International Journal of Applied Environmental Sciences
ISSN 0974-0260 Volume 21, Number 1 (2026) pp. 57-80
Research India Publications
https://www.ripublication.com

Long term time series approach to forecast annual temperatures using ARIMA. A case study of India

Lalit Kumar[1], Vijeta Pundir[2]
[1]Department of Business Economics,
Dr. Bhim Rao Ambedekar College, University of Delhi
[2]Department of Business Economics,
Maharaja Agrasen College, University of Delhi
Corresponding author: Lalit Kumar (lalit.kumar@bramb.du.ac.in)

Abstract

Rising global temperatures have resulted in increased concerns about global warming and its economic and environmental implications. This study examines the long-term time series behaviour of two important temperature components namely, annual mean maximum temperature and minimum temperature and estimates their forecasted values for the next ten years for India. Time series analysis helps in understanding temperature dynamics which is critical for climate change management and measuring its effect on the Indian economy. Global developments have demanded the urgency of such analysis as the years 2023 and 2024 have been officially confirmed as the warmest years since the global temperature data records began in 1850. The global linear trends indicate that the annual maximum temperature has risen by 1.29°C over the last century and the minimum by 1.5°C over the same time horizon. Similar pattern is witnessed in India as well. Between 1901 and 2024, the annual mean maximum temperature showed an increasing trend of 0.70°C per 100 years, while the mean minimum temperatures rose by over 0.27 °C per 100 years. The study employs Box-Jenkins methodology to build Autoregressive Integrated Moving Average (ARIMA) model to forecast and estimate the annual maximum and minimum temperature for India. It provides reliable temperature forecasts and projections for the coming decade. This information is vital for policy makers who are able to take evidence-based climate policy decisions that can be effectively adapted and implemented.

JEL Classification: C32, Q54

Keywords: Climate change, annual mean maximum temperature, annual mean minimum temperature, ARIMA, forecasting

1. Introduction

Sustained temperature increases and its variation is the leading cause of climate change at the global level. The global average temperature has increased by almost 0.86°C (0.65–1.06 °C) from 1901 to 2024 (NOOA, 2025). According to several Assessment Reports of the Intergovernmental Panel on Climate Change, the rate of increase of annual global temperature over the last 50 years is almost double that over the last 100 years (0.14°C vs 0.07°C per decade) (IPCC 2007, 2014). Year 2024 was 0.1°C warmer than 2016, which is the warmest year on record and 1.43°C warmer than the pre-industrial period from 1850-1950. India's average annual mean temperature has increased by around 0.7°C during the period 1901–2022. Average annual mean maximum and minimum temperatures increased by 0.72 and 0.27°C during 1901–2020 respectively. This trend has got accelerated and annual mean maximum and annual mean minimum temperatures have risen by 0.15 °C and 0.13 °C per decade over the period from 1980–2020. The top five warmest years recorded are 2009 (+0.71 °C anomaly), 2016 (+0.55 °C), 2017 (+0.54 °C), 2021 (+0.53 °C) and 2023(+63 °C). Since 2010, 10 of the 15 warmest years have been recorded within the last 15 years (2010–2024). This translates into an overall trend of increase of 0.63 °C per 100 years with annual maximum temperatures 0.99 °C per 100 years increasing more rapidly than minimum temperatures 0.26 °C per 100 years (Khaiwal et al, 2024). Last five years from 2020-2024 have experienced the highest five-year average global temperature. In India, the average of the mean annual temperature was 25.30 from 1901 to 2024. In the last ten years, the yearly temperature was above the average in nine years. On the regional basis, the parts of southern and western India show an increasing rising trend of 1.06 and 0.36°C per 100 years respectively, while region of north Indian show a falling trend of –0.38°C per 100 years (Nanditha et al, 2020). The annual mean temperature has increased by 0.94°C for the post-monsoon season and by 1.1°C for the winter season. Annual mean temperature, mean maximum temperature and mean minimum temperature have risen at the rate of 0.42, 0.92 and 0.09°C per 100 years respectively (Arora et al, 2005). The trend in India followed the global trend where the annual mean temperature has increased by 0.204°C per decade which was more than annual mean maximum temperature of 0.141°C per decade during the last 100 years (Vose et al, 2005). The trend has accelerated in the recent years with temperatures reaching record level with the average anomaly in 2015-2024 being higher by +0.44 °C above the 1951–2010 average.

Since 1990, annual mean minimum temperature is increasing more rapidly and its rate of increase is comparatively more than that of mean maximum temperature (IMD, 2024). The Intergovernmental Panel on Climate Change (IPCC, 2014) released a special report on global warming of 1.5°C above the mean level. It reignites emphasis on the ways to reduce the global emissions and the appropriate response to climate change within the overall context of sustainable development and sustained efforts to combat poverty. There is an urgent need to take urgent action and the world must act decisively to transition to a low carbon economy to stay below 1.5°C target.

Rise in temperature is causing erratic rainfall patterns increasing the risk of extreme weather across the globe which is impacting agricultural production and fresh water availability. Higher temperatures are the leading cause of rise in variability in production and productivity of rice and wheat crops across the world (Stern, 2007).

Climate change and the associated uncertainty have added layers of new dimensions for long-term investment decisions and quantifying these multiple challenges of climate change has gained vital importance in becoming a critical input for risk assessment and informed policy decision making at the regional level (Hallegatte, 2009). Getting accurate and timely future forecast values for the temperature is now an essential feature and of vital importance for effective planning, adaptation and mitigation policies. The earliest collaboration at the global level produced General Circulation Models (GCMs) which allowed for the global forecast to be downscaled and then applied to the regional level. But its success was limited and did not reproduce accurate result in several case studies. These global models have in-built model-related uncertainties, unstable low scale down techniques and reduced spatial resolution capability which limits its ability to analyse actual status of clouds and temperature. This when combined with uncertain climatic conditions and increasing greenhouse gases makes them an unreliable tool for regional and local analyses. Getting climate estimates at local levels involves many sequential steps and at each step, many assumptions and approximations are made which increases the uncertainties which are inherent in projections of dynamic changes in climate and their impacts (Trzaska and Schnarr, 2014; Ahmad, 2011).

Accurate yearly temperature forecasts are crucial inputs for governments to plan and strategize timely actions in order to minimize climate change losses. Time-series properties of global temperatures trends reveal a steady, consistent and non-stationary and nonlinear deterministic trends, based on various stochastic and probabilistic model in line with the underlying science of climate science. Non-stationary stochastic processes in time series data have been modelled using trend stationery and difference stationary process (Estrada et al, 2013). Some other efforts including the claim by Gay et al., 2009, that “surface temperature can be better described as a trend stationary process with a one-time permanent shock” and that complex models between temperature and radiative forcing have not been able to estimate statistically significant results.

ARIMA models have gained vast acceptance and prominence for accurate forecasting from different and varied type of time series data mainly due its inherent simple requirement of its own data. Many ARIMA papers have involved variables such as household electric consumption (Chujai et al, 2013; Contreras et al, 2003), sugar prices (Suresh and Priya, 2011), stock market prices (Almasarweh and Sadam, 2018) to get accurate future forecast values.

2. Literature Survey

ARIMA has been increasingly being used by researchers to forecast and analyze the long-term trend of the climatic variables like temperature and annual rainfall. This is mainly possible due to availability of accurate data for more than 100 years which makes the extrapolation based on past data robust and statistically significant. Here some studies at the global and Indian level are discussed to show the applicability of using ARIMA for the present study.

Chaudhuri and Dutta, 2014, have estimated the impact of climatic parameters on the city of Kolkata. The authors have analysed and identified the trends in the concentrations of atmospheric pollutants like sulphur dioxide (SO_2), nitrogen dioxide

(NO_2) etc. and climatic parameters like the surface temperature and relative humidity. They used the result to develop models for precise forecast of the concentration of the pollutants and the meteorological parameters over the city of Kolkata. They used the time series approach (ARIMA) and concluded that that ARIMA (0, 2, 2) is the best fit model for forecasting the daily concentration of pollutants as well as the climatic parameters over Kolkata.

In a study of climatic variables in Bangladesh by Sultana and and Hasan, 2015, the authors have utilized data on the monthly maximum and minimum temperature data (1949-2012) for the city of Chittagong. They used seasonal ARIMA model SARIMA (1, 1, 1) (2, 0, 0) [12] for maximum temperature and for minimum temperature SARIMA (1, 1, 1) (1, 0, 1) [12] since they used monthly data for the two variables. They came to the conclusion that forecasted temperature for the city is in line with the global trend of increasing temperature. The variation is even more pronounced for minimum temperature but shows less marked change for maximum temperature. Thus, they conclude that fitted forecasted model can be used to generate the forecast value of temperature in others areas of Bangladesh.

Climatic variability and its impact on the rice production is the focus of a study in Sri Lanka by Partheepan and Jayakumar, 2006. The study has used a forecasting approach which included factors causing climatic variability in the rice production in Batticaloa district. They took rainfall time series data to estimate the annual climatic trend and variability. They found that incorporation of long-term 100 years rainfall and temperature data in the ARIMA model was instrumental in explaining and forecasting climatic trends for overall planning for agriculture planning and management. They fitted a bivariate ARIMA (1, 1, 0) and (0, 1, 1) to analyse the rainfall time-series data. The results further estimated that rainfall has becomes more variable due to climate change and better forecast of drought and flood years could be done with the help of a drought index. They were able to show how annual rainfall behaves on a 10 year cyclical pattern and maximum mean annual temperature will increase by nearly l degree centigrade after next seven years. In another important study by Tularam and Mahbub, 2010, the authors examined the relationships between rainfall, minimum temperature and maximum temperature using time series techniques. They also examined the existence of a cyclic nature in the three variables and trend relationship between rainfall and temperature. They used 50 years data of rainfall and temperature data using ARIMA methodology to analyze climatic trends and interactions. Their analyses showed that rainfall and temperature varied with each other and the cyclical nature could be detected using ARIMA models which broadened the understanding of the complexity of the climate change.

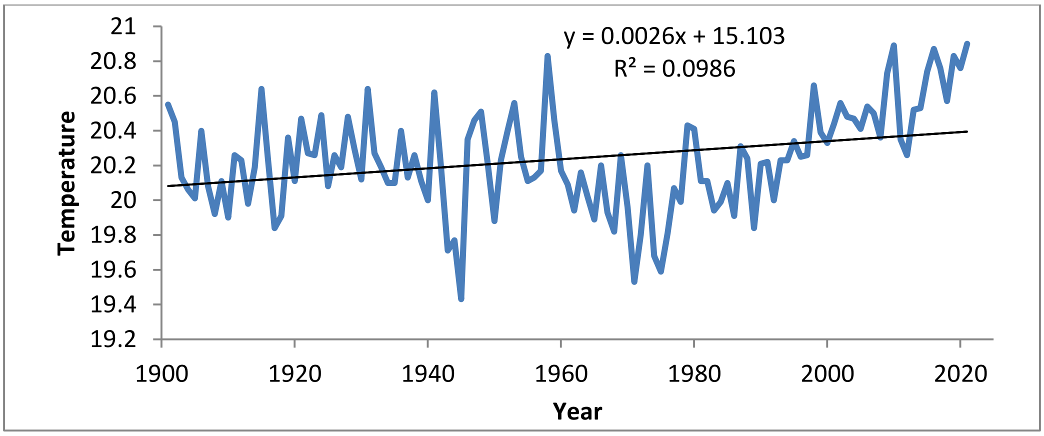
In a study by Ye Liming et al., 2013, the authors have used time series modeling framework to analyze existing records of surface temperature to study the variation patterns and the cyclic oscillations of temperature at the global level. They also showed the relative superiority of the time series models like ARIMA over other available modelling techniques. They made short-term prediction of the monthly global surface temperature using the seasonally adjusted SARIMA (3, 1, 1) x (0, 1, 1) process. They estimated that the global temperature would increase at a rate of 0.12 degree centigrade per decade, which is almost double the rate of 0.52 degree centigrade for the past 100 years.

In a study by Babu et al., 2015, the authors have demonstrated the use of ARIMA and ANIFIS (Adaptive Network Based Fuzzy Inference System) model to determine the climatic prediction using four climatic variables- Relative Humidity, Ambient Air Temperature, Barometric Pressure and Wind Direction. They concluded that in ARIMA model the output of average of testing errors weather forecasting predictions is 0.858 and in ANFIS model average of testing errors is 10.356. This shows the superior predicting quality of ARIMA as compared to more sophisticated and mathematical models used for climatic forecasting.

Mishra and Desai, 2005, have used ARIMA to forecast drought conditions as its early detection can help to adopt mitigation strategies and corrective measures in advance. They applied ARIMA and multiplicative Seasonal Autoregressive Integrated Moving Average (SARIMA) models to forecast droughts in the Kansabati river basin in the Purulia district of West Bengal. They estimated results using the ARIMA model and showed that drought could be predicted with reasonable accuracy 1–2 months ahead of the growing season. The predicted value decreases with increase in lead-time.

Babu and Reddy, 2014, combined linear and nonlinear models for arriving at a more robust prediction as compared to a standalone linear and nonlinear model for more accurate forecasting derived from time series data. They used a new hybrid ARIMA and ANN model for the prediction of time series data of sunspot, electricity price and stock market. The results obtained through this hybrid ARIMA-ANN gave better and accurate result furthering the advances in the time series forecasting.

3. Description of the two climatic variables



Source: created by authors

Fig. 1: Annual Mean Minimum Temperature (degree Celsius) in India (1901-2024)

3.1 Annual Mean Minimum Temperature (1901-2024)

The trend for the annual minimum mean temperature is also unambiguously in the upward direction. The data in the fig.1 shows that the minimum temperature is showing a clear increasing trend in line with the global average trend in the last 100 years. The long-term average value of minimum temperature for the last 100 years is 19.33 degree Celsius.

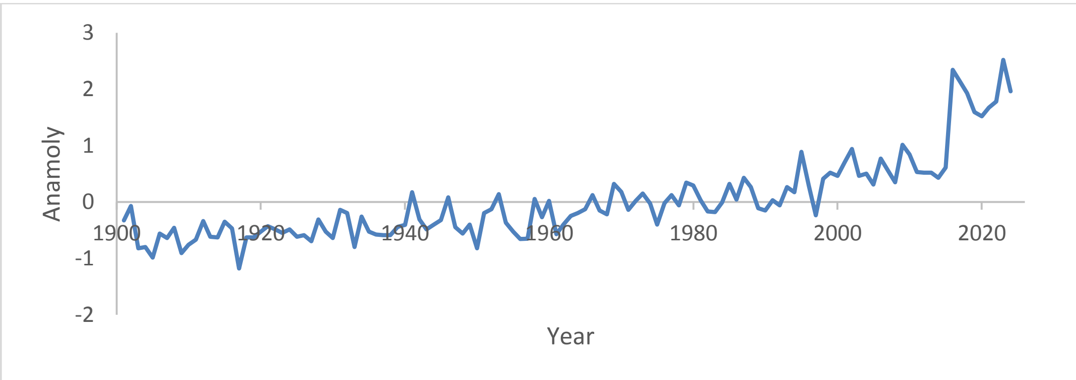
The data in the Table 1 reveals the increasing trend of the annual mean minimum temperature in India in the past 100 years. The average value for the same period is 19.33 Degree Celsius. But in the latter period, the average value for the period 1980-2024 increased to 19.80 Degree Celsius. The increase in mean minimum temperature in the period 1981-2012 is 1.5% as compared to the period 1950-1980. This increase is similar to the increase in the mean annual temperature for the two periods. The absolute and incremental change is positive which is having a negative impact on the agricultural sector whose growth rate has slowed down appreciably in the last 20 years.

Table 1: Descriptive statistics of Annual Mean Minimum Temperature (Degree Celsius) in India

Time period	**Annual Mean Minimum Temperature**	**Standard deviation**
1901-2024	19.46	0.2925
1950-2024	19.56	0.3172
1950-1980	19.22	0.2888
1981-2024	19.80	0.2810
Time Period	**% age change**	
1901-2024 and 1981-2024	+0.2%	
1950-1980 and 1981-2024	+1.5%	

Source: Statistics Related to Climate Change – India 2015

Fig 2: **Anomaly of Annual Mean Minimum Temperature (1900-2024)**



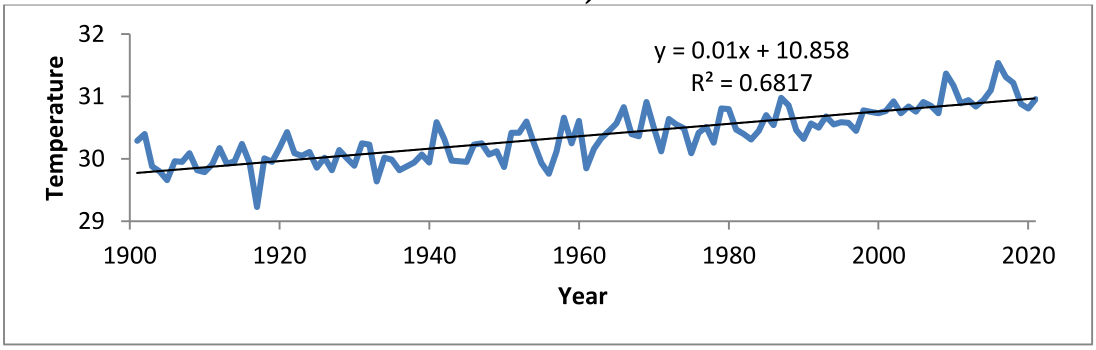
Source: created by authors

In the fig.2, the anomaly for the Annual Mean Minimum Temperature (1900-2024) shows a sharp upward trend specially after 1980's. This is particularly relevant when the anomaly was in the negative territory till the middle of the 20th century. The upward trend got accelerated after the year 2000 and has remained in the positive space with increasing pace, especially in the last 20–30 years. This implies significant warming of the annual mean temperature in the recent years along with increasing variability.

3.2 Annual Mean Maximum Temperature (1901-2024)

The trend for the annual mean maximum temperature also shows an upward trend. The data in the fig.3 shows that the maximum temperature is similarly showing a clear increasing trend in line with the global average trend in the last 100 years. The long-term average value of maximum temperature for the last 124 years is 29.29 degree Celsius.

Fig. 3: Annual Mean Maximum Temperature (degree Celsius) in India (1901-2024)



Source: created by authors

In table 2, the data from the year 1901 shows a clear and consistent increasing trend in annual mean maximum temperatures. The annual mean maximum temperature increased from 28.96°C in 1901 to 31.25 in 2024, an increase of 2.29°C.

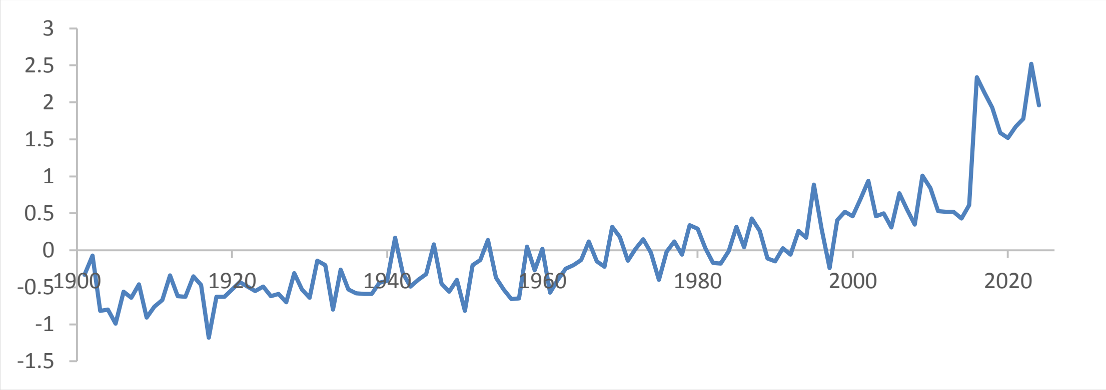
Table 2: Descriptive statistics of Annual Mean Maximum Temperature (Degree Celsius) in India

Time period	**Annual Mean Maximum Temperature**	**Standard deviation**
1901-2024	29.29	0.4277
1950-2024	29.62	0.3436
1950-1980	29.14	0.2067
1981-2024	29.95	0.2580
Time Period	**% age change**	
1901-2021 and 1981-2024	1.3	
1950-1980 and 1981-2024	1.2	

Source: created by authors

The period from 1981-1924 shows an average temperature of 29.95 degree Celsius which is almost 0.6 degree Celsius more than of the period 1901-2024. The percentage increase of 1.2–1.3% is highly significant since positive changes in average temperature is having substantial environmental impacts. This has happened with increasing temperature variability as witnessed in the late 20th century and in the last 30 years.

Fig 4: **Anomaly of Annual Mean Maximum Temperature (1900-2024)**



Source: created by authors

In fig.4, the data shows that temperature anomalies have shown a clear rising trend after the middle of 20th century. Up to the 1950's, the anomalies were in negative range but since have remained in the positive territory sharply rising after the year 2000. The worrying trend after 2000 reveals a sharp increase, with anomalies scaling new heights indicating significantly higher temperatures as compared to the earlier baseline period. The variability also indicates a clear long-term upward trend.

4. Source and Description of Data

The most comprehensive source of climatic and other environmental date in India is available on the website of India Meteorological Department (IMD), Pune, with data available on www.data.gov.in. and www.mausam.imd.gov.in. The data for minimum and maximum mean annual temperatures from 1901-2024 were obtained from these websites.

5. Methodology of the paper

The purpose of this paper is to conduct a statistical analysis of time series data of two climatic variables, namely, annual mean maximum temperature and annual mean minimum temperature from the period 1901 to 2024 using ARIMA analysis. Climate change is not a homogenous factor and differs with latitude and longitudes spatially. Climatic variables like maximum and minimum annual average temperature and annual monsoon rainfall have shown a linear upward trend in the last 100 years. The other climatic factors like El Nino, Sea Surface Temperature, and Low-Pressure Areas etc. have also shown a non-linear irregular trend in the last 50 years. It has thus become important to analyze these variables in Time Series set up to obtain a link between the past and future predicted values. In this approach the past year variable and the lagged variable are the explanatory variables. The simplest models which link past and present values are the autoregressive (AR) models, moving average (MA) models, autoregressive moving average (ARMA) models and autoregressive integrated moving average (ARIMA) models. AR and MA models are combined to produce a special class of time series models known as ARMA models. The climatic variables have to be made stationary before they can be used in the autoregressive models. ARIMA models are expressed in order (p, d, q), where p is the order of the AR part, d is the degree of first differencing involved and q is the order of the moving average. To quantify the impacts, we need to show clear link between econometric models with empirical climate variables. The study looks at the performance of 123 years of (1901-2024) temperature data and estimate their future forecast using ARIMA method. Past trends of rainfall and temperature exert a strong influence on the pattern of present and future climate change. Informed policy decisions to tackle climate change must rely on both estimates of the physical quantitative estimates and uncertainties of the climate change response to increasing anthropogenic activities and its economic impacts

5.1 ARIMA Methodology

The ARIMA method and its application to the time series analysis essentially consists of three main steps: testing for stationarity of the data, lag order selection, and model selection (Jenkins et al 2015). First, the order of differencing is determined using the Augmented Dicky Fueller (ADF) and Philips-Perron stationary tests. The building block of ARIMA is to verify the stationarity of the given series since ARIMA model estimation works only for the stationary series. Non-stationary is removed by differencing the series once or twice. Let X be the given original series which is non-stationary, then the first difference is given by $X_t - X_{t-1}$. Second, the lag order is chosen using the criterion of AIC and BIC values. Lowest value of the Akaike Information Criteria (AIC) and Bayesian Information Criteria (BIC) obtained is usually selected for further analysis of the given variables. Thirdly, in this paper we have used the LjungeBox Q-statistical tests to choose the best available ARIMA model. ARIMA model results are held significant only after the mandatory diagnostic required to check the adequacy of the model. This test estimates the independence and significance of the residuals generated in the ARIMA process and behave as white noise process using chi square $\square^2$ statistics. This implies that series of the two annual temperatures are random and independent. The LjungeBox Q-statistics is given by $Q = (n+2)(n)\sum_{k=1}^{k} p_k^2$/(n-k), where where $n = N - 1$ (sample size) and d (degree of differencing) which in this paper is equal to 1, k is the maximum number of lags and p_k^2 is the sample Auto Correlation Function ACF at lag k. LjungeBox Q-statistical test is tested under the null hypothesis that all residuals are independently distributed or has no auto-correlation as against the alternate hypothesis that residuals are serially correlated. The value of Q is then compared to critical $\square^2$ with k-degree of freedom. If Q < $\square^2_{1-\alpha,k}$ with k degrees of freedom and α is the selected level of significance, we fail to reject the null hypothesis and state that residuals are independently distributed. If Q > $\square^2_{1-\alpha,k}$, then the null hypothesis is rejected and state that residuals generated are not independently distributed. Lastly, the best selected ARIMA model is used for forecasting values of temperatures for ten years ahead.

6. Material and Methods: Empirical Analysis

6.1 Trend analysis of climatic variables using Autoregressive Integrated Moving Average
(ARIMA) analysis

In econometrics, forecasting helps in providing estimates of the variable in interest based on the past time series data. The estimation of the future/expected value of a dependent/endogenous variable on the basis of past and current data and information is extremely usefull concept. Generally, the linear regression techniques are used to forecast the dependent variable with the help of independent variables and then calculate the predicted values. ARIMA method is a linear forecasting technique that ignores independent variables in making forecasts and uses current and past values of the dependent variable to estimate short term forecasts of that variable. An example includes forecasts of stock market price predictions based on past patterns of movement of only the stock prices and ignores all other independent variables which could impact the dependent variable. ARIMA is used when little is known about the dependent variable being forecasted and the independent variables cannot be independently forecasted. The ARIMA approach combines two different processes into one equation: an autoregressive process (AR) which expresses a dependent variable as a function of past values of the dependent variable and a moving average process (MA) which expresses a dependent variable as a function of past values of the error term (Gujarati and Sangeetha, 2007). In an autoregressive process a dependent variable Y_t is a function of past values of the dependent variable as given below:

$$Y_t = f(Y_{t-1}, Y_{t-2}, \dots Y_{t-p}) \text{ or } Y_t = b_o + b_1 Y_{t-1} + b_2 Y_{t-2} \dots + b_p Y_{t-p} + \mu_t$$

where Y_t is the variable being forecasted and as there are p different lagged values of Y in this equation, it is referred to as a "pth order" or AR(p) autoregressive process (p is the number of past values used). Here μ_t is the white noise error term. The value of p is determined empirically by Akaike information criterion (AIC) (Akaike, 1974) used commonly in time series analysis. A moving-average process expresses a dependent variable Y_t as a function of past values of the error term, as given below:

$$Y_t = f(e_{t-1,} e_{t-2,} \dots e_{t-q,}) \text{ or}$$
$$Y_t = C_0 + C_1 e_t + C_2 e_{t-1} \dots + C_q e_{t-q}$$

where e_t is the error term associated with Y_t and q is the number of past values of the error term used. This function is a moving average of past error terms that can be added to the mean of Y to obtain a moving average of past values of Y_t. Such an equation is of qth order moving-average process MA (q). ARIMA means Auto Regressive Integrated Moving Average. Thus, in an ARMA (p, q) process, there are p autoregressive and q moving average terms. Collating the two we get an ARIMA (p, d, q) process which is called an autoregressive integrated moving average time series, where p denotes the number of autoregressive terms, d the number of times the series has to be differenced before it becomes stationary, and q the number of moving average terms. For e.g. the ARIMA (2, 1, 1) time series has to be differenced once (d = 1) before it becomes stationary and becomes fit for forecasting. In case of d = 0 (stationary series), ARIMA (p, d = 0, q) = ARMA (p, q). An ARIMA (p, 0, 0) process implies a purely AR (p) stationary process and an ARIMA (0, 0, q) implies a purely MA (q) stationary process. ARIMA model adds structural flexibility to the regression approach. The three main characteristics of ARIMA modeling technique useful for its applicability in this study are a) it takes into account the serial correlation which is the most important characteristic of time series data and is invariable present in the temperature and hydrological data b) the ARIMA model uses less parameter to describe a time series data and c) the ability of ARIMA to forecast results when the variable exhibits strong seasonal behavior (Brockwell and Davis, 2002). This makes it superior to other techniques like exponential smoothing and neural network when the data is of long duration and the correlation between past observations is stable (Mahmud et al. 2016). Minimum number of 50 observations is needed to build reasonable ARIMA model, which is reasonably correct in the present study (Wei, W. W. S. 1990). The data requirements for conducting an ARIMA analysis are far less than those required by other techniques such as (multiple regression, cointegration etc.) and is even more true for studying complex phenomena of climate change. The use of ARIMA technique is being increasingly employed by researchers and analyst to enhance their understanding of the time series phenomenon of climate change and climatic variables. Afrifa-Yamoah, 2015, has identified and tested ARIMA (1,0,0) (0,1,2) (12) to be an appropriate model for predicting monthly average surface temperature for Ghana for the time period 1975 to 2009. The suitability of using ARIMA technique for forecasting the mean surface temperature trend was tested by Tanja Van Hecke, 2010, for Belgium using data from 1850-1899. He came to the conclusion that ARIMA (1, 1, 0) is suitable to predict the temperature evolution in Belgium and that the temperature would rise over the reference period. To capture and highlight the underlying trend and to develop a model to forecast the global mean temperature for short term was estimated in a time series modeling of global mean temperature by Peter Romillay, 2005. He concluded that SARIMA and GARCH modeling techniques provides for highly robust results with most of its coefficients being statistically significant at the 1% level significance. Sultana and Hasan, 2015, have forecasted the temperature of the city of Chittagong by using SARIMA (1, 1, 1) (2, 0, 0) [12] for maximum temperature and SARIMA (1, 1, 1) (1, 0, 1) for minimum temperature. Partheepan et al., 2005, have used a bivariate ARIMA (1, 1, 0) and (0, 1, 1) using long-term 100 years rainfall and temperature data. They were able to forecast temperature accurately the long-term forecast for agri-environmental planning. Tularam and Mahbub, 2010, have used ARIMA models to show that rainfall and temperature vary with each with a cycle of about 2-3 years. For estimating global temperature trends, Ye Liming et al, 2013, used ARIMA technique and concluded that global temperature would increase at a rate of 0.12 degree centigrade per decade. Mishra and Desai, 2005, have used ARIMA modelling to forecast drought occurrence in the Kansabati river basin in the Purulia district of West Bengal. They showed that predicted values gradually fall with increase in lead time of two months which provides a suitable base to forecast droughts. Babu et al, 2015, have demonstrated the superiority of ARIMA over other forecasting models like ANIFIS (Adaptive Network Based Fuzzy Inference System).

7. Results and Interpretations

The Box–Jenkins model is based on the following steps process:

*(i) Model identification and Parameter estimation***:** Firstly, the ARIMA (p, d, q) values orders are chosen on the basis of minimum Akaike information criterion (AIC) or Bayesian information criterion (BIC) (Akaike 1974).

(ii) Validation To check for the significance the selected models, diagnostic checking models are used to select the best model. This implies selecting the model on the basis of residuals (i.e., errors between the actual and predicted values) depicting no autocorrelation and are random in nature.
(iii) Forecasting Finally, the best fit model is utilized for prediction and the accuracy of prediction is checked by forecasting errors. If the results are found to be unacceptable, model parameters are estimated again and the entire process is reiterated.

7.1 Selection of lag length

(i) Annual Mean Minimum Temperature

The first step before carrying out unit root test to test stationarity of the data is to determine the optimum lag length. The general rule is to select the optimum lag length on the basis of the information obtained about the following variables: final prediction error (FPE), Akaike information criterion (AIC), Hannan and Quinn information criterion (HQIC) and Schwarz Bayesian information criterion (SBIC). Out of these four variables, the tendency amongst researchers is to use the lowest values of AIC and BIC more frequently. The model with the lowest AIC and BIC is selected. Minimum value of FPE is required to minimize the prediction error to select the optimum lag length.

Table 3: Lag Selection: Annual Mean Minimum Temperature

Lag	**LL**	**LR**	**dF**	**P**	**FPE**	**AIC**	**HQIC**	**SBIC**
0	-86.0165				.249673	1.45028	1.45971	1.4735
1	-30.323	111.39	1	0.000	.100343	.538717	.557584	.585175
2	-28.6244	3.3972	1	0.065	.099183	.527073	.555373	.59676
3	-21.5367	14.175*	1	0.000	.089614*	.425612*	.463346*	.518529*
4	-21.2798	.51378	1	0.474	.090733	.437997	.463346*	.554143

Source: calculated by authors

The likelihood-ratio (LR) test statistics works on the following rule: maximize the likelihood-ratio or minimize the negative of the log-likelihood, The estimated results are shown in the following table 3. The '*' sign next to the estimated statistics indicates the optimal lag length.
In the given table 3 for Annual Mean Minimum Temperature the optimum lag length is 3. All the relevant variables like FPE, AIC and HQIC and SBIC have chosen a model with 3 lags.

Table 4: Lag Selection: Annual Mean Maximum Temperature

Lag	LL	LR	dF	P	FPE	AIC	HQIC	SBIC
0	278.22				.000577	-4.62033	-4.6109	-4.5971
1	364.446	172.45	1	0.000	.000139	-6.04077	-6.0219	-5.99431

2	367.885	6.8785	1	0.009	.000134	-6.08142	-6.05312	-6.01173
3	370.717	5.663	1	0.017	.00013	-6.11195	-6.07421	-6.01903*
4	372.672	3.9099*	1	0.048	.000128*	-6.12786*	-6.0807*	-6.01172

Source: calculated by authors

In the given table 4 for Annual Mean Minimum Temperature the optimum lag length is 4. All the relevant variables like FPE, AIC and HQIC and SBIC have chosen a model with 4 lags.

7.2 Unit Cointegration Test

In the next step, the stationarity of time series data for the two variables is verified. The results of unit root test for the two temperature variables are given in table 5. The result of the Augmented Dicky Fuller Test indicate that both the variables have unit root at the levels.

Table 5: Annual Mean Minimum Temperature

	Level Z(t)	**1st Difference Z(t)**
Test Statistics	-1.954	-8.231
1% Critical Value	-4.033	-4.033
5% Critical Value	-3.447	-3.447
10% Critical Value	-3.147	-3.147
MacKinnon approximate p-value for Z(t)	0.6261	0.0000

Source: calculated by authors

The test statistics of -1.954 (absolute) is less than the critical values at 1%, 5% and 10% significance level. MacKinnon approximate p-value is also 0.6241 which is greater than 0.05 thus validating the null hypothesis of unit root or non-stationarity. After differencing the series, the values obtained of test statistics of -8.231 (absolute) is greater than all the respective critical values. MacKinnon approximate p-value of 0.0000 is also less than 0.05 which further corroborate the fact that differenced series is stationary.

Table 6: Annual Mean Maximum Temperature

	Level Z(t)	**1st Difference Z(t)**
Test Statistics	-2.355	-8.687
1% Critical Value	-4.033	-3.504
5% Critical Value	-3.447	-2.889
10% Critical Value	-3.147	-2.579

MacKinnon approximate p-value for Z(t)	0.4040	0.0000

Source: calculated by authors

In table no. 6, the test statistics of -2.355 (absolute) is less than the critical values at 1%, 5% and 10% significance level. MacKinnon approximate p-value is also 0.4040 which is greater than 0.05 thus validating the null hypothesis of unit root or non-stationarity. After differencing the series, the values obtained of test statistics of -8.687 (absolute) is greater than all the respective critical values. MacKinnon approximate p-value of 0.0000 is also less than 0.05 which further corroborate the fact that differenced series is stationary. This implies that both the temperature variables namely, Annual Mean Minimum Temperature and Annual Mean Maximum Temperature are non-stationary at level i.e. both series are integrated of order zero I (1).

7.3 ARIMA regression

The ARIMA modelling process in the Box Jenkins Methodology is an iterative process. The different models are compared according to the minimum values of AIC, BIC and maximum values of Log likelihood values. These two most widely used model for the selection criteria in literature are the Akaike Information Criterion (AIC) and the Bayesian Information Criterion (BIC). They are defined as follows:

AIC = T ln(sum of squared residuals) + 2n

BIC = T ln(sum of squared residuals) + n ln(T)

where n = number of parameters estimated (p + q + possible constant term)

T = number of usable observations.

Both the variables should as small as possible and they could be in the negative zone also and in can approach $-\infty$ as the model fit improves.

7.4 Identification and Estimation of ARIMA models

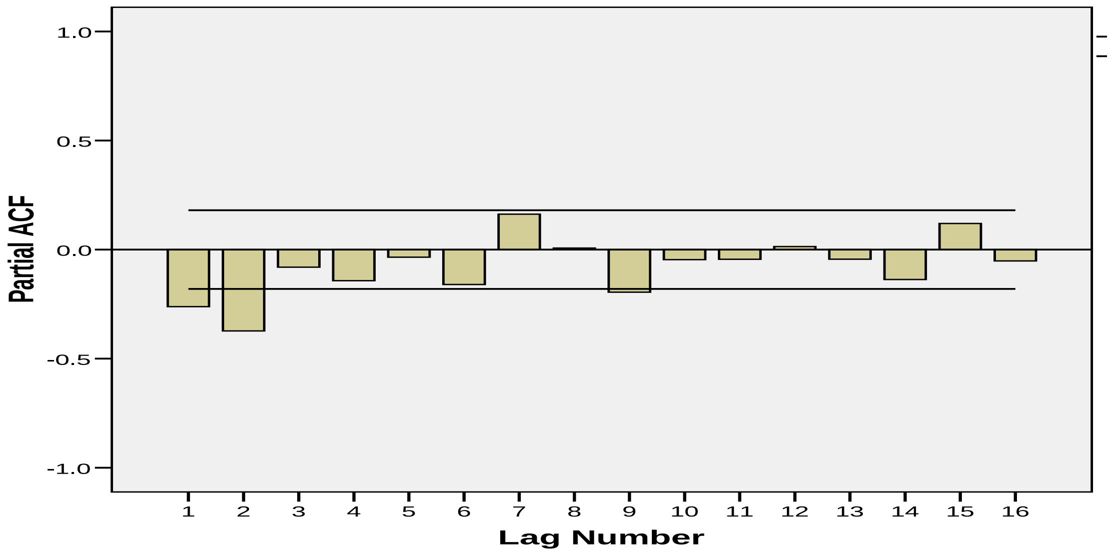
Table 7: Annual Mean Minimum Temperature

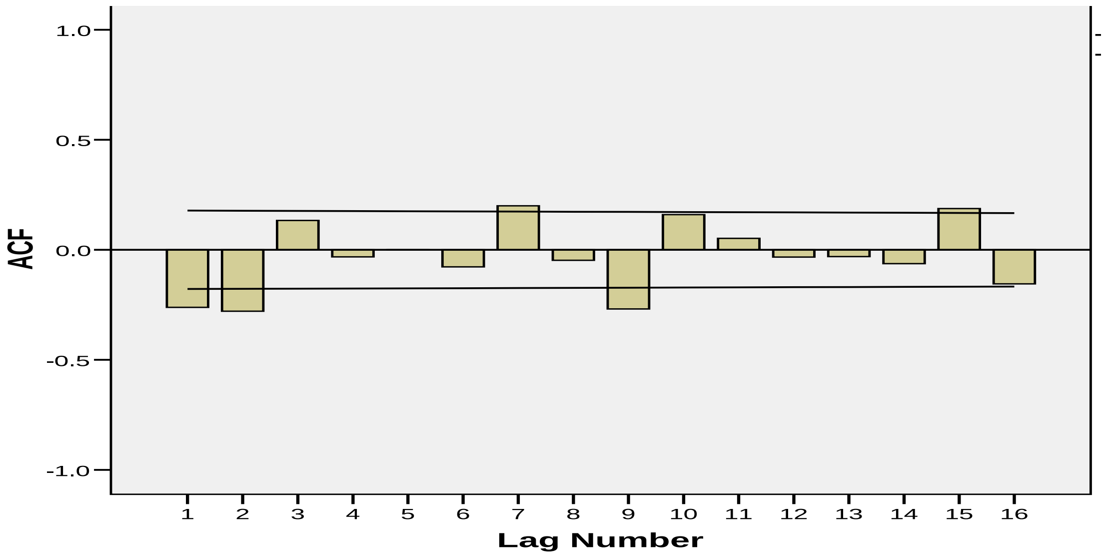
ARIMA (p, d, q)	**AIC**	**BIC**	**Log likelihood**
(2, 1, 2)	48.91132	62.73572	-18.45566
(1, 1, 2) *	47.34781	58.86814	-18.67391
(3, 1, 2	50.66083	66.78929	-18.33041
(2, 1, 1)	48.25766	59.77799	-19.12883
(1, 1, 1)	50.63305	59.84931	-21.31652

Source: calculated by authors

In table 7, the ARIMA (1, 1, 2) denoted by * is chosen as the model for the variable Annual Mean Minimum Temperature for this paper. This model has the lowest values of AIC and BIC and highest values of log likelihood ratio of all the possible values for the temperature variable.

Fig. 5: ACF and PCAF plot for **Annual Mean Minimum Temperature**





In fig. 5, the graphs ACF of the mean minimum annual temperature clearly shows that Lag 1 and Lag 2 are significantly negative and then decays gradually but showing no significant pattern which irrefutably indicates an MA (q) process with probable of 2. This is also corroborated by the PACF plot showing significant lags. The sharp cutoff after lag 2 in PACF also indicates an AR process with likely value of 2. Thus, both AR and MA processes appear possible strongly indicating the possible values of ARIMA (2,1,0), ARIMA (1,1,2) ARIMA (2,1,2). The ACF and PCAF plots of the residuals in figures 5&6 also show no significant autocorrelations. This implies that the residuals of the selected model ARIMA (1, 1, 2) for the annual mean minimum temperature are random or white noise.

Table 8: Annual Mean Maximum Temperature

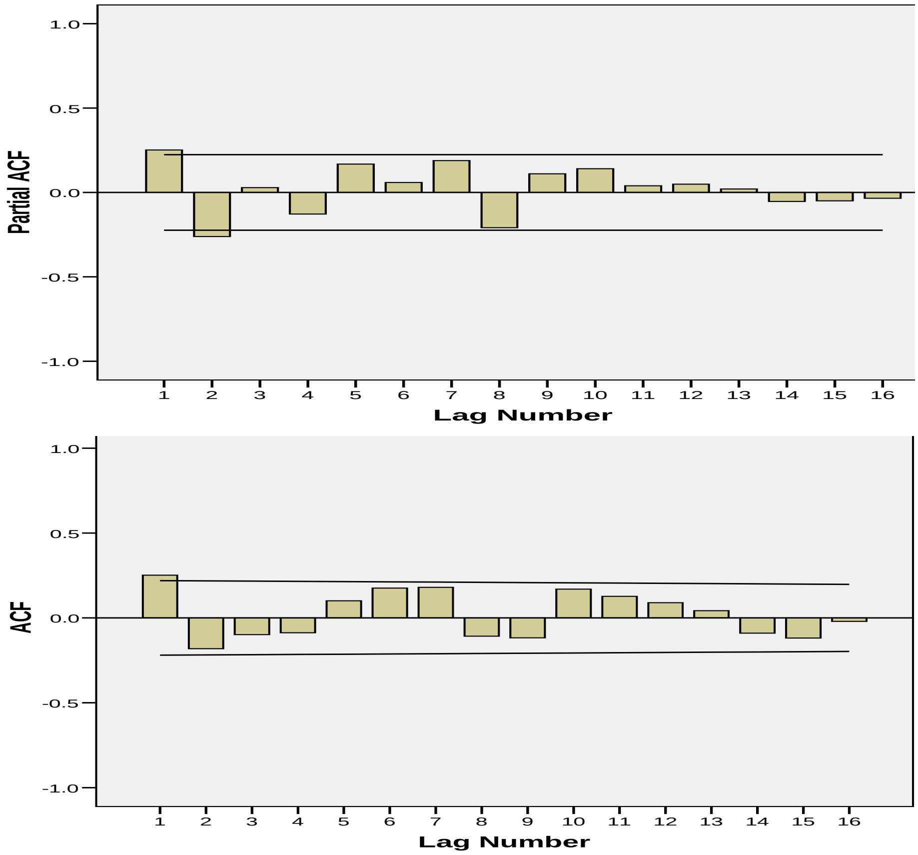
ARIMA (p, d, q)	AIC	BIC	Log likelihood

(2, 1, 2)	60.00905	73.83344	-24.00453
(1, 1, 2)	60.77214	72.29246	-25.38607
(3, 1 2)	59.38321	75.51167	-22.69161
(2, 1, 1)	60.44406	71.96439	-25.22203
(1, 1, 1)*	59.5687	68.78496	-25.78435

Source: calculated by authors

In table no. 8, the ARIMA (1, 1, 1) denoted by (*) is chosen as the model for the variable Annual Mean Maximum Temperature for this paper. This model has the lowest values of BIC, second lowest values of AIC and highest values of log likelihood ratio of all the possible values for the temperature variable. The model (1, 1, 1)* is chosen as two out of three parameters are for it.

Fig. 6: ACF and PCAF plot for **Annual Mean Maximum Temperature**



In fig. 6, the ACF Plot of the annual mean maximum temperature shows a definite positive spike at lag 1 which is followed by alternating small positive and negative values which then gradually decays towards zero. This pattern of gradual decay positively indicates the presence of an AR (autoregressive) process. The PACF plot also shows a positive spike at lag 1 and a negative spike at lag 2. Accordingly, ARIMA (1, 1, 2) is chosen is the chosen model for Annual Mean Minimum Temperature and

ARIMA (1, 1, 1) is chosen is the chosen model for Annual Mean Maximum Temperature.

7.5 Validation of the Annual Mean Minimum Temperature variable

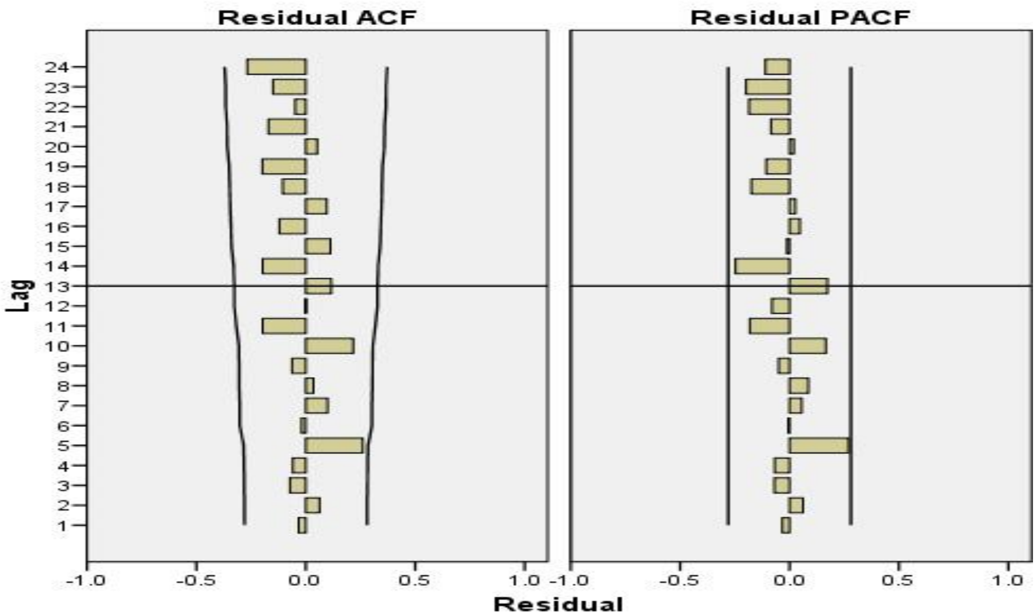
Table 9: Annual Mean Minimum Temperature

Test	Model	Test Statistics	p-value
Jarque-Bera	ARIMA (1, 1, 2)	1.56	0.4590
Ljung-Box	ARIMA (1, 1, 2)	12.892	0.611

Source: calculated by authors

The Ljung-Box test is based on the Null hypothesis that the residuals are distributed independently implying that there is no autocorrelation in the variable. In table no.9, the ARIMA regression analysis gave the Ljung-Box test statistic chi-square value of 12.892 with a corresponding *p-value* of 0.611 which clearly confirms that there are no autocorrelations present in the residual of the ARIMA model. Thus, we fail to reject the null hypothesis and the null hypothesis of no autocorrelation in the residual can be accepted. The Jarque–Bera (JB) test is a goodness-of-fit test used to confirm whether the data used in the study confirms to a normal distribution. This test is based on the Null Hypothesis which states that data sample comes from a normal distribution. Since if the p-value is greater than 0.05, we can accept the null hypothesis confirming normalcy of the sample data set.

Fig. 7: Residual plots of ACF and PACF by using Box-Jenkins Method



The residual plots of ACF and PACF in fig.7 shows the model is adequate as it shows a random variation from the origin zero (0), the points below and above are all uneven, hence the model fitted is adequate

7.6 Validation of the Annual Mean Maximum Temperature variable

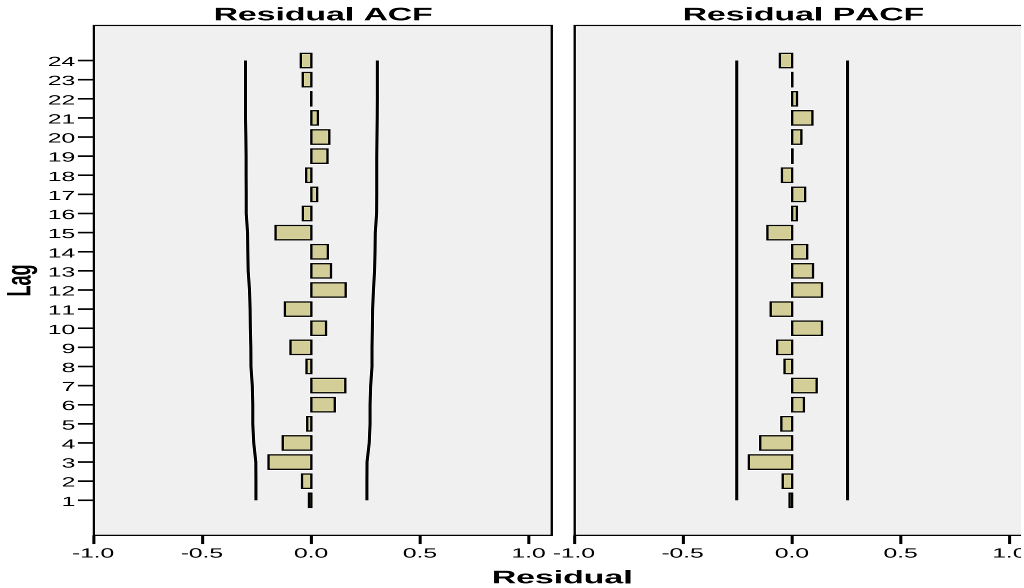
Table 10: Annual Mean Maximum Temperature

Test	Model	Test Statistics	p-value
Jarque-Bera	ARIMA (1, 1, 1)	3.61	0.1641
Ljung-Box	ARIMA (1, 1, 1)	7.707	0.957

Source: calculated by authors

In table no. 10, the ARIMA regression gave the Ljung-Box test statistic chi-square value of 7.707 with a corresponding *p-value* of 0.957 which again clearly confirms to the fact that that there are no autocorrelations present in the residual of the ARIMA model. This implies that the null hypothesis of no autocorrelation in the residual can be accepted. The Jarque–Bera (JB) test is a goodness-of-fit test used to confirm whether the data used confirms to a normal distribution. This test is based on the Null Hypothesis which states that data sample comes from a normal distribution. As in the above case, since the p-value is greater than 0.05, we can accept the null hypothesis confirming normalcy of the sample data set.

Fig. 8: Residual plots of ACF and PACF by using Box-Jenkins Method



In fig. no. 8, the residual ACF (Autocorrelation Function) and PACF (Partial Autocorrelation Function) plots are employed to understand the relationship between a variable and its lagged values over two points of time. If the residuals of a ARIMA model behaves like white noise, then all the autocorrelations should lie within the 95% confidence bands or between the two horizontal lines. As

7.7 Forecast of temperature values

Table 11: Annual Mean Maximum Temperature

Year	2025	2026	2027	2028	2029	2030	2031	2032	2033	2034	2035

Forecast	31.19	31.19	31.2	31.22	31.24	31.26	31.28	31.3	31.32	31.34	31.64

Source: calculated by authors

Fig. 9: Annual Mean Maximum Temperature

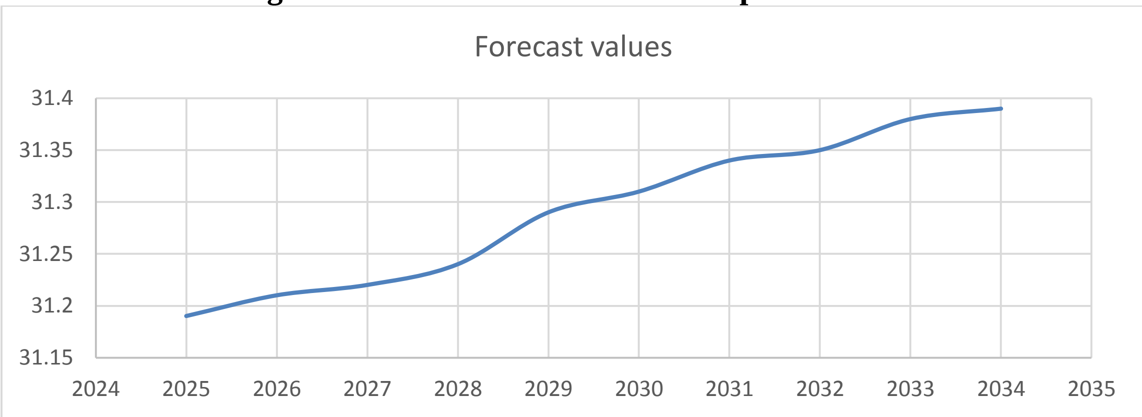
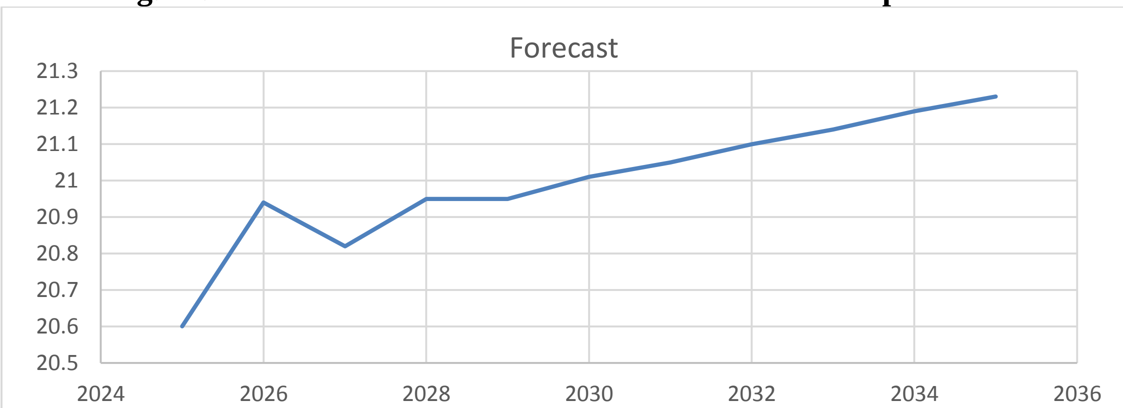


Table 12: Forecast values for Annual Mean Minimum Temperature

Year	2025	2026	2027	2028	2029	2030	2031	2032	2033	2034
Forecast	20.6	20.94	20.82	20.95	20.95	21.01	21.05	21.1	21.14	21.19

Fig. 10: Forecast values for Annual Mean Minimum Temperature



Source: calculated by authors

The table no. 11 and 12 clearly shows that in the period from 1901-2024, the Annual Mean Minimum Temperature has increased by 3.75% or 0.58% per year while Annual Mean Maximum Temperature increased by 7.9% or 1.8% per year. In the next ten years from 2025-2034, the Annual Mean Minimum Temperature as per the Box -Jenkins methodology employed in this study is expected to increase from 20.6 to 21.19 degree centigrade. This translates into 0.59^{o}C or 2.89% rise during the next ten years. Yearly increase is 5.9% per year. In case of Annual Mean Maximum Temperature, the increase is from 31.19 to 31.64 ^{o}C which implies an annual increment of 4.5% per annum. The consistent yearly increase in both the temperatures is leading to extreme weather as witnessed recently in different parts of the India at the global level. The worrying finding is that the minimum mean temperature is increasing more than maximum mean temperature across the Indian sub-continent. The data for initial period of 2026 indicates that while the maximum temperature are within normal range, the minimum temperature were above normal by almost 0.5 degree centigrade. Thia has wider implications for power demand and specially for agricultural sector.

Table 13: Comparison of Temperature change in the past 125 years and prediction for next 10 years

Period	% change	
	Annual Mean Minimum Temperature	Annual Mean Maximum Temperature
1901-2024	+3.75%	+7.9%
2000-2024	+3.91%	+0.5%
2025-2034 (prediction as per this study)	+2.8%	+1.4%

Source: calculated by author

In table no. 13, the trend as witnessed in the past years is reinforced for the next ten years. While the maximum temperature rose by 7.4% in the period from 1901 to 2024, the minimum temperature rose by 3.75% in comparison. However, this trend changed dramatically in the last 30 years. While Annual Mean Maximum Temperature increased by 0.5%, the Annual Mean Minimum Temperature rose by 3.91%.

Conclusions

This uneven increase in the two temperatures is one of the leading cause of unexpected climatic changes because the night does not cool down and the heat residue is carried over to the next day adding to the day temperature. More than two lakh people have died in Europe in the last four year due to extreme summer heat. The positive temperature anomalies are even more indicative of the average temperature are greater than the baseline benchmark average. The major worry is that while before 1980 the anomalies varied between positive and negative range. It has stayed into positive territory since 1980 putting pressure on the increasing overall temperature. The night time anomalies have stayed higher than the daytime anomalies thereby reducing the gap between them. The increase in minimum temperature implies warmer nights which is aggravating the heat stress of the daytime. This is leading to warmer winters and hotter summer. The future forecast of both the minimum and maximum temperature are clearly showing an upward trend in this paper.

References

[1] National Oceanic and Atmospheric Administration, National Centers for Environmental Information. (2025). *Monthly global climate report for annual 2024*. Retrieved July 21, 2025, from

https://www.ncei.noaa.gov/access/monitoring/monthly-report/global/202413

[2] Intergovernmental Panel on Climate Change. (2007). *Climate change 2007: Impacts, adaptation and vulnerability*. Contribution of Working Group II to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change (M. L. Parry, O. F. Canziani, J. P. Palutikof, P. J. van der Linden, & C. E. Hanson, Eds.). Cambridge University Press.

[3] Intergovernmental Panel on Climate Change. (2014). *Climate change 2014: Impacts, adaptation, and vulnerability*. Cambridge University Press.

[4] Ravindra, Khaiwal et al, 2024., Bhardwaj, S., Ram, C., Goyal, A., Singh, V., Venkataraman, C., Bhan, S. C., Sokhi, R. S., & Mor, S. (2024). Temperature projections and heatwave attribution scenarios over India: A systematic review. *Heliyon*, *10*(4), e26431. https://doi.org/10.1016/j.heliyon.2024.e26431

[5] Nanditha, J. S., et al. (2020). Temperature analysis over India. *Environmental Research Letters*, *15*, 044028.

[6] Arora, M., Goel, N. K., & Singh, P. (2005). Evaluation of temperature trends over India. *Hydrological Sciences Journal*, *50*(1), 81–93. https://doi.org/10.1623/hysj.50.1.81.56330

[7] Vose, R. S., Easterling, D. R., & Gleason, B. (2005). Maximum and minimum temperature trends for the globe: An update through 2004. *Geophysical Research Letters*, *32*, L23822. https://doi.org/10.1029/2005GL024379

[8] India Meteorological Department. (2024). *Annual climate summary 2024*.

[9] Intergovernmental Panel on Climate Change. (2014). *Climate change 2014: Impacts, adaptation, and vulnerability*. Cambridge University Press.

[10] Stern, N. (2006). *The economics of climate change: The Stern review*. Cambridge University Press.

[11] Hallegatte, S. (2007). Do current assessments underestimate future damages from climate change? *World Economics*, *8*(3), 131–146.

[12] Trzaska, S., & Schnarr, E. (2014). *A review of downscaling methods for climate change projections*. Tetra Tech ARD.

[13] Ahmed, K. (2011). *Bias correction and downscaling of climate model outputs required for impact assessments of climate change in the U.S. Northeast*. University of Connecticut Graduate School.

[14] Estrada, F., Perron, P., & Martínez-López, B. (2013). Statistically derived contributions of diverse human influences to twentieth-century temperature changes. *Nature Geoscience*, *6*, 1050–1055.

[15] Gay, C., Estrada, F., & Sanchez, A. (2009). Global and hemispheric temperature revisited. *Climatic Change, 94,* 333–349. https://doi.org/10.1007/s10584-008-9524-8

[16] Chujai, P., Kerdprasop, N., & Kerdprasop, K. (2013). Time series analysis of household electric consumption with ARIMA and ARMA models. *Lecture Notes in Engineering and Computer Science, 2202*, 295–300.

[17] Contreras, J., Espínola, R., Nogales, F., & Conejo, A. (2003). ARIMA models to predict next-day electricity prices. *IEEE Transactions on Power Systems, 18*(3), 1014–1020. https://doi.org/10.1109/TPWRS.2002.804943

[18] Suresh, K., & Priya, K. (2011). Forecasting sugarcane yield of Tamil Nadu using ARIMA models. *Sugar Tech, 13,* 23–26. https://doi.org/10.1007/s12355-011-0071-7

[19] Almasarweh, M., & Alwadi, S. (2018). ARIMA model in predicting banking stock market data. *Modern Applied Science, 12*(11), 309. https://doi.org/10.5539/mas.v12n11p309

[20] Chaudhuri, S., & Dutta, D. (2014). Mann–Kendall trend of pollutants, temperature and humidity over an urban station of India with forecast verification using different ARIMA models. *Environmental Monitoring and Assessment*. https://doi.org/10.1007/s10661-014-3733-6

[21] Sultana, N., & Hasan, M. M. (2015). Forecasting temperature in the coastal area of Bay of Bengal: An application of Box–Jenkins seasonal ARIMA model. *Civil and Environmental Research, 7*(8).

[22] Partheepan, K., & Jeyakumar, P. (2006). Development of a time-series model to forecast climatic data in the Batticaloa District, Sri Lanka. *Water Professionals' Day Symposium*.

[23] Tularam, A., & Ilahee, M. (2010). Time series analysis of rainfall and temperature interactions in coastal catchments. *Journal of Mathematics and Statistics, 6*, 372–380.

[24] Ye, L. M., Yang, G. X., Van Ranst, E., & Tang, H. J. (2013). Time-series modeling and prediction of global monthly absolute temperature for environmental decision making. *Advances in Atmospheric Sciences, 30*(2), 382–396. https://doi.org/10.1007/s00376-012-1252-3

[25] Babu et al. (2015). Comparison of ANFIS and ARIMA model for weather forecasting. *Indian Journal of Science and Technology, 8*(S2), 70–73.

[26] Mishra, A., & Desai, V. (2005). Drought forecasting using stochastic models. *Stochastic Environmental Research and Risk Assessment, 19,* 326–339. https://doi.org/10.1007/s00477-005-0238-4

[27] Narendra Babu, B. Eswara Reddy. (2014). A moving-average filter based hybrid ARIMA–ANN model for forecasting time series data, Applied Soft

Computing, Volume 23, Pages 27-38, ISSN 1568-4946, https://doi.org/10.1016/j.asoc.2014.05.028.

[28] Box, G. E. P., Jenkins, G. M., Reinsel, G. C., & Ljung, G. M. (2015). *Time series analysis: Forecasting and control*. Wiley.

[29] Gujarati, D. and Sangeetha, N. (2007). Basic Econometrics. Fourth Edition, Tata McGraw-Hill, New Delhi.

[30] Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control, 19*(6), 716–723.

[31] Brockwell, P. J. and Davis R. A. 2002. Introduction to time series and forecasting. Springer texts in statistics.

[32] Mahmud, I. et al. 2016. Monthly rainfall forecast of Bangladesh using autoregressive integrated moving average method. Environmental Engineering Research; 22(2): 162-168

[33] Wei, W. W. S. 1990. Time series analysis. Addison-Wesley Publishing, Reading, MA

[34] Afrifa-Yamoah E. 2015. Application of ARIMA Models in Forecasting Monthly Average Surface Temperature of Brong Ahafo Region of Ghana, International Journal of Statistics and Applications, Vol. 5 No. 5, pp. 237-246

[35] Van Hecke, Tanja. 2010. Time series analysis to forecast temperature change. The Mathematical Scientist, 35

[36] Romilly, P. 2005. Time series modelling of global mean temperature for managerial decision-making. J Environ Manage. 2005 Jul;76(1):61-70. doi: 10.1016/j.jenvman.2005.01.008. PMID: 15854737.

[37] Sultana, N. and Hasan, M. M. 2015. Forecasting Temperature in the Coastal Area of Bay of Bengal-An Application of Box-Jenkins Seasonal ARIMA Model. Civil and Environmental Research. ISSN 2224-5790 (Paper) ISSN 2225-0514 (Online), Vol.7, No.8, 2015

[38] Partheepan et al., 2005. Development of a Time-Series Model to Forecast Climatic Data in the Batticaloa District, Sri Lanka. Water Professionals' Day Smposium, Octuber, 2005

[39] Tularam, A. and Ilahee, M. 2010. Time Series Analysis of Rainfall and Temperature Interactions in Coastal Catchments. Journal of Mathematics and Statistics. 6. 10.3844/jmssp.2010.372.380.

[40] Mishra, A. and Desai, V. 2005. Drought forecasting using stochastic models. Stochastic Environmental Research and Risk Assessment. 19. 326-339. 10.1007/s00477-005-0238-4.

[41] Babu, R., Babu, C., Reddy, D. and Gowtham, M. 2015. Comparison of ANFIS and ARIMA Model for Weather Forecasting. Indian Journal of Science and Technology, 8(S2), pp. 70–73, 2015.