MAUSAM, **77**, 1 (January 2026), 55-70

MAUSAM

DOI: https://doi.org/10.54302/zn4ewx26
Homepage: https://mausamjournal.imd.gov.in/index.php/MAUSAM

UDC No.551.581.1:551.524(540.35)

A study on precision of statistical models of temperature in various zones of Assam

DIPANJALI RAY[1], TANUSREE DEB ROY[1*] and SEBUL ISLAM LASKAR[2]

[1]*Department of Statistics, Assam University, Silchar, Assam, India*

[2]*India Meteorological Department, New Delhi, India*

(*Received 5 February 2024, Accepted 13 October 2025*)

***Corresponding author's email : tanusree.deb.roy@gmail.com**

सार – विश्व स्तर पर तापमान में उल्लेखनीय वृद्धि देखी जा रही है, जो जलवायु परिवर्तन को प्रतिकूल रूप से प्रभावित कर रही है। अनेक शोधकर्ताओं द्वारा वार्षिक एवं मौसमी आँकड़ों के आधार पर तापमान की विभिन्न विशेषताओं का अध्ययन किया गया है। यह शोधपत्र असम के विभिन्न क्षेत्रों (गुवाहाटी, डिब्रूगढ़, सिलचर एवं तेजपुर) में 1985-2022 की अवधि के लिए तापमान के सांख्यिकीय मॉडलों की सटीकता का तुलनात्मक मूल्यांकन प्रस्तुत करता है। अध्ययन में दो मापदंडों–मासिक औसत न्यूनतम तापमान तथा मासिक औसत अधिकतम तापमान–का विश्लेषण वार्षिक एवं मौसमी आँकड़ों के आधार पर किया गया। परिणाम स्पष्ट रूप से दर्शाते हैं कि मौसमी आँकड़ों पर आधारित सांख्यिकीय मॉडलों की सटीकता, वार्षिक आँकड़ों पर आधारित मॉडलों की तुलना में अधिक है। अध्ययन से यह भी ज्ञात होता है कि मासिक औसत अधिकतम तापमान में समग्र रूप से वृद्धि की प्रवृत्ति विद्यमान है, जबकि मासिक औसत न्यूनतम तापमान में महीनों के अनुसार वृद्धि एवं कमी दोनों प्रकार की प्रवृत्तियाँ देखी गईं, विशेष रूप से 1985-2018 की अवधि के दौरान गुवाहाटी, जो असम की राजधानी है।

ABSTRACT. Temperature is increasing significantly, and its influence on climate change is unfavorable. Many researchers have studied several characteristics of temperature all over the world in various frames of data usually in yearly data or season wise data. This paper focuses on the comparison of accuracy percentage of statistical model of temperature in two dimensions, one is monthly mean minimum temperature and another one is monthly mean maximum temperature based on yearly data and season wise data for the period of 1985-2022 in various zones of Assam (Guwahati, Dibrugarh, Silchar, Tezpur). The analysis clearly showed that accuracy percentage of statistical model of temperature season wise is more than that of yearly based data. The study shows that there is an increasing trend in case of monthly mean maximum temperature and in case of monthly mean minimum temperature, there is both increasing and decreasing trends in accordance to months for the period 1985-2018 in Guwahati, the capital of Assam.

Key words – Minimum temperature, Maximum temperature, Statistical model, Forecast, Accuracy.

1. Introduction

The survival of species on Earth is intricately linked to the climate, as it shapes the ecosystems and environmental conditions essential for their existence. The changes in mean climate condition, that is, the average of hundreds or thousands of events over the span of decades in the atmosphere, is defined as climate change (Huber & Gulledge, 2011). The Intergovernmental Panel on Climate Change (IPCC, 2022) in its sixth report stated that there are observed impacts on climate change that are human-induced, which have accelerated recently with the advent of new extreme events in nature. Globally, calamitous environmental changes have had serious effects on human health, agricultural productivity, and natural systems. This climate is a set of a number of parameters, and out of those parameters, temperature has a higher impact on changes in weather patterns and precipitation amounts in a particular area (Deschenes *et al.*, 2009). Amongst the various climatic parameters, Temperature emerges as the most pivotal and influential factor (Yáñez-López *et al.*, 2012). Many studies have been conducted throughout the globe looking at how urban area's physical health is affected by a rise in temperature. The most important impact of temperature rise on physical health are the aggravation of pre-existing chronic conditions, notably cardiovascular and respiratory problems, respiratory infections as well as outright heat exhaustion and heat stroke. Besides these, the continuous trend of rise in temperature and climate change have a significant impact on vector-borne illness including malaria and dengue (Wong *et al.,* 2018). Temperature varies with respect to time in a whole day.

Several researchers have studied one of the most significant parameters of climate *i.e.* temperature (surface air temperature) in various dimensions. Research indicates warming during the several past decades, daily maximum (day-time) temperatures, daily minimum (night-time) temperatures have increased. Greater quantities of anthropogenic greenhouse gases, greater cloud cover, and urbanisation may all have contributed to the rise in night-time temperatures (Turkes *et al.,* 1996). The monthly, season annual variation of temperature in Northeast Region of India is studied considering trend analysis to check the pattern of growth and decline of temperature (Jain *et al.,* 2012). To analyse the variation between the daily maximum and minimum temperature, different Assam zones were chosen.

Climate change is one of the biggest environmental threats to water availability, food production, the forest ecosystems and livelihoods for many countries in the world. Furthermore, it is widely believed that developing countries in tropical regions of the world, e.g. India, will be affected more severely than the developed ones (Machiwal & Jha, 2006**;** Shamsnia *et al.,* 2011). Due to population growth, there is a stress of some factors such as water resources, economic changes, land uses and in particular urbanization. The population growth has a significant impact on climate change (Gutiérrez, 2013). Most of emissions come from the burning of fossil fuels like gas, oil and coal, therefore the country's contributions to global warming are significant (Raihan & Tuspekova, 2022). Cong and Brady have presented a copula-based method for modelling the joint distribution of temperature and rainfall, which are utmost important for agricultural production especially in the context of climate change (Cong & Brady, 2012). Lopez *et al.,* have proposed that temperature is one of the main factors in conjunction with the rain to determine the incidence and severity of disease, but the effect could be positive and negative (Yáñez-López *et al.,* 2012). Johnson *et al.,* have discussed that extreme temperature events bring significant efforts on the environment and society (Johnson *et al.,* 2007). Roy and Das have studied time series analysis on air temperature for particular regions of Assam (Roy & Das, 2012). Yildiz and Imanov have proposed that climate change is a reality that leads to rise in temperatures. This climate change will cause disruptions in the weather, which will negatively affect the performance of aircraft engines and flight performance (Yıldız & Imanov, 2021). Karl *et al.,* have studied thoroughly that how temperature altered on a particular region and observed that there is a increase in minimum temperature and little overall change on maximum temperature (Karl *et al.,* 1991).

The North East region of India, which includes Assam, is anticipated to be extremely vulnerable to the effects of climate change due to its delicate geo-ecological structure, strategic location with international boundary, presence of the Eastern Himalayan ranges, transboundary river systems, inhabitation of ecosystem by people of different ethnic groups and inherent socio-economic differences (ASTEC, 2011). In Assam, climate is predominantly humid subtropical with hot, humid summers, severe monsoons and mild winters. The winter temperature varies from 10℃ to 22℃. The summer temperature varies between 30℃ to 36℃. Assam receives 29% of its total rainfall during the south west monsoon in July and August and 24% during September. (Guhathakurta *et al.,* 2020). Assam is divided into various agro-climatic zones namely Lower Brahmaputra Valley Zone, Upper Brahmaputra Valley Zone, North Bank Plain zone, Barak Valley Zone etc (Mandal, 2014).

TABLE 1

Selected stations from various zones of Assam

Name of Zones	Name of Stations
Lower Brahmaputra Valley Zone	Guwahati
North Bank Plain Zone	Tezpur
Upper Brahmaputra Valley Zone	Dibrugarh
Barak Valley Zone	Silchar

2. Data and methodology

2.1. *The data and climatological features of study area*

The requisite information about minimum and maximum temperature (monthly mean minimum and maximum temperature in Celsius) for the various zones of Assam has been collected from National Data Centre (NDC), India Meteorological Department (IMD), Pune (http://dsp.imdpune.gov.in/). The data covers monthly data for 38 years, from January 1985 to September 2022. Fig. 1 represents the selected geological location of this study.

The largest city in Assam and one of the largest metropolis cities in Northeastern India is Guwahati station. **"**The station, Guwahati is located within the North-East region of India, approximately between the latitudes 91°33′ E and 91°52′6″ E, and longitudes 26°4′45" N and 26°14′ N, spanning both sides of the Brahmaputra River and covers an area of around 328 square km. The region has a subtropical climate, with hot humid summers, severe monsoons, and mild winters" (Sarma *et al.*, 2020). Because of its unique geographic position, Guwahati station has played a significant role throughout history as a hub for tourism as well as vigorous port of trade, commerce, administrative headquarters and political hub. There are areas inside the study area that are thought to be the world's rainiest. Because of its extreme vulnerability to landslides, floods, and river bank erosion in Guwahati (Hemani & Das, 2016). In North-East India, Guwahati is the one of the largest city of Assam (Sultana, 2020). Dibrugarh station is one of the emerging cities in Assam. It is often referred to as North East India's industrial and communication powerhouse. It is also called "tea city of India" (Hazarika & Lahkar, 2019). "Dibrugarh station has a total size of 3381 square kilometres. It spans 94°30' E to 95°30' E longitude and 27°0' N to 27°45' N latitude. Dibrugarh is a defensive town on the south bank of the river 'Dibru' centering a 'Garh' (fort) at the intersection of 27°28' N latitude and 94°35' east (E) longitudes at 104.24 mt. above mean sea level" (Mili & Acharjee, 2014). Dibrugarh experiences subtropical weather, with relatively mild, November through February are the dry winter months, while April through mid-October is the long, hot, rainy season. The monsoon season continues from around June to early to mid-October, although showers and thunderstorms are more common from March to May (IMD). Silchar is located within the North-East Region at 92.51° E longitude and 24.5° north (N) latitude and at a height of 114.68 m above the sea level. Mostly, the weather remains dry from October to March, while the rainy season starts near about the beginning of April and that lasts up to the end of September. The winter season runs mostly from November to the end of February, while the summer months are mostly throughout the wet months (Gupta & Biswas, 2010). Silchar is the location of the world's first polo club and competitive match. The tea town of Silchar serves as a gathering place for tea planters, and different bamboo crafts are produced commercially throughout Assam's diverse regions. Silchar which is located in the Cachar district of Assam is famous for its bamboo craft (Hazarika, 2020). One of the most beautiful cities on the north bank of the Brahmaputra River in Assam is Tezpur station, is located in 26º38" N latitude and 92º48" E longitude. A variety of lower hills and hillocks, as well as a luxuriant growth of ever-green and semi-deciduous forests, adorn the physiography of this lovely town in Assam. The dynamic city of Tezpur continues to be an important hub for Assamese education, tourism, and cultural endeavors while fusing its rich historical past with contemporary growth. In case of climate, Tezpur station sees typical summer and winter temperatures of 36 ℃ and 13 ℃ , respectively, with 1836 millimeters of annual rainfall (Sonowal *et al.*, 2022).

2.2. *The statistical model*

Time Series analysis does not deal with non-stationary data. Determination of number of differences to make a series stationary is the first step to start a time series analysis to forecast a statistical model. There are number of tests to check the stationarity of the chosen data, Augmented Dicky-Fuller test (ADF test), Phillips-Perron test ("Phillips-Perron Test", 2022) and Kwiatkowski-Phillips-Schmidt-shin (KPSS) are most common test to check the stationary of the given data (V. Kumar, 2021).

The Moving Average (MA) and Auto Regressive (AR) component has significant role to determine the optimal number of order of differences to make a non-stationary data to stationary. There are no set guidelines for choosing the proper AR and MA components. In general, determination of time series statistical model is carried out by generating a set of residuals and dictating if they satisfy the requirements of a white noise procedure (Nyoni & Nyoni, 2020). The order of MA and AR can be elaborated by Autocorrelation function (ACF) and Partial Autocorrelation function (PACF). The similarity between observations of a random variable as a function of the lag of time between them is called the ACF ("Autocorrelation," 2023) and PACF is the a partial correlation of a stationary time series with its own lagged values ("Partial Autocorrelation Function," 2022).

The Box-Jenkins test is considered to determine the best suitable statistical model on the provided data. In 2012, Gerretsadikan and Sharma mentioned that the Box and Jenkins method is the most general way of approaching to forecast time series model than that of others. In 1976, Montgomery and Johnson also considered Box and Jenkins methodology as probably the most accurate method for forecasting of time series data (Roy & Das, 2012). In 1970s, Box and Jenkins developed AutoRegressive Integrated Moving Average **ARIMA (p, d, q)** model for forecasting performance, which outperforms multivariate models. The performance

TABLE 2

Table of M-K test for monthly mean minimum temperature of Guwahati for the period of 1985-2018

Month	Tau-value(τ)	p-value	Sen's Slope
January	-0.1773	0.1490	-0.0173
February	1.8068e-02	0.8937	0
March	1.2706e-02	0.9289	0
April	7.2540e-02	0.5622	0.0111
May	2.7328e-02	0.8350	0
June	1.6520e-02	0.9051	0
July	0.2907	**0.0186**	**0.0200**
August	0.1677	0.1758	0.0129
September	0.2040	0.1032	0.0142
October	-0.0173	0.1495	-0.0173
November	-0.0900	0.4670	-0.1785
December	7.4148e-02	0.5524	0.0083

of this model is better than smoothing and naïve model. In ARIMA (p,d,q) model, p denotes the order of autoregressive part of the model, q denotes the order of moving average part of the model, d denotes the order of integration accounts for differentiating to the time series data stationary (Nyoni *et al.*, 2019).

The ARIMA (p,d,q) model is a combination of AR (autoregressive) and MA (moving average). The AR(p) can be represented as

$$y_t = \phi_0 + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \cdots + \phi_p y_{t-p} + \epsilon_t \quad (1)$$

where $y_t, y_{t-1}, \ldots, y_{t-p}$ are stationaries and $\phi_1, \phi_2, \ldots, \phi_p$ are constants. $\boldsymbol{\epsilon_t}$ is a Gaussian white noise series with mean zero with p lags.

The $MA(q)$ can be represented as

$$y_t = \varepsilon_t - \theta_1 \epsilon_{t-1} - \theta_2 \epsilon_{t-2} - \cdots - \theta_q \epsilon_{t-q} \quad (2)$$

where $\theta_1, \theta_2, \ldots, \theta_q$ are parameters, $\epsilon_1, \epsilon_{t-1}, \ldots, \epsilon_{t-q}$ is a Gaussin white noise series with mean zero with q lags and differencing (d) can be represented as

$$\Phi(\mathrm{B})\Delta^d y_t = \Theta(B)\epsilon_t \quad (3)$$

where ∇^d is difference operators: $(1-B)^d$, $\Theta(B)$ is the moving average polynomial; $1 - \theta_1 B - \cdots - \theta_q B^q$, $\Phi(B)$ is an autoregressive polynomial: $1 - \phi_1 B - \cdots - \phi_p B^p$.

The Seasonal AutoRegressive Integrated Moving Average $(SARIMA)$ model is generally a $ARIMA\ (p,d,q) \times ARIMA(P,D,Q)S$ represents by an equation below

$$\nabla^d \nabla_S^D y_t = \frac{\Theta(B) \times \Theta_S(B)}{\Phi(B) \times \Phi_S(B)} \in_t \quad (4)$$

where, ∇_S^D is the seasonal difference operator: $(1 - B^S)^D$, $\Theta_s(B)$ is the seasonal moving average polynomial: $1 - \Phi_1 B^S - \cdots - \Phi_p B^{PS}$(Liu *et al.*, 2023).

3. Results and analysis

The data is collected for the study is from NDC, Pune for the period of 1985-2022. The whole data from 1985-2022 is breakdown up into two parts, the period from 1985-2018 is used for model identification and the period 2019-2022 is used to check the accuracy of the statistical time series model. In the analysis portion, at first the linearity is checked by Mann-Kendall (M-K) test. Based on automatic ARIMA model, which is one of the forecast models (Awan & Aslam, 2020), the diagnostic of statistical model for the period of 1985-2018 is done. The data set considered is both for yearly and season wise to check the deviation in accuracy measurement of the arrangement of data to forecast a best time series model on the available data. Here four seasons have been considered, namely Winter (January, February), Pre-Monsoon (March, April, May), Monsoon (June, July, August, September) and Post-Monsoon (October, November, December) for the period of 1985-2018.

TABLE 3

Table of M-K test for monthly mean maximum temperature of Guwahati for the period of 1985-2018

Month	Temperature$_{(max)}$ (τ)	p-value	Sen's Slope
January	0.2288	**0.0614**	**0.0375**
February	0.3723	**0.0022**	**0.0851**
March	0.3564	**0.0034**	**0.0789**
April	0.1892	0.1226	0.0285
May	0.3149	**0.0101**	**0.0466**
June	0.4497	**0.0002**	**0.0551**
July	0.5926	**1.194e-06**	**0.0933**
August	0.5199	**2.137e-05**	**0.0642**
September	0.5841	**1.727e-06**	**0.0769**
October	0.2288	0.0614	0.0375
November	0.5420	**8.976e-06**	**0.0625**
December	0.3294	**0.0071**	**0.0428**

3.1. *Trend for monthly mean minimum and maximum temperature of Capital of Assam, Guwahati for the period of 1985-2018*

To confirm the presence of trend or linearity the Mann-Kendall (M-K) test is adopted. For this test the null hypothesis is that the data set is normally distributed on the other hand the alternative hypothesis is that the provided data is not normally distributed *i.e*. there is a trend.

Table 2 elaborates the positive or negative association between the month and observations of that month by mentioning the tau values for each month, p-value shows the significant presence of trend or not and Sen's slope value indicates the upward or downward trend according to positive & negative values of it. Among the all months, it is clearly observed that only July month has a significant p-value 0.0186 (<0.05, p-value of 0.05 denotes significance at a 95% level). The month of July month indicates a significant increasing trend. Since this month has a positive Sen's slope value. From the Sen's slope value, it is clear that the provided data on minimum temperature is fluctuating, in some months it is decreasing (negative) or increasing (positive). The positive tau value (0.2907) indicates the positive association between the month July & values of minimum temperature of that month.

In Table 3, it is observed that all months carries positive values of Sen's Slope, which depicts there is a increasing trend in all months. But considering all months January, February, March, May, June, July, August, September, November and December month has significant trend, since their p-values are less than 0.05.

Statistical Models for Monthly Mean Minimum Temperature on Yearly data and Season wise data for the period of 1985-2018

In Fig. 2, the time series plot (a) of monthly mean minimum temperature of Guwahati for the period of 1985-2018 is defined. From (b), It is clearly viewed that the there is a definite fluctuation in data. From (c), the required best statistical model of yearly data *i.e*. SARIMA (0,1,2)(2,1,0)[12] from 1985-2018 is established with the criterions of goodness of fit to get best model i.e. AIC (Akaike Information Criterian) = 977.97, AICC (Corrected Akaike Information Criterian) = 978.13, BIC (Bayesian Information Criterian) = 997.87 with p-value 0.0334 (<0.05), which depicts that there is a auto-correlation in residuals, which can be interpreted from the visualization of ACF (Autocorrelation Function) plot of residuals in (c). In the ACF plot of residuals up to 36th lag, some lags are breaching the blue dotted line *i.e*., the variation line with p-value 0.0334 interprets that there is an auto-correlation up to 36th lag and up to 24th lag, none of the autocorrelation coefficients are breaching the significant limits. All the ACF values are well within the significant bounds. Hence the interpretation can be drawn as there is no non-zero autocorrelations in the forecast residuals (or standard errors) up to 24th lag in the fitted model. The histogram with bell shaped curve in residual plot indicates residuals are normally distributed.

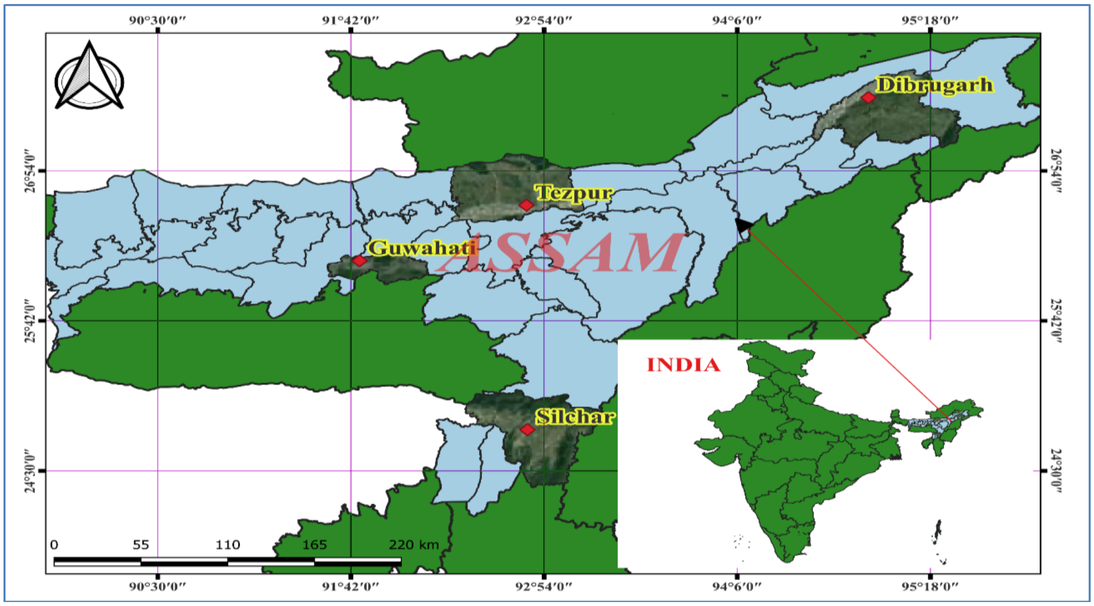
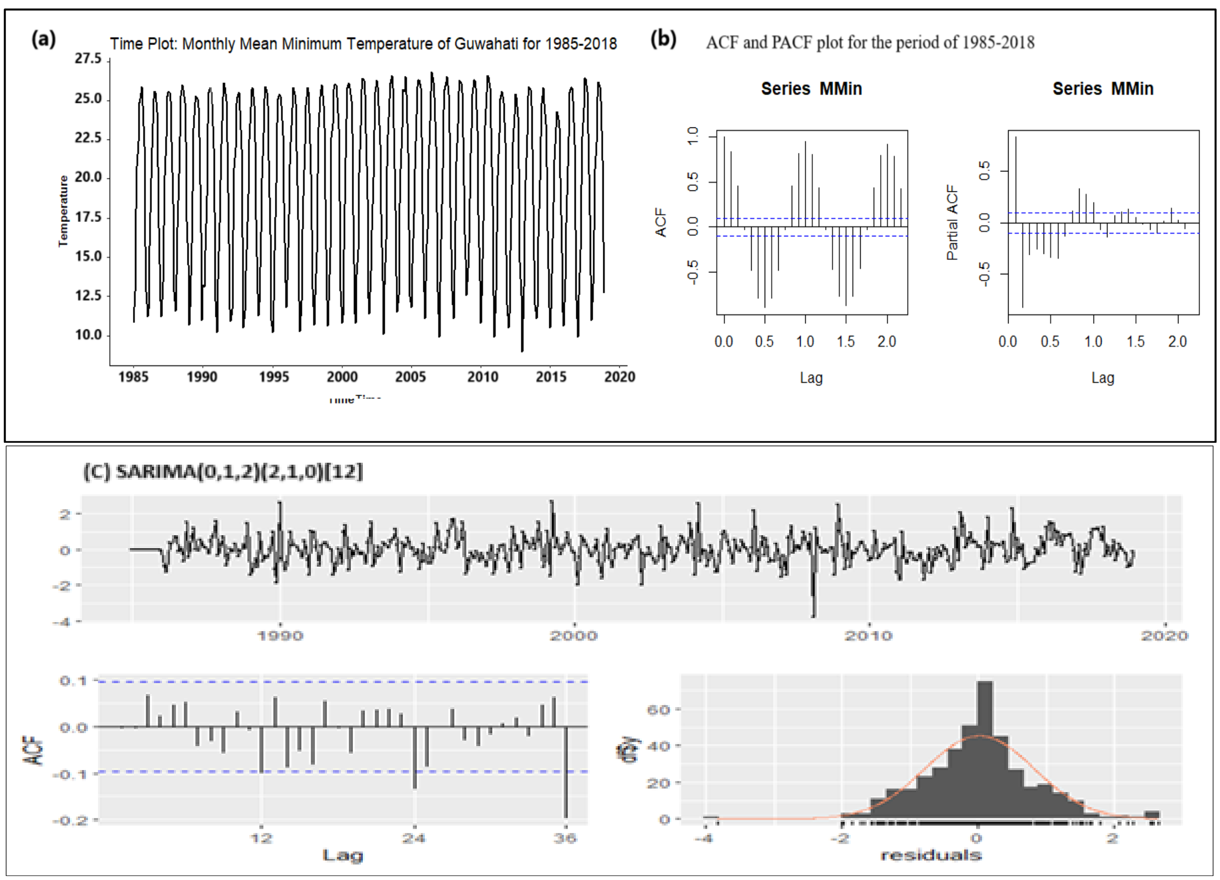
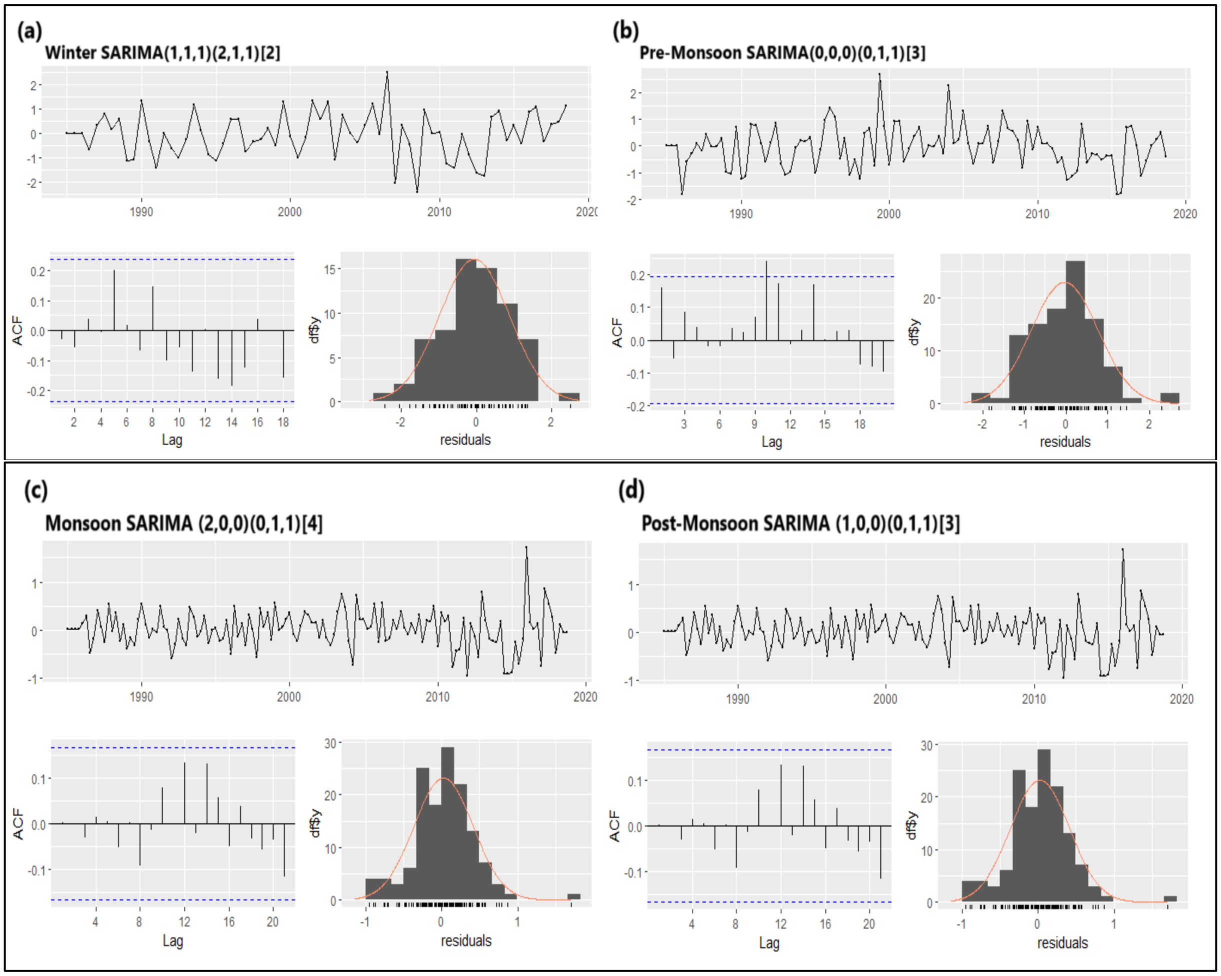


Fig. 1. Geographical location of Study areas in Assam



Figs. 2(a-c). Time Series, ACF and PACF plot of monthly mean minimum temperature of Guwahati with its statistical model for the period of 1985-2018



Figs. 3(a-d). Statistical model for Winter, Pre-Monsoon, Monsoon and Post-Monsoon seasons of monthly mean minimum temperature of Guwahati

In Fig. 3, the required best statistical model for each season *i.e.*, SARIMA (1,1,1) (2,1,1) [2] for winter with p-value 0.1386, SARIMA (0,0,0) (0,1,1) [3] for Pre-Monsoon with p-value 0.5376, SARIMA (2,0,0) (0,1,1) [4] for Monsoon with p-value 0.8762 and SARIMA (1,0,0) (0,1,1) [3] for Post-Monsoon with p-value 0.5099.

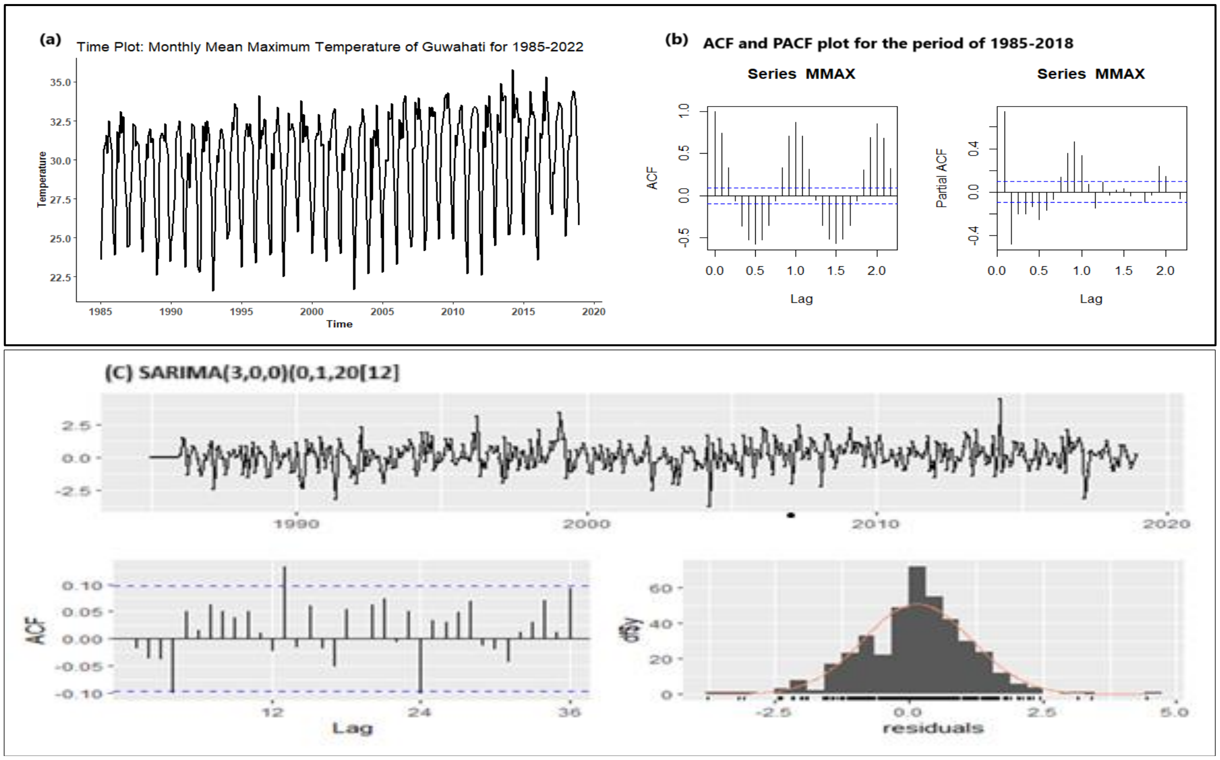
All the p-values for Winter, Pre-Monsoon, Monsoon and Post-Monsoon are not significant since all are greater than 0.05 which depicts that there is no auto-correlation in residuals, which can be interpreted from the visualization of ACF plots of residual of four seasons. In all seasons, residual plot shows their normality by their bell shaped graphical representation.

Similarly, the Statistical Model of Monthly Mean Maximum Temperature on Yearly data and Season wise data for the period of 1985-2018 is explained.

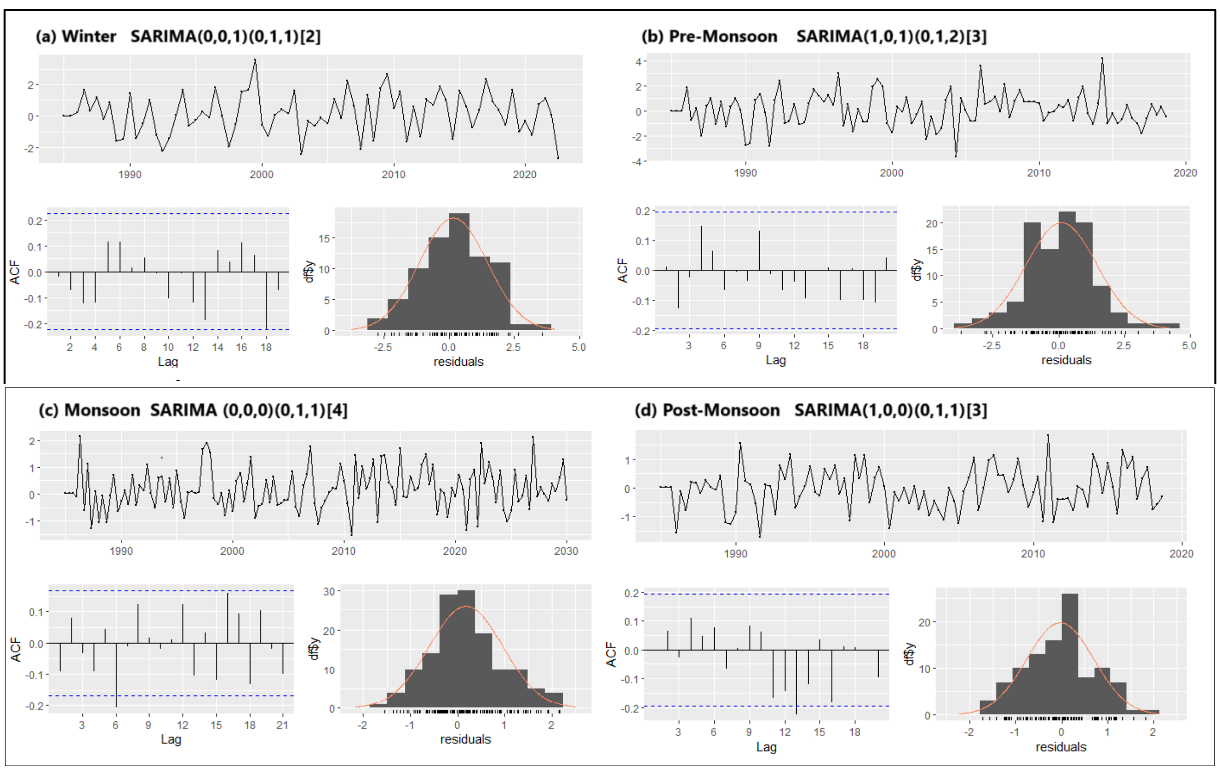
Fig. 4 represents same interpretation as of Fig. 2 with different values of AIC = 1178.95, AICC = 1179.17, BIC=1202.84 with p-value 0.0257. The best model is SARIMA (3,0.0) (0,1,2) [12] which is based on yearly data for the period of 1985-2018.

In Fig. 5, the required best statistical model for each season is identified *i.e.*, SARIMA (0,0,1) (0,1,1) [2] for winter with p-value 0.2559, SARIMA (1,0,1) (0,1,2) [3] for Pre-Monsoon with p-value 0.1675, SARIMA (0,0,0) (0,1,1) [4] for Monsoon with p-value 0.5417 and SARIMA (1,0,0) (0,1,1) [3] for Post-Monsoon with p-value 0.7187. All the p-values for Winter, Pre-Monsoon, Monsoon and Post - Monsoon are significant since all are greater than 0.05.

3.2. *Comparison of accuracy of statistical model of Yearly Data and Seasonwise Data of Monthly Mean Minimum Temperature and Maximum Temperature of Guwahati for the period of 1985-2018.*



Figs. 4(a-c). Time Series, ACF and PACF plot of monthly mean maximum temperature of Guwahati with its statistical model for the period of 1985-2018



Figs. 5(a-d). Statistical model for Winter, Pre-Monsoon, Monsoon and Post-Monsoon seasons of monthly mean maximum temperature of Guwahati

TABLE 4

The forecasted values of monthly mean minimum temperature in winter by the above model i.e. SARIMA (0,1,2) (2,1,0) [12]

Months	Oi	Ei	Accuracy (%)
Jan-19	10.30	11.03	92.94
Feb-19	13.30	13.75	96.59
Jan-20	11.20	11.06	98.73
Feb-20	12.30	13.66	88.96
Jan-21	12.70	11.54	90.87
Feb-21	13.50	14.05	95.94
Jan-22	11.50	11.71	98.19
Feb-22	11.00	14.33	69.70
	Average		91.49

TABLE 5

The forecasted values of monthly mean minimum temperature in winter by the above model i.e. SARIMA (3,0,0)(0,1,2)[12]

Months	Oi	Ei	Accuracy (%)
Jan-19	26.10	24.48	93.80
Feb-19	27.30	27.98	97.51
Jan-20	24.30	24.69	98.42
Feb-20	26.50	27.99	94.39
Jan-21	25.20	24.68	97.92
Feb-21	28.90	27.98	96.81
Jan-22	25.20	24.68	97.92
Feb-22	25.10	27.98	88.53
	Avg		**95.67**

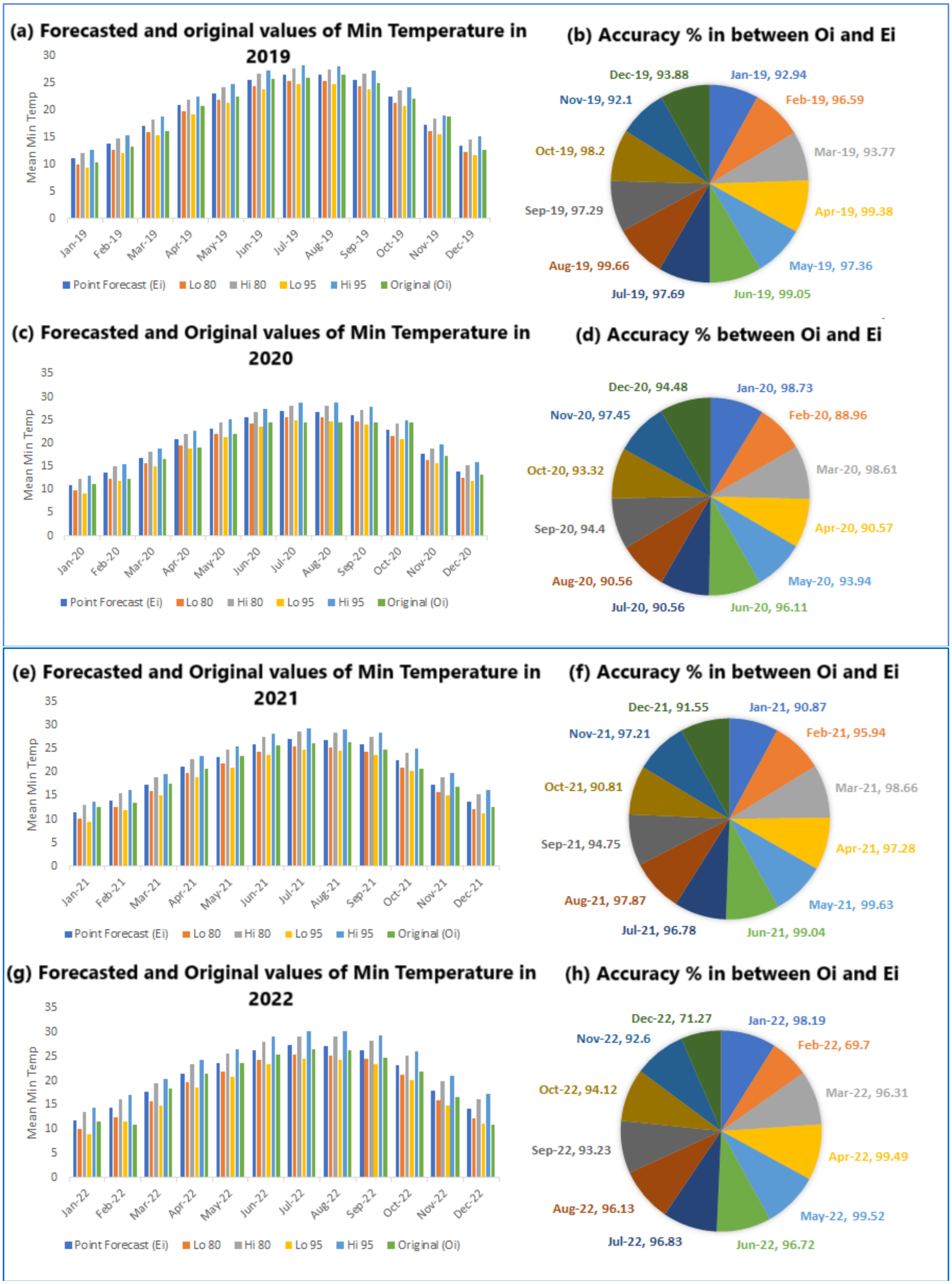
Forecasted values of monthly mean minimum temperature and its accuracy

The monthly mean minimum temperature for 2019, 2020, 2021 and 2022 year are forecasted by the best identified model SARIMA (0,1,2) (2,1,0) [12] considering the whole data yearly for the period of 1985-2018. The accuracy of this model is evaluated by comparing the original values and forecasted values of monthly mean minimum temperature values of 2019, 2020, 2021 and 2022

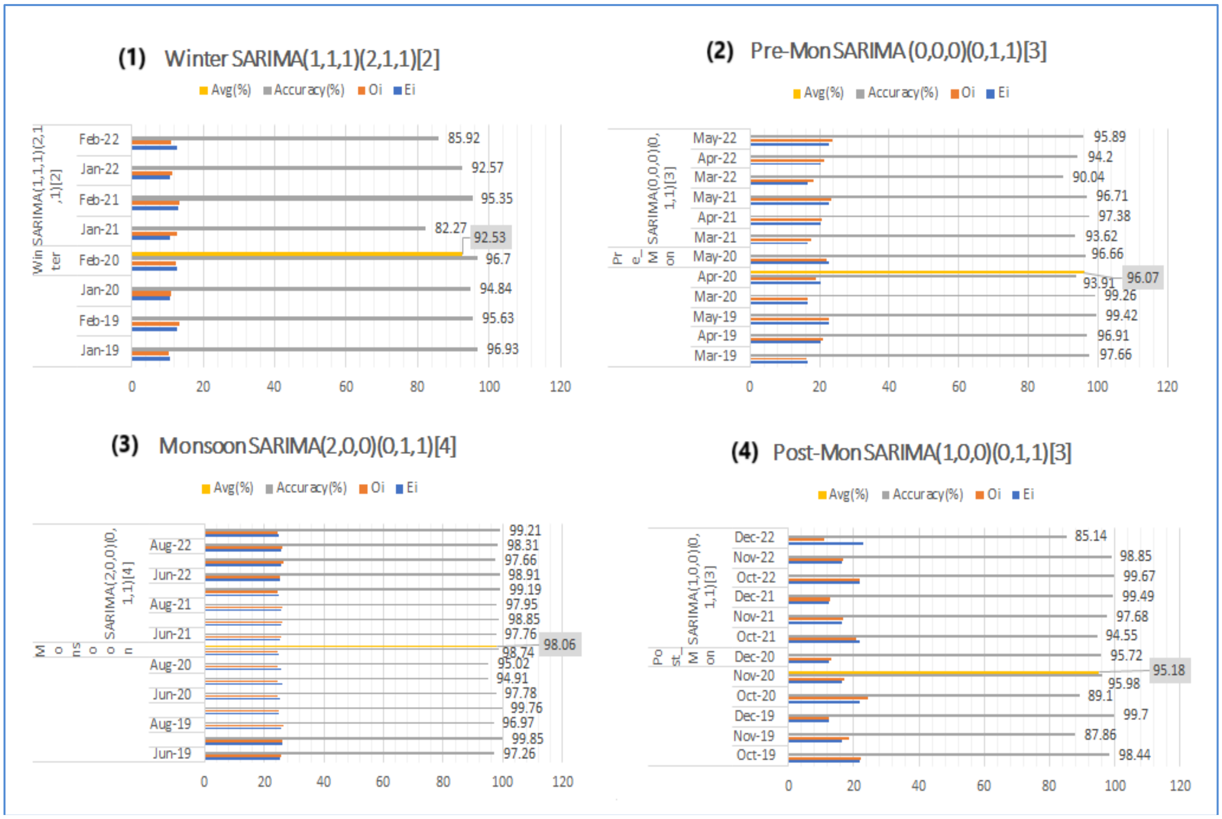
In Fig. 6, both original values and forecasted values of monthly mean minimum temperature of 2019-2022 are plotted monthly. Here (a) represents the original (Oi) and forecasted (Ei) values of minimum temperature in various percentage levels (Low 80%, High 80%, Low 95% and High 95%). (b) represents accuracy percentage in between observed and expected values minimum temperature for each month in 2019 and similarly (c), (d), (e), (f), (g) and (h) can be interpreted. The accuracy of the whole forecasted values to the original values is found to be 94.59% by calculating the average of all accuracy percentages of four years *i.e.*, the accuracy of the SARIMA (0,1,2) (2,1,0) [12] model is 94.59%.

The SARIMA (0,1,2) (2,1,0) [12] model is established by considering the whole data yearly wise for the period of 1985-2018 and mean minimum temperature values are projected by this model for the year 2019, 2020, 2021 and 2022. The accuracy of this model is checked by comparing the values of original and projected values. By considering the projected or forecasted values for January and February months from each year *i.e*., 2019, 2020, 2021 and 2022 which is for winter season the accuracy can be checked by this model. Similarly, it can be done for Pre-Monsoon, Monsoon and Post-Monsoon.

From Table 4, it is observed that the accuracy of the model SARIMA (0,1,2) (2,1,0) [12] is 91.49% for the January and February month from each four year, *i.e*. for winter.



Figs. 6(a-h). Forecasted values of monthly mean minimum temperature for 2019-2022 from SARIMA(0,1,2)(2,1,0)[12] with its accuracy percentage between forecasted (Ei) and Original (Oi) values of monthly mean minimum temperature in Guwahati



Figs. 7(1-4). Season wise forecasted values and original values of monthly mean minimum temperature by season wise model with its accuracy percentage

Similarly, by combining each month of 2019, 2020, 2021 and 2022, the accuracy of the selected model can be elaborated for Pre-Monsoon, Monsoon and Post-Monsoon. The accuracy of the SARIMA (0,1,2) (2,1,0) [12] model for Pre-Monsoon is 97.04%, for Monsoon is 96.04% and for Post-Monsoon is 92.25%.

In Fig. 7, all the original and forecasted values of monthly mean minimum temperature are represented season wise with the help of season wise statistical model. The accuracy is calculated season wise by comparing the original values and forecasted values from where it is observed that the accuracy percentage of season wise model is quite more than that of the statistical model SARIMA (0,1,2) (2,1,0) [12] considered yearly for the period 1985-2018. The accuracy percentage of forecasted values of monthly mean minimum temperature is more in Winter (92.53%), Monsoon (98.06%) and Post-Monsoon (95.18%) defined by the season wise model than that of defined by the statistical model, which was based on yearly data for the period of 1985-2018.

In Fig. 8, by considering the statistical model SARIMA (3,0,0) (0,1,2) [12], all the original values and forecasted values of monthly mean maximum temperature for the period 2019-2022 are represented in (a), (c), (e) and (g). The accuracy percentage in between observed and expected values of maximum temperature for each month of the years 2019, 2020, 2021 and 2021 is represented in (b), (d), (f) and (h) like Fig. 6. The accuracy of this model is 96.87% in prediction for monthly mean maximum temperature of Guwahati.

Table 5 represents forecasted values of monthly mean maximum temperature only for January and February of 2019-2022 *i.e.* for winter season by the SARIMA (3,0,0) (0,1,2) [12] model, which was identified depending on yearly based data from 1985-2018 and 95.67% is found accurate for the winter season Similarly, 97.18% accuracy for pre-monsoon, 97.91% accuracy for monsoon and 95.99% accuracy for post-monsoon season is found.

In Fig. 9, all the expected and forecasted values of monthly mean maximum temperature are represented season wise with the help of season wise statistical model and it is found that accuracy percentage of the forecasted values are more in winter season (95.76%) and post-monsoon season (96.17%) considering the season wise

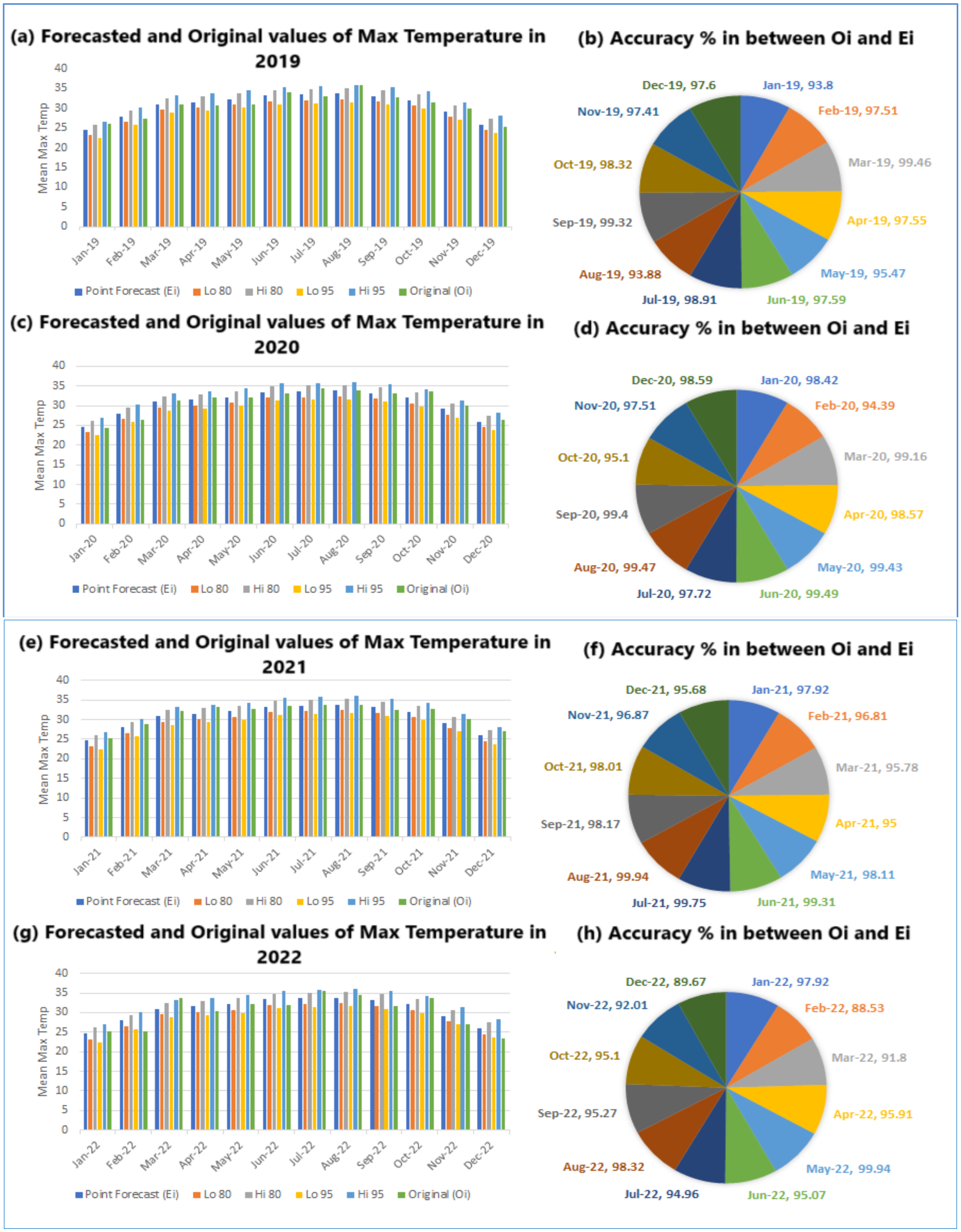
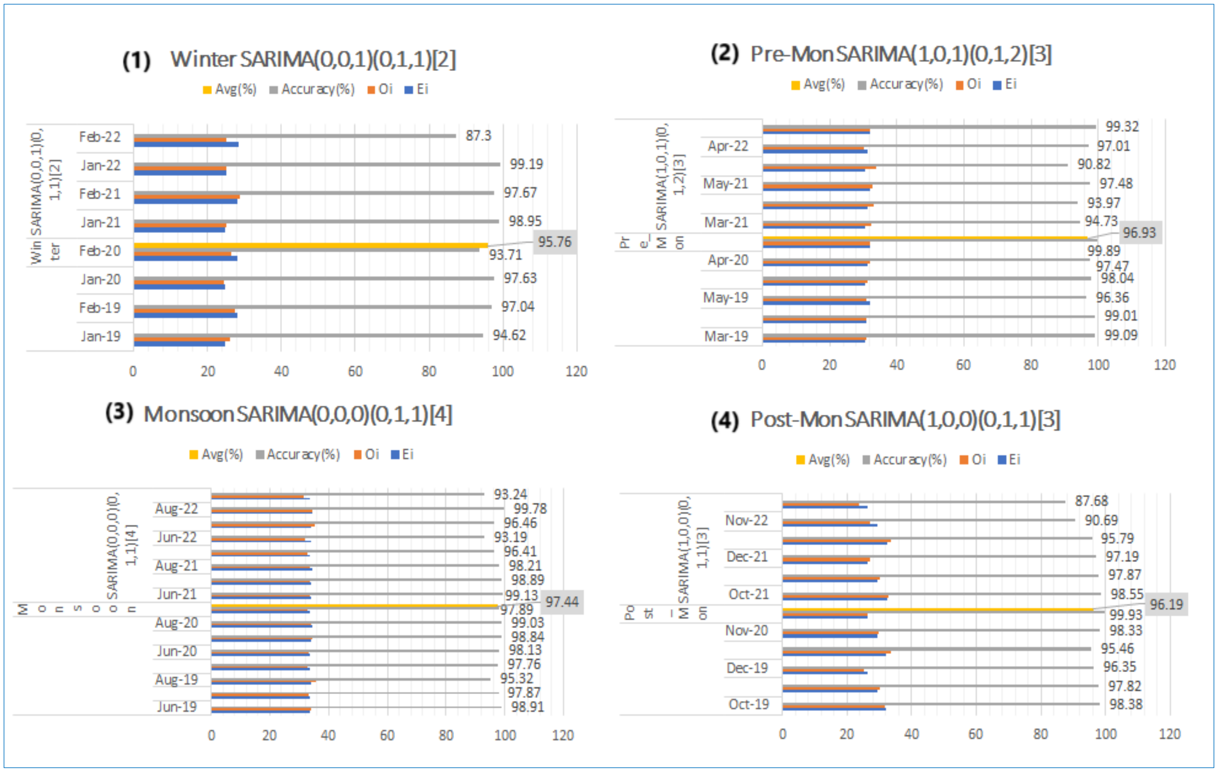


Fig. 8(a-h). Forecasted values of monthly mean maximum temperature for 2019-2022 by SARIMA (3,0,0)(0,1,2)[12] (yearly based data)



Figs. 9(1-4). Season wise forecasted values by season wise model with its accuracy

TABLE 6

Table for the accuracy percentage of yearly and season wise based data of Guwahati station

Monthly Mean Minimum Temperature			**Monthly Mean Maximum Temperature**		
Model	Period	Accuracy (%)	Model	Period	Accuracy (%)
	Winter	91.49		Winter	95.67
SARIMA (0,1,2) (2,1,0) [12]	Pre-Monsoon	97.04	SARIMA (3,0,0) (0,1,2) [12]	Pre-Monsoon	97.18
	Monsoon	96.04		Monsoon	97.91
	Post-Monsoon	92.25		Post-Monsoon	95.99
SARIMA (1,1,1) (2,1,1) [2]	Winter	92.53	SARIMA (0,0,1) (0,1,1) [2]	Winter	95.76
SARIMA (0,0,0) (0,1,1) [3]	Pre-Monsoon	96.07	SARIMA (1,0,1) (0,1,2) [3]	Pre-Monsoon	96.93
SARIMA (2,0,0) (0,1,1) [4]	Monsoon	98.06	SARIMA (0,0,0) (0,1,1) [4]	Monsoon	97.44
SARIMA (1,0,0) (0,1,1) [3]	Post-Monsoon	95.18	SARIMA (1,0,0) (0,1,1) [3]	Post-Monsoon	96.19

models *i.e.* SARIMA (0,0,1) (0,1,1) [2] and SARIMA (1,0,0) (0,1,1) [3] respectively compared to the accuracy percentage of that of defined by yearly data based model SARIMA (3,0,0) (0,1,2) [12]. The whole analysis for Guwahati can be framed in a tabular form in Table 6.

Similarly, the models for monthly mean maximum temperature and minimum temperature can be elaborated in yearly based and season wise data of Dibrugarh, Silchar and Tezpur. Which is given in tabular form in Table 7.

TABLE 7

Table for the accuracy percentage of yearly and season wise based data of Dibrugarh, Silchar and Tezpur

Dibrugarh (1985-2018)					
Monthly Mean Minimum Temperature			**Monthly Mean Maximum Temperature**		
Model	**Period**	**Accuracy (%)**	**Model**	**Period**	**Accuracy(%)**
	Winter	72.02%		Winter	93.24%
SARIMA(2,0,0)(2,1,1)[12]	Pre-Monsoon	86.71%	SARIMA (1,0,1) (0,1,2)[12]	Pre-Monsoon	92.18%
	Monsoon	91.28%		Monsoon	94.35%
	Post-Monsoon	72.07%		Post-Monsoon	93.79%
SARIMA(0,0,0)(2,1,1)[2]	Winter	93.27%	SARIMA (0,0,1)(0,1,1)[2]	Winter	95.68%
SARIMA(0,0,0)(0,1,1)[3]	Pre-Monsoon	69.41%	SARIMA (1,0,0)(0,1,1)[3]	Pre-Monsoon	94.51%
SARIMA(2,0,0)(0,1,1)[4]	Monsoon	97.41%	SARIMA (1,0,0)(1,1,1)[4]	Monsoon	96.74%
SARIMA(1,0,0)(0,1,1)[3]	Post-Monsoon	88.47%	SARIMA (0,0,1)(0,1,2)[3]	Post-Monsoon	96.52%
Silchar (1985-2018)					
Monthly Mean Minimum Temperature			**Monthly Mean Maximum Temperature**		
Model	**Period**	**Accuracy(%)**	**Model**	**Period**	**Accuracy(%)**
	Winter	91.50%		Winter	95.45%
SARIMA(1,0,1)(2,1,0)[12]	Pre-Monsoon	94.01%	SARIMA(1,0,0)(1,1,1)[12]	Pre-Monsoon	94.01%
	Monsoon	94.74%		Monsoon	94.74%
	Post-Monsoon	92.15%		Post-Monsoon	92.15%
SARIMA(1,1,0)(2,1,2)[2]	Winter	92.28%	SARIMA(0,0,1)(0,1,1)[2]	Winter	95.86%
SARIMA(1,0,2)(2,1,0)[3]	Pre-Monsoon	94.01%	SARIMA(1,0,1)(0,1,2)[3]	Pre-Monsoon	97.04%
SARIMA(1,0,0)(2,1,2)[4]	Monsoon	96.84%	SARIMA(0,0,0)(1,1,1)[4]	Monsoon	96.94%
SARIMA(1,0,0)(0,1,2)[3]	Post-Monsoon	93.89%	SARIMA(1,0,0)(0,1,1)[3]	Post-Monsoon	97.30%
Tezpur (1985-2018)					
Monthly Mean Minimum Temperature			**Monthly Mean Maximum Temperature**		
Model	**Period**	**Accuracy (%)**	**Model**	**Period**	**Accuracy (%)**
	Winter	94.10%		Winter	96.41%
SARIMA(1,0,2)(2,1,0)[12]	Pre-Monsoon	96.01%	SARIMA(0,0,3)(1,1,1)[12]	Pre-Monsoon	96.08%
	Monsoon	98.50%		Monsoon	97.18%
	Post-Monsoon	95.74%		Post-Monsoon	97.19%
SARIMA(1,0,0)(0,1,2)[2]	Winter	94.99%	SARIMA(1,0,0)(0,1,1)[2]	Winter	96.42%
SARIMA(1,0,0)(2,1,1)[3]	Pre-Monsoon	95.12%	SARIMA(1,0,1)(0,1,2)[3]	Pre-Monsoon	96.05%
SARIMA(1,0,1)(1,1,2)[4]	Monsoon	96.97%	SARIMA(0,0,0)(0,1,1)[4]	Monsoon	96.85%
SARIMA(0,0,2)(0,1,1)[3]	Post-Monsoon	93.06%	SARIMA(1,0,0)(0,1,1)[3]	Post-Monsoon	96.85%

4. Conclusions

The paper focuses in establishing an optimal framework for data representation, which will yield a more suitable and accurate statistical model. To achieve this, data the period from 1985 to 2022 has been sourced from NDC, IMD, Pune. The accuracy of the statistical model can be assessed by comparing the forecasted temperature values with the actual observed values. To obtain the original values, the dataset is split into two segments: 1985-2018 and 2019-2022. The statistical time series model is developed for the 1985-2018 period, using both annual and seasonal data frameworks. Forecasts are then generated from 2019 through 2022. The accuracy percentage is derived by comparing the forecasted values with the actual monthly mean minimum and maximum temperatures, based on both the yearly and seasonal data frameworks.

In alignment with the prescribed methodology, the initial step entails evaluating the linearity of the monthly mean minimum and maximum temperature data for Guwahati, the capital of Assam. To assess this linearity, the Mann-Kendall (M-K) test was applied, revealing both upward and downward trends in the monthly mean minimum temperatures across different months, while the monthly mean maximum temperatures displayed a consistent upward trend across all months during the period from 1985 to 2018. The accuracy percentage of the statistical model, derived from both annual and seasonal frameworks for the monthly mean minimum and maximum temperatures at Guwahati Station, is further elaborated.

It is evident that the accuracy percentage of the season-wise statistical model surpasses that of the model based on annual data. Specifically, for the monthly mean minimum temperature, the accuracy percentages for winter, monsoon, and post-monsoon seasons using season-based data are higher compared to those derived from yearly data. Similarly, for the monthly mean maximum temperature, the accuracy during winter and post-monsoon seasons exceeds that of the annual model for Guwahati. This trend is also observed across other stations, including Dibrugarh, Silchar and Tezpur, where the season-wise model consistently demonstrates higher accuracy compared to the yearly model. These findings lead to the conclusion that season-wise data framing is superior to annual data in evaluating statistical time series models and forecasting future values. This approach has practical benefits extending beyond agriculture, influencing various aspects of daily activities.

Data Availability

The data has been collected from National Data Centre (NDC), India Meteorological Department (IMD), Pune (http://dsp.imdpune.gov.in/).

Funding

The authors declare that no funds were received during the preparation of this manuscript.

Acknowledgements

The authors would like to thank the India Meteorological Department (IMD), Pune for supplying the required study's data. Authors would like to thank the editor and the referees for their support and assistance to improve the manuscript.

Author's Contributions

Tanusree Deb Roy: Material preparation
Sebul Islam Laskar: Data collection (Email-drsebul@gmail.com)
Dipanjali Ray: Draft of the manuscript (Email-dipanjali0298@gmail.com**)**

Disclaimer: The contents and views presented in this research article/paper are the views of the authors and do not necessarily reflect the views of the organizations they belong to.

References

ASTEC, 2011, "Assam's strategy and action plan on climate change—Recommendations—First draft—ASTEC, 2011, "*India Water Portal Hindi*", https://https://www.indiawaterportal. org/articles /assams-strategy-and-action-plan-climate-change-recommendations-first-draft-astec-2011.

Autocorrelation., 2023, In Wikipedia. https://en.wikipedia.org/w/ index.php?title=Autocorrelation&oldid=1131369009.

Awan, T. M., and Aslam, F., 2020, "Prediction of Daily Covid-19 Cases in European Countries Using Automatic Arima Model". *Journal of Public Health Research*, **9**, 3.

Gutiérrez, W. B., 2013, "Population growth and climate change" *Science & Society Notes* , **3**, 1.

Cong, R. G. and Brady, M., 2012, "The interdependence between rainfall and temperature: copula analyses", *The Scientific World Journal*, 2012, **1**, e405675.

Roy, T. D. and Das, K. K., 2012, "Time series analysis of Dibrugarh air temperature", *Journal of Atmospheric and Earth Environment*, **1**, 1, 30-34.

Deschenes, O., Greenstone, M. and Guryan, J., 2009, "Climate Change and Birth Weight", *American Economic Review*, **99**, 2, 211–217.

Guhathakurta, P., Kumar, S., Menon, P., Prasad, A. K., Sable, S. and Advani, S., 2020, "Observed Rainfall Variability and Changes over Kerala State Observed Rainfall Variability and Changes Over Kerala State", *Climate Research and Services India Meteorological Department Ministry of Earth Sciences Pune*, https://doi.org/10.13140/RG.2.2.25201.10087

Gupta, R., and Biswas, A., 2010, "Wind data analysis of Silchar (Assam, India) by Rayleigh's and Weibull methods", *Journal of Mechanical Engineering Research*, **2**, 1, 010-024.

Hazarika, D. R., and Lahkar, B., 2019, "Waitresses in an Emerging City of North East India: Status, Challenges and Perceptions", *Restaurant Business*, **118**, 9, 230–244.

Hazarika, K., 2020, "Bamboo an agro-based industry: History of bamboo crafts and its importance in our day-to-day life*", IOSR Journal of Humanities and Social Science*, **25**, 11, 8, 26–30.

Hemani, S., and Das, A. K., 2016, "City profile: Guwahati", Cities, **50**, 137–157.

Huber, D. and Gulledge, J., 2011, "Extreme Weather and Climate Change: Understanding the Link and Managing the Risk", *Center for Climate and Energy Solutions*.

IPCC, 2022, IPCC Report 2022: How climate change impacts vulnerable countries / IDR. *India Development Review*. https://idronline.org/article/climate-emergency/ipcc-report-2022-how-climate-change-impacts-vulnerable-countries.

Jain, S. K., Kumar, V., and Saharia, M., 2012, "Analysis of rainfall and temperature trends in northeast India", *International Journal of Climatology*, **33**, 4, 968–978.

Johnson, H., Lloyd, A., Mur, L., Smith, A., and Causton, D., 2007, "The application of MANOVA to analyse Arabidopsis thaliana metabolomic data from factorially designed experiments", *Metabolomics*, **3**, 517–530.

Karl, T. R., Kukla, G., Razuvayev, V., Changery, M., Quayle, R., Heim, R., and Easterling, D., 1991, "Global warming: Evidence for asymmetric diurnal temperature change", *Geophysical Research Letters*, **18**, 2253–2256.

Liu, J., Yu, F., and Song, H., 2023, "Application of SARIMA model in forecasting and analyzing inpatient cases of acute mountain sickness", *BMC Public Health*, 23-56.

Machiwal, D., and Jha, M. K., 2006, "Time Series Analysis of Hydrologic Data for Water Resources Planning and Management: A Review", *Journal of Hydrology and Hydromechanics*, **3**, 54, 237-257.

Mandal, R., 2014. “Flood, Cropping Pattern Choice and Returns in Agriculture: A Study of Assam Plains, India”, *Economic Analysis and Policy*, **44**, 2014, 333-344.

Mili, N., and Acharjee, S., 2014, "Urbanisation in Dibrugarh District: An important driver of environmental degradation", *Asian Journal of Spatial Sciences*, **2**, 2014, 57-67.

Nyoni, S. P., and Nyoni, M. T., 2020, "Analyzing Trauma Cases at Gweru Provincial Hospital: What Do Arima Models Reveal", *JournalNX*, **6**, 7, 152–160.

Nyoni, T., Muchingami, Muchingami, L., Jambani, L., and Robert, N., 2019, "Foreign Direct Investment (FDI) Dynamics in India: What Do Arima Models Tell Us?", *Greener Journal of Economics and Accountancy*, **7**, 1, 01-10.

Partial autocorrelation function, 2022. In Wikipedia. https://en.wikipedia.org/w/index.php?title=Partial_autocorrelation_function&oldid=1130008279#Definition.

Phillips–Perron test, 2022. In Wikipedia. https://en.wikipedia.org/w/index.php?title=Phillips%E2%80%93Perron_test&oldid=1080038074.

Raihan, A. and Tuspekova, A., 2022, "Nexus between energy use, industrialization, forest area, and carbon dioxide emissions: New insights from Russia", *Journal of Environmental Science and Economics*, **1**, 4, Article 4.

Sarma, C. P., Dey, A. and Krishna, A. M., 2020, "Influence of digital elevation models on the simulation of rainfall-induced landslides in the hillslopes of Guwahati", *India. Engineering Geology*, **268**, 105523.

Shamsnia, S. A., Shahidi, N., Liaghat, A., Sarraf, A., and Vahdat, S. F., 2011, "Modeling of Weather Parameters Using Stochastic Methods (ARIMA Model) (Case Study: Abadeh Region, Iran)", In *International Conference on Environment and Industrial Innovation*, **12**, 1, 282-285, IPCBEE.

Sonowal, G., Doley, J., Dutta, N., and Borah, R., 2022, "Greenhouse Industry! Role of Nursery Service in Environmental Restoration of Tezpur Town, Assam (India)", Eco. *Env. & Cons*, **28**, S176-S180.

Sultana, M., 2020, "The Effect of Urbanization on Environment: With Special Reference to the City of Guwahati, Assam", *PalArch’s Journal of Archaeology of Egypt / Egyptology*, **17**, 9, Article 9.

Turkes, M., Sumer, U., and Kilic, G., 1996, "Observed changes in maximum and minimum temperatures in Turkey", *International Journal of Climatology*, **16**, 463–477.

V. Kumar, G., 2021, "Statistical tests to check stationarity in Time Series", *Analytics Vidhya*, https://www.analyticsvidhya.com/blog/2021/06/statistical-tests-to-check-stationarity-in-time-series-part-1.

Wong, L. P., Alias, H., Aghamohammadi, N., Aghazadeh, S., and Nik Sulaiman, N. M., 2018, "Physical, Psychological, and Social Health Impact of Temperature Rise Due to Urban Heat Island Phenomenon and Its Associated Factors", *Biomed Environ Science*, **31**, 7, 545-550.

Yáñez-López, R., Torres-Pacheco, I., Guevara-González, R. G., Hernández-Zul, M. I., Quijano-Carranza, J. A., & Rico-García, E., 2012, "The effect of climate change on plant diseases", *African Journal of Biotechnology*, **11**, 10, Article 10.

Yıldız, M. and Imanov, T., 2021, “Climate change effects on temperature elevation and evaluation of engine performance”, *Journal of Aerospace Science and Management*, **1**, 15–35.