SUMMER INTERNSHIP REPORT

ON

Rainfall as a Driver of Electricity Demand Variability: State and Sector-Level Evidence from India

And

Development of the National Trilemma Tool Web Page

submitted in partial fulfilment for the award of the

B. Tech Bioinformatics

By

SANSKRITI GOYAL

A0504222058

Under the guidance of

Dr. Abhishek Sengupta

(Faculty Mentor)

Vivek Kumar Gupta

(Industry Mentor)

AMITY INSTITUTE OF BIOTECHNOLOGY

Amity University Uttar Pradesh, Sector – 125 | NOIDA - 201 303| INDIA

STUDENT DECLARATION

I, hereby declare that the report titled **Rainfall as a Driver of Electricity Demand Variation in India's Agriculture-Dominated Regions: A Sectoral and State-Level Analysis** has been prepared by me as part of the credit requirement for the course work of the program towards the partial fulfilment of the degree B. Tech Bioinformatics, NOIDA.

This report is an original work done by me and has not been copied or duplicated from any source. I have given due reference and credit to the original authors / entities from where relevant data & inputs have been taken as part of my report.

I, also declare that no part of this report has not been submitted by me in any form or manner to any other institute or organization as part of any field of study or purpose.
In case of any objections or observations are made pertaining to the information / details mentioned in my report, I take the ownership to be held responsible for the same and request an opportunity to defend my work.

Name of the Student: Sanskriti Goyal

Signature of the Student:
Enrollment No: A0504222058
Term:
Date:

Place: NOIDA

ACKNOWLEDGMENTS

A summer project is a valuable opportunity for learning and growth. I am profoundly fortunate and honoured to have been supported by numerous outstanding individuals throughout this endeavour.

I would first like to express my sincere gratitude to **Mr. Vivek Kumar Gupta,** Assistant General Manager, NTPC for providing me with this valuable opportunity to contribute to such a meaningful project at **WEC India**. I am truly grateful for the trust he placed in me, as well as for his constant support, approachability, and thoughtful guidance throughout.

I extend my heartfelt gratitude to **Mr. Nishant Kumar Sharma**, Deputy General Manager, who, despite his extensive responsibilities, found time to guide and steer me on the right path. His invaluable counsel has been crucial to my journey, and for that, I am immensely thankful. Similarly, I am deeply appreciative of **Mr. Shubham Kumar**, Research Consultant who has diligently monitored, helped, and provided me with essential domain knowledge throughout this project.

I am deeply thankful to Prof. **Dr. Abhishek Sengupta**, whose patience I have undoubtedly tested. His involvement, knowledge sharing, and encouragement to think critically have been pivotal to my learning process.

Lastly, I am grateful to the many individuals who shared their valuable insights and information, contributing to the successful completion of this.

Table of Contents

PHASE 2: SECTOR-WISE ANALYSIS OF ELECTRICITY CONSUMPTION AND ECONOMIC ACTIVITY

PHASE 3: MACHINE LEARNING-BASED PREDICTION OF ELECTRICITY DEMAND

PROJECT 2

Development of the National Trilemma Tool Web Page

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1: INTRODUCTION

1.1 Background and Motivation

India's electricity demand is intricately linked to its climatic and economic landscape. As a country where agriculture accounts for a significant share of rural livelihoods and national GDP, the monsoon's variability directly influences both crop output and the energy required for irrigation. According to the Institute of Economic Growth, "agriculture, known to be particularly sensitive to rainfall, helps determine GDP by its own performance and through its linkages with the industrial sector"[1].The monsoon season, which delivers over 70% of India's annual rainfall, is thus a critical determinant of agricultural productivity and, by extension, the country's energy consumption patterns. [2]

Electricity demand in India has grown rapidly over the past two decades, with the agricultural sector emerging as a major consumer, primarily due to the proliferation of electric pumps for groundwater irrigation[3]. The relationship between rainfall and agricultural electricity demand is complex. While increased rainfall typically reduces the need for irrigation—and thus electricity consumption—in agriculture-heavy states, the expansion of irrigated area and the rising number of pump sets have made the sector less sensitive to short-term rainfall fluctuations[3], [4]. Nevertheless, regional studies and empirical analyses continue to find that rainfall variability remains a significant driver of year-to-year changes in agricultural electricity demand, especially in states with limited irrigation infrastructure or high dependence on groundwater extraction[1], [4].

The importance of understanding this dynamic is underscored by the ongoing depletion of groundwater resources and the need for efficient energy and water management. Recent research in Hydrology and Earth System Sciences notes a "very strong negative correlation with the electricity consumption for agricultural usage, which may also be considered as a proxy for groundwater pumped for irrigation in a region"[4]. This nexus between water, food, and energy is increasingly central to India's sustainable development agenda.

1.2 Research Objective and Scope

The primary objective of this study is to develop a predictive framework for forecasting electricity demand in India based on rainfall forecasts, with the aim of enabling more effective electricity planning and allocation. This is particularly important in agriculture-heavy regions, where monsoon rainfall significantly influences irrigation needs and, consequently, electricity consumption for groundwater pumping and related agricultural activities. The study also seeks to analyse the sectoral and state-level variations in this rainfall-electricity relationship to inform targeted planning and policy decisions.

Specific objectives include:

- **Quantitative Assessment of Rainfall–Electricity Demand Linkages:**
  To empirically measure the correlation between rainfall deviations and sectoral electricity demand, especially in agriculture. Previous studies have shown that rainfall shocks significantly affect farmers' income and their adaptation strategies, which in turn influence energy use in irrigation[5], [8].

- **Regional and Sectoral Heterogeneity:**
  To identify and compare the sensitivity of electricity demand to rainfall across different states and sectors (agriculture, commercial, industrial, residential). For example, the Central Electricity Authority (CEA) tracks the number of agricultural electricity consumers at the state level, highlighting the regional variation in dependence on electric irrigation[9].

- **Integration of Economic Indicators:**
  To explore the interplay between rainfall, electricity use, and economic variables such as sectoral Gross State Domestic Product (SGDP). The water-energy-food (WEF) nexus literature emphasizes the interconnectedness of these domains, especially in groundwater-dependent, rice-based production systems[6].

- **Advanced Forecasting and Analytical Methods:**
  To apply machine learning and improved statistical models for forecasting electricity demand using rainfall, SGDP, and other relevant variables. The Government of India is currently revamping its demand forecasting, emphasizing the need for granular, weather-linked models to ensure grid stability and better planning[10].

- **Policy-Relevant Recommendations:**
  To generate actionable insights for policymakers, grid operators, and agricultural planners, aiming to improve demand forecasting, grid reliability, and climate resilience in the power

sector[11], [12].

Scope:

The study leverages high-resolution, multi-year datasets from the India Meteorological Department (IMD) for rainfall[7], the CEA for electricity sales and consumer data, and PPAC for economic indicators. The analysis spans at least two decades and focuses on both direct and indirect effects of rainfall on electricity demand, with a primary emphasis on agriculture but also considering commercial, industrial, and residential sectors.[4], [5], [12]

This comprehensive approach ensures that the study not only quantifies the direct effects of rainfall on electricity demand but also situates them within the broader economic and policy context.

1.3 Hypothesis

Central Hypothesis:

"In agriculture-intensive states of India, increased rainfall generally leads to a reduction in electricity demand for agricultural activities, as natural precipitation decreases the need for electric-powered irrigation. Accurately predicting this relationship can support more effective electricity planning during the monsoon season [5], [6], [12]."

Supporting Evidence:

- Empirical studies confirm that rainfall shocks significantly affect agricultural income and energy use, as farmers adjust irrigation practices in response to precipitation variability[5], [8].
- Field-based analyses in Andhra Pradesh and other states have shown that electricity consumption for irrigation is closely tied to crop patterns and rainfall, with higher rainfall years resulting in lower power demand for pumping[5].
- The WEF nexus literature highlights that India's agriculture is highly dependent on groundwater and energy, making it especially sensitive to rainfall variability and climate change[4].
- Official data from the CEA demonstrates significant variation in the number of agricultural electricity consumers across states, reflecting regional differences in irrigation dependency and rainfall sensitivity[9].

Recent media and government reports show that early or above-normal monsoon rainfall can lead to a measurable decline in national and regional electricity demand, especially in the agricultural sector[10], [11]

CHAPTER 2: LITERATURE REVIEW

2.1 Climate–Agriculture–Energy Nexus in India

The interdependence between climate, agriculture, and energy systems in India is increasingly recognized as a central challenge for sustainable development. India's agricultural sector is the largest consumer of both water and energy, with groundwater-based irrigation systems heavily reliant on electricity. This reliance has intensified as traditional surface water sources become less dependable due to erratic rainfall and climate change impacts[4], [13].

Recent research highlights that more than 80% of India's freshwater withdrawals are allocated to agriculture, and a significant portion of this is used for irrigation powered by subsidized electricity[4], [13]. The resulting feedback loop—where increased irrigation drives higher energy demand, and energy subsidies encourage further groundwater extraction—has led to widespread groundwater depletion, particularly in water-deficit, rice-based production systems in northwestern India. These trends are further exacerbated by changes in rainfall patterns, with climate change contributing to more frequent droughts, unpredictable monsoons, and declining groundwater recharge rates[4].

The water-energy-food (WEF) nexus approach has emerged as a valuable framework for understanding and addressing these interlinked challenges. By considering water, energy, and food sectors together, the nexus approach supports integrated solutions that enhance resource efficiency and sustainability. Conservation agriculture, for example, is gaining traction as a means to reduce water and energy use in crop production without compromising yields, offering a pathway to harmonize the goals of food security, resource conservation, and climate resilience[4], [13].

Policy interventions reflect this growing awareness. Initiatives such as the Pradhan Mantri Krishi Sinchayee Yojana (PMKSY) and solar irrigation programs aim to improve irrigation efficiency and promote renewable energy adoption in agriculture[14]. However, these efforts must be carefully managed to avoid unintended consequences, such as further incentivizing groundwater over-extraction in the absence of robust water governance frameworks[13], [14]. Economic and demographic drivers—such as population growth, dietary shifts, and the minimum support price (MSP) regime—continue to shape resource demand and sectoral interlinkages, underscoring the need for coordinated policy action.

In summary, the literature consistently demonstrates that India's climate, agriculture, and energy systems are deeply interconnected, with groundwater irrigation and energy use at the heart of this nexus. Addressing the sustainability challenges posed by these linkages requires integrated approaches, informed by robust data and cross-sectoral governance, to ensure long-term food, water, and energy security in the face of climatic and socio-economic change.

2.2 Rainfall Variability and Irrigation Electricity Demand

Rainfall variability is a defining feature of India's climate, with the southwest monsoon accounting for nearly 75% of the country's annual precipitation. The spatial and temporal distribution of monsoon rainfall is highly uneven, resulting in significant year-to-year fluctuations in water availability for agriculture. As a consequence, irrigation requirements—and by extension, electricity demand for groundwater pumping—are closely tied to the vagaries of the monsoon [4].

Empirical studies have shown that in regions with high dependence on groundwater irrigation, deficient rainfall years lead to a marked increase in electricity consumption as farmers compensate for the lack of surface water by extracting more groundwater. Conversely, years with above-average rainfall reduce the need for supplemental irrigation, resulting in lower electricity usage. This dynamic is particularly pronounced in states such as Punjab, Haryana, and Gujarat, where agricultural electricity demand is highly sensitive to monsoon performance.

The proliferation of subsidized electricity for agriculture has further intensified this relationship. While subsidies have improved access to irrigation and supported rural livelihoods, they have also contributed to the over-extraction of groundwater, especially during drought years. This feedback loop not only threatens the sustainability of water resources but also places additional strain on state electricity utilities, which must accommodate large and unpredictable spikes in demand during periods of rainfall deficit[14].

Recent policy initiatives, such as the promotion of solar-powered irrigation pumps, aim to decouple electricity demand from rainfall variability by providing a more sustainable and predictable energy source for farmers. However, experts caution that without complementary measures to regulate groundwater use, these interventions may inadvertently exacerbate resource depletion [15].

In summary, the literature underscores the critical role of rainfall variability in shaping irrigation practices and electricity demand in Indian agriculture. Addressing the challenges posed by this nexus requires integrated solutions that balance the needs of food production, water sustainability, and energy efficiency.

2.3 Sectoral Patterns of Electricity Demand

India's electricity demand is distributed across several key sectors, with agriculture, industry, residential, and commercial users representing the largest shares. The agricultural sector, in particular, has become a dominant consumer due to the widespread adoption of groundwater irrigation, which relies on electric pumps. Recent data indicate that agriculture accounts for more than 20% of total electricity consumption in India, with states like Punjab, Haryana, and Rajasthan exhibiting even higher shares due to their intensive irrigation practices [16] [17].

This pattern is closely linked to the country's groundwater crisis. As water tables fall, especially in over-exploited regions, farmers are compelled to use more powerful pumps and operate them for longer hours, leading to surges in electricity demand [17], [18]. The situation is further exacerbated by subsidized electricity tariffs, which, while supporting rural livelihoods, have inadvertently encouraged inefficient water and energy use[16], [19].

Industrial and urban sectors also contribute significantly to national electricity consumption, but their demand is more stable and less climate-dependent compared to agriculture. However, rapid urbanization and economic growth are expected to increase the electricity needs of these sectors in the coming decades [17].

In summary, the sectoral distribution of electricity demand in India is heavily influenced by agricultural practices and groundwater availability, making the sector uniquely sensitive to both policy and environmental changes.

2.4 Economic and Policy Drivers of Energy Use

The evolution of India's energy use, particularly in agriculture, is deeply intertwined with economic incentives and government policy frameworks. Subsidies for electricity and irrigation have historically been used to promote food security and rural development, but they have also contributed to unsustainable groundwater extraction and rising energy demand [17], [19]
Policy interventions such as the Pradhan Mantri Krishi Sinchayee Yojana and Atal Bhujal Yojana have sought to improve water use efficiency and regulate groundwater withdrawal, yet challenges persist due to fragmented governance and inconsistent enforcement [16], [20]. The legal and institutional landscape is evolving, with recent efforts to treat groundwater as a common pool resource and to introduce metering and rationing of electricity for agricultural use [17], [21].

Economic growth, population pressures, and climate variability continue to drive up demand for both water and energy. As groundwater becomes increasingly scarce, the cost of extraction rises, placing further strain on both farmers and state utilities [19]. The need for integrated policy approaches that balance agricultural productivity, energy efficiency, and water sustainability is now widely recognized in both research and practice[21]

2.5 Advances in Electricity Demand Forecasting

Recent advances in electricity demand forecasting for Indian agriculture have increasingly leveraged machine learning and data-driven models. Traditional forecasting methods often struggled to account for the high variability introduced by monsoon-dependent irrigation demand. In contrast, modern approaches integrate weather data, crop cycles, and economic variables to improve prediction accuracy and grid reliability [4]

Random Forest algorithms excel at capturing complex, non-linear interactions among variables such as rainfall, cropping patterns, and historical electricity use [22], [23], [24].Studies demonstrate that Random Forest-based approaches can deliver highly accurate predictions, outperforming many conventional and even some advanced machine learning models in both agricultural yield and energy consumption contexts[23], [24], [25]

The Random Forest Regressor's ability to handle missing data, categorical variables, and large feature sets makes it especially suitable for real-world energy and agricultural datasets, where data quality and completeness can vary[26]. As a result, this method is increasingly adopted by researchers and practitioners for demand forecasting, resource allocation, and policy planning in India's agricultural sector.

PHASE 1: STATE-LEVEL CORRELATION ANALYSIS

CHAPTER 3: METHODOLOGY

3.1 Data Collection

3.1.1 Rainfall data

For this study, yearly gridded rainfall data were obtained from the Indian Meteorological Department (IMD) Pune website, covering the period from 2014 to 2024. The data are provided in NetCDF format at a high spatial resolution of 0.25° × 0.25°, offering daily rainfall values for each grid point across the Indian subcontinent [27]. This granular dataset enables precise spatial and temporal analysis of rainfall patterns relevant to agricultural and hydrological studies [7], [28].

To process and extract relevant information from these large NetCDF files, Python libraries such as **xarray** and **pandas** were employed. These tools facilitated efficient data handling, aggregation to annual values, and spatial averaging over specific regions of interest [29]. The high resolution and quality control standards of the IMD gridded dataset make it a reliable source for climate and water resource research in India.[30]

day	lat	lon	rf	year	month	month_name
01-12-2024 00:00	8.25	77	0.038272	2024	12	Dec
01-12-2024 00:00	8.5	76.75	0.597107	2024	12	Dec
01-12-2024 00:00	8.5	77	0.324981	2024	12	Dec
01-12-2024 00:00	8.75	76.5	1.227204	2024	12	Dec
01-12-2024 00:00	8.75	76.75	15.33703	2024	12	Dec
01-12-2024 00:00	8.75	77	8.011066	2024	12	Dec

Table 1Extracted NetCDF file

3.1.2 Electricity Data

State-wise monthly electricity requirement data for the years 2014–2024 were obtained from the Central Electricity Authority (CEA) official executive summaries. These reports, published monthly, provide detailed tables of electricity demand across all Indian states. [30]

To efficiently compile the required data, a semi-automated workflow was developed using Python. Scripts were employed to systematically download the relevant PDF or Excel files for each month and year. Python libraries such as **pdfplumber, requests, tabula-py** and **pandas** were used to extract the necessary tables from these files and convert them into structured Excel spreadsheets. The final compiled dataset was organized with multiple sheets—one for each year—where each sheet contained a matrix with states as rows and months as columns, facilitating straightforward temporal and spatial analysis of electricity demand patterns.

This approach ensured a high degree of accuracy and consistency in data extraction, while significantly reducing manual effort and the risk of transcription errors.

Region	jan	feb	mar	apr	may	june	july	aug	sep	oct	nov	dec	TOTAL
Andaman-	29	29	29	29	29	29	29	29	29	29	29	29	348
Andhra Pr	5364	5,699	5,791	4788	5606	5075	4751	5054.00	4997	4963	4842	4,820	61,750
Arunachal	69	62	60	36	44	56	53	60	63	68	65	70	706
Assam	684	661	690	616	741	911	996	1101	1034	1020	757	745	9,956
Bihar	2,450	2,153	2,163	2317	2753	2999	3295	3440	3313	3285	2217	2,461	32,846
Chandigar	138	109	90	76	115	161	185	172	170	113	91	112	1,532

Table 2 Extracted Electricity Data

3.1.3 Agriculture share in electricity

Since electricity data were initially taken as aggregate figures encompassing all sectors, it was necessary to isolate the agricultural share to accurately identify states with a significant agricultural electricity load. To achieve this, sector-wise electricity sales data were extracted from the **Central Electricity Authority (CEA) Annual Reports**, which provide detailed state-level breakdowns of electricity consumption by sectors including agriculture, Traction and others. This allowed us to distinguish agriculture-heavy states based on the proportion of electricity consumed specifically for agricultural purposes.

Formula used was

$$\frac{Agri\ Sales}{Total\ Sales} \times 100$$

3.1.4 Economic Relevance of Agricultural Electricity Use

To assess the economic relevance of agricultural electricity use, we examined the correlation between actual energy consumption in the agriculture sector and state-level GDP .The data was collected from WEC India This approach allowed us to identify agriculture-heavy states more precisely and to evaluate the alignment between electricity demand patterns and the sector's economic significance, which is crucial for understanding the broader impact of agricultural energy consumption on state economies.

3.2 Data Cleaning and Preprocessing

3.2.1 Rainfall data

Spatial Mapping to States:

The original rainfall dataset, obtained as daily gridded data at specific latitude and longitude points, required aggregation to the state level for meaningful analysis. To achieve this, the GADM level-1 shapefile (gadm41_IND_1.shp) was used to define state boundaries[31], [32]. The shapefile was reprojected to the WGS 84 coordinate system (EPSG:4326) to ensure alignment with the coordinate system of the IMD NetCDF rainfall data [33]. This spatial overlay and mapping were performed using Python libraries such as rioxarray, geopandas, and shapely.geometry, enabling the assignment of each rainfall grid point to its corresponding state polygon.

Temporal Aggregation:

To maintain temporal fairness and facilitate state-wise analysis, the following steps were taken:

- **Daily Area Average:** For each state, the average rainfall over its entire area was computed for each day.
- **Monthly Mean:** The mean of all daily values was calculated for each month, yielding a monthly rainfall average per state.
- **Yearly Aggregation:** Monthly data were further aggregated to obtain yearly rainfall values for each state.

Handling Missing Values:

Missing values in the rainfall data were addressed by imputing the average rainfall for the corresponding month and state, computed across all available years in the dataset. This approach preserved the integrity of the temporal patterns while minimizing bias.

Calculation of Normals and Summary Statistics:

For each state, the rainfall "normal" was calculated as the average of the ten-year period (2014–2024). These normals served as baseline values for anomaly detection and further analysis. The final summary dataset consisted of yearly rainfall statistics for each state, ready for subsequent analysis.

Columr	Columr	Columr	Columr	Columr	Columr	Columr	Columr	Columr	Column1	Column1	Column1	Column13	Column14
State	2014	2015	2016	2017	2018	2019	2020	2021	2022	2023	2024		normal
ndhra Prades	23.447605	33.967369	25.383689	29.604035	22.21023	30.297066	41.794587	39.138355	34.8996063	28.575449	34.5901643		31.2643779
ınachal Prad	74.92852	85.602322	107.67415	106.581	74.47466	83.082756	93.290145	66.661799	95.3968393	69.7641061	70.3717814		84.3480068
Assam	60.193748	64.867138	69.624128	76.082098	59.671406	69.021619	74.375926	52.336568	78.1934323	62.2391256	67.906463		66.7737865
Bihar	36.373865	31.927199	39.915958	38.374931	28.906546	40.260255	51.703733	50.888098	31.6429364	32.0447208	32.1879981		37.6569311
Chandigarh	27.231516	26.961523	21.033321	30.200604	35.409847	30.607931	30.655779	27.726841	43.4627503	51.8859568	30.5480319		32.3385547
Chhattisgarh	42.165078	37.204635	42.699796	36.581063	39.642985	45.82169	49.681925	41.666167	45.1457735	41.4128935	46.1022015		42.556746
ı and Nagar I	78.934278	56.697844	97.600453	107.3816	73.076018	122.95652	80.895888	96.836584	118.166134	88.7261032	109.553418		93.7113493
Goa	126.91499	82.589956	84.951289	90.94455	120.51152	157.82408	149.57936	138.74544	126.837904	124.125466	179.920639		125.72229
Gujarat	20.76439	20.408613	21.116588	27.361189	16.344988	35.238473	36.199573	26.39992	30.8898842	29.6163083	37.3858749		27.4296183
Haryana	12.164108	16.127289	13.474645	15.217558	17.285844	13.1286	19.021218	23.614144	23.0102952	19.8860224	17.0422656		17.2701808

Table 3 Rainfall Summary Table

3.2.2 Electricity data

The state-wise monthly electricity demand data, sourced from the Central Electricity Authority (CEA), underwent a thorough **manual cross-verification** process to ensure accuracy and consistency. Any discrepancies or anomalies identified during this step were resolved by consulting original CEA reports and, where necessary, corroborating with secondary sources. Missing values were addressed by imputing the average demand for the corresponding month and state across available years, thereby preserving the temporal and spatial integrity of the dataset [30].

Following the data cleaning and validation process, the monthly electricity demand data for each state were aggregated to produce **annual electricity demand totals**. This aggregation enabled a clearer understanding of yearly consumption patterns and facilitated comparison across states and years.

Column1	2014	2015	2016	2017	2018	2019	2020	2021	2022	2023	2024
A&N	240	240	240	302	348	348	348	339	343	376	4
AP	72128	49939	53628	56700	63930	64476	61750	67522	71367	79423	7861
Arunachal	664	600	716	788	843	781	706	839	908	994	103
Assam	8360	8677	9082	9179	9604	9769	9956	10852	11321	12289	1277
Bihar	18121	22973	25908	26280	30361	31430	32846	36220	39278	41328	4416
Chandigarh	1605	1613	1654	1623	1569	1677	1532	1602	1777	1748	198
Chattisgarh	20722	24917	23960	25702	26017	29638	29281	31794	37321	39785	4231
Delhi	29050	28261	30888	29814	32183	33124	29275	30898	35022	34878	3864
DVC	18123	18465	12312	21199	21530	22136	20840	23502	25950	26597	2627

Table 4 Electricity Summary Table

3.2.3 Agricultural Electricity Share & SGDP Alignment

The agricultural share in electricity sales and the corresponding State Gross Domestic Product by Activity (SGDPA) data underwent a rigorous cleaning and preprocessing procedure to ensure accuracy and consistency before analysis. The sector-wise electricity demand data, sourced primarily from the Central Electricity Authority (CEA) Annual Reports, were first examined for missing or anomalous values. Missing data points were addressed through imputation techniques, such as replacing gaps with the average consumption for the respective state and sector across available years, preserving both temporal and spatial integrity. Outliers and inconsistencies were identified using statistical methods and cross-verified against original CEA reports and secondary sources to confirm data validity.

Similarly, the SGDPA data, obtained from official government publications, were checked for completeness and uniformity. Any discrepancies between reported values across different sources were reconciled by consulting authoritative datasets from the Ministry of Statistics and Programme Implementation (MoSPI). To facilitate meaningful comparison, the agricultural electricity consumption data and SGDPA figures were aligned temporally and normalized where necessary, ensuring compatibility for correlation analysis.

3.3 Data analysis

3.3.1 Rainfall data

Calculation of Excess and Deficit Rainfall

After aggregating the rainfall data to yearly values for each state, further processing was conducted to quantify rainfall anomalies and interannual variability.
For each state and year, the rainfall anomaly was computed using the following formula:

$$Rainfall\ Anomaly\ (\%) = \frac{Rainfall\ in\ Year - Normal\ Rainfall}{Normal\ Rainfall} \times 100$$

where "Normal Rainfall" is the 10-year average (2014–2023) for each state. This classification aligns with the Indian Meteorological Department’s standards, which define rainfall as:

- **Excess:** > +20% of normal
- **Normal:** ±20% of normal

- **Deficient:** -20% to -59% of normal
- **Scanty:** < -60% of normal

Year-on-Year Change:

To capture interannual variability, the year-on-year (YoY) change in rainfall was calculated for each state as:

$$YOY\ Change = Rainfall\ in\ 2024 - Rainfall\ in\ 2023$$

These calculations facilitated the identification of years with significant excess or deficit rainfall and enabled further analysis of their impact on electricity demand and agricultural outcomes. The processed data, with computed anomalies and YoY changes, formed the final summary dataset used for subsequent statistical and modelling analyses.

3.3.2 Electricity data

Calculation of Year-on-Year Growth

To analyse trends in electricity demand, year-on-year (YoY) growth rates were calculated for each state using the formula:

$$YOY\ Growth(2024) - \frac{Demand\ 2024 - Demand\ 2023}{Demand\ 2023}$$

This metric quantifies the relative change in electricity demand from one year to the next, providing a standardized measure of growth independent of the absolute size of the state or its energy consumption.

Calculation of Change in Growth Rate

To further capture the dynamics of electricity demand, the change in YoY growth rate (ΔYoY) was computed as:

$$\Delta\, YOY\ 2024 = YOY\ 2024 - YOY\ 2023$$

This value highlights whether the growth in electricity demand is accelerating or decelerating, which can be particularly informative when assessing the impact of external factors such as rainfall variability [30].

3.3.3 Importance of Analysing Relative Changes and Normalization

Analyzing **relative changes** and applying **normalization** are crucial steps in the processing of both rainfall and electricity demand data, especially when making cross-state or temporal comparisons in a country as diverse as India.

3.3.3.1 Rainfall Data

Absolute rainfall amounts differ greatly across Indian states due to variations in geography and climate. For example, states like Kerala or Assam typically receive much higher annual rainfall than Rajasthan or Gujarat. However, for effective drought or flood assessment, agricultural planning, and resource allocation, it is not the absolute rainfall that matters most, but how much the rainfall deviates from the long-term average (the "normal") for each state, or how it changes from one year to the next.

By focusing on **relative changes**—such as percentage deviations from normal rainfall or year-on-year anomalies—we can:

- Identify which states are experiencing drought or excess rainfall relative to their usual climate, not just in absolute terms.
- Enable fair comparisons between states, regardless of their typical rainfall volumes.
- Detect the real impact of climate variability and change on agricultural outcomes, groundwater recharge, and water resource management.

For instance, a 20% rainfall deficit in Rajasthan may have far more severe agricultural implications than a 100 mm deficit in a high-rainfall state. Similarly, consecutive years of negative rainfall anomalies can signal the onset of drought conditions, even if the absolute rainfall seems adequate by national standards [34]

3.3.3.2 Electricity Demand Data

A similar logic applies to electricity demand. The absolute electricity consumption varies dramatically between states due to differences in size, population, and economic activity. For example, Maharashtra reported a demand of 2,01,866 GWh in 2024, while Sikkim's demand was only 578 GWh . While these raw values indicate total consumption, they do not reveal how rapidly demand is changing or how external factors like rainfall are influencing demand.

Normalization through growth rates and other relative metrics is essential for:

- Capturing underlying trends in electricity consumption, independent of absolute state size or baseline demand.
- Allowing meaningful cross-state comparisons, so that both large and small states can be assessed on the same scale.
- Detecting the true impact of external influences such as rainfall, policy interventions, or economic developments. For example, even if a state's absolute demand increases, a decline in its growth rate may indicate that improved rainfall has reduced the incremental need for electricity, especially for agricultural irrigation.

This approach not only supports fairer and more insightful comparisons, but also enhances our ability to detect and interpret the effects of key drivers on electricity demand and water usage.

3.3.3.3 Integrated Perspective

By systematically applying normalization and focusing on relative changes for both rainfall and electricity data:

- We obtain a clearer picture of trends and anomalies, regardless of the underlying scale of each state.
- We can better attribute observed changes to specific causes—such as climate variability, technological adoption, or policy shifts—rather than to structural differences between states.
- This methodology aligns with best practices recommended by national and international agencies for robust, evidence-based decision making in the water-energy nexus [3], [35].

3.3.4 Correlation Analysis

Correlation analysis is a **statistical technique used to measure and evaluate the strength and direction of the relationship between two variables**[36].In this study, I conducted a correlation analysis using Microsoft Excel to explore the relationship between the selected variables at both the all-India and state levels for the years 2016 to 2024. The result is summarized by the correlation coefficient, which ranges from -1 (perfect negative correlation) to +1 (perfect positive correlation), with 0 indicating no linear relationship.

For this analysis, I used Excel's built-in statistical functions to calculate the correlation coefficients for each year and for each state. This allowed me to efficiently process the data and systematically compare trends over time and across regions.

It is important to note that while correlation analysis reveals whether and how strongly variables are related, it does not imply causation. The findings from this analysis provide valuable insights into patterns and associations in the data, which are further discussed in the results and interpretation section.

To deepen our understanding of the relationship between rainfall, electricity demand, and the economic significance of agriculture across Indian states, we conducted a comparative analysis integrating three key variables: the state-wise correlation between rainfall and electricity demand, the agriculture sector's share in total electricity sales, and the agriculture sector's contribution to state GDP (SGDPA).

States were systematically grouped into 6 groups based on their agriculture share in electricity sales and the strength of the rainfall–electricity correlation. This classification enabled us to distinguish states where the expected relationship—high agriculture share coupled with a strong negative correlation—was observed, from those where the patterns deviated. Additionally, by overlaying economic data (SGDPA), we assessed whether states with high agricultural electricity use also demonstrated a corresponding economic reliance on agriculture.

G1	High agri share (>=20%) & strong negative corr (<= -0.4)
G2	Low agri share (<10%) & positive corr (>= 0.1)
G3	Moderate agri share (10–20%) & weak corr (-0.4 to 0.1)
G4	High agri share (>=20%) & weak/positive corr (> -0.4)
G5	Low/moderate agri share (<20%) & strong negative corr (<= -0.4)
G6	Very low agri share (<5%) & weak corr

Figure 1. Criteria for classifying states into groups

Visualizations, including scatter plots and bar charts, were used to map these relationships and highlight both typical cases and outliers. This multi-dimensional approach provided a nuanced view of how sectoral composition and economic structure influence the sensitivity of electricity demand to climatic variability.

CHAPTER 4: RESULT AND INTERPRETATION

4.1 National Level Correlation

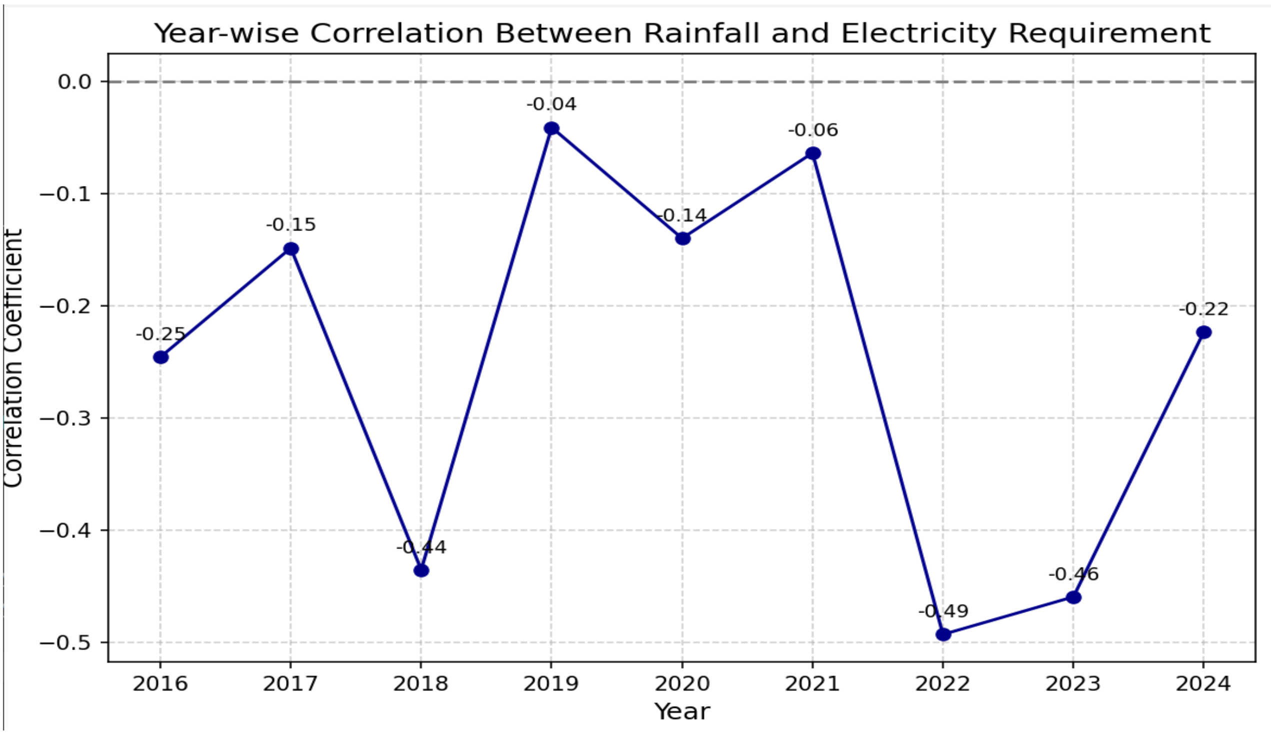


Figure 2. Line chart showing all India corelation between rainfall and electricity demand across years

STATES	POVERGF	RAINFALI	POVER(	RAINFA	POVERG	RAINFAI	POVERGI	RAINFAI	POVER(	RAINFAI	POVERGR(	RAINFAI	POVERG	RAINFAI	POVERG	RAINFALL	POVERGR	RAINFALL(	OVTH2024
AP	38.15%	-27.75%	-1.66%	13.64%	7.02%	-23.90%	-11.90%	26.14%	-5.08%	37.17%	13.58%	-8.59%	-3.65%	-13.70%	5.59%	-20.45%	-12.31%	19.45%	
Arunachal	28.97%	25.74%	-9.28%	-1.27%	-3.08%	-37.44%	-14.33%	10.04%	-2.25%	11.90%	28.44%	-31.06%	-10.61%	33.51%	1.25%	-29.89%	-5.85%	0.71%	
Assam	0.88%	7.14%	-3.60%	9.69%	3.56%	-24.62%	-2.91%	14.03%	0.20%	8.03%	7.09%	-33.06%	-4.68%	38.79%	4.23%	-23.93%	-4.56%	8.50%	
Bihar	-14.00%	20.91%	-11.34%	-4.03%	14.09%	-24.78%	-12.01%	29.72%	0.98%	29.95%	5.77%	-2.13%	-1.83%	-50.37%	-3.22%	1.05%	1.64%	0.38%	
Chandigarh	2.04%	-18.23%	-4.42%	28.19%	-1.45%	16.02%	10.21%	-14.77%	-15.53%	0.15%	13.22%	-9.01%	6.35%	48.39%	-12.56%	25.90%	15.19%	-65.62%	
Chattisgarh	-24.08%	13.02%	11.11%	-14.50%	-6.04%	7.26%	12.69%	14.64%	-15.12%	9.15%	9.79%	-18.99%	8.80%	8.25%	-10.78%	-8.85%	-0.24%	11.11%	
Delhi	12.01%	-23.35%	-12.77%	0.71%	11.42%	15.80%	-5.02%	-22.88%	-14.54%	20.67%	17.16%	53.99%	7.80%	-34.92%	-13.76%	8.37%	11.22%	-7.89%	
Goa	-15.93%	1.96%	-13.64%	4.98%	20.91%	24.58%	-7.69%	31.02%	-3.42%	-6.85%	14.59%	-9.01%	-6.40%	-9.90%	1.22%	-2.25%	5.24%	46.38%	
Gujarat	-5.37%	2.68%	2.42%	23.62%	3.02%	-41.67%	-10.02%	71.47%	-2.43%	3.64%	17.81%	-37.07%	-2.25%	16.99%	-2.07%	-4.82%	-6.87%	29.39%	
Haryana	2.51%	-15.34%	-2.10%	10.08%	4.61%	11.96%	-4.74%	-24.04%	-7.99%	34.08%	12.64%	26.56%	2.68%	-3.49%	-7.08%	-18.07%	10.16%	-16.44%	

	POWERGROW	LLGROWTH2016
POWER	1	
RAINFAL	-0.2455	1

	POWERGROW	LLGROWTH2017
POWER	1	
RAINFAL	-0.1483	1

	POWERGROW	LLGROWTH2018
POWER	1	
RAINFAL	-0.4351	1

Table 5 All India-Yearly Correlation

The correlation coefficients for All-India from 2016 to 2024 consistently show negative values. The average correlation coefficient over these nine years is **-0.2231**, indicating a **weak negative correlation** between the variables at the national level.

This weak association suggests that, overall, there is only a minor inverse relationship between the variables when aggregated for the entire country. However, this national average may be **diluted due to significant regional differences across Indian states**. India is characterized by pronounced regional disparities in economic development, literacy, industrialization, and infrastructure[38]. States vary widely in their economic structure, population growth rates, investment levels, and social indicators, all of which can influence the strength and direction of correlations at the local level.

For example, some regions may experience strong negative or positive correlations due to specific local factors, but when these are averaged at the national level, the distinct patterns get masked, resulting in a weaker overall correlation.

Therefore, to obtain a clearer and more actionable understanding of the relationships between variables, it is essential to conduct a deeper, state-wise or regional analysis. This approach allows for the identification of strong patterns and significant associations that are not visible in the aggregated national data. By analysing data at the state level, we can better understand the unique challenges and opportunities in different regions, enabling more targeted and effective policy interventions.

4.2 State-Wise Correlation

The correlation coefficients calculated for each state reveal a wide range of relationships, from strong negative to weak positive correlations. The top 10 states exhibiting the strongest negative correlations are as follows:

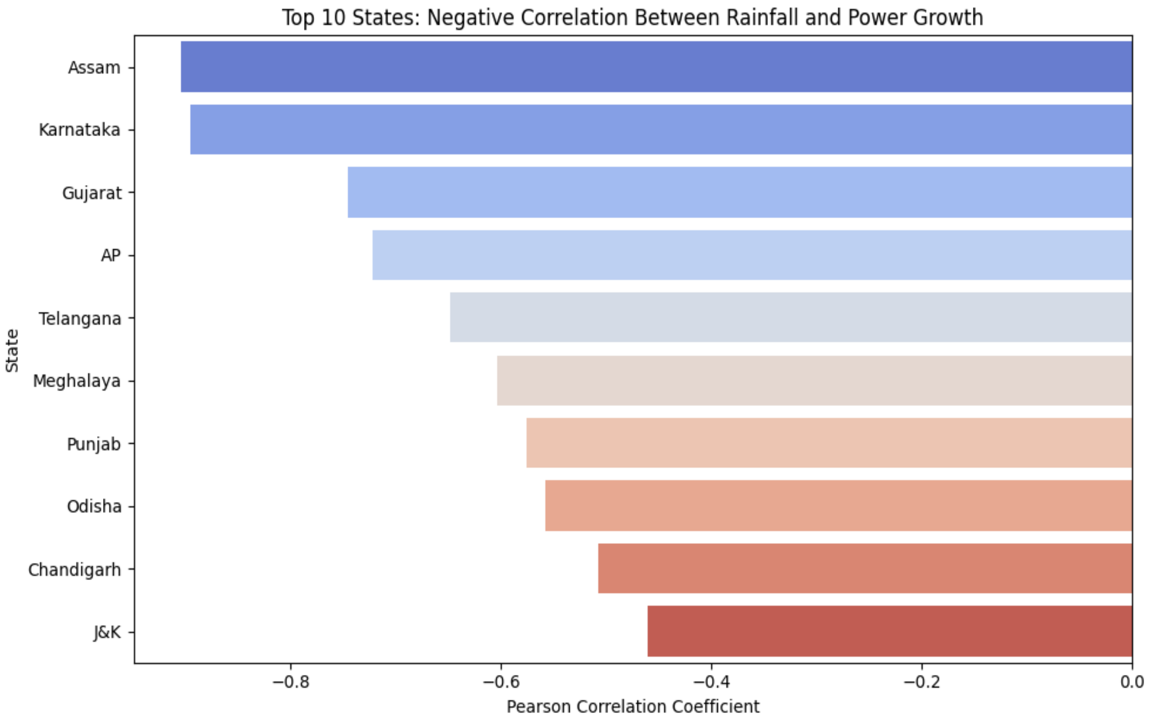


Figure 3. Top 10 states with strongest negative correlation

State	correlati
Assam	-0.903715
Karnatak.	-0.895111
Gujarat	-0.744949
Andhra P	-0.721333
Telangan	-0.648349
Meghalay	-0.603824
Punjab	-0.575173
Odisha	-0.55798
Chandiga	-0.507578
Jammu &	-0.460532
Maharash	-0.44438
Tamil Nac	-0.429551
Chhattisg	-0.339129
Mizoram	-0.307804

Table 6 State-wise Correlation

The average correlation coefficient for top 8 states is **-0.7063**, indicating a strong negative correlation.

The state-wise analysis highlights substantial variation in the strength and direction of correlations across different regions. Unlike the weak negative correlation observed at the national level, several states show **strong and significant negative correlations**

4.3 Agriculture Share and Economic Relevance

The comparative analysis of state-wise correlations between rainfall and electricity demand, agriculture’s share in electricity sales, and the correlation of agricultural electricity consumption with state GDP (SGDPA) reveals significant heterogeneity across Indian states.

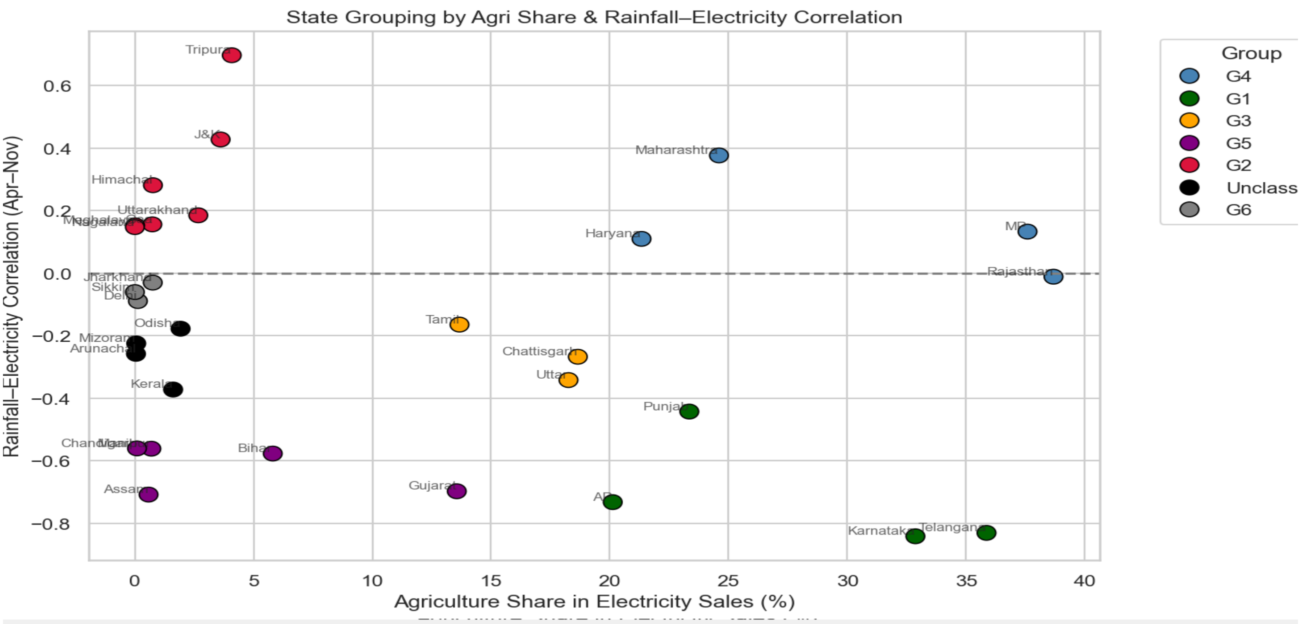


Figure 4. Scatter Plot with State Grouping by Agriculture Share and Rainfall–Electricity Correlation

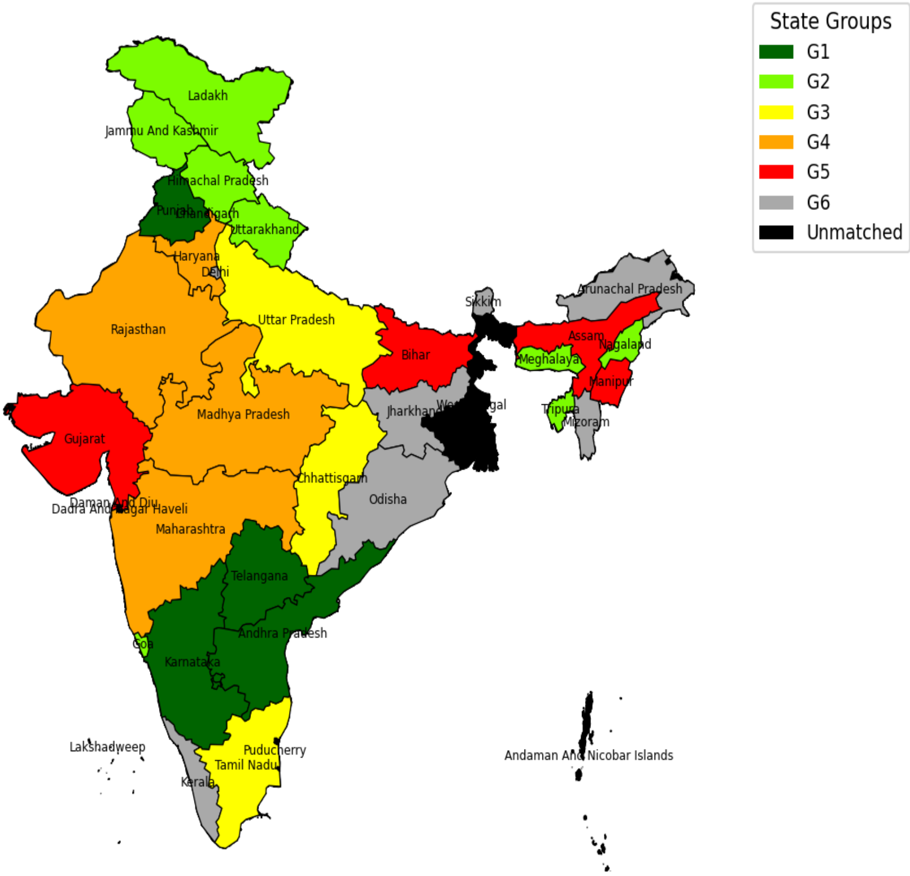


Figure 5. Map Illustrates the classification of Indian states into six groups

States that strongly support the central hypothesis—that increased rainfall reduces electricity demand in agriculture-heavy states—are found in Group G1. These include Karnataka, Andhra Pradesh, Telangana, and Punjab, all of which exhibit both a high agriculture shares in electricity sales (above 20%) and a strong negative correlation between rainfall and electricity demand (correlation ≤ -0.4). In these states, the correlation between agricultural electricity consumption and state GDP is also high, indicating that electricity use for irrigation is both rainfall-sensitive and economically significant. This aligns with established findings that electricity demand for groundwater pumping is closely linked to agricultural output in regions reliant on irrigation.

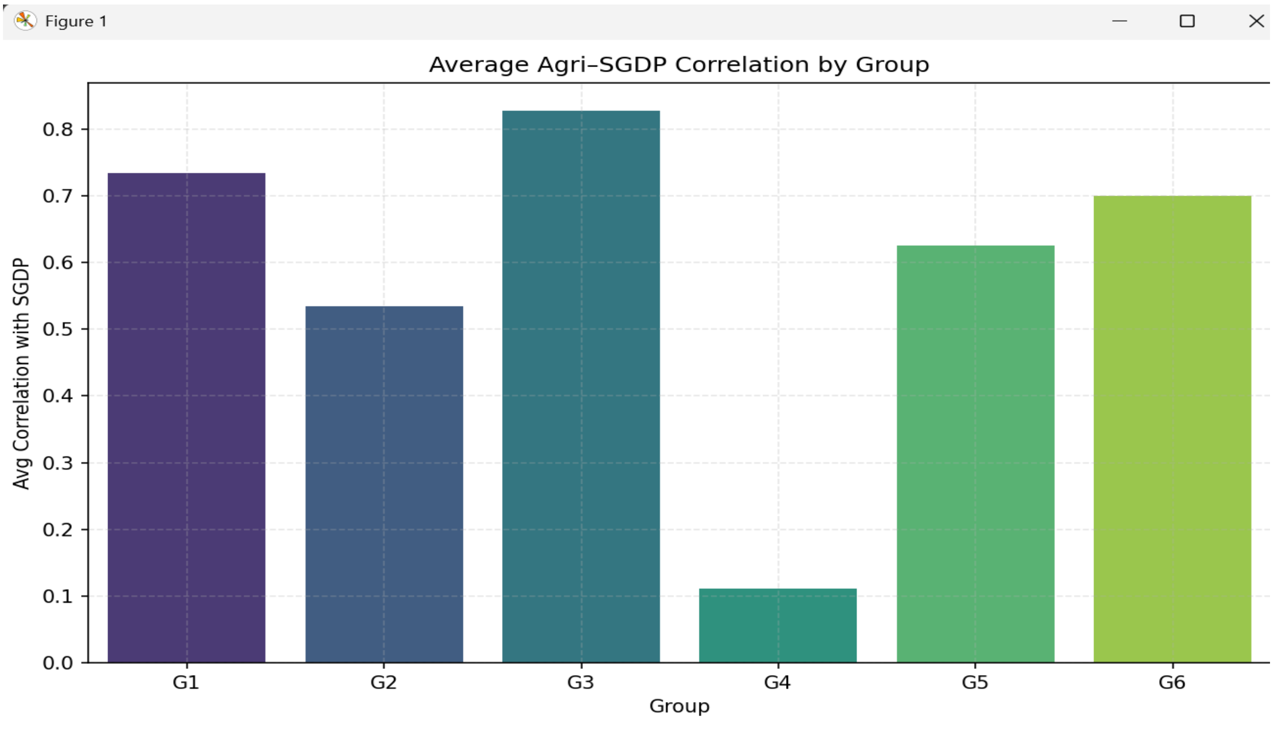


Figure 6. Average Correlation between Agriculture Electricity Share and Sector Corelation with SGDP by State Group

Group G5 states—such as Assam and Gujarat—present important exceptions. While Assam has a moderate agriculture share (0.6%) and a very strong negative correlation, Gujarat has a moderate agriculture share (13.6%) but also a strong negative correlation. These cases suggest that factors beyond just the agriculture share, such as the prevalence of groundwater-based irrigation or state-specific policies, may amplify the sensitivity of electricity demand to rainfall even in states not classified as agriculture-heavy.

Group G4 states—including Maharashtra, Madhya Pradesh, Rajasthan, and Haryana—are agriculture-heavy (agriculture share above 20%) but show weak or even positive correlations with rainfall. This divergence may be attributed to the use of canal irrigation, less rainfall-sensitive groundwater extraction, or significant policy interventions (such as subsidized or free electricity) that decouple electricity use from actual rainfall variability. In Haryana, for example, the correlation between agricultural electricity uses and state GDP is notably negative, suggesting structural or policy-driven disconnects.

Group G2 and G6 states—such as Meghalaya, Delhi, Sikkim, and Goa—have very low agriculture shares in electricity sales and as expected, show weak or positive correlations with rainfall. Their electricity demand is largely driven by non-agricultural sectors, so rainfall variability does not significantly affect overall consumption.

Group G3 states—like Tamil Nadu, Chhattisgarh, and Uttar Pradesh—have moderate agriculture shares (10–20%) and weak correlations. These states likely have a more diversified electricity demand profile, with agriculture playing a less dominant but still relevant role.

The correlation between agricultural electricity consumption and state GDP is generally strongest in G1 states, confirming that where agriculture is both a major electricity consumer and a key economic sector, rainfall-induced changes in electricity demand have clear economic implications. In contrast, states with weak or inconsistent correlations often have less economic dependence on agriculture or display patterns shaped by other dominant sectors.

In summary, the results demonstrate that while the central hypothesis holds for many agriculture-dominated states, there are notable exceptions driven by local practices, irrigation infrastructure, and policy environments. These findings reinforce the need for state-specific strategies in electricity demand management and agricultural policy, as well as the importance of considering sectoral and economic context in interpreting the impact of climatic factors on energy use.

PHASE 2: SECTOR-WISE ANALYSIS OF ELECTRICITY CONSUMPTION AND ECONOMIC ACTIVITY

CHAPTER 5: INTRODUCTION

Building on the insights and limitations identified in the initial aggregate and agriculture-focused analyses, the next phase of this study shifts to a detailed sector-wise examination of electricity consumption and its relationship with economic activity at the state level. While previous results highlighted important patterns and anomalies in the agriculture sector, they also revealed that aggregate data can obscure the unique drivers and sensitivities present in other major sectors—such as industry, residential, and commercial.

This phase aims to disaggregate electricity sales data by sector for each state and systematically analyse the correlation between sectoral electricity consumption and the corresponding sectoral contribution to state GDP (SGDPA). By doing so, the study seeks to uncover sector-specific demand patterns, assess the economic relevance of electricity use across different parts of the economy, and identify state-level variations in these relationships. This approach will help clarify whether the trends observed in agriculture also apply to other sectors, or if distinct sectoral dynamics are at play.

CHAPTER 6: METHODOLY

6.1 Data Collection

Sector-wise electricity sales data were obtained from the official WEC India website using their dashboard, covering the financial years (FY) 2016 to 2024. The data include detailed figures for each major sector—domestic, commercial, industrial, railway, electric vehicles (EV), and others—at the state level. To ensure temporal alignment, the original monthly rainfall data were also converted into financial year format. Additionally, for each sector, the share of electricity sales in total consumption was calculated, and the correlation between sectoral energy consumption and State Gross Domestic Product by Activity (SGDPA) was determined. This comprehensive dataset provides the necessary foundation for a robust sector-wise analysis of electricity demand and its relationship with both climatic and economic factors.

STATES	AGRI	COMM	DOMESTIC	INDUSTRIAL	OTHERS	PUBLIC SERVICI	RAILWAYS
Punjab	-0.84773165	0.6211773	-0.67673785	0.58556803	-0.1654708	0.003812206	0.60719507
WB	-0.80598048	-0.0277505	-0.02323952	-0.54992606	0.62794465	-0.345307857	0.62738581
Bihar	-0.54737256	0.1687085	-0.1366286	-0.067168	-0.0635511	0.682050249	-0.01293596
Telangana	-0.46008621	-0.9129609	-0.14611306	-0.6278239	0.55445812	0.643315278	0.20818838
Haryana	-0.40078684	-0.2523585	-0.75073787	0.56487338	0.11311772	0.089927125	0.65367779
Maharasht	-0.18494999	0.0278814	-0.15291893	-0.21555419	0.29206506	-0.661820472	0.47804823
Chhattisga	0.00995836	-0.0076405	-0.00162093	-0.13327387	0.84103081	0.358023465	-0.36199018

Table 7Sector-Wise Correlation Between Electricity Sales and Rainfall

6.2 Data Analysis

Consistent with the approach used in phase one, both the sector-wise electricity sales data and the rainfall data were converted into year-on-year (YoY) growth values using the same methodology. This transformation allowed for a standardized comparison of trends and variability across states and sectors. For each sector, the share in total electricity sales was calculated using the same formula as before.

Next, the correlation between sector-wise electricity sales and rainfall was computed for each sector and each state. This step enabled an assessment of how sensitive electricity demand in different sectors is to rainfall variability, building on the insight that rainfall can have varying impacts across sectors due to differences in operational drivers and economic linkages.

To deepen the analysis, the calculated correlations were compared with both the sector's share in total electricity sales and the correlation between sectoral electricity sales and sectoral SGDPA. This comparative approach provided a more nuanced understanding of whether sectors with higher electricity usage are also more responsive to rainfall, and how this sensitivity aligns with the sector's economic relevance in each state.

Line charts were generated to visually track how the correlation between rainfall and electricity demand varies across sectors and states. This visualization helped to identify sector-specific patterns, outliers, and the potential influence of economic structure and policy on electricity consumption trends.

This analytical framework allows for a comprehensive, sector-wise exploration of the interplay between climatic variability, electricity demand, and economic activity, setting the stage for more targeted policy insights and recommendations.

CHAPTER 7: RESULT AND INTERPRETATION

7.1 Introduction

While the Phase 1 analysis provided initial insights into the correlation between rainfall and overall electricity demand at the state level, it was limited in its ability to isolate the true drivers of this relationship. Electricity consumption at the state level is a composite of multiple sectors—agriculture, domestic, industrial, commercial, public services, and railways—each with varying sensitivity to rainfall. Consequently, the overall correlation observed in Phase 1 may be diluted, exaggerated, or even misinterpreted due to the influence of non-agricultural sectors.

To resolve this ambiguity, Phase 2 focuses on **sector-wise disaggregation of electricity consumption**, allowing for a more precise identification of how different sectors respond to rainfall variations.

A critical part of this phase is a **closer focus on states in Groups G4 and G5**, which were **not fully aligned with the central hypothesis in Phase 1**. These states either have a high agricultural share but showed weak or positive correlations with rainfall (G4), or exhibited strong negative correlations despite low agricultural shares (G5). Their seemingly contradictory behavior warrants deeper investigation, as it could be explained by factors like the dominance of other sectors, irrigation methods (e.g., canal vs. groundwater), or state-specific electricity subsidy regimes.

7.2 STATE WISE INTERPRETAION

7.2.1 Group G5 States

7.2.1.1 Gujarat

In Phase 1, Gujarat showed a strong negative correlation (-0.74) between rainfall and overall electricity demand, suggesting that increased monsoon rainfall generally coincides with reduced electricity consumption. However, sector-wise analysis (Phase 2) reveals that this overall trend is not uniformly distributed across sectors. Agriculture, which accounts for 20.1% of electricity sales, shows a **positive correlation (0.26)** with rainfall—contrary to the hypothesis—implying that rainfall does not reduce electricity use in this sector, possibly due to irrigation methods or policies that disconnect electricity demand from rainfall variability. Conversely, the **industrial sector**, which constitutes the largest share (43.5%), shows a **moderate negative correlation (-0.24)**, indicating that rainfall may reduce industrial activity, perhaps due to supply chain disruptions or power constraints during heavy rains. The domestic (16.9%), public service (2.6%), and others (0.1%) sectors all show **moderate positive correlations (0.35, 0.34, 0.28 respectively)**, suggesting limited rainfall sensitivity. Notably, the **railways sector** shows a strong **negative correlation (-0.46)**, which could reflect operational vulnerabilities during the monsoon. Overall, the strong Phase 1 negative correlation is likely driven by industrial and railway consumption dips, rather than agriculture, challenging the central hypothesis in Gujarat's case.

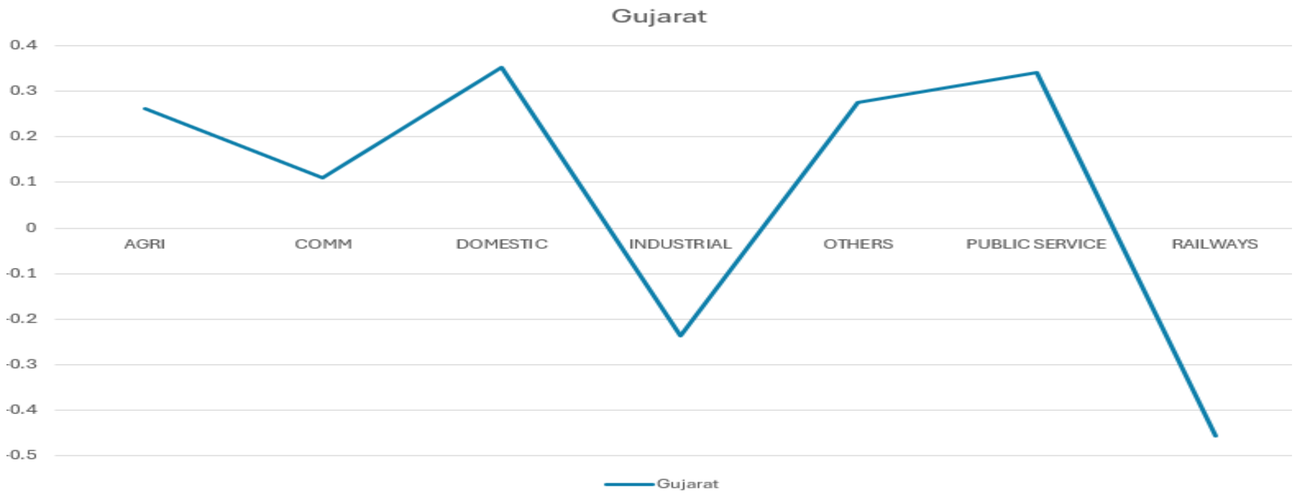


Figure 7 Correlation between Rainfall and Sector-Wise Electricity Sales in Gujarat

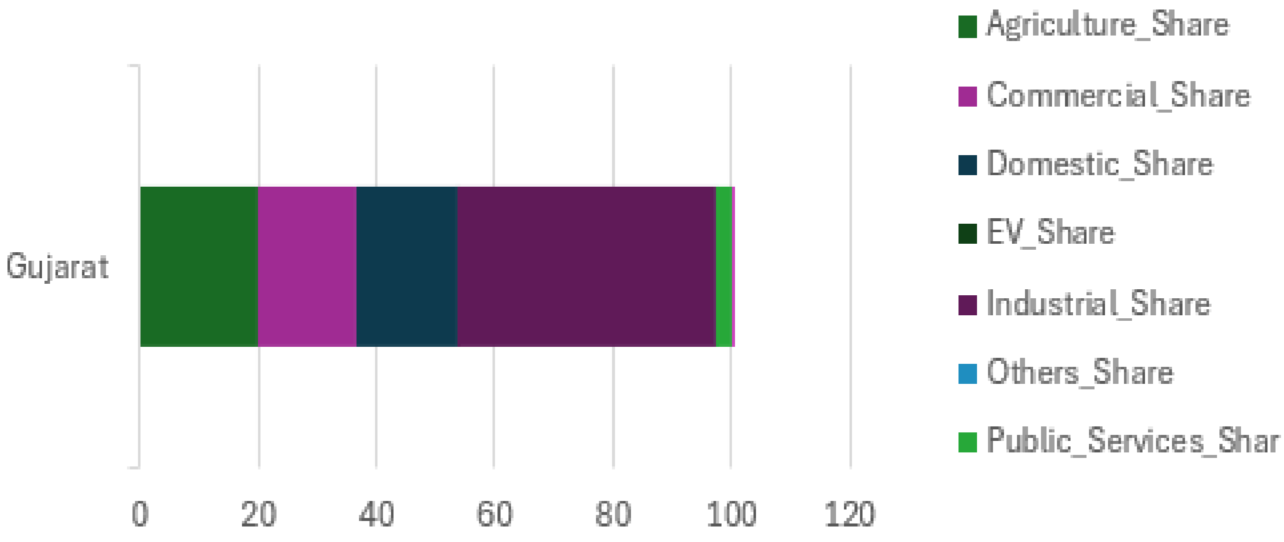


Figure 8. Sector-Wise Electricity Consumption Share in Gujarat

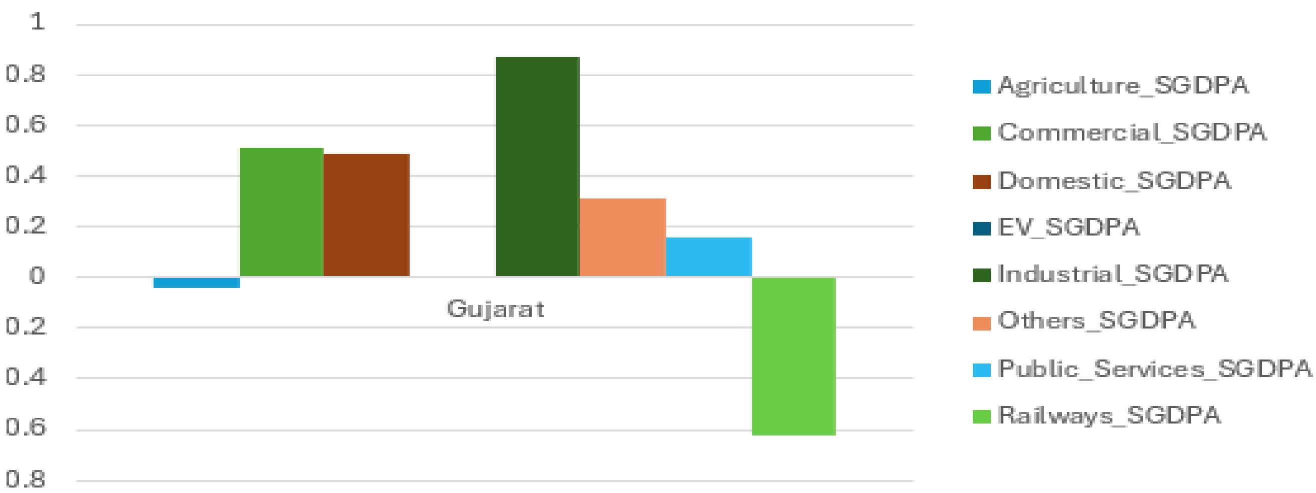


Figure 9 Sectoral Contribution to State Gross Domestic Product

7.2.1.2 Assam

Assam displayed the strongest negative correlation in Phase 1 (-0.90), supporting the hypothesis that rainfall sharply reduces electricity demand. However, Phase 2 reveals an important nuance: the agriculture sector's share in electricity usage is minimal (1.7%), and its rainfall correlation is **positive (0.67)**. This suggests agriculture is not a key driver here. Instead, **industrial (-0.96)** and **domestic (-0.30)** sectors show significant negative correlations with rainfall. With over **53.6% of total electricity going to domestic consumption**, even a moderate negative correlation in that sector contributes heavily to the overall pattern. Industrial usage (21.3%) also amplifies the effect due to its strong negative link. Thus, Assam's overall negative correlation is not explained by agriculture but rather by rainfall-sensitive industrial and domestic demand, perhaps affected by infrastructure limitations or seasonal slowdown

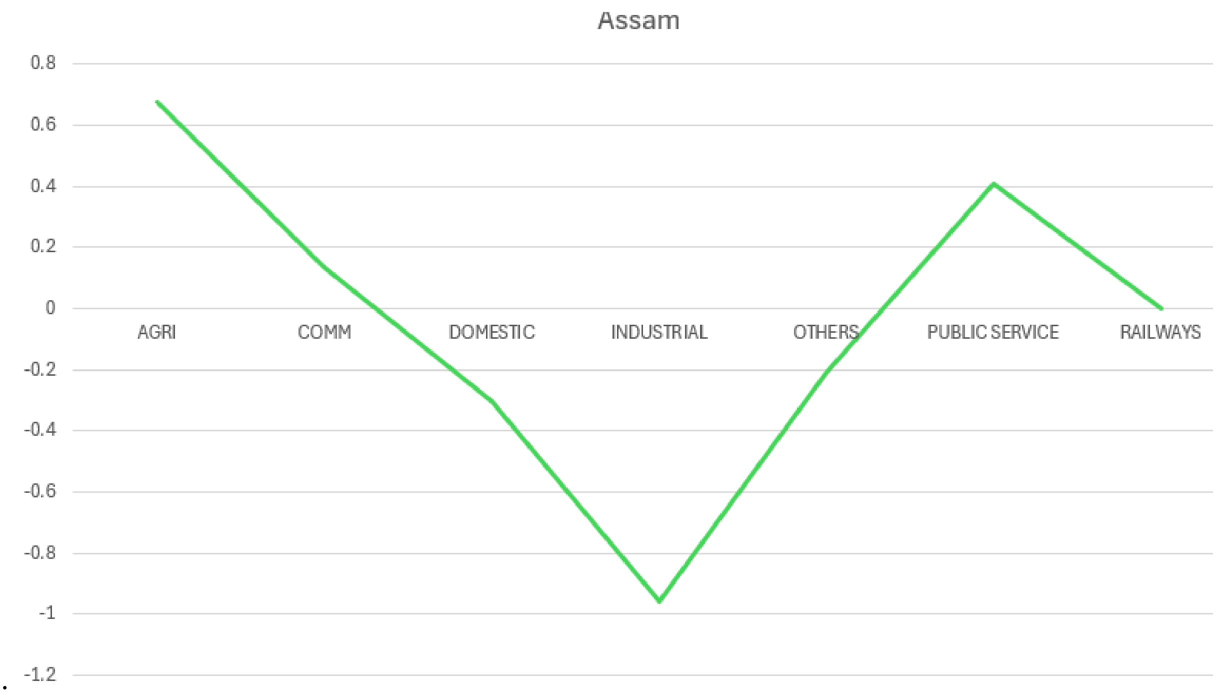


Figure 10. Correlation between Rainfall and Sector-Wise Electricity Sales in Assam

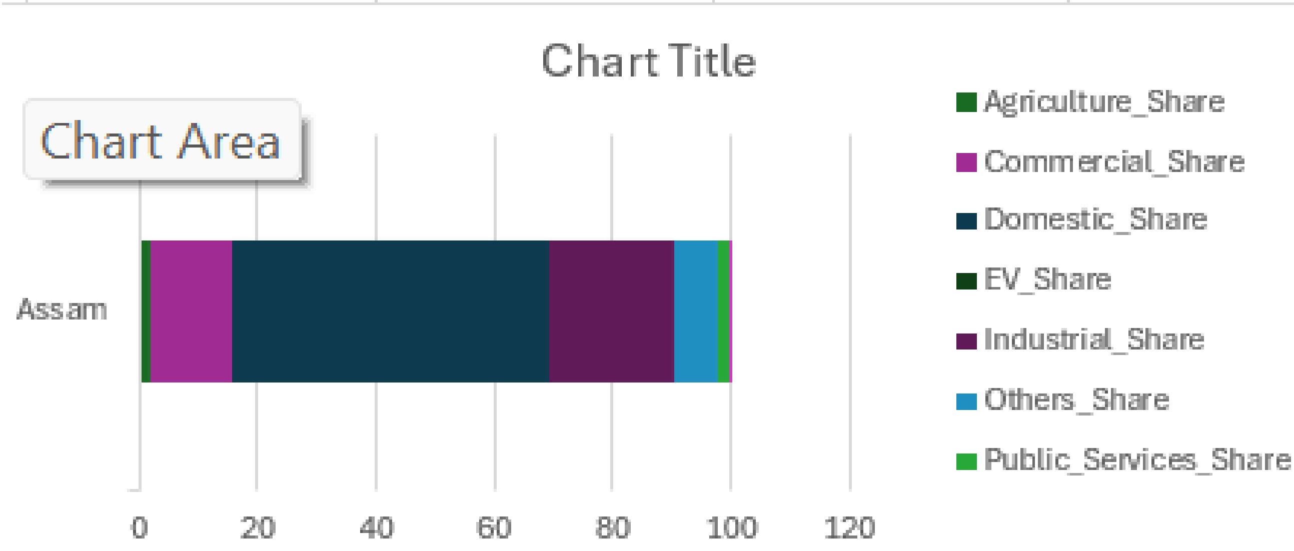


Figure 11. Sector-Wise Electricity Consumption Share in Assam

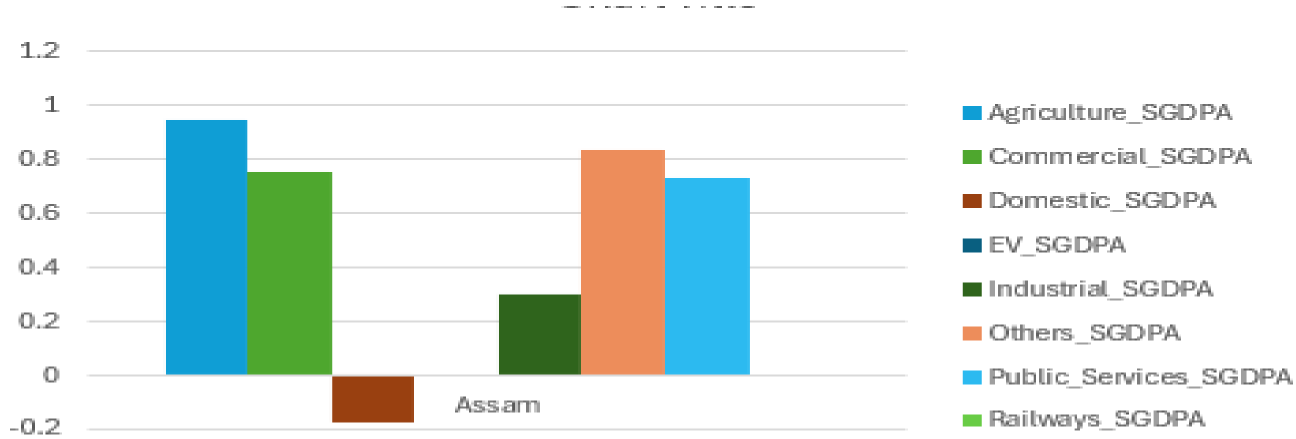


Figure 12. Sectoral Contribution to State Gross Domestic Product in Assam

7.2.2 Group G4 States

7.2.2.1Maharashtra

In Phase 1, Maharashtra had a weak negative correlation (-0.44) between rainfall and overall electricity consumption. Phase 2 reveals that agriculture, which contributes about **23.7%** of electricity usage, has a slightly **positive correlation (0.14)** with rainfall—suggesting that irrigation demand might not decrease with more rainfall, possibly due to free or subsidized electricity policies. Meanwhile, the **industrial sector** (34.9%)—a major consumer—has a **slight positive correlation (0.13)**, suggesting limited rainfall impact. **Domestic (25.3%)** and **public service (5.3%)** sectors show moderate **positive correlations (0.29 and 0.56 respectively)**, which could indicate increased electricity use during the monsoon, perhaps due to higher indoor activity. These opposing sectoral trends dilute any strong rainfall sensitivity, explaining the weak overall correlation observed in Phase 1.

7.2.2.2 Madhya Pradesh

Madhya Pradesh showed a near-zero overall correlation in Phase 1, despite agriculture holding a large share of electricity consumption (20.2%). In Phase 2, agriculture shows a **moderate positive correlation (0.18)** with rainfall, contrary to expectations. Similarly, **commercial (0.73)** and **domestic (0.34)** sectors also show **positive correlations**, with domestic use comprising over **48%** of the state's electricity sales. The **industrial sector (15.9%)** has a mild **positive correlation (0.13)**, offering no offset. The lack of any sector showing a strong negative correlation with rainfall explains the overall neutrality observed in Phase 1. This suggests that in Madhya Pradesh, rainfall does not systematically reduce electricity consumption, possibly due to a policy environment that buffers agricultural users or diversified demand sources.

7.2.2.3 Rajasthan

Rajasthan also displayed a very weak overall correlation (-0.06) in Phase 1, despite having the **highest agriculture share (41%)** among G4 states. However, Phase 2 shows a **slightly positive correlation (0.06)** between agriculture and rainfall, again contradicting the central hypothesis. Industrial (23.9%) and domestic (20.3%) sectors both show **weak negative correlations (-0.34 and -0.28 respectively)**. Public services (-0.86) and railways (-0.40) show stronger negative correlations, but their consumption shares are relatively small. The dominance of agriculture combined with a lack of strong negative correlations in any major sector likely explains why the Phase 1 result is weak, despite expectations.

7.2.2.4 Haryana

In Phase 1, Haryana showed a **slight positive correlation (0.09)** overall. Agriculture, which accounts for **23.9%** of consumption, shows a **moderate positive correlation (0.11)** in Phase 2, suggesting electricity use for irrigation is not rainfall-sensitive. Industrial (33.1%) and domestic (26.6%) sectors both show **positive correlations (0.56 and 0.75 respectively)**, reflecting increased usage during the monsoon, likely due to routine or residential demand. Railways and public services show weak negative correlations. Thus, Haryana's overall positive correlation is driven largely by non-agricultural sectors whose consumption patterns are either unaffected or enhanced by rainfall, again challenging the main hypothesis.

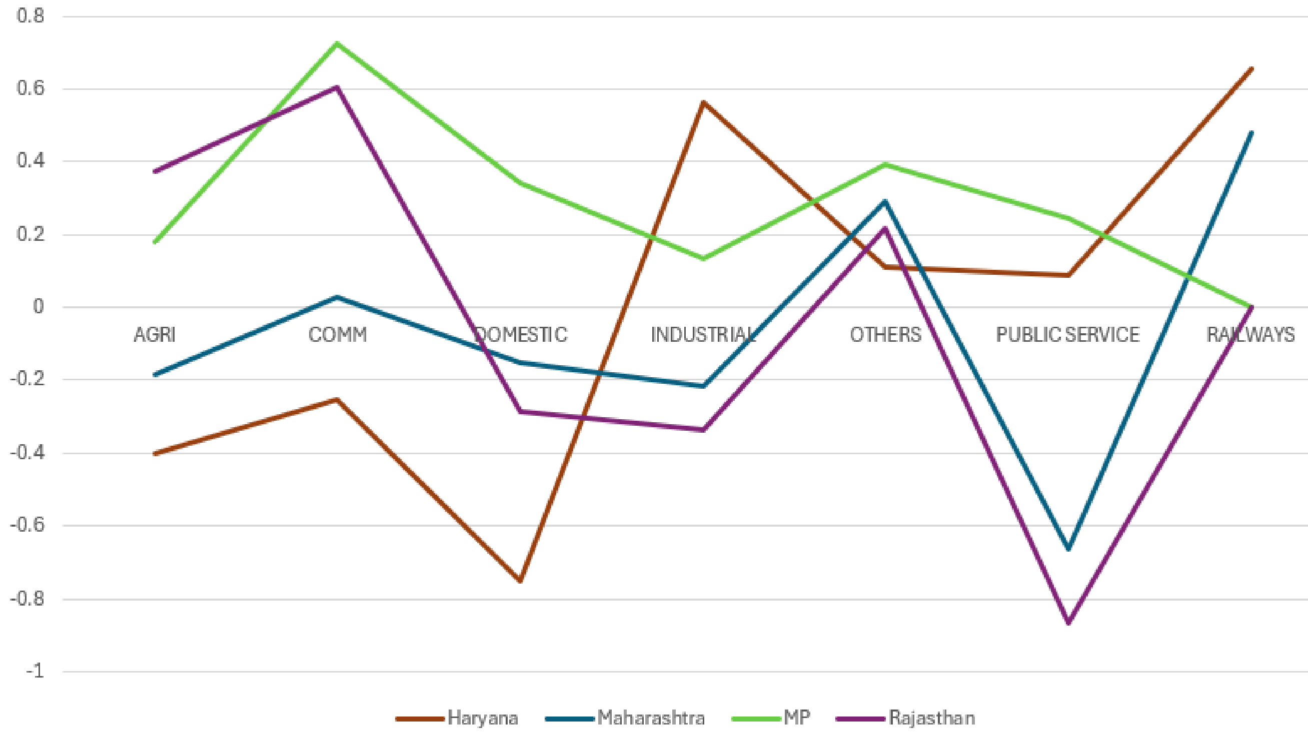


Figure 13. Correlation between Rainfall and Sector-Wise Electricity Sales in Haryana, Maharashtra, MP and Rajasthan

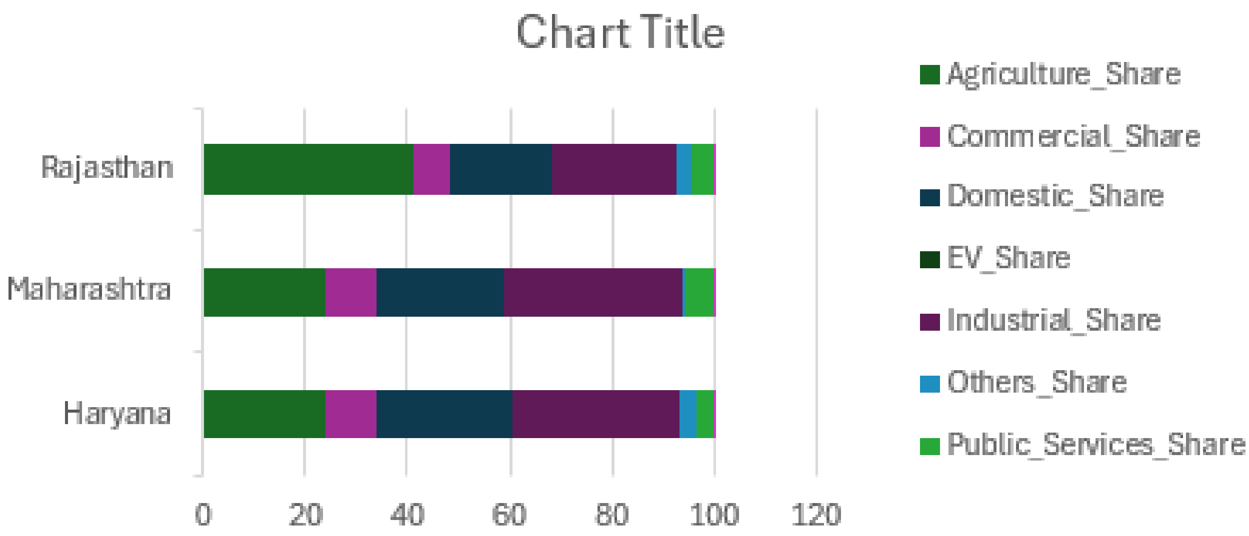


Figure 14. Sector-Wise Electricity Consumption Share in Rajasthan, Maharashtra and Haryana

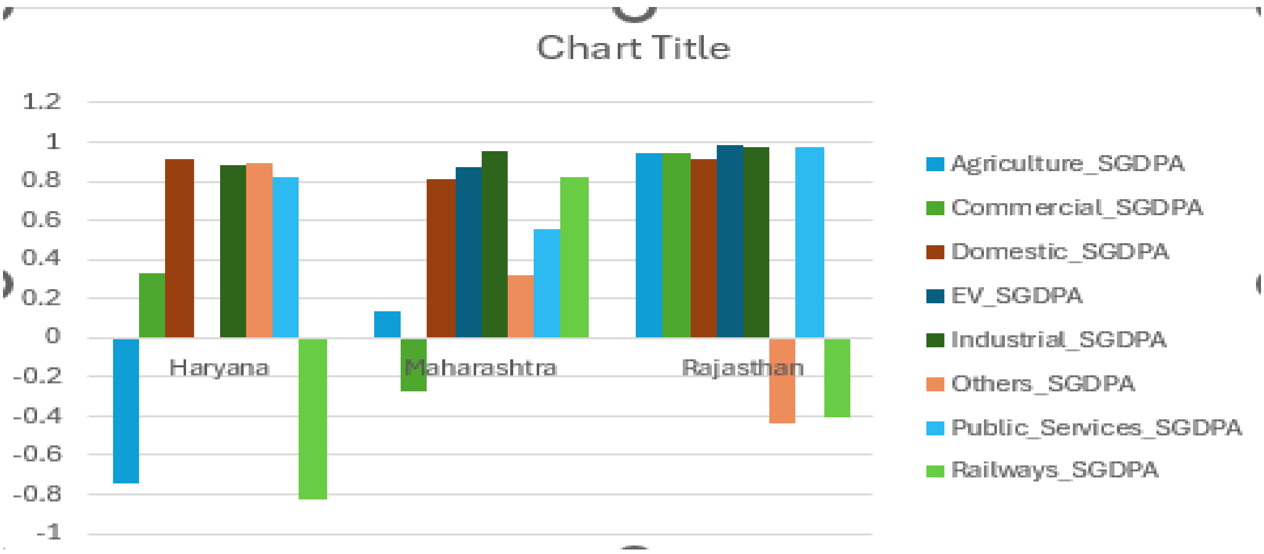


Figure 15. Sectoral Contribution to State Gross Domestic Product in Haryana, Maharashtra and Rajasthan

7.2.3 Group G1 States

7.2.3.1 Karnataka

the Phase 1 analysis revealed a strong negative correlation of –0.895 between rainfall and total electricity demand, with agriculture accounting for nearly 37% of electricity sales. The Phase 2 findings confirm this relationship, showing that the agriculture sector alone exhibits a significant negative correlation with rainfall. Interestingly, the domestic sector also reveals a strong negative correlation, likely influenced by seasonal energy use patterns or rural electricity access dynamics. On the other hand, commercial and industrial sectors show positive correlations, suggesting their electricity demand may be less rainfall-dependent or perhaps even counter-seasonal. Nevertheless, the dominant agricultural trend ensures that Karnataka remains a clear supporter of the hypothesis.

7.2.3.2 Andhra Pradesh

also demonstrated a strong negative Phase 1 correlation of –0.72, supported by an agriculture electricity share of around 26%. In the sectoral analysis, agriculture's negative correlation persists, though moderately, indicating that rainfall does affect irrigation-related consumption but not as sharply as in Karnataka. This could be due to differences in irrigation infrastructure, energy efficiency, or cropping patterns. Other sectors such as industry and domestic use appear to exert only mild influence on the total correlation, leaving agriculture as the primary explanatory factor.

7.2.3.3 Telangana

exhibited a Phase 1 correlation of –0.65 and an agriculture share of 23%. In Phase 2, agriculture once again shows a clear negative relationship with rainfall, albeit less intense than Karnataka or Punjab. Notably, commercial and industrial sectors in Telangana also exhibit strong negative correlations. This may reflect seasonal variations in energy-intensive agricultural processing or rural enterprise activities tied to the agricultural calendar. Despite the complexity, the results reaffirm that agriculture plays a key role in the overall rainfall-electricity relationship in Telangana.

7.2.3.4 Punjab

while slightly lower in overall correlation (–0.575), presents one of the clearest and strongest sectoral signals in Phase 2. The agricultural sector alone shows a very strong negative correlation of –0.85 with rainfall—one of the highest observed across all states. This finding is consistent with Punjab's heavy dependence on electric irrigation, especially for water-intensive crops like paddy. Other sectors, such as domestic and public services, also show negative or neutral trends, while the commercial and industrial sectors exhibit mildly positive correlations. Despite these minor opposing forces, Punjab's total correlation remains significantly negative, reinforcing the role of rainfall in shaping electricity demand in this agriculture-dominant state.

Figure 16. Correlation between Rainfall and Sector-Wise Electricity Sales in Punjab, Karnataka, Telangana and Andhra Pradesh

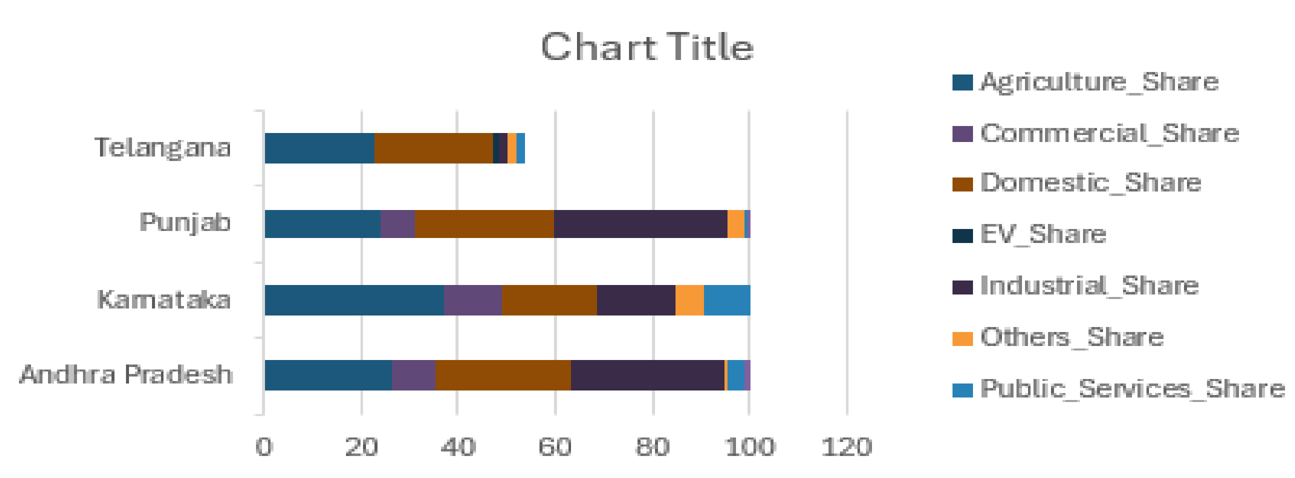


Figure 17. Sector-Wise Electricity Consumption Share in Telangana, Punjab, Karnataka and Andhra Pradesh

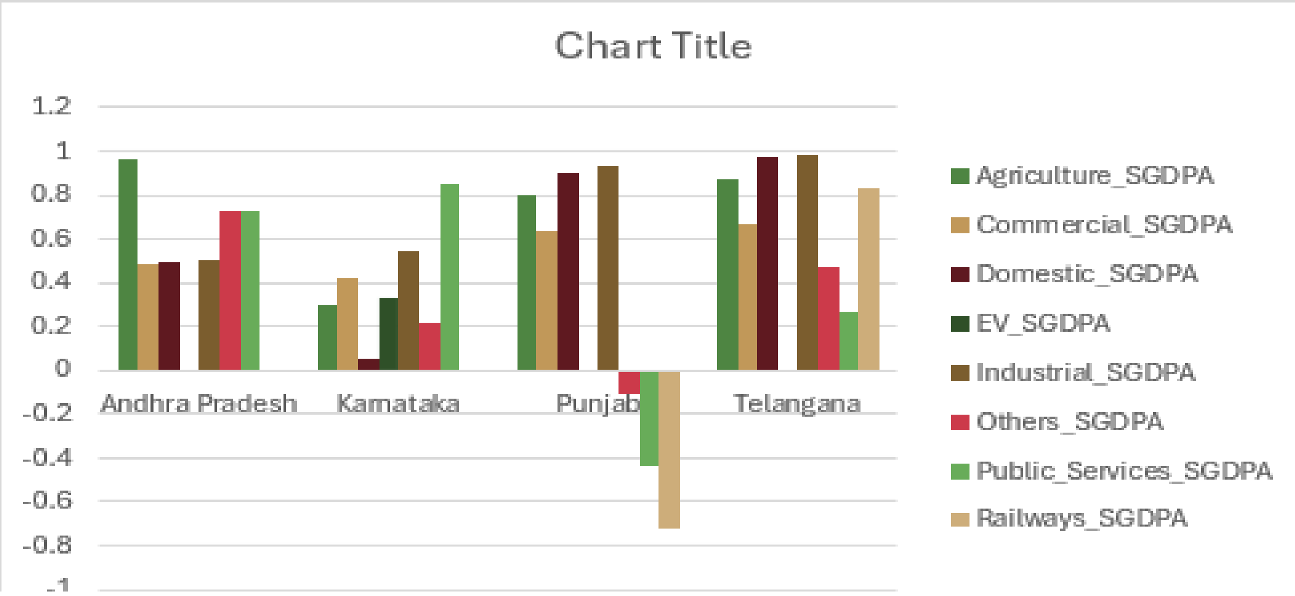


Figure 18. Sectoral Contribution to State Gross Domestic Product in Andhra Pradesh, Karnataka, Punjab and Telangana

7.2.4 Group G2 States

7.2.4.1 Delhi

In Phase 1, Delhi showed a near-zero correlation (0.05) between rainfall and overall electricity demand. This aligns with its very low agricultural share (~0.13%), making rainfall impact negligible on irrigation-based electricity use.

Phase 2 confirms that the bulk of Delhi's electricity is consumed by the domestic sector (56%), which shows a moderate negative correlation with rainfall (–0.34), and the commercial and public service sectors, both of which are positively correlated with rainfall. This suggests that in Delhi, electricity usage responds more to urban lifestyle or economic activity patterns than agricultural needs.The positive commercial correlation could indicate increased demand during rainy months due to indoor commercial activity, while the public service sector's correlation might reflect government infrastructure usage patterns.

7.2.4.2 Sikkim

Sikkim showed a slight positive overall correlation (0.08) in Phase 1, with zero agricultural electricity usage, reaffirming that rainfall has no direct impact on electricity demand here.
Sector-wise data reveals dominance by industrial (55%) and domestic (28%) sectors, both of which show negative correlations with rainfall, though modest. The result validates Phase 1: in a state with no agriculture-related demand, rainfall is not a primary influencing factor on energy consumption.

7.2.4.3 Goa

Goa recorded a moderate positive correlation (0.21) in Phase 1, despite having less than 1% agricultural share. Phase 2 shows the industrial sector (50%) dominates and is negatively correlated with rainfall (–0.64), while the domestic and others sectors show neutral to weak negative correlations.
The overall positive correlation likely arises from short-term increases in indoor energy use (residential and commercial) during the rainy season, overcoming the slight industrial dip. Again, rainfall is not driving demand due to agriculture, but affecting urban behavior and energy consumption patterns.

7.2.4.4 Himachal Pradesh

With a weak positive correlation (0.14) and negligible agriculture share (0.76%), Himachal Pradesh's demand patterns are largely climate-independent in terms of irrigation. Sector-wise, it is driven by industrial (58%) and domestic (24%) usage, both moderately negatively correlated with rainfall.
The correlation is likely affected by hydropower generation patterns or domestic heating needs, making it more climate-sensitive, but not agriculture-linked, affirming the central hypothesis.

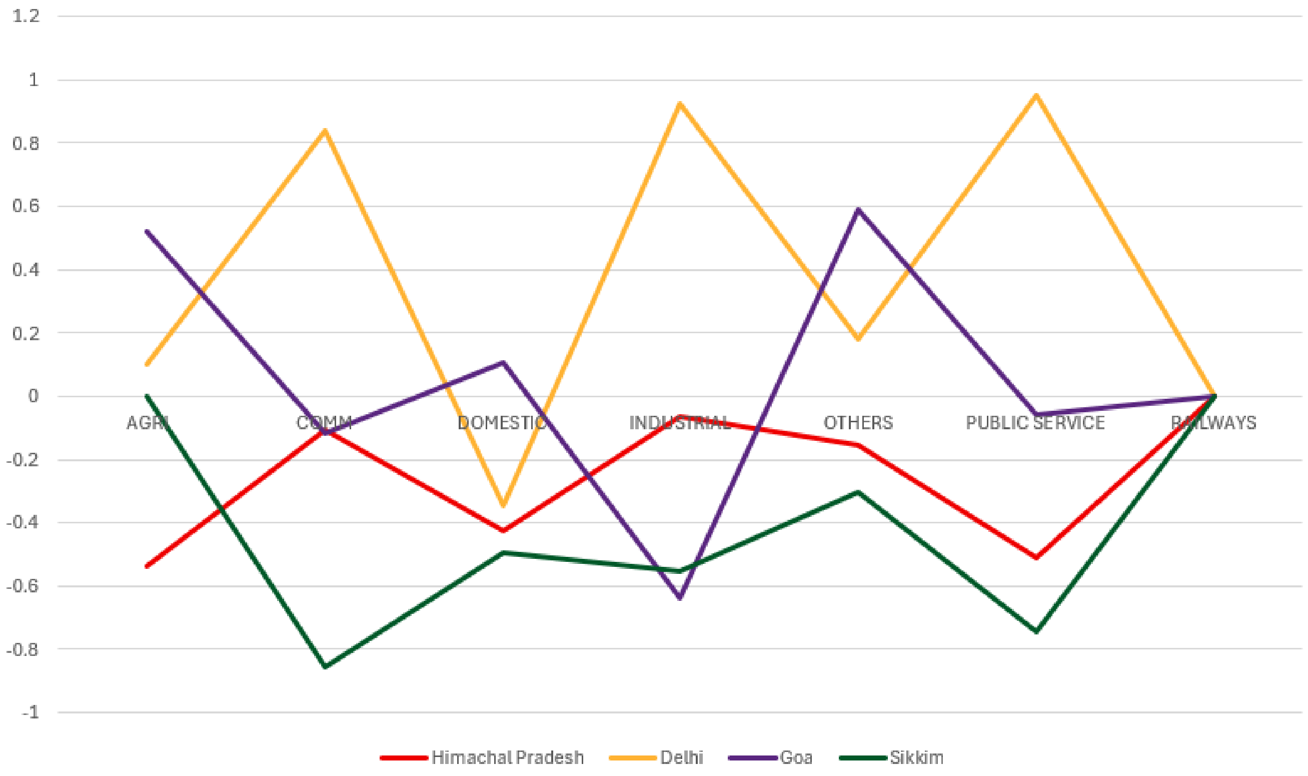


Figure 19. Correlation between Rainfall and Sector-Wise Electricity Sales in Himachal Pradesh, Delhi, Goa and Sikkim.

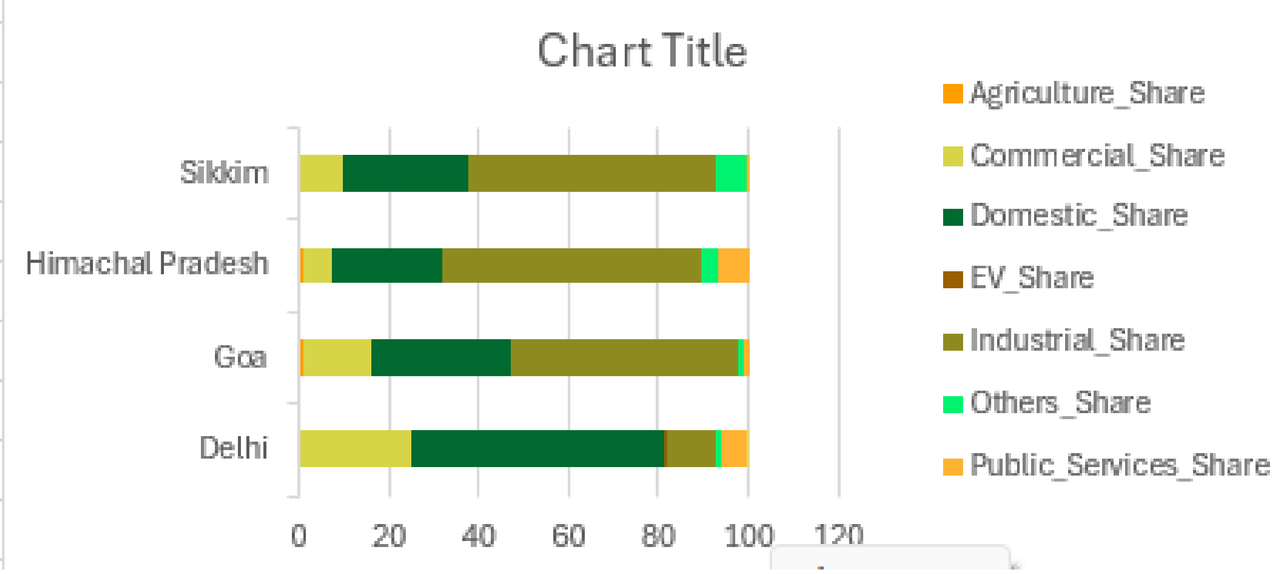


Figure 20. Sector-Wise Electricity Consumption Share in Himachal Pradesh, Delhi, Goa and Sikkim

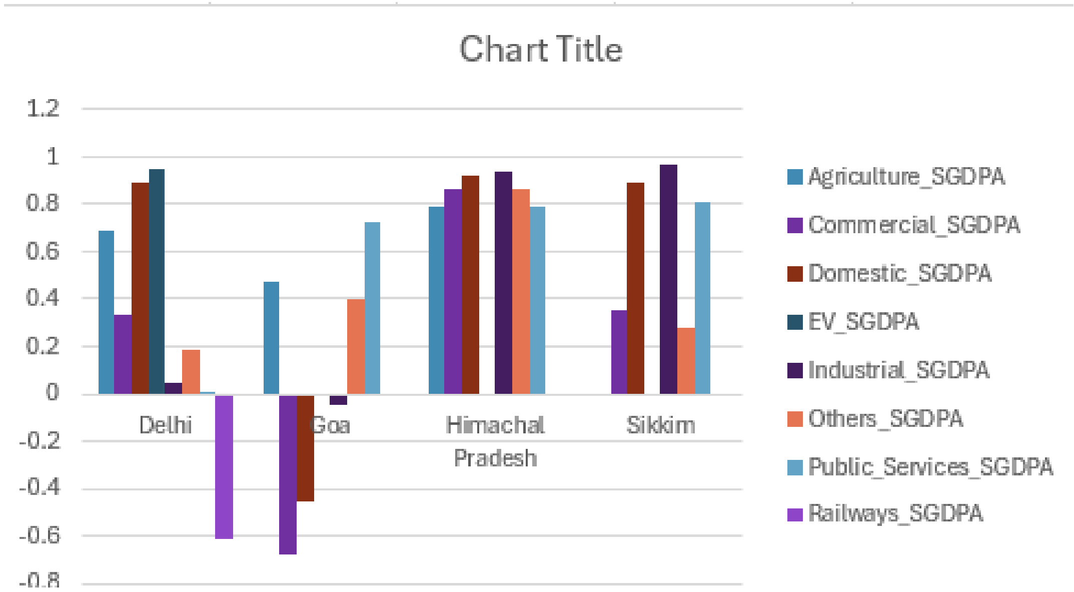


Figure 21. Sectoral Contribution to State Gross Domestic Product in Delhi, Goa, Himachal Pradesh and Sikkim

7.2.4.5 Nagaland, Meghalaya, and Tripura

These smaller northeastern states exhibit a range of weak correlations, with low agricultural shares across the board (all <4%).

- Nagaland shows an overall neutral correlation (0.04) and is dominated by domestic (54%) and commercial (14%) sectors. No strong pattern with rainfall is seen.
- Meghalaya had a negative correlation (–0.60) in Phase 1, which is somewhat surprising given agriculture's tiny share (0.02%). However, Phase 2 reveals high domestic (37%) and industrial (30%) shares, which may reflect seasonal fluctuations in rural electricity supply or infrastructure response to weather rather than irrigation needs.
- Tripura showed a small negative correlation overall but interestingly, both the commercial (35%) and domestic (3.8%) sectors show strong positive correlations. This suggests rainfall could be influencing electricity demand through residential/commercial patterns, not agriculture.

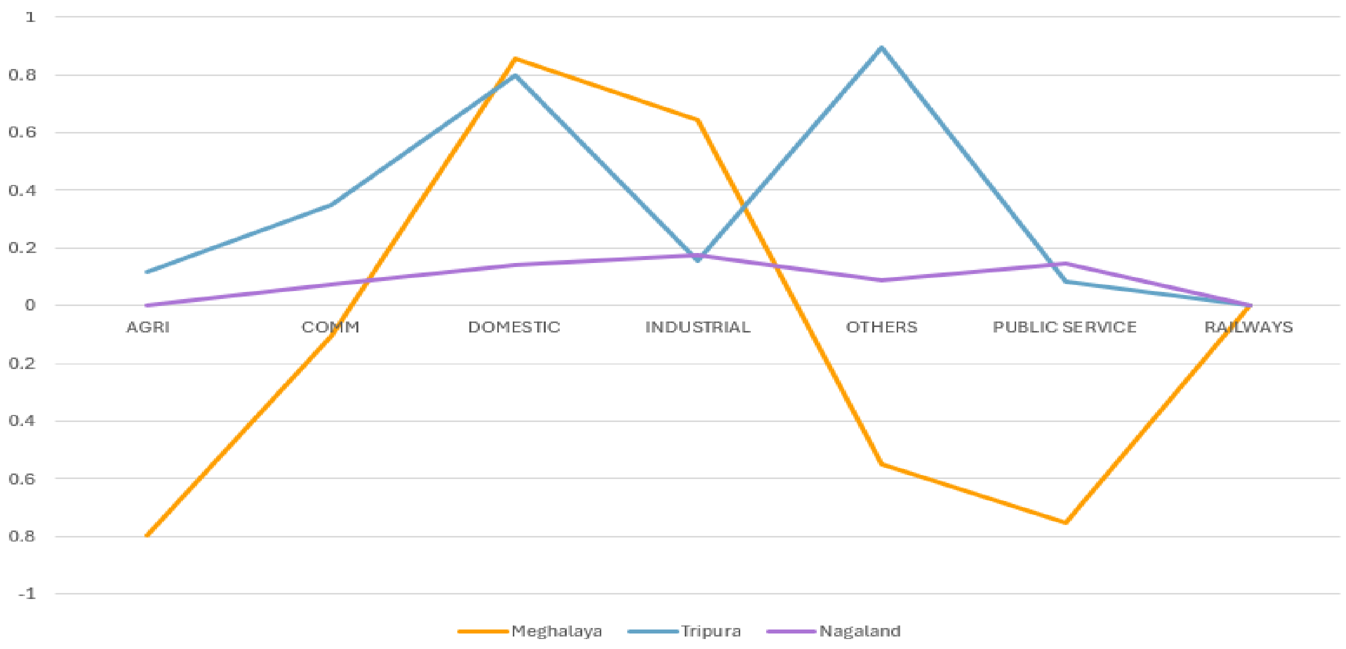


Figure 22. Correlation between Rainfall and Sector-Wise Electricity Sales in Meghalaya, Tripura and Nagaland

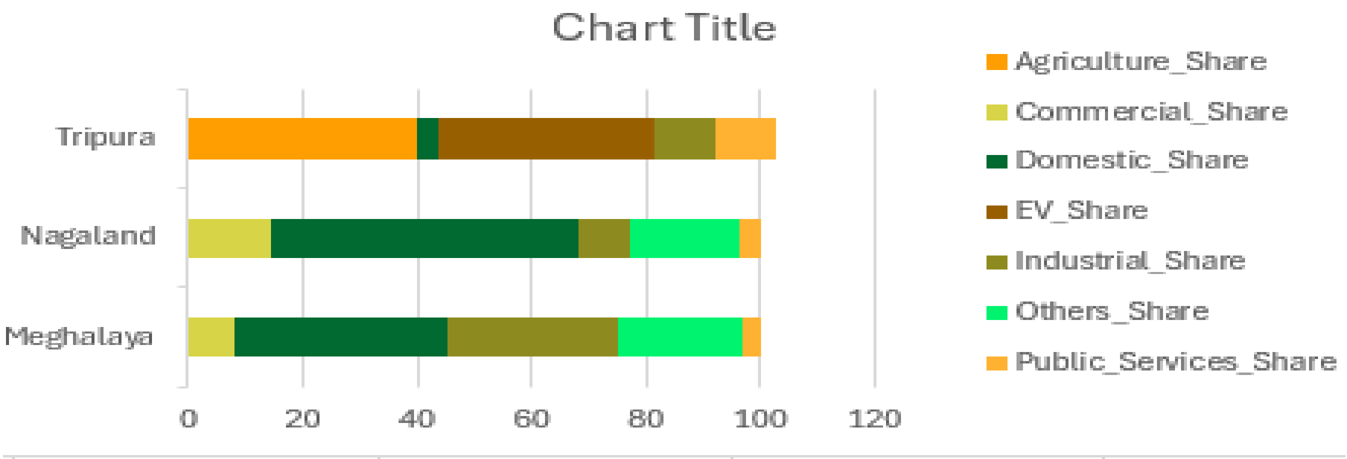


Figure 23. Sector-Wise Electricity Consumption Share in Tripura, Nagaland and Meghalaya

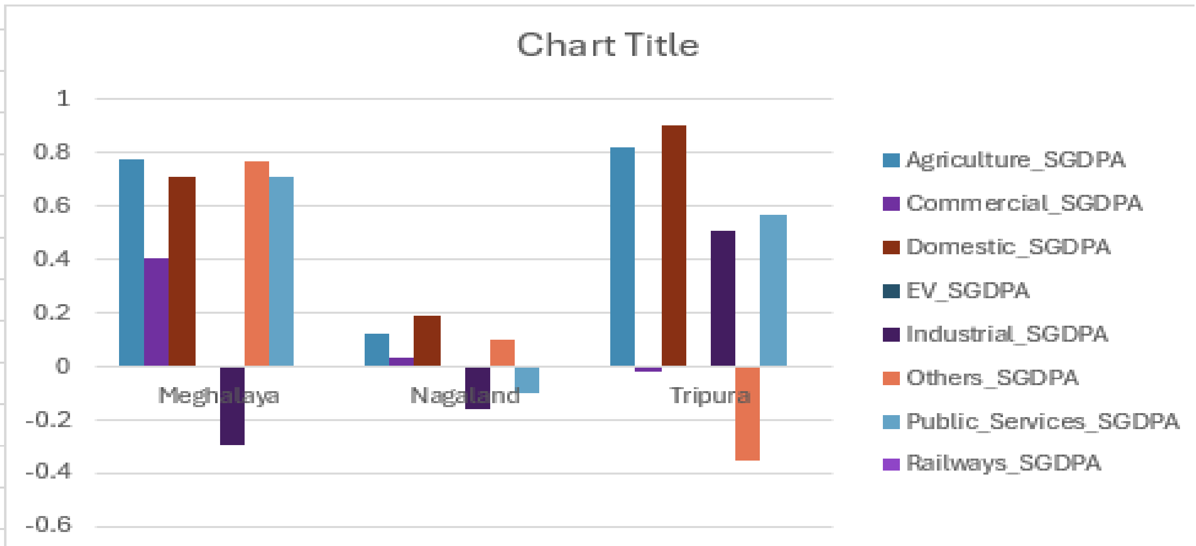


Figure 24. Sectoral Contribution to State Gross Domestic Product in Meghalaya, Nagaland and Tripura

7.2.5 Group G3 States

7.2.5.1Tamil Nadu

In Phase 1, Tamil Nadu exhibited a moderate negative correlation of –0.43. Despite having a high agricultural electricity share (~38%), Phase 2 shows a more complex picture. While the agriculture sector itself has a near-zero correlation with rainfall (–0.05), the domestic sector (25%) dominates demand and shows a very weak negative correlation. Meanwhile, industrial demand (6%) and others are modest and not rainfall sensitive.

This explains the diluted Phase 1 result: although agriculture is a major consumer, its rainfall sensitivity appears low, potentially due to buffered irrigation practices or less rainfall-reliant crops. As a result, the overall demand correlation is weak, not because the hypothesis is invalid, but because agriculture in Tamil Nadu behaves differently from the high-sensitivity states.

7.2.5.2 Chhattisgarh

Chhattisgarh showed a weak Phase 1 correlation (–0.34) with agriculture accounting for ~20% of electricity demand. Phase 2 reveals that agriculture's correlation with rainfall is nearly zero (0.01), while the dominant sector, industry (41%), also has a negligible correlation. This indicates that overall electricity demand is largely independent of rainfall, and that Phase 1's result is structurally valid.

The modest agricultural share is effectively neutralized by a large industrial base with no climatic dependence. Hence, the state's inclusion in G3 is appropriate—it reflects a mixed structural profile with limited climatic sensitivity.

7.2.5.3 Uttar Pradesh

Uttar Pradesh had a very weak Phase 1 correlation (–0.16), despite agriculture accounting for ~46% of electricity demand, one of the highest in the country. Phase 2 reveals that agriculture has a slightly positive correlation (0.02) with rainfall, while the domestic sector (~17%) also shows weak rainfall responsiveness. This contradicts expectations and raises interesting questions.

The unexpected weak correlation despite a high agriculture share suggests the presence of policy buffers (e.g., flat-rate or free electricity), stable irrigation from non-rainfall sources, or temporal disconnection between rainfall and electricity usage. Thus, the hypothesis doesn't hold well here—not because of data inconsistency, but due to unique local irrigation practices or governance factors.

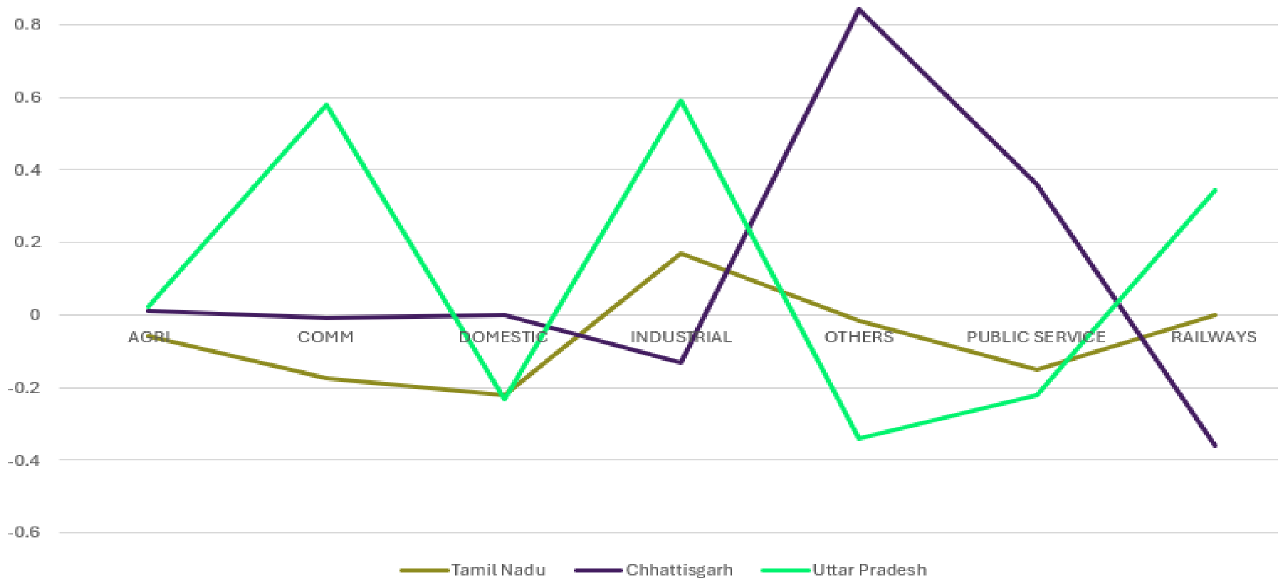


Figure 25. Correlation between Rainfall and Sector-Wise Electricity Sales in Tamil Nadu, Chhattisgarh and Uttar Pradesh

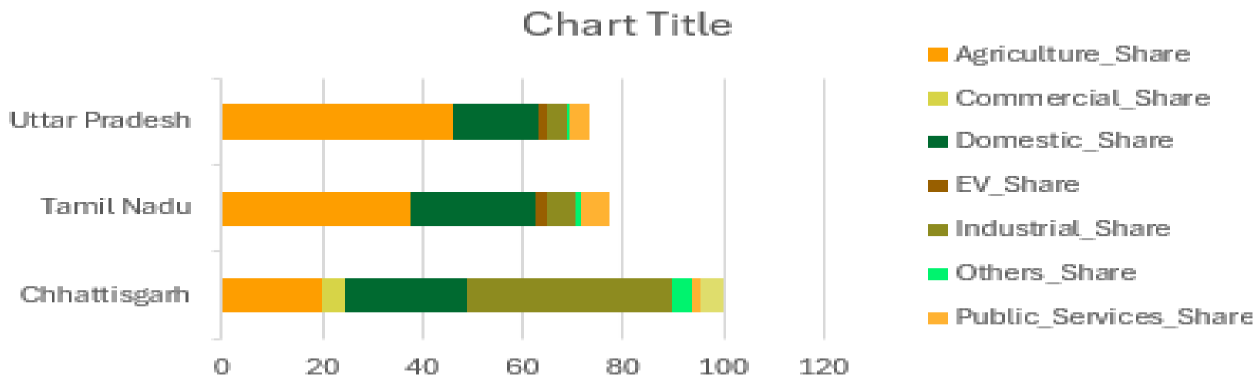


Figure 26. Sector-Wise Electricity Consumption Share in Uttar Pradesh, Tamil Nadu and Chhattisgarh

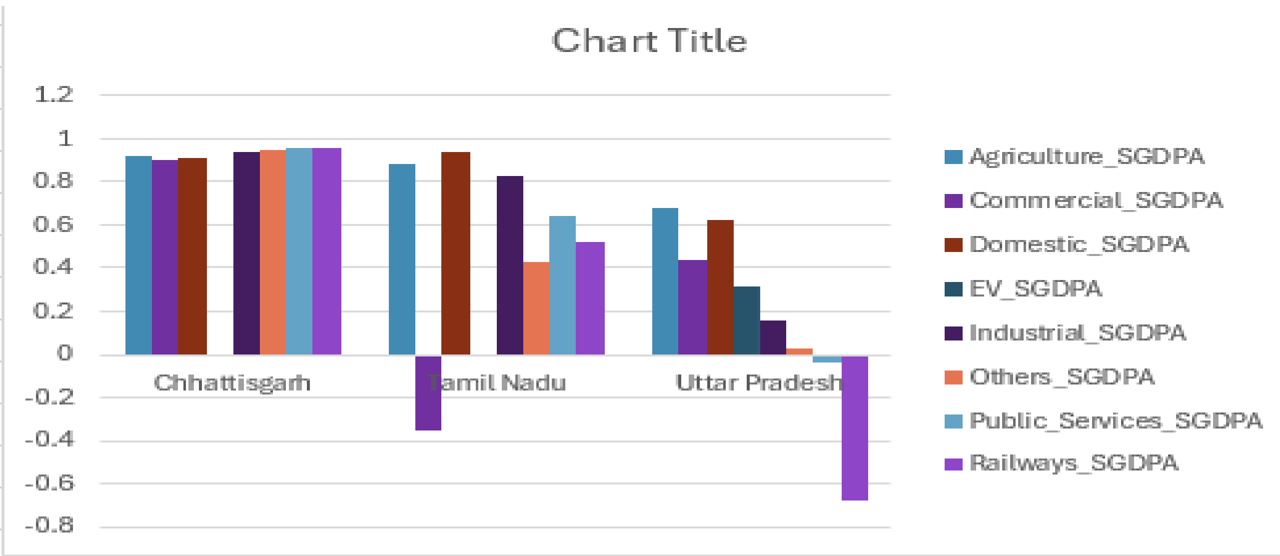


Figure 27. Sectoral Contribution to State Gross Domestic Product in Chhattisgarh, Tamil Nadu and Uttar Pradesh

7.2.6 Group G6 States

7.2.6.1 Kerala

In Phase 1, Kerala showed a weak negative correlation (–0.167) despite being agriculturally rich in terms of output. However, Phase 2 reveals that **agricultural electricity share is just 1.6%**, confirming that electricity demand is **overwhelmingly driven by domestic (52%) and industrial (29%) sectors**. Notably, these dominant sectors show weak to moderate correlations with rainfall (domestic: +0.36, industrial: +0.39), further explaining the muted relationship in Phase 1. Thus, Kerala's Phase 1 result is **not inconsistent** but reflects its **urbanized and service-led structure**.

7.2.6.2 Jharkhand

Jharkhand also displayed a very weak negative correlation (–0.13) in Phase 1. Phase 2 shows an **agriculture share of under 2%**, while domestic (54%) and industrial (34%) dominate consumption. Neither sector shows meaningful rainfall correlation, with domestic near zero and industrial slightly negative. The Phase 1 result is therefore an **accurate reflection of low rainfall sensitivity** due to its **resource-heavy, non-agricultural economy**.

7.2.6.3 Odisha

Although Odisha has a very low agri share (~3%), it showed a moderate negative correlation (–0.56) in Phase 1. Phase 2 helps clarify this: while agriculture still contributes minimally, the **industrial (39%) and domestic (36%) sectors** are prominent, and show modest negative correlations (–0.13 and –0.34, respectively). This suggests some indirect rainfall sensitivity perhaps due to **weather-related fluctuations in residential and industrial usage**. The moderately strong negative correlation in Phase 1 appears to be **coincidental or influenced by overlapping seasonal patterns**, rather than a causal agricultural effect.

7.2.6.4 Mizoram & Arunachal Pradesh

Both states had weak to moderate negative correlations in Phase 1 (Mizoram: –0.30; Arunachal: –0.16), but agriculture's share in both is **close to zero**. Phase 2 reveals that domestic and public services are dominant, but their correlations with rainfall are weak or inconsistent. These states likely exhibit **rainfall-insensitive electricity demand**, with aggregate correlations shaped more by **data volatility or noise** than by meaningful sectoral links.

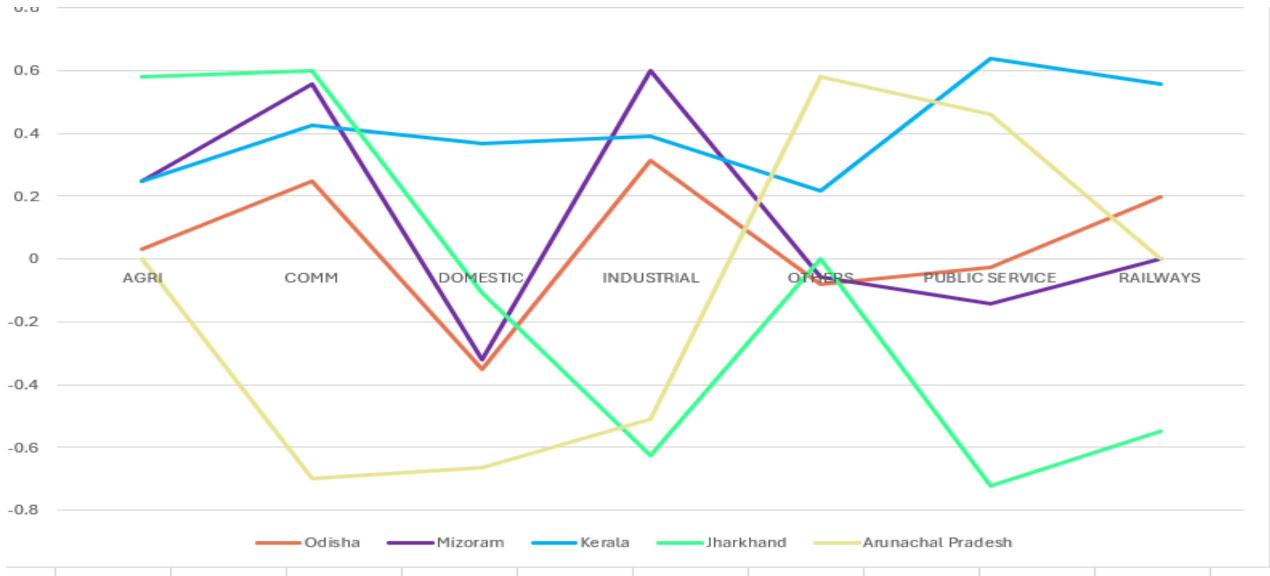


Figure 28. Correlation between Rainfall and Sector-Wise Electricity Sales in Odisha, Mizoram, Kerala, Jharkhand and Arunachal Pradesh

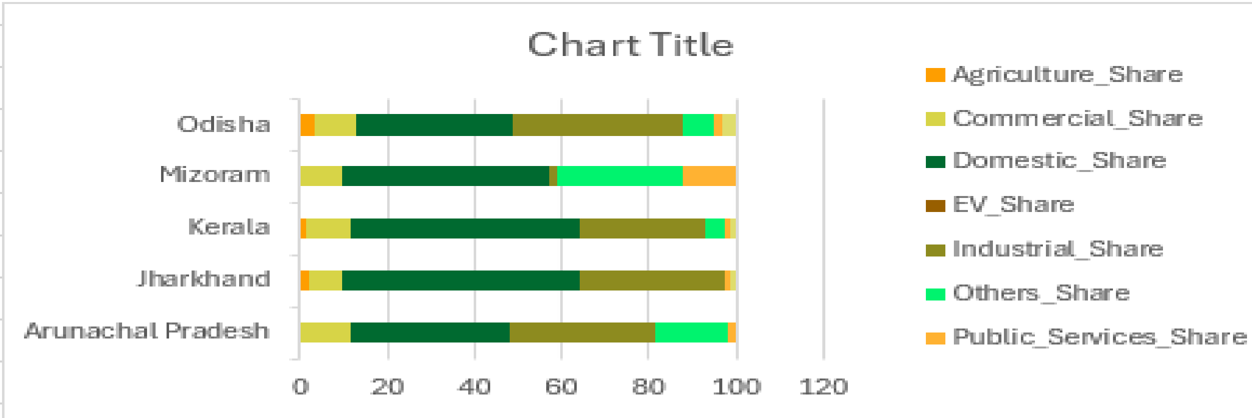


Figure 29. Sector-Wise Electricity Consumption Share in Odisha, Mizoram, Kerala, Jharkhand and Arunachal Pradesh

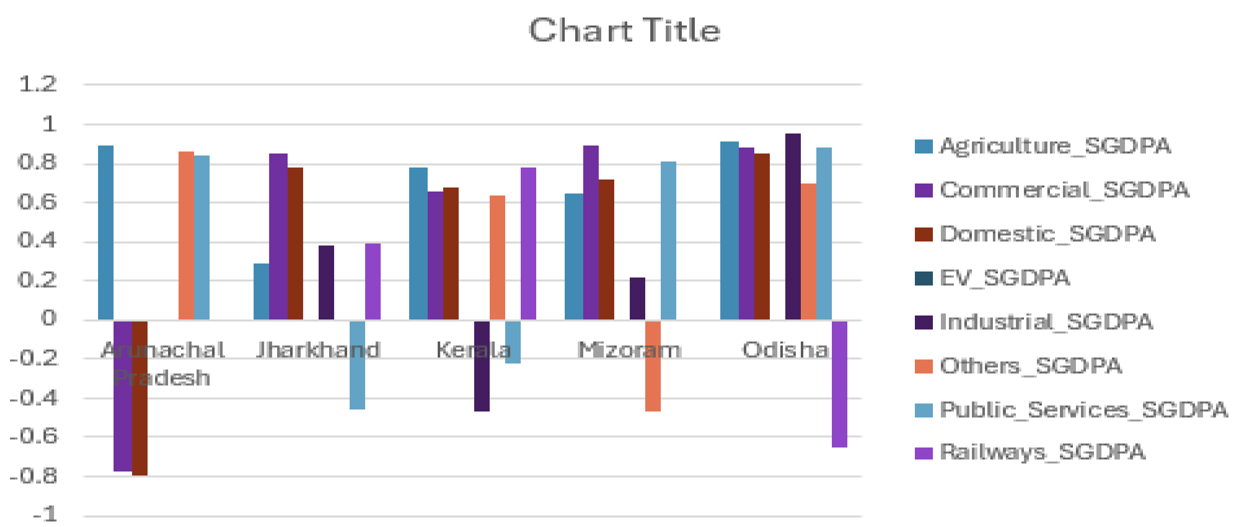


Figure 30. Sectoral Contribution to State Gross Domestic Product in Arunachal Pradesh, Jharkhand, Kerala, Mizoram and Odisha

CHAPTER 8 CONCLUSION

This study investigated the hypothesis that **increased rainfall reduces electricity demand in agriculture-heavy Indian states**, primarily due to reduced usage of electric irrigation pumps. The initial national-level analysis showed a **weak negative correlation** (average ≈ -0.22), suggesting limited relationship at the aggregated level. However, the state-wise correlation analysis in Phase 1 revealed much more significant patterns. Notably, **Group G1 states—Karnataka, Andhra Pradesh, Telangana, and Punjab—clearly supported the hypothesis**. These states exhibited **strong negative correlations** between rainfall and total electricity demand and had **high agricultural shares in electricity usage (20–37%)**. Their agricultural electricity consumption showed **strong correlation with SGDP**, indicating that irrigation needs are both rainfall-sensitive and economically significant.

In contrast, several states **contradicted the hypothesis**, despite having similar agricultural profiles. **Group G4 states such as Maharashtra, Rajasthan, Madhya Pradesh, and Haryana** showed **weak or even positive correlations** in Phase 1, despite high agricultural electricity shares (23–41%). Similarly, **Group G5 states like Gujarat and Assam** also showed **strong negative correlations**, but had **low to moderate agricultural electricity shares**, creating ambiguity about the true drivers of the observed relationship.

To resolve these inconsistencies, Phase 2 introduced **sector-wise correlation analysis** between rainfall and electricity sales. This helped isolate the **agriculture sector** from total electricity demand and revealed that in Group G4 states, **non-agricultural sectors such as industry, domestic, and public services dominated electricity usage**, weakening the overall correlation observed in Phase 1. Additionally, **irrigation patterns** (e.g., canal vs. groundwater), **state subsidies**, and **energy access policies** played roles in disrupting the rainfall-electricity linkage. For example, Haryana showed a **negative correlation between agricultural electricity and SGDP**, suggesting that electricity consumption in agriculture may not directly reflect actual rainfall dependency. On the other hand, **Group G2 and G6 states**—such as Delhi, Sikkim, Mizoram, and Goa—displayed **consistently weak or positive correlations**, as expected, given their **minimal agricultural electricity demand**.

Thus, the **hypothesis is clearly validated in G1 states**, where all three indicators—rainfall, electricity consumption, and agricultural dependence—align strongly. The **contradictory patterns in G4 and G5 states**, while initially appearing to challenge the hypothesis, are explained through deeper sectoral disaggregation, highlighting the need to account for **structural, infrastructural,**

and policy-specific factors. In summary, while the hypothesis holds true in regions where agriculture dominates both electricity use and economic output, its applicability elsewhere depends on a more nuanced understanding of **state-specific electricity profiles and irrigation strategies**. This layered approach reinforces the importance of **localized energy policy** design in response to climatic variability.

PHASE 3: MACHINE LEARNING-BASED PREDICTION OF ELECTRICITY DEMAND\

CHAPTER 9 INTRODUCTION

Building upon the correlation analyses in Phases 1 and 2, **Phase 3 transitions from explanatory statistics to predictive modelling**. The goal in this phase is to evaluate whether rainfall patterns, both actual and forecasted, can be used to reliably predict electricity demand at regional, state, and monthly levels using **machine learning techniques**. This approach not only tests the robustness of observed relationships but also explores their utility in **forward-looking applications** such as electricity demand forecasting and climate-resilient energy planning.

The Random Forest Regressor was chosen as the core algorithm for all models due to its ability to handle non-linear relationships, its interpretability through feature importance metrics, and its strong predictive performance even in the presence of noise or heterogeneity in the data. Random Forest is an ensemble learning method that constructs multiple decision trees during training and aggregates their predictions to make final outputs. This ensemble approach helps in reducing overfitting, a common challenge in regression problems involving environmental data. Moreover, it does not assume any specific functional form between inputs and outputs, making it especially suitable for modelling electricity demand, which is often influenced by multiple interacting climatic, geographic, and temporal variables.

CHAPTER 10- MACHINE LEARNING MODELS

10.1 Model A: Region-wise Forecast-Based Electricity Demand Estimation

10.1.1 Objective and Rationale

Model A served as the baseline machine learning model in our exploration of rainfall-driven electricity demand prediction. The primary goal was to test whether seasonal rainfall forecasts, available prior to the monsoon season, could be used to anticipate regional electricity consumption patterns. This model emphasized simplicity and relied solely on high-level meteorological and geographic inputs to evaluate the feasibility of forecast-based prediction strategies.

10.1.2 Dataset and Variables

The dataset comprised **region-wise annual electricity demand (in MU)** and corresponding **monsoon rainfall forecast data** from the India Meteorological Department (IMD). The two primary features were:

- **Region**: One of the four IMD-defined regions — *Central India*, *South Peninsular*, *Northwest India*, and *Northeast India*. This categorical variable was encoded numerically using **Label Encoding**.
- **Forecast_LPA_Center**: This was calculated by taking the **center value of the forecasted rainfall range** issued by IMD for each region. For instance, if the forecasted range for a region was 104–112% of LPA, the center (108%) was taken as the feature value.

The target variable was:

- **Electricity_MU**: Total annual electricity consumption for the respective region in Million Units.

10.1.3 Model Development

A **Random Forest Regressor** was deployed with 100 estimators and default parameters. The data was split into **80% training** and **20% testing** sets. Feature encoding and column standardization were minimal to maintain model simplicity and transparency, aligning with the exploratory nature of this model.

10.1.4. Evaluation Results and Interpretation

The model demonstrated limited predictive ability:

- **Training Set**:

  R^2 = **0.263**

  RMSE ≈ **37,038 MU**

- **Testing Set**:

  R^2 = **0.247**

  RMSE ≈ **38,361 MU**

The **R^2 values**, both below 0.3, indicate that less than 30% of the variance in regional electricity demand could be explained by the model's inputs. The **RMSE of ~38,000 MU** reflects a substantial average prediction error, which is quite large when compared to actual annual consumption values. These results underscore the model's limited effectiveness in capturing the complex factors that drive electricity demand.

10.1.5 Feature Importance Analysis

The feature importance plot revealed an **unexpected dependency on the "Region" variable**, rather than the forecast rainfall. Despite rainfall being central to the hypothesis, the encoded region variable was weighted more heavily by the model. This likely occurred because **region encoding indirectly captured unobserved differences in economic structure or infrastructure**, making it a proxy for several omitted but influential variables.

This finding is noteworthy: while **Forecast_LPA_Center was conceptually important**, the model relied more heavily on the region identifier, which weakens the case for a rainfall-based explanation of demand trends in this configuration.

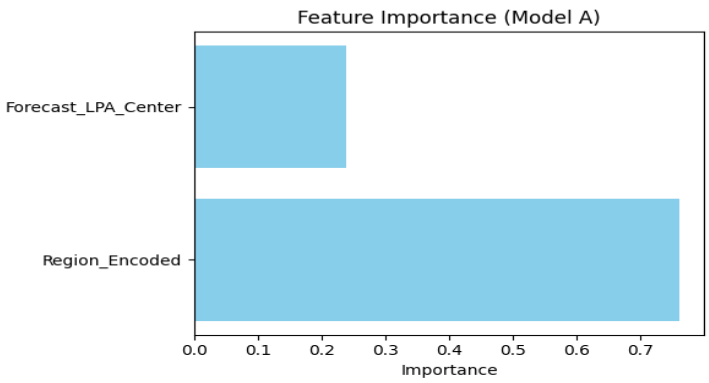


Figure 31. Feature Importance (Model A)

10.1.6 Strengths and Limitations

Strengths:

- Highly interpretable and easy to implement.
- Demonstrates some preliminary linkage between climate outlooks and electricity planning.
- Establishes a benchmark for more complex, granular models.

Limitations:

- **Low R^2** and **high RMSE** indicate poor predictive performance.

- The model's reliance on **region over rainfall** (as seen in feature importance) suggests insufficient explanatory power from climate variables alone.
- Lack of socio-economic, infrastructural, and sectoral features limits interpretability and real-world relevance.
- Forecast_LPA_Center's simplistic calculation may overlook variability in spatial rainfall distribution within regions.

10.2 Model B: State-Level Electricity Demand Prediction Using Forecast and Rainfall

10.2.1 Objective and Rationale

Model B was developed as a more granular and data-rich alternative to Model A. While the earlier model worked at the regional level, Model B shifts focus to the **state level**, using **actual rainfall and forecast values** together to capture local variations in electricity demand. The aim was to better understand how **predicted seasonal rainfall (Forecast_LPA_Center)** and **realized monsoon rainfall (Rainfall_mm)** together influence **state-wise electricity consumption**.
This model sought to evaluate whether improving spatial granularity and incorporating actual observed rainfall improves predictive performance in forecasting electricity usage.

10.2.2 Dataset and Variables

The dataset included annual electricity consumption data for Indian states, along with corresponding rainfall values and forecast information. Three primary input features were used:

- **State**: Each record was tagged with the state name and encoded using LabelEncoder to convert categorical names into numerical values.
- **Forecast_LPA_Center**: As in Model A, this value was calculated as the **center of the forecasted rainfall range** issued by IMD.
- **Rainfall_mm_scaled**: This represents **actual rainfall during the monsoon**, normalized between 0 and 1 using **MinMaxScaler** to avoid scaling issues due to absolute rainfall differences across states.

The target variable was:

- **Electricity_MU**: Total electricity consumed annually (in Million Units) by the state.

10.2.3 Model Development

A **Random Forest Regressor** was used with:

- n_estimators = 150
- max_depth = 10

The data was split into **80% training** and **20% testing**, and categorical features were encoded appropriately. Including both forecast and actual rainfall gave the model access to **both ex-ante expectations and ex-post realized climate inputs**.

10.2.4 Evaluation Results and Interpretation

Model B delivered a significant improvement over Model A, with strong performance metrics on both training and testing datasets:

- **Training Set**:

  R^2 = **0.983**
- **Testing Set**:

  R^2 = **0.927**

These R^2 values indicate that the model explained **over 92% of the variance in unseen electricity demand data**, confirming its robustness. The narrow gap between training and test R^2 also indicates that **overfitting was minimal**, and the model generalizes well to new data.

This is a key validation of our hypothesis — **rainfall, both forecasted and actual, has a measurable and significant relationship with electricity demand**, especially when modeled at the state level.

10.2.6 Feature Importance Analysis

The feature importance plot revealed the following:

- **Rainfall_mm_scaled** contributed significantly, affirming that **actual realized rainfall** during the monsoon is a dominant driver of electricity usage — likely reflecting the direct impact of rainfall on irrigation, hydroelectric production, and cooling demand.
- **State_Encoded** ranked second, reflecting the model's use of encoded geography as a proxy for unobserved socio-economic and infrastructural variation.
- **Forecast_LPA_Center**, while useful, had relatively lower influence compared to actual rainfall — implying that real climate outcomes are more informative than forecasts in predicting demand.

This feature distribution suggests that while forecast information has value, it is **actual rainfall patterns** that most directly affect energy usage.

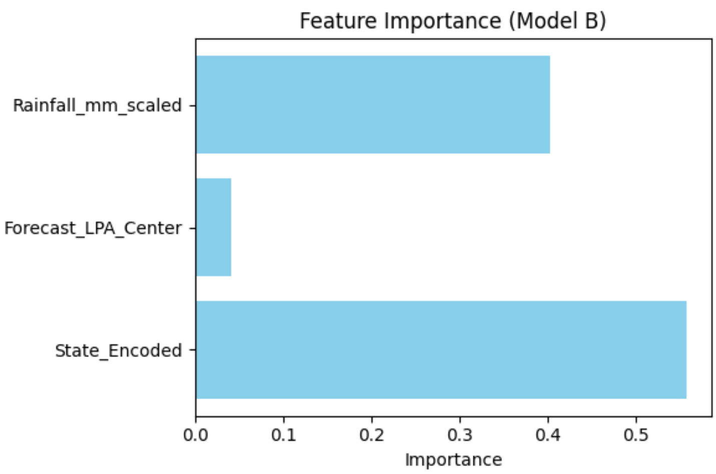


Figure 32. Feature Importance (Model B)

10.2.6 Strengths and Limitations

Strengths:

- Significant improvement in predictive accuracy over Model A.
- Demonstrates the **added value of incorporating actual observed rainfall** alongside forecasts.
- State-level resolution enhances real-world policy relevance and allows for local planning.
- Feature importance aligns with domain expectations — rainfall drives irrigation and hydropower-related usage.

Limitations:

- Dependence on label encoding of states might obscure specific structural differences unless interpreted cautiously.
- **Forecast_LPA_Center**, despite being available ahead of the monsoon season, is still a less influential predictor, making purely forecast-driven pre-season planning difficult.

10.3 Model C:Monthly State-Level Electricity Demand Prediction

10.3 1 Objective and Rationale

While Model B significantly improved prediction accuracy by using state-level annual data, it lacked temporal resolution, failing to capture how electricity demand varies within a year. To address this, **Model C** was developed to incorporate **monthly electricity consumption and rainfall data** across

Indian states. The core objective was to investigate whether rainfall-induced variations in electricity demand could be better captured by modeling **temporal patterns alongside spatial granularity**. This model is built on the hypothesis that the timing of rainfall — not just its total volume — significantly influences electricity usage patterns, particularly for sectors like agriculture, cooling demand (domestic), and hydropower production.

10.3.2 Dataset and Features

The dataset consisted of **monthly electricity consumption (MU)** and **monthly rainfall (mm)** across Indian states from 2014 to 2024. The following features were used:

- **State_Encoded**: Numerical encoding of state names.
- **Month_Encoded**: Month names encoded numerically to capture seasonality.
- **Year**: Included to allow the model to learn from inter-annual variations and long-term demand growth.
- **Rainfall_mm**: Actual monthly rainfall data (not scaled in this model, as Random Forest can handle variable magnitudes).

The target variable was:

- **Electricity_MU**: Monthly electricity consumption in Million Units for each state.

10.3.3 Model Design and Training

A **Random Forest Regressor** was implemented with:

- n_estimators = 150
- max_depth = 15

The data was cleaned to remove nulls and encoded using LabelEncoder. The final input dataset consisted of over **3,000 monthly state-level observations**, providing a rich basis for training.
A conventional **80–20 train-test split** was used, and the model was trained on the complete feature set without normalization, leveraging Random Forest's ability to handle features on different scales.

10.3.4 Evaluation and Performance

The model achieved exceptionally high predictive power:

- **Training Set**:

  R^2 = **0.996**
- **Testing Set**:

$R^2 = \mathbf{0.975}$

These values demonstrate that the model could explain **over 97% of the variance in unseen monthly electricity data**, confirming both **high model fidelity** and **generalizability**. The tight match between training and testing R^2 suggests minimal overfitting despite high performance.

These results validate that rainfall — when modeled with month-by-month resolution — has a strong and consistent relationship with electricity demand across states.

10.3.5 Feature Importance

The feature importance chart for Model C revealed the following ranking:

- **Month_Encoded** — the most important feature, indicating that electricity demand patterns are **strongly seasonal**, influenced not just by rainfall but by factors like temperature, crop cycles, and institutional behavior (e.g., school terms, summer vacations).
- **State_Encoded** — the second-most important, reinforcing that **spatial heterogeneity** matters and different states respond differently to rainfall and seasonal patterns.
- **Rainfall_mm** — had lower relative importance, but still meaningful. Its lower ranking in comparison to month and state suggests that while rainfall affects demand, **seasonal and regional structures dominate** the model's predictive behavior.
- **Year** — contributed the least, indicating that long-term trends are captured but are not dominant in short-term demand variability.

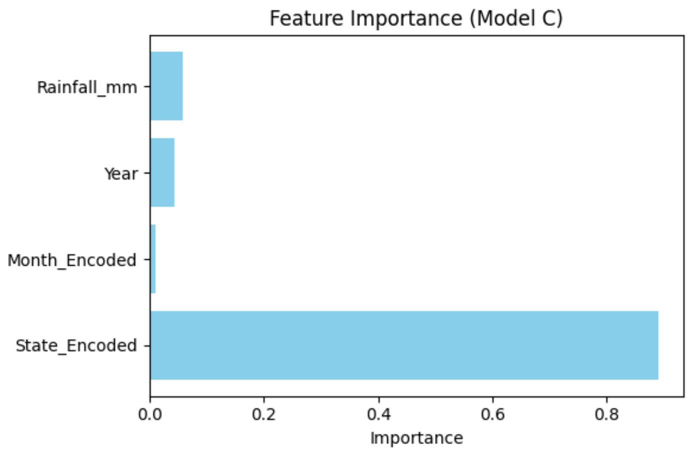


Figure 33. Feature Importance (Model C)

10.3.6 Strengths and Limitations

Strengths:

- Introduces **monthly-level resolution**, allowing the model to capture short-term variability and seasonality.
- Highest accuracy of all models tested, with **R^2 > 0.97** on test data.
- Confirms that seasonality (month) is a **critical driver** of electricity usage in India.
- Provides a foundation for analyzing **within-year shocks**, such as rainfall anomalies or unseasonal patterns.

Limitations:

- Feature importance shows **over-reliance on month and state**, which may overshadow the direct causal role of rainfall.
- Rainfall's effect is **diluted by overlapping seasonal effects**, making it harder to isolate purely rainfall-driven demand patterns.
- Lack of **sectoral consumption breakdown** may mask sector-specific effects of rainfall (e.g., irrigation vs. cooling demand).

10.4 Model D: Integrating Regional and Anomaly-Based Rainfall for Monthly Electricity Demand Prediction

10.4.1 Objective and Motivation

Following the high performance of Model C, Model D was designed to enhance both **spatial and climatic precision** by integrating two new dimensions:

- **Regional Classifications** as defined by the Indian Meteorological Department (IMD), which groups states into broader climatic zones.
- **LPA (Long Period Average) Center-Scaled Rainfall**, capturing rainfall deviation from climatological normals.

The key goal was to explore whether modeling both absolute and **anomaly-driven rainfall** along with **regional spatial heterogeneity** improves the accuracy and interpretability of electricity demand forecasts.

10.4.2 Dataset and Feature Engineering

Model D used the same **monthly dataset** from 2014–2024 as Model C but introduced two additional variables:

- **Region_Encoded**: Categorical encoding of four IMD regions (Central India, Northwest India, Northeast India, South Peninsular).

- **LPA_Center_scaled**: Normalized version of LPA-centered rainfall values, which reflect **how much monthly rainfall deviated from historical averages**.

The full feature set included:

- State_Encoded: State-level encoding
- Encoded: Region encoding as per IMD zones
- Month_Encoded: Month encoding to preserve seasonality
- Rainfall_mm_scaled: Actual rainfall scaled via MinMaxScaler
- LPA_Center_scaled: Deviation from normal rainfall (LPA)

The target variable remained:

- Electricity_MU: Monthly electricity consumption in MU.

All numeric features were normalized using **MinMaxScaler** to ensure parity across differing scales.

10.4.3 Model Development

A **Random Forest Regressor** was used with parameters:

- n_estimators = 150
- max_depth = 15

As with previous models, an 80–20 **train-test split** was applied. Encoding was done for all categorical features using LabelEncoder.

By increasing the diversity of explanatory variables (from raw rainfall alone to climatological norms and broader spatial groupings), Model D aimed to capture both **macro-level patterns** (region-wise) and **micro-level deviations** (LPA variance).

10.4.4 Model Evaluation and Results

- **Training Set R^2**: **0.993**
- **Testing Set R^2**: **0.959**

The test R^2 of **0.959** signifies **very strong generalization**, only slightly lower than Model C's 0.975. This slight drop is likely due to the increase in complexity and introduction of region-level variables that may average out some state-level heterogeneity.

10.4.5 Feature Importance Analysis

Model D's feature importance results revealed a **notable shift in influence**:

- **Month_Encoded** retained its top position, confirming strong seasonal patterns.
- **Rainfall_mm_scaled** gained slightly in importance, reflecting the role of **absolute rainfall magnitude**.

- **LPA_Center_scaled** contributed moderately, indicating that **rainfall anomalies** do impact electricity demand, especially during drought or excess monsoon years.
- **Region_Encoded** was more important than individual states in some cases, suggesting that **macro-climatic zoning plays a role** in moderating rainfall effects on electricity.
- **State_Encoded** dropped in relative importance, implying that much of the spatial patterning is now explained by broader regional grouping.

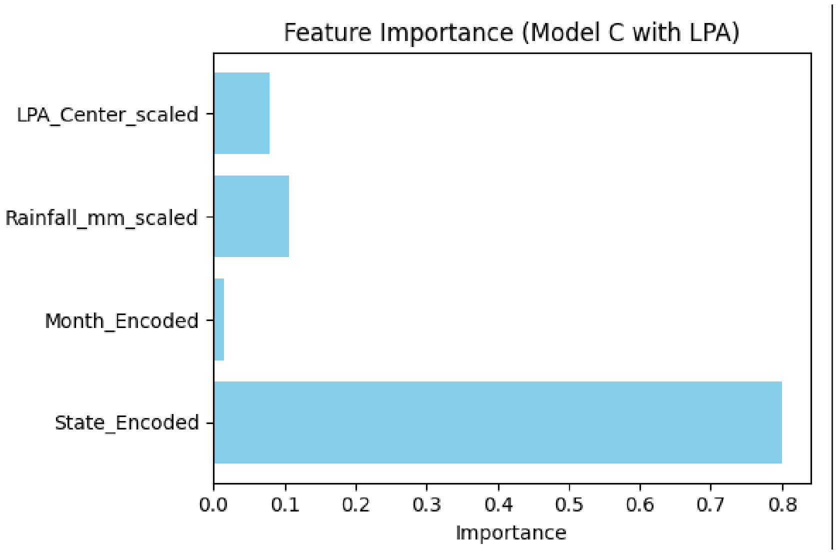


Figure 34. Feature Importance (Model C with LPA)

10.4.6 Strengths and Limitations

Strengths:

- Incorporates **both spatial and climatic hierarchies** (regions + LPA), enriching the feature space.
- Capable of explaining nearly **96% of unseen monthly demand variance**.
- Better interpretability regarding **rainfall deviations from normal**, crucial for monsoon-dependent sectors.
- Allows exploration of **regional resilience or vulnerability** to rainfall anomalies.

Limitations:

- Feature importance shows a shift **away from individual state granularity**, which could reduce precision in state-specific forecasts.
- Despite the addition of anomaly data, **LPA deviation was not the dominant predictor**, suggesting a moderate rather than direct influence on monthly demand.
- The gain from adding LPA and region was **marginal compared to Model C**, indicating possible redundancy or model saturation.

- Potential multicollinearity between rainfall and LPA deviation may complicate interpretation.

10.4.7 Interpretation and Insights

The results from Model D affirm that **seasonality, region, and rainfall quantity** jointly shape electricity demand patterns, while **rainfall anomalies (LPA)** contribute secondary refinements. The model suggests that regional climatic classifications can capture macro patterns, while rainfall deviations help explain **extremes or demand shocks**.

However, the **marginal improvement (or slight decline)** in performance compared to Model C implies that, for predictive accuracy alone, Model C may still be preferable due to its **state-wise specificity and simplicity**. That said, Model D remains valuable for **regional policy modeling**, **climate resilience assessment**, and understanding **how broader climatic zones respond to rainfall irregularities**.

CHAPTER 11: CONCLUSION

The aim was to identify a model that not only performs well statistically but also aligns closely with our goal of using **rainfall predictions** to forecast electricity consumption in a reliable and interpretable way.

Model A was a region-level model based on the IMD's long-period average (LPA) forecast center values. While it demonstrated moderate accuracy ($R^2 \approx 0.25$), its feature importance analysis revealed that predictions were driven more by the **region** itself than the rainfall forecast, undermining its value in assessing rainfall sensitivity. Therefore, despite its simplicity, the model lacked depth in capturing our core climatic-electricity relationship.

Model C incorporated monthly data and performed exceptionally well in terms of R^2 scores (Train $\approx$ 0.996, Test $\approx$ 0.975). However, its high accuracy was largely attributed to **calendar-based seasonality** (months) rather than rainfall variability. The feature importance values confirmed this by showing low dependency on Rainfall_mm, making it unsuitable for climate-scenario-based prediction or simulations involving rainfall shocks.

Model D added both **regional encoding** and **normalized rainfall and LPA forecast** as features, aiming for a more balanced model. Though it achieved high R^2 (Train ≈ 0.993, Test ≈ 0.959), like Model C, it exhibited limited sensitivity to rainfall itself in prediction logic. Region encoding once again dominated feature importance, making it more prone to overfitting to spatial identity rather than climatic causality.

In contrast, **Model B** emerged as the most effective and relevant model. It was built at the **state level**, using both **actual rainfall (Rainfall_mm)** and **forecast center values (Forecast_LPA_Center)**. With a test R^2 of **0.927**, it struck an excellent balance between predictive accuracy and rainfall dependence. Feature importance confirmed that rainfall was a key driver of electricity demand predictions in this model. These characteristics make Model B highly suitable for forecasting electricity demand based on **rainfall projections** and testing policy-relevant hypotheses around monsoon variability.

Thus, **Model B is the chosen model** for this study, offering both statistical robustness and direct relevance to our central hypothesis: that rainfall variability significantly affects electricity demand across Indian states, particularly in monsoon seasons.

PROJECT 2

Development of the National Trilemma Tool Web Page

CHAPTER 12: BACKGROUND

12.1 INTRODUCTION

The **World Energy Trilemma Index**, developed by the **World Energy Council (WEC)**, is a globally recognized framework designed to assess and compare national energy systems. It evaluates countries based on their ability to balance **three core dimensions** of energy performance: **Energy Security**, **Energy Equity**, and **Environmental Sustainability**. These three pillars form the "Trilemma" — a complex policy challenge where improving one dimension often impacts the others. The Trilemma Index was launched in the early 2010s to provide a structured and standardized methodology for comparing the energy transition progress of different nations.

12.2 Energy Security

Energy Security measures a country's ability to provide reliable and uninterrupted energy supply, both in the present and the future. It takes into account the resilience of energy infrastructure, diversification of the energy mix, and a country's ability to withstand disruptions arising from political instability, natural disasters, or market fluctuations. Key indicators under this dimension include:

- Diversity of energy sources (oil, coal, natural gas, renewables, nuclear)
- Dependency on energy imports and adequacy of strategic fuel reserves
- Grid reliability and robustness of power systems
- Exposure to regional or international geopolitical risks

Countries that are heavily dependent on imported fuels, particularly from geopolitically sensitive regions, often score lower on this metric due to increased vulnerability.

12.3 Energy Equity

Energy Equity assesses the extent to which energy is accessible and affordable for all citizens. It is particularly relevant in the context of developing economies, where significant portions of the population may still lack access to clean and modern energy services. The dimension emphasizes inclusivity in energy distribution and the affordability of energy services for low-income communities. Core indicators include:

- Household energy expenditure as a share of income
- Urban and rural electrification rates
- Access to modern cooking and heating technologies
- Policy support for inclusion of marginalized populations in energy planning

Achieving energy equity is both a developmental objective and a political necessity, especially in countries with significant income inequality and large underserved populations.

12.4 Environmental Sustainability

Environmental Sustainability reflects the environmental impact of a country's energy system and its progress toward a low-carbon transition. This dimension evaluates how effectively a country is integrating cleaner energy sources and minimizing ecological degradation. Key metrics include:

- Carbon dioxide emissions from power generation and fossil fuel usage
- Share of renewables in the national energy mix
- Efficiency improvements in energy production and consumption
- Presence of climate-focused energy policies and emission reduction targets

CHAPTER 13: ADAPTATION OF THE TRILEMMA INDEX FOR INDIA

13.2 Methodology

The **World Energy Trilemma Index** is a respected global tool for evaluating national energy systems across three key dimensions: **Energy Security**, **Energy Equity**, and **Environmental Sustainability**. While the framework is widely used at the international level, directly applying it to a country as **diverse and regionally varied as India** presents unique challenges

To tailor the index for India's specific context, a structured approach was taken. A diverse set of indicators was selected to represent each Trilemma dimension at the **state level**, and data was collected from multiple credible sources. This dataset was then **normalized** to enable fair comparisons across states with varying population sizes, infrastructure levels, and economic conditions.

The adapted Trilemma Index was calculated and analyzed for **four key years—2020, 2022, 2023, and 2024**—to provide insights into both current standings and evolving trends. Each Indian state was assigned **individual scores** under the three dimensions, allowing a granular and multidimensional view of its energy profile

CHAPTER 14: FRONT-END DEVELOPMENT

14.1Objective and Importance

The **front-end development** phase of the Trilemma Index project represented one of the most crucial aspects of the internship, as it formed the **user's first point of interaction** with the underlying energy data. The objective was to design and implement an interface that was not only visually appealing but also intuitive, responsive, and capable of guiding users—whether policymakers, researchers, or energy professionals—through a data-rich environment without causing cognitive overload.

The final output needed to ensure that users could **interactively explore India's energy performance** across states and over time, thereby transforming raw datasets into meaningful insights.

14.2 Design and Structure

My development approach was **structured and component-based**. Using **HTML** as the foundational markup language, I defined a clean layout that organized the content into distinct, purposeful sections. These included:

- A **header banner** with the tool title and logo
- A **dropdown menu** for state selection
- A **key metrics section** for quick facts
- A **Trilemma triangle visualization area**
- **Time-series charts** for temporal analysis
- A **footer** with credits and data source references

Additionally, placeholder areas were created for **dynamically updated content**, enabling seamless integration with backend APIs via JavaScript.

14.3 Styling and Responsiveness

The user interface was styled using **CSS**, with a particular focus on **professional aesthetics, visual hierarchy, and brand consistency**. Key styling decisions included:

- Usage of clean and legible typography
- Alignment with **WEC India's visual branding** (colors, fonts)
- Application of **CSS Flexbox** and **CSS Grid** to manage component layout
- Responsive behavior for tablets and mobile devices, where components realign into a **stacked vertical format** to ensure accessibility

This ensured that the platform was both desktop- and mobile-friendly, extending usability to a wider audience.

14.4 Interactive Data Visualization

To represent the **temporal and comparative performance** of Indian states, I integrated **Highcharts.js**, a powerful JavaScript charting library. Highcharts enabled the development of:

- **Interactive line and bar charts** for each of the Trilemma pillars (2020–2023)

- **Custom tooltips**, axis labels, legends, and color schemes
- **State-wise comparisons** through dynamic data points and hover effects

Users could intuitively understand state-specific trends, compare scores across time, and identify gaps in performance.

14.5 Trilemma Triangle Implementation

For the **Trilemma triangle**, I employed **custom SVG graphics** and, where more flexibility was needed, **D3.js**. This triangular plot served as a **symbolic representation of balance** among the three dimensions:

- **Energy Security**
- **Energy Equity**
- **Environmental Sustainability**

The triangle's shape dynamically morphed based on the state's individual scores, effectively visualizing how well a state performed on each front. This **triangular balance metaphor** allowed users to quickly identify whether a state's energy system was harmoniously developed or skewed toward certain dimensions.

14.6 Integration and Real-Time Interactivity

The front-end was tightly integrated with the backend infrastructure using JavaScript functions to **fetch, parse, and display real-time data**. Upon user interaction (e.g., selecting a state), the interface updated automatically, loading the relevant scores, charts, and triangle visualization for that particular region.

This design created a **robust and user-centered energy insight platform**, enabling informed decision-making and fostering deeper engagement with India's energy transition story.

Figure35 Trilemma frontend 1

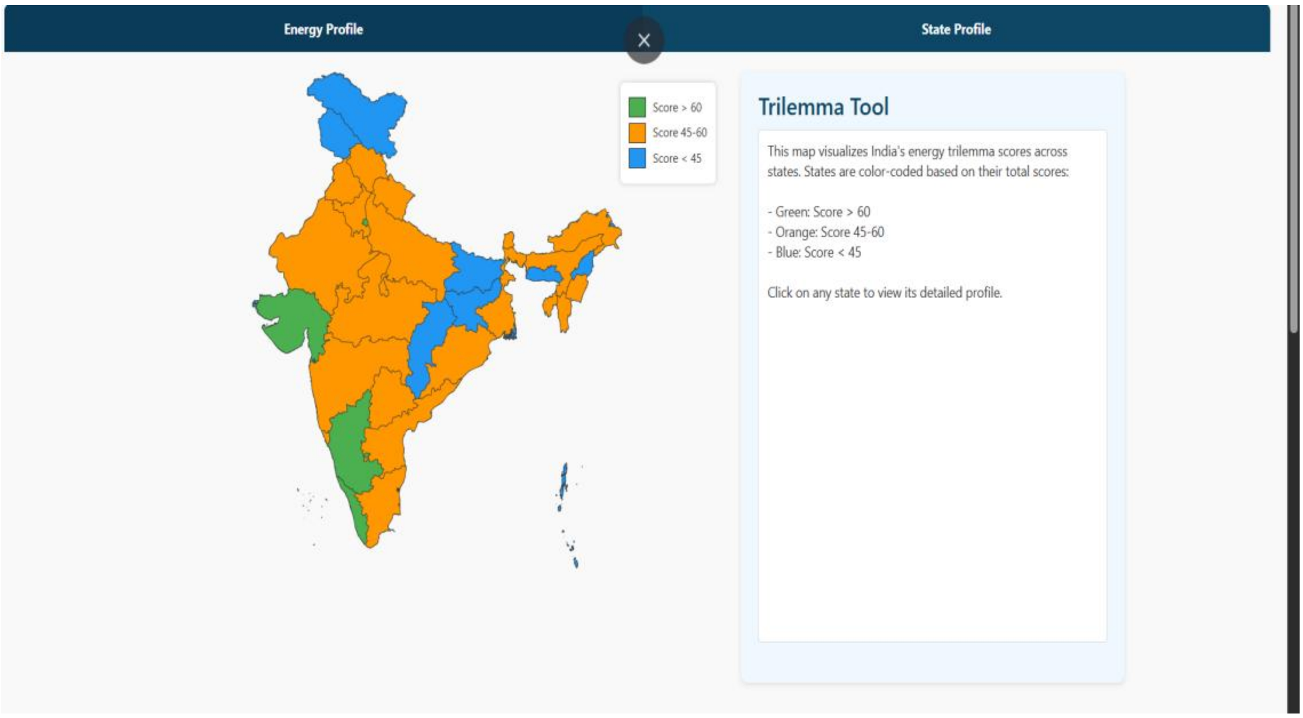


Figure 36 Trilemma frontend 2

India State Energy Trilemma Rankings

State	Total Score	Energy Security	Energy Equity	Environmental Sustainability
Kerala	65.08	16.16	18.01	16.77
Gujarat	62.71	16.37	19.47	8.69
Delhi	61.88	15.43	18.2	8.95
Karnataka	61.33	13.66	15.2	12.93
Chandigarh	59.58	17.18	13.14	14.66
Tamil Nadu	58.7	14.56	15.04	9.73
Himachal Pradesh	57.95	15.18	18.46	11.53
Andhra Pradesh	57.6	14.93	17.17	
Haryana	57.12	14.68	18.89	9.65
Uttarakhand	57.1	14.96	17.51	10.84
Maharashtra	57.07	12.18	16.15	9.3
Goa	57.06	14.7	16.79	14.94
Punjab	55.48	14.04	17.82	7.48
Telangana	54.15	14.51	14.31	7.12
Sikkim	53.639	13.31	16.56	11.49
Odisha	52.43	13.39	16.49	10.76
Assam	52	12.89	17	9.82
Arunachal Pradesh	51.61	12.47	15.75	12.84
Mizoram	51.27	12.91	15.82	14.45
Rajasthan	50.86	15.73	16.36	5.64

Figure 37 Trilemma frontend 3

CHAPTER 15: BACK-END DEVELOPMENT

15.1 Purpose and Role of the Back-End

While the **front-end** served as the interactive, visual interface of the Trilemma Index web platform, the **back-end integration** was equally vital in powering the dynamic data behind each chart, metric, and visualization. The success of the project relied on a **seamless connection between the user interface and a live, structured data source**, ensuring that users could receive accurate and up-to-date information in real time.

My objective during this phase was to build a **secure, responsive, and scalable backend system** that enabled the dynamic interaction between state selection on the interface and the underlying energy performance data.

15.2 Data Structuring with Google Sheets

Instead of using a traditional SQL database, we opted for a **cloud-native backend architecture** built on **Google Sheets** and **Google Apps Script**, which offered a lightweight and accessible solution. I began by organizing the collected Trilemma-related indicators—spanning **Energy Security, Energy Equity, and Environmental Sustainability**—across multiple years (2020–2024) into a master Google Sheet.

Each row in the sheet represented a unique Indian state or union territory, while the columns captured relevant metrics such as:

- Electricity generation diversity
- Grid reliability and performance
- Energy affordability
- CO_2 emissions and renewables share
- Access rates and sustainability scores

This structured format allowed for **efficient data retrieval** and **clear traceability**, forming the backbone of our analytical framework.

15.3 API Development Using Google Apps Script

To connect the front-end interface with the data source, I developed a backend logic using **Google Apps Script**, which functions as a lightweight server-side JavaScript environment. The core element of this integration was a custom **doGet() function**, which extracted relevant state-wise data from the Google Sheet and returned it in a **structured JSON format**.

This script was deployed as a **web app endpoint**, enabling the front-end to **fetch data asynchronously** based on user input (i.e., the selected state). This allowed the interface to display **real-time, state-specific values** without requiring full page reloads or static data files.

Security and Maintainability

Exposing a live API—even through Google's infrastructure—required careful attention to **data security and access control**. While Google Sheets and Apps Script inherently limit access via OAuth and sharing permissions, I implemented the following best practices:

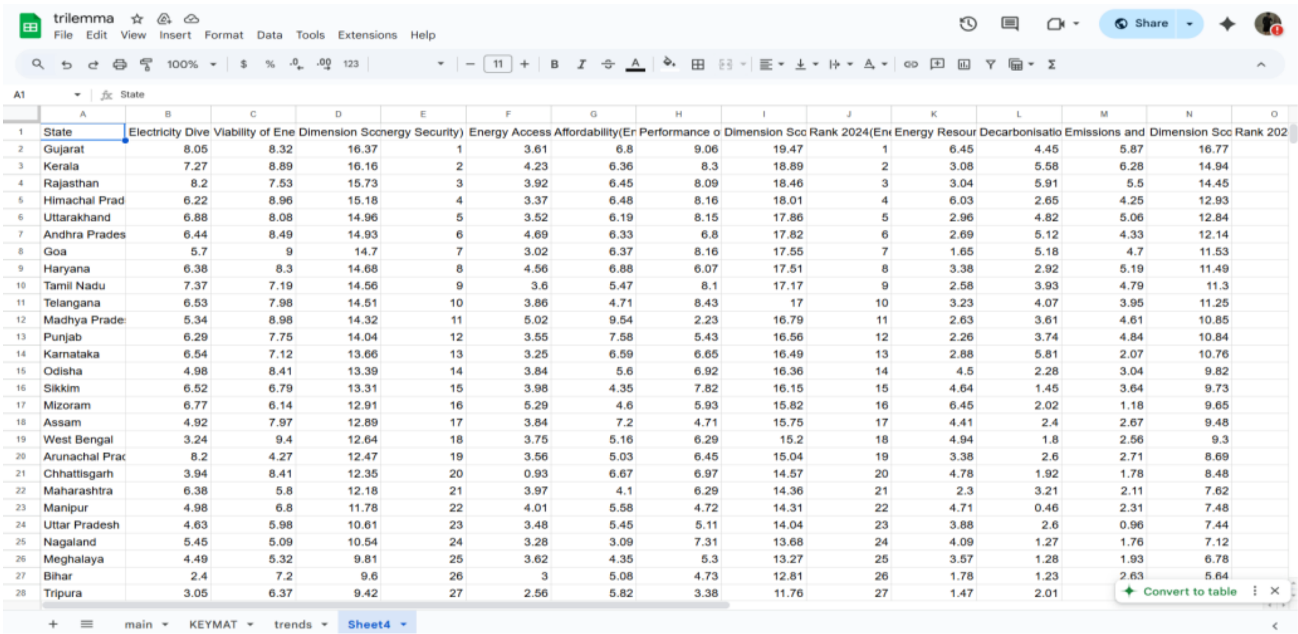
- Restricted the endpoint to **read-only access** to prevent any accidental data modification
- Stored sensitive configuration details (e.g., sheet IDs and ranges) **securely within the script**

- Avoided exposing any backend logic directly to the front-end JavaScript code

To ensure maintainability, I also created **internal documentation** for the backend system, describing:

- The structure of the script and logic flow
- The format of the JSON response
- How to add or modify indicators in the future

This enabled future contributors or team members to build upon or migrate the setup to more advanced backend platforms such as **Firebase**, **RESTful APIs**, or **cloud-hosted databases**, if required.



State	Electricity Dive	Viability of Ene	Dimension Sc	cnergy Security)	Energy Access	Affordability(En	Performance o	Dimension Scc	Rank 2024(Ene	Energy Resour	Decarbonisatio	Emissions and	Dimension Scc	Rank 202
Gujarat	8.05	8.32	16.37	1	3.61	6.8	9.06	19.47	1	6.45	4.45	5.87	16.77	
Kerala	7.27	8.89	16.16	2	4.23	6.36	8.3	18.89	2	3.08	5.58	6.28	14.94	
Rajasthan	8.2	7.53	15.73	3	3.92	6.45	8.09	18.46	3	3.04	5.91	5.5	14.45	
Himachal Prad	6.22	8.96	15.18	4	3.37	6.48	8.16	18.01	4	6.03	2.65	4.25	12.93	
Uttarakhand	6.88	8.08	14.96	5	3.52	6.19	8.15	17.86	5	2.96	4.82	5.06	12.84	
Andhra Prades	6.44	8.49	14.93	6	4.69	6.33	6.8	17.82	6	2.69	5.12	4.33	12.14	
Goa	5.7	9	14.7	7	3.02	6.37	8.16	17.55	7	1.65	5.18	4.7	11.53	
Haryana	6.38	8.3	14.68	8	4.56	6.88	6.07	17.51	8	3.38	2.92	5.19	11.49	
Tamil Nadu	7.37	7.19	14.56	9	3.6	5.47	8.1	17.17	9	2.58	3.93	4.79	11.3	
Telangana	6.53	7.98	14.51	10	3.86	4.71	8.43	17	10	3.23	4.07	3.95	11.25	
Madhya Prade:	5.34	8.98	14.32	11	5.02	9.54	2.23	16.79	11	2.63	3.61	4.61	10.85	
Punjab	6.29	7.75	14.04	12	3.55	7.58	5.43	16.56	12	2.26	3.74	4.84	10.84	
Karnataka	6.54	7.12	13.66	13	3.25	6.59	6.65	16.49	13	2.88	5.81	2.07	10.76	
Odisha	4.98	8.41	13.39	14	3.84	5.6	6.92	16.36	14	4.5	2.28	3.04	9.82	
Sikkim	6.52	6.79	13.31	15	3.98	4.35	7.82	16.15	15	4.64	1.45	3.64	9.73	
Mizoram	6.77	6.14	12.91	16	5.29	4.6	5.93	15.82	16	6.45	2.02	1.18	9.65	
Assam	4.92	7.97	12.89	17	3.84	7.2	4.71	15.75	17	4.41	2.4	2.67	9.48	
West Bengal	3.24	9.4	12.64	18	3.75	5.16	6.29	15.2	18	4.94	1.8	2.56	9.3	
Arunachal Prac	8.2	4.27	12.47	19	3.56	5.03	6.45	15.04	19	3.38	2.6	2.71	8.69	
Chhattisgarh	3.94	8.41	12.35	20	0.93	6.67	6.97	14.57	20	4.78	1.92	1.78	8.48	
Maharashtra	6.38	5.8	12.18	21	3.97	4.1	6.29	14.36	21	2.3	3.21	2.11	7.62	
Manipur	4.98	6.8	11.78	22	4.01	5.58	4.72	14.31	22	4.71	0.46	2.31	7.48	
Uttar Pradesh	4.63	5.98	10.61	23	3.48	5.45	5.11	14.04	23	3.88	2.6	0.96	7.44	
Nagaland	5.45	5.09	10.54	24	3.28	3.09	7.31	13.68	24	4.09	1.27	1.76	7.12	
Meghalaya	4.49	5.32	9.81	25	3.62	4.35	5.3	13.27	25	3.57	1.28	1.93	6.78	
Bihar	2.4	7.2	9.6	26	3	5.08	4.73	12.81	26	1.78	1.23	2.63	5.64	
Tripura	3.05	6.37	9.42	27	2.56	5.82	3.38	11.76	27	1.47	2.01	[illegible]	[illegible]	

Figure38 Google Sheet API

CHAPTER 16: CONTENT DEPLOYMENT USING WORDPRESS

16.1 Deployment Objective and Context

The final stage of the Trilemma Index project involved **deploying the interactive webpage** on the official website of **WEC India**, which is managed using **WordPress**, a widely adopted Content Management System (CMS). WordPress offers a user-friendly interface, extensive customization options, and a rich plugin ecosystem, making it ideal for publishing both static and dynamic web content.

However, deploying a **custom-coded, data-driven web application** onto an existing WordPress site required careful adaptation to ensure full functionality, compatibility, and maintainability within the platform's infrastructure.

16.2 Preparing Web Assets for Integration

To begin the deployment process, I **packaged all project files**—HTML, CSS, JavaScript, and image assets—into a consolidated and optimized format. This included:

- Minifying code to reduce file size and improve performance
- Cleaning unnecessary scripts and debugging logs
- Ensuring external libraries (e.g., Highcharts.js, Google Fonts, and CDNs) were **correctly referenced**
- Testing dynamic components locally to confirm **responsive behavior and functional integrity**

This preparatory step ensured that once uploaded to the server, the application would retain its interactivity while remaining **visually consistent** with the existing WordPress theme and design.

16.3 Custom Shortcode Integration

To integrate the web page into WordPress, I leveraged **shortcodes**, a native WordPress feature that enables the inclusion of **custom code blocks within content areas**. I created a custom shortcode function by modifying the site's functions.php file, allowing me to reference the full HTML content using a simple shortcode such as:

- php
- CopyEdit
- [trilemma_index_page]

This approach provided a clean, maintainable solution for embedding the interactive dashboard on any desired WordPress page. It also allowed for **easy future updates** without needing to modify the core content structure of the site.

16.4 Testing, Responsiveness, and Compatibility

After embedding the content, I conducted **comprehensive testing** across different browsers—**Google Chrome, Mozilla Firefox, Microsoft Edge, and Safari**—as well as across multiple device types (desktops, tablets, and smartphones). Key focus areas included:

- Verifying **responsive layout restructuring** on smaller screens
- Testing **dropdown menus, charts, and tooltips** for accurate behavior
- Ensuring **real-time data fetches** from the backend API were functioning correctly

- Confirming **load performance** and script execution did not conflict with existing WordPress plugins or theme settings

Adjustments were made where needed to improve mobile usability and visual alignment

CHAPTER 18: CONCLUSION

This project presented a unique opportunity to translate a globally recognized energy assessment framework into a **functioning digital tool tailored to India's federal structure**. By breaking down the World Energy Trilemma Index into state-level indicators, the project addressed the challenge of regional disparity in energy access, infrastructure, and sustainability, making the index more actionable at the subnational level.

The development process covered the **entire life cycle of a digital public tool**—from data collection and normalization to interactive front-end development, real-time backend integration, and final deployment within a content management system. Using technologies such as **Google Sheets, Apps Script, Highcharts.js**, and **WordPress**, the project demonstrated how low-cost, flexible platforms can be leveraged to create **robust, real-time energy visualization dashboards**.

Beyond its technical achievement, the project also emphasized **usability, scalability, and policy relevance**. It provides a foundation not only for tracking India's progress toward a balanced energy system but also for adapting similar tools in other developing regions. This hands-on experience enhanced my ability to think across disciplines—bridging energy systems knowledge, digital development, and user-oriented design.

REFERENCES

[1] N. Ghosh, M. Rajeshwor, and H. Narayan, “Economy’s response to Rainfall and Economic factors”.

[2] “IMD forecasts above-normal monsoon rainfall: Here’s what it means,” LiveMint. Accessed: Jul. 02, 2025. [Online]. Available: https://www.livemint.com/news/imd-monsoon-rainfall-forecast-long-period-average-rural-economy-kharif-crops-11748354197323.html

[3] Spencer, Thomas; Awasthy, Aayushi, “Analysing and Projecting Indian Electricity Demand to 2030.” The Energy and Resources Institute (TERI).

[4] B. Barik, S. Ghosh, A. S. Sahana, A. Pathak, and M. Sekhar, “Water–food–energy nexus with changing agricultural scenarios in India during recent decades,” *Hydrol. Earth Syst. Sci.*, vol. 21, no. 6, pp. 3041–3060, Jun. 2017, doi: 10.5194/hess-21-3041-2017.

[5] “Analysis on electrical energy consumption of agricultural sector in Indian context,” vol. 4, no. 2, 2009.

[6] A. K. Sikka, “Conservation Agriculture: Towards Managing the Water-Energy- Food Nexus in India,” *Conserv. Agric.*, vol. 21.

[7] D. S. Pai, M. Rajeevan, O. P. Sreejith, B. Mukhopadhyay, and N. S. Satbha, “Development of a new high spatial resolution (0.25° × 0.25°) long period (1901-2010) daily gridded rainfall data set over India and its comparison with existing data sets over the region,” *MAUSAM*, vol. 65, no. 1, pp. 1–18, Jan. 2014, doi: 10.54302/mausam.v65i1.851.

[8] Y. Chuang, “Climate variability, rainfall shocks, and farmers’ income diversification in India,” *Econ. Lett.*, vol. 174, pp. 55–61, Jan. 2019, doi: 10.1016/j.econlet.2018.10.015.

[9] “India | Electricity: Consumption: Utilities: Number of Consumers | CEIC.” Accessed: Jul. 03, 2025. [Online]. Available: https://www.ceicdata.com/en/india/electricity-consumption-utilities-number-of-consumers

[10] www.ETCFO.com, “India to revamp power demand forecasting for better planning - ETCFO,” ETCFO.com. Accessed: Jul. 03, 2025. [Online]. Available: https://cfo.economictimes.indiatimes.com/news/india-to-revamp-power-demand-forecasting-for-better-planning/114964260

[11] “Early monsoon cools June power demand target - The Economic Times.” Accessed: Jul. 03, 2025. [Online]. Available: https://economictimes.indiatimes.com/industry/energy/power/early-monsoon-cools-june-power-demand-target/articleshow/121631177.cms

[12] Bhandari, Laveesh; Dwivedi, Aasheerwad, “On India’s Electricity Consumption.” Centre for Social and Economic Progress (CSEP), Oct. 2023.

[13] N. Kholod, M. Evans, Z. Khan, M. Hejazi, and V. Chaturvedi, “Water-energy-food nexus in India: A critical review,” *Energy Clim. Change*, vol. 2, p. 100060, Dec. 2021, doi: 10.1016/j.egycc.2021.100060.

[14] The World Bank; NITI Aayog, “Energy-Water-Agriculture Nexus: Grow Solar, Save Water, Double the Farm Income.” The World Bank and NITI Aayog, Dec. 18, 2019.

[15] “Frontiers | Water-energy-food-ecosystem nexus in India—A review of relevant studies, policies, and programmes.” Accessed: Jul. 03, 2025. [Online]. Available: https://www.frontiersin.org/journals/water/articles/10.3389/frwa.2023.1128198/full

[16] “India’s Groundwater Crisis,” Drishti IAS. Accessed: Jul. 03, 2025. [Online]. Available: https://www.drishtiias.com/daily-updates/daily-news-editorials/india-s-groundwater-crisis

[17] A. Kapoor and M. Anand, “Addressing Groundwater Depletion Crisis in India: Institutionalizing Rights and Technological Innovations”.

[18] Central Ground Water Board Department of Water Resources, River Development & Ganga Rejuvenation, “National Compilation on DYNAMIC GROUND WATER RESOURCES OF INDIA, 2023.” Ministry of Jal Shakti Government of India, Sep. 2023.

[19] "Groundwater crises threaten the poor in India | SIWI - Leading expert in water governance." Accessed: Jul. 03, 2025. [Online]. Available: https://siwi.org/latest/groundwater-crises-threaten-the-poor-in-india/
[20] "A History of India's Groundwater Crisis - Grewal - 2024 - Groundwater - Wiley Online Library." Accessed: Jul. 03, 2025. [Online]. Available: https://ngwa.onlinelibrary.wiley.com/doi/10.1111/gwat.13371
[21] A. Prakash, N. Kumar, A. Chhatre, S. Thakkar, and A. Dar, "Navigating India's groundwater crisis: legal and institutional perspectives on regulation and conservation," *Water Policy*, vol. 26, no. 8, pp. 835–855, Jul. 2024, doi: 10.2166/wp.2024.123.
[22] G. V. Reddy, L. J. Aitha, Ch. Poojitha, A. N. Shreya, D. K. Reddy, and G. Sai. Meghana, "Electricity Consumption Prediction Using Machine Learning," *E3S Web Conf.*, vol. 391, p. 01048, 2023, doi: 10.1051/e3sconf/202339101048.
[23] M. Sharma, N. Mittal, A. Mishra, and A. Gupta, "An efficient approach for load forecasting in agricultural sector using machine learning," *E-Prime - Adv. Electr. Eng. Electron. Energy*, vol. 6, p. 100337, Dec. 2023, doi: 10.1016/j.prime.2023.100337.
[24] DR. B. RAVI PRASAD1 , MR. DUDEKULA SIDDAIAH2 , PROF. TS. DR. YOUSEF A.BAKER EL-EBIARY3 , DR. S. NAVEEN KUMAR4, DR. K SELVAKUMAR5, "FORECASTING ELECTRICITY CONSUMPTION THROUGH A FUSION OF HYBRID RANDOM FOREST REGRESSION AND LINEAR REGRESSION MODELS UTILIZING SMART METER DATA." Journal of Theoretical and Applied Information Technology.
[25] S. Gupta, R. Jasrasaria, S. Kalonia, and A. Khurana, "Farmer's Guide: Crop Prediction using Random Forest Regression," vol. 6, no. 12, 2023.
[26] M. P. Ahire, S. Katkar, R. Barhate, and S. Asodekar, "Smart City Energy Prediction Using Random Forest," vol. 13, no. 3, 2025.
[27] "Climate Monitoring and Prediction Group." Accessed: Jul. 04, 2025. [Online]. Available: https://imdpune.gov.in/cmpg/Griddata/Rainfall_25_NetCDF.html
[28] M. Rajeevan, Jyoti Bhate, J.D. Kale & B. Lal, "DEVELOPMENT OF A HIGH RESOLUTION DAILY GRIDDED RAINFALL DATA FOR THE INDIAN REGION." National Climate Centre India Meteorological Department Pune 411 005, India., 2005.
[29] C. M. and S. Khaiyum, "Data Collection, Pre-Processing, and Visualization of IMD Gridded Climate Data for Spatio-Temporal Exploratory Analysis:," in *Proceedings of the 1st International Conference on Intelligent and Sustainable Power and Energy Systems*, Bangalore, India: SCITEPRESS - Science and Technology Publications, 2023, pp. 136–143. doi: 10.5220/0012528900003808.
[30] C. E. Authority, "थर्मल पावर स्टेशनों के लिए मासिक/वार्षिक पर्यावरण डेटा प्रारूप," Central Electricity Authority. Accessed: Jul. 04, 2025. [Online]. Available: https://cea.nic.in/थर्मल-पावर-स्टेशनों-के-लि/
[31] "GADM." Accessed: Jul. 04, 2025. [Online]. Available: https://gadm.org/data.html
[32] "India - Subnational Administrative Boundaries | Humanitarian Dataset | HDX." Accessed: Jul. 04, 2025. [Online]. Available: https://data.humdata.org/dataset/geoboundaries-admin-boundaries-for-india
[33] J. Acharya, "Download India Administrative Boundary Shapefiles - States, Districts, Pincodes and more." Accessed: Jul. 04, 2025. [Online]. Available: https://www.igismap.com/download-india-administrative-boundary-shapefiles-states-districts-sub-districts-pincodes-constituencies/
[34] "Home | India Meteorological Department." Accessed: Jul. 04, 2025. [Online]. Available: https://mausam.imd.gov.in/
[35] "Perform, Achieve and Trade (PAT) | BUREAU OF ENERGY EFFICIENCY, Government of India, Ministry of Power." Accessed: Jul. 04, 2025. [Online]. Available: https://beeindia.gov.in/en/programmes/perform-achieve-and-trade-pat

[36] "What is Correlation Analysis? Definition, Types & How to Measure." Accessed: Jul. 04, 2025. [Online]. Available: https://www.questionpro.com/features/correlation-analysis.html

Annexture

STATE WISE ELECTRICITY DATA											
Column1	2014	2015	2016	2017	2018	2019	2020	2021	2022	2023	2024
A&N	240	240	240	302	348	348	348	339	343	376	411
AP	72128	49939	53628	56700	63930	64476	61750	67522	71367	79423	78615
Arunachal	664	600	716	788	843	781	706	839	908	994	1030
Assam	8360	8677	9082	9179	9604	9769	9956	10852	11321	12289	12779
Bihar	18121	22973	25908	26280	30361	31430	32846	36220	39278	41328	44162
Chandigarh	1605	1613	1654	1623	1569	1677	1532	1602	1777	1748	1985
Chattisgarh	20722	24917	23960	25702	26017	29638	29281	31794	37321	39785	42315
Delhi	29050	28261	30888	29814	32183	33124	29275	30898	35022	34878	38648
DVC	18123	18465	12312	21199	21530	22136	20840	23502	25950	26597	26274
Goa	4153	4780	4740	4054	4315	4261	4062	4465	4622	4841	5324
Gujarat	93983	101173	103477	108340	116704	114016	108623	122836	136149	148086	150897
Haryana	45844	46798	48947	50168	53730	54998	51902	55539	60919	62505	70481
Himachal Pradesh	8872	8748	8830	9274	9737	10463	9823	11926	12627	12672	13463
India	1064773	1089786	1141988	1191607	1156336	1163689	1045774	1365434.37	3875208	1607055	1683257
J&K	16184	16344	17362	18344	19430	19825	19608	20159	19722	19913	20115
Jharkhand	7581	7566	7982	7784	8744	8818	8956	10870	12906	14534	15024
Karnataka	63236	63413	66597	66748	72025	71498	69541	71232	74284	90491	91697
Kerala	22411	22870	24338	24664	25284	25940	25114	26239	27746	30022	31661
Lakshadweep	48	48	48	48	48	48	52	54	61	63	64
Maharashtra	136599	137073	140285	146622	158689	152576	147503	169327	183529	202967	201866
Manipur	668	805	777	830	915	926	965	998	1026	996	1071
Meghalaya	1871	1886	1739	1626	1896	2109	2022	2205	2266	2262	2117
Mizoram	455	455	504	497	649	654	733	656	644	663	704
MP	53693	56390	66219	69013	74169	74875	80913	85705	90494	103078	102491
Nagaland	645	752	762	788	864	836	836	839	914.3887	923	931
Odisha	26550	26327	26789	31704	28783	29912	28622	36222	42894	40990	42671
Puducherry	2396	2394	2559	2605	2745	2863	2612	2879	3022	3353	3551
Punjab	48552	49095	52705	53992	55446	57025	56552	62457	68321	68731	76979
Rajasthan	65494	65651	68534	70100	77895	81525	83139	88576	99786	104906	113716
Sikkim	414	382	460	486	520	523	542	595	602	527	578
Tamil Nadu	95155	95606	105051	104168	109463	108634	101977	107974	113932	122821	130012
Telangana	52456	50368	50817	57897	67065	65984	65004	72121	74106	84618	85991
Tripura	1222	1224	1407	1588	1884	1677	1474	1555	1553.2761	1676	1951
Uttar Pradesh	103032	104579	106578	118554	120961	122746	121277	127589	144119	146397	164300
Uttarakhand	12344	12820	13060	13300	12979	14488	13514	15151	15918	15402	16842
WB	46699	46510	48913	49204	52312	54010	49743	54645	58945	66798	69953

State-Wise Electricity Data

RAINFALL												
State	2014	2015	2016	2017	2018	2019	2020	2021	2022	2023	2024	NORMAL
daman and Nicol	0	0	0	0	0	0	0	0	0	0	0	0
Andhra Pradesh	23.44760523	33.96736943	25.384	29.604	22.21	30.297	41.795	39.138	34.9	28.575	34.59	30.9318
runachal Prades	74.92852021	85.60232204	107.67	106.58	74.475	83.083	93.29	66.662	95.397	69.764	70.372	85.7456
Assam	60.19374795	64.86713833	69.624	76.082	59.671	69.022	74.376	52.337	78.193	62.239	67.906	66.6605
Bihar	36.37386537	31.92719885	39.916	38.375	28.907	40.26	51.704	50.888	31.643	32.045	32.188	38.2038
Chandigarh	27.23151643	26.96152255	21.033	30.201	35.41	30.608	30.656	27.727	43.463	51.886	30.548	32.5176
Chhattisgarh	42.16507787	37.20463519	42.7	36.581	39.643	45.822	49.682	41.666	45.146	41.413	46.102	42.2022
lra and Nagar Ha	78.9342782	56.69784419	97.6	107.38	73.076	122.96	80.896	96.837	118.17	88.726	109.55	92.1271
Daman and Diu	0	0	0	0	0	0	0	0	0	0	0	0
Goa	126.914986	82.58995592	84.951	90.945	120.51	157.82	149.58	138.75	126.84	124.13	179.92	120.302
Gujarat	20.76438982	20.40861309	21.117	27.361	16.345	35.238	36.2	26.4	30.89	29.616	37.386	26.434
Haryana	12.16410803	16.12728876	13.475	15.218	17.286	13.129	19.021	23.614	23.01	19.886	17.042	17.293
limachal Prades	36.161913	42.0494865	33.872	40.269	38.122	40.078	32.146	34.632	36.29	42.208	33.836	37.5829
ammu and Kashm	26.31470514	30.21754641	17.259	25.686	20.538	30.304	21.082	19.43	20.182	18.959	17.195	22.9971
Jharkhand	37.86264038	37.19240533	40.549	40.198	31.609	38.197	41.313	47.953	33.187	36.233	42.548	38.4294
Karnataka	37.66946948	32.9294543	26.364	34.716	33.538	42.189	43.495	44.95	47.488	28.1	40.232	37.1438
Kerala	90.45354902	78.50754438	54.285	78.819	111.96	94.744	89.586	109.52	94.728	75.379	92.442	87.7981
Lakshadweep	0	0	0	0	0	0	0	0	0	0	0	0
Madhya Pradesh	29.60505362	32.59371247	38.984	25.628	28.243	46.1	35.896	35.951	42.988	35.587	39.87	35.1574
Maharashtra	32.40293088	28.64751476	40.086	35.546	31.354	47.8	42.443	44.814	44.657	34.798	44.682	38.2548
Manipur	37.25715421	45.80979481	53.694	67.883	45.209	33.544	44.774	30.569	45.793	30.627	46.559	43.5159
Meghalaya	134.3391589	162.4727867	128.7	152.51	95.604	132.97	179.58	124.81	184.53	109.15	132.46	140.466
Mizoram	54.51945572	66.04437941	75.963	126.18	84.805	59.88	56.51	60.121	65.297	55.796	59.375	70.5117
Nagaland	47.26448491	52.97437363	66.06	64.573	58.896	48.51	51.392	34.624	52.827	58.334	64.999	53.5455
NCT of Delhi	12.61973875	20.66451931	16.111	16.25	19.331	14.869	18.901	29.432	22.62	24.252	22.714	19.5049
Odisha	50.4837834	39.62191947	41.23	43.979	53.027	52.316	52.675	48.342	47.601	46.043	43.972	47.5319
Puducherry	0	0	0	0	0	0	0	0	0	0	0	0
Punjab	15.82717785	19.74047358	15.676	17.956	21.896	23.21	20.695	20.569	22.845	23.531	15.512	20.1944
Rajasthan	14.9267337	17.4189276	18.365	15.451	13.123	21.14	16.139	19.237	21.207	20.193	23.167	17.7201
Sikkim	87.23466696	104.8577309	101.3	102.38	107.52	107.31	116.05	105.72	122.22	100.79	105.84	105.538
Tamil Nadu	30.1974254	40.88094732	17.838	32.296	25.964	31.235	34.187	47.625	37.908	34.71	38.967	33.2841
Telangana	25.36782413	28.21046641	35.568	28.766	30.745	36.232	46.019	40.004	43.605	37.065	39.89	35.1582
Tripura	59.82898719	74.02942909	78.164	99.013	73.359	74.294	70.475	60.513	65.234	63.921	79.37	71.8831
Uttar Pradesh	21.11658513	19.95249226	26.603	22.388	26.949	28.434	27.604	33.696	27.329	25.015	27.197	25.9086
Uttarakhand	42.61157411	40.58498249	45.646	51.472	51.485	49.919	48.607	58.728	52.154	51.31	54.257	49.2518
West Bengal	51.81853947	61.58588898	57.696	61.834	48.983	56.443	67.171	67.869	51.543	53.464	62.237	57.8405

Rainfall Data

column	state	correlatio	corr_2_(apr-nov	agriculture_share_in_electricity_sales	agrr_corr_with_sgd	group	descriptio
29	Assam	-0.903715	-0.708562092	0.576	0.948095898	G5	
28	Karnataka	-0.895111	-0.842018987	32.87	0.301174252	G1	
27	Gujarat	-0.744949	-0.697952851	13.556	-0.041624475	G5	
26	AP	-0.721333	-0.732787675	20.124	0.961066715	G1	
25	Telangana	-0.648349	-0.831089724	35.864	0.875200691	G1	
24	Meghalaya	-0.603824	0.152798076	0.012	0.777761604	G2	
23	Punjab	-0.575173	-0.442991466	23.348	0.796184424	G1	
22	Odisha	-0.55798	-0.177846467	1.924	0.911613178	G6	
21	Chandigarl	-0.507578	-0.56075557	0.09	NOT AVAILABLE	G5	
20	Jammu An	-0.460532	0.427645832	3.61	0.8302328	G2	
19	Maharasht	-0.44438	0.376703085	24.598	0.138863671	G4	
18	Bihar	-0.443329	-0.577272324	6.17	0.863574386	G5	
17	Tamil	-0.429551	-0.164691573	13.67	0.881862331	G3	
16	Chattisgar	-0.339129	-0.267544792	18.654	0.920950903	G3	
15	Mizoram	-0.307804	-0.224895868	0.056	0.642276963	G6	
14	Kerala	-0.167741	-0.372442364	1.61	0.77888517	G6	
13	Arunachal	-0.166864	-0.258226253	0.046	0.891279218	G6	
12	Tripura	-0.159461	0.697223566	4.08	0.819549499	G2	
11	Uttar	-0.15928	-0.342120584	18.26	0.679627824	G3	
10	Jharkhand	-0.132938	-0.030322528	0.752	0.286920073	G6	
9	MP	-0.062644	0.132887604	37.592	NOT AVAILABLE	G4	
8	Rajasthan	-0.058242	-0.011426388	38.686	0.938267528	G4	
7	Nagaland	0.043055	0.147243964	0	0.125249336	G2	
6	Delhi	0.049659	-0.089308675	0.126	0.689985009	G6	
5	Sikkim	0.086515	-0.06059173	0	NOT AVAILABLE	G6	
4	Haryana	0.096315	0.109684121	21.336	-0.741301012	G4	
3	Himachal	0.144935	0.281311804	0.76	0.790800343	G2	
2	Manipur	0.206359	-0.561661679	0.692	0.732073867	G5	
1	Goa	0.209212	0.156021221	0.734	0.471593486	G2	
0	Uttarakhar	0.546257	0.184973386	2.67	-0.075015948	G2	

state-correlation-agriculture share-agriculture correlation with SGDP

STATES	AGRI	COMM	DOMESTIC	INDUSTRIAL	OTHERS	PUBLIC SERVICI	RAILWAYS
Punjab	-0.84773165	0.6211773	-0.67673785	0.58556803	-0.1654708	0.003812206	0.60719507
WB	-0.80598048	-0.0277505	-0.02323952	-0.54992606	0.62794465	-0.345307857	0.62738581
Meghalaya	-0.79605953	-0.1039194	0.85458407	0.64537651	-0.5495582	-0.75119834	#DIV/0!
Karnataka	-0.5826045	0.7753711	-0.89939119	0.59966337	-0.166058	-0.534346189	#DIV/0!
Bihar	-0.54737256	0.1687085	-0.1366286	-0.067168	-0.0635511	0.682050249	-0.01293596
Himachal F	-0.53970527	-0.1081066	-0.42632373	-0.06360976	-0.1548487	-0.509670954	#DIV/0!
Telangana	-0.46008621	-0.9129609	-0.14611306	-0.6278239	0.55445812	0.643315278	0.20818838
Haryana	-0.40078684	-0.2523585	-0.75073787	0.56487338	0.11311772	0.089927125	0.65367779
Andhra Pra	-0.31033047	-0.0066446	-0.05813292	0.3690265	#DIV/0!	#DIV/0!	#DIV/0!
Maharasht	-0.18494999	0.0278814	-0.15291893	-0.21555419	0.29206506	-0.661820472	0.47804823
Tamil Nadu	-0.05815928	-0.1728887	-0.22190316	0.1687941	-0.0169683	-0.152516037	#DIV/0!
Chhattisga	0.00995836	-0.0076405	-0.00162093	-0.13327387	0.84103081	0.358023465	-0.36199018
Uttar Prade	0.02168285	0.5790786	-0.23147798	0.59228829	-0.3411116	-0.222836605	0.34218758
Odisha	0.0300384	0.2481156	-0.34979233	0.31370134	-0.082504	-0.02740234	0.19845875
Delhi	0.10168542	0.8402181	-0.34337527	0.92637681	0.18185043	0.949917841	#DIV/0!
Tripura	0.11628974	0.3506421	0.8001797	0.15689542	0.89551966	0.084751867	#DIV/0!
UTTRAKHA	0.13332114	0.4795085	-0.57240521	0.91917727	0.17537603	-0.660822308	0.51965413
MP	0.18188095	0.7251226	0.34147885	0.13381389	0.39009645	0.245878731	#DIV/0!
Chandigarl	0.2173244	0.273008	0.24742015	0.34985599	-0.1255208	0.27232794	#DIV/0!
Mizoram	0.24765439	0.5574756	-0.32100544	0.60002854	-0.0560887	-0.143910819	#DIV/0!
Kerala	0.24922059	0.4262754	0.36673309	0.39045812	0.21602785	0.638115181	0.55802137
Gujarat	0.26245989	0.1093537	0.3535321	-0.23680306	0.27596113	0.342470181	-0.45642047
Rajasthan	0.37449576	0.6059188	-0.28370109	-0.33625171	0.21644874	-0.864485631	#DIV/0!
Goa	0.52299337	-0.1193557	0.10554497	-0.63606243	0.58814802	-0.060653129	#DIV/0!
Manipur	0.56946687	0.8338564	0.8155044	0.53815578	-0.4926181	0.665652261	#DIV/0!
Jharkhand	0.58144526	0.6015268	-0.10723627	-0.62428596	#DIV/0!	-0.721331333	-0.54795067
Assam	0.67494472	0.1370696	-0.3012873	-0.95930499	-0.2102301	0.405733184	#DIV/0!
Jammu & K	0.73184246	0.6872004	0.64168083	0.65884421	0.68947982	0.705360082	#DIV/0!
Arunachal	#DIV/0!	-0.6992427	-0.66446855	-0.50805004	0.58106454	0.460703233	#DIV/0!
Nagaland	#DIV/0!	0.0727234	0.14047492	0.17364393	0.08653459	0.145656678	#DIV/0!
Sikkim	#DIV/0!	-0.8588394	-0.49241714	-0.55100447	-0.3016567	-0.743507799	#DIV/0!

Sector-Wise Correlation Between Electricity Sales and Rainfall

RF EL PRED > A.PY > ...

```python
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
from sklearn.ensemble import RandomForestRegressor
from sklearn.model_selection import train_test_split
from sklearn.preprocessing import LabelEncoder
from sklearn.metrics import mean_squared_error, r2_score

# === STEP 1: Load and clean the dataset ===
df = pd.read_excel(r"C:\Users\Sanskriti Goyal\RF  EL PRED\forecast_center_cleaned.xlsx")
df.columns = df.columns.astype(str).str.strip().str.replace(" ", "_").str.replace("\xa0", "")

# === STEP 2: Define features and target ===
features = ["Region", "Forecast_LPA_Center"]
target = "Electricity_MU"
df = df.dropna(subset=features + [target])

# === STEP 3: Encode 'Region' column using IMD regions ===
all_regions = ["Central India", "South Peninsular", "Northwest India", "Northeast India"]
le = LabelEncoder()
le.fit(all_regions)
df["Region_Encoded"] = le.transform(df["Region"])

# === STEP 4: Prepare features and labels ===
X = df[["Region_Encoded", "Forecast_LPA_Center"]]
y = df[target]
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2, random_state=42)

# === STEP 5: Train the model ===
model = RandomForestRegressor(n_estimators=100, random_state=42)
model.fit(X_train, y_train)

# === STEP 6: Evaluate ===
y_pred_train = model.predict(X_train)
y_pred_test = model.predict(X_test)
rmse_train = np.sqrt(mean_squared_error(y_train, y_pred_train))
r2_train = r2_score(y_train, y_pred_train)
rmse_test = np.sqrt(mean_squared_error(y_test, y_pred_test))
r2_test = r2_score(y_test, y_pred_test)

print("\n✅ Model A Evaluation:")
print(f"Train Set - RMSE: {rmse_train:.2f}, R²: {r2_train:.3f}")
print(f"Test Set  - RMSE: {rmse_test:.2f}, R²: {r2_test:.3f}")

# === STEP 7: Feature Importance ===
importances = model.feature_importances_
feature_names = X_train.columns
plt.figure(figsize=(6, 4))
plt.barh(feature_names, importances, color='skyblue')
plt.xlabel("Importance")
plt.title("Feature Importance (Model A)")
plt.tight_layout()
plt.show()

# === STEP 8: Predict for 2025 ===
future_forecast = pd.DataFrame({
    "Region": ["Central India", "South Peninsular", "Northwest India", "Northeast India"],
    "Forecast_LPA_Center": [108, 110, 100, 92]
})
future_forecast["Region_Encoded"] = le.transform(future_forecast["Region"])
X_future = future_forecast[["Region_Encoded", "Forecast_LPA_Center"]]
future_forecast["Predicted_Electricity_MU"] = model.predict(X_future)

print("\n🔮 Forecast-Based Electricity Demand Predictions for 2025:")
print(future_forecast[["Region", "Forecast_LPA_Center", "Predicted_Electricity_MU"]])
```

Model A

```python
import pandas as pd
import numpy as np
from sklearn.ensemble import RandomForestRegressor
from sklearn.model_selection import train_test_split
from sklearn.preprocessing import LabelEncoder
from sklearn.metrics import r2_score, mean_squared_error

# === Load and clean ===
df = pd.read_excel(r"C:\Users\Sanskriti Goyal\RF  EL PRED\forecast_center_cleaned.xlsx")
df.columns = df.columns.astype(str).str.strip().str.replace(" ", "_").str.replace("\xa0", "")
df["State"] = df["State"].astype(str).str.strip()

# Drop missing values
df = df.dropna(subset=[
    "State", "Year", "Electricity_MU", "Forecast_LPA_Center",
    "Rainfall_mm"
])

# Optional: Normalize rainfall
df["Rainfall_mm_scaled"] = df["Rainfall_mm"] / df["Rainfall_mm"].max()

# Encode state
le_state = LabelEncoder()
df["State_Encoded"] = le_state.fit_transform(df["State"])

# Features and target
features = [
    "State_Encoded", "Forecast_LPA_Center", "Rainfall_mm_scaled",

]
target = "Electricity_MU"

X = df[features]
y = df[target]

# Train-test split
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2, random_state=42)
model_b = RandomForestRegressor(n_estimators=150, max_depth=10, random_state=42)
model_b.fit(X_train, y_train)

# Evaluation
print("\n✅ Model B Evaluation:")
print(f"Train R²: {r2_score(y_train, model_b.predict(X_train)):.3f}")
print(f"Test  R²: {r2_score(y_test, model_b.predict(X_test)):.3f}")


# ⚠️ Make sure this comes after you've trained `model_b` and encoded `State`

# Example prediction data for 2025
data_2025 = pd.DataFrame({
    "State": ["Gujarat", "Tamil Nadu", "Uttar Pradesh", "Punjab"],
    "Forecast_LPA_Center": [100, 110, 96, 92],          # from forecast report
    "Rainfall_mm": [450, 720, 400, 380],                # assumed/actual rainfall


})

# Normalize Rainfall_mm
data_2025["Rainfall_mm_scaled"] = data_2025["Rainfall_mm"] / df["Rainfall_mm"].max()

# Encode State
data_2025["State_Encoded"] = le_state.transform(data_2025["State"])

# Select features
X_2025 = data_2025[[
    "State_Encoded", "Forecast_LPA_Center", "Rainfall_mm_scaled",

]]

# Predict
data_2025["Predicted_Electricity_MU"] = model_b.predict(X_2025)

# View result
print("\n📊 Predicted 2025 Electricity Demand (MU):")
print(data_2025[["State", "Forecast_LPA_Center", "Predicted_Electricity_MU"]])

import matplotlib.pyplot as plt

importances = model_b.feature_importances_
feature_names = X_train.columns

plt.figure(figsize=(6,4))
plt.barh(feature_names, importances, color='skyblue')
plt.xlabel("Importance")
plt.title("Feature Importance (Model B)")
plt.tight_layout()
```

Model B

```python
import pandas as pd
import numpy as np
from sklearn.ensemble import RandomForestRegressor
from sklearn.model_selection import train_test_split
from sklearn.preprocessing import LabelEncoder
from sklearn.metrics import mean_squared_error, r2_score
import matplotlib.pyplot as plt

# === STEP 1: Load the preprocessed combined dataset ===
df = pd.read_excel("monthly_rainfall_electricity_combined.xlsx")

# === STEP 2: Clean and encode ===
df["State"] = df["State"].astype(str).str.strip()
df["Month"] = df["Month"].astype(str).str.upper()

df_model = df.dropna(subset=["State", "Month", "Electricity_MU", "Rainfall_mm"])

# Encode state and month
le_state = LabelEncoder()
df["State_Encoded"] = le_state.fit_transform(df["State"])

le_month = LabelEncoder()
df["Month_Encoded"] = le_month.fit_transform(df["Month"])

# === STEP 3: Define features and target ===
features = ["State_Encoded", "Month_Encoded", "Year", "Rainfall_mm"]
target = "Electricity_MU"

# Select features and target, drop rows with any NaNs
model_data = df[features + [target]].dropna()
X = model_data[features]
y = model_data[target]

# === STEP 4: Train-test split ===
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2, random_state=42)

# === STEP 5: Train the model ===
model_c = RandomForestRegressor(n_estimators=150, max_depth=15, random_state=42)
model_c.fit(X_train, y_train)

# === STEP 6: Evaluation ===
y_train_pred = model_c.predict(X_train)
y_test_pred = model_c.predict(X_test)

r2_train = r2_score(y_train, y_train_pred)
r2_test = r2_score(y_test, y_test_pred)

print("\n✅ Model C Evaluation:")
print(f"Train R²: {r2_train:.3f}")
print(f"Test  R²: {r2_test:.3f}")

# === STEP 7 (Optional): Feature Importance ===
importances = model_c.feature_importances_
feature_labels = X.columns

plt.figure(figsize=(6, 4))
plt.barh(feature_labels, importances, color='skyblue')
plt.xlabel("Importance")
plt.title("Feature Importance (Model C)")
plt.tight_layout()
plt.show()
```

Model C

```python
import pandas as pd
import numpy as np
from sklearn.ensemble import RandomForestRegressor
from sklearn.model_selection import train_test_split
from sklearn.preprocessing import LabelEncoder, MinMaxScaler
from sklearn.metrics import mean_squared_error, r2_score
import matplotlib.pyplot as plt

# === STEP 1: Load dataset with LPA_Center and Region included ===
df = pd.read_excel("monthly_rainfall_electricity_LPA_added.xlsx")

# === STEP 2: Clean and convert ===
df["Electricity_MU"] = pd.to_numeric(df["Electricity_MU"], errors="coerce")
df["Rainfall_mm"] = pd.to_numeric(df["Rainfall_mm"], errors="coerce")
df["LPA_Center"] = pd.to_numeric(df["LPA_Center"], errors="coerce")
df = df.dropna(subset=["Electricity_MU", "Rainfall_mm", "LPA_Center", "Region", "State", "Month"])

# === STEP 3: Encode categorical columns ===
df["State"] = df["State"].astype(str).str.strip()
df["Region"] = df["Region"].astype(str).str.strip()
df["Month"] = df["Month"].astype(str).str.upper()

le_state = LabelEncoder()
le_region = LabelEncoder()
le_month = LabelEncoder()

df["State_Encoded"] = le_state.fit_transform(df["State"])
df["Region_Encoded"] = le_region.fit_transform(df["Region"])
df["Month_Encoded"] = le_month.fit_transform(df["Month"])

# === STEP 4: Normalize continuous features ===
scaler = MinMaxScaler()
df["Rainfall_mm_scaled"] = scaler.fit_transform(df[["Rainfall_mm"]])
df["LPA_Center_scaled"] = scaler.fit_transform(df[["LPA_Center"]])

# === STEP 5: Define features and target ===
features = [
    "State_Encoded",
    "Region_Encoded",
    "Month_Encoded",
    "Rainfall_mm_scaled",
    "LPA_Center_scaled"
]
target = "Electricity_MU"

X = df[features]
y = df[target]

# === STEP 6: Train-test split ===
X_train, X_test, y_train, y_test = train_test_split(
    X, y, test_size=0.2, random_state=42
)

# === STEP 7: Train the model ===
model_c = RandomForestRegressor(n_estimators=150, max_depth=15, random_state=42)
model_c.fit(X_train, y_train)

# === STEP 8: Evaluate ===
y_train_pred = model_c.predict(X_train)
y_test_pred = model_c.predict(X_test)

r2_train = r2_score(y_train, y_train_pred)
r2_test = r2_score(y_test, y_test_pred)

print("\n✅ Model d Evaluation (with Region and LPA_Center):")
print(f"Train R²: {r2_train:.3f}")
print(f"Test  R²: {r2_test:.3f}")

# === STEP 9: Feature Importance ===
importances = model_c.feature_importances_
feature_labels = X.columns

plt.figure(figsize=(7, 4))
plt.barh(feature_labels, importances, color='lightseagreen')
plt.xlabel("Importance")
plt.title("Feature Importance (Model D with Region + LPA)")
plt.tight_layout()
plt.show()
```

Model D