JETIR.ORG ISSN: 2349-5162 | ESTD Year : 2014 | Monthly Issue
JOURNAL OF EMERGING TECHNOLOGIES AND INNOVATIVE RESEARCH (JETIR)
An International Scholarly Open Access, Peer-reviewed, Refereed Journal

A Systematic Approach for Minimum and Maximum Temperature Prediction Using SARIMA Model

[1]G. S. Phad and [2]A. A. Muley

[1,2]Assistant Professor,
[1]Department of Statistics, KRT Arts, BH Commerce and AM Science (KTHM) College, Nashik (M.S.), India
[2]School of Mathematical Sciences, Swami Ramanand Teerth Marathwada University, Nanded (M.S.), India
E-mail: aniket.muley@gmail.com

Abstract : This study concern with the systematic approach for Seasonal Autoregressive Integrated Moving Average (SARIMA) model fitting by dealing with the time series of monthly average minimum and maximum temperature dataset. Here, the data is collected from the Nashik meteorological centre/hydrology project of Maharashtra state in India. Initially, base model has been fitted and simulation technique is used to obtain the best model among the all others. Finally, we have obtained SARIMA (1, 1, 2)(3, 1, $1)_{12}$ model for monthly average minimum temperature and SARIMA(0, 1, 2)(2, 1, $1)_{12}$ model for monthly average maximum temperature as the optimum models for given study area. The selected optimum models have the highest R-square with lowest Root Mean Squared Error (RMSE) and Mean Absolute Percentage Error (MAPE) values. The residual analysis shows good prediction accuracy of the selected optimum models. In a nutshell, in future, these models can useful to deal with the prediction of maximum and minimum temperature so that, it will be helpful in the smart city planning development and to decide administrative policies for future agriculture.

***Keywords* -Time series, temperature, SARIMA, modelling, R-square, RMSE, MAPE, prediction.**

I. INTRODUCTION

Meteorological parameters have great influence on environmental changes and which also results in natural. Temperature is one of the most important meteorological parameters and commonly used to detect fluctuations in climate. Temperature and solar radiation affect the normal growth of the plant in agricultural domain because of their linkages to the changes in cloudiness and soil moisture. Hence forecasting the temperature accurately is important and it also helps to minimize damages.

Ho *et al* (2002) compared neural network and Box-Jenkins ARIMA model in time series prediction. Hyndman and Koehler (2006) propose that the mean absolute scaled error (MASE) become the standard measure for comparing forecast accuracy across multiple time series. Wakaura and Ogata (2007) used a Fourier form AR model on the seasonality of air temperature anomalies and this model fits substantially better than an ordinary AR model for normalized datasets of the surface air temperature obtained from almost all of the stations in Japan, and exhibits a significant seasonal structure in their auto-correlation. Alysha *et al* (2011) forecasted time series with complex seasonal patterns using exponential smoothing. Chawsheen and Broom (2017) used the SARIMA model to predict the mean temperature of the Erbil Kurdistan Region, Iraq. The pattern of mean temperatures in the Erbil Kurdistan Region from January 1992 to November 2015 was observed to be non-stationary and increasing over time. Sample ACF and PACF plots are used to verity the non-stationarity of the series. By following the procedures of the Box–Jenkins SARIMA model building the SARIMA(0, 1, 2)(0, 1, 1)12 model has been selected as best model. The model diagnostics were performed using model residuals. The model residual were found to be uncorrelated and follows white-noise process with zero mean and a stable variance. Forecasting shows the mean temperature of the Erbil, Iraq will be stable for the next 10 years. Tadesse and Dinka (2017) forecast monthly flows in Waterval River of South Africa using SARIMA model. Xin *et al* (2018) forecast the indoor air temperature and relative humidity based on cloud database by using an improved backpropagation (BP) neural networks.

Athiyarath and Krishnaswamy (2020) performed the comparative study and analysis of time series forecasting techniques. They compares various forecasting algorithmic approaches and explores their usefulness and limitations for different types of time series data in different domains. Shad *et al.* (2020) carried out study of seasonal autoregressive integrated moving average (SARIMA) and artificial neural network (ANN) with multi-layer perceptron (MLP) models to forecast the month to month relative humidity in Delhi, India during 2017-2025. The average month to month relative humidity data for the period 2000-2016 have been utilized to complete the goals of the proposed study. The prediction pattern in relative humidity decreases from 2017 to 2025. The exactness of the models has been estimated by utilizing RMSE and MAE. Louloudis *et al.* (2021) dealt with the prediction of the spatiotemporal development of water levels in a mined-out pit by generating forecasts of the dependent variables (rainfall and temperature) via linear (ARIMA) and non-linear (ANN) models. Do *et al.* (2022) examined the wastewater inflow conduct and foster an seasonal autoregressive integrated moving normal (SARIMA) estimating model at low transient goal (hourly) for a momentary time of 7 days for a genuine organization in South Australia, the Murray Bridge wastewater organization/wastewater treatment plant (WWTP). Verifiable wastewater inflow information gathered for a 32-month time span was pre-handled and afterward isolated into two sections for training (80%) and testing (20%). Results reveal that there is irregularity presence in the wastewater inflow time series information, as it is vigorously subject to time and day of the week. Additionally, the SARIMA $(1,0,3)(2,1,2)_{24}$ was seen as the best model to anticipate wastewater inflow and its estimating exactness was resolved in light of the assessment measures including the RMSE, the MAPE and the R^2.

Aghelpour and Norooz-Valashedi (2022) assessed the performance of two stochastic and AI models in forecasting evapotranspiration of Mazandaran region, which is one of the main places of rice development in Iran. They have collected the data of air temperature, relative humidity, wind speed, and sunshine duration from the Iranian Meteorological Organization during the period 2003-2018. Then, at that point, these factors and the FAO-56 Penman Monteith model are utilized to compute every day evapotranspiration rates. Additionally, stochastic models including AR, MA, ARMA, and ARIMA, and AI models including least square support vector machine (LSSVM), adaptive neuro-fuzzy inference system (ANFIS), and generalized regression neural network (GRNN) are utilized to forecast evapotranspiration. The time series models of ARMA and ARIMA, and the AI model of LSSVM furnish the most accurate prediction. Subsequently, it is observed that stochastic models are better than AI models because of their more precise prediction and less complexity. The ARMA model showed the most noteworthy prediction accuracy.

In this study, our main objective is to propose the systematic procedure for SARIMA model fitting. To obtain the best SARIMA model for monthly average minimum and maximum temperature and to predict the next three years temperature of Nashik region. In the next subsequent sections, methodology, result and discussion and conclusion are discussed in detail.

II. METHODOLOGY

In this study, we have collected secondary data. The data is collected from Nashik Meteorological Centre/Hydrology Project (State Data Storage Centre) of Maharashtra state in India. The daily minimum and maximum temperature data was available. We processed and converted daily data into month to month average minimum and maximum temperature data by taking average for each month. Monthly data of average minimum and average maximum temperature for 228-months from January 2003 to December 2021 is taken for analysis. The monthly average minimum and monthly average maximum temperature has been recorded in °C unit.

2.1 ARIMA/SARIMA MODEL

2.1.1 ARIMA model

An autoregressive integrated moving average (ARIMA) is a generalization of an autoregressive moving average (ARMA) model. ARIMA model is classified as an ARIMA (p, d, q) model, where, p is the number of autoregressive terms, d is the number of non-seasonal differences needed for stationarity and q is the number of moving average terms. The model can be written as:

$$(1-B)^d \phi(B) X_t = \theta(B) Z_t$$

Where, $\phi(B) = 1 - \phi_1 B - \cdots - \phi_p B^p$ is termed as autoregressive polynomial,

$\theta(B) = 1 + \theta_1 B + \cdots + \theta_q B^q$ is termed as moving average polynomial, and

Z_t is the white noise process.

2.1.2 SARIMA Model

Seasonal Autoregressive Integrated Moving Average (SARIMA or Seasonal ARIMA) is an extension of ARIMA that explicitly supports univariate time series data with a seasonal component. There are four seasonal elements such as:

P: Seasonal autoregressive order,

D: Seasonal difference order,

Q: Seasonal moving average order and

S: The number of time steps for a single seasonal period.

The general SARIMA (p, d, q)(P, D, Q)s model can be written as:

$$\phi(B)\Phi(B^S)\nabla^d \nabla_S^D X_t = \theta(B)\Theta(B^S) Z_t$$

Where, $\phi(B) = 1 - \phi_1 B - \cdots - \phi_p B^p$: Non-stationary AR operator,

$\theta(B) = 1 + \theta_1 B + \cdots + \theta_q B^q$: Non-stationary MA operator,

$\Phi(B^s) = 1 - \Phi_1 B^s - \cdots - \Phi_P B^{PS}$: Stationary AR operator,

$\Theta(B^s) = 1 + \Theta_1 B^s + \cdots + \Theta_Q B^{QS}$: Stationary MA operator,

$\nabla^d = (1-B)^d$: Non-seasonal differencing;

$\nabla_S^D = (1-B^s)^D$: Seasonal differencing

$B^S X_t = X_{t-s}$;

Zt is the white noise process.

2.2 Accuracy measures

It is essential to deal with the accuracy, we used the following measure (Table 1) and with the help of that the efficiency of the model is defined. Let, $e_t = X_t - \hat{X}_t$ is the error, where X_t is the actual value and $\hat{X}_t$ is the forecasted value.

Table 1: Accuracy Measures

Error	Short Form	Formula
Root Mean Squared Error	RMSE	$\sqrt{\frac{\sum_{t=1}^{n} e_t^2}{n-1}}$
Mean Absolute Percentage Error	MAPE	$\frac{\sum_{t=1}^{n} \left\| \frac{e_t}{X_t} \right\|}{n}$

Fig. 1 represents the flowchart for fitting the SARIMA model and data analysis workflow of our study.

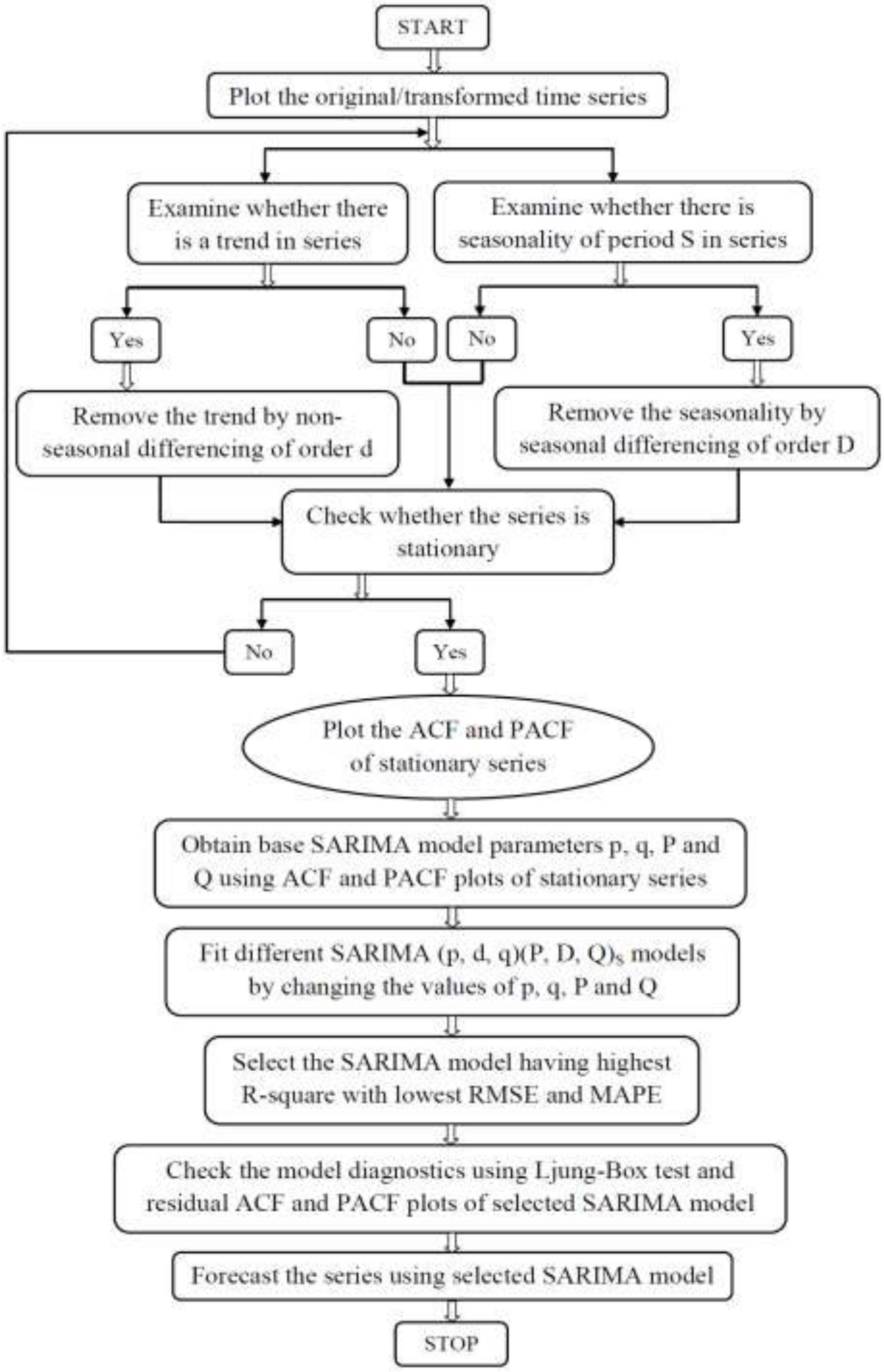


Fig. 1: Flowchart for fitting the SARIMA model

III. RESULTS AND DISCUSSION

In this study, we have studied monthly time series dataset. Initially, it is essential to identify the nature of the data so that, we have performed decomposition of original data into trend, seasonal and irregular component as shown in Fig. 2. The monthly average minimum temperature is falling to the least in December/January during winter season and rising to attain a peak in May/June during summer season. While the monthly average maximum temperature is rising to attain a peak in April/May during summer season, falling down in July/August during rainy season, and again it shows some increment in October (it is due to Indian weather October heat which is period of transition from hot rainy season to dry winter) and attain lowest in December/January during winter season. Fig. 2 (c-d) indicates there is slight upward trend in both the series. Fig. 2(e-f) clearly shows the seasonality of period of 12 months in both the series. The original time series of monthly average minimum temperature and monthly average maximum temperature are shown in Fig. 3(a-b) respectively.

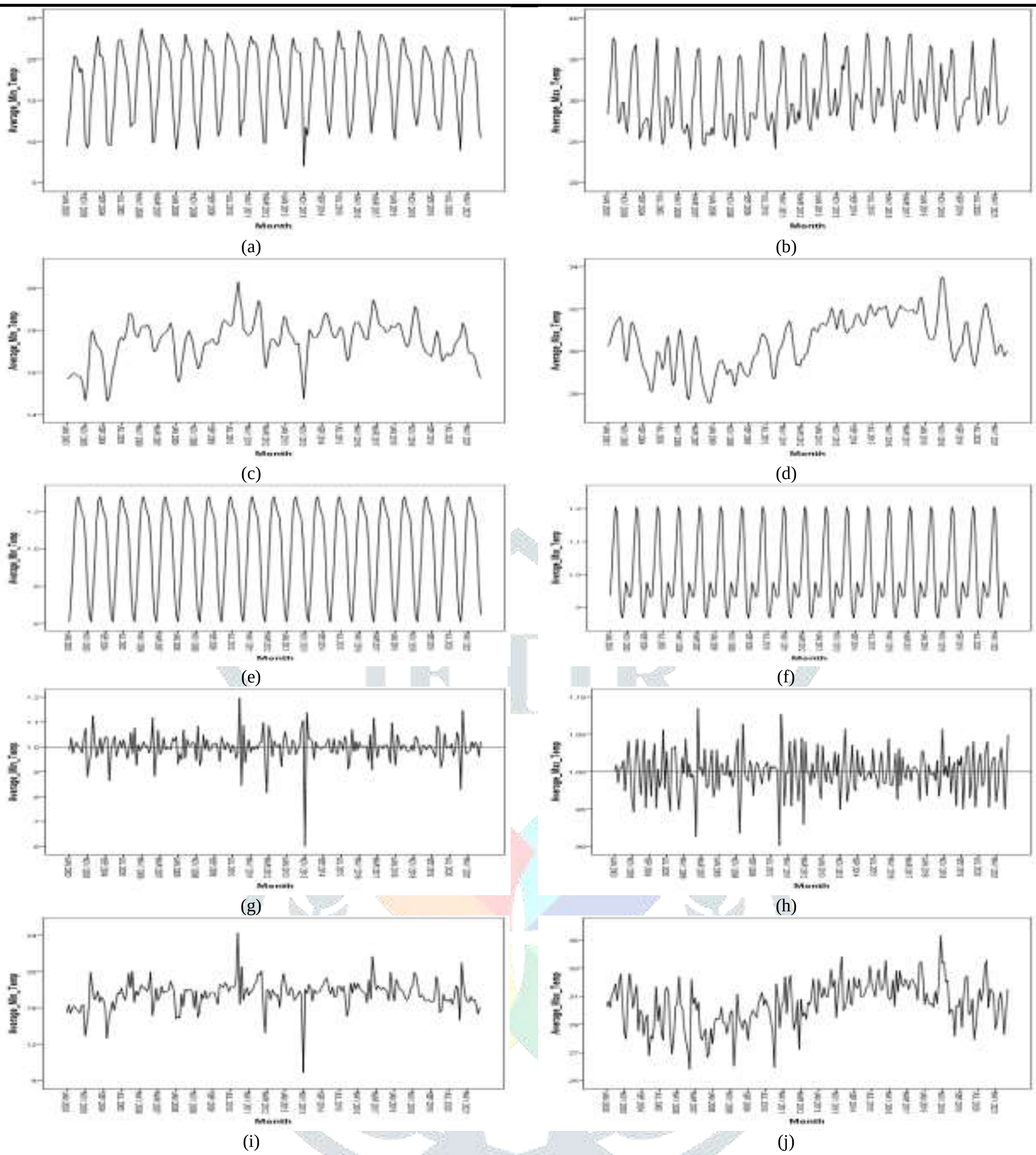
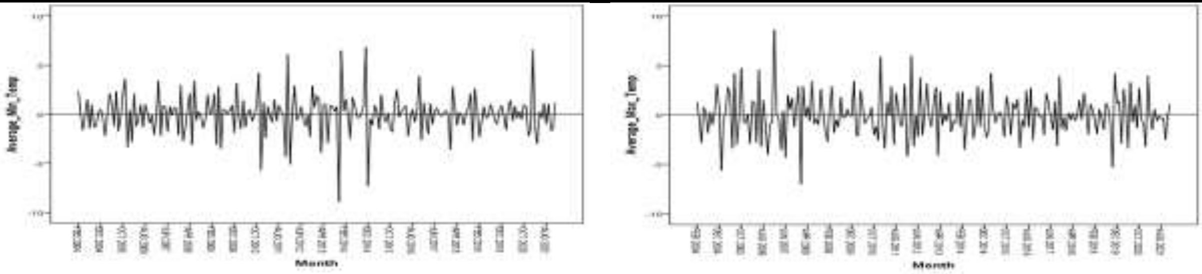


Fig. 2: Time-series plots and decomposed plots of monthly average minimum and maximum temperature: (a) time-series plot of monthly average minimum temperature (b) time-series plot of monthly average maximum temperature (c) plot of trend cycle component for monthly average minimum temperature (d) plot of trend cycle component for monthly average maximum temperature (e) plot of seasonal indices for monthly average minimum temperature (f) plot of seasonal indices for monthly average maximum temperature (g) plot of irregular component for monthly average minimum temperature (h) plot of irregular component for monthly average maximum temperature (i) plot of seasonal adjusted series for monthly average minimum temperature (j) plot of seasonal adjusted series for monthly average maximum temperature.

3.1 Fitting of SARIMA model

The development of model begins by studying the ACF and PACF plots for original series of monthly average minimum temperature and monthly average maximum temperature Fig. 4(a-c) represents the ACF and PACF plots of monthly average minimum temperature. Fig. 4b, Fig. 4d represents monthly average maximum temperature shows seasonal variation with decay of the spikes in the form of sinusoidal wave. Fig. 4(a-d) strongly indicates the seasonality of period 12 in both the series. The next step is to convert the original time series into stationary series by removing the trend and seasonality. To remove the seasonal component from both the series, one seasonal difference of order 12 is taken called differencing at lag 12 and then to eliminate the remaining trend from the series one regular difference of order 1 is taken called differencing at lag 1 (Fig. 3).



(a) (b)

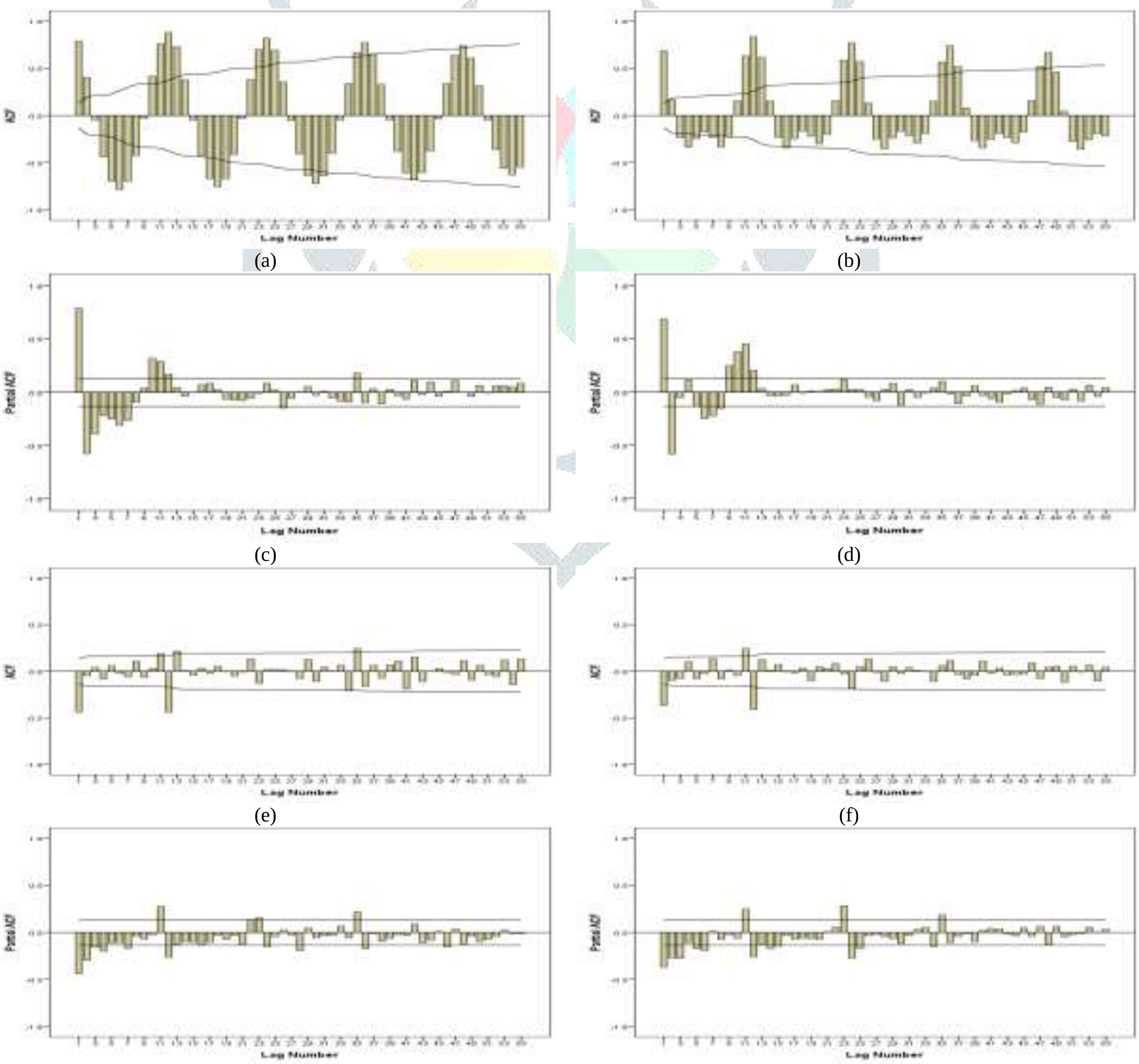
Fig. 3: Monthly average Plots of differencing at lag 12 and lag 1 series: (a) Minimum temperature after removal of trend and seasonal components (b) Maximum temperature after removal of trend and seasonal components.

To check whether the differenced series shown in Fig. 3 are stationary or not, we used the Augmented Dickey-Fuller test and the summarized results are shown in Table 2. The p-value (<0.01) for both the series less than 0.05 as well, indicates both the differenced series are stationary.

Table 2: Augmented Dickey-Fuller Test

Temperature	Test	Test statistic value	Lag order	p-value	Remark
Average Minimum Temperature	Augmented Dickey-Fuller	-9.0739	5	<0.01	The series is stationary
Average Maximum Temperature	Augmented Dickey-Fuller	-10.234	5	<0.01	The series is stationary

Fig. 4(e-f) are the ACF and PACF plots of stationary series obtained after removing trend and seasonality from monthly average minimum temperature. Figure 3f and 3h are the ACF and PACF plots of stationary series obtained after removing trend and seasonality from monthly average maximum temperature.



(a) (b) (c) (d) (e) (f) (g) (h)

Fig. 4: ACF and PACF plots: (a) ACF plot for original series of monthly average minimum temperature (b) ACF plot for original series of monthly average maximum temperature (c) PACF plot for original series of monthly average minimum temperature (d) PACF plot for original series of monthly average maximum temperature (e) ACF plot for stationary series of monthly average

minimum temperature (f) ACF plot for stationary series of monthly average maximum temperature (g) PACF plot for stationary series of monthly average minimum temperature (h) PACF plot for stationary series of monthly average maximum temperature.

The plots of ACF and PACF of stationary time series give the rough idea about the selection of SARIMA model parameters. First we obtain the SARIMA model parameters for the monthly average minimum temperature series. To obtain the non-seasonal AR term (p), we look at the PACF of stationary series of monthly average minimum temperature, which shows clear spikes at lag 1 and 2 with exponential decay. Therefore, the order of the non-seasonal AR terms will be 2 (p=2). To obtain the non-seasonal MA term (q), we look at the ACF of stationary series of monthly average minimum temperature, which shows clear spikes at lag 1 and 12. Therefore, the order of the non-seasonal MA terms will be 2 (q=2). For the seasonal part, we look at ACF and PACF plots of stationary time series at the lags 12, 24, 36 etc. The plot of PACF for stationary series of monthly average minimum temperature indicates significant spikes at lags 12, 24 and 36. Therefore the order of the seasonal AR is 3 (P=3) while the plot of ACF for stationary series of monthly average minimum temperature indicates significant spikes at lags 12 and 36. Therefore the order of the seasonal MA is 2 (i.e.Q=2). Thus our base model is SARIMA(2, 1, 2)(3, 1, 2)$_{12}$. Starting from SARIMA(2, 1, 2)(3, 1, 2)$_{12}$ model, we fit different models (See Table 3) and compared all these models using R-square, RMSE and MAPE. Among these models, SARIMA(1, 1, 2)(3, 1, 1)$_{12}$ has highest R-square (0.918) with lowest RMSE (1.258) and MAPE (5.986) values and appeared to be the best model for monthly average minimum temperature. Thus, the proposed SARIMA(1, 1, 2)(3, 1, 1)$_{12}$ model considered a best model for forecasting the monthly average minimum temperature and the model equation is:

$$\phi(B)\Phi(B^{12})\nabla\nabla_{12}\mathrm{X_t} = \theta(\mathrm{B})\Theta(\mathrm{B}^{12})\mathrm{Z_t}$$

where, $\phi(B) = 1 - \phi B = 1 - 0.46B$

$$\theta(B) = 1 + \theta_1 B + \theta_2 B^2 = 1 + 1.191B - 0.191B^2$$
$$\Phi(B^{12}) = 1 - \Phi_1 B^{12} - \Phi_2 B^{24} - \Phi_3 B^{36} = 1 - 0.047B^{12} + 0.056B^{24} + 0.128B^{36}$$
$$\Theta(B^{12}) = 1 + \Theta B^{12} = 1 + 0.999B^{12}$$
$$\nabla = (1 - \mathrm{B})\ ;\ \nabla_{12} = (1 - \mathrm{B}^{12})$$

After simplification the final equation of SARIMA(1, 1, 2)(3, 1, 1)$_{12}$ model is:

$$X_t = 1.46X_{t-1} - 0.46X_{t-2} + 1.047X_{t-12} - 1.5286X_{t-13} + 0.4816X_{t-14} - 0.103X_{t-24} + 0.1504X_{t-25} - 0.0474X_{t-26} - 0.07X_{t-36} + 0.1022X_{t-37} - 0.0322X_{t-38} + 0.126X_{t-48} - 0.184X_{t-49} + 0.058X_{t-50} + Z_t + 1.191Z_{t-1} - 0.191Z_{t-2} + 0.999Z_{t-12} + 1.189Z_{t-13} - 0.1908Z_{t-14}$$

Now, we have obtained the SARIMA model parameters for the monthly average maximum temperature series. To obtain the non-seasonal AR term (p), we look at the PACF of stationary series of monthly average maximum temperature, which shows clear spikes at lag 1, 2 and 3 with exponential decay. Therefore, the order of the non-seasonal AR terms will be 3 (p=3). To obtain the non-seasonal MA term (q), we look at the ACF of stationary series of monthly average maximum temperature, which shows clear spikes at lag 1, 11 and 12. Therefore, the order of the non-seasonal MA terms will be 3 (q=3). For the seasonal part, we look at ACF and PACF plots of stationary time series at the lags 12, 24, 36 etc. The plot of PACF for stationary series of monthly average maximum temperature indicates significant spikes at lags 12, 24 and 36. Therefore the order of the seasonal AR is 3 (P=3) while the plot of ACF for stationary series of monthly average minimum temperature indicates significant spikes at lags 12. Therefore the order of the seasonal MA is 1 (Q=1). Thus our base model is SARIMA(3, 1, 3)(3, 1, 1)$_{12}$. Starting from SARIMA(3, 1, 3)(3, 1, 1)$_{12}$ model, we have tried different models (See Table 3) and compare all these models using R-square, RMSE and MAPE. Among these models, SARIMA(0, 1, 2)(2, 1, 1)$_{12}$ has highest R-square (0.852) with lowest RMSE (1.44) and MAPE (3.606) values and appeared to be the best model for monthly average maximum temperature. Thus, the fitted SARIMA(0, 1, 2)(2, 1, 1)$_{12}$ model equation is:

$$\phi(B)\Phi(B^{12})\nabla\nabla_{12}\mathrm{X_t} = \theta(\mathrm{B})\Theta(\mathrm{B}^{12})\mathrm{Z_t}$$

where, $\phi(B) = 1$

$$\theta(B) = 1 + \theta_1 B + \theta_2 B^2 = 1 + 0.628B + 0.24B^2$$
$$\Phi(B^{12}) = 1 - \Phi_1 B^{12} - \Phi_2 B^{24} = 1 + 0.183B^{12} + 0.255B^{24}$$
$$\Theta(B^{12}) = 1 + \Theta B^{12} = 1 + 0.64B^{12}$$
$$\nabla = (1 - \mathrm{B})\ ;\ \nabla_{12} = (1 - \mathrm{B}^{12})$$

After simplification the final equation of SARIMA (0, 1, 2)(2, 1, 1)$_{12}$ model is:

$$X_t = X_{t-1} + 0.817X_{t-12} - 0.817X_{t-13} - 0.072X_{t-24} + 0.072X_{t-25} - 0.255X_{t-36} + 0.255X_{t-37} + Z_t + 0.628Z_{t-1} + 0.24Z_{t-2} + 0.64Z_{t-12} + 0.4019Z_{t-13} + 0.1536Z_{t-14}$$

Table 3: R-square, RMSE and MAPE of different SARIMA Models for monthly average minimum and maximum temperature

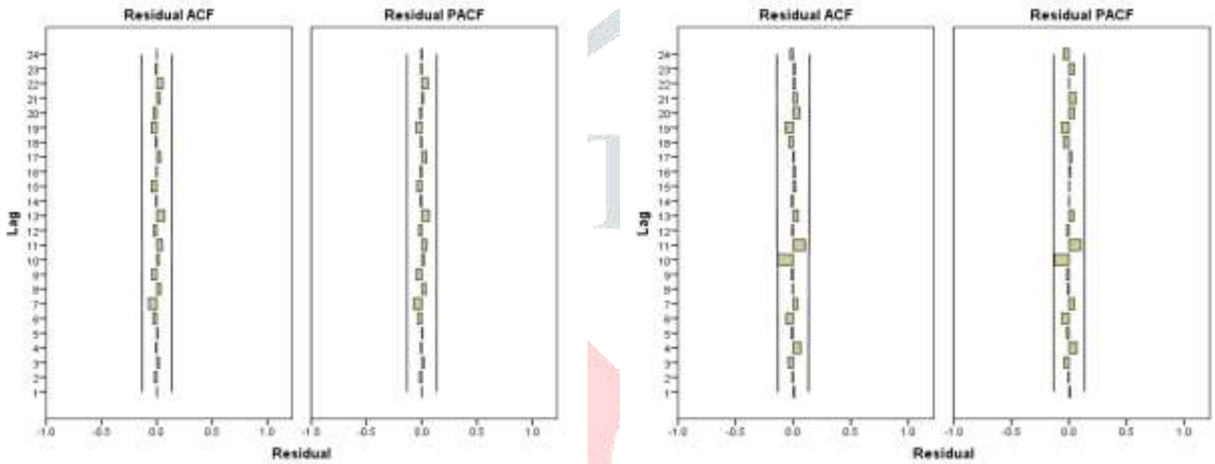
Sr. No.	Monthly Average Minimum Temperature				Monthly Average Minimum Temperature			
	Model	R-square	RMSE	MAPE	Model	R-square	RMSE	MAPE
1	SARIMA(2, 1, 2)(3, 1, 2)12	0.918	1.264	5.990	SARIMA(3, 1, 3)(3, 1, 1)12	0.847	1.482	3.786
2	SARIMA(0, 1, 2)(0, 1, 1)12	0.914	1.269	6.120	SARIMA(0, 1, 2)(0, 1, 1)12	0.848	1.452	3.755
3	SARIMA(0, 1, 2)(1, 1, 1)12	0.915	1.272	6.122	SARIMA(2, 1, 1)(0, 1, 1)12	0.849	1.453	3.738
4	SARIMA(0, 1, 2)(3, 1, 1)12	0.917	1.262	6.038	SARIMA(0, 1, 2)(3, 1, 0)12	0.849	1.454	3.740
5	SARIMA(0, 1, 2)(3, 1, 2)12	0.917	1.263	6.039	SARIMA(1, 1, 3)(0, 1, 1)12	0.850	1.453	3.755
6	SARIMA(1, 1, 2)(0, 1, 1)12	0.916	1.261	5.992	SARIMA(2, 1, 3)(0, 1, 1)12	0.850	1.453	3.672
7	SARIMA(1, 1, 2)(0, 1, 2)12	0.916	1.264	5.991	SARIMA(2, 1, 2)(3, 1, 0)12	0.851	1.454	3.760
8	SARIMA(1, 1, 2)(2, 1, 1)12	0.916	1.270	6.015	SARIMA(1, 1, 3)(3, 1, 0)12	0.851	1.453	3.740
9	**SARIMA(1, 1, 2)(3, 1, 1)12**	**0.918**	**1.258**	**5.986**	SARIMA(1, 1, 1)(2, 1, 1)12	0.851	1.445	3.724
10	SARIMA(2, 1, 1)(0, 1, 1)12	0.916	1.262	5.988	SARIMA(1, 1, 1)(3, 1, 1)12	0.852	1.446	3.700
11	SARIMA(2, 1, 1)(0, 1, 2)12	0.915	1.269	6.008	SARIMA(1, 1, 2)(2, 1, 1)12	0.852	1.443	3.707
12	SARIMA(2, 1, 1)(1, 1, 1)12	0.915	1.271	6.021	SARIMA(0, 1, 3)(2, 1, 1)12	0.852	1.443	3.706
13	SARIMA(2, 1, 1)(1, 1, 2)12	0.915	1.271	5.990	**SARIMA(0, 1, 2)(2, 1, 1)12**	**0.852**	**1.440**	**3.606**
14	SARIMA(2, 1, 1)(2, 1, 1)12	0.915	1.272	6.001	SARIMA(1, 1, 2)(3, 1, 1)12	0.852	1.446	3.691
15	SARIMA(2, 1, 1)(3, 1, 1)12	0.917	1.259	6.000	SARIMA(0, 1, 3)(3, 1, 1)12	0.852	1.445	3.688
16	SARIMA(2, 1, 2)(0, 1, 1)12	0.916	1.264	5.988	SARIMA(2, 1, 1)(2, 1, 1)12	0.851	1.441	3.693
17	SARIMA(2, 1, 2)(2, 1, 1)12	0.916	1.270	5.998	SARIMA(3, 1, 1)(2, 1, 1)12	0.849	1.445	3.691

3.2 Model Diagnostic Check

The assumptions of the SARIMA model is that, for a good model, the residuals must be uncorrelated with past values and must follow a white noise process with zero mean and constant variance. For the selected SARIMA models the Ljung–Box test is used to check the assumption of white noise. For the SARIMA(1, 1, 2)(3, 1, 1)$_{12}$ model of monthly average minimum temperature the Ljung-Box test statistic p-value is 0.908 (>0.05) and for the SARIMA(0, 1, 2)(2, 1, 1)$_{12}$ model of monthly average maximum temperature the Ljung-Box test statistic p-value is 0.657 (>0.05) as shown in table 4, which suggest that the behavior of the residuals is like white noise and there is no indication of autocorrelation in residuals of the selected models which reveal that both the selected model does not exhibit the lack of fit. Further, by looking at figure 4 of residual ACF and PACF plot, shows a random pattern with zero mean and constant variance which indicates both the fitted models are adequate and good. Therefore, SARIMA(1, 1, 2)(3, 1, 1)$_{12}$ model is used to forecast the future monthly average minimum temperature and SARIMA(0, 1, 2)(2, 1, 1)$_{12}$ model is used to forecast the future monthly average maximum temperature.

Table 4: Ljung-Box Test

Monthly Average Minimum Temperature Ljung-Box Q(18)			**Monthly Average Minimum Temperature Ljung-Box Q(18)**		
Statistics	DF	Sig.	Statistics	DF	Sig.
5.432	11	0.908	10.445	13	0.657

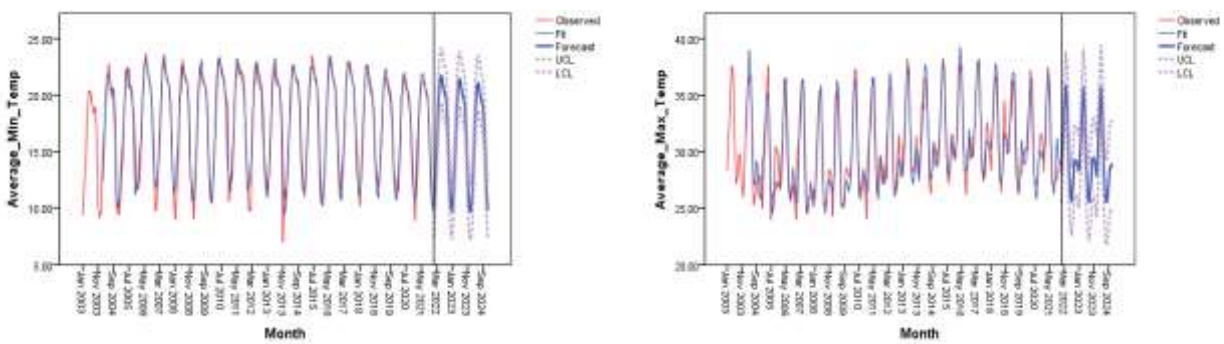


(a) (b)

Fig. 5: Plots of ACF and PACF of residuals: (a) ACF and PACF plots of residuals of SARIMA (1, 1, 2)(3, 1, 1)$_{12}$ model for monthly average minimum temperature (b) ACF and PACF plots of residuals of SARIMA(0, 1, 2)(2, 1, 1)$_{12}$ model for monthly average minimum temperature.

3.3 Forecasting using SARIMA model

The selected SARIMA (1, 1, 2)(3, 1, 1)$_{12}$ model has been used to forecast future values of monthly average minimum temperature and selected SARIMA(0, 1, 2)(2, 1, 1)$_{12}$ model has been used to forecast future values of monthly average maximum temperature. Fig. 6 (a-b) shows the observed verses fitted values with forecast of 36 months from January 2022 to December 2024 with 95% confidence bounds for both the series. It shows up from Fig. 6 that the selected models are very appropriate for anticipating the future improvement in the monthly average minimum temperature and monthly average maximum temperature, as the distinction between the observed series and fitted series is tiny. By seeing future forecast in Fig. 6, it does not show a clear trend in both the series. Thus, monthly average minimum temperature and monthly average maximum temperature might show increment or decrement in future inside the confidence bounds. At last, Table 5 shows the predicted values from January-2022 to December-2024 with 95% confidence bounds.



(a) (b)

Fig. 6: Plots of observed verses fitted values with forecast from January 2022 to December 2024 of (a) monthly average minimum temperature (b) monthly average maximum temperature.

Table 5: Predicted value with 95% confidence bounds

Month	Monthly Average Minimum Temperature			Monthly Average Minimum Temperature		
	Predicted Value	LCL	UCL	Predicted Value	LCL	UCL
Jan-22	9.69	7.4	11.99	28.16	25.46	30.86
Feb-22	12.36	9.97	14.74	30.99	28.11	33.87
Mar-22	15.57	13.17	17.98	33.55	30.64	36.45
Apr-22	19.24	16.83	21.65	35.93	33.01	38.86
May-22	21.34	18.92	23.75	34.73	31.78	37.68
Jun-22	21.77	19.35	24.18	29.17	26.2	32.14
Jul-22	20.86	18.44	23.27	26.37	23.38	29.37
Aug-22	20.11	17.7	22.53	25.57	22.56	28.58
Sep-22	19.75	17.33	22.16	27.23	24.2	30.26
Oct-22	17.32	14.9	19.73	29.23	26.17	32.28
Nov-22	12.8	10.39	15.22	29.2	26.12	32.28
Dec-22	10.17	7.76	12.59	28.87	25.78	31.97
Jan-23	9.7	7.27	12.13	28.25	25.04	31.45
Feb-23	11.97	9.54	14.4	30.29	27.04	33.54
Mar-23	15.22	12.78	17.65	33	29.72	36.28
Apr-23	18.92	16.49	21.35	35.74	32.43	39.04
May-23	21.12	18.68	23.55	33.96	30.63	37.29
Jun-23	21.47	19.04	23.91	28.77	25.41	32.13
Jul-23	20.51	18.08	22.94	26.07	22.68	29.45
Aug-23	19.8	17.37	22.23	25.49	22.08	28.9
Sep-23	19.44	17	21.87	26.98	23.55	30.42
Oct-23	16.98	14.55	19.42	29.29	25.82	32.75
Nov-23	12.47	10.04	14.91	29.45	25.96	32.93
Dec-23	10.02	7.58	12.45	29.03	25.52	32.54
Jan-24	9.64	7.2	12.07	27.83	24.27	31.4
Feb-24	11.27	8.83	13.7	30.26	26.66	33.86
Mar-24	14.6	12.17	17.04	33.14	29.51	36.77
Apr-24	18.62	16.19	21.06	35.76	32.11	39.42
May-24	20.74	18.31	23.17	33.51	29.82	37.19
Jun-24	21.13	18.69	23.56	28.54	24.83	32.25
Jul-24	20.07	17.63	22.5	25.93	22.19	29.66
Aug-24	19.46	17.02	21.89	25.49	21.73	29.25
Sep-24	18.97	16.54	21.4	26.66	22.87	30.45
Oct-24	16.7	14.26	19.13	28.45	24.64	32.27
Nov-24	12.37	9.94	14.81	28.84	25	32.68
Dec-24	9.82	7.38	12.25	28.7	24.83	32.57

IV. CONCLUSION

In this study the data of monthly average minimum temperature and monthly average maximum temperature of 228-months from January-2003 to December-2021 of Nashik (M.H.), India was gathered and analyzed. The plot of both the series saw to be non-stationary and shown slight vertical pattern. The sample ACF- PACF plots and decomposed plots are utilized to verify the non-stationarity of both the series. Both the series contains slight upward trend and slight vertical pattern with seasonality, so to change over both the series into stationary series one seasonal differencing and one non-seasonal differencing was applied. Utilizing ACF and PACF plots of stationary series the SARIMA model boundaries was acquired. Beginning from the base model SARIMA(2, 1, 2)(3, 1, 2)$_{12}$ for monthly average minimum temperature we show up at the model SARIMA(1, 1, 2)(3, 1, 1)$_{12}$ having most elevated R-square (0.918) with least RMSE (1.258) and MAPE (5.986) esteems and seemed, by all accounts, to be the best model. Beginning from the base model SARIMA(3, 1, 3)(3, 1, 1)$_{12}$ for monthly average maximum temperature we show up at the model SARIMA(0, 1, 2)(2, 1, 1)$_{12}$ having most noteworthy R-square (0.852) with least RMSE (1.44) and MAPE (3.606) esteems and seemed, by all accounts, to be the best model. The model diagnostic checking was performed by utilizing Ljung-Box test and residual ACF-PACF plots, and it shows the residuals are uncorrelated and follows repetitive sound with zero mean and steady change. Thus the models are great fitted to the data and don't show the absence of fit. Further, the selected models SARIMA (1, 1, 2)(3, 1, 1)$_{12}$ for monthly average minimum temperature and SARIMA (0, 1, 2)(2, 1, 1)$_{12}$ for monthly average maximum temperature were utilized for prediction. The assessed conjecture in both the series was extremely near to the original series. The future predicted values with 95% confidence bounds of monthly average minimum temperature and monthly average maximum temperature from January-2022 to December-2024 were given. The prediction of monthly average minimum and average maximum temperatures will be useful in rural perspective to decide administrative policies for future agriculture. The systematic procedure examined in this paper can be suggested for the determination of SARIMA model for the comparable datasets.

REFERENCES

[1] Aghelpour, P., & Norooz-Valashedi, R. 2022. Predicting daily reference evapotranspiration rates in a humid region, comparison of seven various data-based predictor models. Stochastic Environmental Research and Risk Assessment: 1-23.

[2] Alysha, M., Hyndman, R. J., & Snyder, R. D. 2011. Forecasting Time Series With Complex Seasonal Patterns Using Exponential Smoothing, J. of the American Statistical Association, 106:496, 1513-1527.

[3] Athiyarath, S., Paul, M., and Krishnaswamy, S. 2020. Comparative Study and Analysis of Time Series Forecasting Techniques. SN Computer Science 1: 175.

[4] Chawsheen, T.A., & Broom, M. 2017. Seasonal time-series modeling and forecasting of monthly mean temperature for decision making in the Kurdistan Region of Iraq. Journal of Statistical Theory and Practice 11(4): 604-633.

[5] Das, S., Das, J., & Umamahesh, N. V.2022. Investigating seasonal drought severity-area-frequency (SAF) curve over Indian region: incorporating GCM and scenario uncertainties. Stochastic Environmental Research and Risk Assessment, 36(6): 1597-1614.

[6] Do, P., Chow, C. W., Rameezdeen, R., & Gorjian, N. 2022. Wastewater inflow time series forecasting at low temporal resolution using SARIMA model: a case study in South Australia. Environmental Science and Pollution Research: 1-16.

[7] Ho, S.L., Xie, M., & Goh, T.N. (2002). A comparative study of neural network and Box-Jenkins ARIMA modeling in time series prediction. Computers and Industrial Engineering, 42: 371-375.

[8] Hyndman, R., and Khandakar, Y. 2008. Automatic time series forecasting: the forecast package for R. Journal of Statistical Software, 26(3): 1–22.

[9] Louloudis, G., Louloudis, E., Roumpos, C., Mertiri, E., Kasfikis, G., & Chatzopoulos, K. 2021. Forecasting Development of Mine Pit Lake Water Surface Levels Based on Time Series Analysis and Neural Networks. Mine Water and the Environment: 1-17.

[10] Shad, M., Sharma, Y. D., & Singh, A. 2022. Forecasting of monthly relative humidity in Delhi, India, using SARIMA and ANN models. Modeling earth systems and environment: 1-9.

[11] Shi, X., Lu, W., Zhao, Y., & Qin, P. .2018. Prediction of Indoor Temperature and Relative Humidity Based on Cloud Database by Using an Improved BP Neural Network in Chongqing. IEEE Access, 6: 30559–30566.

[12] Tadesse, K.B., and Dinka, M.O. 2017. Application of SARIMA model to forecasting monthly flows in Waterval River. South Africa. J. Water L. Dev., 35(1): 229–236.

[13] Wakaura, M., & Ogata, Y. 2007. A time series analysis on the seasonality of air temperature anomalies. Meteorol. Appl. 14: 425–434.