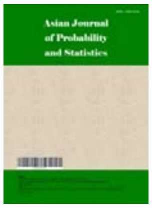

Asian Journal of Probability and Statistics

12(3): 37-45, 2021; Article no.AJPAS.66808
ISSN: 2582-0230

Sarima Modeling of Monthly Temperature in the Northern part of Ghana

Emmanuel Ayitey[1*], Justice Kangah[2] and Frank B. K. Twenefour[3]

[1]Department of Mathematics and ICT, Shama Senior High School, P.O. Box 30, Shama, Western Region, Ghana.
[2]Department of Mathematics and Statistics, Adiembra Senior High School, P.O. Box 16, Sekondi, Western Region, Ghana.
[3]Department of Mathematics, Statistics and Actuarial Science, Takoradi Technical University, P.O.Box 256, Takoradi, Ghana.

Authors' contributions

This work was carried out in collaboration among all authors. Authors JK, EA and FBKT designed the study, performed the statistical analysis and wrote the protocol and the first draft of the manuscript. All authors read and approved the final manuscript.

Article Information

DOI: 10.9734/AJPAS/2021/v12i330287
Editor(s):
(1) Dr. Dariusz Jacek Jakóbczak, Koszalin University of Technology, Poland.
Reviewers:
(1) Mu Mu Han, Myanmar Maritime University, Myanmar.
(2) Magda M.M. Haggag, Damanhour University, Egypt.
Complete Peer review History: http://www.sdiarticle4.com/review-history/66808

Review Article

Received 29 January 2021
Accepted 01 April 2021
Published 24 April 2021

Abstract

The Sarima model is used in this study to forecast the monthly temperature in Ghana's northern region. The researchers used temperature data from January 1990 to December 2020. The temperature data was found to be stationary using the Augmented Dickey Fuller (ADF) test. The ACF and PACF plots proposed six SARIMA models: SARIMA (1,0,0) (1,0,0) (12), SARIMA (2,0,0) (1,0,0) (12), SARIMA (1,0,1) (1,0,0) (12), SARIMA (0,0,1) (1,0,0) (12), SARIMA (0,0,1) (0,0,1) (12), SARIMA (0,0,1) (0,0,1) (12). The best model was chosen based on the lowest Akaike Information Criteria (AICs) and Bayesian Information Criteria (BIC) values. The Ljung-Box data, among others, were used to determine the model's quality. All diagnostic tests are passed by the SARIMA (1,0,0) (1,0,0) (12) model. As a result, the SARIMA (1,0,0) (1,0,0) (12) is the best-fitting model for predicting monthly temperatures in Ghana's northern region.

**Corresponding author: Email: corelaw33@gmail.com;*

Keywords: SARIMA; ADF; KPSS; AC; Box-jenkins; weather report northern region; PACF.

1 Introduction

During the dry seasons, the northern region of Ghana has one of the highest levels of sunlight in the country. The climate in the Northern Region is tropical. The Northern Region receives significantly less rainfall in the winter than in the summer [1]. The climate in Tamale is tropical. In Tamale, the summers are much rainier than the winters. Köppen and Geiger [2] classify this location as Aw. The average annual temperature in Tamale is 28.4 °C | 83.2 °F. The annual rainfall is around 893 mm | 35.2 inch [3].

Due to the African monsoon, Ghana's climate is tropical, with a dry season in winter and a rainy season in summer. The rainy season in the north lasts from May to September, in the center from April to October, and in the south from April to November [4]. The rainy season is shorter along the east coast, lasting from April to June, with a break in July and August and a minor recovery in September and October. [5]. The wettest areas are in the south, where annual precipitation exceeds 1,500 millimeters (60 inches), and even more so on the small west coast, where annual precipitation exceeds 2,000 millimeters (80 inches). The driest regions are the north, where rainfall averages around 1,000 mm (40 in) per year, and the eastern coast, including Accra, where rainfall averages less than 800 mm (31.5 in) [6]. Anyway, as previously stated, there is only one rainy season in the north, which peaks during the summer months, while the rainy season on the coast is split into two.

Winter in the center and north is hot: daytime temperatures in December and January are usually around 35 °C (95 °F), despite the dry air and cool nights [7]. However, nights in the center-north can get a little chilly at times, with lows of around 10 °C (50 °F). Harmattan, a dry, dusty wind, frequently blows from the desert.

Then there's the monsoon season. Early showers and thunderstorms may occur in March in the center and April in the north, usually in the afternoon or evening, and become more significant the following month, reaching 100 mm (4 in) per month (ie April in the center and May in the north) [8]. The temperature gradually drops as a result of thunderstorms and clouds brought by ocean currents, and from July to September, when clouds and rains are more common, it drops to around 30/31 °C (86/88 °F), but air humidity rises [9]. The rains stop in October in the north and November in the center, and hot, sunny weather returns with the dry wind from the north. Rainfall totals 1,000 millimeters (40 inches) in the north and 1,200/1,300 millimeters (47/51 inches) in the center, where the rainy season is longer [10].

2 Methodology

2.1 Source of Data

For this study, monthly temperature data from the northern part of Ghana was gathered from 1990 to 2020, and the Sarima Model was used to assess the rate of temperature change in the region.

2.2 Stationary Time Series

Before the analysis can be carried out, the time series must be stationary (i.e., the mean and variance must be constant as a function of time). This is an important condition for ARIMA models and potentially their weak point [14]. We used the Augmented Dickey-Fuller test, Phillip-Perron (P-P) test, and Kwiatkowski-Phillips-Schmidt-Shin (KPSS) test to determine if the time series was stationary. This test's p-value should be compared to the significance level [19]. The null hypothesis that the time series is stationary will be dismissed if the p-value is too small [15].

2.3 Identification and Evaluation of the Model

To find the right SARIMA model for a time series, start by figuring out what order of differencing you'll need to get a stationary series and get rid of the main seasonality. If you stop after the first difference and expect that the differenced series will remain constant, you've just fitted a random walk or random trend model [16]. The autocorrelation and partial autocorrelation functions are graphs that contain correlations with various time lags [11]. The ACF and PACF formulas can be used to calculate whether a series is stationary or not, as well as the number of components in an ARMA model. The number of significant spikes in the ACF corresponds to the number of MA terms in the model, while the number of significant spikes in the PACF corresponds to the number of AR terms [12, 20].

3 Results and Discussion

Table 1 Shows the descriptive statistics for the temperature series. The temperature's mean and standard deviation were 34.5 and 3.3, respectively, while the minimum and maximum values were 25.6 and 44.2

Table 1. Descriptive statistics for the temperature data (Jan 1960-Dec 2020)

Variable	Mean	Std.Dev	Minimum	Maximum	Skewness	Kurtosis
Temperature	34.5	3.3	25.6	44.2	-0.23	-0.47

The behavior of time series is studied using monthly temperature data from Ghana's Northern Region.
Fig. 1 depicts the monthly temperature plot, which is extremely fluctuating.

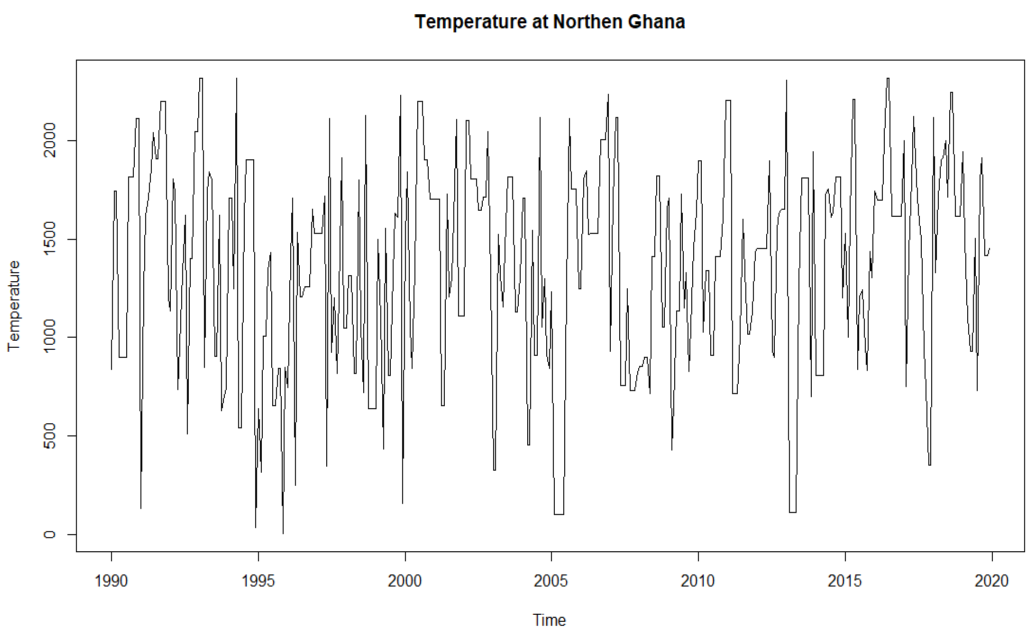

Fig. 1. Time series plot of monthly temperature

3.1 SARIMA (Box-Jenkins) Modeling

Monthly temperatures in the Northern part of Ghana from January 1960 to December 2020 are included in the data. Training data was used in the Box-Jenkins modeling [13]. Before estimating and developing a model, the data must be tested to see if it is stationary using the KPSS (Kwiatkowski-Phillips-Schmidt-Shin) test. Table 2 shows the corresponding results.

Table 2. Result of unit test for temperature

KPSS test	Null Hypothesis: Temp is not stationary
ADF test	Null Hypothesis: Temp is not stationary
KPSS test statistic	0.025
ADF test statistic	0.031

The null hypothesis that the series is not stationary can be refuted, as shown in Table 2, since the test statistic of the KPSS test is less than the critical values at the 5% significance level. As a result, the temperature series has reached a point of equilibrium.

Because the stationary has been reached at the level, models from the ARIMA (p, 0, q) family must be searched, where p and q are the possible order of AR and MA components. The plot of sample ACF and PACF of temperature series (Fig. 2 and Fig. 3) was considered for identification of suitable values for p and q [17].

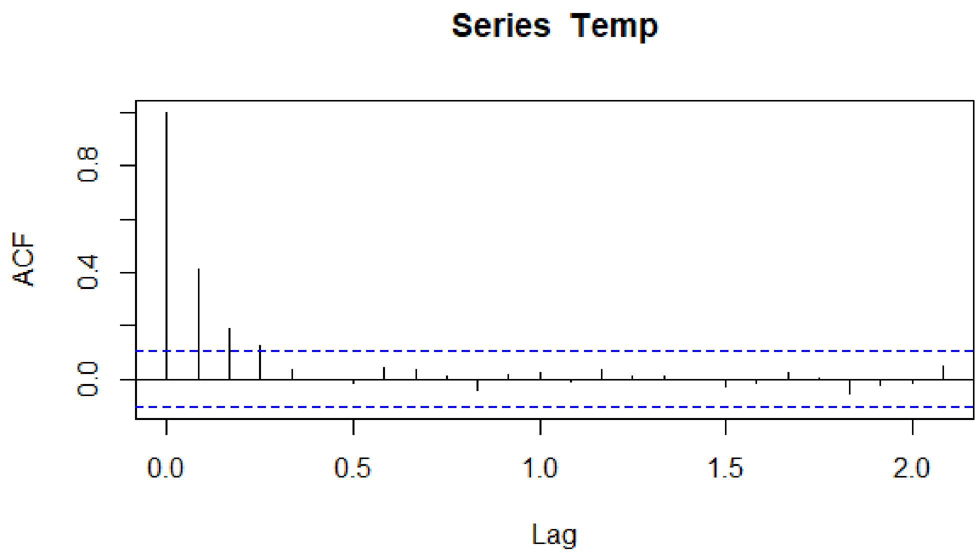

Fig. 2. Sample of a plot of ACF temperature series

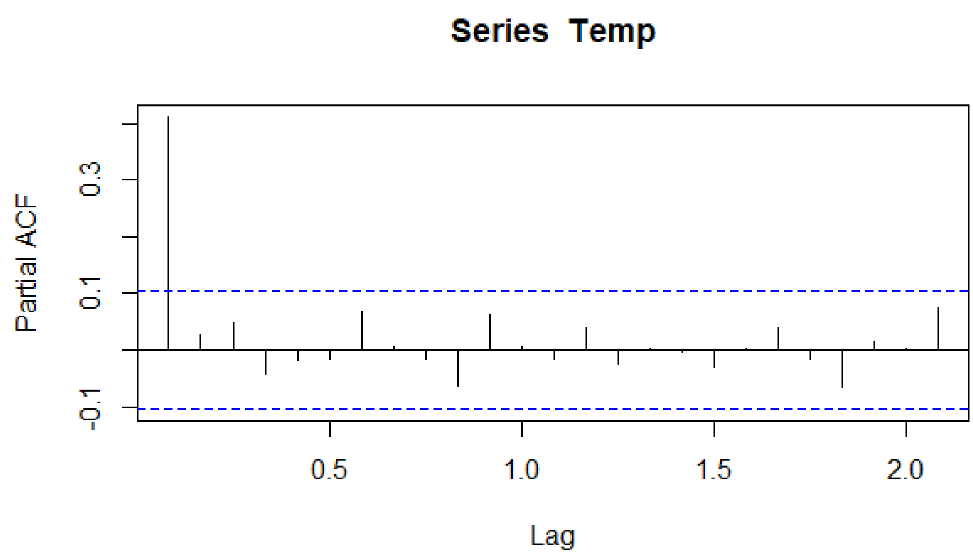

Fig. 3. Sample of a plot of PACF temperature series

Fig. 3 shows that at lag 1, the sample PACF has one major autocorrelation. As a result, it is possible to hypothesize that the AR order in the ARMA model will be 1. Due to the seasonality of the data, the following models were considered as possible models to represent the original series. They are: (i) SARIMA (1,0,0) (1,0,0) (12), (ii) SARIMA (2,0,0) (1,0,0) (12), (iii) SARIMA (1,0,1) (1,0,0) (12), (iv) SARIMA (0,0,1) (1,0,0) (12), (v) SARIMA (0,0,1) (0,0,1) (12) and (vi) SARIMA (0,0,2) (0,0,1) (12).

Table 3. Results of model estimation

Model	Log likelihood	AIC
SARIMA(1,0,0)(1,0,0)(12)	-2745.09	5498.19
SARIMA(2,0,0) (1,0,0) (12)	-2744.94	5499.89
SARIMA(1,0,1) (1,0,0) (12)	-2749.86	5507.72
SARIMA(0,0,1) (1,0,0) (12)	-2749.86	5507.72
SARIMA(0,0,1) (0,0,1) (12)	-2749.85	5507.71
SARIMA(0,0,2) (0,0,1) (12)	-2746.67	5503.94

Table 3 shows that the SARIMA (1,0,0) (1,0,0) (12) model had the highest log likelihood estimates and the lowest AIC values of the six models. As a result, SARIMA (1,0,0) (1,0,0) (1,0,0) (1,0,0) (1,0,0) (1,0,0) (1,0,0) (1,0,0) (1,0,0) (1 (12). This model was confirmed with auto. arima function in R studio software in addition, Table 4 indicates that constant and AR (1) terms are significant in the selected model

Table 4. Parameter estimation of SARIMA (1,0,0) (1,0,0) (12), alpha=0.05

Parameter	Coefficient	Standard error	Confidence interval
Intercept(C)	1358.04	45.1636	1269.52-1446.56
AR(1)	0.4101	0.0480	0.31601-0.50416
SAR(1)	0.0219	0.0534	-0.0826-0.12651

3.2 Model diagnostics

3.2.1 Randomness

The residuals of the chosen model's ACF plot (Fig. 4) shows that the residuals are small and not statistically significant. At the 5% significance level, the p-value derived from Ljung –box Statistics (0.03) is also less than critical values. As a result, the model fit can be considered satisfactory [18].

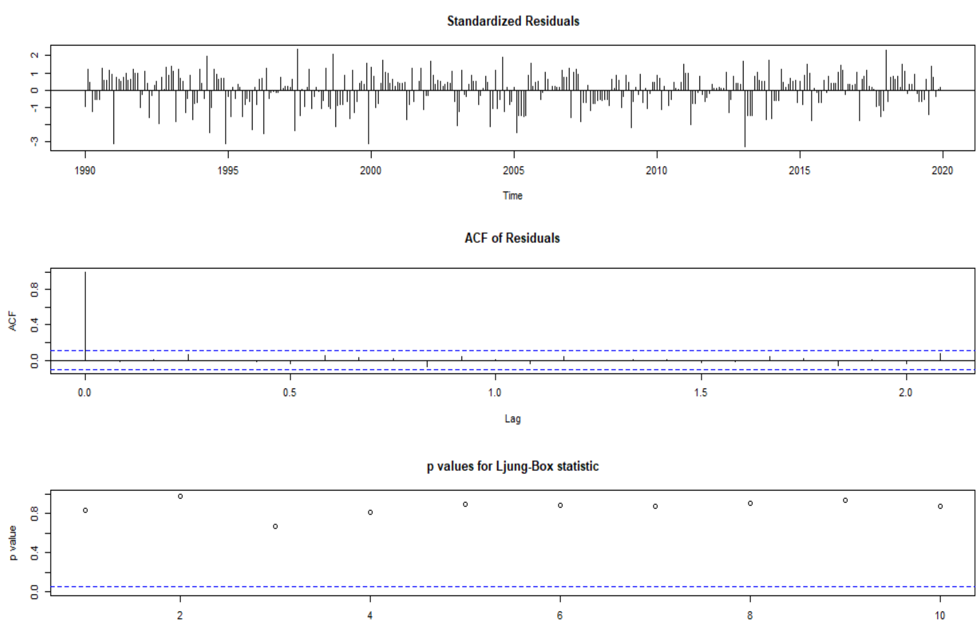

Fig. 4. ACF plot of Residual and p-values for Ljung-Box Statistics

3.2.2 Normality

The normal probability plot of the residuals was carried out to check whether residuals are normal or not.

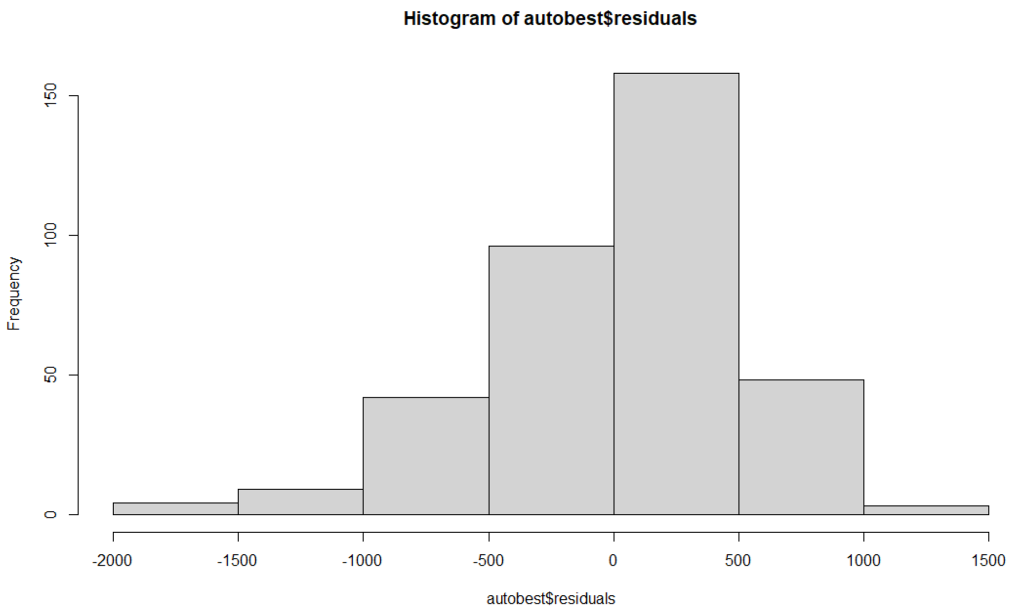

Fig. 5. Normality test using histogram

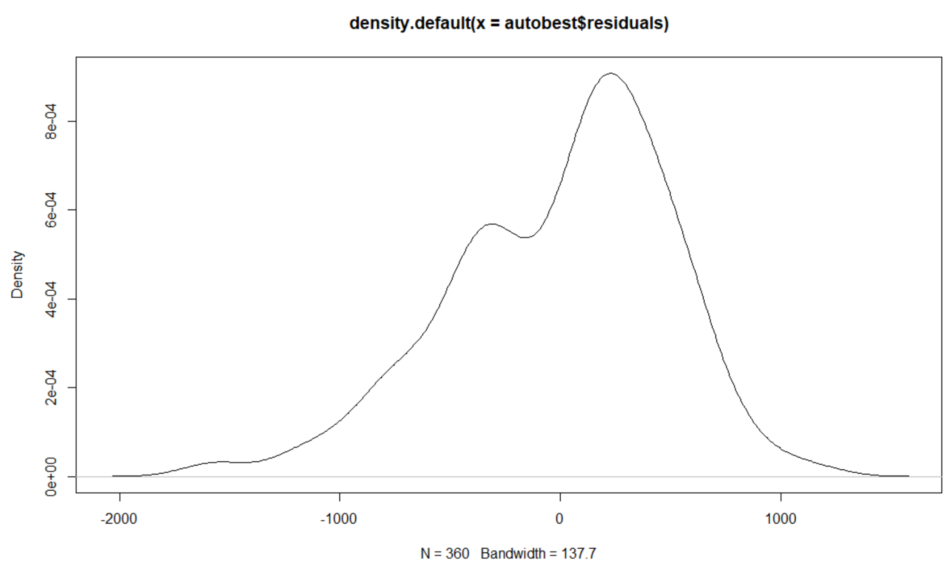

Fig. 6. Normality test using density plot

Fig. 5 and 6, show that the residuals are normally distributed and also from Jarque-Bera (p=2.125e-5) and Shapiro-wilk normality (3.605e-6) test with H_o=Data is not normal ,also confirmed that the residual series is normally distributed indicating that the model is good fit.

3.2.3 Heteroscedasticity

The ARCH LM test was conducted to observe the heteroscedasticity of the developed model. The results of heteroscedasticity of residuals are shown in Table 5 and it shows that no heteroscedasticity exist at 5% significance level

Table 5. ARCH LM test

H_o: No ARCH Effects	
p-value	0.07199

Based on the above detailed analysis of residuals, it can be confirmed that the developed SARIMA (1,0,0) (1,0,0) (12) model satisfies all the diagnostic tests. Hence, the SARIMA (1,0,0) (1,0,0) (12) is considered as the best fitted model for forecasting the monthly temperature of Northern part of Ghana

3.3 Forecasting temperature using SARIMA (1,0,0) (1,0,0) (12) model

Before forecasting the values, it is useful to validate the present model with observed data (training set) as well as an independent data set (validation set). For the validation purpose 20 weeks observed temperature data from August ,2020 to December, 2020 is used.

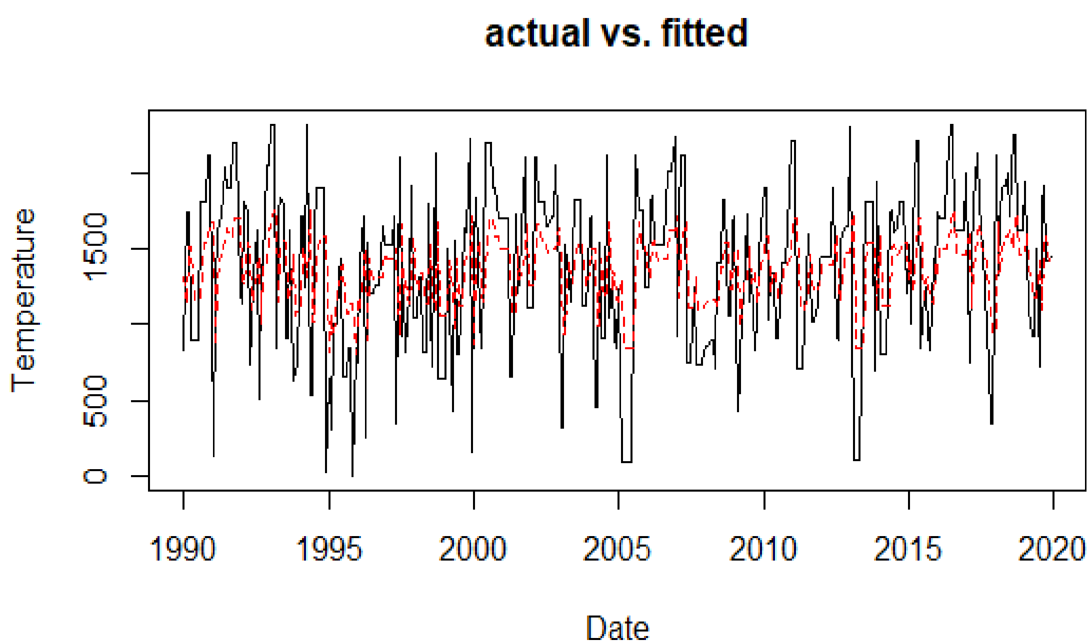

Fig. 7. Time series plot for observed and fitted values

Model Accuracy:

Table 6. Results of forecast performances of SARIMA (1,0,0) (1,0,0) (12) model

Type of data	Period	RMSE	MAE
Validation set	Jan 1960-Dec 2020	495.743	402.367

Table 6 indicates that the RMSE and MAE for validation set from SARIMA (1,0,0) (1,0,0) (12) model deviates from the observed data are 495.743 and 402.367, respectively, which would be considered as being within acceptable range. The estimated and observed values are shown in

SARIMA (1,0,0) (1,0,0) (12) Forecast:

The derived model was used forecasting and the results are shown in the following table.

Table 7. Forecast value for 6 months' temperature for 2020 and their corresponding 95% confidence limits

Year 2020	Point forecast	95%LB	95%UB
Jan	40.469	39.756	41.182
Feb	37.176	36.607	38.745
Mar	35.551	34.293	36.808
Apr	35.192	34.815	36.568
May	34.702	33.969	35.434
June	36.689	35.866	36.481

4 Conclusion

In temperature modeling and forecasting, time series analysis is an important method. The SARIMA model was used to predict monthly temperatures in Ghana's northern region in this study. SARIMA (1,0,0) (1,0,0) (1,0,0) (1,0,0) (1,0,0) (1,0,0) (1,0,0) (1,0,0) (1,0, (12) Nonetheless, diagnostic tests were used to validate the developed model. As a result, the developed model will assist in determining future agriculture and tourism strategies in the respective region. However, due to fluctuations in the data series, this research needs to be extended.

Competing Interests

Authors have declared that no competing interests exist.

References

[1] Bélanger M, Gray-Donald K, O'loughlin J, Paradis G, Hanley J. Influence of weather conditions and season on physical activity in adolescents. Annals of epidemiology. 2009;19(3):180-186.

[2] Nguyen DQ, Renwick J, McGregor J. Variations of surface temperature and rainfall in Vietnam from 1971 to 2010. International Journal of Climatology. 2014;34(1):249-264.

[3] P Fister C. 6 Monthly temperature and precipitation in central Europe 1525-1979: Quantifying documentary evidence on weather and its. Climate Since AD. 1995;1500:118.

[4] Abdul-Razak M, Kruse S. The adaptive capacity of smallholder farmers to climate change in the Northern Region of Ghana. Climate Risk Management. 2017;17:104-122.

[5] Stokke OS, Hønneland G. (Eds.). International cooperation and Arctic governance: regime effectiveness and northern region building, Routledge. 2006;50.

[6] Nantui MF, Bruce SD, Yaw OA. Adaptive capacities of farmers to climate change adaptation strategies and their effects on rice production in the northern region of Ghana. Russian Journal of Agricultural and Socio-Economic Sciences. 2012;11(11).

[7] Larocque I, Hall RI, Grahn E. Chironomids as indicators of climate change: A 100-lake training set from a subarctic region of northern Sweden (Lapland). Journal of Paleolimnology. 2001;26(3):307-322.

[8] Mabe FN, Sarpong DB, Osei-Asare Y. Empirical evidence of climate change: Effects on rice production in the Northern Region of Ghana. Journal of Economics, Management and Trade. 2014;551-562.

[9] Etwire PM, Al-Hassan RM, Kuwornu JK, Osei-Owusu Y. Application of livelihood vulnerability index in assessing vulnerability to climate change and variability in Northern Ghana. Journal of Environment and Earth Science. 2013;3(2):157-170.

[10] Azumah SB, Donkoh SA, Ansah IGK. Contract farming and the adoption of climate change coping and adaptation strategies in the northern region of Ghana. Environment, Development and Sustainability. 2017;19(6):2275-2295.

[11] Chen TC, Murakami M. The 30–50 day variation of convective activity over the western Pacific Ocean with emphasis on the northwestern region. Monthly Weather Review. 1988;116(4):892-906.

[12] Bañón M, Justel A, Velázquez D, Quesada A. Regional weather survey on Byers Peninsula, Livingston Island, South Shetland Islands, Antarctica. Antarctic Science. 2013;25(2):146.

[13] Bridges AS, Peterson MJ, Silvy NJ, Smeins FE, Ben Wu X. Differential influence of weather on regional quail abundance in texas. The Journal of Wildlife Management. 2001;10-18.

[14] Screen JA, Simmonds I. Amplified mid-latitude planetary waves favour particular regional weather extremes. Nature Climate Change. 2014;4(8):704-709.

[15] Woodruff RE, Guest CS, Garner MG, Becker N, Lindesay J, Carvan T, Ebi K. Predicting Ross River virus epidemics from regional weather data. Epidemiology. 2002;384-393.

[16] Murat M, Malinowska I, Gos M, Krzyszczak J. Forecasting daily meteorological time series using ARIMA and regression models. International Agrophysics. 2018;32(2).

[17] Abdul-Aziz AR, Anokye M, Kwame A, Munyakazi L, Nsowah-Nuamah NNN. Modeling and forecasting rainfall pattern in Ghana as a seasonal ARIMA process: The case of Ashanti region. International Journal of Humanities and Social Science. 2013;3(3):224-233.

[18] Liu H, Li C, Shao Y, Zhang X, Zhai Z, Wang X, Jiao M. Forecast of the trend in incidence of acute hemorrhagic conjunctivitis in China from 2011–2019 using the Seasonal Autoregressive Integrated Moving Average (SARIMA) and Exponential Smoothing (ETS) models. Journal of infection and Public Health. 2020;13(2):287-294.

[19] Owusu-Sekyere E, Harris E, Bonyah E. Forecasting and planning for solid waste generation in the Kumasi metropolitan area of Ghana: An ARIMA time series approach; 2013.

[20] Lal A, Ikeda T, French N, Baker MG, Hales S. Climate variability, weather and enteric disease incidence in New Zealand: Time series analysis. PLoS One. 2013;8(12):e83484.

Peer-review history:
The peer review history for this paper can be accessed here (Please copy paste the total link in your browser address bar)
http://www.sdiarticle4.com/review-history/66808