MASTER OF TECHNOLOGY
IN
GEOINFORMATICS

Submitted by

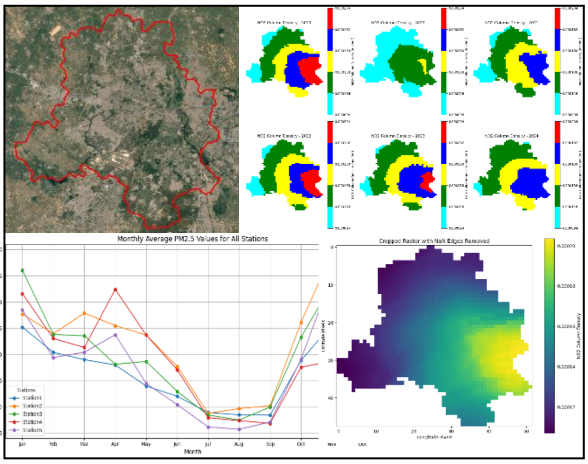
NEELAM SHANMUKHA SREE JAYA PRAKASH

(roll no: 231030052)

NATIONAL CENTRE OF GEODESY (NCG)

IIT KANPUR

2023-2025



Project 1:

Title: Analysis of NO_2 Column Number Density Over Delhi Using Sentinel-5P Data (2019–2024)

Objective

This Project aimed to analyse the temporal and spatial variations of NO_2 column number density over Delhi during the summer months (March, April, May and June) from 2019 to 2024. The key objectives were:

1. Downloading and processing Sentinel-5P OFFL NO_2 data.
2. Trimming the dataset to the Delhi state boundary.
3. Creating spatial plots for NO_2 variations across the years.
4. Developing a methodology to identify NO_2 hotspots.
5. Drawing conclusions based on the analysis.

Methodology

1. Data Collection

- **Source**: Sentinel-5P OFFL NO_2 data was downloaded using Google Earth Engine.
- **Region of Interest**: Delhi state boundary (28 N, 29 N, 76.5 E, 77.75 E).
- **Temporal Scope**: Summer months (March to June) for each year from 2019 to 2024.
- **Dataset Format**: The mean NO_2 column number density for each summer season was extracted and stored as NetCDF files.

Preprocessing: The dataset was clipped to the Delhi shapefile boundary to remove unwanted regions outside Delhi. The shapefile was processed using Python libraries like geopandas and rioxarray. The clipped GeoTIFF files were loaded into an Xarray dataset to facilitate analysis and visualization.

Visualization: Spatial maps for each year (2019–2024) were created to observe NO_2 distribution. The plots were color-coded with a consistent scale for easy comparison.

Hotspot Identification: A threshold-based approach was applied to identify NO_2 hotspots. The 90th percentile of NO_2 values was used as the threshold to classify pixels as hotspots. Binary hotspot maps were generated for each year, with 1 indicating hotspots and 0 indicating non-hotspots.

Results

The results displayed are:

1. Spatial Plots of NO_2 (2019–2024)
2. Hotspot Analysis
3. Percentage Change in NO_2 Levels
4. Hotspot Centroid Locations

1. Spatial Plots of NO_2 (2019–2024)

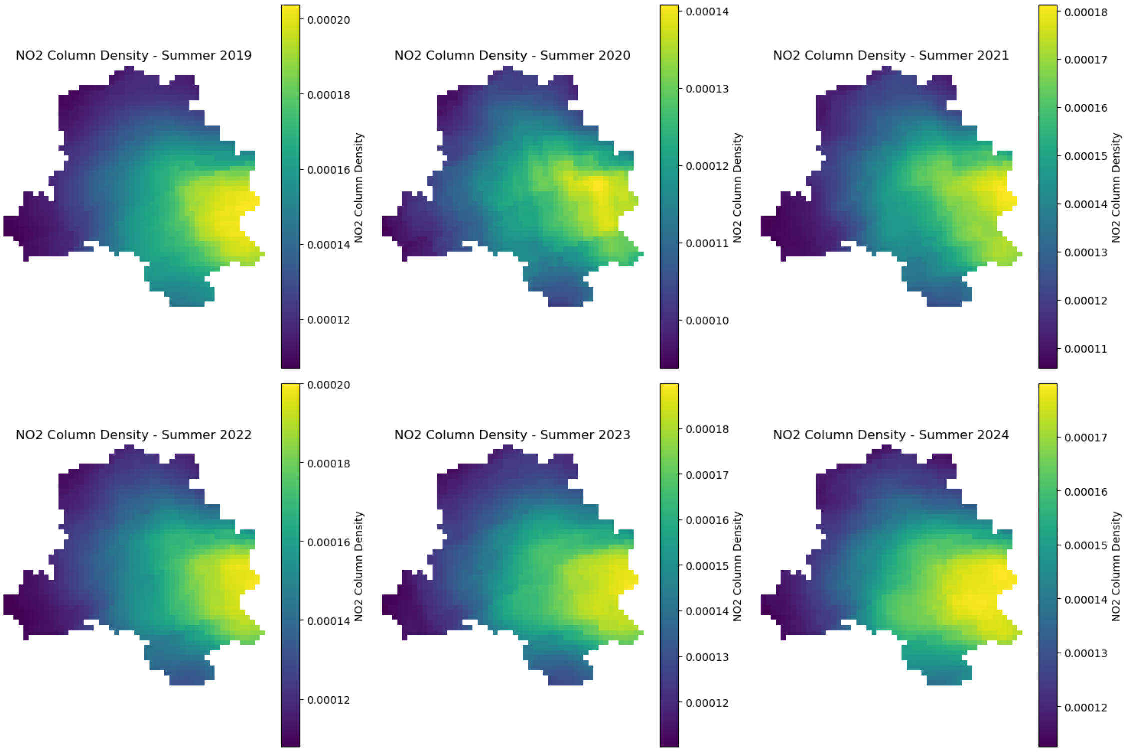


Fig: Spatial plots for 'NO2_column_number_density' over the years during summers.

The following observations were made from the spatial plots:

1. **2019**: Higher NO_2 levels were observed in southern and central Delhi, indicating significant pollution in these regions.
2. **2020**: A notable decline in NO_2 was observed due to the COVID-19 lockdowns, which reduced industrial and vehicular emissions.
3. **2021–2024**: A gradual recovery in NO_2 levels was observed, with some hotspots persisting over the years.

2. Hotspot Analysis

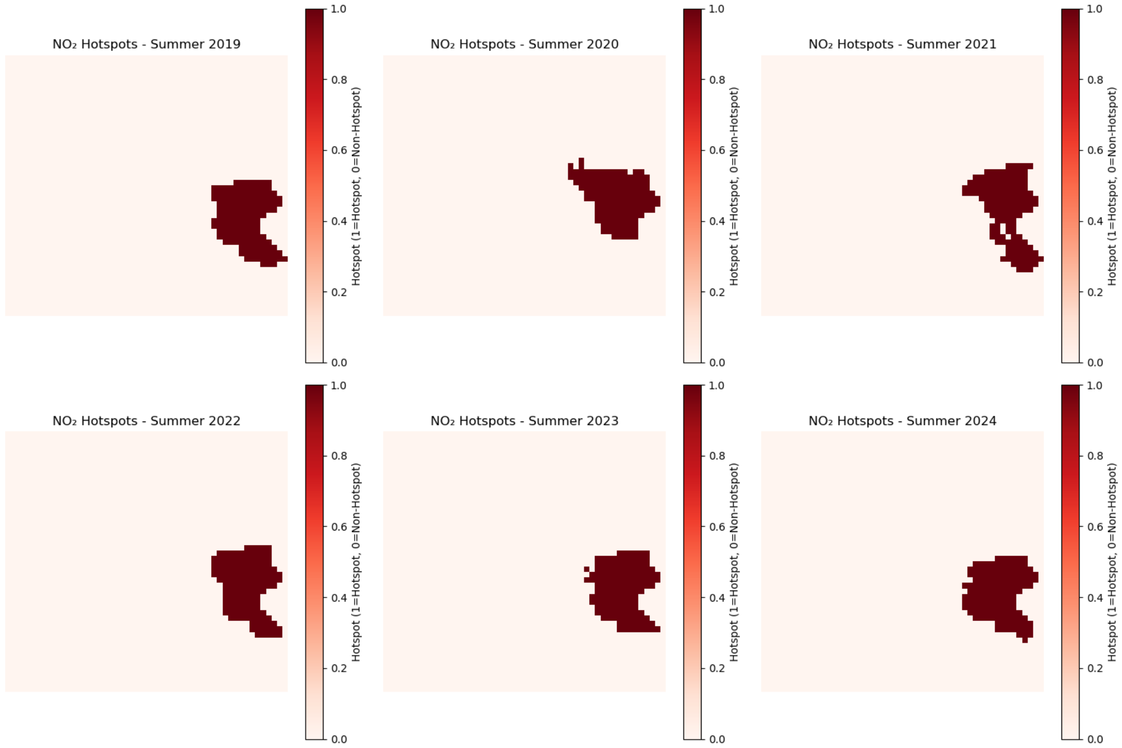


Fig: Hotspots plot for 'NO2_column_number_density' over the years during summers

The **90th percentile threshold** was selected to classify pixels with exceptionally high NO_2 values. Rasters were converted into binary maps, where 1 represented hotspot and 0 represented non-hotspots.

The centroid for each hotspot region was calculated using scipy.ndimage to track their spatial shift over the years. Hotspot maps were visualized using Python and exported to QGIS for detailed analysis with administrative boundaries.

Persistent hotspots were observed in industrial areas and high-traffic zones, particularly in south and west Delhi. Some hotspots showed a temporal shift, indicating changing sources of emissions.

2020–2021: Hotspot centroids were concentrated in south Delhi. 2023–2024: Centroids shifted slightly toward industrial zones.

3. Percentage Change in NO_2 Levels

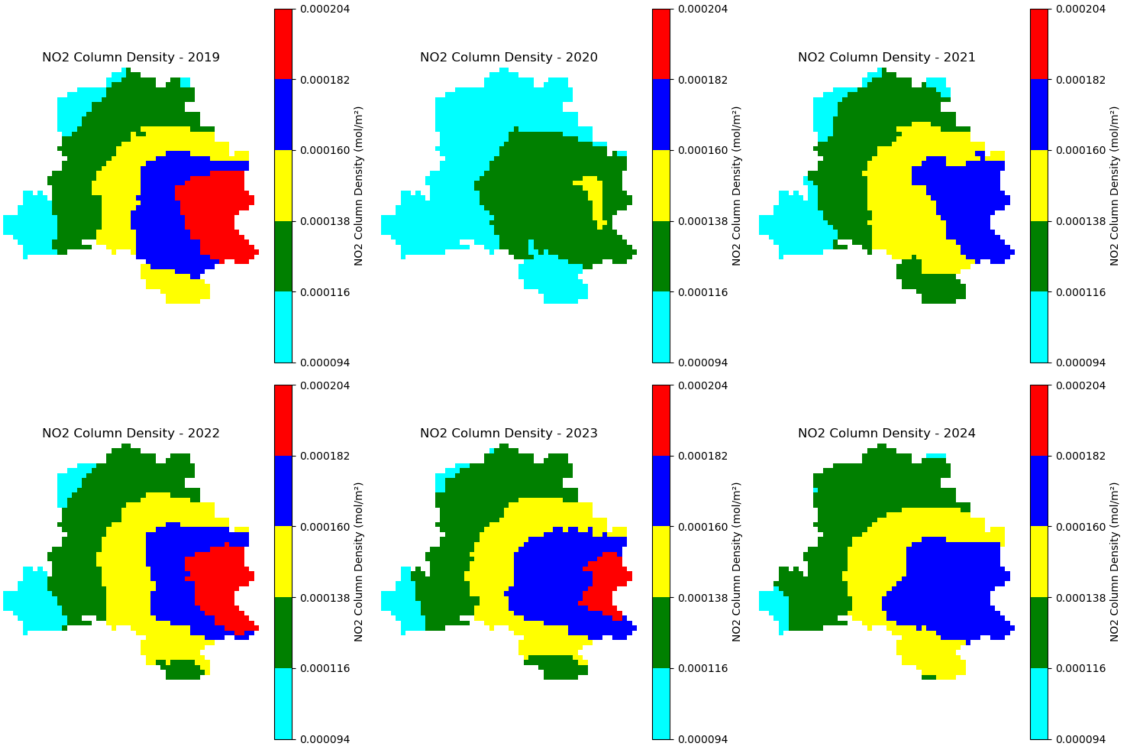


Fig: Change in NO_2 level over the years during summers

Year	% Change
2019 to 2020	-22.21%
2020 to 2021	+20.38%
2021 to 2022	+5.27%
2022 to 2023	+0.07%
2023 to 2024	-0.67%

Table: Percentage change in NO_2 level over the years during summers

Project 2:

Air Quality Analysis Report: Comparison of WUSTL PM2.5 and Observed PM2.5 Data

---

Objective

This report provides a comprehensive analysis of PM2.5 data obtained from satellite-derived datasets (WUSTL, 1x1 km resolution) and observed data from Continuous Ambient Air Quality Stations (AQS) in Delhi for the year 2022. The study involves interpolating satellite PM2.5 values to station coordinates, comparing them with observed values and calculating performance metrics such as Root Mean Square Error (RMSE), Mean Absolute Percentage Error (MAPE), Mean Absolute Error (MAE) and R^2. Diurnal, daily, weekly and monthly trends of PM2.5 are analysed across five stations and key insights are derived to evaluate the reliability of satellite data for air quality monitoring.

Methodology

The methodology followed for this analysis is outlined below:

Data Collection: Satellite-derived PM2.5 data (1x1 km resolution) for the year 2022 was downloaded from WUSTL. PM2.5 data for five Continuous Ambient Air Quality Stations (AQS) within Delhi was obtained from the respective website.

Data Processing: The coordinates of the five stations were identified. Satellite PM2.5 values were interpolated to the station coordinates for each month.

Performance Evaluation: Calculated performance metrics such as RMSE, MAE, MAPE and R^2 to evaluate the satellite data against the observed data.

Data Visualization: Created time series plots comparing WUSTL PM2.5 with observed PM2.5. Generated diurnal, daily, weekly and monthly plots for all five stations.

Analysis and Interpretation: Analysed the plots and performance metrics to derive insights. Identified stations with best and worst performance and explored possible reasons.

Results

This section includes the plots, performance metrics and insights derived from the analysis.

1. Time Series Plots
2. Performance Metric
3. Diurnal, Daily, Weekly and Monthly Trends
4. Insights

1. Time Series Plots: Comparing WUSTL PM2.5 data with observed PM2.5 for each station.

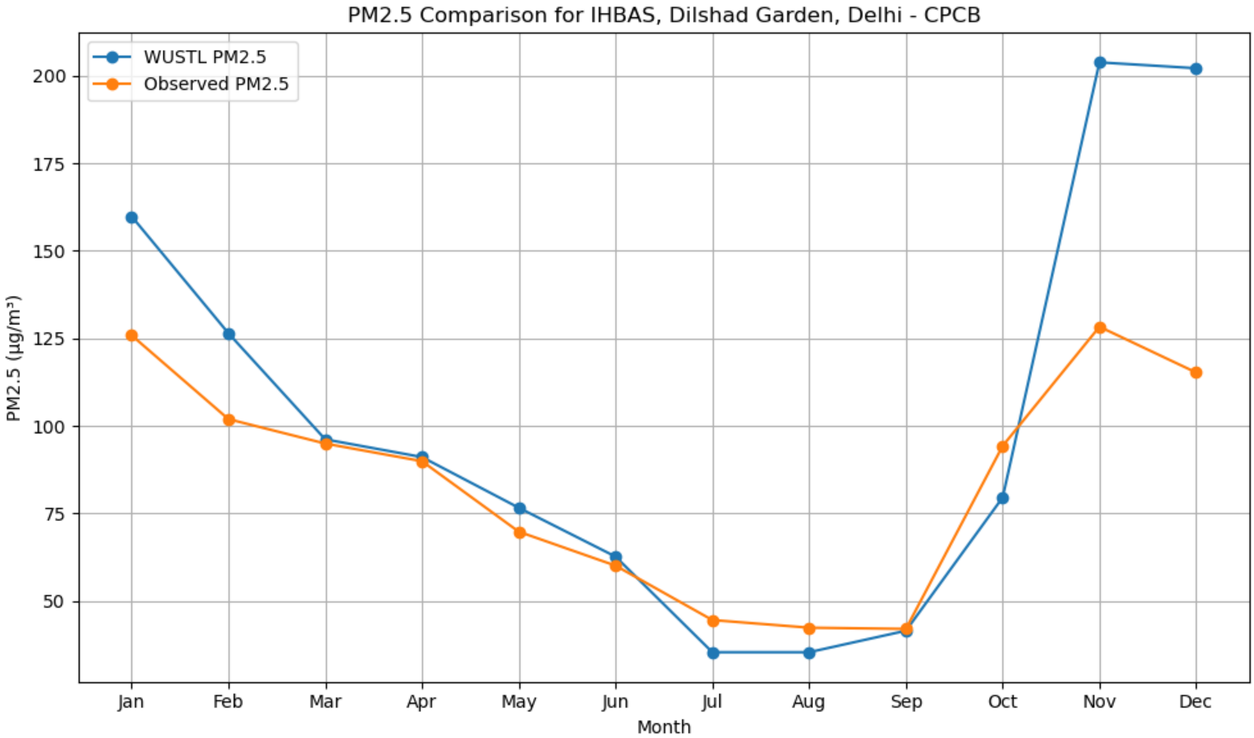


Fig: Time series plot for IHBAS station

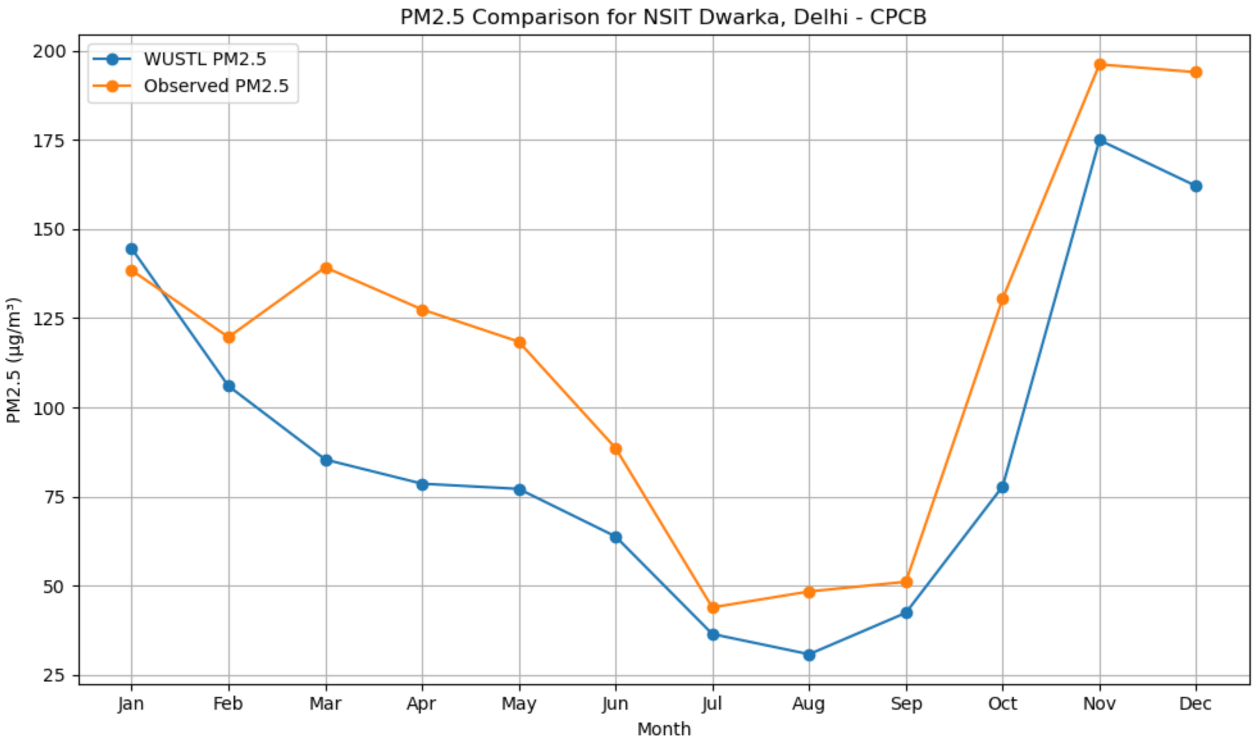


Fig: Time series plot for NSIT station

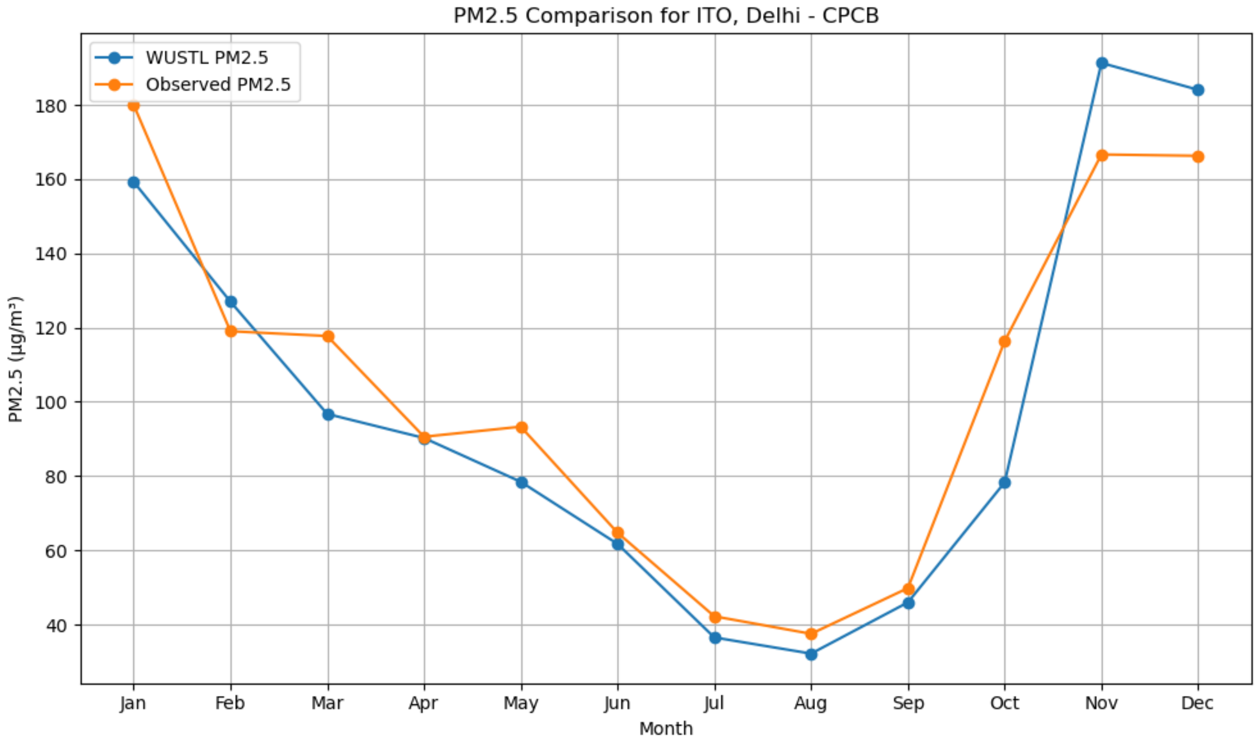


Fig: Time series plot for ITO station

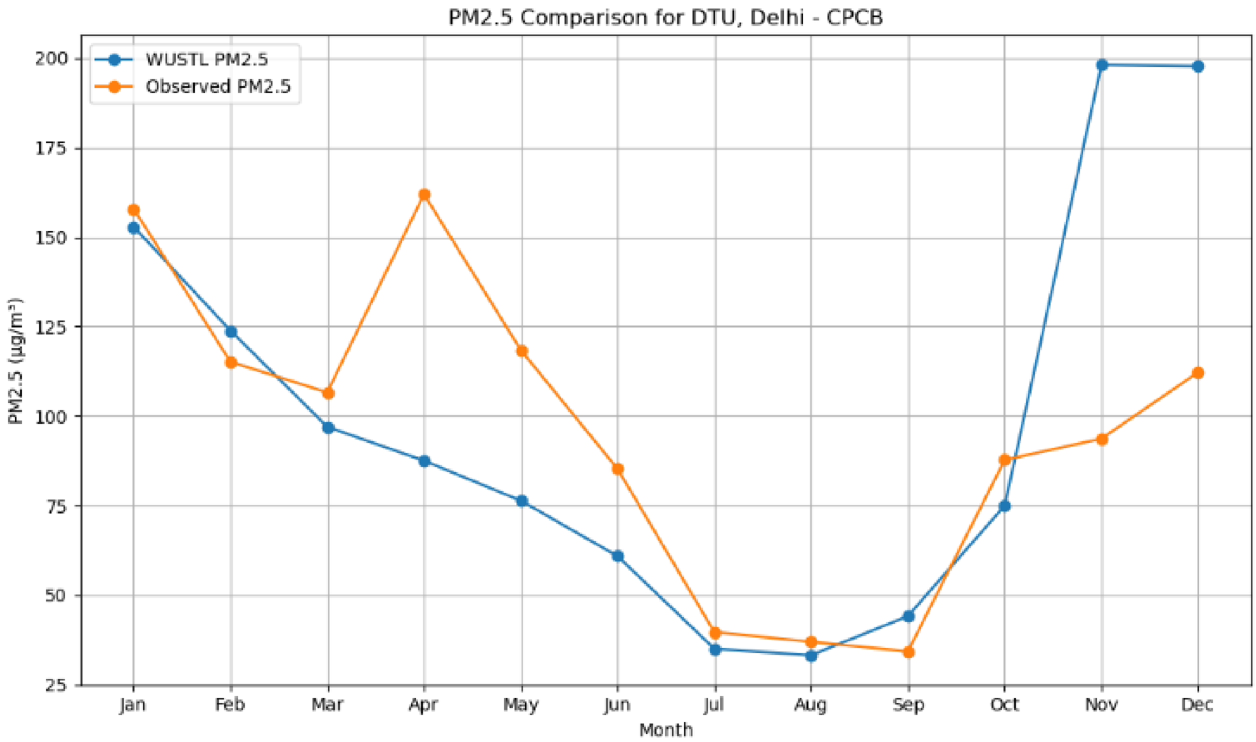


Fig: Time series plot for DTU station

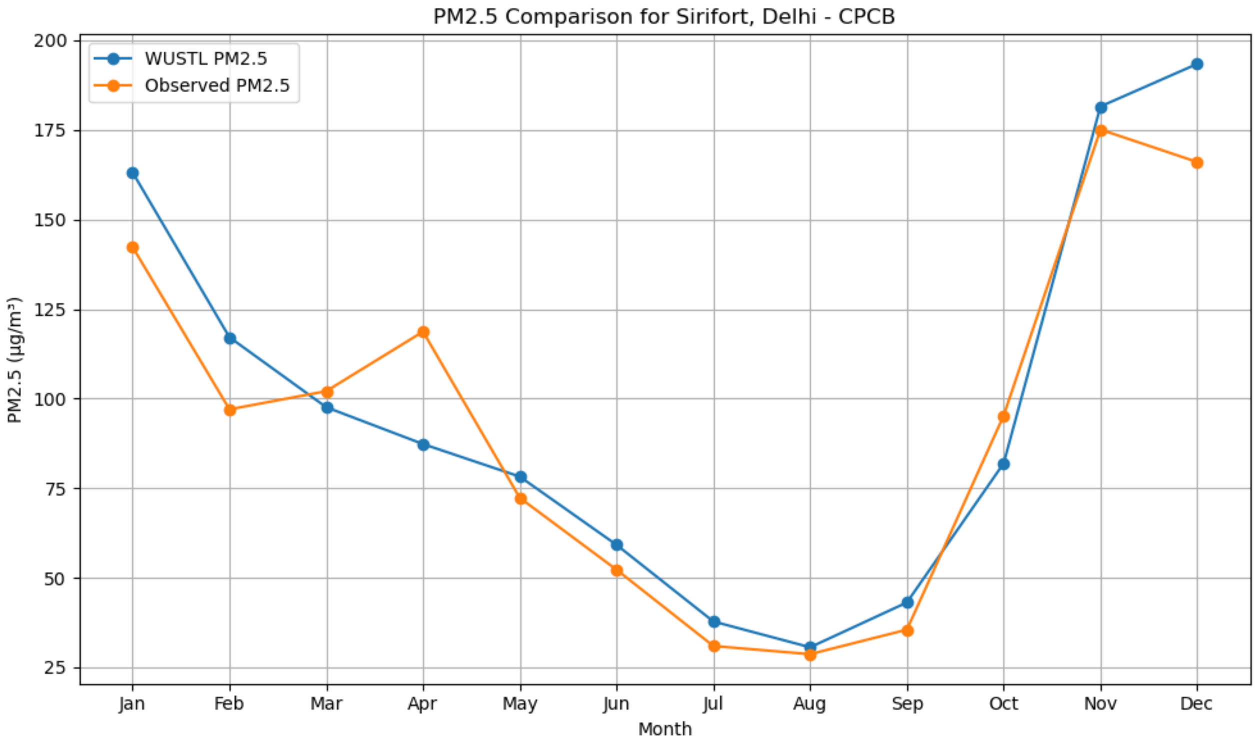


Fig: Time series plot for Siri fort station

Both WUSTL and observed PM2.5 data show similar seasonal trends, with peaks during winter (October–December) and troughs during the monsoon (July–August).

WUSTL data generally overestimates PM2.5 levels during cleaner months and underestimates them during peak pollution months (e.g., November–December).

Station 5 (Siri fort) shows the best alignment between WUSTL and observed data, while Station 4 (NSIT Dwarka) exhibits significant deviations, especially in winter.

The satellite-based model performs well during low-pollution periods but struggles to capture localized high-pollution events, particularly in areas like NSIT Dwarka and ITO.

These observations highlight the need for integrating satellite data with ground-based observations to improve accuracy in capturing localized pollution trends.

2. **Performance Metrics: Present the calculated RMSE, MAE, MAPE and R^2 values in tables or graphs**.

Station	RMSE	MAE	MAPE (%)	RÂ²
DTU Delhi - CPCB	47.065	32.081	31.873	-0.324
IHBAS Dilshad Garden Delhi - CPCB	35.831	22	21.318	-0.388
ITO Delhi - CPCB	17.351	13.600	12.551	0.866
NSIT Dwarka Delhi - CPCB	32.312	27.364	24.494	0.562
Siri fort Delhi - CPCB	15.826	12.771	14.347	0.895

Negative R^2 indicates that the WUSTL PM2.5 predictions are worse than a simple horizontal line at the mean of the observed PM2.5 values. The model is not capturing the relationship between WUSTL and observed PM2.5 values due to Systematic bias in satellite data for these locations.

An R^2 of ~0.56 indicates that the WUSTL PM2.5 predictions explain 56% of the variance in observed PM2.5 data. The relationship between WUSTL and observed PM2.5 is better but still leaves significant unexplained variance.

High R^2 values (close to 1) indicate that WUSTL PM2.5 data explain most of the variance in observed PM2.5 values. The satellite-derived data is capturing the temporal and spatial trends accurately for these stations.

3. **Diurnal, Daily, Weekly and Monthly Trends: Include the respective plots for all five stations.**

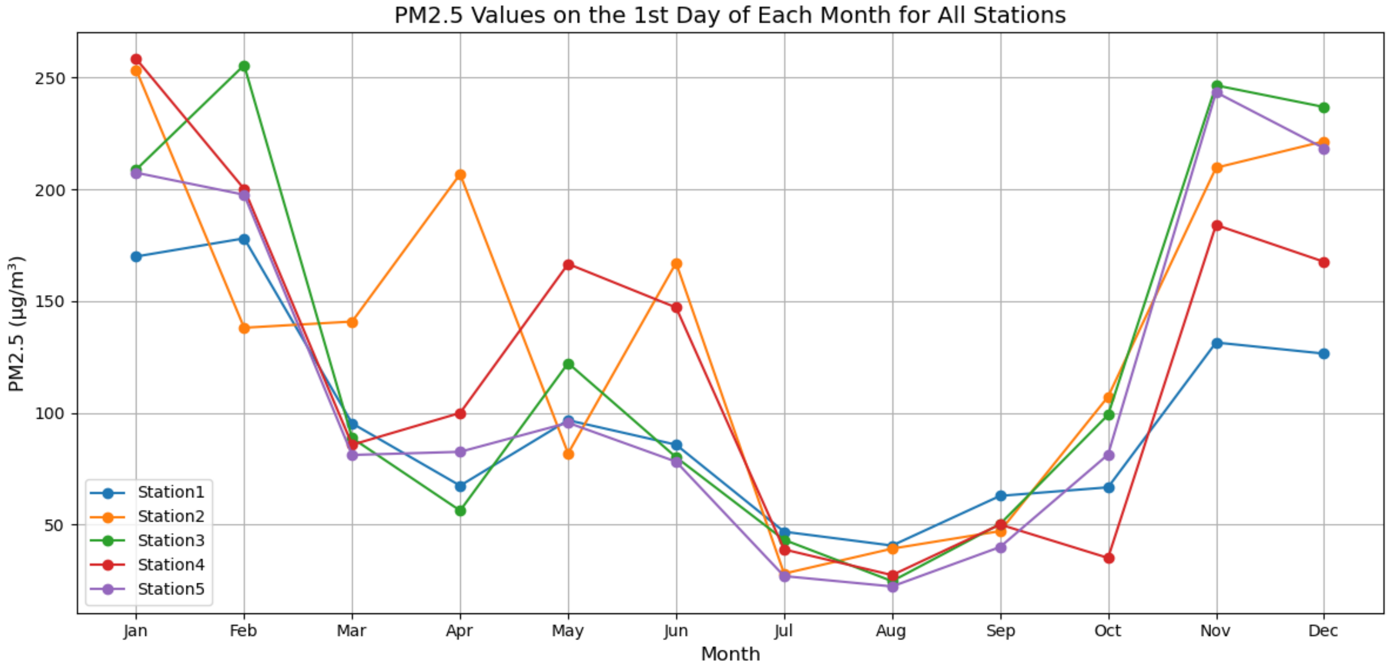


Fig: Daily Trend for all the stations over the 1st day of each month

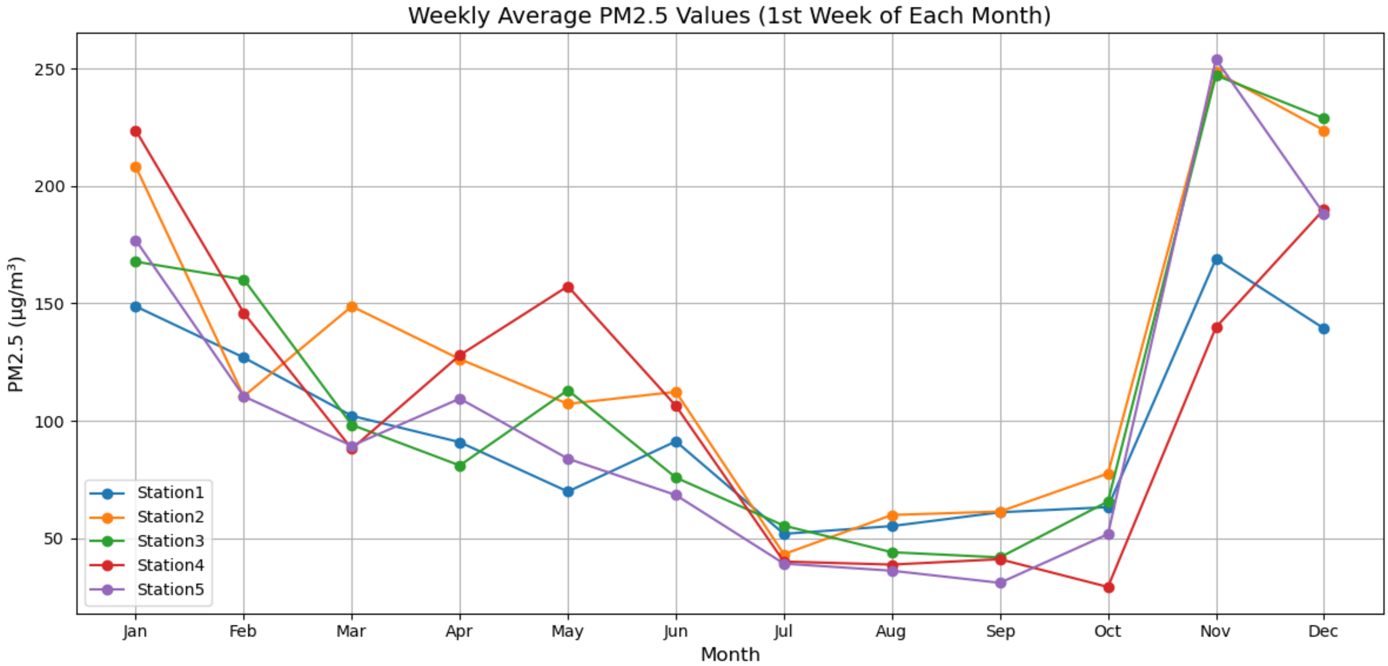


Fig: Weekly Trend for all the stations over the 1st week of each month

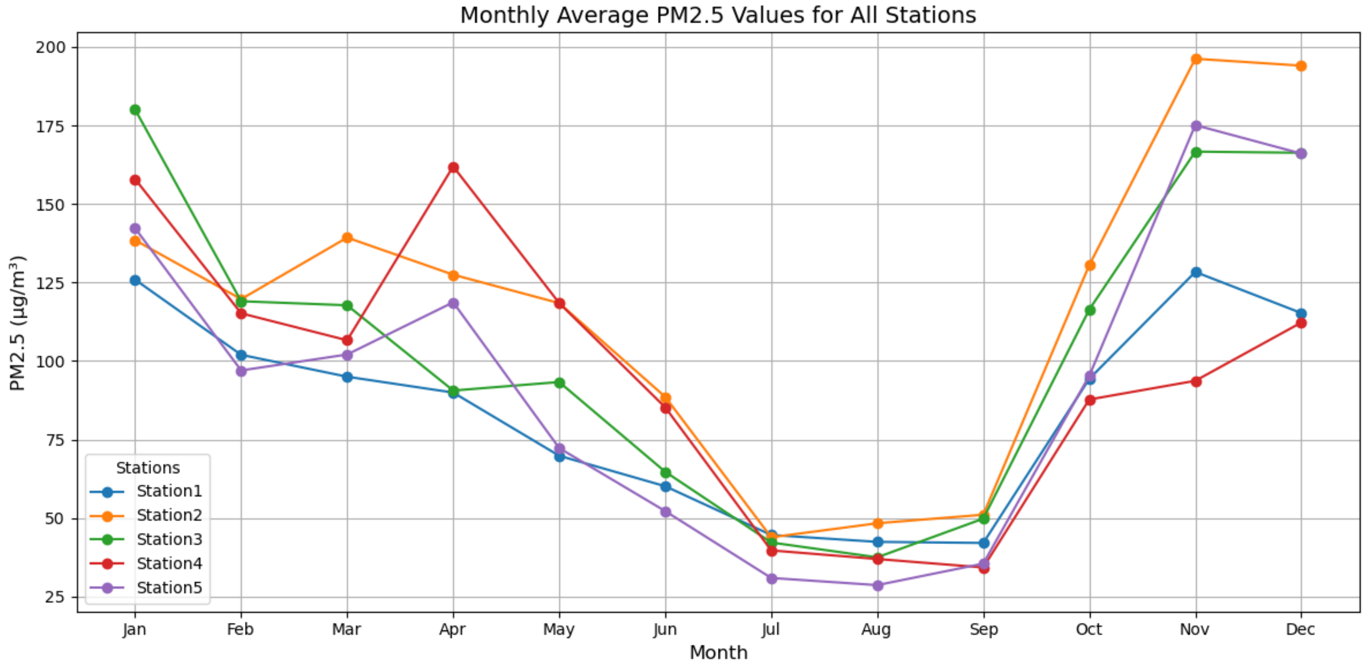


Fig: Monthly Trend for all the stations

Note:

"Station1": IHBAS_Dilshad_Garden_Delhi_CPCB

"Station2": NSIT_Dwarka_Delhi_CPCB

"Station3": ITO_Delhi_CPCB

"Station4": DTU_Delhi_CPCB

"Station5": Sirifort_Delhi_CPCB

Higher PM2.5 levels occur in the winter months (November–January) and Lower PM2.5 levels occur in the monsoon months (July–August). Starting in October, PM2.5 levels increase sharply, peaking around November–December.

From the plots, I have observed that in the daily plot, Station 4 and Station 3 show significant deviations compared to the other stations. In the weekly plot, Station 4 again shows the highest deviation from the others. Similarly, in the monthly plot, Station 4 exhibits noticeable deviations as well. This indicates that there might be some issues with the data from Station 4.

Station 2 (orange line) consistently reports higher PM2.5 levels compared to others, Station 5 (purple line) shows relatively lower values, indicating it may be in a cleaner or less densely populated area.

Station 2 and Station 4 often peak higher than others, indicating consistent exposure to elevated pollution levels. Station 1 and Station 5 tend to have lower PM2.5 levels overall, possibly due to less urban or industrial influence.

4. Insights that you derive from the above analyses

Station 4 (NSIT Dwarka, Delhi - CPCB) consistently records higher PM2.5 values in daily, weekly and monthly plots. Deviates significantly from other stations, indicating higher localized pollution sources.

Station 5 (Siri fort, Delhi - CPCB) consistently records lower PM2.5 levels compared to other stations.

Station 3 (ITO) reflects moderate pollution levels with reliable data.

Station 1 (DTU) and Station 2 (IHBAS) may have data inconsistencies or localized pollution spikes that satellite data struggles to capture.

Thank you,

NSS Jaya Prakash,

MailID: shanmukha23@iitk.ac.in,

IIT Kanpur,

Contact: 8919258152.