Open Access Library Journal
2022, Volume 9, e9496
ISSN Online: 2333-9721
ISSN Print: 2333-9705

Statistical Analysis of Aerosols Characteristics from Satellite Measurements over East Africa Using Autoregressive Moving Average (ARIMA)

Geoffrey W. Khamala[1*], John W. Makokha[1], Richard Boiyo[2,3]

[1]Department of Science Technology and Engineering, Kibabii University, Bungoma, Kenya
[2]Department of Physical Sciences, Meru University of Science and Technology, Meru, Kenya
[3]Department of Environment, Energy and Natural Resources, County Government of Vihiga, Maragoli, Kenya
Email: *khamalawanjala@gmail.com

How to cite this paper: Khamala, G.W., Makokha, J.W. and Boiyo, R. (2022) Statistical Analysis of Aerosols Characteristics from Satellite Measurements over East Africa Using Autoregressive Moving Average (ARIMA). *Open Access Library Journal*, **9**: e9496.
https://doi.org/10.4236/oalib.1109496

Received: October 28, 2022
Accepted: November 21, 2022
Published: November 24, 2022

Copyright © 2022 by author(s) and Open Access Library Inc.
This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).
http://creativecommons.org/licenses/by/4.0/
Open Access

Abstract

Aerosols have become a major subject of concern at global, regional and local scales. They influence Earth's radiation budget by scattering and absorbing solar energy resulting in atmospheric cooling and warming respectively. However, immense efforts have been devoted to monitoring atmospheric aerosols using various techniques ranging from *in-situ*, ground and satellite-based remote sensing and modeling techniques. Thus, time series analysis and forecasting have gained momentum over recent decades. The current study performed a time series analysis using Box-Jenkins procedure-based ARIMA (Autoregressive Integrated Moving Average) model for aerosol properties (Total Aerosol Optical Depth, TAOD; Absorption Aerosol Optical Depth, AAOD; Scattering Aerosol Optical Depth, SAOD and Direct Aerosol Radiative Forcing, DARF) over EA derived from satellite platforms. The formulation process in MATLAB followed by the current study has been outlined with a view to generating the best fitting seasonal ARIMA $(p, q, d) \times (P\ Q\ D)$ model. The finding for the forementioned characteristics reveals clear seasonal variation, hence, differencing was done. The Autocorrelation Function (ACF) and Partial Autocorrelation Function (PACF) of differenced series are estimated and the significant lags are used to find out the order of the model. The statistical parameters (RMSE, MAE, MAPE, MASE and normalized BIC) were estimated for testing the validity of ARIMA models so formulated. The current study found that ARIMA $(1, 0, 0) \times (2, 1, 2)_{12}$ model is adequate for forecasting and was therefore used to forecast aerosol characteristics for the year 2022-2025 over EA domain. ARIMA model ascertained can be applied to other fields of study such as climatology, and climate change among other areas to predict future values so that timely control measures can effectively be planned.

Subject Areas

Atmospheric Sciences

Keywords

Series, ARIMA, Differencing, Forecast, ACF, PACF

1. Introduction

Aerosols have become a major subject of concern over recent decades in varied domains because of their significant influence on Earth's radiation budget. They influence Earth's radiation budget by scattering or absorbing incoming and outgoing radiation [1] [2] [3]. Scattering of incoming solar radiation results in the cooling of the Earth's atmosphere, whereas absorption by aerosols yields a warming effect. On the other hand, aerosols also modify cloud properties, by enhancing cloud droplet concentration, decreasing the mean droplet size and hence increasing cloud lifetime [3] [4] [5]. These aerosol effects depend on the amount, type and size of dominant aerosol over a given domain.

Due to aerosol effects outlined, massive efforts have been dedicated to monitoring atmospheric aerosols using various techniques ranging from *in-situ* and ground-based remote sensing (AERONET [6]; EARLINET [7]; CARSNET [8]) to satellite-based remote sensing (MODIS [9]; MERRA [10]). Recent efforts have also focused on statistical modeling and prediction [11] [12] and [13] which is intended to determine various atmospheric processes such as sources, transformation processes, transport, trends and sinks of aerosols and their precursors.

In this regard, time series analysis and forecasting deal with understanding the past relationship among the variables by using various modeling techniques with the ultimate goal of obtaining accurate predictions of future values. The time series prediction and forecasting of atmospheric aerosols are hence theoretically and practically of great importance. The current study, therefore, used Box-Jenkins-based ARIMA (Autoregressive Integrated Moving Average) model to perform a time series analysis on aerosol properties over EA. The model was chosen because of covering a wide range of patterns, varying from stationary to non-stationary and seasonal time series. However, in formulating the best fitting ARIMA model, fundamental assumption that the prediction of the future values is dependent on historical sequences of the observed variables was made.

The ARIMA model provides a framework for steering uncertainty into predictions and thus, has widely been used in different fields [14] and [15]. The use of stochastic methods has been drawn-out to various fields of climatology such as wind speed [16]; precipitation [17]; air and water temperature [17] and aerosols [13]. As such, several studies have been conducted in various fields on statistical analysis using ARIMA model. For instance, Soltani *et al.* [18] developed the time-series model to forecast climatic fluctuations. Autoregressive

(AR) models, Moving Average (MA) models or Autoregressive Moving Average (ARIMA) models were used in air-pollution modeling to predict and analyze the time series data [19] [20]. Soni *et al.* [21] performed a statistical analysis of MODIS AODs over the Gangetic-Himalayan region using ARIMA and indicated that AODs followed a Brownian time series motion. Further, the autocorrelation structure indicated a deterministic pattern in the time series over the region. Jere and Moyo [22] performed a residual analysis for autocorrelated econometric model using Box-Jenkins procedure in determining ARIMA model. The study identified that ARIMA (0, 1, 0) is most appropriate for analysis. Zhang [23] presented modelling of epidemiological data using Box-Jenkins procedure and further applied it in the forecasting malaria data over Zambia. Results indicated that the suitable model is ARIMA (1, 0, 0) since ACF has an exponential decay and the PACF has a spike at lag 1 which is an indication of the said model. Khan and Gupta [24] determined a time series forecasting using a hybrid ARIMA and neural network model. Bhatnagar *et al.* [25] used ARIMA based prediction model for time series analysis of COVID-19 cases in India and proposed that ARIMA (1, 1, 0) was the best fitting model based on BIC and regression values (R^2).

1.1. Instrumentation and Data

The ARIMA method is majorly known based on the Box-Jenkins method [26]. The Box-Jenkins method relates to the fitting of ARIMA model to a given data set. It is fabricated to replicate time series values through three steps: identification, estimation, and diagnostic check [13] [21] [27]. Besides the three steps in model advancement, it can further be used to forecast future trends.

AOD data used to characterize and assess fundamental models in the TAOD, AAOD, SAOD and DARF time series, traverses from 2001-2019 and were sourced from various satellite platforms such as moderate Moderate-resolution Imaging Spectroradiometer (MODIS, TAOD), Ozone Monitoring Instrument (OMI, AAOD) and Modern-Era Retrospective Analysis for Research and Applications (MERRA, SAOD and DARF)-2 model. However, these platforms and how to obtain data have been described in detail by [3].

1.2. Scope of the Study

Statistical analysis in the present study was conducted over East Africa which lies between latitudes (12˚S, 5˚N) and between longitudes (28˚E, 42˚E). The figure showing the study domain for the present study has been outlined vividly by Khamala [3]. The region has a great diversity of landforms that includes glaciated mountain peaks with permanent snow cover, plateaus and coastal plain [3]. The domain is bounded by Ethiopia and Sudan to the north, Somalia and the Indian Ocean to the east, Rwanda, Burundi, and DRC to the west, and Mozambique and Zambia to the south. Further description of the study domain and the climatology has been outlined by Khamala *et al.* [3].

2. Methodology

2.1. Arima Model Formulation

The Box-Jenkins methodology for forecasting was first described by Box *et al.* [26]. Mounting a time series model using the Box-Jenkins modeling procedure allows us to determine an ARIMA (p, d, q) model that is simple in providing sufficiently accurate description of the behavior of the data. To formulate the best ARIMA model, at least 50 observations (n) of data are set in Box-Jenkins. However, this data is edited to remove the fill values or other distortions through the use of log or inverse to achieve stabilization. Figure 1 illustrates the forecasting flowchart formulated for forecasting aerosol parameters using commands in MATLAB following Box-Jenkins procedure.

The process begins with downloading raw data for the studied system (input file). The pre-processed input file containing the fill values are processed and imported into MATLAB using the load command and import file command. After the model is trained and calibrated, data is visualized using interactive plots, including ACF and PACF following the process outlined in Figure 1. The number of ACF and PACF to be calculated is denoted as n/4 with n as the number of observations. The calculated ACF and PACF are used in identifying the orders of p and q by matching the patterns with the theoretical patterns of known models.

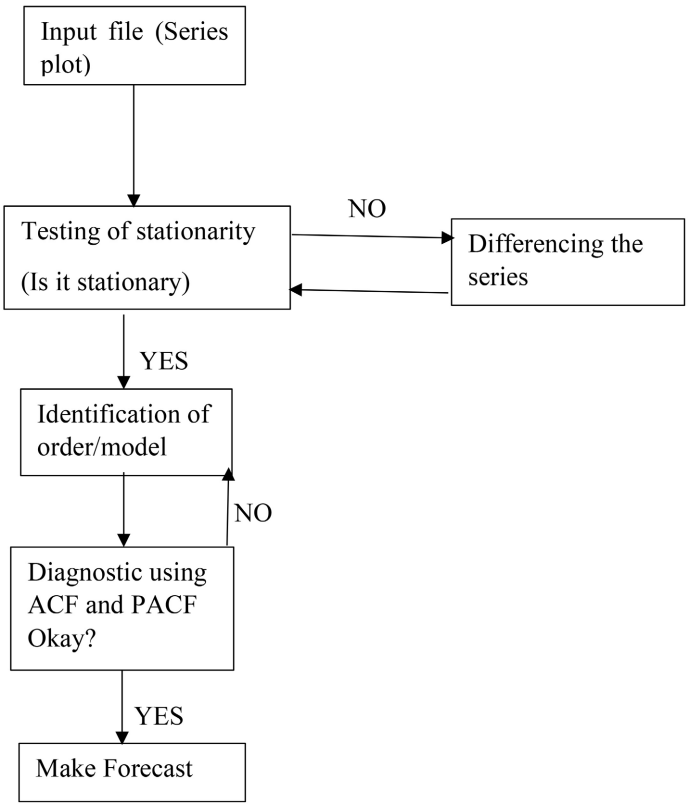

Figure 1. Formulated flowchart for trend simulation and forecast using Box-Jenkins procedure.

2.2. Time Series Prediction and Forecast

A time series in AOD, AAOD, SAOD and DARF is viewed like a series of random variables comprising one or more of the subsequent elements: a trend component (T_t), a seasonal component (S_t), and a random component (R_t). These comportments can be perceived by examining the Autocorrelation Function (ACF) plot and the Partial Autocorrelation Function (PACF) plot. When TAOD, AAOD, SAOD or DARF time series gives a fixed mean and variance, then it is a fixed time series. The erraticism of a fixed time series can thoroughly be explained by an Autoregressive Moving Average (ARMA) model. ARMA is a collective mode of the autoregressive model (AR) that has a p-order and the moving average (MA) model that has a q-order, defined as ARMA (p, q) *i.e.*:

$$\phi_p(B)y_t = \theta_q(B)\varepsilon_t \tag{1}$$

where, y_t signifies MODIS AOD at time point t; B the backshift operator on time series as $By_t = y_{t-1}$, $B^2y_t = BBy_t = By_{t-1} = y_{t-2}$, $\cdots$, $B^k y_t = y_{t-k}$; $\phi_p(B)$ denotes the AR operator as $\phi_p(B) = 1 - \phi_1 B - \phi_2 B^2 - \cdots - \phi_p B^p$; while $\theta_q(B)$ marks the MA operator as $\theta_q(B) = 1 + \theta_1 B + \theta_2 B^2 + \cdots + \theta_q B^q$ and $\varepsilon_t\left(0, \sigma^2\right)$ is the random shocks.

When gradual decay is noted in ACF map, the time series is non-stationary and differencing procedure (d) is required to change it to a stationary one. This time series is outlined as ARIMA (p, d, q) and defined by Equation (2).

$$\phi_p(B)(1-B)^d y_t = \theta_q B\varepsilon_t \tag{2}$$

To justify the annual, seasonal and non-stationary comportments in TAOD, AAOD, SAOD and DARF time series, an extension of multiplicative seasonal ARIMA (SARIMA) model is introduced. It is denoted as ARIMA (p, d, q) × (P, D, Q)$_s$ and given by:

$$\phi_P(B)\Phi_p\left(B^s\right)(1-B)^s\left(1-B^s\right)^D y_t = \theta_q(B)\Theta_Q B^s\varepsilon_t \tag{3}$$

where, D signifies seasonal differencing; s seasonal period (e.g., 2, 3, 4, 12); $\Phi_p\left(B^s\right)$ seasonal AR operator as $\Phi_p\left(B^s\right) = 1 - \Phi_1 B^s - \Phi_2 B^{2s} - \cdots - \Phi_p B^{ps}$ and $\Theta_Q B^s$ is the seasonal MA operator as $\Theta_Q B^s = 1 + \Theta_1 B^s + \Theta_Q B^{2s} + \cdots + \Theta_Q B^{Qs}$. After differencing the time series, the ACF and PACF plots are used to settle parameters p, q, P, and Q. Then, a maximum likelihood estimate is used to fit the model.

To check if the constructed model is suitable to describe variations in TAOD, AAOD, SAOD and DARF time series, fitting accuracy measures will used. If a model fits well, the residuals of the model would be random and the goodness of fitting should be lowermost. Accordingly, this study will use Bayesian Information Criterion (BIC), Root Mean Square Error (RMSE), Mean Absolute Error (MAE), and Mean Absolute Percentage Error (MAPE). These principles were used to reveal a model's abilities in explaining variances of TAOD, AAOD and SAOD time series. The best-designed ARIMA model is the one with least BIC, RMSE, MAE, MASE and MAPE [13]. The best-fitted model ascertained is ap-

plied in forecasting TAOD, AAOD, SAOD and DARF values in sixty seasons (five years). RMSE, MAE, and MAPE are additionally used to measure the accurateness of the model projections. Furthermore, to authenticate with MODIS AOD retrievals, the AERONET AOD measurements available at the given time period are also deemed for site validation.

2.3. Model Parameters

The formulated methodology consists of a three-step iterative procedure. The first step is the identification of a suitable box-Jenkins model where stationarity of the time series is established. To obtain a stationary time series, differencing process is utilized until the seasonality in data diminishes. The AR and MA terms of stationary time series data are then obtained by examining the patterns of ACF and PACF plots, which involves much trial and error. Once tentative models are identified, the next step is estimation of model parameters. This is followed by the diagnostic step to verify the adequacy of the identified models in order to select the best fit model. The model is affirmed adequate if residuals (errors between actual and model predicted values) are random. In this regard, the preferred model is the one that yields minimum Bayesian Information Criterion or BIC [28], which is denoted as

$$\mathrm{BIC} = -2\ln f\left(y \mid \theta k\right) + k\ln\left(n\right) \tag{4}$$

where, "y" is the observed data; "θ" embodies model parameters; "n" signifies sample size and "k" is the number of estimation parameters.

Accordingly, root mean square error (RMSE; Equation (5)), mean absolute error (MAE; Equation (6)) and mean absolute percentage error (MAPE; Equation (7)) were used in determining the best model developed. These principles will used to reveal a model's abilities in explaining variances of TAOD, AAOD and SAOD time series.

$$\mathrm{RMSE} = \frac{1}{n}\sum_{i=1}^{n}\left(y_i - \hat{y}_i\right)^2 \tag{5}$$

$$\mathrm{MAE} = \frac{1}{n}\sum_{i=1}^{n}\left|y_i - \hat{y}_i\right| \tag{6}$$

And

$$\mathrm{MAPE} = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{y_i - \hat{y}_i}{y_i}\right| \tag{7}$$

The best-formulated ARIMA model is the one with least BIC, RMSE, MAE, MASE and MAPE (Li *et al.*, 2019). The best-fitted model ascertained is applied in forecasting TAOD, AAOD, SAOD and DARF values in sixty months/seasons (five years). RMSE, MAE, and MAPE are additionally used to measure the accurateness of the model projections. To authenticate with MODIS AOD, SAOD, OMI and DARF retrievals, the AERONET AOD, extinction AOD, AAOD and DARF measurements available at the given time period are deemed for site validation.

3. Results and Discussion

Time Series Prediction

The long-term series in monthly $TAOD_{550\ nm}$ (actual series-dotted black line and simulated series-dotted red line) average over EA from January 2001 to December 2019 is depicted in Figure 2. The seasonality in the time series can be used in the statistical model, in order to determine the future trends in TAOD. There is a common trend in TAOD, SAOD, AAOD and DARF observed in the seasons of each year. Based on the spatiotemporal trend in forementioned parameters, prediction of future trend is achieved using ARIMA model. The results obtained by following the iterative procedure of ARIMA model estimation is given in the subsequent sections.

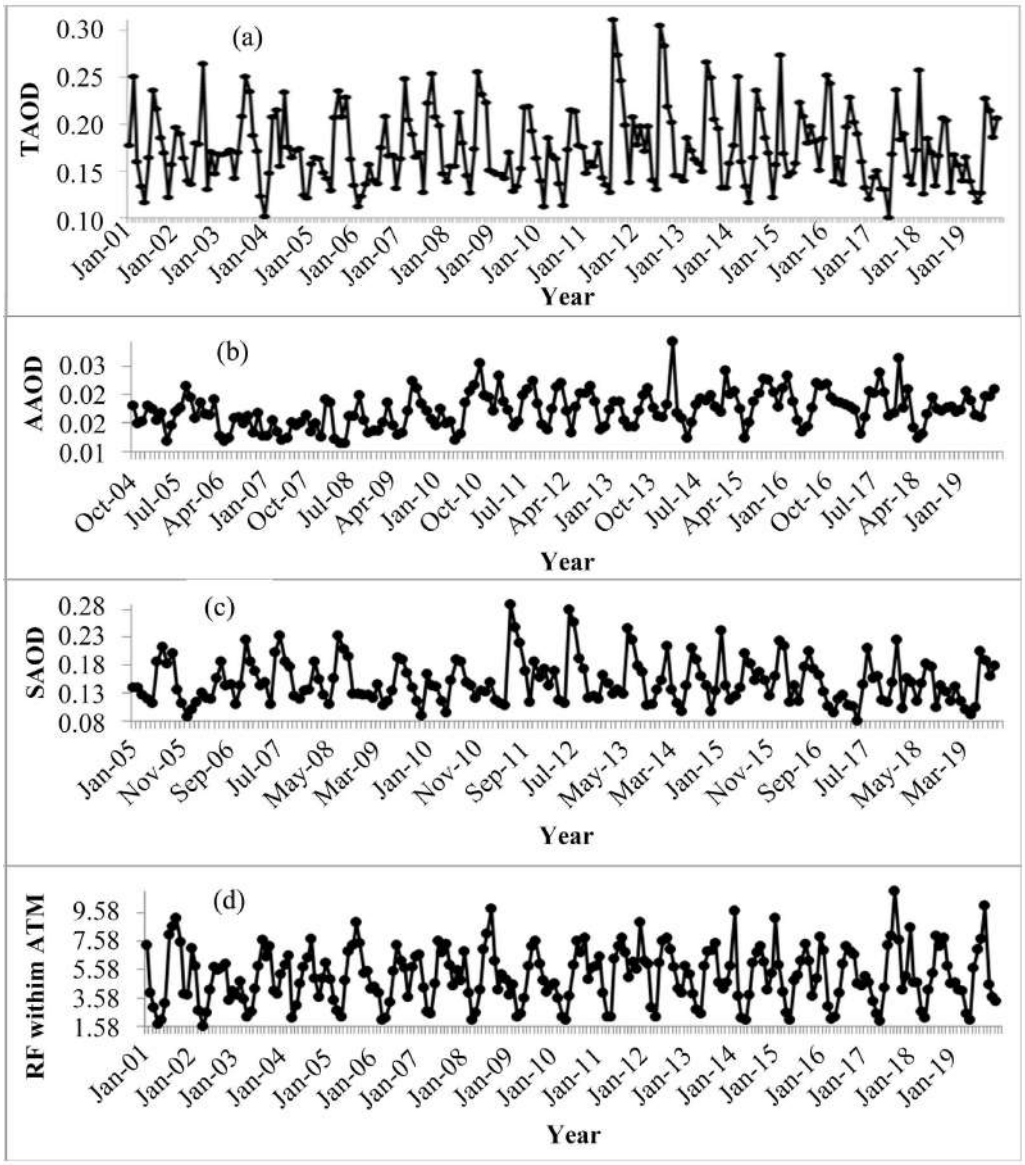

Figure 2. Monthly time series for (a) MODIS TAOD, (b) OMI AAOD, (c) MERRA SAOD, and (d) derived MERRA DARF during 2001 to 2019 over EA.

First, is the identification of stationary time series: Since the time series has varying mean, variance, and autocorrelation over time, it is considered non-stationary. The variations are attributed to atmospheric variables that interfere with the identification of correlation behavior in time series [13]. Therefore, there was need to transform to a stationary series by taking 12-month differencing of data to remove seasonal effect. The patterns of the autocorrelation (ACF) and partial autocorrelation (PACF) plots were then studied for stationary state. In this regard, autocorrelation includes the correlations between time series and its own lag. On the other hand, partial autocorrelation entails correlation coefficients between the time series and its own lag along with elimination of transmission between the individual values, for example between y_t and y_{t-2} with the elimination of influence of the observation y_{t-1}. The sequence obtained after evaluating ACF and PACF is considered stationary if it rapidly converges to zero with the increasing value of lag.

Figure 3 is ACF and PACF correlation graphs for monthly mean TAOD with different selections of seasonal and non-seasonal differencing. The two horizontal lines in the plots denote a p-value of 0.05 (95% confidence intervals) of the estimated autocorrelation and partial autocorrelation coefficients. The x-axis indicates time lag (k), signifying the number of times steps one value is separate from another; the y-axis represents value of the correlation that lies between +1 and −1.

Secondly, is the estimation of the order of ARIMA model: In this step, the orders for ARIMA model were tentatively identified by examining the ACF and PACF graphs of stationary time series. **Figure 3(a)** indicates ACF and PACF plots without any differencing ($d = 0$, $D = 0$). The seasonal autocorrelation relation is significantly shown in this plot. As such, a big spike is noticed at lag 1 in ACF in **Figures 3(a)-(d)**, which reveals a strong correlation of each value of time series with the preceding value. Another significant spike is noted at lag 12 in **Figures 3(a)-(d)**, indicating a strong seasonal correlation between each series value and the value occurring 12 points formerly. It is equally detected that the PACF has significant spike at lag 1 (**Figures 3(a)-(d)**) and dies off thereafter. At the seasonal level, PACF has large spike at lag 12 and cuts off thereafter. Thus, the orders of non-seasonal (p, d, q) and seasonal (P, D, Q) autoregressive and moving average parameters have been determined from the ACF and PACF plots of the stationary time series.

Thirdly, the assessment of ARIMA model parameters: **Table 1** shows the twelve possible combinations for model parameters. The identified models fitted to the data with parameter estimates have been summarizes based on Root Mean Squared Error (RMSE); Mean Absolute Percentage Error (MAPE); (MASE); Mean Absolute Error (MAE) and normalized Bayesian Information Criteria (BIC). The assessed autocorrelations and partial autocorrelations between the residuals at various lags using only two best combinations out of the twelve ARIMA models are depicted in **Figure 4**.

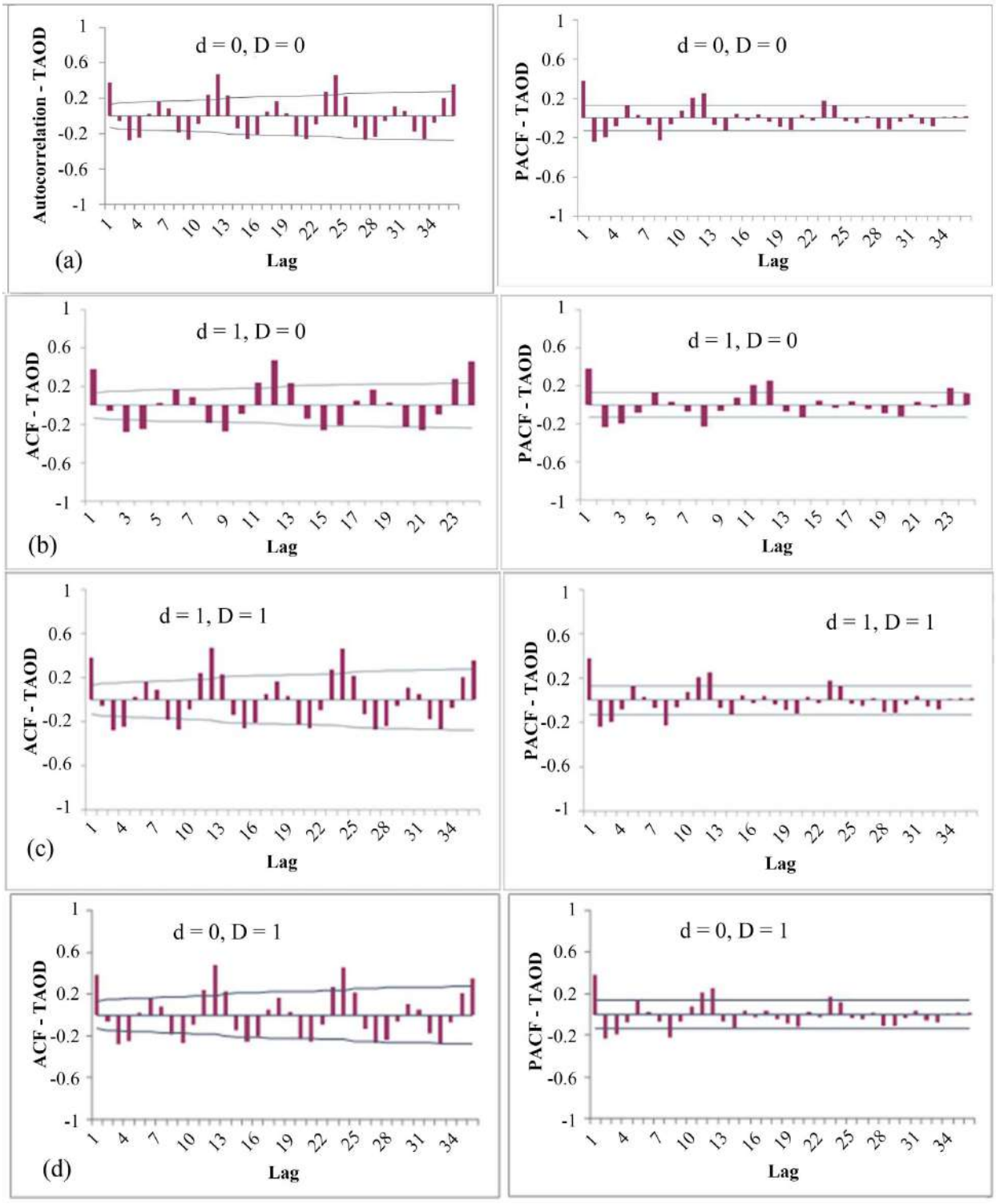

Figure 3. ACF and PACF plots: (a) without differencing, (b) with non-seasonal differencing only, (c) with seasonal differencing only, and (d) with both seasonal and non-seasonal differencing both.

It is distinct that all the lags are within the 95% confidence level at a *p*-value of 0.05. This infer those residuals are random (white noise), specifying that the models are a good fit. It is also observed that all 34 autocorrelation coefficients

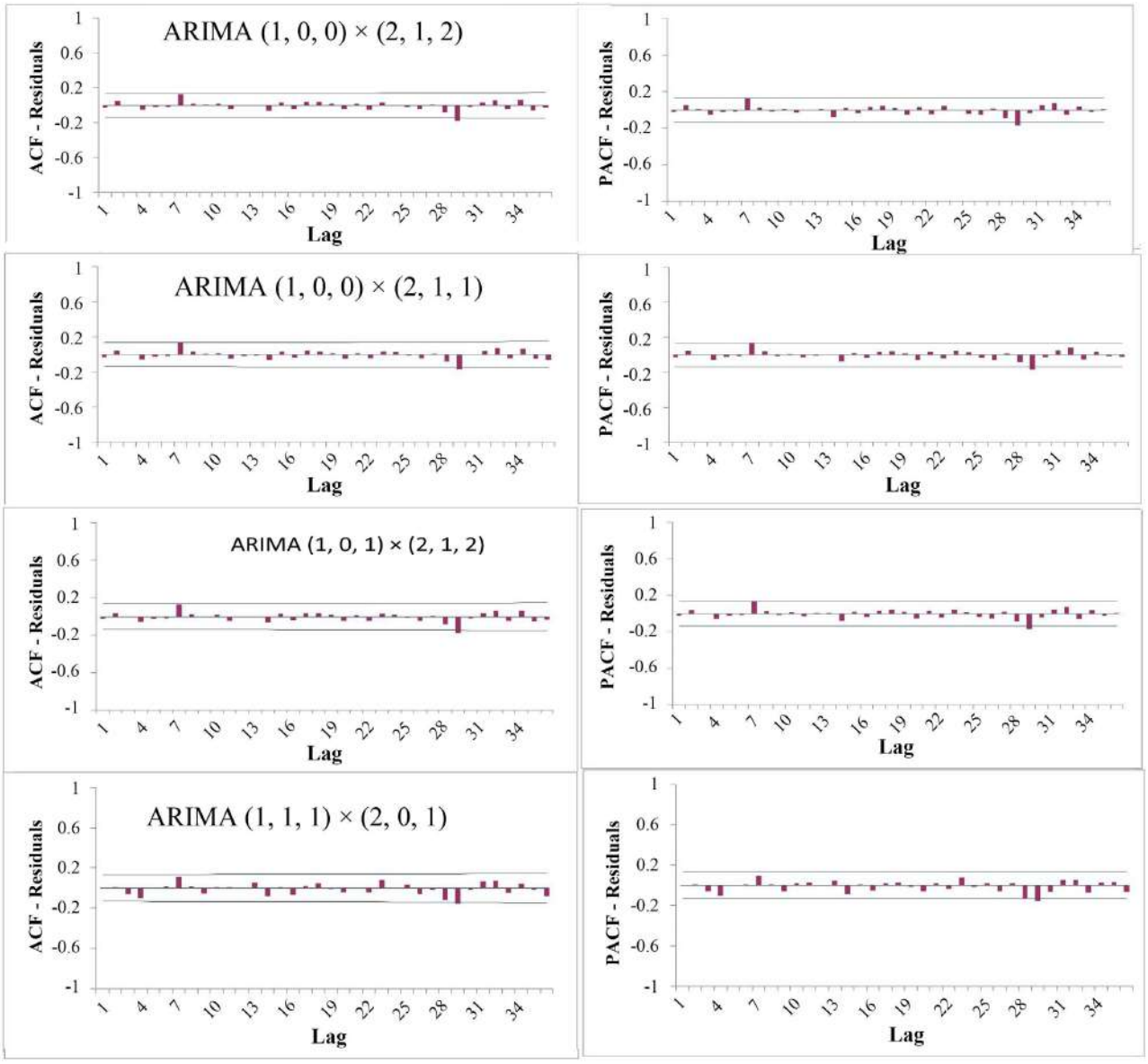

Figure 4. Assessed residual ACF and PACF using different ARIMA models.

Table 1. Possible ARIMA model combinations of parameters.

Model	RMSE	MAPE	MASE	MAE	Normalized BIC
ARIMA $(1, 0, 1) \times (2, 0, 1)_{12}$	0.0647	17.050	0.9186	0.0286	−68.429
ARIMA $(1, 0, 0) \times (2, 1, 1)_{12}$	0.0314	13.415	0.7373	0.0230	−114.175
ARIMA $(1, 0, 1) \times (2, 1, 2)_{12}$	0.0314	13.423	0.7477	0.0230	−103.797
ARIMA $(1, 0, 1) \times (2, 1, 2)_{12}$	0.0314	13.367	0.7366	0.0230	−99.135
ARIMA $(1, 0, 0) \times (2, 1, 2)_{12}$	0.0313	13.379	0.7350	0.0228	−114.380
ARIMA $(1, 0, 2) \times (2, 1, 2)_{12}$	0.0314	13.425	0.7393	0.0231	−92.7053
ARIMA $(1, 1, 1) \times (2, 0, 1)_{12}$	0.0319	13.833	0.7555	0.0236	−107.305
ARIMA $(1, 0, 0) \times (0, 1, 2)_{12}$	0.0315	13.435	0.7403	0.0231	−113.429
ARIMA $(2, 0, 2) \times (2, 1, 2)_{12}$	0.0313	13.481	0.7421	0.0231	−87.3830

Continued

ARIMA (2, 1, 1) × $(2, 0, 1)_{12}$	0.0321	13.692	0.7489	0.0234	−93.115
ARIMA (2, 1, 2) × $(2, 0, 2)_{12}$	0.0320	13.848	0.7557	0.0236	−90.7291
ARIMA (2, 0, 0) × $(2, 1, 2)_{12}$	0.0315	13.465	0.7410	0.0231	−103.291

RMSE—Root Mean Square Error; MAPE—Mean Absolute Percentage Error; MASE—Mean Absolute Scaled Error; MAE—Mean Absolute Error; and BIC—Bayesian Information Criterion.

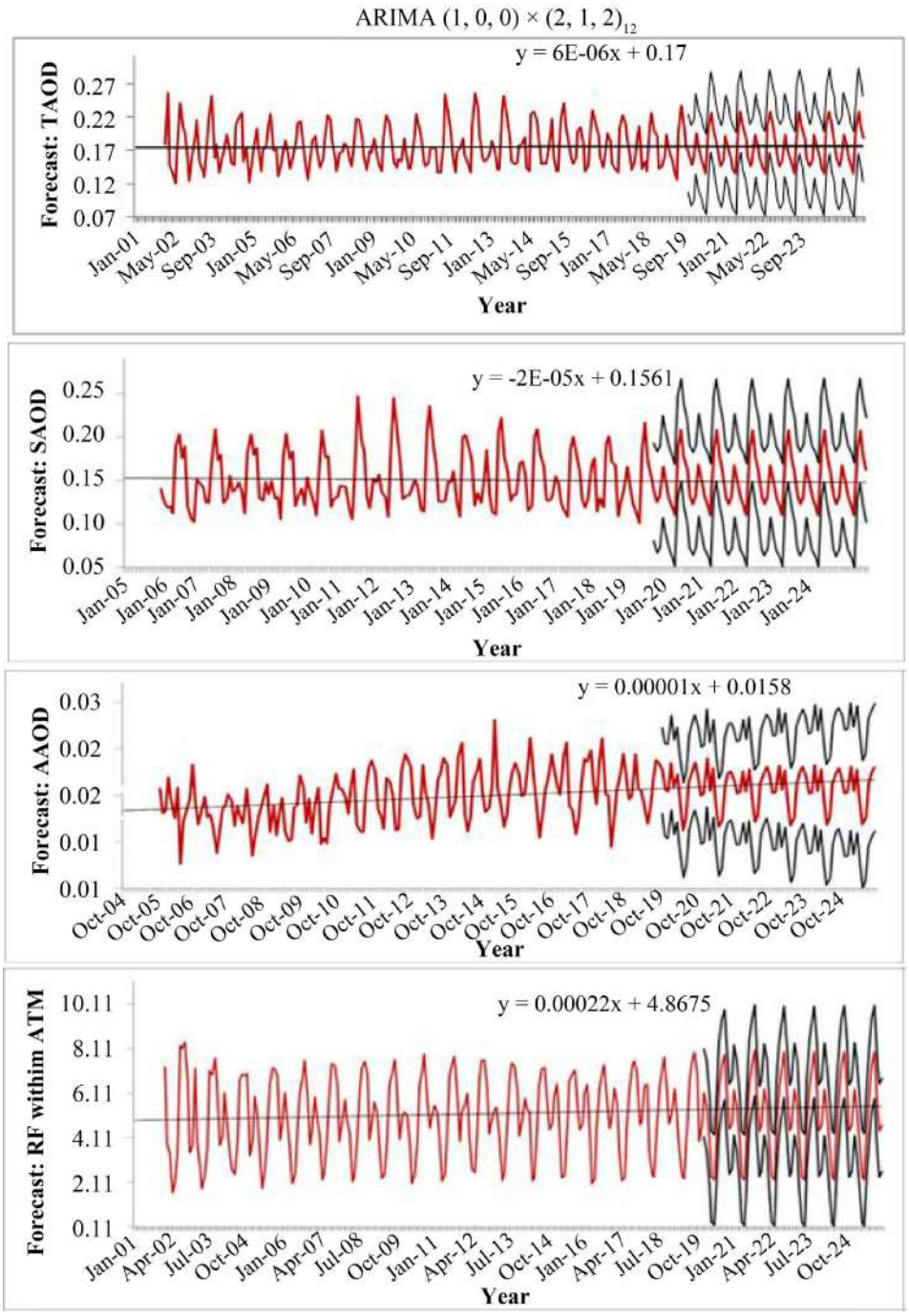

Figure 5. Forecast for the best models.

are statistically insignificant, an indication that residuals are not autocorrelated with each other. A model with the minimum value of normalized BIC, RMSE, MASE MAE and MAPE is realized to be ARIMA $(1, 0, 0) \times (2, 1, 2)_{12}$, hence considered as the best fit model which can be used to generate any required forecasts.

Lastly, is the forecasting of TAOD, SAOD, AAOD and DARF: Once the model adequacy was established, the time series was forecasted during the definite period, using the best fitting model but keeping a track on the forecasting errors and consequently re-evaluating the model. Therefore, the best model has been used for forecasting and hence evaluated for the magnitude of errors. The time series analysis of the forecasted TAOD, SAOD, AAOD and DARF using the best model are illustrated in **Figure 5**.

The trends in the forecasted series conform to the finding by previous researchers such as [3] during the spatiotemporal analysis of TAOD, SAOD as well as AAOD. As such, the probable implication of developing forecasting model for predicting the expected trends in AOD and its component in advance is that timely control measures in AOD can effectively be planned. Also, ARIMA model ascertained in this study can be applied to other fields of study such as climatology, climate change among other areas to predict the future values.

4. Conclusion and Recommendation

The study has presented the stochastic behaviour of TAOD, SAOD, AAOD and DARF over EA. Based on the forementioned parameters, a set of ARIMA models have been identified and evaluated to replicate and reproduce their time series. The time series in AOD and radiative characteristics portray clear seasonal variation with both showing a Brownian time series motion. Therefore, seasonal differencing is performed to make it stationary. The ACF and PACF of the transformed series are estimated and the significant lags are used to find out the order of the model. The statistical parameters (RMSE, MAE, MAPE, MASE and normalized BIC) are estimated for testing the validity of ARIMA models so formulated. Based on the model parameters it is found that ARIMA $(1, 0, 0) \times (2, 1, 2)_{12}$ model is the best fitting model hence adequate for forecasting aerosol parameters. Also, after finding the mean percentage error (MPE) between the actual and forecasted values of the models shows the least error is realized by ARIMA $(1, 0, 0) \times (2, 1, 2)_{12}$, indicating the suitability of this model as well for the forecasting purpose. Forecasts at a p-value of 0.05 have also been carried out based on the best fitting model till the year 2025 and are clear from the results that the past influences the future values of parameters. The best model ascertained is herein recommended to be applied to other fields of study such as climatology, and climate change among other areas, to predict future values.

Conflicts of Interest

The authors declare no conflicts of interest.

References

[1] Charlson, R.J., Schwartz, S.E., Hales, J.M., Cess, D., Coakley, J.A., Hansen, J.E. and Hofmann, D.J. (1992) Climate Forcing by Anthropogenic Aerosols. *Science*, **255**, 423-430. https://doi.org/10.1126/science.255.5043.423

[2] Satheesh, S.K, Ramanathan, V., Holben B.N., Moorthy, K.K., Loeb, N.G., Maring, H., Prospero, J.M. and Savoie, D. (2002) Chemical, Microphysical, and Radiative Effects of Indian Ocean Aerosols. *Journal of Geophysical Research*: *Atmospheres*, **107**, AAC 20-1-AAC 20-13. https://doi.org/10.1029/2002JD002463

[3] Khamala, G.W., Makokha, J.W., Boiyo, R. and Kumar, R.K. (2022) Long-Term Climatology and Spatial Trends of Absorption, Scattering and Total Aerosol Optical Depths over East Africa during 2001-2019. *Environmental Science and Pollution Research*, **29**, 61283-61297. https://doi.org/10.1007/s11356-022-20022-6

[4] Twomey, S.A., Piepgrass, M. and Wolfe, T.L. (1984) An Assessment of the Impact of Pollution on the Global Cloud Albedo. *Tellus B*: *Chemical and Physical Meteorology*, **36**, 356-366.https://doi.org/10.1111/j.1600-0889.1984.tb00254.x

[5] Ramanathan, V., Crutzen, P.J., Kiehl, J.T. and Rosenfeld, D. (2001) Aerosols, Climate, and the Hydrological Cycle. *Science*, **294**, 2119-2124. https://doi.org/10.1126/science.1064034

[6] Holben, B.N., Eck, T.F., Slutsker, I., Tanré, D., Buis, J.P., Setzer, A., Vermote, E., Reagan, J.A., Kaufman, Y.J., Nakajima, T., Lavenu, F., Jankowiak, I. and Smirnov, A. (1998) AERONET—A Federated Instrument Network and Data Archive for Aerosol Characterization. *Remote Sensing of Environment*, **66**, 1-16. https://doi.org/10.1016/S0034-4257(98)00031-5

[7] Amiridis, V., Balis, D.S., Kazadzis, S., Bais, A., Giannakaki, E., Papayannis, A. and Zerefos, C. (2005) Four-Year Aerosol Observations with a Raman Lidar at Thessaloniki, Greece, in the Framework of European Aerosol Research Lidar Network (EARLINET). *Journal of Geophysical Research*: *Atmospheres*, **110**, D21203. https://doi.org/10.1029/2005JD006190

[8] Che, H., Zhang, X.Y., Chen H.B., Damiri, B., Goloub, P., Li, Z.Q., *et al.* (2009) Instrument Calibration and Aerosol Optical Depth Validation of the China Aerosol Remote Sensing Network. *Journal of Geophysical Research*: *Atmospheres*, **114**, D03206. https://doi.org/10.1029/2008JD011030

[9] Remer, L.A., Kaufman, Y.J., Tanre, D., Matto, S., Chu, D.A., Martins, J.V., *et al.* (2005) The MODIS Aerosol Algorithm, Products, and Validation. *Journal of the Atmospheric Sciences*, **62**, 947-973. https://doi.org/10.1175/JAS3385.1

[10] Rienecker, M.M., Suarez, J.M., Gelaro, R., Todling, R., Bacmeister, J., Liu, E., Bosilovich, G.M., Schubert, D.S., Takacs, L., Kim, G., Bloom, S., Chen, J., Collins, D., Conaty, A., Dasilva, A., Gu, W., Joiner, J., Koster, R.D., Lucchesi, R., *et al.* (2011) MERRA: NASA's Modern-Era Retrospective Analysis for Research and Applications. *Journal of Climate*, **24**, 3624-3648. https://doi.org/10.1175/JCLI-D-11-00015.1

[11] Liu, Z., Liu, Q., Lin, H.C., Schwartz, C.S., Lee, Y.H. and Wang, T. (2011) Three-Dimensional Variational Assimilation of MODIS Aerosol Optical Depth: Implementation and Application to a Dust Storm over East Asia. *Journal of Geophysical Research*: *Atmospheres*, **116**, D23206. https://doi.org/10.1029/2011JD016159

[12] Liu, D. and Li, L. (2015) Application Study of Comprehensive Forecasting Model Based on Entropy Weighting Method on Trend of PM2.5 Concentration in Guangzhou, China. *International Journal of Environmental Research and Public Health*, **12**, 7085-7099. https://doi.org/10.3390/ijerph120607085

[13] Taneja, K., Ahmad, S., Ahmad, K. and Attri, S.D. (2016) Time Series Analysis of

Aerosol Optical Depth over New Delhi Using Boxe-Jenkins ARIMA Modeling Approach. *Atmospheric Pollution Research*, **7**, 585-596. https://doi.org/10.1016/j.apr.2016.02.004

[14] Ahmad, S., Khan, I.H. and Parida, B.P. (2002) Performance of Stochastic Approaches for Forecasting River Water Quality. *Water Research*, **35**, 4261-4266. https://doi.org/10.1016/S0043-1354(01)00167-1

[15] Wang, Z.F., Li, J., Wang, Z., Yang, W.Y., Tang, X., Ge, B.Z., Yan, P.Z., Zhu, L.L., Chen, X.S., *et al*. (2014) Modeling Study of Regional Severe Hazes over Mid-Eastern China in January 2013 and Its Implications on Pollution Prevention and Control. *Science China Earth Sciences*, **57**, 3-13. https://doi.org/10.1007/s11430-013-4793-0

[16] Cadenas, E. and Rivera, W. (2010) Wind Speed Forecasting in Three Different Regions of Mexico, Using a Hybrid ARIMA-ANN Model. *Renewable Energy*, **35**, 2732-2738. https://doi.org/10.1016/j.renene.2010.04.022

[17] Kripalani, R.H. and Kulkarni, A. (2001). Monsoon Rainfall Variations and Teleconnections over South and East Asia. *International Journal of Climatology*, **21**, 603-616. https://doi.org/10.1002/joc.625

[18] Soltani, S., Modarres, R. and Eslamian, S.S. (2007) The Use of Time Series Modeling for the Determination of Rainfall Climates of Iran. *International Journal of Climatology*, **27**, 819-829. https://doi.org/10.1002/joc.1427

[19] Liang, W.-M., Wei, H.-Y. and Kuo, H.-W. (2009) Association between Daily Mortality from Respiratory and Cardiovascular Diseases and Air Pollution in Taiwan. *Environmental Research*, **109**, 51-58. https://doi.org/10.1016/j.envres.2008.10.002

[20] Chattopadhyay, G. and Chattopadhyay, S. (2009) Autoregressive Forecast of Monthly Total Ozone Concentration: A Neurocomputing Approach. *Computers & Geosciences*, **35**, 1925-1932. https://doi.org/10.1016/j.cageo.2008.11.007

[21] Soni, K., Kapoor, S., Parmar, K.S. and Kaskaoutis, D.G. (2014) Statistical Analysis of Aerosols over the Gangetic-Himalayan Region Using ARIMA Model Based on Long-Term MODIS Observations. *Atmospheric Research*, **149**, 174-192. https://doi.org/10.1016/j.atmosres.2014.05.025

[22] Jere, S. and Moyo, E. (2016) Modelling Epidemiological Data Using Box-Jenkins Procedure. *Open Journal of Statistics*, **6**, 295-302. https://doi.org/10.4236/ojs.2016.62025

[23] Zhang, G.P. (2003) Time Series Forecasting Using a Hybrid ARIMA and Neural Network Model. *Neurocomputing*, **50**, 159-175. https://doi.org/10.1016/S0925-2312(01)00702-0

[24] Khan, M.F. and Gupta, R. (2020) ARIMA and NAR Based Prediction Model for Time Series Analysis of COVID-19 Cases in India. *Journal of Safety Science and Resilience*, **1**, 12-18. https://doi.org/10.1016/j.jnlssr.2020.06.007

[25] Bhatnagar, S., Lal, V., Gupta, S.D. and Gupta, O.P. (2021) Forecasting Incidence of Dengue in Rajasthan, Using Time Series Analyses. *Indian Journal of Public Health*, **56**, 281-285. https://doi.org/10.4103/0019-557X.106415

[26] Box, G.E.P., Jenkins, G.M. and Reinsel, G.C. (1994) Time Series Analysis Forecasting and Control. 3rd Edition, Prentice-Hall, Englewood Cliffs.

[27] Schwarz, G. (1978) Estimating the Dimension of a Model. *The Annals of Statistics*, **6**, 461-464. https://doi.org/10.1214/aos/1176344136

[28] Elsayir, H.A. (2019) Residual Analysis for Auto-Correlated Econometric Model. *Open Journal of Statistics*, **9**, 48-61. https://doi.org/10.4236/ojs.2019.91005