"येथे बहुतांचे हित"

Welfare of Masses: Peer-Reviewed Multidisciplinary Research Journal (WMRJ)

Journal Website - www.journal.mmcc.edu.in/wmrj

Temperature Trend Analysis For All The State Of India Between (2000-2019)

Asst. Prof. Komal Galande

Department of Computer Science,
MMCC, Pune, Maharashtra
Author Email ID: komalg2897@gmail.com

ARTICLE INFO

Article history:

***Received**: 29 June 2026*

***Received in revised form**: 28 July 2026*

***Accepted**: 4 August 2026*

Keywords:
India,Mann-kendall trend test , Temperature, Machine learning, Test, Graph

ABSTRACT

Temperature across India varies widely by region. Generally, the average temperature across India ranges from 24°C to 27°C annually. Summers (March to June) can see temperatures soaring from 30°C to 45°C in many regions, while winters (December to February) can range from 5°C to 25°C. Here we use Machine learning (ML) techniques for time series prediction, By using the time series prediction it becomes accurate and helpful. Here weather forecasting, climate prediction and accurate temperature prediction is important. We have studied 20 years of data sets for the year 2000-2019. The statistical analysis here collected the secondary data from the Government website.

I INTRODUCTION:

Temperature data is crucial for understanding various environmental, industrial, and societal phenomena. Weather forecasting, climate research, agriculture, and energy management are just a few of the scientific, industrial, and societal pursuits that depend critically on temperature prediction. Making wise decisions.[1-2]

Regression is a statistical method used to examine the relationship between one dependent variable and one or more independent variables. It's commonly used for prediction & forecasting. In simple terms, regression analysis helps to understand how the value of a dependent variable changes over time Linear regression is the regression in which we study the relationship between dependent & independent variables. In a given data temperature is the dependent variable, by using this method we can predict short-term and long-term temperature.[3-5]

This research paper shows how to the improve temperature forecast by using machine learning Linear regression, as a classical algorithm in statistics, is widely used in various data analysis

tasks because of its simplicity and explanatory power [6-8] Additionally, this research supports continuing initiatives to use cutting- edge technology to address the urgent problems brought on by climate change and extreme weather.

The Mann-Kendall (MK) trend test is a non-parametric statistical method used to identify monotonic upward or downward trends in time series data. It is highly valued in fields like hydrology, climate science, and environmental studies because it doesn't require data to be normally distributed or linear, and it resists extreme outliers. [9-11]

Objective of Study

1. The objective of this project is to create and assess temperature prediction models using Linear Regression (LR)
2. To find trend by using scatter plot
3. To analysis change of temperature season wise by using simple bar diagram
4. To check trend analysis by using the Mannkendall test.

TABLE 1. DATA OF MONTHLY, SEASONAL ANNUAL MEAN TEMPERATURE (2000-2019)

YEAR	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC	ANNUAL
2000	18.87	19.78	23.22	27.27	28.92	28.02	27.34	26.98	26.53	25.58	22.75	19.66	24.6
2001	18.5	21	24.12	26.9	29.46	28.13	27.63	27.62	26.86	25.49	22.51	19.55	24.73
2002	18.78	20.52	24.44	27.66	29.56	28.77	28.47	27.27	26.43	25.44	23.17	19.87	25
2003	18.6	21.03	23.62	27.29	28.64	29.01	27.53	27.25	26.75	24.92	22.17	19.59	24.72
2004	18.83	21.03	26.19	27.38	28.12	27.94	27.44	27.16	26.8	24.34	22.43	19.93	24.74
2005	19.07	20.51	24.14	26.62	28.22	28.99	27.34	27.41	26.82	24.84	21.73	19.23	24.58
2006	19.96	23.02	23.91	26.83	28.82	28.13	27.92	27.15	26.42	25.41	22.59	19.99	25.06
2007	19.24	20.97	23.52	27.72	28.84	28.31	27.66	27.37	26.61	24.92	22.4	19.65	24.77
2008	18.49	19.83	24.43	26.54	28.42	28.1	27.5	27	26.44	25.47	22.51	20.62	24.61
2009	19.79	21.66	24.55	27.35	28.71	28.77	27.83	27.85	27.11	25.2	22.31	20.22	25.11
2010	19.15	21.23	26.53	28.4	29.19	28.51	27.55	27.33	26.6	25.58	22.98	19.22	25.13
2011	18.32	20.79	24.11	26.1	28.92	28.6	27.6	27.2	26.7	25.51	22.84	19.84	24.67
2012	18.25	20.43	23.98	26.89	28.72	28.91	27.98	27.31	26.65	24.85	22.26	19.91	24.69
2013	18.88	21.07	24.53	26.97	29.06	28.24	27.5	27.22	26.87	25.63	22.18	19.69	24.82
2014	18.81	20.35	23.34	26.91	28.45	29.42	28.07	27.42	26.61	25.38	22.53	19.5	24.73
2015	19.02	21.23	23.52	26.52	28.82	28.15	28.03	27.64	27.04	25.82	22.95	20.21	24.91
2016	20.92	23.58	26.61	29.56	30.41	29.7	28.18	28.17	27.72	26.81	23.9	21.89	26.45
2017	20.59	23.08	25.58	29.17	30.47	29.44	28.31	28.12	28.11	27.24	23.92	21.47	26.29
2018	19.9	22.58	26.01	28.48	29.99	29.58	28.19	27.73	27.63	26.42	23.84	20.41	25.9
2019	19.62	21.8	24.92	28.82	30.27	30.26	28.65	27.92	27.56	26.22	23.93	20.32	25.86

METHODOLOGY

Single Model ML Methods:

Linear Regression Analysis: Linear regression analysis is used to predict the value of a dependent variable based on the value of an independent variable. The variable you want to predict is called the dependent variable. The variable you are using to predict the other variable's value is called the

independent variable. Linear regression fits a straight line that minimizes the discrepancies between predicted and actual output values. Linear regression calculators use a “least squares” method to discover the best-fit line.

Here the linear regression equation is y= mx+c

Here,

m is slope of equation is intercept, y is dependent variable and x is independent variable. Linear regression analysis is used to predict the value of a variable based on the value of another variable.

Descriptive statistics:

Descriptive statistics in research are used to summarize, organize, and describe the main features of a dataset. Descriptive statistics are the summary of the given data. In measures of central tendency we study mean, mode and median, In measures of dispersion, we study variability of the given data, we visualize the data by using diagrammatic as well as graphical representation

Scatter diagram:

A scatter diagram is a graph that displays the relationship between two variables by plotting pairs of data as points on a two-dimensional plane, with one variable on the X-axis and the other on the Y-axis.

Inferential Statistics:

Mann - Kendall Test:

The Mann-Kendall (MK) trend test is a non-parametric statistical method used to identify monotonic upward or downward trends in time series data. It is highly valued in fields like hydrology, climate science, and environmental studies because it doesn't require data to be normally distributed or linear, and it resists extreme outliers

The study concludes that Mann-Kendall statistical tests in TREND are relatively easy to understand and users can gain a good implementation of the tests by following the descriptions and the examples presented in the study. Trend analysis based on Mann Kendall statistics helps to understand how the underlying function has performed and predict where the current state of the selected function will guide the dataset .

II. RESULT AND ANALYSIS

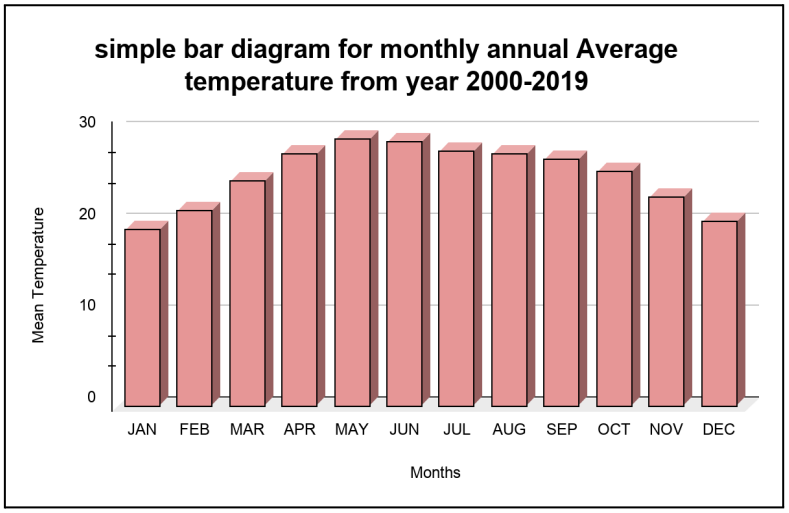

Fig1. Graphical representations of annual temperature of all india monthly fo r years 2000-2019

Winter temperatures typically involve a decrease in both maximum and minimum temperatures, with variations depending on location and proximity to bodies of water or mountains. The coldest months are generally December and January. The above graph shows the lowest temperature in the months of December & January. Summer months are typically characterized by high temperatures and increased daylight hours; the above graph shows high temperatures in May.The rainy season in

India, also known as the monsoon season, typically lasts from June to September. During this period, the southwest monsoon brings heavy rainfall to most of the country, with the heaviest rainfall occurring in July and August.

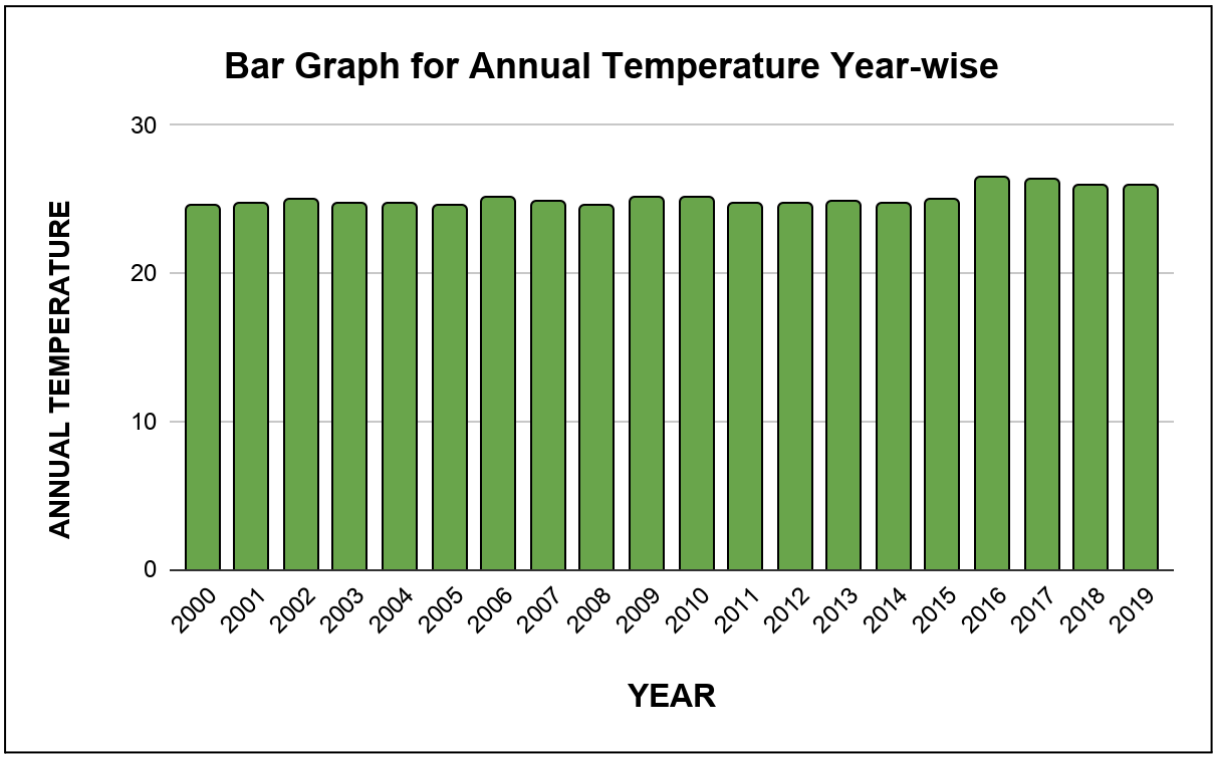

Fig 2. Annual Temperature of India for the year 2000- 2019

The temperature remains fairly stable between 24°C and 26°C throughout the 20-year period. There are small year-to-year fluctuations. From around 2016 onwards, the temperatures are slightly higher than in the earlier years. The highest temperatures are observed around 2016–2017.Overall, the graph indicates a slight increasing trend in annual temperature, suggesting gradual warming over the years. The annual temperature of India remained mostly consistent from 2000 to 2019, but there is a small rise in temperature in the later years, indicating a gradual increase in average temperature over time.

Linear Regression By using Python:

Linear regression uses the relationship between the data-points to draw a straight line through all them. This line can be used to predict future values.

```
import pandas as pd
df = pd.read_csv('C:/Users/Admin/Desktop/data_for_lr.csv')
print(df.to_string())
    X     Y
0   2000  24.60
1   2001  24.73
2   2002  25.00
3   2003  24.72
4   2004  24.74
5   2005  24.58
6   2006  25.06
7   2007  24.77
8   2008  24.61
9   2009  25.11
10  2010  25.13
11  2011  24.67
12  2012  24.69
13  2013  24.82
14  2014  24.73
15  2015  24.91
16  2016  26.45
17  2017  26.29
18  2018  25.90
19  2019  25.86
import matplotlib.pyplot as plt
from scipy import stats
x =
[2000,2001,2002,2003,2004,2005,2006,2007,2008,2009,2010,2011,2012,2013,2014,2015,2016,
2017,2018,2019]
y =
[24.60,24.73,25.00,24.72,24.74,24.58,25.06,24.77,24.61,25.11,25.13,24.67,24.69,24.82,24.73,24
.91,26.45,26.29,25.90, 25.86]
slope, intercept, r, p, std_err = stats.linregress(x, y)

def myfunc(x):
  return slope * x + intercept

mymodel = list(map(myfunc, x))
plt.scatter(x, y)
plt.plot(x, mymodel)
plt.show()
```

SCATTER DIAGRAM

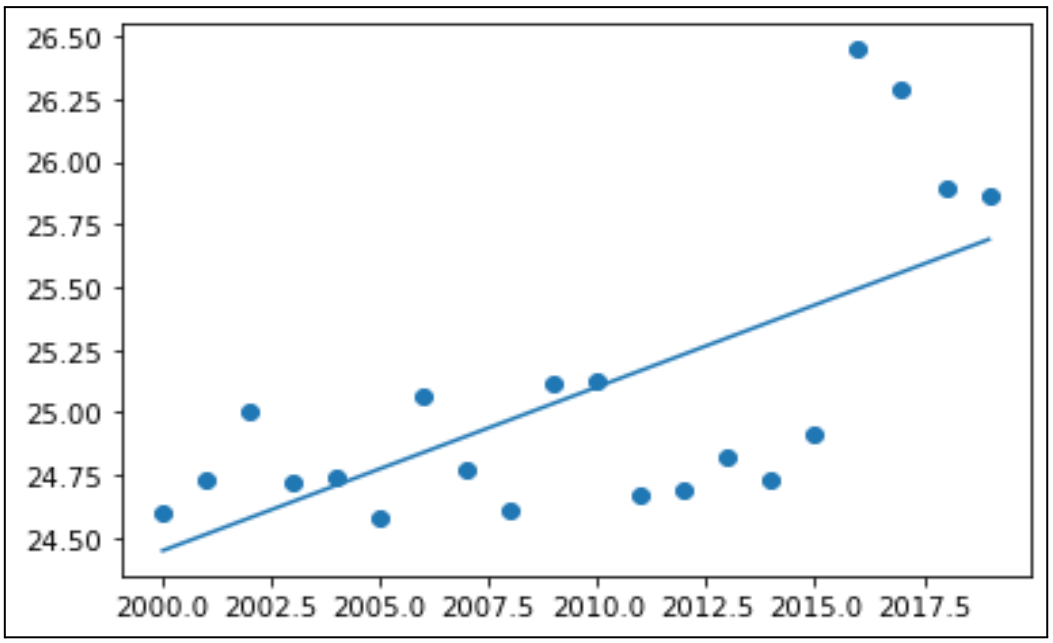

Fig 3. Scatter Plot with a fitted trend line (regression line) for annual temperature year 2000-2019

The above image shows a fitted trend line. Here we study 20 years of data for temperature. Here on X-axis(Horizontal) Years ranging from 2000 to 2019, On Y-axis(Vertical) some measured annual temperature of India year wise. The data points i.e. blue dot shows data points scattered over time with some change year by year but not perfectly linearly. Here we conclude that The trend line shows a moderate upward trend, suggesting an overall rise in values over the two-decade period.

Mannkendall test by using python programming:

The Mann-Kendall Trend Test **is** a non-parametric statistical test used to determine whether a time series has a monotonic upward or downward trend

```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import pymannkendall as mk
data=[24.6,24.73,25,24.72,24.58,25.06,24.77,24.61,25.11,25.13,24.67,24.69,24.82,24.82,24.73,24.91,26.45,26.29,
25.9,25.86]
mk.original_test(data)
```

Output of the given data:

Mann_Kendall_Test(trend='increasing', h=True, p=0.007023780103332822, z=2.6957146802157683, Tau=0.4421052631578947, s=84.0, var_s=948.0, slope=0.036250000000000115, intercept=24.475625)

Output Interpretation:

The output interpretation can be done in the following different ways:

Trend indicates that is increasing, decreasing or no trend present here, If h is True then trend is present. If h is False then no trend is present. Here p shows the p-value of the test. Z is The normalized test statistic. Tau is Kendall Tau,s is the Mann-Kendal's score, var_s is Variance of S which shows variation between the data. Slope is Theil-Sen Estimator/slope & Intercept of Kendall-Theil Robust Line

Since in the above example, the p-value is 0.007 which is less than the threshold (0.5) which is the alpha(0.5)then we fail not to reject the accepted hypothesis i.e. we do have sufficient evidence to say that sample has a trend present. Then we reject the accepted hypothesis and it is also increasing. Theil–Sen estimator (slope) is positive; it indicates an increasing trend in temperature over time.

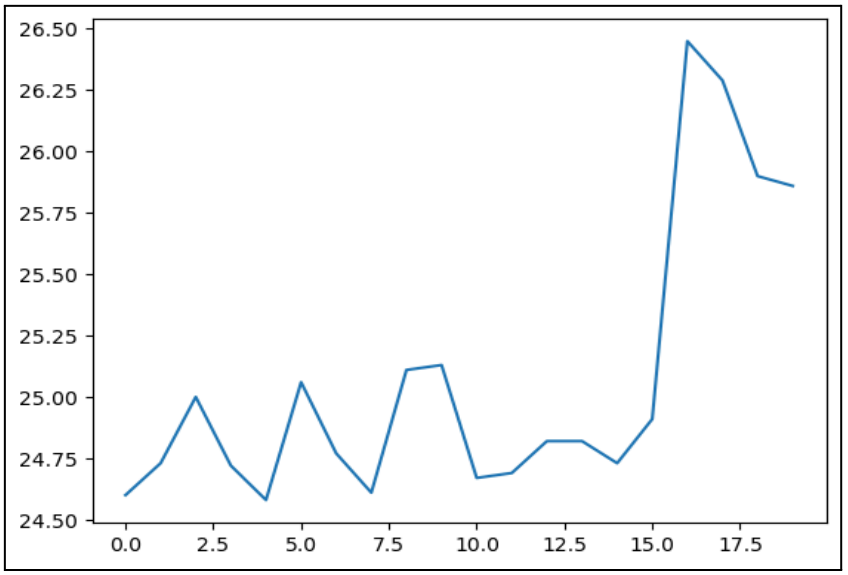

Fig 4. Visualizing the trend of the data

The temperature data for India moderate fluctuations during the initial period, the above graph shows values ranging between 24.6°C and 25.1°C Here the above graph shows maximum temperature at the end of the period. This graph generally shows an increasing trend.

CONCLUSION

Here we conclude for overall study, lowest temperature in the month of December & January due to winter season, there are the lowest temperatures, also we conclude that the highest temperature in the month of may due to the summer monsoon. Heaviest rainfall occurs in the months of July & August. From around 2016 onwards, the temperatures are slightly higher than in the earlier years. The highest temperatures are observed around 2016–2017.Overall, the graph indicates a slight increasing trend in annual temperature, suggesting gradual warming over the years. The annual temperature of India remained mostly consistent from 2000 to 2019, but there is a small rise in temperature in the later years, indicating a gradual increase in average temperature over time. Also we conclude that every three years the temperature increases.

REFERENCES

[1] Azari B, Hassan K, Pierce J, Ebrahimi S. Evaluation of machine learning methods

application in temperature prediction. Environ Eng. 2022;8:1-2.

[2] Research paper on 'Climatic Temperature Trends: Analyzing Past Data and Predicting Future Temperature'.

[3] Gupta S. C. and Kapoor, V. K. (1983). Fundamentals of Mathematical Statistics. 8th Edition, Sultan Chand and Sons Publishers, New Delhi.

[4] Goon A. M., Gupta, M. K. and Dasgupta, B. (1983). Fundamentals of Statistics, Vol. 1. 6th Revised Edition, The World Press Pvt. Ltd., Calcutta. Vikram Singh, S.K. Gupta and Jagdish Prasad

[5] Doblas-Reyes FJ, Garcia-Serrano J, Lienert F,Biescas AP, Rodrigues LR. Seasonal climate Predictability and forecasting: status and prospects. Wiley Interdisciplinary Reviews: Climate Change. 2013 Jul;4(4):245-68.

[6] Research Paper on 'Weather Prediction Using Machine Learning Algorithms'

[7]https://www.geeksforgeeks.org/machine-learning/how-to-perform-a-mann-kendall-trend-test-in-python/

[8] Research Paper on "Exploring climate trends and extremes in India: A study of temperature and rainfall from 1980 to 2023"

[9] Review of trend detection methods and their application to detect temperature changes in India

[10]Kendall, M.G., 1975. Rank Correlation Techniques, Charles Griffen. London ISBN, 195205723.

[11]. Kendall, M.G., 1955. Further contributions to the theory of paired comparisons. Biometrics, 11(1), pp.43-62.