Journal of Climate Change, Vol. 8, No. 3 (2022), pp. 31-39.
DOI 10.3233/JCC220019

Forecast Modeling and Trend Analysis of Temperature in India

Mamta Meena
Department of Geography, University of Rajasthan, Jaipur, India
✉ mamtamanutal@gmail.com

Received June 15, 2022; revised and accepted June 22, 2022

Abstract: Since historical times, the climate of the earth has been changing. In terms of diversity and population, India is a vast country. Any significant change in its climatic conditions can have an impact on the country's entire economic and cultural structure. Temperature is the most important variable when focussing on climate change study. This study aims to recognise the changing trend of temperature in India. For this purpose, temperature data from 1901 to 2020 has been analysed. To find out the variability in temperature, various statistical tools were applied. The Mann-Kendall method has been used along with Sen's slope estimator for determining the trend and slope magnitude. ARIMA modelling is applied to the data series to forecast the temperature trend. The study has shown that variability in temperature is increasing in recent years and the temperature trend is increasing in all the months, especially in the cold months of November, December and January. The average annual temperature of India in the year has increased significantly in the past 120 years. The trend analysis of temperature is tested significantly statistically at a 95% confidence level.

Keywords: Temperature; Variability; Mann-Kendall Test; ARIMA modelling; Sen's slope estimator.

Introduction

Before the industrial revolution, climate change occurred due to natural aspects; however, during the latter half of the 20th century, global warming accelerated and became more severe as a result of human activities. The planet's temperature has risen by about 1.62 degree Fahrenheit (0.9 degrees Celsius) since the 19th century. The current warming trend has a special significance because of extreme weather phenomena occurring throughout the world. These climatic events are increasing at an unprecedented rate since the mid-20th century. Temperature and precipitation are basic components for the assessment of climate change. The variability in temperature and climate change has a considerable impact on agriculture, ecosystem, health and economic activities. At present, the whole world community is facing extreme climatic phenomena.

India being a diverse country is at higher risk of climatic extreme events. Thus, understanding the variability and the prediction of weather on the basis of time series is significant for future adaptation strategies (Muthoni et al., 2018). The study of time series involves the analysis of trends over the period and looking into the variability in the data. Trends are defined as the long-term decrease or increase in the time series. On surveying the literature study, temperature has been found to be a crucial variable to study the climate. Nasher et al. (2013) studied the temperature trends of two important cities of Bangladesh, the coastal city of Chittagong and Rajshahi. In Chittagong city, an increasing trend has been observed post-monsoon, highest maximum temperature and no significant decline in trend were found in the city. In Rajshahi city, falling trends are been found in the annual highest maximum temperature and lowest minimum temperature. The

climate variability in Rajshahi city is more evident compared to Chittagong city. The climate of Chittagong is becoming warmer while the climate of Rajshahi city is more prone to extreme climate events. Sultana and Hasan (2015) conducted a study on forecasting temperature in the coastal areas of the Bay of Bengal by using the ARIMA modelling method. The Mann-Kendall model was used for testing the temperature time series data. It indicates the increasing trend in maximum temperature and minimum temperature but minimum increase significantly. The results observed an increasing pattern of temperature on upcoming days in the region. In one study, Dash and Hunt (2017) have suggested trends of increasing temperature over the past century. This also shows the variation during different seasons and over different regions of India. Likewise, the present study also discussed the variation in minimum temperature in north and south India. The study also discussed how airborne dust on clouds and precipitation impact climate change's effect on agriculture and human. Mudelsee (2018) writes in his research article, "Trend analysis of climate time series: A review of methods" that statistical trend estimation methods have advanced to include change-points, faster increases, nonlinear behaviour, and non-parametric descriptions. Simulation algorithms account for the non-Gaussian distributional shape and autocorrelation of climate data. Monte Carlo simulation methods have demonstrated the dependability of such computer-age statistical methods. This shows that inferior data quality and wrong selection of fit intervals show errors in the results and show false peaks. Alemu and Dioha (2020) in their study, i.e., "Climate change and trend analysis of temperature: The case of Addis Ababa, Ethiopia" applied the MK test and Sen's Slope estimate to find out climate variations. On the basis of climatic variations and geographical locations, they choose two stations: Bole and Entoto in Addis Ababa. The station Bole shows an increase in temperature, which can cause extreme weather events in the capital city.

This study analyses the trends and variability in temperature data in India over the period 1901-2020 using various statistical tools. Another objective is the prediction of temperature change through forecast modelling of temperature data by using the ARIMA method.

Study Area

India is the most diverse country in terms of climate. Various parts of the country are observed to show different climatic occurrences. Stamp (1958) says that "India is basically a tropical country although its northern part is situated in the temperate belt". (2014) suggested that "India is par excellence, a tropical monsoon country". The recent developments in global climate due to global warming have adverse impacts on the country. A major part of the population depends on agriculture. It is said that the Indian budget is a gamble during the monsoon. We can say that monsoon is the pivot upon which the economic life of a country like India swings. Now irrigation facilities have a better expansion, however, still in many parts of our country, agriculture is dependent on rainfall and temperature. So studies related to climate change have greater relevance.

India lies to the north of the equator between 8 degrees 4 minutes and 37 degrees 6 minutes north latitude and 68 degrees 7 minutes and 97 degrees 25 minutes east longitude. The Tropic of Cancer 23°30′ N divides India into almost two halves. It has a length measuring 3,287,263 sq. km of area. It has a monsoon climate. Marsden said that the climate of India is diverse that can also be found in different regions around the world. If we talk about the state of Rajasthan, in the month of June, the day temperature in some districts may soar from 48°C to 50°C and at the very same time in Kashmir, the temperature may hardly reach 22°C. In the month of December, the night temperature at Dras may drop to minus 40°C and at the very same time, it may be as high as 22°C in Chennai. In the South, the Indian coasts are washed away by the Indian Ocean showing a typical tropical climate. The northern boundaries of India are bordered by the Himalayas, which protect India from the cold masses of central Asia. After a state-level analysis of the Indian climate using the data of temperature time series, Mishra and Mishra (2017) found that the temperature has witnessed a significant rise over almost all parts of the country.

Data and Research Methodology

Temperature data for the study was obtained from the Indian Meteorological Department (IMD) on a monthly basis. For calculating the temperature, the deviation formula used is (Actual temperature-Normal temperature)/Normal temperature)*100, which expresses a percentage deviation from the normal. The normal temperature is the long-term average of the actual temperature. The average, standard deviation and coefficient of variance were also calculated for analysis. The estimation of the trend in the temperature changes is done on the basis of the Mann-Kendall

methods. According to this test, the null hypothesis H0 assumes that there is no trend (the data is independent and randomly ordered), and this is tested against the alternative hypothesis H1, which assumes that there is a trend. For calculating the statistical data, Mann-Kendall test Addinsoft's XLSTAT 2018 software is used.

The computational procedure for the Mann-Kendall test considers the time series of n data points and T_i and T_j as two subsets of data where i = 1, 2, 3, …, n–1 and $j = i$+1, i+2, i+3, …, n. The data values are evaluated as an ordered time series. Each data value is compared with all subsequent data values. If a data value from a later time period is higher than a data value from an earlier time period, the statistic *S is incremented by 1.* On the other hand, if the data value from a later time period is lower than a data value sampled earlier, *S is decremented by 1. The net result of all such increments and* decrements yields the final value of *S.* The Mann-Kendall S Statistic is computed as follows: where T_j and T_i are the annual values in years j and i, $j > i$, respectively.

$$S = \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} \text{sign}(T_j - T_i)$$

$$\text{sign}(T_j - T_i) = \begin{cases} 1 \; if \; T_j - T_i > 0 \\ 0 \; if \; T_j - T_i = 0 \\ -1 \; if \; T_j - T_i < 0 \end{cases}$$

If n < 10, the value of $|S|$ is compared directly to the theoretical distribution of S as derived by Mann and Kendall. The two-tailed test is used. At a certain probability level, H0 is rejected in favour of H1 if the absolute value of S equals or exceeds a specified value $S\alpha/2$, where $S\alpha/2$ is the smallest S, which has the probability less than α/2 to appear in case of no trend. A positive (negative) value of S indicates an upward (downward) trend.

For $n \geq 10$, the statistic S is approximately normally distributed with the mean and variance as follows:

$$E(S) = 0$$

The variance (σ^2) for the S-statistic is defined by:

$$(\sigma^2) = \frac{n(n-1)(2n+5) - \Sigma t_i(i)(i-1)(2i+5)}{18}$$

In time series analysis, it is essential to consider autocorrelation or serial correlation, defined as the correlation of a variable with itself over successive time intervals, prior to testing for trends. Autocorrelation increases the chances of detecting significant trends even if they are absent and vice versa. Hamed and Rao (1998) suggest a modified Mann-Kendall test, which calculates the autocorrelation between the ranks of the data after removing the apparent trend. The adjusted variance is given by:

$$\text{Var}[S] = \frac{1}{18}[N(N-1)(2N+5)]\frac{N}{NS'}$$

$$\text{Where } \frac{N}{NS'} = 1 + \frac{2}{N(N-1)(N-2)}$$

$$\sum_{i=1}^{p} (N-i)(N-i-1)(N-i-2)Ps(i)$$

where N is the number of observations in the sample, NS^* is the effective number of observations to account for autocorrelation in the data, ps (i) is the autocorrelation between ranks of the observations for lag i, and p is the maximum time lag under consideration. P value is compared with alpha.

Kendall's Tau

Another statistic obtained on running the Mann-Kendall test is Kendall's tau, which is a measure of correlation and therefore measures the strength of the relationship between the two variables. Kendall's tau, like Spearman's rank correlation, is carried out on the ranks of the data. That is, for each variable separately, the values are put in order and numbered, 1 for the lowest value, 2 for the next lowest and so on. In common with other measures of correlation, Kendall's Tau will take values between ±1 and +1, with a positive correlation indicating that the ranks of both variables increase together whilst a negative correlation indicates that as the rank of one variable increases, the other decreases.

Sen's Slope Estimation Test

The altering trend of temperature for each period is assessed by using Sen's method. The slope of all data is prepared as:

$$T_i = \frac{X_j - X_k}{j - k} \text{ for i = 1,2, … … … } N$$

In this X_j and X_k are considered as the data values at times j and k ($j > k$) correspondingly. The median of these N values of T_i is represented as Sen's estimator of slope that is given below:

$$Q_i = \begin{cases} T_{\frac{N-1}{2}} & N \text{ is odd} \\ \frac{1}{2}\left(T_{\frac{N}{2}} + T_{\frac{N+2}{2}} \right) & N \text{ is even} \end{cases}$$

In Sen's slope estimator a positive value of Q_i indicates an upward or increasing trend and a negative value of Q_i gives a declining trend in the time series.

ARIMA Model is also used for forecasting the time series data. ARIMA stands for the autoregressive integrated moving average and is a statistical analysis model. It is used to better understand the temperature time series data set and to predict future trends. This autoregressive statistical model predicts future values based on past values.

Results and Discussion

M.K. test is commonly used in the analysis of climatic time series because it does not follow the normal distribution. Thus, the MK trend test and Sen's slope estimate were used in this study to better understand the nature of the temperature trend and significance level in the study area.

In the Figure 2 given below the deviation in temperature from normal temperature is analysed. The temperature time series used for the trend analysis is from the year 1901 to the year 2020. Thus 120 years of data are analysed and it has been observed that temperature series have some peaks like in the given diagram the years 1996 and 2016 have shown greater positive deflections from the normal temperature. The year 2016 has more than a 9% deviation from the mean temperature.

Trend analysis is used to estimate trends in time series data of temperature in India. The analysis has been done on the basis of Mann-Kendall and Sen's slope estimate method.

The statistical characteristics like average, standard deviation and coefficient of variance (CV) are given in Table 1 and Figure 1, which are also showing the trendline and variability. The coefficient of variance (CV) shows the extent of variability of data in relation to the mean of the population. It has been found that though the values of CV are low but showing an increasing trend in recent years that indicates an increase in variability. The months of February, November and December are showing greater variability in the temperature time series. The equations are also given in the diagram.

Table 1: Temperature: Statistical characteristics during 1901-2020

Months	*Minimum*	*Maximum*	*Mean*	*SD*	*CV %*
January	17.25	20.92	18.42	0.601	3.263
February	17.79	23.58	20.142	0.915	4.543
March	21.78	26.61	23.442	0.887	3.784
April	24.84	29.56	26.512	0.743	2.803
May	26.97	30.78	28.378	0.624	2.199
June	27.33	29.88	28.294	0.449	1.587
July	26.48	28.47	27.363	0.335	1.224
August	26.21	28.17	26.932	0.332	1.233
September	25.47	27.72	26.328	0.351	1.333
October	23.52	26.81	24.717	0.514	2.080
November	20.59	23.9	21.744	0.602	2.769
December	10.445	21.89	19.079	1.003	5.257

The trend line in Figure 3 shows the increasing trend in the temperature data series. The analysis of the temperature data series reveals that during the last 120 years the average temperature in the city has significantly increased.

India is an agrarian country and has a monsoon-type climate. The characteristic of its climate is its rhythm of seasons. In Figure 4 given above the seasonal temperature is also analysed by grouping all twelve months into four groups (Jan-Feb, Mar-May, Jun-Sep and Oct-Dec). All four groups are showing an increasing trend in temperature. It can alter the whole crop cycle of the country.

The results of Mann- Kendall (MK) and Sen's slope are given in Table 2. The M K test statistic (S) indicates that there is an increasing trend in temperature data for all the months specifically the months of November, December, September and February in India. The null hypothesis (H0) states that no trend exists, while the alternative hypothesis (H1) states that there is a trend in the two-sided test or an upward (or downward) trend in the one-sided test. H0 is rejected while H1 is accepted for all months. Thus, there is a trend in the temperature data series of India. The greater the value of *S*, the stronger the trend. Sen's slope has risen in recent years. Sen's slope estimator calculates the slope of any time series (amplitude and direction). Positive values indicate an upward trend, while negative values indicate a downward trend. These numbers represent the rate of change of the time series (via slope).

The *p*-value is less than the significance level (1% level of significance) and α (alpha) = 0.01, H0 is

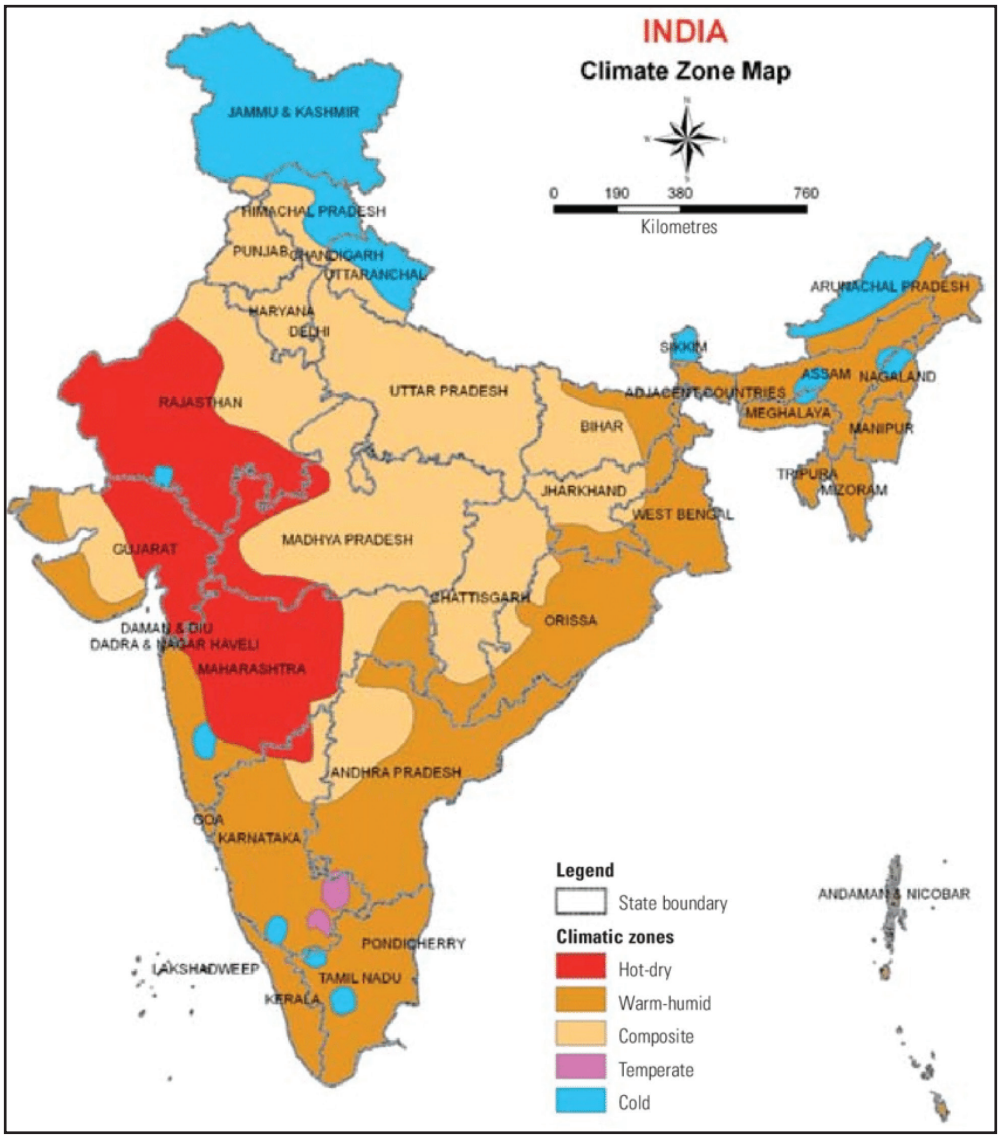

Figure 1: Showing study area with the general classification of climate.

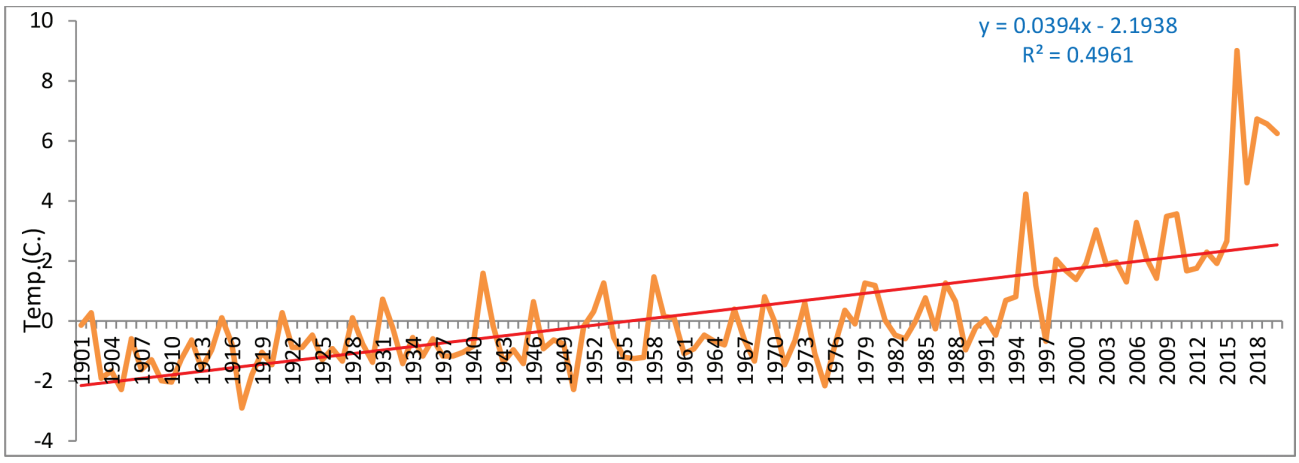

Figure 2: Temperature variability in the study area (1901-2020).

Table 2: Results of Mann Kendall test and Sen's slope estimator

Series/Test	*S value*	*Var (s)*	*Kendall's Tau*	*Sen's Slope*	*p-value*	*Results*
January	1903	180165.667	0.281	0.007	<0.0001	Ho Rejected
February	2523	180177.667	0.373	0.013	<0.0001	Ho Rejected
March	2214	180174.667	0.327	0.011	<0.0001	Ho Rejected
April	1655	180182.333	0.244	0.008	<0.0001	Ho Rejected
May	1055	180153.667	0.156	0.004	0.013	Ho Rejected
June	1184	180157.333	0.175	0.003	0.005	Ho Rejected
July	1523	180126.333	0.225	0.003	0	Ho Rejected
August	2446	180142	0.362	0.005	<0.0001	Ho Rejected
September	3090	180105.333	0.458	0.006	<0.0001	Ho Rejected
October	2134	180148.667	0.316	0.007	<0.0001	Ho Rejected
November	2855	180179	0.421	0.011	<0.0001	Ho Rejected
December	3127	180163	0.462	0.011	<0.0001	Ho Rejected

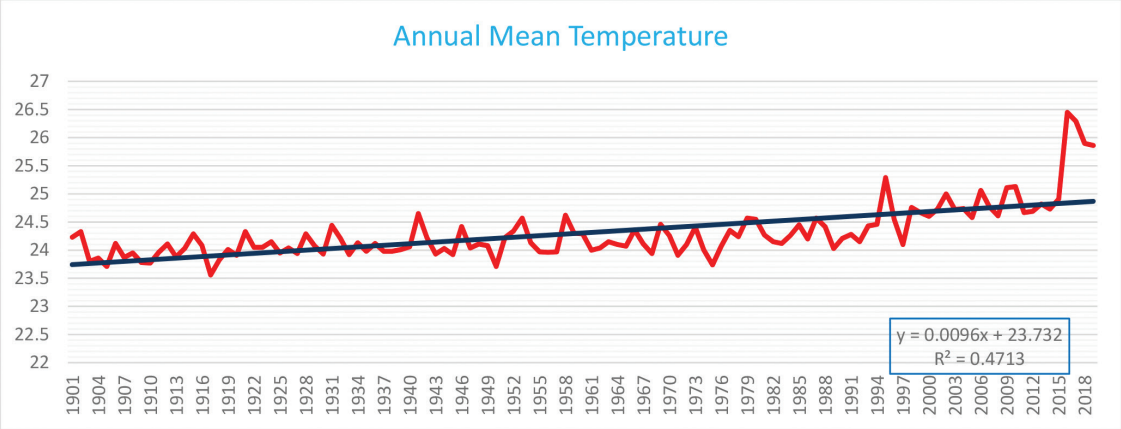

Figure 3: Showing annual mean temperature (1901-2020) with trend line and equation.

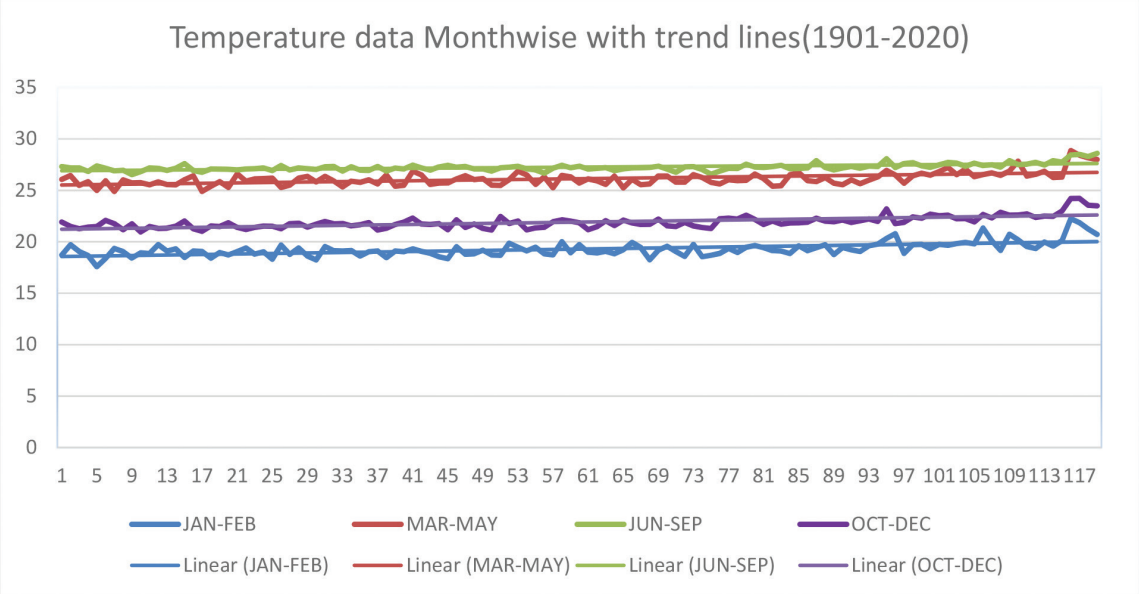

Figure 4: Temperature data in month-wise groups with trend lines (1901-2020).

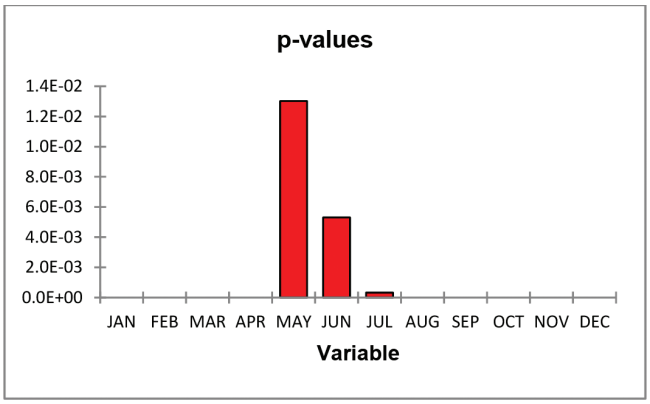

Figure 5: *p* values of the test.

rejected. Rejecting the H0 indicates that there is a trend in the time series while accepting H0 indicates that there was no trend detected. On rejecting the null hypothesis, the result is said to be statistically significant. Figure 5 shows the *p*-value of the test, the null hypothesis was rejected for all the months, except May. While on 5% level of significance, the *p*-value for the month of May is also less than the alpha value of 0.05. We can say with a 95% level of confidence that the May month has a positive trend in temperature data. Thus with a 95% level of confidence, we can say that all the months are showing positive trends in temperature data of India.

ARIMA Modelling

The ARIMA modelling is used for the analysis of time series data on the temperature in India. Given below Figure 6 has shown the changes in the temperature in the coming years. The trend line is indicating an increasing trend that is the result of global warming. It shows that the year 2016 was the warmest and after 2021 decline in the temperature will be observed. Temperature also impacts the rainfall pattern as both are correlated.

The autoregressive moving average method is an effective method of forecasting climatic data series. In the context of India, the forecasting of temperature is showing and continuous increase in the trend line. but after the year 2016, it will show less variability.

Scope for Further Research and Research Gap

The Mann-Kendall test and Sen's slope estimator both are effective tools for analysing the climatic data series but the nature of data and its distribution play a significant role in the selection of the method to be applied. So the data should be studied carefully. The selection of fit intervals and the explanations are also done very carefully for appropriate and accurate results. The most important thing that should be held on is that the methods applied in this study requires no autocorrelation.

The results of the Mann-Kendal test and Sen's slope estimator are proved with the statistical tools used in this study but the forecasting results by ARIMA modelling can be authenticated by applying it to the old data to see the forecast results of recent data. For studying local variability and trends in temperature, the climatic region-wise study will prove to be more comprehensive and useful. This gap can be filled with further research.

Conclusion

From the above study, particularly from the section Results and Discussions, it can be concluded that the Mann-Kendall test, Sen's slope estimator and other statistical tools show an increasing trend in the temperature time series of India. In the study area, the pattern of temperature in the cold months of November, December and February shows a significant increase in the temperature. The variability in temperature in recent

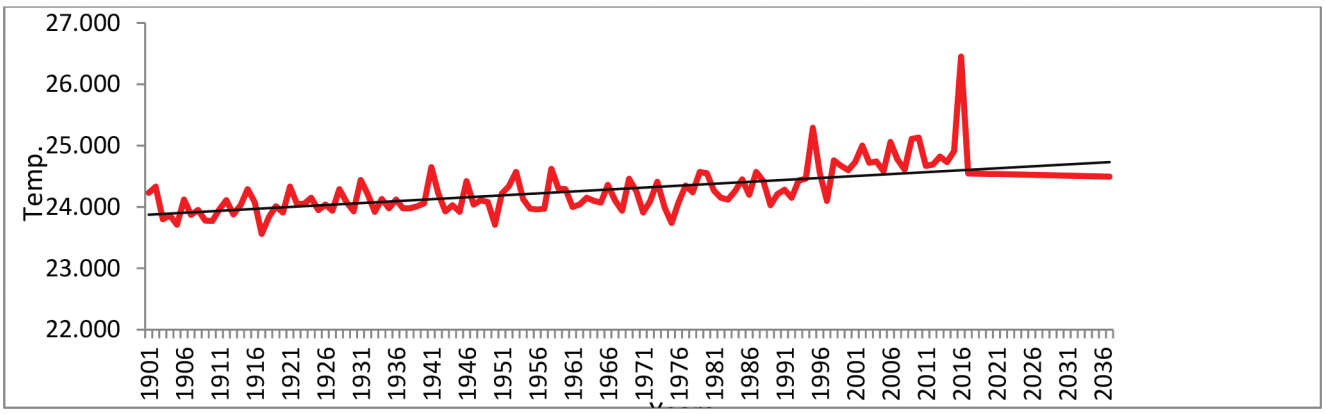

Figure 6: Temperature forecasting through ARIMA modeling.

years is increasing and 2016 was observed to be the warmest year. Data of temperature of 117 years was also analysed into four groups: January and February, March to May, June to September, and October to December. The data was divided into these four groups according to seasons and agrarian significance in the context of India. Trend lines of all the four groups are showing an increasing trend in the temperature of India. And we all know that temperature is the most significant element of the weather and climate of a place. The climate is becoming more prone to extreme weather in India. At the global level, extreme weather phenomena are occurring due to global warming. For India, the significant increase in temperature is an alarming situation, this can lead to change in the crop cycle, it can make India more vulnerable to vector-borne disease and can adversely impact crop productivity. The ecosystem of North India can get seriously affected by this increase in temperature. Melting ice, and melting life can be seen everywhere due to this change. Cyclonic systems can affect our country more impactfully and frequently. The coastal life and marine biotic life can get adversely affected. Coral bleaching can happen and the droughts can be more frequent.

The study, therefore, offers remarkable insights and a new perspective for policymakers and planners in helping them take proactive measures in the context of this trend. Timely measures and institutional changes can certainly help in reducing the irreparable damages that can be caused by the increase in temperature. More than 50% population of India is rural population and depends upon agriculture so it is necessary to expand irrigation facilities. Because the increase in temperature can affect the pattern of rainfall and further impact the other aspects of human life.

We have to be prepared for the disease that is caused because of the increase in temperature. We have to be prepared for the new crop cycle and crop pattern, we have to accept climate change and make policies to mitigate its negative impacts. It would be worthy if the gap in the research is filled byfurther research.

References

Alemu, Z.A. and Dioha, M.O., 2020. Climate change and trend analysis of temperature: The case of Addis Ababa, Ethiopia. *Environmental Systems Research*, **9:** 27.

Carro-Calvo, L., Jaume-Santero, F., Gracia-Herrera, R. and Salcedo-Sanz, S., 2021. k-Gaps: A novel technique for clustering incomplete climatological time series. *Theoretical and Applied Climatology*, **143:** 447-460.

Cheng, W.L., Saleem, A. and Sadr, R., 2017. Recent warming trend in the coastal region of Qatar. *Theoretical and Applied Climatology*, **128:** 193-205.

Dash, S.K. and Hunt, J.C.R., 2007. Variability of climate change in India. *Current Science*, **93(6):** 782-788.

Diaz, H.F. and Bradley, R.S., 1997. Temperature variations during the last century at high elevation sites. *Climate Change*, **36:** 253-279.

Dimri, T., Ahmad, S. and Sharif, M., 2020. Time series analysis of climate variables using seasonal ARIMA approach, *Journal of Earth System Science*, **129:** 149.

Fang, X., Luo, S. and Lyu, S., 2019. Observed soil temperature trends associated with climate change in The Tibetan Plateau, 1960-2014. *Theoretical and Applied Climatology*, **135:** 169-181.

Khullar, D.R., 2014. India: A Comprehensive Geography. 3rd ed., Kalyani Publishers. pp.119-167.

Kolendowicz, L., Czernecki, B., Polrolniczak, M., Taszarek, M., Tomczyk, A.M. and Szyga-Pluga, K., 2019. Homogenization of air temperature and its long-term trends in Poznan (Poland) for the period 1848-2016. *Theoretical and Applied Climatology*, **136:** 1357-1370.

Lama, A., Singh, K.N., Singh, H., Shekhawat, R., Mishra, P. and Gurung, B., 2021. Forecasting monthly rainfall of sub-Himalayan region of India using parametric and non-parametric modelling approaches. *Modeling Earth Systems and Environment*, **8:** 837-745.

Latif, Y., Yaoming, M., Yaseen, M., Muhammad, S. and Wazir, M.A., 2020. Spatial analysis of temperature time series over The Upper Indus Basin (UIB), Pakistan. *Theoretical and Applied Climatology*, **139:** 741-758.

Mahmood, R., Jia, S. and Zhu, W., 2019. Analysis of climate variability, trends, and prediction in the most active parts of The Lake Chad Basin, Africa. *Scientific Reports*, **9:** 6317.

Mishra, H.N. and Mishra, A., 2017. Climate change in India: Challenges and resolutions. *Annals of the National Association of Geographers, India*, **1(2):** 6-16.

Mudelsee, M., 2019. Trend analysis of climate time series: A review of methods. *Earth-Science Reviews*, **190:** 310-322.

Murat, M., Malinowska, I., Gos, M. and Krzyszczak, J., 2018. Forecasting daily meteorological time series using ARIMA and regression models. *International Agrophysics*, **32:** 253-264.

Nasher, N.R. and Uddin, M.N., 2013. Maximum and minimum temperature trends variation over Northern and Southern Part of Bangladesh. *Journal of Environment Sciences & Natural Resources*, **6(2):** 83-88.

Oyamakin, S.O. and Ayoola, F.J., 2010. Time series analysis of rainfall and temperature in South West Nigeria. *The Pacific Journal of Science and Technology*, **11(2):** 372-380.

Semenoglou, A.A., Spiliotis, E., Makridakis, S. and Assimakopoulos, V., 2021. Investigating the accuracy of cross-learning time series forecasting methods. *International Journal of Forecasting*, **37(3):** 1072-1084.

Shi, X. and Chen, J., 2021. Trends in the differences between homogenized ground surface temperature and surface air temperature in China during 1961-2016 and its possible causes. *Theoretical and Applied Climatology*, **144:** 41-54.

Stamp, L.D., 1958. Asia: Regional and Economic Geography.

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 Environment Research*, **7(8):** 149-159.

Szwed, M., 2019. Variability of precipitation in Poland under climate change. *Theoretical and Applied Climatology*, **135:** 1003-1015.