Climatic Analysis & its effect on Ground water level of major cities

A Project report submitted in partial fulfillment of the requirement for

POST-GRADUATE DIPLOMA IN
STATISTICAL METHODS AND ANALYTICS

Submitted By

[Manashjyoti Lahkar]
[DST-24/25-014]

[Saptashi Sen]
[DST-24/25-022]

[Utkarsh Singh Rajput]
[DST-24/25-024]

June, 2025

INDIAN STATISTICAL INSTITUTE
North-East Centre

<u>CERTIFICATE</u>

This is to certify that **Utkarsh Singh** (DST-24/25-024), **Manashjyoti Lahkar** (DST-24/25-014) and **Saptashi Sen** (DST-24/25-022) have done the project under my supervision (from 06^{th} January 2025 to 30^{th} May 2025). This is an original project report based on work carried out by him/her/them in partial fulfillment of the requirement for the Post-Graduate Diploma in Statistical Methods and Analytics programme of the Indian Statistical Institute, North-East Centre, Tezpur, Assam.

Date: ....

Dr. Darpa Saurav Jyethi

ACKNOWLEDGEMENT

We would like to express our sincere gratitude to our supervisor, **Dr. Darpa Saurav Jyethi**, of the **Indian Statistical Institute, North-East Centre, Tezpur**, for his invaluable guidance, continuous encouragement, and unwavering support throughout the course of this project. His insightful suggestions and mentorship were instrumental in shaping our research and bringing this work to completion.

We also extend our thanks to the faculty members and staff of the Indian Statistical Institute, North-East Centre, Tezpur, for providing a supportive academic environment and the necessary resources for carrying out our project.

We are grateful to our friends and colleagues for their cooperation, motivation, and helpful discussions, which enriched our learning experience.

Utkarsh Singh ……………………

Manashjyoti Lahkar ……………………

Saptashi Sen ……………………

TABLE OF CONTENT

INTRODUCTION

1.1 Climate Change

Climate, defined as the long-term statistical behavior of weather, emerges from complex interactions among the atmosphere, hydrosphere, lithosphere, and biosphere. In recent decades, global and regional climate patterns have undergone significant transformations, largely driven by anthropogenic activities. These shifts are evident in rising temperatures, erratic precipitation patterns, and a growing frequency of extreme weather events.

According to the Intergovernmental Panel on Climate Change (IPCC, 2007), the increase in floods and droughts globally is likely to intensify the environmental nexus linking water, food, and energy security. Such climatic disruptions have serious implications for agriculture, infrastructure, and water availability, especially in climate sensitive regions like South Asia. For the Indian subcontinent, these changes may also lead to heightened water stress, reduced crop yields, and damage to critical infrastructure in the near future (Cruz et al., 2007).

1.3 Groundwater

Groundwater is the largest distributed store of freshwater, essential for sustaining life and economic development. In India, it provides over 60% of irrigation water and nearly 85% of rural drinking supply. Despite its importance, groundwater resources are increasingly under stress due to both natural variability and anthropogenic pressures.

The groundwater level (GWL) in any region is governed by a complex interplay of factors, which can be broadly categorized into following factors-

- **Climatic Factors**: Rainfall variability (total amount, frequency, intensity, and seasonal distribution), rising temperature, delayed or erratic monsoon, and extreme events (like droughts or heavy downpours) can lead to either insufficient recharge or excessive runoff.
- **Hydrogeological Factors**: Aquifer type & depth, soil texture, slope and topography determine how easily water can be soaked in and stay underground.
- **Anthropogenic Factors:** Over-extraction for irrigation, urbanization and land use change (e.g., increase in paved surfaces), unregulated borewell drilling all drain groundwater faster than it can refill.

1.4 Hydro-Climatic Vulnerability in Urban India: Focus on Guwahati, Patna, and Bhubaneswar

India's rapidly urbanizing districts are increasingly exposed to hydro-climatic stress due to the combined impacts of climate variability, population growth, and unsustainable groundwater extraction. The synergy between climate change and groundwater dynamics is complex—altered rainfall patterns, rising temperatures, and urban sprawl all influence groundwater recharge and extraction rates. These shifts are particularly evident in eastern and northeastern India, where cities rely heavily on the monsoon for aquifer replenishment.

This study focuses on **Guwahati (Assam), Patna (Bihar), and Bhubaneswar (Odisha)**—three urban centers with differing geographies but shared challenges of groundwater depletion and climate sensitivity:

- **Guwahati**, despite high annual rainfall, suffers from uneven aquifer recharge due to complex hydrogeology, seasonal flooding, and growing urban demand.
- **Patna**, located in the Ganga floodplain, experiences significant groundwater stress due to high irrigation demand and urban dependency, with marked pre-monsoon depletion.
  Bhubaneswar, situated in the coastal plains, faces reduced infiltration from increased built-up surfaces and rising heat stress, which disrupt monsoonal recharge.

By examining climatic and groundwater trends in these three cities, the study aims to understand how regional climate dynamics and local land use changes interact to drive urban groundwater stress.

LITERATURE REVIEW

S.N.	Title Name	Key-points
1.	Statistical investigation of long-term meteorological data to understand the variability in climate: a case study of Jharkhand, India Mahato et al. (2021)	Used to detect long-term climatic trends (temperature, rainfall) in Jharkhand over a century.
2.	Forecasting Maximum Temperature Trends with SARIMAX Mdpi.com (2024)	Model effect of temperature and rainfall on groundwater; includes exogenous inputs.
3.	Comparison of SARIMA and SARIMAX on Groundwater in Sulaymaniyah ResearchGate.net (2024)	SARIMAX provided better forecasting of groundwater when climatic variables included.
4.	Anomaly Detection using Isolation Forest GeeksforGeeks.org	Applies well to temperature and rainfall datasets for spotting anomalies.
5.	Climate Zone Classification Using Base Temperature IBPSA (2019)	Proposed climate zones using statistical indices and classification techniques.
6.	Application and comparison of different statistical methods for the analysis of groundwater levels over time: Response to rainfall and resource evolution in the Piedmont Plain (NW Italy) Susanna Mancini, et al. (2022)	Measures the magnitude of detected climate trends on Groundwater level.

OBJECTIVE

1. To detect long-term trends in temperature and rainfall data from 1970 to 2020 for Patna, Guwahati, and Bhubaneswar.
2. To classify and compare climatic regimes across five decades and to identify anomalous climatic events such as extreme rainfall or unusual temperature spikes.
3. To analyze the temporal dynamics of groundwater levels from 1996 to 2017 across the same cities and to interpret the spatial and seasonal response of groundwater to climatic variability across pre-monsoon, monsoon, post-monsoon kharif, and post-monsoon rabi periods.

METHODOLOGY

4.1 Study Area

This study focuses on three urbanizing and climatically sensitive districts in eastern and northeastern India: Guwahati (Assam), Patna (Bihar), and Bhubaneswar (Odisha). These cities represent distinct hydro-climatic zones but share growing dependence on groundwater and are increasingly vulnerable to climate variability.

Climatic Variables

Three key climate variables were analyzed:

- Maximum Temperature (°C): Monthly mean values
- Minimum Temperature (°C): Monthly mean values
- Rainfall (mm): Monthly cumulative totals

The data spans a 50-year period (1970–2019) to identify long-term trends using non-parametric statistical methods.

Groundwater Level Data

Seasonal groundwater level (GWL) data were obtained from dug well observations for each district. GWL data is categorized by four hydrological seasons:

- Pre-Monsoon
- Monsoon
- Post-Monsoon Kharif
- Post-Monsoon Rabi

The groundwater dataset covers the years 1996–2017.

To assess the climate–groundwater interaction, correlation analysis was conducted between seasonal GWL and corresponding monthly climatic parameters for the overlapping period (1996–2017). This approach helps quantify the relationship between groundwater fluctuations and climatic variability across urban centers experiencing growing water stress.

4.2 Methods Used

The climatic data was processed and analyzed statistically on monthly basis at different significant levels (<0.0001%, 1.0%, 5.0% and 10.0%) for each of the above-mentioned district. A pre-whitening test was employed to remove any serial correlation present in the climatic data. The overall trend (z-statistics) for the climatic data was estimated using Mann–Kendall (MK) test. The magnitude of trend was obtained using Sen's slope estimator. The relative change (RC) in all the climate variables was obtained at seasonal and annual time step at district level. All the aforementioned statistical tests utilized in the current study are discussed in the following sections.

4.2.1 Pre-whitening test:

The climatic data was processed using a pre-whitening test prior to applying the MK test. This was done to remove the presence of any serial correlation (persistence) present in the climatic data since the MK test assumes that the time series is uncorrelated.

4.2.2 Mann Kendall:

In this study, we aim to detect monotonic trends in long-term climatic variables over the period 1970–2019. To accomplish this, we employ the Mann–Kendall (MK) test, a widely used non parametric statistical method for trend analysis in climatology and hydrology. The MK test is particularly suitable for environmental time series data as it does not assume normality, is robust against outliers, and can effectively handle missing values. Its applicability to non-normally distributed datasets makes it a preferred choice for analyzing hydro-climatic trends.

The Mann-Kendall test statistic (S) was calculated according to:

$$S = \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} sgn(x_j - x_i)$$

where, $sgn(x_j - x_i) = \begin{Bmatrix} -1, if\ x_j - x_i > 0 \\ +1, if\ x_j - x_i < 0 \end{Bmatrix}$; 0 (otherwise)

and, n is the number of observed data series, x_j and x_i are the values at j and i periods respectively, j > i.

The statistic S is approximately normally distributed with the mean zero and a variance is expressed as follows:

$$Var(S) = \frac{n(n-1)(2n+5) - \sum_{i=1}^{m} t_i(2t_i+5)}{18}$$

where, n is the number of data points, m is the number of tied groups and t_i is the number of data points in the ith group.

The standardized test statistic Z is computed as follows:

$$Z = \begin{Bmatrix} \frac{S-1}{\sqrt{Var(S)}}, if\ S>0 \\ \frac{S+1}{\sqrt{Var(S)}}, if\ S<0 \end{Bmatrix}; 0\ \text{(otherwise)}$$

Positive and negative values of S indicate increasing and decreasing trends, respectively. The null hypothesis H_o, meaning that no significant trend is present, is accepted if the test statistic Z is not statistically significant, i.e., $-Z_{\alpha/2} < Z < +Z_{\alpha/2}$, where $Z_{\alpha/2}$ is the standard normal deviation. In this study, four different significance levels i.e., < 0.001%, 1%, 5% and 10% is considered.

4.2.3 Sen's Slope Estimator

To quantify the magnitude of climatic trends, Theil and Sen's slope estimator (Sen, 1968; Theil, 1950) was employed. This non-parametric method calculates the median slope of all possible pairwise comparisons in the time series, providing a robust estimate of the rate of change over time.

For a time series with values x_j and x_i at time steps j>i, the individual slopes T_i are computed as:

$$T_i = \frac{x_j - x_i}{j - i}$$

The Sen's slope (β) is the median of all T_i values:

- If N (number of slope values) is odd: $\beta = T_{\frac{(N+1)}{2}}$
- If N is even: $\beta = \frac{1}{2}\left(T_{\frac{N}{2}} + T_{\frac{(N+1)}{2}}\right)$

A positive β indicates an increasing trend, whereas a negative β suggests a declining trend in the time series.

4.2.4 Relative Change (RC)

The Relative Change (RC) quantifies the percentage change in seasonal and annual climate variables over the study period. It offers a normalized measure of trend magnitude relative to the average of the observed variable. The RC was calculated as follows:

$$RC = \frac{n \times \beta}{|x|}$$

where;

- n = length of the trend period (in years)
- β = trend magnitude estimated using Sen's slope
- |x| = absolute mean of the climatic variable over the period

This approach was applied to log-transformed seasonal and annual climate variables to ensure scale comparability across time and parameters.

4.2.5 Isolation Forest

Isolation Forest is an unsupervised machine learning algorithm used for anomaly detection, particularly effective for high-dimensional datasets. Unlike distance or density-based methods, it isolates anomalies instead of profiling normal data. The algorithm builds an ensemble of random binary trees (isolation trees). Each tree is created by randomly selecting a feature and a split value within its range. Since anomalies are fewer and different, they tend to be isolated in fewer splits, resulting in shorter average path lengths.

There is no traditional hypothesis test like a p-value or t-test in Isolation Forest. Instead, the model outputs an anomaly score typically ranging from 0 to 1:

Score ≈ 1.0 → Strong anomaly

Score ≈ 0.5 → Borderline

Score ≈ 0.0 → Normal observation

4.2.6 Chi-Square test

The Chi-Square Test is a statistical method used to determine whether there is a significant association between two categorical variables. It compares the observed frequencies in each category to the expected frequencies under the assumption of no change (null hypothesis).

H_0: There is no association between the two categorical variables.

H_1: There is a significant association between the two categorical variables.

The test statistic is calculated as:

$$\chi^2 = \sum \frac{\left(O_{ij} - E_{ij}\right)^2}{E_{ij}}$$

where,

O_{ij}: Observed frequency

E_{ij}: Expected frequency

If the Chi-square statistic is large, and p-value is small, it means the observed distribution differs significantly from the expected i.e., there is likely a relationship between the variables

4.2.7 STL decomposition

STL decomposition is a time series analysis method that separates the original series into trend, seasonal, and residual (irregular) components. The residuals representing short-term deviations from expected seasonal and long-term behavior were retained to isolate anomaly driven variability.

The additive model can be expressed as:

$$Y_t = T_t + S_t + R_t$$

where,

T_t= Trend component

S_t= Seasonal component

R_t= Residual component

4.2.8 Augmented Dickey-Fuller (ADF) test

The Augmented Dickey-Fuller (ADF) test is a statistical method used to assess the stationarity of a time series. Stationarity implies that the properties of the series such as mean, variance, and autocorrelation remain constant over time.

- Null Hypothesis (H_0): The time series is non-stationary.

- Alternative Hypothesis (H_1): The time series is stationary.

If the **p-value** is less than the chosen significance level (commonly 0.05), the null hypothesis is rejected, indicating that the time series is stationary. If not, the series may require differencing or detrending before further analysis.

4.2.9 Ljung-Box test

The Ljung-Box test is a diagnostic statistical tool used to examine the presence of autocorrelation in a time series. It evaluates whether the residuals (errors) from a fitted model (such as ARIMA/SARIMAX) are independently distributed or exhibit significant correlation across lags.

This test is particularly important for validating the adequacy of time series models. If residuals are autocorrelated, the model may not have captured all patterns in the data.

- H_0: The residuals are independently distributed (no autocorrelation).
- H_1: The residuals are not independently distributed (autocorrelation exists).

The Ljung-Box Q-statistic is defined as:

$$Q = n(n+2)\sum_{k=1}^{h} \frac{\left(\dot{\rho}_k\right)^2}{n-k}$$

where,

n = number of observations

h = number of lags tested

$\dot{\rho}_k$ = sample autocorrelation at lag k

A p-value greater than the significance level (e.g., 0.05) suggests that the residuals are uncorrelated, confirming model adequacy. A small p-value indicates that autocorrelation remains in the residuals and further model refinement is needed.

4.2.10 SARIMAX model

The SARIMAX model is an extension of the classical ARIMA model that includes both seasonality and exogenous variables. It is widely used for modeling and forecasting time series data where external regressors (climatic or other influencing variables) and seasonal patterns significantly influence the target variable.

- Captures both trend and seasonality in groundwater level data.
- Models the impact of climatic variables directly.
- Suitable for handling multivariate and seasonal time series.

DATA DESCRIPTION

This study investigates long-term climatic and groundwater data from three urbanizing Indian cities—**Guwahati (Assam)**, **Patna (Bihar)**, and **Bhubaneswar (Odisha)**—to analyze hydro-climatic variability and groundwater trends.

5.1 Climatic Data

- **Source:** Daily meteorological data from the India Meteorological Department (IMD), aggregated to monthly values for trend analysis.
- **Station Indexes:**
  - Guwahati: 42410
  - Patna: 42492
  - Bhubaneswar: 42971
- **Time Period:** 1970–2019
- **Parameters:**
  - Maximum Temperature (°C): Monthly mean of daily maximum temperatures
  - Minimum Temperature (°C): Monthly mean of daily minimum temperatures
  - Rainfall (mm): Monthly cumulative rainfall

Trend detection was performed on these monthly datasets. Additionally, seasonal (Premonsoon, Monsoon, Postmonsoon Kharif, Postmonsoon Rabi) and annual aggregations were used for relative change analysis and correlation studies.

5.2 Groundwater Level (GWL) Data

- **Source:** Central Ground Water Board (CGWB), Ministry of Jal Shakti, Government of India
- **Time Period:** 1996–2017
- **Data Type:** Seasonal dugwell groundwater level observations (measured in meters below ground level)
- **Frequency:** Seasonal readings corresponding to Premonsoon, Monsoon, Postmonsoon Kharif, and Postmonsoon Rabi periods

Seasonal groundwater levels were correlated with climatic parameters to evaluate the influence of hydro-climatic variability on aquifer recharge and depletion.

RESULT AND CONCLUSION

6.1 Plots of the linear trends in Maximum temperature, Minimum temperature and Rainfall of Guwahati, Patna and Bhuvneshwar.

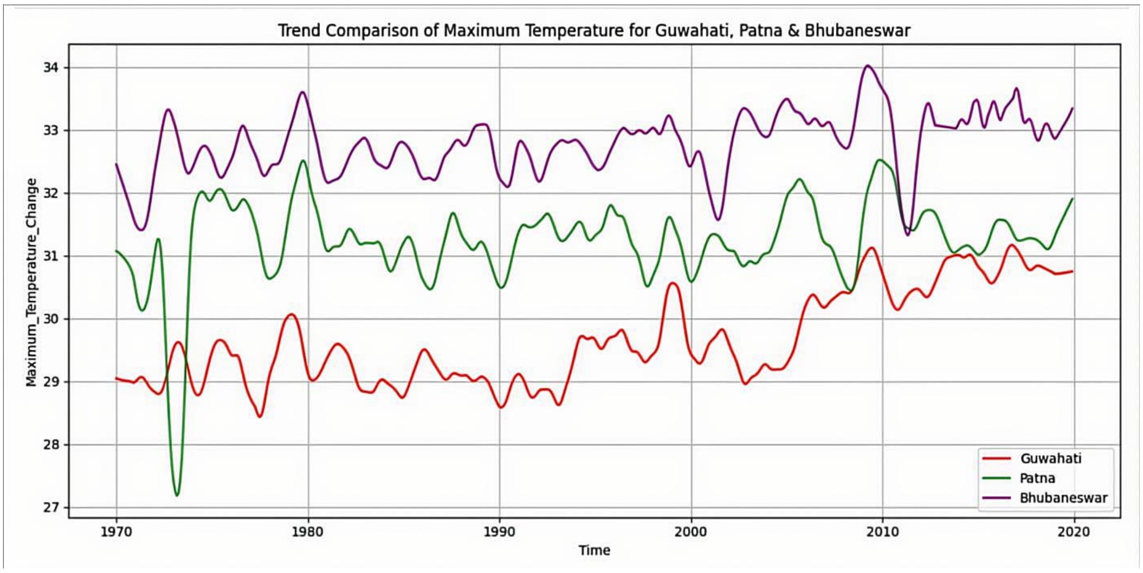


Fig 6.1: Trend comparison of Maximum temperature

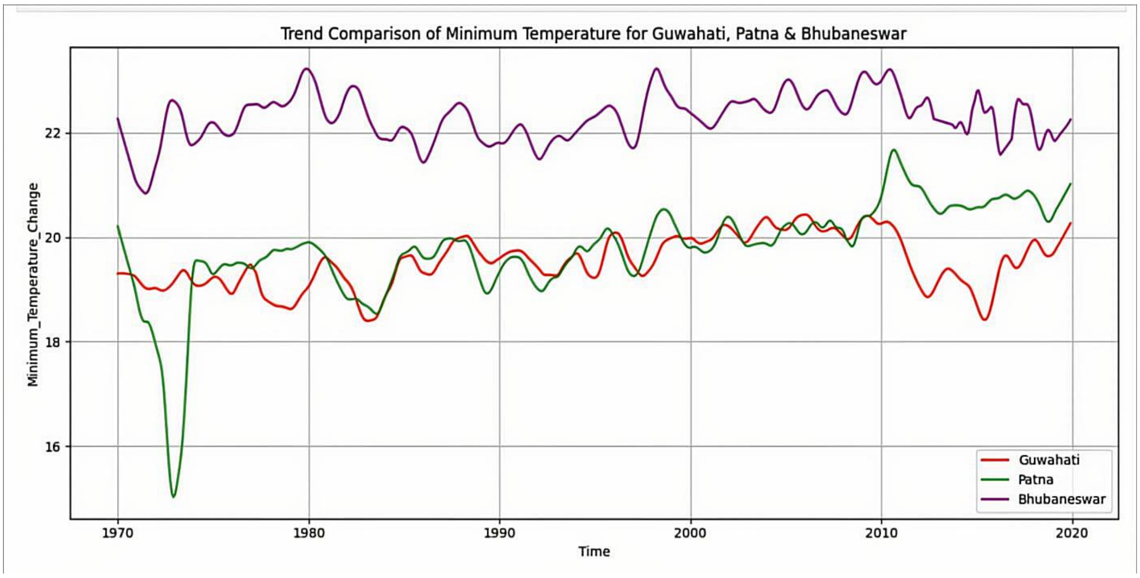


Fig 6.2: Trend comparison of Minimum temperature

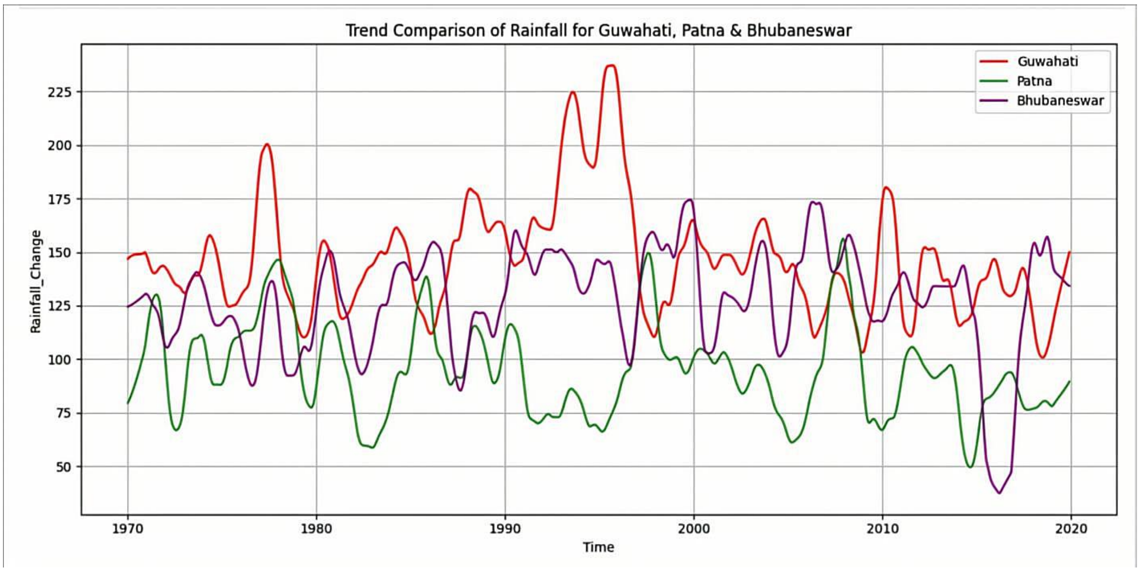


Fig 6.3: Trend comparison of Rainfall

6.2 Statistical summary of monthly Maximum temperature, Minimum temperature and Rainfall of Guwahati, Patna and Bhuvneshwar.

Month_Name	MAX_MKz	MAX_sig	MAX_slope	MAX_RC	MIN_MKz	MIN_sig	MIN_slope	MIN_RC	RAINFALL_MKz	RAINFALL_sig	RAINFALL_slope	RAINFALL_RC
January	4.89	5%	0.03	0.17	1.19	NS	0.01	0.2	-0.06	NS	-0.06	0.06
February	6.45	5%	0.04	0.2	3.69	5%	0.03	0.21	-0.16	NS	-0.17	0.01
March	4.23	10%	0.02	0.04	3.88	1%	0.03	0.17	0.04	NS	0.04	-0.05
April	2.36	NS	0.01	0.02	1.06	NS	0.01	-0.18	1.19	NS	1.24	-0.05
May	4.23	5%	0.02	0.3	0.93	NS	0.01	0.44	0.62	NS	0.64	-0.04
June	5.35	1%	0.03	0.27	1.87	5%	0.01	0.3	-0.78	NS	-0.82	0.07
July	4.99	1%	0.03	0.47	2.27	1%	0.02	0.34	-2.85	1%	-2.96	0.06
August	3.67	1%	0.02	0.46	2.10	5%	0.01	0.49	-0.95	NS	-0.99	0.02
September	4.05	1%	0.02	0.48	1.87	5%	0.01	0.51	0.14	NS	0.14	0.01
October	4.64	5%	0.03	0.49	1.30	NS	0.01	0.2	0.33	NS	0.35	-0.13
November	5.10	1%	0.03	0.41	1.43	NS	0.01	0.29	-0.32	5%	-0.34	0.04
December	4.20	1%	0.02	0.39	3.03	10%	0.02	0.24	-0.05	NS	-0.06	-0.21

Table 6.1: The MKz value, significance level, slope and relative change of monthly climate of Guwahati

Guwahati exhibits a consistent warming trend in both maximum and minimum temperatures throughout the year, with many months showing statistically significant increases. Rainfall trends are mostly non-significant, except for a notable decrease in July, which may indicate emerging dry spells during monsoon. These findings suggest a shift toward warmer and potentially drier conditions, especially during the rainy season.

Month_Name	MAX_MKz	MAX_sig	MAX_slope	MAX_RC	MIN_MKz	MIN_sig	MIN_slope	MIN_RC	RAINFALL_MKz	RAINFALL_sig	RAINFALL_slope	RAINFALL_RC
January	-1.58	0.05	-0.02	-0.05	1.66	0.10	0.02	0.11	0.02	NS	0.01	-0.09
February	1.10	NS	0.02	0.02	2.44	0.05	0.04	0.37	0.01	NS	0.01	-0.26
March	0.00	NS	0.00	-0.08	3.59	< 0.001	0.05	0.16	0.00	NS	0.00	0.01
April	-0.68	NS	-0.01	-0.17	3.93	< 0.001	0.06	0.36	0.05	NS	0.03	-0.02
May	1.91	0.10	0.03	0.35	2.53	0.01	0.04	0.31	-0.51	NS	-0.30	0.12
June	1.51	NS	0.02	-0.12	1.85	0.01	0.03	0.15	1.62	NS	0.97	-0.10
July	0.42	NS	0.01	0.24	1.29	0.01	0.02	0.30	-2.71	NS	-1.62	0.20
August	0.22	NS	0.00	0.26	1.33	0.01	0.02	0.46	-1.01	NS	-0.60	-0.08
September	0.51	NS	0.01	0.12	1.08	0.01	0.02	0.39	-0.03	NS	-0.02	0.20
October	-0.28	NS	0.00	0.05	1.52	0.05	0.02	0.20	-0.84	NS	-0.50	0.19
November	-0.21	NS	0.00	0.10	0.62	NS	0.01	0.35	0.00	NS	0.00	0.29
December	-0.81	0.10	-0.01	0.12	2.42	< 0.001	0.04	0.19	0.00	NS	0.00	0.16

Table 6.2: The MKz value, significance level, slope and relative change of monthly climate of Patna

Patna shows a consistent and significant rise in minimum temperatures throughout the year, especially in spring and winter months, highlighting warming nights. Maximum temperature trends are weaker, except for a significant warming in May. Rainfall trends are mostly non-significant, but there's a strong decreasing signal in July, which could impact monsoon season planning. The climate pattern suggests warming conditions with possible dry spells in peak monsoon.

Month_Name	MAX_MKz	MAX_sig	MAX_slope	MAX_RC	MIN_MKz	MIN_sig	MIN_slope	MIN_RC	RAINFALL_MKz	RAINFALL_sig	RAINFALL_slope	RAINFALL_RC
January	1.35	NS	0.02	-0.02	-0.03	NS	0.00	-0.06	0.00	NS	0.00	0.13
February	3.72	0.01	0.06	0.17	-0.26	NS	0.00	-0.11	-0.14	0.10	-0.15	-0.01
March	1.87	0.05	0.03	-0.19	2.47	0.05	0.02	-0.02	-0.05	NS	-0.05	0.00
April	1.12	NS	0.02	-0.19	1.91	0.10	0.01	0.15	0.02	NS	0.02	-0.05
May	0.44	NS	0.01	-0.08	2.30	0.05	0.02	0.05	1.21	0.05	1.29	-0.11
June	0.25	NS	0.00	0.11	2.45	0.05	0.02	0.02	-0.08	NS	-0.09	-0.09
July	0.46	NS	0.01	-0.17	1.49	0.05	0.01	0.35	1.56	NS	1.66	0.01
August	1.39	0.01	0.02	0.00	1.06	NS	0.01	0.49	-1.74	NS	-1.85	0.11
September	0.77	NS	0.01	-0.03	1.21	0.05	0.01	0.44	2.13	0.10	2.27	0.13
October	-0.21	NS	0.00	0.07	1.38	NS	0.01	0.04	-0.31	NS	-0.33	0.08
November	1.11	NS	0.02	0.11	1.94	NS	0.01	0.18	0.00	NS	0.00	0.05
December	1.19	NS	0.02	-0.25	3.09	NS	0.02	0.20	0.00	NS	0.00	-0.15

Table 6.2: The MKz value, significance level, slope and relative change of monthly climate of Bhuvneshwar

Bhubaneswar exhibits a strong warming trend, especially in minimum temperatures, confirming increasing nighttime heat. Significant maximum temperature increases occur in February, March, and August, suggesting early and prolonged heat conditions. Rainfall patterns are shifting, with September and May showing significant upward trends, potentially altering monsoon dynamics. Overall, the region is experiencing notable warming and changing rainfall timing, with implications for urban heat, water stress, and crop calendars.

6.3 Temporal shifts in regional climate types across Guwahati, Patna and Bhuvneshwar.

region	Cool & Dry	Cool & Wet	Hot & Dry	Hot & Wet
decade				
1970	49	3	21	46
1980	50	3	18	49
1990	49	3	12	56
2000	46	3	23	48
2010	49	1	21	49

Table 6.4: Regional climate frequencies of Guwahati

region	Cool & Dry	Cool & Wet	Hot & Dry	Hot & Wet
decade				
1970	51	0	26	43
1980	48	2	36	34
1990	49	0	34	37
2000	48	0	34	37
2010	45	0	37	38

Table 6.5: Regional climate frequencies of Patna

region decade	Cool & Dry	Cool & Wet	Hot & Dry	Hot & Wet
1970	41	6	38	35
1980	41	3	38	38
1990	41	6	28	45
2000	41	4	38	37
2010	36	3	26	37

Table 6.6: Regional climate frequencies of Bhuvneshwar

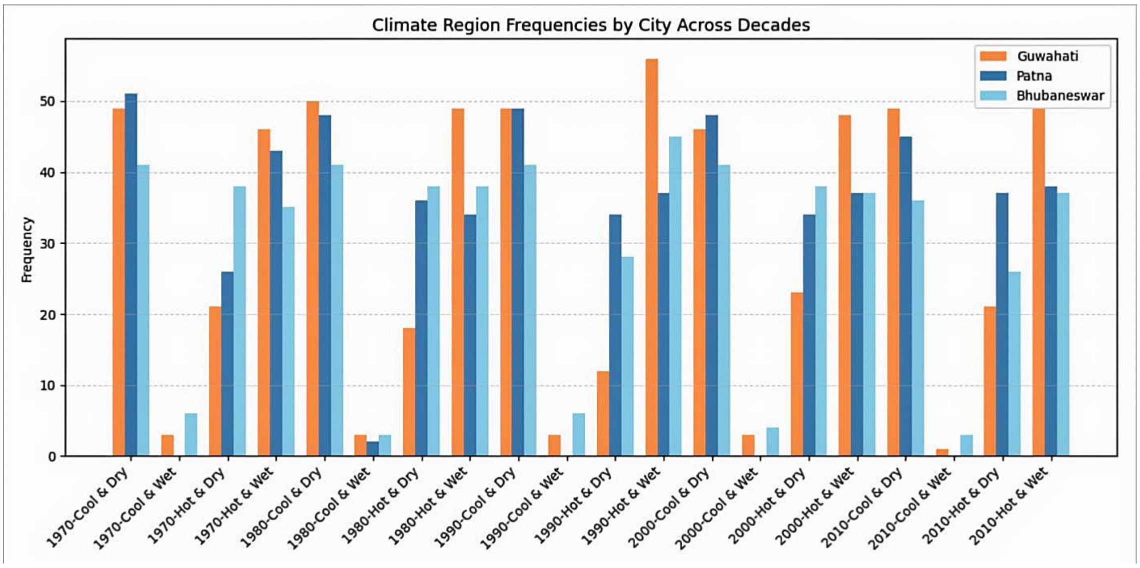


Fig 6.4: Regional climate frequencies

Decade-wise Analysis:

•1970s: Guwahati has a high frequency of Cool & Dry climate, while Patna and Bhubaneswar have a relatively lower frequency. The Hot & Wet climate is more frequent in Patna and Bhubaneswar.

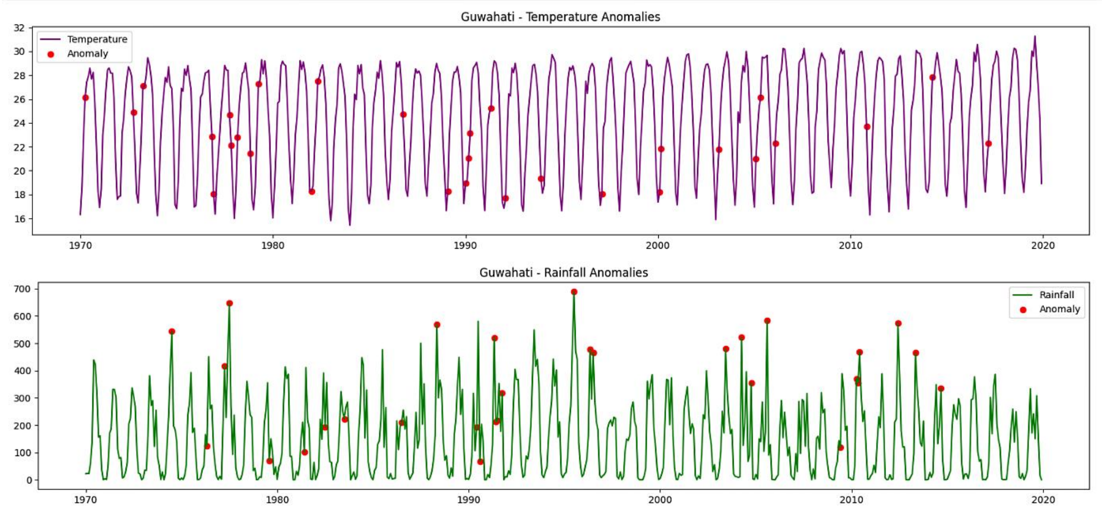
•1980s: The frequency of Cool & Dry climate decreases in Guwahati, while it increases in Patna and Bhubaneswar. The Hot & Wet climate remains relatively high in Patna and Bhubaneswar.

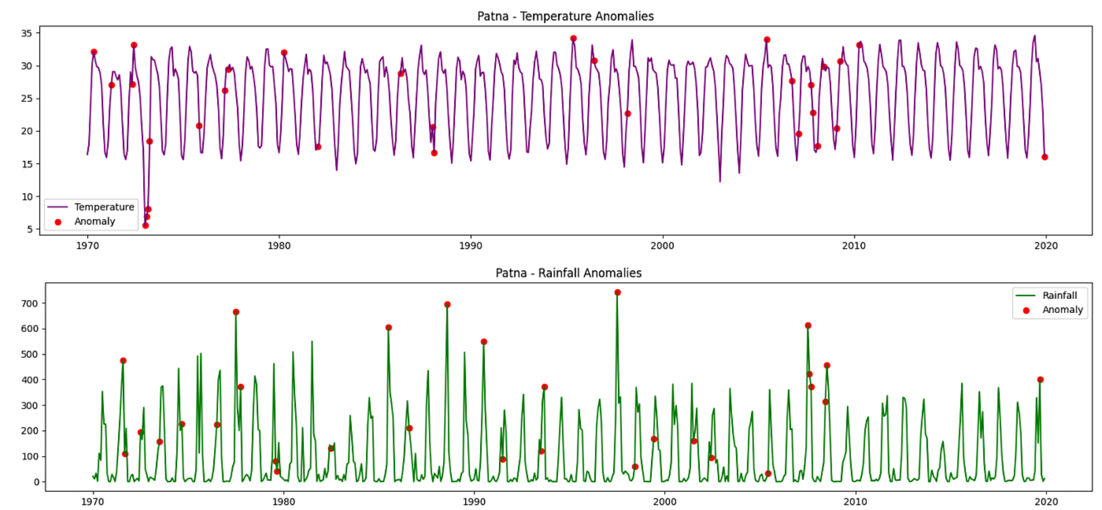
•1990s: The frequency of Hot & Dry climate increases in all three cities. Guwahati shows a significant increase in Hot & Dry climate.

•2000s: The frequency of Cool & Wet climate increases in all three cities. Bhubaneswar shows a significant increase in Cool & Wet climate.

•2010s: The frequency of Hot & Wet climate remains relatively high in Patna and Bhubaneswar. Guwahati shows a decrease in Cool & Dry climate.

6.4 Detection of temperature and rainfall anomalies (1970-2019)





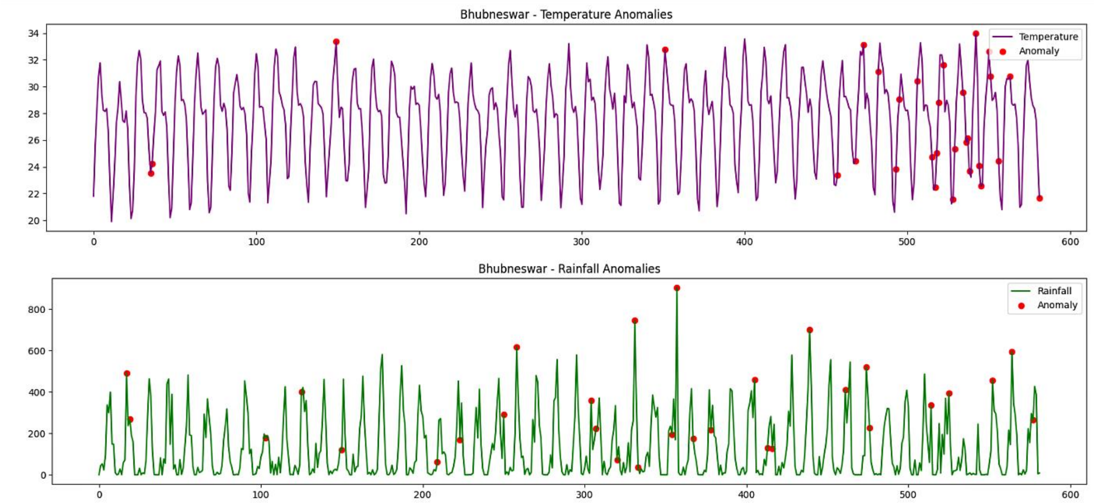


Fig 6.5(c): Temperature and Rainfall anomalies of Guwahati, Patna and Bhuvneshwar

Top 5 temp. anomalies in Guwahati:

- April 2014 – 27.86 °C
- November 1976 – 22.84 °C
- March 1978 – 22.83 °C
- February 2006 – 22.30 °C
- January 1990 – 18.97 °C

Top 5 rainfall anomalies in Guwahati:

- July 1995 – 688.56 mm
- August 2005 – 583.53 mm
- May 1988 – 569.70 mm
- April 2004 – 520.90 mm
- April 1977 – 416.94 mm

Top 5 temp. anomalies in Patna:

- April 2010 – 33.21 °C
- May 1970 – 32.16 °C
- April 2009 – 30.69 °C
- November 2007 – 22.78 °C
- January 1988 – 20.62 °C

Top 5 rainfall anomalies in Patna:

- July 1997 – 742.10 mm
- August 1988 – 694.90 mm
- July 1985 – 605.10 mm
- July 2008 – 456.97 mm
- September 2019 – 401.36 mm

Top 5 temp. anomalies in Bhubaneswar:

- March 2015 – 33.97 °C

Top 5 rainfall anomalies in Bhubaneswar:

- September 1999 – 903.3 mm

- April 2012 – 31.64 °C
- February 2011 – 30.77 °C
- December 2013 – 29.56 °C
- August 2018 – 21.69 °C

- July 1996 – 745.0 mm
- September 2017 – 594.8 mm
- June 2012 – 457.0 mm
- September 2006 – 394.3 mm

6.5 Response of the Temperature and Rainfall on Groundwater Level

```
Best ARIMA order: (4, 0, 4)
                               SARIMAX Results
==============================================================================
Dep. Variable:                 gw_res   No. Observations:                   88
Model:               SARIMAX(4, 0, 4)   Log Likelihood                  14.634
Date:                Sat, 31 May 2025   AIC                             -7.268
Time:                        20:46:23   BIC                             19.983
Sample:                             0   HQIC                             3.711
                                 - 88
Covariance Type:                  opg
==============================================================================
                 coef    std err          z      P>|z|      [0.025      0.975]
------------------------------------------------------------------------------
rf_res         0.0003      0.000      0.831      0.406      -0.000       0.001
temp_res       0.1185      0.123      0.966      0.334      -0.122       0.359
ar.L1         -0.6241      0.090     -6.951      0.000      -0.800      -0.448
ar.L2         -0.6203      0.113     -5.503      0.000      -0.841      -0.399
ar.L3         -0.4550      0.108     -4.209      0.000      -0.667      -0.243
ar.L4         -0.4322      0.100     -4.322      0.000      -0.628      -0.236
ma.L1         -0.0200      0.495     -0.040      0.968      -0.990       0.950
ma.L2          0.0054      0.478      0.011      0.991      -0.932       0.943
ma.L3         -0.0615      0.464     -0.133      0.894      -0.970       0.847
ma.L4         -0.9067      0.438     -2.070      0.038      -1.765      -0.048
sigma2         0.0362      0.017      2.100      0.036       0.002       0.070
===================================================================================
Ljung-Box (L1) (Q):                   0.05   Jarque-Bera (JB):                 3.14
Prob(Q):                              0.82   Prob(JB):                         0.21
Heteroskedasticity (H):               1.71   Skew:                             0.38
Prob(H) (two-sided):                  0.15   Kurtosis:                         3.54
===================================================================================

Warnings:
[1] Covariance matrix calculated using the outer product of gradients (complex-step).
```

Fig 6.6: SARIMAX model fitting for groundwater level of Guwahati

We have applied a SARIMAX (4, 0, 4) model using STL residuals of groundwater as the dependent variable, and STL residuals of rainfall and temperature as exogenous variables. This approach isolates short-term variability by removing seasonal and trend effects.

The results show that rainfall and temperature residuals are not statistically significant predictors of short-term groundwater fluctuations ($p > 0.05$). However, the autoregressive terms (AR1–AR4) are highly significant ($p < 0.001$), indicating strong influence of past groundwater levels on current values.

```
Best ARIMA order: (4, 0, 4)
                               SARIMAX Results
==============================================================================
Dep. Variable:                 gw_res   No. Observations:                   88
Model:               SARIMAX(4, 0, 4)   Log Likelihood                   3.883
Date:                Sat, 31 May 2025   AIC                             14.234
Time:                        20:47:41   BIC                             41.485
Sample:                             0   HQIC                            25.213
                                 - 88
Covariance Type:                  opg
==============================================================================
                 coef    std err          z      P>|z|      [0.025      0.975]
------------------------------------------------------------------------------
rf_res        -0.0013      0.001     -2.259      0.024      -0.002      -0.000
temp_res      -0.0277      0.142     -0.195      0.845      -0.306       0.251
ar.L1         -0.4404      0.124     -3.554      0.000      -0.683      -0.198
ar.L2         -0.5749      0.120     -4.799      0.000      -0.810      -0.340
ar.L3         -0.4895      0.146     -3.343      0.001      -0.777      -0.203
ar.L4         -0.4098      0.135     -3.047      0.002      -0.673      -0.146
ma.L1         -0.1125      0.137     -0.820      0.412      -0.382       0.157
ma.L2          0.0733      0.155      0.474      0.635      -0.230       0.377
ma.L3          0.1887      0.178      1.060      0.289      -0.160       0.538
ma.L4         -0.8437      0.209     -4.038      0.000      -1.253      -0.434
sigma2         0.0500      0.008      6.262      0.000       0.034       0.066
===================================================================================
Ljung-Box (L1) (Q):                   0.01   Jarque-Bera (JB):                 5.51
Prob(Q):                              0.94   Prob(JB):                         0.06
Heteroskedasticity (H):               0.52   Skew:                             0.60
Prob(H) (two-sided):                  0.08   Kurtosis:                         3.27
===================================================================================

Warnings:
[1] Covariance matrix calculated using the outer product of gradients (complex-step).
```

Fig 6.7: SARIMAX model fitting for groundwater level of Patna

In this SARIMAX (4, 0, 4) model, we used STL residuals of groundwater as the dependent variable and STL residuals of rainfall and temperature as exogenous predictors to focus on short-term variability.

Rainfall residuals show a small but statistically significant negative effect on groundwater residuals ($p = 0.024$), suggesting a weak short-term inverse relationship. Temperature residuals remain non-significant ($p = 0.848$), indicating minimal direct short-term influence. Again, the autoregressive terms (AR1–AR4) are highly significant, showing strong temporal dependence in groundwater levels.

```
Best ARIMA order: (4, 0, 4)
                               SARIMAX Results
==============================================================================
Dep. Variable:                 gw_res   No. Observations:                   88
Model:               SARIMAX(4, 0, 4)   Log Likelihood                  10.783
Date:                Sat, 31 May 2025   AIC                              0.435
Time:                        20:49:18   BIC                             27.686
Sample:                             0   HQIC                            11.413
                                 - 88
Covariance Type:                  opg
==============================================================================
                 coef    std err          z      P>|z|      [0.025      0.975]
------------------------------------------------------------------------------
rf_res        -0.0018      0.000     -5.543      0.000      -0.002      -0.001
temp_res       0.0950      0.062      1.537      0.124      -0.026       0.216
ar.L1         -0.6273      0.140     -4.477      0.000      -0.902      -0.353
ar.L2         -0.4975      0.126     -3.957      0.000      -0.744      -0.251
ar.L3         -0.5587      0.095     -5.905      0.000      -0.744      -0.373
ar.L4         -0.4595      0.101     -4.555      0.000      -0.657      -0.262
ma.L1         -0.1803      3.219     -0.056      0.955      -6.489       6.129
ma.L2         -0.1435      2.677     -0.054      0.957      -5.390       5.103
ma.L3          0.1167      2.210      0.053      0.958      -4.214       4.448
ma.L4         -0.7898      2.549     -0.310      0.757      -5.786       4.206
sigma2         0.0417      0.134      0.311      0.756      -0.221       0.305
===================================================================================
Ljung-Box (L1) (Q):                   0.25   Jarque-Bera (JB):                 1.58
Prob(Q):                              0.62   Prob(JB):                         0.45
Heteroskedasticity (H):               0.57   Skew:                             0.16
Prob(H) (two-sided):                  0.13   Kurtosis:                         3.58
===================================================================================

Warnings:
[1] Covariance matrix calculated using the outer product of gradients (complex-step).
```

Fig 6.8: SARIMAX model fitting for groundwater level of Bhuvneshwar

This SARIMAX (4, 0, 4) model analyzes the relationship between STL residuals of groundwater (dependent variable) and residuals of rainfall and temperature (exogenous inputs).

Rainfall residuals have a statistically significant negative effect on groundwater residuals ($p < 0.001$), suggesting that short-term rainfall variability may lead to immediate but inverse fluctuations in groundwater, possibly due to delayed infiltration or groundwater recharge processes. Temperature residuals, however, are not statistically significant ($p = 0.124$). Autoregressive terms (AR1–AR4) are all highly significant, reinforcing that past groundwater anomalies strongly influence future values. The moving average (MA) components are not significant, suggesting limited short-term shock correction in this setup.

REFERENCES

- ICRISAT level of groundwater Dataset: International Crops Research Institute for the Semi-Arid Tropics (ICRISAT). (n.d.). Additional Datasets. https://data.icrisat.org/dld/src/additional.html
- IMD temperature and rainfall dataset: India Meteorological Department. (n.d.). Mausam: Weather Information Portal. https://mausam.imd.gov.in/
- Arora, M., Goel, N. K., & Singh, P. (2005). Evaluation of temperature trends over India. Hydrological Sciences Journal, 50(1), 81–93.
- Ashok, K., Behera, S. K., Rao, S. A., Weng, H., & Yamagata, T. (2007). El Niño Modoki and its possible teleconnection. Journal of Geophysical Research: Oceans, 112(C11), 1–27. https://doi.org/10.1029/2006JC003798
- Bahadur, R., Jaiswal, R. K., Nema, A. K., Gangwar, A., & Kumar, S. (2020). Trends analysis of rainfall and temperature over Nagwan Watershed, Hazaribagh District, Jharkhand. Current Journal of Applied Science and Technology, 66, 112–128.
- Kothyari, U. C., & Singh, V. P. (1996). Rainfall and temperature trends in India. Hydrological Processes, 10(3), 357–372.
- Peixeiro, M., Santos, J., & Pereira, M. J. (2024). Time-series forecasting of particulate organic carbon on the Sunda Shelf using SARIMAX. Environmental Technology & Innovation, 38, 103496. https://doi.org/10.1016/j.eti.2024.103496
- Wang, L., & Zhang, Y. (2024). Forecasting maximum temperature trends with SARIMAX. Sustainability, 16(16), 7183. https://doi.org/10.3390/su16167183
- Singh, R., & Kumar, A. (2023). A novel WD-SARIMAX model for temperature forecasting using wavelet decomposition. Sustainability, 15(1), 757. https://doi.org/10.3390/su15010757
- DataScientest. (2024). SARIMAX model: What is it? How can it be applied to time series? Retrieved from https://datascientest.com/en/sarimax-model
- Sharma, P., & Gupta, S. (2024). Forecasting temperature trends using SARIMAX: A case study in India. Preprints, 2024041234. https://doi.org/10.20944/preprints202404.1234.v1