CLIMATE DATA ANALYSIS OF SWACH-CITY INDIA-INDORE USING STATISTICAL METHODS

Lokesh Avlasia, **Charchit Jain***, **Abhinandan Mahajan**

Department of Computer Science and Engineering, Medi-Caps University, Indore, India

* corresponding author: Charchit Jain
E-mail: charchitjain33@gmail.com

Abstract

Science has shown that climate change impacts every corner of our planet's environment. It is, therefore, very important to study the predictions and analyses of change in critical climate variables such as temperature, humidity, and rainfall.

To find patterns in temperature, humidity, precipitation, and visibility, an analytical study has been conducted. 48 years, from 1973 to 2021, have been covered through statistical approaches. In addition to temperature-related variables (mean, mean maximum, and mean minimum), rainfall, visibility factors, and humidity were also taken into account for study on a monthly and annual basis. Each year, we showed the percentage of statistically significant trends and the regression equations we found for each parameter over the previous 12 months.

Researchers can use this study as a starting point when evaluating current and future climate hazards to the region, notably Indore.

Keywords: climate change, statistical methods, temperature, humidity, rainfall analysis

Introduction

To date, climate change is widely considered one of the most crucial and intricate global environmental problems (Kandlikar and Sagar 1999). Alarming scientific and media reports over the past few decades have confirmed the undeniable fact that the Earth's average temperature is gradually rising (Verma 2021). By the year 2100, the average global temperature may rise by an unprecedented 1.4 to 5.8 °C, according to the Intergovernmental Panel on Climate Change. Such an increase is expected to have severe consequences on various global systems, such as the hydrological and ecosystem systems, sea levels, crop production, and related processes. The impacts of these changes would be particularly severe in tropical regions, which primarily comprise developing countries, including India (Balasubramanian and Birundha 2012).

The climate in India is notably diverse, shaped by the country's varied topography and geography. The country is known for its diverse climate types, ranging from a desert climate in the west with rainfall less than 300 mm per year, to a wet climate in the tropics with an average rainfall of around 2000 mm per year. In addition, there is a wet-dry climate in the tropics characterized by temperatures above 18 ℃ and rainfall above 1500 mm per year. The subtropical regions of India experience a humid climate with temperatures ranging from 0-27 ℃ and rainfall ranging from 1000-2500 mm per year. The northern regions of India have an alpine climate type with glaciers. and cool winters (The climate in India n.d.) (Nayak and Takemi 2021).

In India, a trend of increasing temperatures has been noticed, with the year 2016 being the hottest on record since nationwide data collection began in 1901 (as reported by IMD in 2016). The last 16 years (2000-2016) have seen the top five warmest years on record (Goyal and Surampalli 2018). The effects of climate change that are now being felt in the country. The increasing temperatures have shattered previous records, with the mercury climbing above 49 degrees Celsius in the capital city. Heatwaves have become more frequent, causing significant distress to a large portion of the Indian population. Failure to take immediate action to address climate change could have serious consequences, particularly in India, as emphasized by the Intergovernmental Panel on Climate Change's (IPCC) latest Sixth Assessment Report from Working Group II (Singh 2022). The climate in India is presently undergoing a warming trend, and there have been noticeable alterations in rainfall patterns. Specifically, a decrease in monsoon rainfall has been detected since the 1950s. According to the results of the rainfall forecast, there will be a 15% decrease in rainfall by 2030, which suggests a concerning situation for both the environment and the living world (Praveen, Talukda and Shahfahad 2020).

There has been an increase in the occurrence of heavy rainfall events. It has been observed that certain regions of South Asia have experienced a decrease in precipitation since the 1970s, increasing the frequency of droughts. The impact of droughts can be significant, as demonstrated by the significant reduction in crop production due to droughts in 1987 and 2002-2003, which affected over 50% of India's crop area. The most severe drought in the past century is believed to be the 2002 drought, which was labeled as the most acute drought because there was a complete absence of monsoon rainfall (only 64mm) up until August of that year (Gautam and Bana 2014). India is highly reliant on groundwater because more than 60% of its agriculture depends on rainfall. About 15% of the country's groundwater resources are already overused, even without the impact of climate change. As a result, many regions of India are already experiencing water scarcity. The increasing need for water in the agricultural, industrial, and residential sectors has resulted in a range of issues including the excessive use of groundwater resources, persistent reduction in groundwater levels, intrusion of seawater in coastal regions, and contamination of groundwater in various parts of the nation (Shrivastava n.d.). India's health is predicted to be significantly impacted by climate change. Some health problems affecting people directly or indirectly related to climate change include respiratory infections, asthma, allergies, hyperthermia, and dehydration (Kaur and Pandey 2021).

The purpose of this study is to evaluate how climate change would affect Indore's humidity, rainfall, air temperature, and visibility. A retrospective study is carried out and Statistical Analysis have been done on climate data covering a period of 48 years 1973 - 2021 to identify trends and patterns of climate in Indore and prediction of climate variables such as humidity, rainfall, temperature, and average visibility up to 5 years is also done. The use of statistical methods in weather forecasting is based on the assumption that future weather patterns will be a continuation of the patterns observed in the past. This assumption relies on the analysis of historical weather data (Didal, et al. 2017). Researchers can use this study as a starting point when evaluating current and future climate hazards to the region, notably Indore.

Study Area

The city of Indore holds the title of being the biggest and most heavily populated among all cities located in the Indian state of Madhya Pradesh (Indore n.d.). Indore is listed as one of the 20 most polluted cities in India (Panday and Dohare 2016). Indore's 2023 population is now estimated at 4,361,747 (Indore Population 2023 2023). The Indore district, which is in the catchment basin of the river Chambal, has been chosen as the study location for climate change. Its latitude and longitude are 22.2-23.05 degrees North and 75.25-76.16 degrees East, respectively. It is in the rain-shadow region of the Western Ghats and mostly receives rainfall during the southwest monsoon (Kawadia and Tiwari 2017).

Fig.1. Study Area

Methodology And Data

Data Sources

For this study, data from Indore spanning 48 years (1973-2021) was examined. The data included a time series of monthly mean humidity, mean temperature, average visibility, maximum annual mean temperature per day, minimum annual mean temperature per day, annual mean rainfall, and average relative humidity. This data has been extracted from website tutiempo (htt). The humidity data was in percentage, the temperature is in Celsius, Rainfall in mm, and Visibility in KM.

Time Series

A time series is a group of data items that are measured consecutively and often cover a range of periods. The measurements taken during an event are arranged chronologically in a time series (Adhikari and Agrawal 2013). The process of time series analysis entails constructing models that can effectively capture the dynamic characteristics of the data. These models can be utilized to generate predictions or to evaluate hypotheses regarding the fundamental mechanisms responsible for generating the data (Baltagi 1995).

Linear Regression

It is a statistical method for figuring out and quantifying how variables in a set of data relate to one another (Kumari and Yadav 2018).

$$Y_i = f(X_i, \beta) + e_i$$

Y_i = dependent variable

f = function

X_i = independent variable

β = unknown parameters

e_i = error terms

Fig.2. Linear Regression

Ordinary Least Square

Statistically, OLS regression is used to estimate the relationship between one or more independent variables and a dependent variable. It accomplishes this by minimizing the total squared difference between the dependent variable's actual values and its predicted values as calculated by a straight-line model (Ordinary Least Squares Regression n.d.).

To conduct the analysis, the following approach has been utilized.

$$\hat{x}_i = a + bt_i \ (i = 1,2,3\ldots.n)$$

Fig.3.

The offered equation shows a linear relationship between a climate variable (either temperature or humidity) x_i during the selected research period and the time series t_i that is given.

The following equations, which take 't_i' as the independent variable and 'x_i' as the dependent variable, are used to determine the regression coefficient 'b' and the regression constant 'a' of least-squares estimation.

$$b = \frac{\sum_{i=1}^{n} x_i t_i - \frac{1}{n}(\sum_{i=1}^{n} x_i)(\sum_{i=1}^{n} t_i)}{\sum_{i=1}^{n} t_i^2 - \frac{1}{n}(\sum_{i=1}^{n} t_i)^2} + \cdots$$

$$a = \frac{1}{n}\sum_{i=1}^{n} x_i - \frac{b}{n}\sum_{i=1}^{n} t_i$$

Fig.4.

PERSON'S MOMENT CORRELATION COEFFICIENT

Correlation measures the degree of the monotonic link between two variables. When both variables increase simultaneously or when one variable drops as the other increases, there is a monotonic relationship between the two variables.

A specific illustration of a linear relationship between two variables is a monotonic relationship. Typically, when referring to the correlation between two continuous, random variables, the term "correlation" specifically denotes a Pearson product-moment correlation in the context of a linear relationship (Schober, Boer and Schwarte 2018).

$$r = \frac{n(\sum x_i y_i) - (\sum x_i)(\sum y_i)}{\sqrt{n(\sum x_i^2) - (\sum x_i)^2}\sqrt{n(\sum y_i^2) - (\sum y_i)^2}}$$

Fig.5. person's moment correlation coefficient

Where x is the annual mean temperature, y is the annual mean humidity, and i = 1,2,3……n element of the time series.

Coefficient Of Determination

The amount of variance in the dependent variable that can be explained by the independent variable is expressed as a percentage by a statistical measure in a regression model called the coefficient of determination (R2). In other words, the coefficient of determination can be used to assess how well the data match the model.

The Yi element in the formula below represents the actual ith value, whereas x_i represents the anticipated ith value. (Chicco, Warrens and Jurman 2021).

$$R^2 = 1 - \frac{\sum_{i=1}^{m}(X_i - Y_i)^2}{\sum_{i=1}^{m}(\bar{Y} - Y_i)^2}$$

Fig.6. Coefficient of Determination

Result

Climate Variables

Mean temperatures of months are plotted in Fig.7. to Fig.18. for each year to understand the variations in temperatures in the same months throughout the years.

Jupyter Notebook was utilized to develop linear trends and plot the time series of the annual mean humidity, annual mean rainfall, annual mean temperature, and annual average visibility in Indore over the period from 1973 to 2021. These time series are illustrated in Fig.19. to Fig.23. The regression equations and coefficients of determination (R^2) were computed using the ordinary least squares method and presented in Table.1.

Fig.19. presents two climate variables, the annual mean minimum temperature, and the annual mean maximum temperature, depicted by the orange and blue lines, respectively. The black line represents the linear trend in the data over the analyzed period.

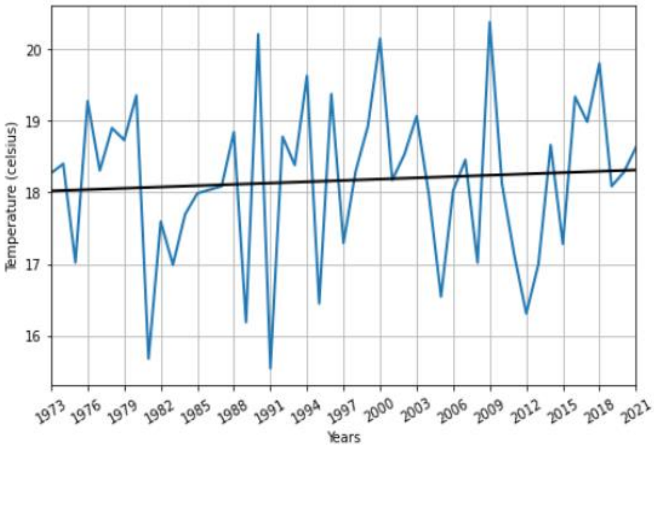


Fig.7. January

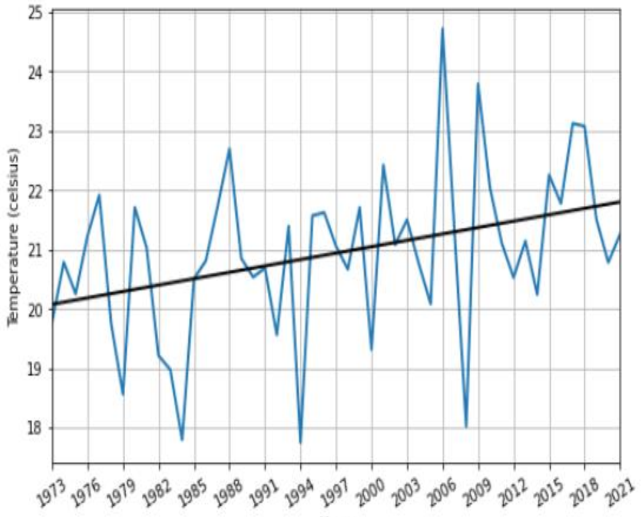
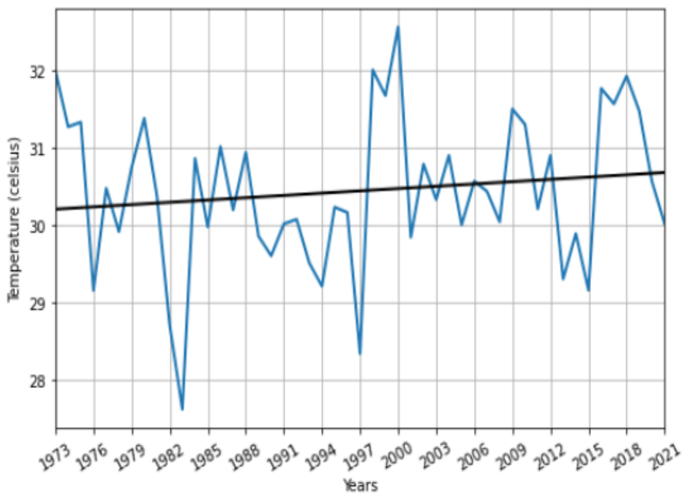


Fig.8. February



5

Fig.9. March

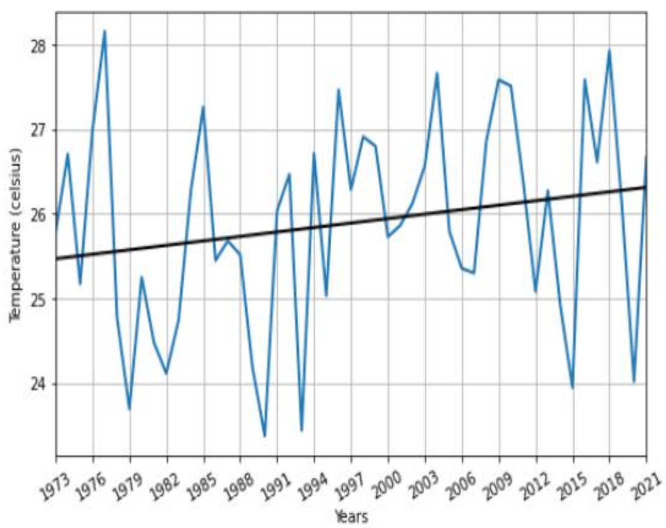


Fig.10. April

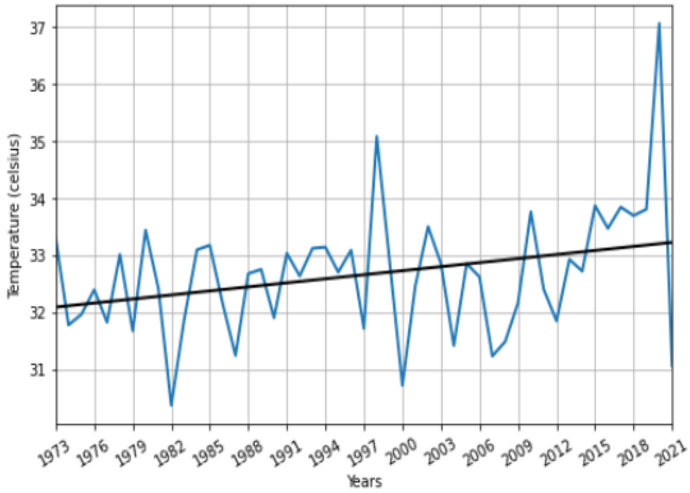


Fig.11. May

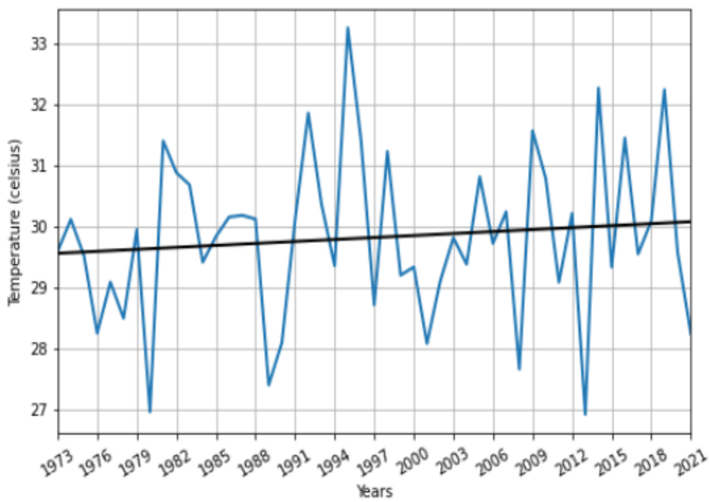


Fig.12. June

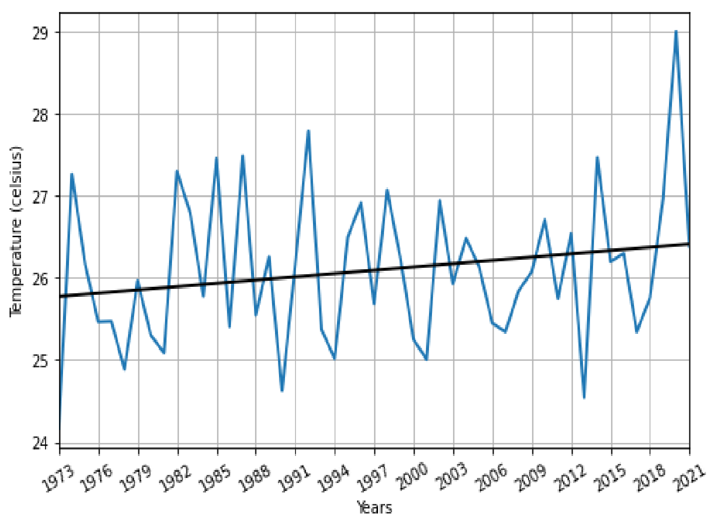


Fig.13. July

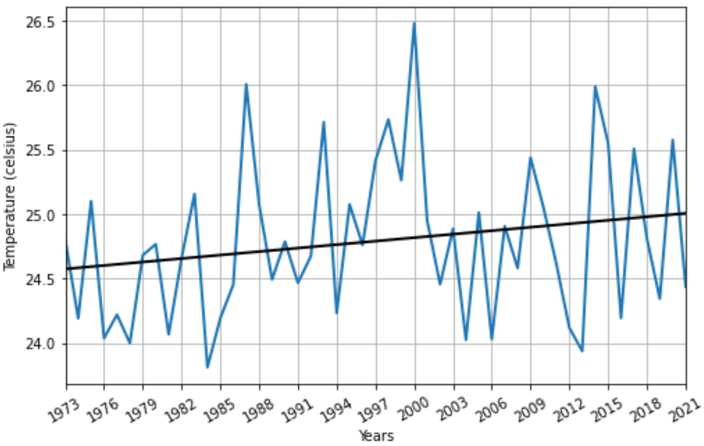


Fig.14.August

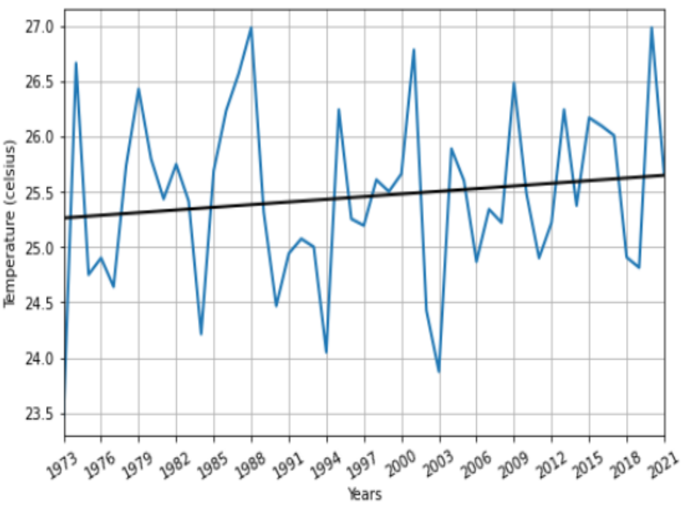


Fig.15. September

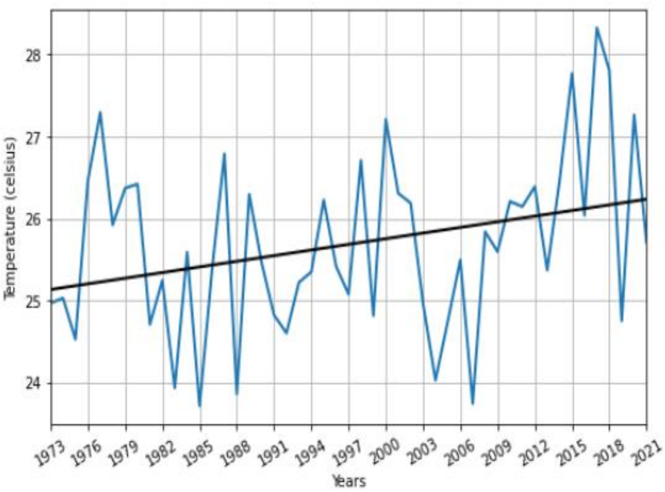


Fig.16. October

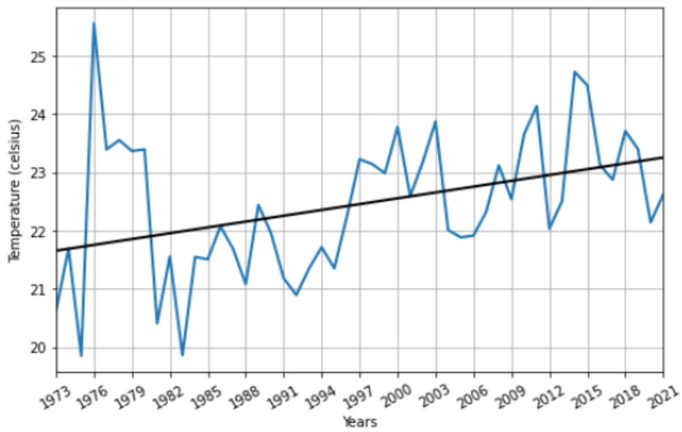


Fig.17. November

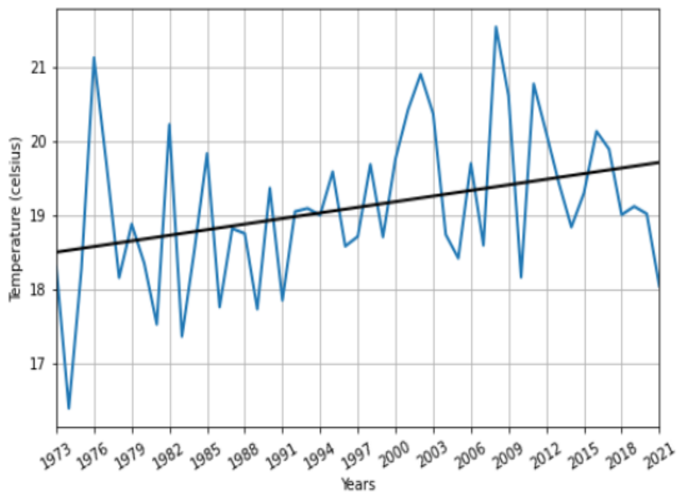


Fig.18. December

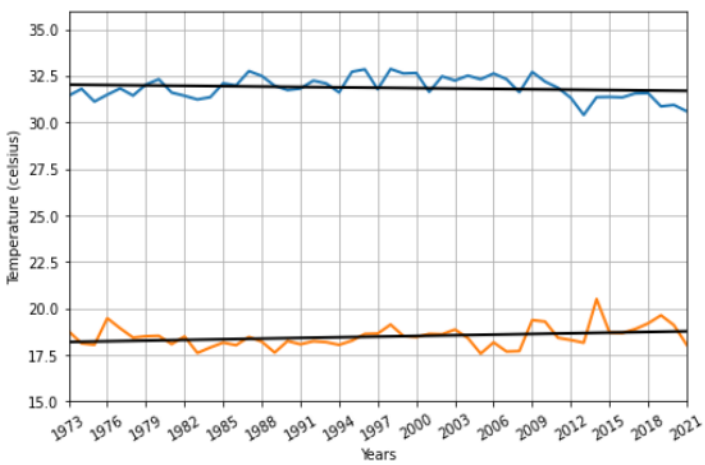


Fig.19. Annual mean max. temperature and annual mean min. temperature

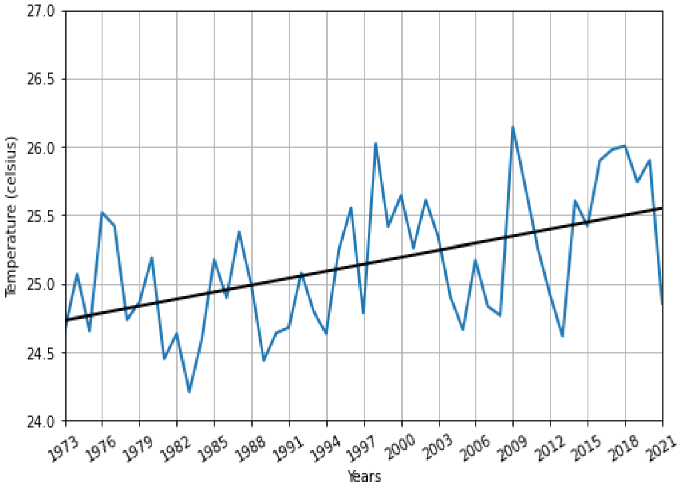


Fig.20. Annual Mean Temp

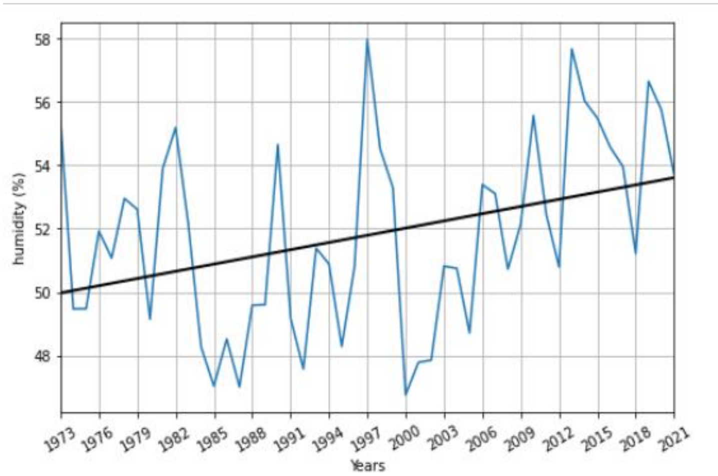


Fig.21. Average relative Humidity

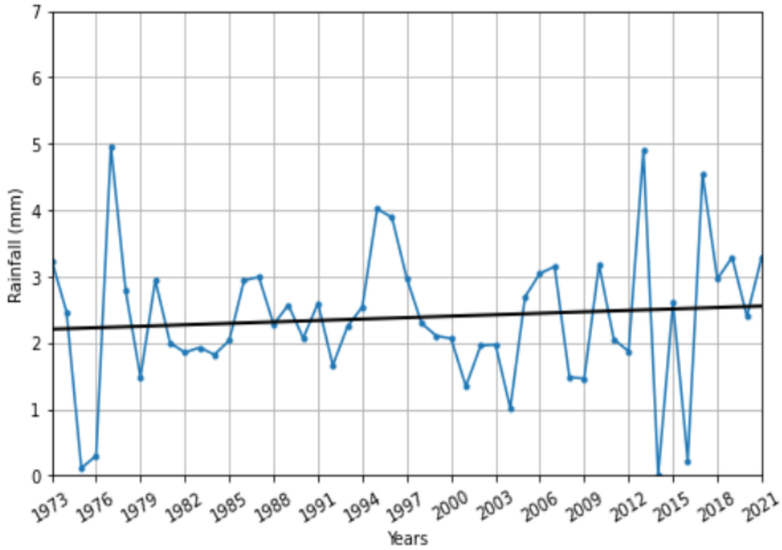
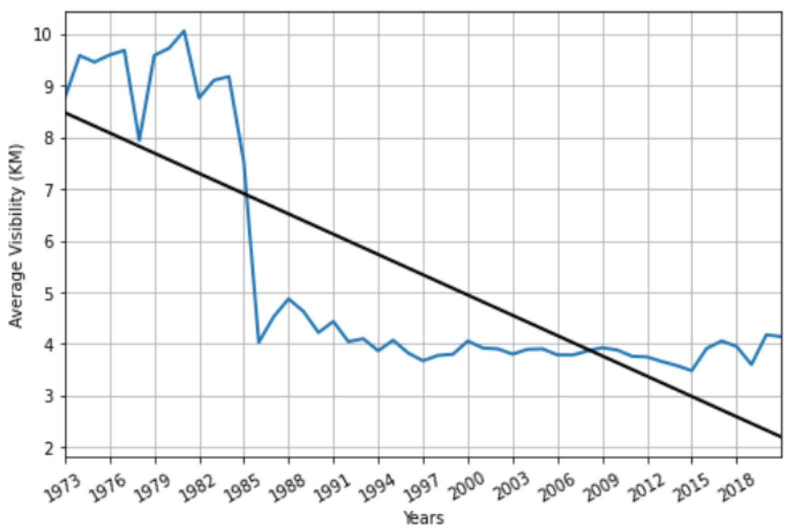


Fig.22. Annual Mean Rainfall



	Months and annual	Equation	Coefficient of determination
0	January	y = 0.0061x + 18.0213	0.005806546
1	February	y = 0.0359x + 20.075	0.131018655
2	March	y = 0.0176x + 25.4691	0.041763726
3	April	y = 0.0099x + 30.2098	0.01973311
4	May	y = 0.0237x + 32.0872	0.090271416
5	June	y = 0.0108x + 29.5516	0.012681863
6	July	y = 0.0133x + 25.7756	0.040280048
7	August	y = 0.009x + 24.5746	0.043069756
8	September	y = 0.008x + 25.2615	0.020703069
9	October	y = 0.023x + 25.1321	0.091414943
10	November	y = 0.0334x + 21.6535	0.152304966
11	December	y = 0.0253x + 18.4948	0.117489588
12	annual mean temperature	y = 0.017x + 18.0213	0.246653848
13	annual mean max temperature	y = -0.0069x + 20.075	0.026622795
14	annual mean min temperature	y = 0.0119x + 25.4691	0.08861623
15	annual mean humidity	y = 0.0756x + 30.2098	0.129412633
16	annual mean rainfall	y = 0.0072x + 32.0872	0.009089147
17	annual mean visibility	y = -0.1308x + 29.5516	0.629741197

Fig.23. Average Visibility

Table.1. show the regression equation and coefficient of determination

Climate Variables Projected In Future

Table.1. provides the regression equation used to estimate the mean temperatures of a given month for each year. The study also employed regression equations to project the annual mean temperature, minimum and maximum daily temperatures, humidity percentage, rainfall, and average visibility for Indore spanning the period from 1973 to 2021. The outcomes of these projections are tabulated in Table.2.

Pearson Correlation Coefficient

The correlation between mean temperature and humidity was determined using the Pearson product-moment correlation coefficient, which produced a value of (-0.24) that represents the strength of their association

	Months and Annual	2021	2022	2023	2024	2025	2026
0	January	18.31385	18.31994	18.32603	18.33213	18.33822	18.34432
1	February	21.79813	21.83402	21.86992	21.90582	21.94172	21.97761
2	March	26.31409	26.3317	26.3493	26.3669	26.3845	26.40211
3	April	30.68465	30.69454	30.70444	30.71433	30.72422	30.73411
4	May	33.22368	33.24736	33.27104	33.29471	33.31839	33.34207
5	June	30.06917	30.07996	30.09074	30.10152	30.11231	30.12309
6	July	26.41545	26.42878	26.44211	26.45544	26.46877	26.4821
7	August	25.00588	25.01486	25.02385	25.03283	25.04182	25.0508
8	September	25.64684	25.65487	25.6629	25.67093	25.67896	25.68699
9	October	26.23376	26.25671	26.27966	26.30262	26.32557	26.34852
10	November	23.2544	23.28775	23.3211	23.35445	23.38781	23.42116
11	December	19.70694	19.73219	19.75745	19.7827	19.80795	19.83321
12	Annual Mean Temperature	25.55167	25.56872	25.58576	25.6028	25.61985	25.63689
13	Annual Mean Max Temperature	31.6939	31.68701	31.68012	31.67323	31.66634	31.65945
14	Annual Mean Min Temperature	18.77166	18.7836	18.79554	18.80748	18.81942	18.83135
15	Annual Mean Humidity	53.60337	53.679	53.75462	53.83025	53.90587	53.9815
16	Annual Mean Rainfall	2.553399	2.560648	2.567896	2.575144	2.582393	2.589641
17	Annual Average Visibility	2.19981	2.068994	1.938178	1.807361	1.676545	1.545728

Table.2. shows the prediction of monthly and annual temperature in (Celsius), Humidity, Rainfall, visibility

Conclusion

For the monthly means, table (3) shows that February had the highest Temperature change (1.75), followed by November (1.63) and May (1.16), while September had the lowest change (0.39).
The corresponding annual rates of change likewise differ by month, with September having the lowest annual rate of change (0.01) and February having the highest annual rate of change (0.04).

For the annual means, table.3. shows that the annual mean humidity had the highest change (3.71), followed by annual mean max temperature (-0.34), annual mean min temperature (0.58), annual mean temperature (0.84), annual mean rainfall (0.36), and annual average visibility (-6.41). The corresponding rates of change per year also vary for these annual means, with annual mean humidity having the highest rate of change (0.08), followed by annual mean temperature (0.02) and annual mean min temperature (0.01). Annual mean max temperature and annual mean rainfall show a negative rate of change, while annual average visibility has the lowest rate of change (-0.13) per year.

Months and Annuals	Climate variable change for the period 1973 to 2022	Rate of change of climate variable
January	0.29	0.006
February	1.75	0.035
March	0.86	0.017
April	0.48	0.009
May	1.16	0.023
June	0.52	0.010
July	0.65	0.013
August	0.44	0.008
September	0.39	0.008
October	1.12	0.022
November	1.63	0.033
December	1.23	0.025
Annual mean temperature	0.83	0.017
Annual mean max temperature	-0.33	0.017
Annual mean min temperature	0.58	0.011
Annual mean humidity	3.70	0.075
Annual mean rainfall	0.35	0.007
Annual average visibility	-6.41	-0.130

Table 3. Climate variable change

References

Kandlikar, Milind, and Ambuj Sagar. 1999. "Climate change research and analysis in India: an integrated assessment of a South–North divide." In *Global Environmental Change, Volume 9, Issue 2*, 119-138.

n.d. https://en.tutiempo.net/.

Adhikari, Ratnadip, and R. K. Agrawal. 2013. "An Introductory Study on Time series Modeling and Forecasting." LAP Lambert Academic Publishing.

Balasubramanian, Muniyandi, and V. Dhulasi Birundha. 2012. "Climate Change and its Impact on India." 31-46. The IUP Journal of Environmental Sciences.

Baltagi, Badi H. 1995. "Forecasting and Time Series Analysis." *Journal of Forecasting.*

Chicco, Davide, Matthijs J. Warrens, and Giuseppe Jurman. 2021. "The coefficient of determination R-squared is more informative than SMAPE, MAE, MAPE, MSE and RMSE in regression analysis evaluation." *PeerJ Computer Science.*

Didal, Vijay Kumar , Brij Bhooshan, Kamlesh Choudhary, and Anita Todawat. 2017. "Weather Forecasting in India: A Review." *International Journal of Current Microbiology and Applied Sciences* 577-590.

Gautam, R C, and R S Bana. 2014. "Drought in India: Its impact and mitigation strategies-A review." *Indian Journal of Agronomy* 59: 179-190.

Goyal, Manish Kumar , and Rao Y. Surampalli. 2018. "Impact of Climate Change on Water Resources in India." *Journal of Environmental Engineering.*

n.d. *Indore.* https://en.wikipedia.org/wiki/Indore.

2023. *Indore Population 2023.* https://worldpopulationreview.com/world-cities/indore-population.

Kaur, Rajveer , and Puneeta Pandey. 2021. "Air Pollution, Climate Change, and Human Health in Indian Cities: A Brief Review." 3.

Kawadia, Ganesh, and Era Tiwari. 2017. "Understanding Climate Change in Indore District: An Empirical Investigation of Trends and Shifts." 2 (1).

Kumari, Khushbu, and Suniti Yadav. 2018. "Linear regression analysis study." *Journal of the Practice of Cardiovascular Sciences.* https://www.encyclopedia.com/social-sciences/applied-and-social-sciences-magazines/ordinary-least-squares-regression.

Nayak, Sridhara , and Tetsuya Takemi . 2021. "Assessing the Impact of Climate Change on Temperature and Precipitation Over India." In *Wadi Flash Floods*, 121–142.

n.d. *Ordinary Least Squares Regression.* https://www.encyclopedia.com/social-sciences/applied-and-social-sciences-magazines/ordinary-least-squares-regression.

Panday, Vijayant, and Devendra Dohare. 2016. "Assessment of Present Ambient Air Quality of Indore City using Indian Air Quality Index -A Case Study." *International Journal of Science Technology and Engineering* 3 (6).

Praveen, Bushra, Swapan Talukda, and Shahfahad. 2020. "Analyzing trend and forecasting of rainfall changes in India using non-parametrical and machine learning approaches." *Scientific Reports.*

Schober, Patrick, Christa Boer, and Lothar A Schwarte. 2018. "Correlation Coefficients: Appropriate Use and Interpretation." 126 (5).

Shrivastava, Prashant Kumar. n.d. "Impact of Climate Change on Water Resources of India." 121-127.

Singh, Ishani Kumar. 2022. *IPCC Climate Change Report 2022 and its Implications for India.* June 29. https://india.wcs.org/Newsroom/Blog/ID/17734/IPCC-Climate-Change-Report-2022-and-its-Implications-for-India.

n.d. *The climate in India.* https://www.worlddata.info/.

Verma, Omkar. 2021. "Climate Change and Its Impacts with Special Reference to India." In *Water, Cryosphere, and Climate Change in the Himalayas, A Geospatial Approach*, 39-55. Springer.