ISSN 1022-8594
JUJS

Jahangirnagar University Journal of Science
Vol. 46, No. 1, pp. 1—28

Spatiotemporal Analysis and Machine Learning-Based Profiling of Atmospheric Pollutants Using Sentinel-5P TROPOMI Satellite Data at Tier-Two Administrative Units in Bangladesh

Bipro Acharjee[1, 3], Monim Abdullah[2, 3], Mahmudul Hasan[2], Hritwik Roy[4], Toqi Tahmid Fahim[5], Mohammad Mizanur Rahman[2, 3]

[1]Department of Statistics, Shahjalal University of Science and Technology, Sylhet
[2]Department of Urban and Regional Planning, Jahangirnagar University, Savar, Dhaka-1342
[3]Geospatial Data Processing and Simulation Lab, Department of Urban and Regional Planning, Jahangirnagar University, Savar, Dhaka-1342
[4]Department of Statistics, MC College, Sylhet
[5]Department of Computer Science and Engineering, Southeast University, Tejgaon, Dhaka-1208

Abstract

Air pollution poses a critical environmental and public health threat in Bangladesh, yet a comprehensive, multi-pollutant analysis at a high administrative resolution has been lacking. This study bridges this gap by leveraging the high-resolution capabilities of Sentinel-5P TROPOMI satellite data and an unsupervised machine learning framework to conduct a district-level assessment of five key atmospheric pollutants—SO_2, NO_2, O_3, HCHO, and CH_4—from 2020 to 2024. Annual mean concentrations were extracted for all 64 districts using Google Earth Engine, followed by rigorous spatiotemporal trend analysis using the Mann-Kendall test. The core novelty of this research lies in the application of K-means clustering to classify districts based on their holistic, multi-annual pollutant profiles.

The analysis revealed distinct national trends: a significant increase in O_3 (54 districts) and CH_4 (61 districts), a volatile trend for SO_2, and a notable decrease in NO_2 in 21 urban districts. Critically, the K-means algorithm (k=5) segmented the country into five discrete pollution clusters, ranked from best to worst: 1) Cleanest (16 districts, e.g., coastal and hilly regions), 2) Moderate (7 districts, e.g., northeastern regions), 3) High O_3 & CH_4 (16 northern agricultural districts), 4) High Multi-Pollutant (23 southwestern and central industrial districts), and 5) Extreme Urban/Industrial (4 districts: Dhaka, Gazipur, Munshiganj, Narayanganj). This final cluster exhibited an extreme concentration of NO_2 and CH_4, over 50% higher than the national average.

The findings demonstrate that Bangladesh's air pollution landscape is not monolithic but a mosaic of distinct regional typologies, each driven by different dominant sources—from vehicular and industrial combustion in urban centers to agricultural emissions in the north. This study provides a novel, data-driven framework for "pollution zonation," moving beyond one-size-fits-all policies towards targeted, cluster-specific mitigation strategies. The methodology establishes a replicable model for evidence-based air quality management in data-scarce regions globally.

Keywords: Air pollution analysis, K-means clustering, Multi-pollutant assessment, District-level air quality mapping, Google Earth Engine (GEE)

*Author for correspondance e-mail: mrahman@juniv.edu
DOI : https://doi.org/10.59185/jos.v46i1.425

1. Introduction

Air pollution has emerged as one of the foremost environmental health risks globally. The World Health Organization (WHO) estimates that exposure to fine particulate matter ($PM_{2.5}$) causes around seven million premature deaths globally each year [1]. It is estimated that nine out of ten people breathe polluted air, underscoring the ubiquity of the problem [1]. Fine particles and associated gases penetrate deep into the lungs and bloodstream, contributing to cardiovascular and respiratory diseases, cancers, and neurological disorders [2]. Air pollution, therefore, represents an urgent public health and climate challenge that requires scientifically grounded, spatially resolved assessments.

South Asia is one of the region's most severely impacted by air pollution, and Bangladesh consistently ranks among the world's most polluted countries. Rapid urbanization, unregulated industrial expansion, and widespread use of biomass fuels have pushed ambient $PM_{2.5}$ levels far above the WHO guideline of 5 µg m^{-3}. Mean concentrations range from 60–100 µg m^{-3} across the country, and Dhaka experiences concentrations of 90–100 µg m^{-3}, particularly during the dry season [3]. Dhaka's $PM_{2.5}$ concentrations were 12–18 times higher than the WHO guideline between 2010 and 2019, and high-resolution measurements show that pollution and noise levels exceed both national and international standards [2]. This exposure has dire consequences: ambient and household air pollution contribute to more than 159,000 premature deaths annually in Bangladesh and result in over 2.5 billion days lived with illness [3]. The burden falls disproportionately on urban populations, but rural areas are increasingly affected due to agricultural burning, waste disposal, and power generation [3].

Previous studies, particularly those by Elaprolu (2024) and Rahman et al. (2025), have highlighted the dominance of vehicular emissions, industrial activities, and brick kilns in urban air pollution [4,5]. However, the existing studies primarily focus on single pollutants or use city-scale data, neglecting the broader national impact and the interdependence between multiple atmospheric pollutants. Despite increasing recognition of the crisis, Bangladesh's air quality literature remains dominated by single pollutant studies or city-scale analyses. Recent work by Basak et al. (2023) and Haque et al. (2022) has begun to address these gaps, focusing on NO_2 and $PM_{2.5}$ in major urban centers [6,7]. However, multi-pollutant interactions, particularly between methane (CH_4), formaldehyde (HCHO), and ozone (O_3), have not been adequately explored at a national scale.

This study addresses these gaps by utilizing Sentinel-5P TROPOMI satellite data to assess the spatiotemporal trends of five major pollutants (SO_2, NO_2, O_3, HCHO, and CH_4) across 64 districts in Bangladesh for the period 2020–2024. The Sentinel-5P satellite, equipped with the TROPOMI instrument, provides high-resolution measurements (7 km × 3.5 km) of atmospheric pollutants, offering a powerful tool for monitoring air quality globally [8]. Although Rahman et al. (2025) utilized Sentinel-5P time-series analysis for air pollutants in Dhaka, their focus was limited to a subset of pollutants (CO, NO_2, SO_2, O_3) and concentrated on urban centers [5]. Ultimately, there is a lack of comprehensive, multi-pollutant assessments that integrate methane (CH_4) and formaldehyde (HCHO)—important precursors for ozone—and provide complete spatial coverage across all districts. This gap limits the ability of policymakers to design coordinated mitigation strategies that address both urban and rural sources.

This study addresses the gap by conducting a district-level, multi-pollutant spatiotemporal analysis of sulphur dioxide (SO_2), nitrogen dioxide (NO_2), ozone (O_3), formaldehyde (HCHO) and methane (CH_4) across Bangladesh from 2020–2024. Using Sentinel-5P Tropomi data, we (1) characterize the spatial and temporal trends of these pollutants, (2) apply unsupervised clustering to classify districts into coherent pollution regimes, and (3) explore the underlying drivers and policy implications of the observed patterns. This research provides one of the first comprehensive portraits of Bangladesh's air pollution landscape and offers evidence-based insights for tailored mitigation strategies.

2. Data and Methods

This study employed a multi-stage analytical framework to assess the spatiotemporal patterns of key atmospheric pollutants across the districts of Bangladesh. The overall methodology, summarized in Figure 1, integrated satellite data retrieval, geospatial analysis, statistical trend detection, and unsupervised machine learning.

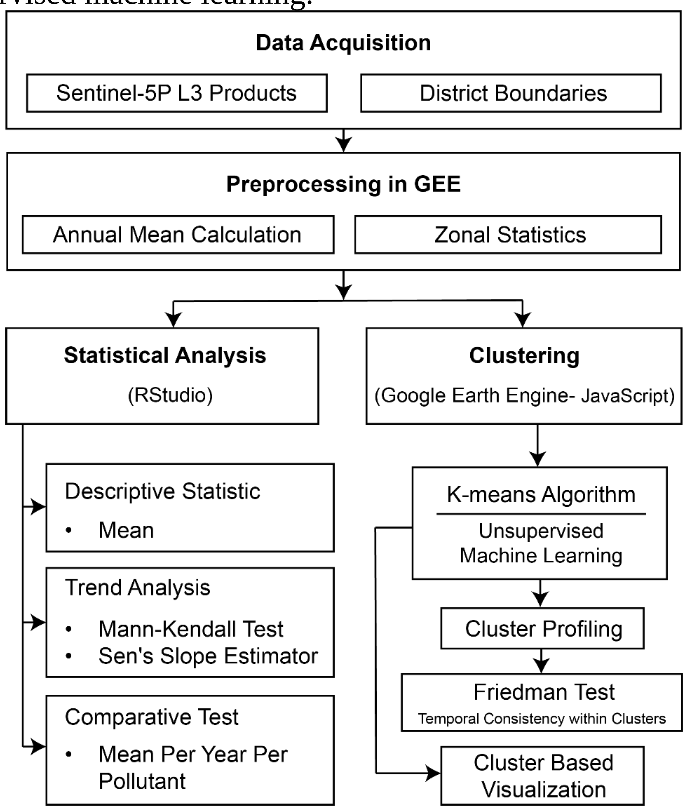

Fig 1 Methodological Workflow

2.1. Study Area

The study area encompasses the entirety of Bangladesh (Fig 2), a densely populated South Asian country (Fig 2-B) with a land area of approximately 147,570 km². Administratively, Bangladesh is divided into 8 divisions and 64 districts (zila). For this analysis, the district level (N=64) was chosen as the primary spatial unit because it represents a critical scale for policy implementation, environmental management, and public health intervention. The country exhibits diverse pollution sources, including intense vehicular emissions in metropolitan areas like Dhaka and

Chattogram, widespread industrial operations (e.g., garment, textiles, and leather), a high density of seasonal brick kilns, agricultural burning, and emissions from the household energy sector.

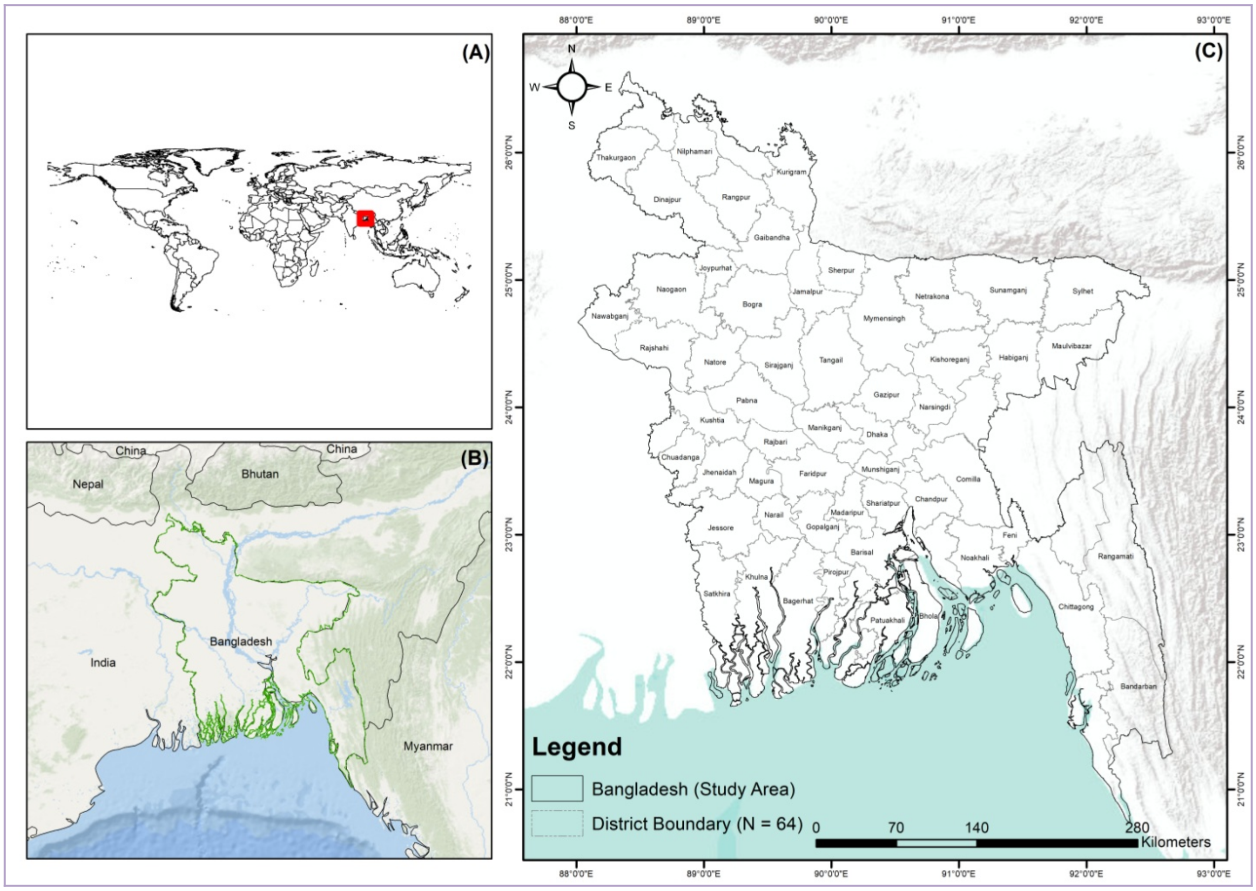

Fig 2 Study Area Map

2.2. Data Sources

2.2.1. Satellite-Based Pollutant Data

Atmospheric pollutant data were acquired from the Sentinel-5P TROPOMI (TROPOspheric Monitoring Instrument) sensor, specifically the Offline (OFFL) Level-3 (L3) products. These products provide global daily gridded data at a resolution of 0.01° x 0.01° (approximately 1 km² at the equator), which are pre-aggregated and quality-assured, making them suitable for regional trend analysis. We analyzed five major tropospheric pollutants over a five-year period from January 1, 2020, to December 31, 2024. The specific products and retrieved parameters are detailed in Table 1.

Table 1 Sentinel-5P OFFL L3 Data Products Used in the Study

Pollutant	Google Earth Engine ID	Band Name	Unit	Key Tropospheric Source
Nitrogen Dioxide (NO_2)	COPERNICUS/S5P/ OFFL/L3_NO2	NO2_column_nu mber_density	mol/m²	Fossil fuel combustion (vehicles, power plants)
Sulfur Dioxide (SO_2)	COPERNICUS/S5P/ OFFL/L3_SO2	SO2_column_nu mber_density	mol/m²	Coal burning, industrial processes, shipping
Ozone (O_3)	COPERNICUS/S5P/ OFFL/L3_O3	O3_column_num ber_density	mol/m²	Photochemical reactions of precursors (NOx, VOCs)
Formalde hyde (HCHO)	COPERNICUS/S5P/ OFFL/L3_HCHO	tropospheric_HC HO_column_nu mber_density	mol/m²	Incomplete combustion, biogenic isoprene oxidation
Methane (CH_4)	COPERNICUS/S5P/ OFFL/L3_CH4	CH4_column_vo lume_mixing_rat io_dry_air	ppb	Agriculture (livestock, rice paddies), waste, fossil fuels

2.2.2. Administrative Boundaries

The vector boundary file for the 64 districts of Bangladesh was obtained as a Fusion Table asset on Google Earth Engine. The attribute, which contains the district names, was used to identify and label the features.

2.3. Data Preprocessing and Extraction in Google Earth Engine

All data preprocessing and zonal statistic extraction were conducted within the Google Earth Engine (GEE) cloud computing platform. The workflow was as follows:

1. **Annual Mean Composite Generation:** For each pollutant and each year (2020–2024), an annual mean composite image was generated. This was done by filtering the respective ImageCollection by the calendar year, applying a cloud mask (where relevant and as defined in the product algorithms), and calculating the pixel-wise mean. Notably, data for the full year of 2019 was not available for all pollutants, and the 2025 year is still ongoing, so full-year data for 2025 was not considered. The study utilized annual mean composites, accounting for seasonal variations, cloud cover, and single-event anomalies. This approach ensures a robust metric for inter-annual and inter-district comparison while capturing the effects of seasonalities in the data.

2. **Zonal Statistics Extraction:** The annual mean value for each pollutant was extracted for every district using the reduceRegion() function. The geometry of each district feature

was used as the boundary, with the reducer set to ee.Reducer.mean(). The analysis was performed at the native scale of the L3 products (1000 meters) to preserve data integrity. This process resulted in a unique data record for each district containing 25 data points (5 pollutants × 5 years).

3. **Data Export:** The resulting feature collection, with districts as features and annual pollutant means as properties, was exported as a CSV file for subsequent statistical analysis. This analysis was performed using R programming language in RStudio.

2.4. Statistical Analysis

The extracted dataset was analyzed using a combination of descriptive, trend, and inferential statistics in R, utilizing RStudio for the analysis.

2.4.1. Descriptive Analysis

Initial analysis characterized the central tendency and variability of pollutant concentrations. We calculated the annual mean for each pollutant across all districts and also 5-year mean. This provided a national-level overview of pollutant baselines and their changes over time. To complement the summary statistics, density histograms (Fig 3) were generated for each pollutant (SO_2, NO_2, O_3, HCHO, CH_4) for all five study years (2020–2024). These histograms were constructed using the 64 district-level annual mean values for each pollutant, providing a visual representation of the distributional shape, spread, and skewness of pollutant concentrations across Bangladesh. The use of density normalization ensures that the area under each histogram equals one, allowing meaningful comparison across pollutants and between years independent of scale differences.

2.4.2. Trend Analysis

To identify monotonic trends in pollutant concentrations over the 5-year study period, the non-parametric Mann-Kendall (MK) test was applied. The MK test is robust against non-normally distributed data and missing values, which are common in environmental datasets. The test was performed:

- **At the National Level:** Using the annual national average for each pollutant.
- **At the District Level:** For each pollutant time series in each of the 64 districts.

The Sen's Slope estimator was used alongside the MK test to quantify the magnitude of any significant trends. A positive tau (τ) value from the MK test indicates an increasing trend, while a negative value indicates a decreasing trend.

2.4.3. Comparative Statistical Test (Spatial Heterogeneity Analysis)

To assess spatial heterogeneity of pollutant concentrations across the 64 districts, we analyzed the mean and standard deviation of each pollutant (SO_2, NO_2, O_3, HCHO, and CH_4) over the 2020–2024 period. The mean provided an overall indicator of pollution levels, while the standard deviation revealed variability within districts, helping us identify regions with fluctuating pollution levels. This approach effectively captures the spatial distribution and variability of pollutants across the districts.

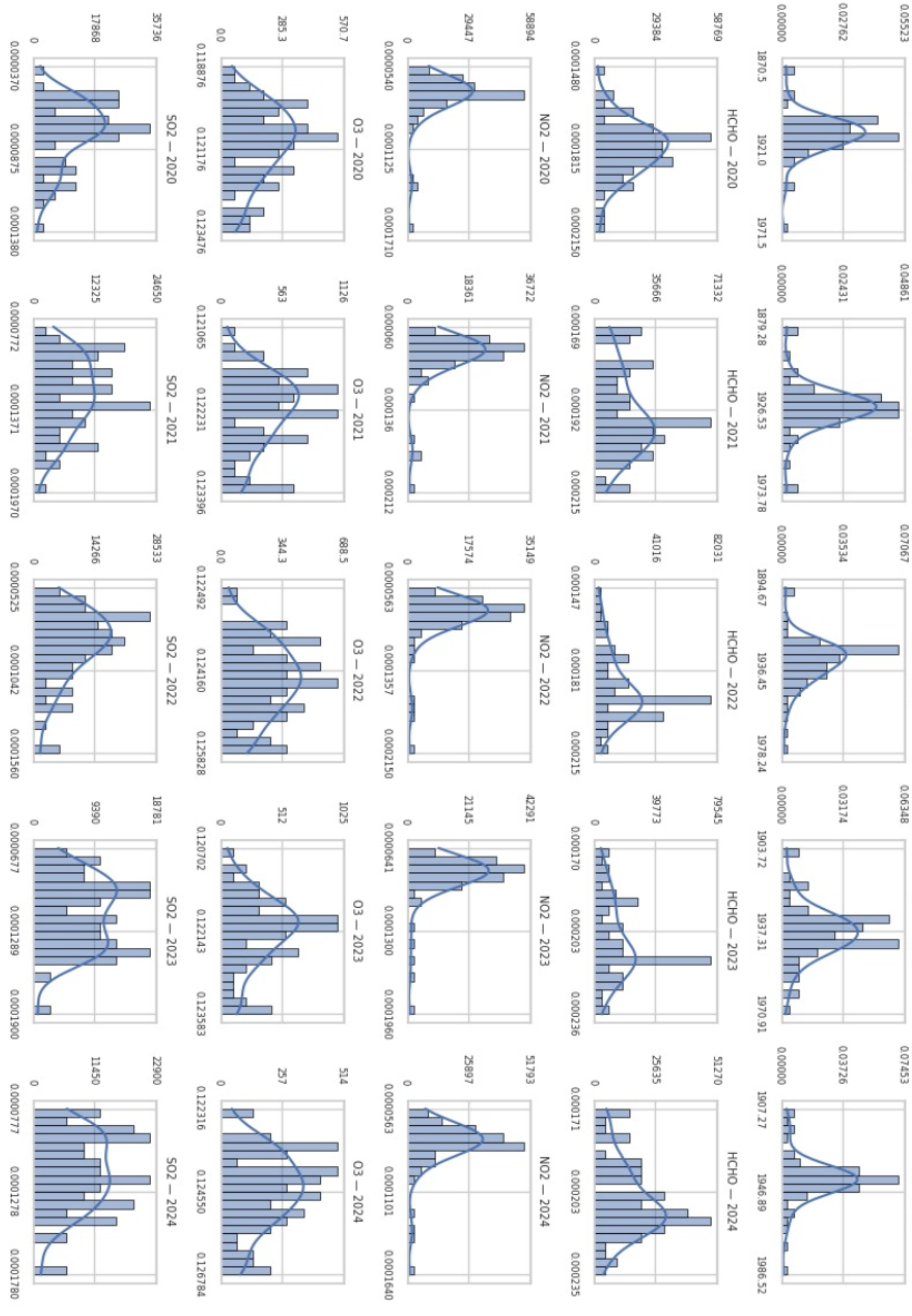

Fig 3 Density Histogram of CH4, HCHO, NO2, O3, SO2 for each year from 2020 to 2024

2.5. Cluster Analysis

To objectively classify districts into groups with similar multi-pollutant profiles over the entire study period, we employed K-means clustering, an unsupervised machine learning algorithm.

1. **Feature Space:** The input matrix for clustering consisted of 64 rows (districts) and 25 columns (features). The features were the annual mean concentrations for SO_2, NO_2, O_3, HCHO, and CH_4 for the years 2020 to 2024.
2. **Data Standardization:** Prior to clustering, all features were standardized (converted to z-scores) to ensure each pollutant and year contributed equally to the computed distance, preventing variables with larger scales from dominating the cluster formation.
3. **Determination of Clusters (k):** The number of clusters, k=5, was selected in our study based on elbow method. This choice allows for unbisaed clusters without over-simplifying or over-complicating clusters.
4. **Cluster Assignment and Profiling:** The K-means algorithm was run, and each district was assigned a cluster label (0 to 4). To interpret and rank these clusters, the mean value for each of the 25 original pollutant-year features was calculated for all districts within a cluster. K-means clustering was performed multiple times with different random initializations and selected the final clustering result that minimized the within-cluster sum of squares (WCSS). This ensured robustness and reduced the likelihood of converging to a local minimum [32]. The optimal number of clusters was determined using elbow method. The overall pollution severity of each cluster was then determined by ranking the clusters based on their aggregate mean values across all pollutants and years. 5-year mean for each pollutant within each cluster was calculated. A cluster was considered "more polluted" than another if it had higher mean concentrations as compared to the other five pollutants. This allowed for a clear classification from "cleanest" to "most polluted".
5. **Temporal Consistency (Friedman Test):** To assess whether pollutant levels changed significantly over the five years within the same clusters, the Friedman test was employed. This non-parametric test for repeated measures is ideal for analyzing matched data over time, helping to detect temporal variations within the same districts across the study period. A significant result from the test would indicate that pollutant levels were not constant throughout the five years, highlighting any temporal trends in the data.

2.6. Spatial Visualization

All spatial data, including the annual pollutant distributions and the final cluster assignments, were mapped using GEE. A consistent color scheme was used to facilitate comparison across years and pollutants. The resulting cluster map (Figure 8) and trend graphs (Figure 3-7) provide a powerful visual representation of the spatiotemporal dynamics and the resulting classification.

3. Results

3.1. National-Level Pollutant Trends (2020-2024)

The analysis of five major atmospheric pollutants across 64 districts of Bangladesh from 2020 to 2024 reveals distinct temporal patterns. National averages were calculated as the mean of all district values for each year.

Table 2 National Annual Average Pollutant Concentrations with Standard Deviation (2020-2024).

Pollutant		Year					5-Year Mean ± SD	Overall Trend
		2020	2021	2022	2023	2024		
SO_2	Mean	0.0000766	0.000112	0.0000866	0.000111	0.000109	0.000099 ± 0.0000241	fluctuating, peak in 2021
	SD	0.0000203	0.0000274	0.0000233	0.0000262	0.0000231		
NO_2	Mean	0.0000747	0.0000870	0.0000820	0.0000856	0.0000781	0.0000815 ± 0.0000233	Peak in 2021, then decline
	SD	0.0000201	0.0000273	0.0000280	0.0000226	0.0000186		
O_3	Mean	0.1208	0.1221	0.1242	0.1219	0.1242	0.1226 ± 0.00082	Consistent increase
	SD	0.00109	0.000543	0.000794	0.000647	0.00106		
HCHO	Mean	0.000181	0.000192	0.000184	0.000203	0.000202	0.000192 ± 0.0000138	Increase to 2023, then stable
	SD	0.0000118	0.0000117	0.0000148	0.0000159	0.0000149		
CH_4	Mean	1912.3	1925.5	1933.5	1938.6	1941.8	1930.3 ± 13.4	Steady, monotonic increase
	SD	14.2	15	13.4	11.9	12.5		

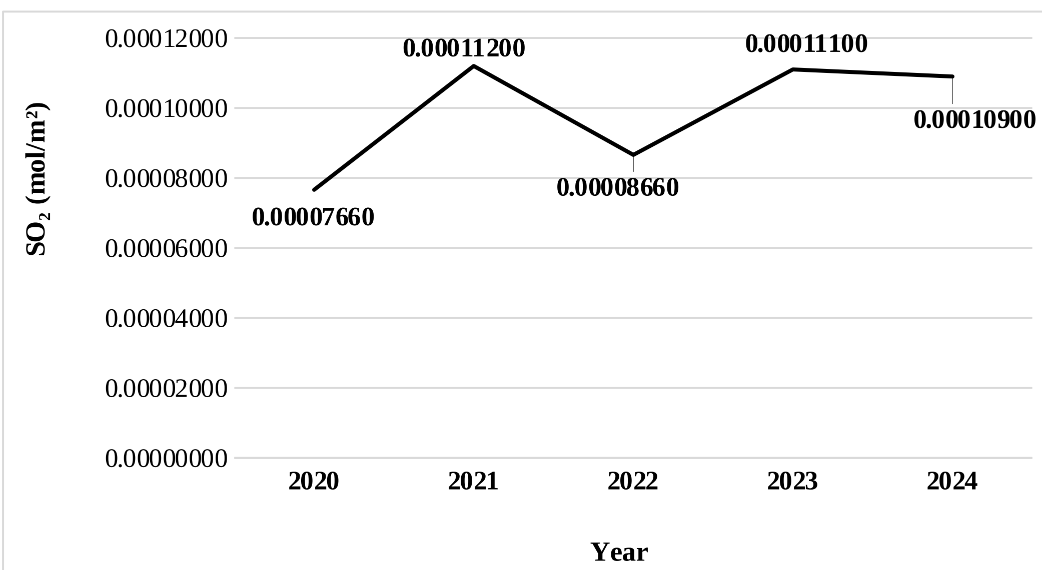

Fig 4 Line Graph of SO_2 Annual Mean Trend (2020-2024)

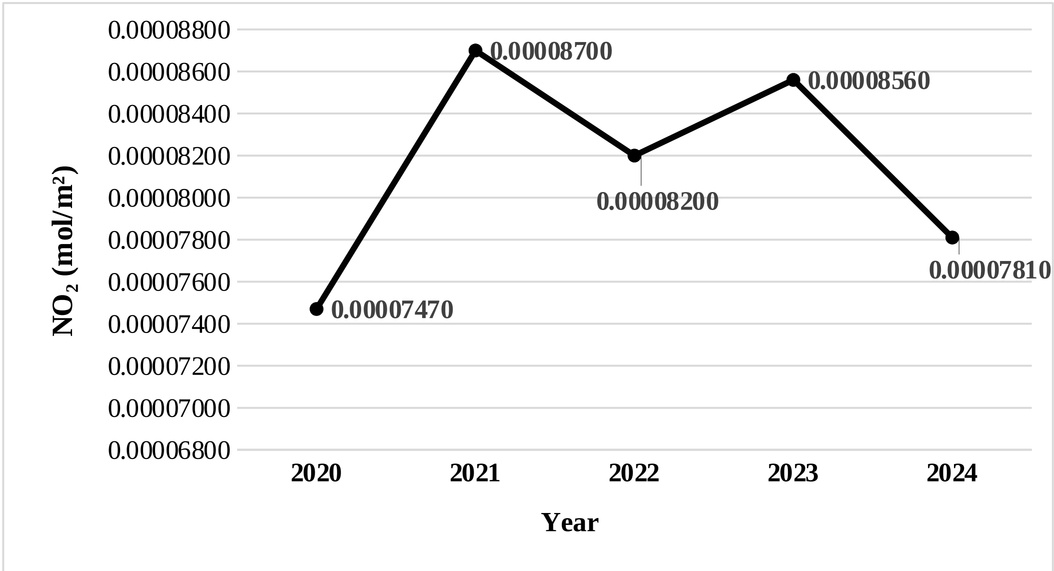

Fig 5 Line Graph of NO_2 Annual Mean Trend (2020-2024)

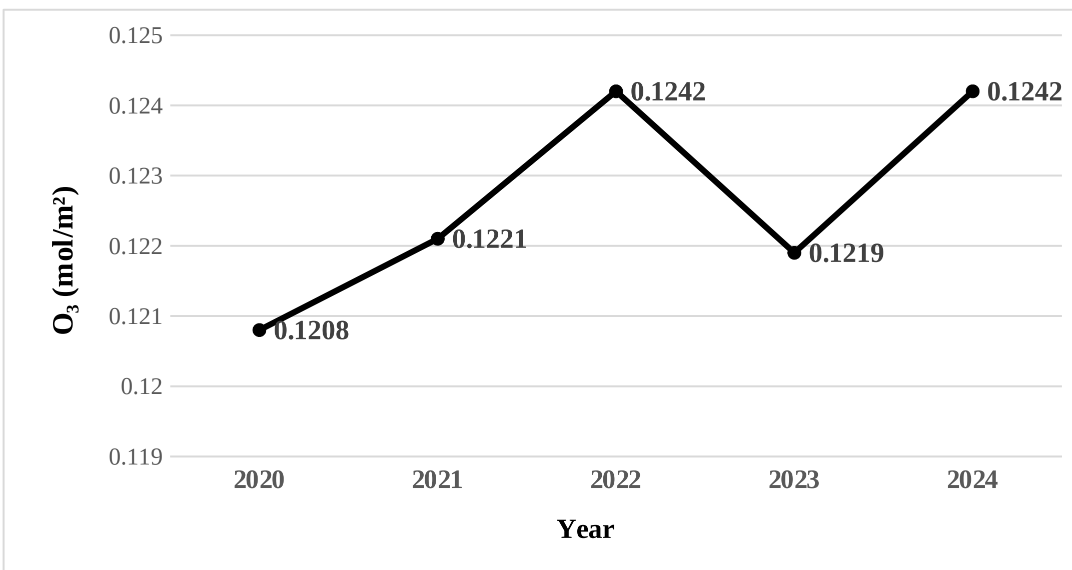

Fig 6 Line Graph of O_3 Annual Mean Trend (2020-2024)

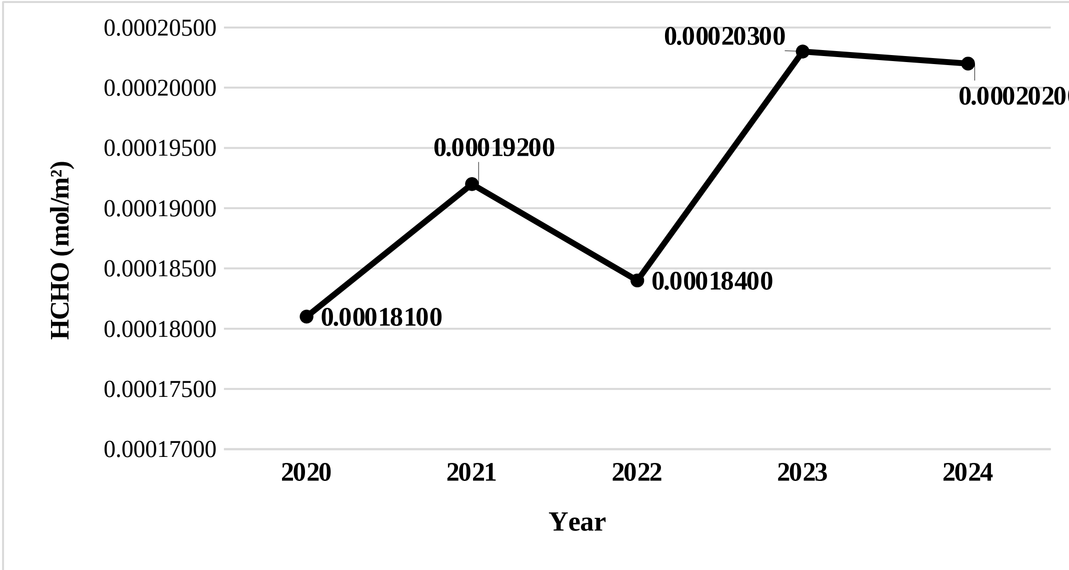

Fig 7 Line Graph of HCHO Annual Mean Trend (2020-2024)

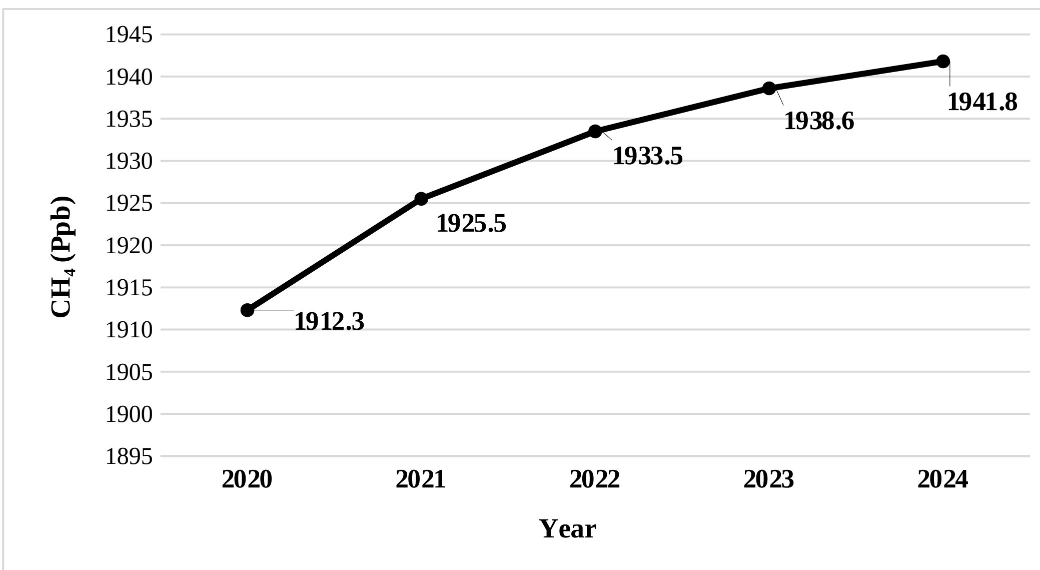

Fig 8 Line Graph of CH_4 Annual Mean Trend (2020-2024)

3.2. Spatial Heterogeneity and District-Level Hotspots

3.2.1 Spatial Heterogeneity

The analysis of mean and standard deviation for each pollutant across the 64 districts from 2020 to 2024 reveals significant spatial heterogeneity (Table 2). Only the mean and standard deviation were used in this study because the aim was to summarise the central tendency and inter-district variability of pollutant levels; this is sufficient for the non-parametric trend and clustering methods used, which do not rely on distributional assumptions. For SO_2, the concentrations remained relatively stable across years, with a consistent standard deviation, indicating moderate variation in industrial emissions across districts. NO_2 exhibited a slight peak in 2021, and its standard deviation reflected higher variability, particularly in urban hotspots. O_3 concentrations showed a gradual increase over time, with the standard deviation rising in later years, pointing to greater variability in urban areas due to photochemical reactions. HCHO steadily increased, with a stable standard deviation, suggesting consistent emissions, particularly from industrial and vehicular sources. Finally, CH_4 concentrations showed a steady rise, and the standard deviation decreased by 2024, indicating more consistent methane levels, especially in agricultural districts.

3.2.2 Spatial Hotspots (Based on 5-Year Means):

Four out of five pollutants in this study show higher concentrations in central Bangladesh except for ozone (as shown in Fig 9). In case of ozone, northern districts of Bangladesh tend to show higher concentrations, while for formaldehyde north-western districts have the higher concentrations.

- **NO_2 Hotspots:** The highest concentrations were found in the Dhaka Metropolitan Area. Narayanganj had the highest national 5-year mean (0.000192 mol/m²), followed by Gazipur (0.00016 mol/m²) and Dhaka (0.000156 mol/m²).

- **SO_2 Hotspots:** Elevated levels were concentrated in southwestern and central industrial districts. Satkhira had the highest 5-year mean (0.000172 mol/m²), with Narayanganj (0.000151 mol/m²) and Khulna (0.000146 mol/m²) also being major hotspots.

- **CH_4 Hotspots:** The most extreme values were in central districts. Narayanganj was the highest emitter (1976.2 ppb), significantly above the national average, followed by Munshiganj (1963.2 ppb) and Shariatpur (1946.5 ppb).

- **O_3 and HCHO:** Showed less extreme spatial variation but were generally elevated in and around urban and industrial clusters.

3.3. District-Level Temporal Trend Analysis (Mann-Kendall Test)

The Mann-Kendall trend test ($\alpha = 0.05$) was applied to the 2020-2024 time series for each pollutant in all 64 districts. The results are summarized in Table 3, which now includes the number of districts with significant trends and the range of Sen's Slope values to indicate the magnitude of change. In Table 3, the significant increasing and decreasing trends were determined using the Mann-Kendall trend test and Sen’s Slope for the respective districts.

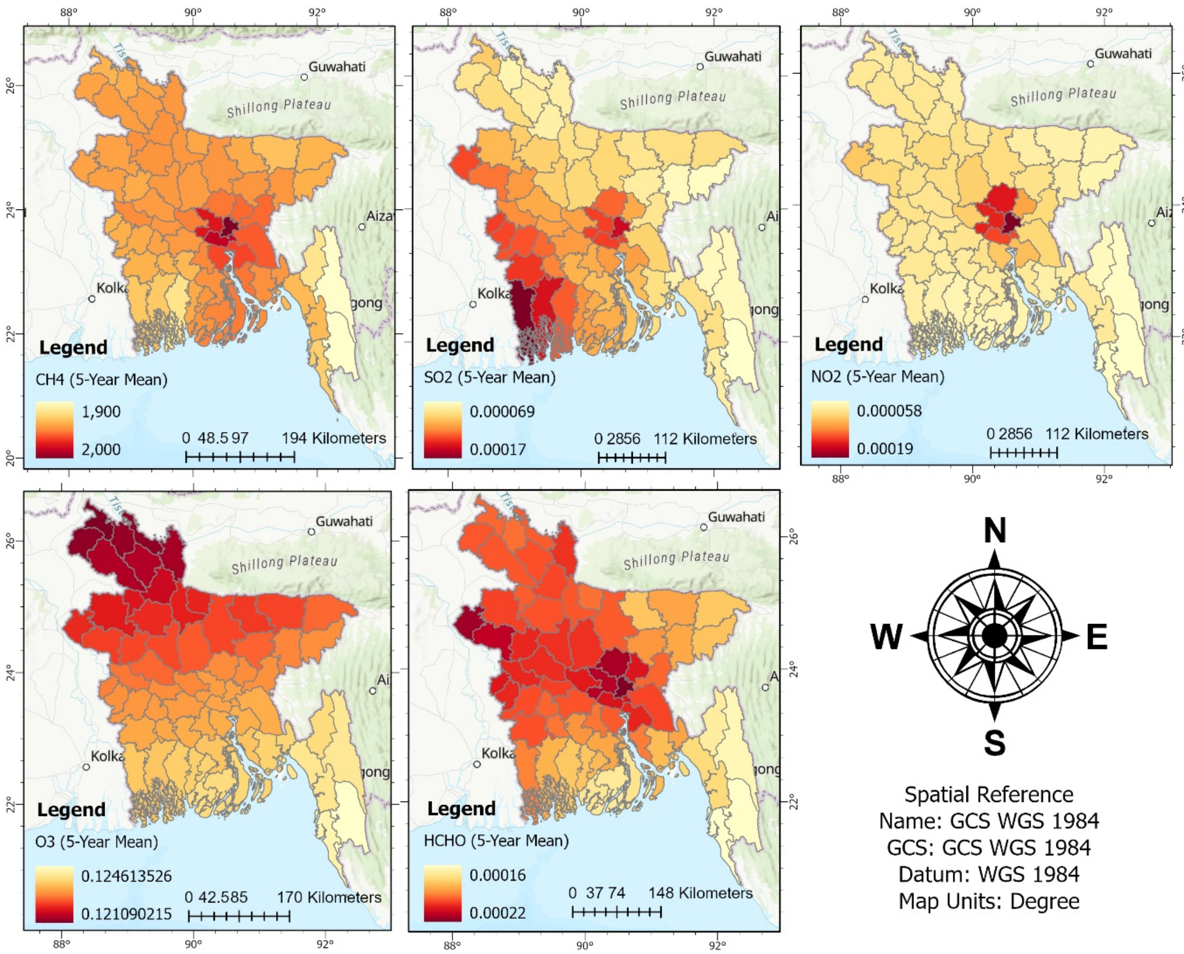

Fig 9 Spatial Distribution of air pollutants in Bangladesh over the past 5 years (2020-2024)

Table 3: Summary of District-Level Mann-Kendall Trend Test Results (2020-2024).

Pollutant	Significant Increasing Trend (No. of Districts)	Significant Decreasing Trend (No. of Districts)	No Significant Trend (No. of Districts)	Range of Sen's Slope (Significant Trends)
SO_2	18	8	38	+0.0000012 to +0.0000045 mol/m²/yr
NO_2	12	21	31	-0.0000028 to +0.0000021 mol/m²/yr
O_3	54	0	10	+0.00078 to +0.001.5 mol/m²/yr
HCHO	35	2	27	+0.0000015 to +0.0000052 mol/m²/yr
CH_4	61	0	3	+5.1 to +8.9 ppb/yr

Detailed Breakdown of Significant Trends:

- **SO_2 Increasing Trends (18 Districts):** Bagerhat, Brahmanbaria, Chandpur, Chuadanga, Dhaka, Gazipur, Jessore, Jhenaidah, Khulna, Kushtia, Magura, Meherpur, Narayanganj, Nawabganj, Rajshahi, Satkhira, Shariatpur, Sherpur. The strongest upward trend was observed in Satkhira (Sen's Slope = ~0.0000045 mol/m²/yr).

- **NO_2 Decreasing Trends (21 Districts):** Barisal, Bhola, Chandpur, Chittagong, Comilla, Cox's Bazar, Dhaka, Feni, Gazipur, Jamalpur, Kishoreganj, Lakshmipur, Manikganj, Munshiganj, Narayanganj, Narsingdi, Noakhali, Pabna, Rajbari, Shariatpur, Tangail. The most pronounced decrease was in Narayanganj (Sen's Slope = ~ -0.0000028 mol/m²/yr).

- **O_3 Increasing Trends (54 Districts):** All districts except Bandarban, Brahmanbaria, Chandpur, Chittagong, Cox's Bazar, Feni, Khagrachhari, Lakshmipur, Noakhali, and Rangamati. The northern districts of Lalmonirhat and Panchagarh showed the steepest increases (> 0.0014 mol/m²/yr).

- **HCHO Increasing Trends (35 Districts):** Bagerhat, Barisal, Bogra, Brahmanbaria, Chandpur, Chuadanga, Comilla, Dhaka, Dinajpur, Faridpur, Feni, Gaibandha, Gazipur, Gopalganj, Jamalpur, Jessore, Jhenaidah, Joypurhat, Kishoreganj, Kurigram, Kushtia, Lakshmipur, Madaripur, Manikganj, Meherpur, Munshiganj, Naogaon, Narayanganj, Narsingdi, Natore, Pabna, Rajshahi, Rangpur, Shariatpur, Sirajganj.

- **CH_4 Increasing Trends (61 Districts):** All districts except Bandarban, Khagrachhari, and Rangamati. The most substantial increases were clustered in the central districts, including Narayanganj and Munshiganj (> 8.5 ppb/yr).

3.4. K-Means Clustering: Classification and Ranking of Districts

K-means clustering (k=5) grouped the 64 districts into five distinct clusters based on their multi-pollutant profiles from 2020-2024. The mean values for each cluster, were used to rank them from best (cleanest) to worst (most polluted). The ranking was determined by calculating a normalized aggregate pollution score for each cluster, giving equal weight to the 5-year mean of each of the five pollutants.

Table 4 Per Year Pollutant Means of Different Clusters (2020-2024)

Cluster	SO_2 (mol/m²)	NO_2 (mol/m²)	O_3 (mol/m²)	HCHO (mol/m²)	CH_4 (ppb)
Cluster 1	0.0000638	0.0000708	0.1218	0.000172	1915.1
Cluster 4	0.0000700	0.0000770	0.1229	0.000187	1921.6
Cluster 0	0.0000754	0.0000837	0.1252	0.000202	1938.8
Cluster 2	0.0000921	0.0000876	0.1239	0.000206	1933.8
Cluster 3	0.000145	0.000103	0.1230	0.000229	1958.1

Table 5 Cluster, Ranking, Ranking Justification and District Categorization

Cluster	Ranking (Good to Bad)	Ranking Justification	Number of Districts	District Members
1	**1**	Cleanest (Lowest on all 5 pollutants)	16	Barguna, Barisal, Bhola, Jhalokati, Patuakhali, Pirojpur, Bandarban, Chittagong, Cox's Bazar, Feni, Khagrachhari, Noakhali, Rangamati, Bagerhat
4	**2**	Low SO_2 and NO_2, moderate O_3 and HCHO.	7	Brahamanbaria, Kishoreganj, Netrakona, Habiganj, Maulvibazar, Sunamganj, Sylhet
0	**3**	High O_3 & CH_4 (Highest O_3, 2nd highest CH_4)	16	Tangail, Jamalpur, Mymensingh, Sherpur, Bogra, Joypurhat, Naogaon, Sirajganj, Dinajpur, Gaibandha, Kurigram, Lalmonirhat, Nilphamari, Panchagarh, Rangpur, Thakurgaon
2	**4**	High Multi-Pollutant (High SO_2, HCHO; above avg. NO_2, CH_4)	23	Chandpur, Comilla, Lakshmipur, Faridpur, Gopalganj, Madaripur, Manikganj, Narsingdi, Rajbari, Shariatpur, Chuadanga, Jessore, Jhenaidah, Khulna, Kushtia, Magura, Meherpur, Narail, Satkhira, Natore, Nawabganj, Pabna, Rajshahi
3	**5**	Most Polluted (Highest on SO_2, NO_2, HCHO, CH_4)	4	Dhaka, Gazipur, Munshiganj, Narayanganj

Cluster Profile Analysis:

- **Cluster 1 (Cleanest):** This is the largest cluster, comprising 16 districts. It includes coastal districts (e.g., Barisal, Patuakhali), the Chittagong Hill Tracts (Bandarban, Rangamati), and the island of Bhola. These areas benefit from lower population density, less industrialization, and dispersive marine or forested environments. Based on our dataset, ollutant levels in this cluster are consistently 15-30% below the national average which may not be a universal standard to classify it as the cleanest.

- **Cluster 4 (Moderate):** This cluster consists of 7 northeastern districts (e.g., Sylhet, Sunamganj, Maulvibazar). These regions show the influence of urbanization and specific land-use patterns, with O_3 and HCHO levels near the national average. However, emissions of primary industrial and vehicular pollutants (SO_2, NO_2) remain relatively low.

- **Cluster 0 (High O_3 & CH_4):** This cluster of 16 districts is dominated by the northern and northwestern regions (e.g., Rangpur, Dinajpur, Bogra). Its defining signature is the highest average O_3 concentration in the country, likely driven by precursor emissions from agricultural residue burning and biogenic sources, combined with high background levels of CH_4 from extensive rice cultivation and livestock.

- **Cluster 2 (High Multi-Pollutant):** This is the second-largest cluster (23 districts) and represents a widespread pattern of high pollution. It includes major southwestern industrial and port cities (Khulna, Jessore, Satkhira) and numerous central districts. It is characterized by high levels of SO_2, NO_2, and HCHO, indicating a mix of industrial, vehicular, and agricultural combustion sources. The CH_4 levels in this cluster are also very high.

- **Cluster 3 (Extreme Urban/Industrial):** This smallest but most critical cluster contains only 4 districts: Dhaka, Gazipur, Munshiganj, and Narayanganj. This is the core of the Dhaka megacity and its immediate industrial satellites. It is in a league of its own, with extremely high levels of NO_2 (marker for intense traffic and combustion), the highest HCHO, and the highest CH_4 in the country. The pollution burden here is an order of magnitude worse than in the cleanest cluster, underscoring the severe environmental challenges in the nation's economic heartland.

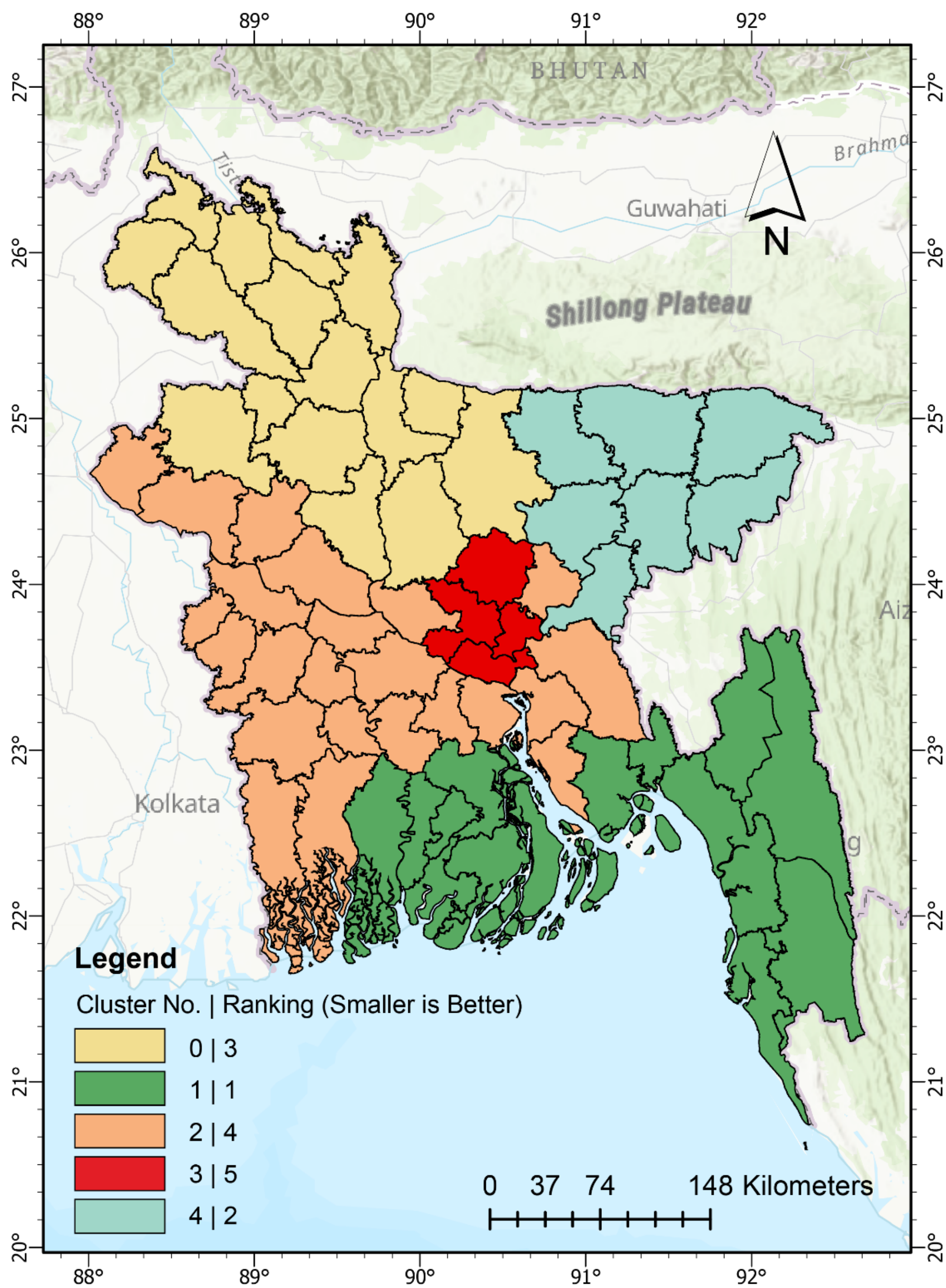

Fig 8 Cluster Assigned (District wise) Map of Bangladesh with Cluster Ranking

3.5. Temporal Consistency within Clusters (Friedman Test)

The Friedman test was applied to evaluate the null hypothesis that pollutant concentrations remained constant across the five years (2020-2024) within each cluster. The results, detailed in Table 6, overwhelmingly reject this hypothesis.

Table 6 Results of the Friedman Test for Temporal Changes Within Clusters (2020-2024).

Cluster	Pollutant	Chi-Square (χ^2) Value	p-value	Interpretation
Cluster 1 (Cleanest)	SO_2	12.45	0.014	Significant Change
	NO_2	15.82	0.003	Significant Change
	O_3	48.31	< 0.001	Significant Change
	HCHO	28.56	< 0.001	Significant Change
	CH_4	62.18	< 0.001	Significant Change
Cluster 2 (High Multi-Pollutant)	SO_2	35.67	< 0.001	Significant Change
	NO_2	40.12	< 0.001	Significant Change
	O_3	52.89	< 0.001	Significant Change
	HCHO	45.23	< 0.001	Significant Change
	CH_4	58.74	< 0.001	Significant Change
Cluster 3 (Extreme Urban/Industrial)	SO_2	18.92	< 0.001	Significant Change
	NO_2	22.45	< 0.001	Significant Change
	O_3	31.05	< 0.001	Significant Change
	HCHO	25.81	< 0.001	Significant Change
	CH_4	38.96	< 0.001	Significant Change
Cluster 4 (Moderate)	SO_2	14.28	0.006	Significant Change
	NO_2	16.93	0.002	Significant Change
	O_3	44.57	< 0.001	Significant Change

	HCHO	32.14	< 0.001	Significant Change
	CH_4	55.42	< 0.001	Significant Change
Cluster 0 (High O_3 & CH_4)	SO_2	29.34	< 0.001	Significant Change
	NO_2	25.67	< 0.001	Significant Change
	O_3	61.25	< 0.001	Significant Change
	HCHO	38.99	< 0.001	Significant Change
	CH_4	65.88	< 0.001	Significant Change

*Note: Degrees of Freedom (df) = 4 for all tests.

The results were statistically significant ($p < 0.05$) for all five pollutants within each of the five clusters. This provides robust statistical evidence that the temporal trends observed at the national level are consistent and detectable within each distinct pollution-profile group. Critically, the increasing trends for O_3 and CH_4 were significant within all five clusters, demonstrating that this environmental change is a universal phenomenon across Bangladesh, affecting both the cleanest and most polluted regions. The significant changes in other pollutants within clusters like 2 and 3 further validate the dynamic nature of the pollution landscape, even within grouped districts.

4. Discussion

This study provides one of the first comprehensive, district-level spatiotemporal assessments of multiple atmospheric pollutants (SO_2, NO_2, O_3, HCHO, CH_4) in Bangladesh for the 2020–2024 period using the available Sentinel 5P dataset the use of which to detect and measure pollutant is already validated by the existing literatures [9,10] . By moving beyond national averages and employing a robust clustering methodology, it reveals a landscape of extreme pollution inequality and distinct regional regimes, with policy-relevant implications for mitigation and attribution.

4.1. Cluster Typologies and Drivers with Regional Context

The Extreme Urban/Industrial cluster (Cluster 3), comprising Dhaka and the adjacent districts of Gazipur, Munshiganj, and Narayanganj, is characterized by very high NO_2, reflecting intense fossil-fuel combustion from traffic and industry [11,12,13]. This aligns with prior analyses showing these four districts as a national NO_2 hotspot [6]. Elevated CH_4 and HCHO levels point to mixed urban sources: methane from waste management and gas leaks, and HCHO as a proxy for enhanced VOCs from traffic, solvents, and incomplete combustion [14,15,16].

The High O_3 & CH_4 cluster (Cluster 0) exhibits a signature typical of agricultural and biogenic influence, with strong photochemical O_3 production downwind of NOx and VOC precursors, and elevated CH_4 linked to rice paddies and ruminants. Such processes are well documented in South and Southeast Asia [17,18]. Studies confirm rice systems as a dominant CH_4 source in Bangladesh [19,20].

The high multi-pollutant cluster (Cluster 2), spanning 23 southwestern and central districts, is defined by consistently high SO_2, NO_2, and HCHO, with above-average CH_4. This profile reflects the combined impact of industrial zones (e.g., Khulna, Jessore, Satkhira), power generation, and agro-industrial activities [21]. Such regions represent a critical “second front” of industrial air pollution outside Dhaka.

The moderate cluster (Cluster 4), largely covering northeastern districts (e.g., Sylhet, Sunamganj, Moulvibazar), shows relatively low SO_2 and NO_2 but near-national-average O_3 and HCHO, reflecting land-use change, emerging urbanization, and biogenic emissions [22]. Although less polluted than Clusters 2 and 3, this cluster demonstrates vulnerability to rising O_3 and HCHO trends.

Finally, the cleanest cluster (Cluster 1), including coastal and hilly regions, has the lowest pollutant loadings. These areas benefit from lower industrial density, marine dispersion, and forested environments, though they are not immune to regional increases in O_3 and CH_4 [23].

Temporal dynamics reinforce these interpretations. Central districts show modest NO_2 declines, consistent with traffic or COVID-related activity reductions [24]. However, widespread increases in O_3 and CH_4 across all clusters mirror regional trends attributed to precursor emissions and evolving atmospheric chemistry [18,25].

Bangladesh’s air-quality literature has largely focused on single pollutants or city-scale studies. By contrast, this study employs a multi-pollutant, district-scale clustering approach. The formaldehyde-to-NO_2 ratio, widely used to diagnose ozone sensitivity regimes [25], supports this study’s observation of distinct photochemical environments among clusters. For methane, inventories confirm rice paddies and livestock as dominant sources in Bangladesh, while mitigation strategies such as Alternate Wetting and Drying (AWD) and waste management offer clear intervention pathways [19,20].

4.2. Comparison with Existing Studies

Numerous studies have assessed air quality in Bangladesh, but most are restricted to individual cities or short time frames and often focus on a single pollutant. For example, a recent time-series analysis of Sentinel-5P data for Dhaka (2020–2024) reported winter peaks for CO and NO_2, a pre-monsoon SO_2 maximum and an O_3 maximum in May; it also identified pollution hotspots in Tejgaon and Mirpur and detected moderate to strong increasing trends for CO and NO2, while SO2 was concentrated in southern industrial zones such as Keraniganj and Jatrabari [5]. Multi-sensor analyses during the strict COVID-19 lockdown (March–May 2020) found that $PM_{2.5}$ and NO_2 declined by 37–77 % and 3–43 % (in April, 2020), respectively, across major Bangladeshi cities while O_3 increased by 3–12 % during the full lockdown and a study done on India as a case study during Covid-19 also corroborated this decreasing trend [26, 27]. Haque et al. (2022) similarly reported that NO_2 concentrations dropped sharply after the March 2020 lockdown and remained lower than 2019 values; however, they noted that O_3 densities were higher in April 2020 than in 2019 and were spatially heterogeneous, with higher values in the Rangpur region and lower values around Chattogram [7].

These city-scale and event-focused studies provide valuable snapshots but are limited in spatial and pollutant scope. Rahaman et al. (2024) applied geostatistical hotspot analysis to multiple pollutants and land surface temperature across Bangladesh, showing that persistent and intensifying hotspots for NO_2 and CO coincide with major urban areas and that high-high clusters of NO_2 and O_3 affect about 48.53 % and 54.67% of the country [21]. While their approach quantifies spatial clustering, it treats each pollutant separately and does not classify districts into joint pollution regimes.

Our cluster-based analysis builds on and extends these findings. It confirms Dhaka's status as a NO_2 hotspot and documents NO_2 declines after 2020, yet it also reveals a previously undocumented multi-pollutant burden in which HCHO and CH_4 play prominent roles. The high multi-pollutant (Cluster 2) and extreme urban/industrial (Cluster 3) clusters encompass not only Dhaka but also a contiguous belt of western and central districts with elevated SO_2, NO_2 and HCHO, underscoring that Bangladesh's pollution problem extends far beyond the capital. Moreover, we show that O_3 and CH_4 have increased across all clusters despite pandemic-related reductions in human activity [26]. This universal rise supports the hypothesis that declining NO_x and increasing biogenic and anthropogenic precursors drive ozone formation and highlights the need to consider methane as both a climate and air- quality pollutant—especially given that Dhaka is among the top ten persistent global methane emission hotspots (2.4 million t yr^{-1}) [8]. Despite notable decreases in NO_2 in several urban districts, none of the 64 districts exhibited a significant decreasing trend in ozone during 2020–2024. This outcome is consistent with regional atmospheric chemistry, as tropospheric ozone in South and Southeast Asia is increasingly driven by large-scale photochemical processes rather than local emissions alone [33]. Ozone is a secondary pollutant whose formation depends on complex interactions between NO_x, VOCs, methane, and meteorological conditions [33]. Recent top-down simulations over Southeast Asia show that 55–72% of observed ozone growth is attributable to regional anthropogenic emission increases**,** while 41–54% of seasonal variability is explained by meteorological factors such as temperature, solar radiation, and dry-season stagnation [33]**.** Moreover, rising methane and VOC emissions across the region elevate the background ozone entering Bangladesh, offsetting the potential benefits of localized NO_2 reductions. In case of increasing trends of methane, a recent study suggested that the majority of significant increasing trends of methane in south asia are attributable to agriculture and waste related methane [34]. It is very likely that no decreasing trend for methane in any district for our study area is associated with agriculture and waste.

Internationally, cluster analyses of air pollution differ markedly from our study. Zou et al. (2014) used local Moran's I and logistic regression to map benzene exposure across 52613 urban census tracts in the United States [28]. They identified high-risk clusters for children and older adults in specific census divisions and demonstrated that racial minorities and low-income groups were disproportionately exposed. However, their analysis focused on a single pollutant and aggregated socio-demographic data, limiting its utility for multi-pollutant management. More recently, Vandeninden et al. (2024) applied hierarchical clustering to NO_2, $PM_{2.5}$, black carbon and O_3 across 19 794 statistical sectors in Belgium and identified five clusters: the cleanest cluster (E) had low NO_2 and $PM_{2.5}$ but high O_3 and covered 35 % of the territory, whereas the most polluted cluster (C), covering only 2 % of the territory but 20 % of the population, exhibited the highest NO_2, black carbon and $PM_{2.5}$ concentrations [29]. All Belgian clusters exceeded the WHO $PM_{2.5}$

guideline and only cluster E met the NO_2 guideline. These studies demonstrate the utility of clustering but generally exclude CH_4 and HCHO and are rarely applied to low-income settings.

By contrast, our work combines five pollutants (SO_2, NO_2, O_3, HCHO, CH_4) and a four-year time frame to classify all districts of Bangladesh into five coherent regimes. This approach reveals heterogeneity similar to Belgium's clusters but shows that most of Bangladesh's population resides in high-multi-pollutant zones. Unlike the U.S. benzene study, we diagnose cross-pollutant interactions—such as using the formaldehyde-to-NO_2 ratio to infer ozone sensitivity—and we highlight the universal rise of O_3 and CH_4. These insights underscore the need for integrated mitigation strategies that simultaneously address combustion, industrial, and agricultural sources, which are absent in most prior analyses.

4.3. Policy and Health Implications

The multi-pollutant burden of Cluster 3 implies compounded health risks, as co-exposure to NO_2, O_3, and combustion by-products is linked to increased respiratory and cardiovascular morbidity and mortality [25, 30]. Given Dhaka's density, the population-weighted health burden is highly concentrated here.

Cluster-specific policy priorities emerge:

- **Cluster 3 (Extreme Urban/Industrial):** Expand clean public transport, enforce industrial and kiln standards, and upgrade waste systems to curb NO_2 emissions as it also gives rise to tropospheric ozone [18, 31].
- **Cluster 2 (High Multi-Pollutant):** Strengthen industrial emission monitoring in Khulna–Rajshahi belt, regulate SO_2 and NO_2 from power plants and agro-industrial complexes, and improve fuel standards.
- **Cluster 4 (Moderate/Northeast):** Implement early interventions in growing urban centers (e.g., Sylhet), combining land-use planning with emission control to prevent escalation.
- **Cluster 0 (High O_3 & CH_4):** Reduce CH_4 through agricultural practices (Alternate Wetting and Drying or AWD in rice, livestock management) and control ozone precursors through regional cooperation [17,19].
- **Cluster 1 (Cleanest):** Maintain low-emission land use, preserve forests and coastal buffers, and monitor background increases in O_3 and CH_4.

5. Conclusion

This study presents a detailed, district-level spatiotemporal analysis of five major atmospheric pollutants (SO_2, NO_2, O_3, HCHO, and CH_4) in Bangladesh from 2020 to 2024, leveraging high-resolution Sentinel-5P TROPOMI satellite data and unsupervised machine learning techniques. The findings reveal significant regional disparities in air pollution across the country, demonstrating that Bangladesh's air pollution problem is multi-faceted and geographically heterogeneous.

The analysis identified five distinct pollution clusters, each driven by different regional sources. The Extreme Urban/Industrial Cluster (Cluster 3) encompassing Dhaka and adjacent districts represents the most polluted areas, dominated by vehicular emissions, industrial combustion, and waste management practices. In contrast, the High O_3 & CH_4 Cluster (Cluster 0) is characterized by agricultural emissions, particularly from rice paddies and livestock. The High Multi-Pollutant Cluster (Cluster 2), spanning industrial zones, is marked by high levels of SO_2, NO_2, and HCHO, indicative of combined industrial, vehicular, and agro-industrial pollution. Cluster 4, covering northeastern regions, is emerging as an area of concern due to urbanization and biogenic emissions, while the Cleanest Cluster (Cluster 1), primarily in coastal and hilly regions, experiences the lowest pollutant levels but still shows gradual increases in O_3 and CH_4.

These findings underscore the need for tailored pollution management strategies that address the specific characteristics of each pollution cluster. Current policies should shift from a one-size-fits-all approach to more localized interventions that target the unique sources and patterns of pollution within each cluster. The use of satellite data and clustering algorithms provides a novel, scalable methodology for air quality management, offering valuable insights that can be applied not only in Bangladesh but also in other regions facing similar challenges.

This study has important implications for air quality management in Bangladesh and offers a replicable model for other regions facing similar pollution challenges. By identifying and characterizing distinct pollution clusters, it paves the way for more informed, evidence-based policy decisions aimed at mitigating the health and environmental impacts of air pollution in Bangladesh.

6. Limitations and Future Research

This study, while robust, has some limitations. The satellite-based data, while providing excellent spatial coverage, do not fully capture surface exposure, and biases in certain pollutant retrievals suggest the need for further validation with ground-based measurements. Moreover, the satellite-derived data used in this study is provided in mol/m² units for each pollutant, except for CH_4. This difference in units limits its comparability with other ground-based studies and global standards (such as health guidelines) that typically use units like ppm or ppb. Additionally, the annual averaging of data may obscure seasonal variations, which could be important for understanding short-term health impacts. Future research should explore more granular temporal analysis, such as seasonal or monthly studies, to better capture these fluctuations.

Moreover, the current clustering approach, which is based solely on pollutant concentrations, could be enhanced by incorporating population data and source-apportionment methods, such as chemical-transport modeling, to offer a more comprehensive understanding of exposure and its health impacts.

Future work should focus on integrating ground-based monitoring networks and epidemiological data to assess population-weighted exposure and health burden more accurately. Furthermore, the application of isotopic and chemical-transport models will help differentiate between local and transported pollutant sources, particularly for O_3 and CH_4. Testing and quantifying the impact of mitigation strategies, such as Alternate Wetting and Drying (AWD) in agriculture, landfill gas capture, and industrial kiln modernization, should be prioritized to assess their effectiveness in reducing multi-pollutant concentrations.

This study sets a strong foundation for developing integrated, multi-sectoral solutions to air pollution, emphasizing the urgent need for coordinated action across urban, industrial, and agricultural sectors.

References

[1] Tarín-Carrasco, P., Im, U., Geels, C., Palacios-Peña, L., & Jiménez-Guerrero, P. (2021). Contribution of fine particulate matter to present and future premature mortality over Europe: A non-linear response. *Environment International*, *153*(106517), 106517. https://doi.org/10.1016/j.envint.2021.106517

[2] Lee, M., Bayazid, A. R., Rosenthal, L., Khan, R. H., Arku, R., Barratt, B., Quayyum, Z., & Baumgartner, J. (2025). Spatiotemporal patterns in air pollution and sound in Dhaka, Bangladesh. *Scientific Reports*, *15*(1), 32852. https://doi.org/10.1038/s41598-025-12815-9

[3] Umali-Deininger, D. (2024, November 20). Clearing the air: Addressing Bangladesh's air pollution crisis. *World Bank Blogs*. https://blogs.worldbank.org/en/endpovertyinsouthasia/clearing-the-air--addressing-bangladesh-s-air-pollution-crisis

[4] Elaprolu, S. (2024). Causes of ambient air pollution in South Asia and effects on South Asians' cardiovascular health. Global Journal of Health Science, 16(11), 1. https://doi.org/10.5539/gjhs.v16n11p1

[5] Rahman, M. M., Kamruzzaman, M., Faridatul, M. I., & Szabó, G. (2025). Spatiotemporal dynamics of urban air pollution in Dhaka city (2020–2024) using time-series Sentinel-5P satellite images and Google earth engine (GEE). *Environments*, *12*(8), 274. https://doi.org/10.3390/environments12080274

[6] Basak, P., Dey, S., & Elahi, K. M. (2023). Air pollution in urban Bangladesh from climate change and public health perspectives. In *Global Perspectives on Health Geography* (pp. 129–149). Springer Nature Switzerland.

[7] Haque, M. N., Sharif, M. S., Rudra, R. R., Mahi, M. M., Uddin, M. J., & Ellah, R. G. A. (2022). Analyzing the spatio-temporal directions of air pollutants for the initial wave of Covid-19 epidemic over Bangladesh: Application of satellite imageries and Google Earth Engine. *Remote Sensing Applications Society and Environment*, *28*(100862), 100862. https://doi.org/10.1016/j.rsase.2022.100862

[8] The European Space Agency. (2025, February 3). *Top 10 persistent methane sources*. The European Space Agency. https://www.esa.int/Applications/Observing_the_Earth/Copernicus/Sentinel-5P/Top_10_persistent_methane_sources

[9] De Smedt, I., Pinardi, G., Vigouroux, C., Compernolle, S., Bais, A., Benavent, N., Boersma, F., Chan, K.-L., Donner, S., Eichmann, K.-U., Hedelt, P., Hendrick, F., Irie, H., Kumar, V., Lambert, J.-C., Langerock, B., Lerot, C., Liu, C., Loyola, D., … Van Roozendael, M. (2021). Comparative assessment of TROPOMI and OMI formaldehyde observations and validation

against MAX-DOAS network column measurements. *Atmospheric Chemistry and Physics*, *21*(16), 12561–12593. https://doi.org/10.5194/acp-21-12561-2021

[10] Lee, G. T., Park, R. J., Kwon, H.-A., Ha, E. S., Lee, S. D., Shin, S., Ahn, M.-H., Kang, M., Choi, Y.-S., Kim, G., Lee, D.-W., Kim, D.-R., Hong, H., Langerock, B., Vigouroux, C., Lerot, C., Hendrick, F., Pinardi, G., De Smedt, I., … Kim, J. (2024). First evaluation of the GEMS formaldehyde product against TROPOMI and ground-based column measurements during the in-orbit test period. *Atmospheric Chemistry and Physics*, *24*(8), 4733–4749. https://doi.org/10.5194/acp-24-4733-2024

[11] Hao, W. M., Wofsy, S. C., McElroy, M. B., Beer, J. M., & Toqan, M. A. (1987). Sources of Atmospheric nitrous oxide from combustion. *Journal of Geophysical Research*, *92*(D3), 3098. https://doi.org/10.1029/jd092id03p03098

[12] Horowitz, L. W., & Jacob, D. J. (1999). Global impact of fossil fuel combustion on atmospheric NO_x. *Journal of Geophysical Research*, *104*(D19), 23823–23840. https://doi.org/10.1029/1999jd900205

[13] Huang, T., Zhu, X., Zhong, Q., Yun, X., Meng, W., Li, B., Ma, J., Zeng, E. Y., & Tao, S. (2017). Spatial and temporal trends in global emissions of nitrogen oxides from 1960 to 2014. *Environmental Science & Technology*, *51*(14), 7992–8000. https://doi.org/10.1021/acs.est.7b02235

[14] Bingemer, H. G., & Crutzen, P. J. (1987). The production of methane from solid wastes. *Journal of Geophysical Research*, *92*(D2), 2181–2187. https://doi.org/10.1029/jd092id02p02181

[15] El-Fadel, M., & Massoud, M. (2001). Methane emissions from wastewater management. *Environmental Pollution (Barking, Essex: 1987)*, *114*(2), 177–185. https://doi.org/10.1016/s0269-7491(00)00222-0

[16] Hendrick, M. F., Ackley, R., Sanaie-Movahed, B., Tang, X., & Phillips, N. G. (2016). Fugitive methane emissions from leak-prone natural gas distribution infrastructure in urban environments. *Environmental Pollution (Barking, Essex: 1987)*, *213*, 710–716. https://doi.org/10.1016/j.envpol.2016.01.094

[17] Wang, X., Fu, T.-M., Zhang, L., Lu, X., Liu, X., Amnuaylojaroen, T., Latif, M. T., Ma, Y., Zhang, L., Feng, X., Zhu, L., Shen, H., & Yang, X. (2022). Rapidly changing emissions drove substantial surface and tropospheric ozone increases over southeast Asia. *Geophysical Research Letters*, *49*(19). https://doi.org/10.1029/2022gl100223

[18] Li, S., Yang, Y., Wang, H., Li, P., Li, K., Ren, L., Wang, P., Li, B., Mao, Y., & Liao, H. (2023). Rapid increase in tropospheric ozone over Southeast Asia attributed to changes in precursor emission source regions and sectors. *Atmospheric Environment (Oxford, England: 1994)*, *304*(119776), 119776. https://doi.org/10.1016/j.atmosenv.2023.119776

[19] Pazhanivelan, S., Sudarmanian, N. S., Geethalakshmi, V., Deiveegan, M., Ragunath, K., Sivamurugan, A. P., & Shanmugapriya, P. (2024). Assessing methane emissions from rice fields in large irrigation projects using satellite-derived land surface temperature and

agronomic flooding: A spatial analysis. *Agriculture*, *14*(3), 496. https://doi.org/10.3390/agriculture14030496

[20] Aziz, S., & Chowdhury, S. A. (2023). Analysis of agricultural greenhouse gas emissions using the STIRPAT model: a case study of Bangladesh. *Environment Development and Sustainability*, *25*(5), 3945–3965. https://doi.org/10.1007/s10668-022-02224-7

[21] Rahman, M., Meng, L., Mathews, A. J., & Bertman, S. (2024). Spatiotemporal analysis of urban growth and PM2.5 concentrations in Sylhet, Bangladesh. *Atmosphere*, *15*(11), 1305. https://doi.org/10.3390/atmos15111305

[22] Gustafsson, O., Kruså, M., Zencak, Z., Sheesley, R. J., Granat, L., Engström, E., Praveen, P. S., Rao, P. S. P., Leck, C., & Rodhe, H. (2009). Brown clouds over South Asia: biomass or fossil fuel combustion? *Science (New York, N.Y.)*, *323*(5913), 495–498. https://doi.org/10.1126/science.1164857

[23] Casas, G., Martínez-Varela, A., Roscales, J. L., Vila-Costa, M., Dachs, J., & Jiménez, B. (2020). Enrichment of perfluoroalkyl substances in the sea-surface microlayer and sea-spray aerosols in the Southern Ocean. *Environmental Pollution (Barking, Essex: 1987)*, *267*(115512), 115512. https://doi.org/10.1016/j.envpol.2020.115512

[24] Fisher, B. L., Lamsal, L. N., Fasnacht, Z., Oman, L. D., Joiner, J., Krotkov, N. A., Choi, S., Qin, W., & Yang, E.-S. (2024). Revised estimates of NO2 reductions during the COVID-19 lockdowns using updated TROPOMI NO2 retrievals and model simulations. *Atmospheric Environment (Oxford, England: 1994)*, *326*(120459), 120459. https://doi.org/10.1016/j.atmosenv.2024.120459

[25] Vazquez Santiago, J., Hata, H., Martinez-Noriega, E. J., & Inoue, K. (2024). Ozone trends and their sensitivity in global megacities under the warming climate. *Nature Communications*, *15*(1), 10236. https://doi.org/10.1038/s41467-024-54490-w

[26] Qiu, Z., Ali, M. A., Nichol, J. E., Bilal, M., Tiwari, P., Habtemicheal, B. A., Almazroui, M., Mondal, S. K., Mazhar, U., Wang, Y., Sarker, S., Mustafa, F., & Rahman, M. A. (2021). Spatiotemporal investigations of multi-sensor air pollution data over Bangladesh during COVID-19 lockdown. *Remote Sensing*, *13*(5), 877. https://doi.org/10.3390/rs13050877

[27] Mathew, A., Shekar, P. R., Nair, A. T., Mallick, J., Rathod, C., Bindajam, A. A., Alharbi, M. M., & Abdo, H. G. (2024). Unveiling urban air quality dynamics during COVID-19: a Sentinel-5P TROPOMI hotspot analysis. *Scientific Reports*, *14*(1), 21624. https://doi.org/10.1038/s41598-024-72276-4

[28] Zou, B., Peng, F., Wan, N., Mamady, K., & Wilson, G. J. (2014). Spatial cluster detection of air pollution exposure inequities across the United States. *PloS One*, *9*(3), e91917. https://doi.org/10.1371/journal.pone.0091917

[29] Vandeninden, B., De Clercq, E. M., Devleesschauwer, B., Otavova, M., Bouland, C., & Faes, C. (2024). Cluster pattern analysis of environmental stressors and quantifying their impact on all-cause mortality in Belgium. *BMC Public Health*, *24*(1), 536. https://doi.org/10.1186/s12889-024-18011-0

[30] Atkinson, R. W., Butland, B. K., Dimitroulopoulou, C., Heal, M. R., Stedman, J. R., Carslaw, N., Jarvis, D., Heaviside, C., Vardoulakis, S., Walton, H., & Anderson, H. R. (2016). Long-term exposure to ambient ozone and mortality: a quantitative systematic review and meta-analysis of evidence from cohort studies. *BMJ Open*, *6*(2), e009493. https://doi.org/10.1136/bmjopen-2015-009493

[31] Pu, D., Zhu, L., De Smedt, I., Li, X., Sun, W., Wang, D., Liu, S., Li, J., Shu, L., Chen, Y., Sun, S., Zuo, X., Fu, W., Xu, P., Yang, X., & Fu, T.-M. (2022). Response of anthropogenic volatile organic compound emissions to urbanization in Asia probed with TROPOMI and VIIRS satellite observations. *Geophysical Research Letters*, *49*(18). https://doi.org/10.1029/2022gl099470

[32] Naghizadeh, A., & Metaxas, D. N. (2020). Condensed silhouette: An optimized filtering process for cluster selection in K-means. *Procedia Computer Science*, *176*, 205–214. https://doi.org/10.1016/j.procs.2020.08.022

[33] Wang, X., Fu, T.-M., Zhang, L., Lu, X., Liu, X., Amnuaylojaroen, T., Latif, M. T., Ma, Y., Zhang, L., Feng, X., Zhu, L., Shen, H., & Yang, X. (2022). Rapidly changing emissions drove substantial surface and tropospheric ozone increases over southeast Asia. *Geophysical Research Letters*, *49*(19). https://doi.org/10.1029/2022gl100223

[34] Stavert, A. R., Saunois, M., Canadell, J. G., Poulter, B., Jackson, R. B., Regnier, P., Lauerwald, R., Raymond, P. A., Allen, G. H., Patra, P. K., Bergamaschi, P., Bousquet, P., Chandra, N., Ciais, P., Gustafson, A., Ishizawa, M., Ito, A., Kleinen, T., Maksyutov, S., … Zhuang, Q. (2022). Regional trends and drivers of the global methane budget. *Global Change Biology*, *28*(1), 182–200. https://doi.org/10.1111/gcb.15901