. 2012; vol. 1, no. I, pp. 14-19.
12. Tiba C, Fraidenraich N. Analysis of monthly time series of solar radiation and sunshine hour in tropical climates. Renewable Energy. 2004; **29**: 1147-1160
13. Rich HI, Hugo TC, Pedro C, Coimbra FM. Solar forecasting methods for renewable energy integration. Progress in Energy and Combustion Science. 2013; 39: 535–576.
14. Guo Z, Dong Y Wang J, Lu H. The Forecasting Procedure for Long-Term Wind Speed in the Zhangye Area. Hindawi Publishing Corporation Mathematical Problems in Engineering Volume 2010, Article ID 684742, 17 pages, doi:10.1155/2010/684742. 2010.
15. SalahShour, F., Moulai Pardeh, A and Khazaee, K (2014). Potential Environmental Assessment Synoptic Stations in Khuzestan Province in Order To Generate Solar Power. Journal of Applied Sciences Research., **10**(6): 565-570.
16. Tijjani BI, Mijinyawa UM, Bala S, Galadanci GSM. Analysis of Auroral Electrojet Magnetic Indices. Bayero Journal of Pure and Applied Science. 2016; **9**(2): 154 – 158.
17. WMO. A Note on Climatological Normal. Technical Note. World Meteorological Organization, Geneva, Switzerland. 1967.
18. Ojo OS, Adeyemi B. Estimation of Solar Radiation using Air Temperature and Geographical Coordinate over Nigeria, The Pacific Journal of Science and Technology. 2014; **15**(2): 78 – 88.
19. Lee C, Ko C. Short-term load forecasting using lifting scheme and ARIMA models. Expert Systems with Applications. 2011; Vol. 38, pp 5902 – 5911.
20. Hendranata A. Water quality trend at the upper part of Johor river in relation to rainfall and run-off pattern. Universiti Teknologi, Malaysia. 2003.