



Forecasting Air Pollution in Aanand Vihar, Delhi through Pollutant Predictive Analysis

Devanshi Dogra - Student

Birla Institute of Technology, Mesra Ranchi, Off campus Noida

Course Name: Masters of computer Application

Mr Anurag Joshi – Head of Department

Abstract -

This paper presents a comprehensive analysis of air pollution in the Aanand Vihar region of Delhi, focusing on predictive analysis to forecast pollutant concentrations. The study aims to address the critical environmental and health issues associated with air pollution by utilizing historical air quality data, statistical examinations, and regression models. The key pollutants considered include PM 2.5, PM 10, SO₂, NO₂, and O₃, known for their adverse effects on human health and the environment.

The research methodology involves gathering relevant air quality data, conducting in-depth data analysis, and utilizing statistical models for forecasting. The results offer valuable insights into pollution trends, which are crucial for policy formulation and implementation of effective pollution control measures. By predicting pollutant concentrations for the next five years, this study contributes to understanding the dynamics of air pollution in the Aanand Vihar region and provides a basis for developing strategies to improve air quality and mitigate the impacts of pollution on public health and the environment.

Introduction

Defining "environment" proves challenging as it refers to the surroundings, which vary based on the object of focus. Einstein once remarked that everything around us, excluding oneself, constitutes the environment. Environment poses a complex and multifaceted challenge impacting human existence. Man stands as nature's greatest asset and adversary. The necessity of industry for societal progress unavoidably leads to pollution. This interconnectedness implies that progress and pollution are inseparable, with no escape from pollution as long as progress continues. Pollution, stemming from the verb "pollutes," signifies making something unclean, impure, or morally tarnished. Moreover, it encompasses the direct or indirect discharge of substances or energy by humans into aquatic environments, posing hazards to human health, harming living resources, ecosystems, amenities, and interfering with water's legitimate uses.

The Kaushambi apartments' residents' welfare association (KARWA) has petitioned the Supreme Court to address various pollution sources affecting the township. The apex court, upon accepting the plea on October 15, 2020, directed issuing notices to respondents, including Uttar Pradesh agencies and

departments from Delhi. The township, situated near East Delhi's Anand Vihar border within Ghaziabad district's jurisdiction, comprises approximately 22 high-rise societies housing an estimated 20,000 residents. Karwa officials have filed three petitions with the National Green Tribunal (NGT) since 2015 to combat pollution in Kaushambi's vicinity. Asha Pushp Vihar Awas Samiti Ltd., Kaushambi, also joins as a co-petitioner. Karwa highlights pollution issues, including air, noise, and waste pollution from garbage dumps in Ghazipur and Kaushambi, along with traffic congestion at the Anand Vihar ISBT and a UP Roadways depot in Kaushambi.

The Delhi Air Pollution crisis gained national significance after an emergency declaration from November 6-10, 2017, drawing global attention through prominent news reports. This paper delves into analysing air pollution levels pre-and-post the emergency period, focusing on Anand Vihar over a year, considering monthly and annual AQI levels concerning key pollutants like NO₂, SO₂, PM_{2.5}, Ozone, and PM₁₀.

Anand Vihar, positioned at the New Delhi border with Uttar Pradesh, faces severe air pollution due to vehicular traffic, waste incineration, and industrial emissions. Pollution monitoring by the Delhi-based Centre for Science and Environment (CSE) revealed higher exposure to harmful gases than previously assumed. Anand Vihar registers among the highest air pollution levels in Delhi-NCR due to numerous pollution sources nearby. Sahibabad and Patparganj areas, adjacent to Anand Vihar, host several industrial units using various polluting fuels. Additionally, National Highway (NH) 24, close to the area, witnesses heavy traffic, contributing significantly to vehicular emissions and chronic congestion.

Objective:

1. Gather all relevant air quality data, including daily Air Quality Index (AQI) readings, average pollutant concentrations, and additional meteorological data.
2. Conduct thorough data analysis, statistical examinations, and modelling to extract valuable insights.
3. Forecast the average levels of specific pollutants for the next five years using the analysed data and statistical models.

Methodology

Overview

This chapter delves into the methodology used for analysing trends and predicting air pollutants, which is crucial for completing this dissertation. We discuss in detail the methods and approaches employed for trend analysis and prognosis of air pollutants. Additionally, we provide a brief overview of the different air pollutants considered in this study.

Research Method

The chapter also discusses the adoption of National Ambient Air Quality Standards (NAAQS) by the Central Pollution Control Board (CPCB), along with the guidelines and standards outlined by NAAQS. Table 1 presents the current NAAQS, while Table 2 displays the AQI Standards for different zones along with the pollutants' breakpoints.

Pollutants	Time weighted Average	Concentration in Ambient Air		
		Industrial, other areas	Rural,	Ecologically Sensitive Area
Sulphur dioxide (SO ₂),	Annual 24 Hour	50 80		20 80
Nitrogen dioxide (NO ₂),	Annual 24 Hour	40 80		30 80
Particulate matter (< 10 µm)	Annual 24 Hour	60 100		60 100
Particulate matter (< 2.5 µm)	Annual 24 Hour	40 60		40 60
Ozone (O ₃)	8 Hours	100		100
	1 hour	180		180
	24 Hour	60		60

Table 1: Present National Ambient Air Quality Standards (Units: µg/m³ Source: CPCB)

AQI Category (Range)	PM10	PM2.5	NO ₂	SO ₂	O ₃
Good (0-50)	0-50	0-30	0-40	0-40	
Satisfactory (51-100)	51-100	31-60	41-80	41-80	
Moderately polluted (100-200)	101-250	61-90	81-180	81-380	
Poor (201-300)	251-350	91-120	81-280	381-800	
Very Poor (301-400)	351-430	121-250	281-400	801-1600	
Severe (401-500)	430+	250+	400+	1600+	

Table 2: AQI Standards with Pollutants Break Points. (Units: µg/m³ Source: CPCB)

Study Area

The study area chosen for this research is Delhi, which holds significant global importance as the capital of India. With a population of 16.3 million people and covering an area of 1483 km², Delhi ranks as the 6th largest metropolis globally. However, it is also renowned for its high pollution levels, especially in terms of airborne particulate matter like PM_{2.5} and PM₁₀. These pollutants have been a major concern due to their adverse health effects, particularly PM_{2.5}, which is notably elevated in Delhi. The city has 20 air monitoring stations, one of which is Anand Vihar (AV), chosen as a representative area due to its dense population and heavy vehicular traffic, as depicted in the study's results.

Parameter Considered and Duration

For this study, we've considered five key air quality parameters based on the data availability for the study period. These parameters include PM₁₀, PM_{2.5}, NO₂, O₃, and SO₂, with their 24-hour average concentrations being analysed. These pollutants were chosen as they are significant indicators of air quality and have a notable impact on human health and the environment. The inclusion of these parameters allows for a comprehensive analysis of air pollution levels in the study area during the specified duration.

Pollutant

PM₁₀: These are granular particles ranging from 2.5 to 10 micrometres in size, including smoke, dirt, dust from industries, agriculture, and roads, as well as Mold, spores, and pollen.

PM_{2.5}: These are even smaller particles, less than 2.5 micrometres in diameter, visible only under an electron microscope. They come from sources like vehicles, power plants, urban wood burning, forest fires, and agricultural burning.

O3: Ozone is a pale blue gas with a strong smell, formed from dioxygen by UV light and electrical discharges in the atmosphere.

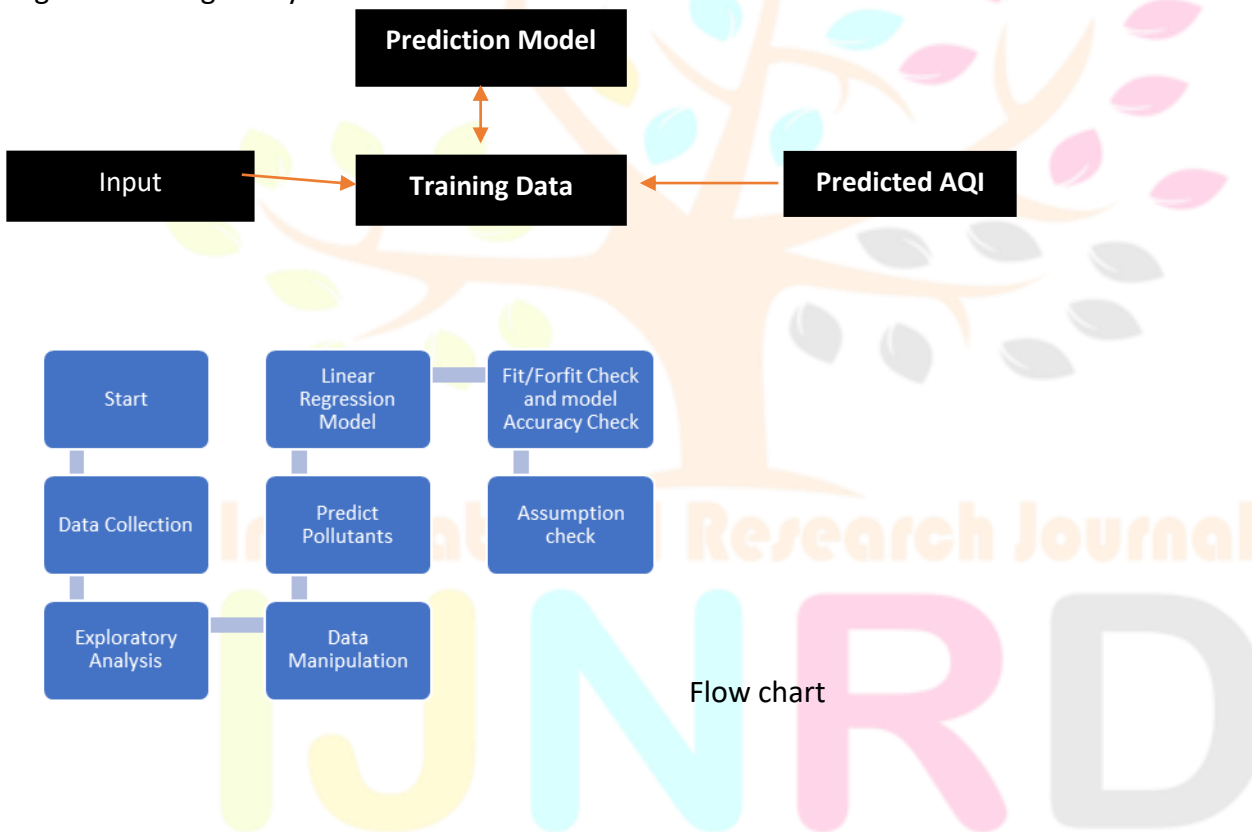
SO2: Sulphur dioxide is a colourless, irritating gas produced when sulphur-containing fuels like coal and oil are burned, often near large industrial areas such as power plants, refineries, and industrial boilers.

NO2: Nitrogen dioxide is part of a group of gaseous air pollutants generated from road traffic and fossil fuel combustion processes.

Design

This project is structured into distinct steps, each with specific components or modules assigned to accomplish tasks in a synchronized sequence. The objective is to predict the Air Quality Index (AQI) of a selected region with utmost precision. Users input their selection of the region and month on a map, and the system generates the predicted AQI for that month as output. This systematic approach ensures accurate and reliable predictions for air quality assessment.

High Level design of system:



Methods and Techniques involved:

The methodology encompasses various techniques for analysing and predicting air quality:

- 1. Descriptive Analysis:** This approach delves into understanding the fundamental traits of the data, offering insights into sample characteristics and measured parameters.
- 2. Predictive Analysis:** Employing statistical and machine learning methods, predictive analytics aims to anticipate future trends based on historical data. In our study, predictive analysis is conducted using time series regression forecasting.

3. Regression Forecasting: This technique involves studying how data patterns evolve over time through a model-based prediction of future values. Widely applied in diverse fields like weather forecasting and pattern recognition, time series regression forecasting plays a crucial role in predicting air quality trends.

Datasets

Months	PM2.5	PM10	NO2	SO2	O3
Jan-16	354.605	581.0767	98.01586	21.07767	17.27333
Feb-16	166.9387	281.6455	86.58696	27.28435	23.44565
Mar-16	119.91	269.1711	79.07053	26.33684	28.25895
Apr-16	148.6352	391.07	95.89478	38.10318	33.87227
May-16	100.8076	408.5776	74.68952	14.72619	27.96524
Jun-16	92.93833	328.1673	56.468	10.70733	24.50633
Jul-16	57.16033	198.089	48.206	10.926	23.44267
Aug-16	46.39103	183.2567	42.00633	10.63267	24.79571
Sep-16	71.14036	345.3325	55.28929	9.566429	30.5575
Oct-16	201.0361	569.2083	103.1971	24.65226	38.91774
Nov-16	333.8697	739.5405	137.3347	30.09667	34.837
Dec-16	294.6113	620.5817	90.73071	23.14032	16.70645
Jan-17	181.8468	400.0282	83.3384	28.3975	19.28714
Feb-17	137.4727	459.7377	103.2531	26.06962	14.79385
Nov-17	327.067667	401.720455	120.775	19.235	51.0833333
Dec-17	298.540667	495.925667	122.121667	23.8943333	30.4196667
Jan-18	261.6381	426.2216	130.6187	22.24065	23.38037
Feb-18	161.2682	348.2989	177.3379	17.015	20.08607
Mar-18	116.1843	325.1883	128.2713	21.74133	30.31692
Apr-18	115.366	317.751	104.17	11.7716667	43.3456667
May-18	104.3718	303.4245	95.77576	8.864242	36.58242
Jun-18	115.904	307.2532	47.43	9.3525	23.1192
Aug-18	45.26563	152.9863	18.515	10.78875	31.47313
Sep-18	49.56269	163.8446	60.0487	18.68231	30.64115
Oct-18	141.4129	391.32	123.6145	21.96645	35.92667
Nov-18	259.218	482.2857	148.893	15.51833	22.86633
Dec-18	280.6997	476.9142	132.4832	17.02839	32.00516
Jan-19	261.6177	397.3258	117.9583	10.64097	19.22742



Nov-19	277.9553	445.4417	109.1477	14.11467	20.889
Dec-19	210.6832	337.5006	140.7467	7.486667	53.724
Jan-20	214.601	335.3097	73.91548	10.05194	48.34774
Feb-20	124.7814	265.5552	71.95931	8.198621	25.71517
Mar-20	75.22839	180.4574	51.12129	9.770645	31.92065
Apr-20	43.72333	96.00167	20.98765	12.90588	68.316
May-20	74.43353	122.2053	35.99133	16.792	94.07412
Jun-20	62.51033	92.541	33.304	14.47667	85.08733
Jul-20	41.28484	79.97903	31.56806	11.51548	65.21613
Aug-20	29.13677	52.71839	49.44129	10.65129	26.63129
Sep-20	45.18933	104.1375	51.715	21.031	32.235
Oct-20	126.223	302.9216	81.63548	21.48968	33.72226
Nov-20	271.0213	451.1467	115.664	18.48333	31.725
Dec-20	228.9769	324.2983	82.666	15.998	24.60457

Approach to Prediction

The prediction approach for air pollutants involves several steps:

- 1. Data Collection:** Daily average concentration data is gathered for long-term forecasting. Daily data is sourced from CPCB, while monthly data is from DPCB. The collected data is organized in spreadsheets using MS Excel.
- 2. Graphical Representation:** The data in the spreadsheet is utilized to create graphical representations, showcasing the concentration of air pollutants over time.
- 3. Forecasting:** Using the collected data, forecasting is performed for air pollutants spanning from January 2021 to December 2025.
- 4. Analysis:** The forecasted data is then analysed to derive R2 values and formulate regression model equations for each pollutant.

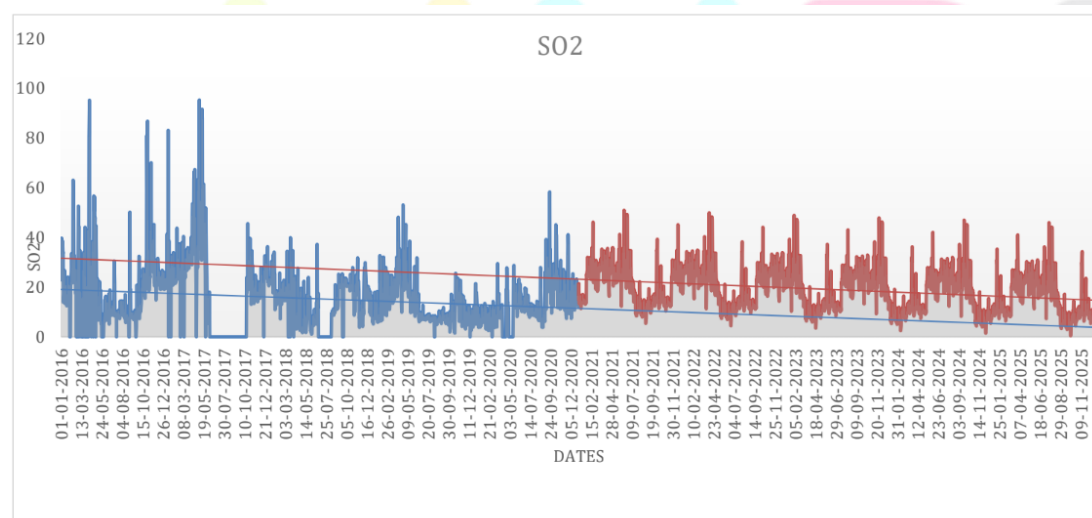
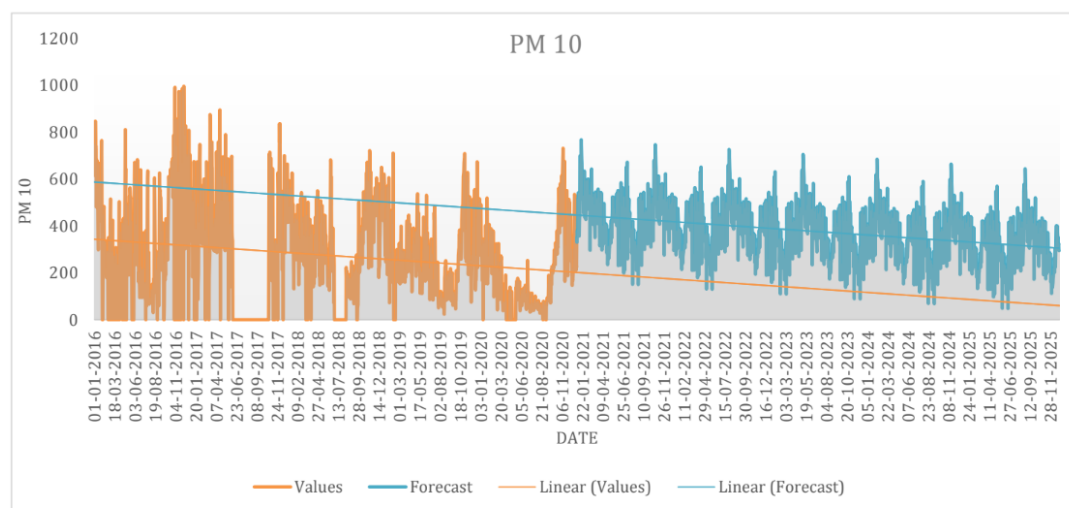
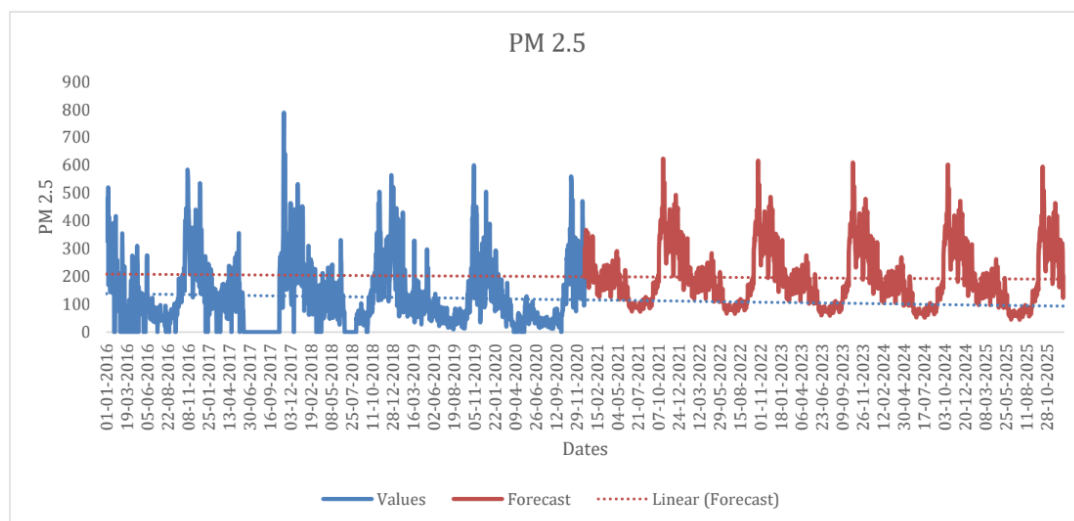
Result and Discussions

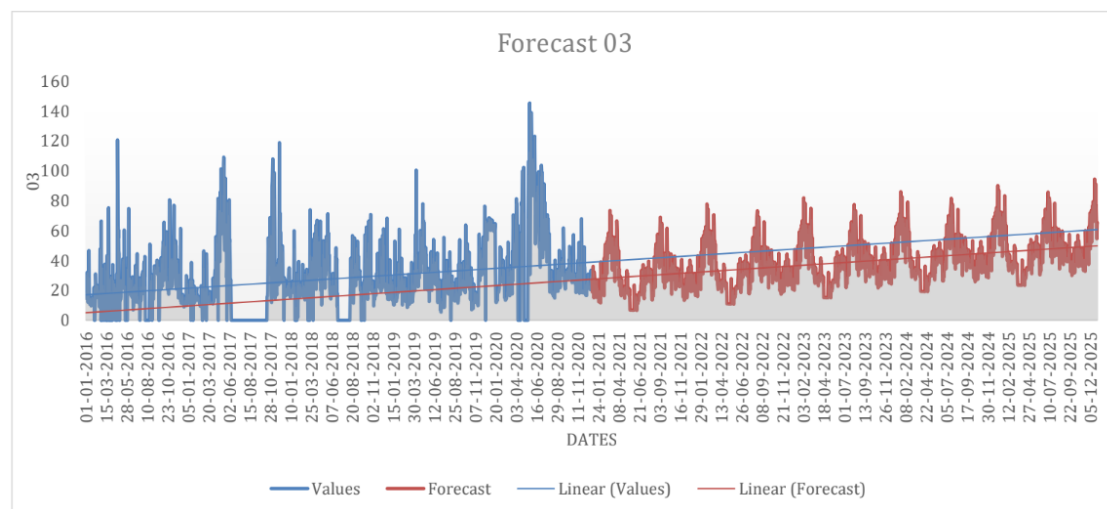
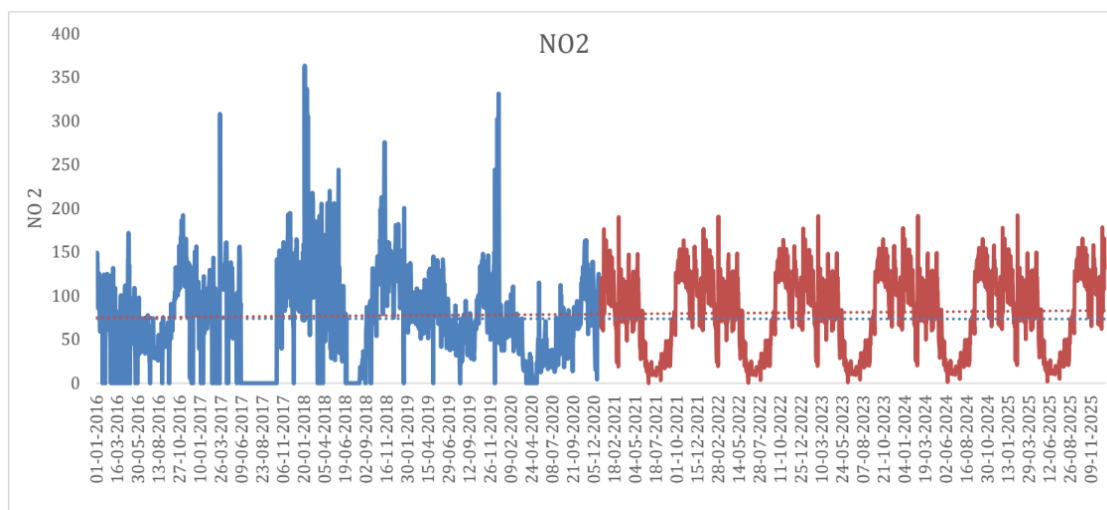
FORECASTING SO₂, NO₂, O₃, PM₁₀ AND PM₂.

In the forecasting process for SO₂, NO₂, O₃, PM₁₀, and PM₂, we treated years as independent variables and air pollutants as dependent variables, assuming a linear

relationship between them. The graphical representation below illustrates the actual versus estimated concentration of each air pollutant over the years:

(*All Values in ug/ m3)





This approach allowed us to analyse and visualize how the concentration of these pollutants' changes over time, aiding in understanding their trends and making informed predictions.

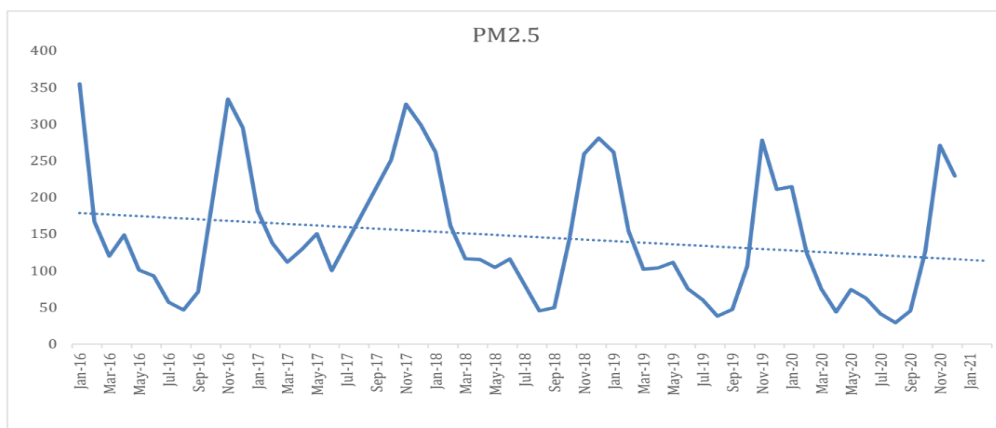
Forecasted Per Day Average Concentration of Pollutants.

Pollutant	Forecasted Values (ug/m3)
PM 2.5	199.97
PM 10	312.66
SO2	106.85
NO2	15.06
O3	45.64

Table 4: Forecasted Per Day Average Concentration of Pollutants

Linear Regression Modelling

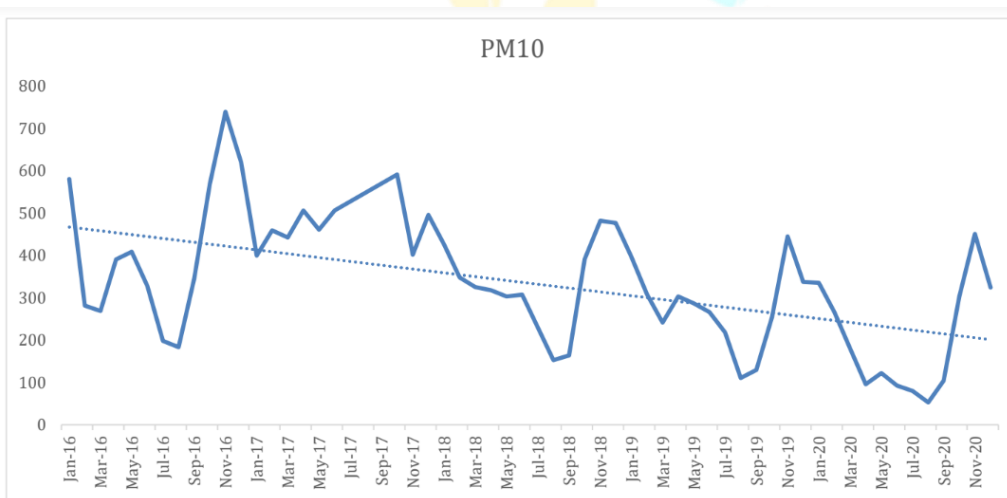
In our linear regression modelling, we analysed the relationship between years (independent variables) and air pollutants (dependent variables) linearly. The graph below shows the actual pollutant concentrations plotted against years, with the linear regression line for comparison.



Regression Equation:

$$[y = -0.035x + 1662.1]$$

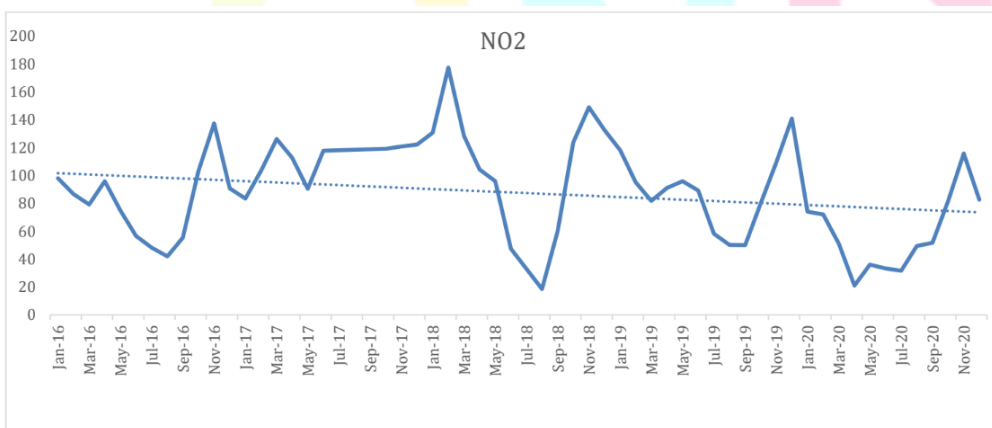
R-squared Value: 0.7302



Regression Equation:

$$[y = -0.1482x + 6746.6]$$

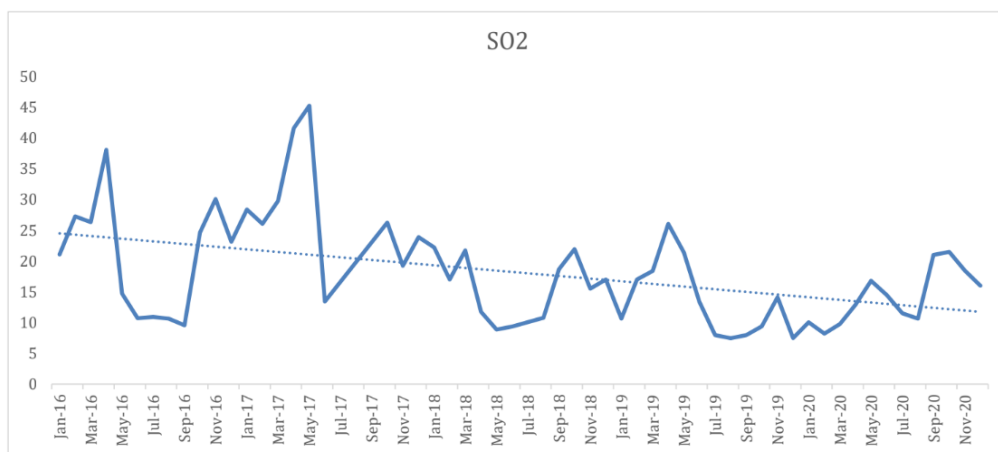
R-squared Value: 0.8222



Regression Equation:

$$[y = -0.0156x + 763.71]$$

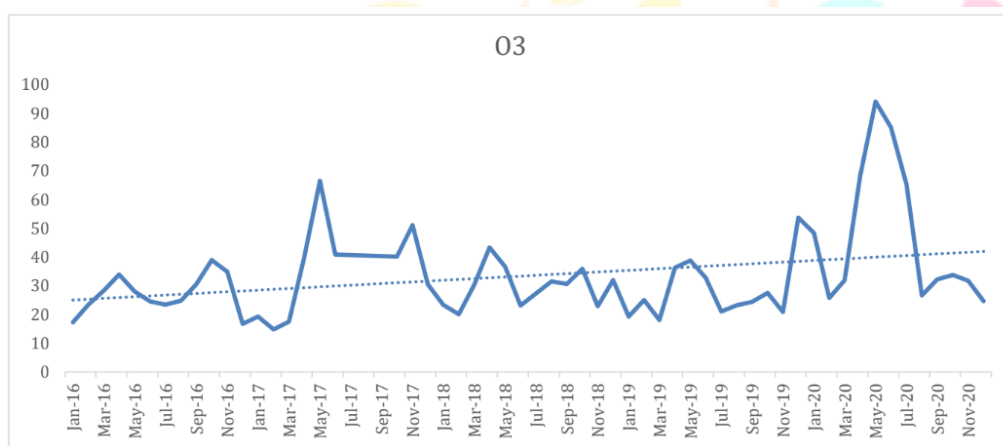
R-squared Value: 0.8601



Regression Equation:

$$(y = -0.0071x + 325.67)$$

R-squared Value: 0.8106



Regression Equation:

$$(y = 0.0094x - 374.14)$$

R-squared Value: 0.8171

3.3 Result at Glance

	PM 2.5	PM10	NO2	SO2	O3
Average Values	149.217183	327.494665	88.25484	17.9051638	33.112785
Forecasted values	199.971201	312.663164	106.854929	15.0626049	45.6490813
Predicted Change (%)	+34%	-4.6%	+21%		+33%

Table 5: Percentage Change in Pollutant Concentrations

	Linear Equation	R squared value	Accuracy level
PM 2.5	$y = -0.035x + 1662.1$	0.7302	Satisfactory
PM10	$y = -0.1482x + 6746.6$	0.8222	Good
SO2	$y = -0.0071x + 325.67$	0.8106	Good
NO2	$y = -0.0156x + 763.71$	0.8601	Good
O3	$y = 0.0094x - 374.14$	0.8171	Good

Table 6: Coefficient of Correlation and Regression Equation Obtained

Conclusion and Future Scope

Conclusion

This study focused on forecasting pollutants in the Anand Vihar Region of Delhi over the next 5 years, employing linear regression forecasting and data mining techniques. Parameters such as PM 2.5, PM 10, SO₂, NO₂, and O₃ were considered. The analysis indicated increasing concentrations of PM 2.5, SO₂, and NO₂, while PM 10 and O₃ concentrations are expected to decrease by 2025. However, PM10 levels may remain above permissible limits due to a gradual decrease. Public policy involvement is crucial for pollution prevention, and dynamic atmospheric variables pose challenges for long-range forecasting. Future verification using different techniques is recommended.

References

The Hindu. (n.d.). Most Polluted Cities In India mapped. Retrieved from <https://www.thehindu.com/news/national/polluted-cities-in-india-mapped/article23747106.ece>

The Hindu. (n.d.). Nearly 15000 premature deaths due to air pollution in Delhi study. Retrieved from <https://www.thehindu.com/news/national/nearly-15000-premature-deaths-due-to-air-pollution-in-delhi-study/article24419003.ece>

Indian express. (n.d.). Indians can live 4 years longer if air meets WHO norms. Retrieved from <https://indianexpress.com/article/india/indians-can-live-4-years-longer-if-air-meets-who-norms-pollution4839262/>

Central Pollution Control Board (CPCB). (2009). National Ambient Air Quality Standards (NAAQS) – 2009.

Central Pollution Control Board (CPCB). (2020). National Air Quality Index.

Central Pollution Control Board (CPCB). (2017). Annual Report - 2017.

Central Pollution Control Board (CPCB). (2000). Air quality status & trends in India.

Gunasekaran, R., Kumaraswamy, K., Chandrasekaran, P., & Elanchizhian, R. (2012). Monitoring of Ambient Air Quality in Salem City, Tamil Nadu, India. *International Journal of Current Research*, 4, 275-280.

Karuna, M. S., Upadhyay, O. P., Saxena, D. K., Swamy, M., & Majumder, S. (2017). *International Journal of Current Research*, 9(12), 63243-63255.

Mohan, M., & Kandya, A. (2007). An Analysis of The Annual and Seasonal Trends of Air Quality Index of Delhi. *Environmental Monitoring and Assessment*, 131, 267-277.

N.T., M., & Reddy, K. S. (2016). Ambient Air Quality Analysis Using Air Quality Index – A Case Study of Bangalore City. *International Journal of Innovative Research in Science, Engineering and Technology*, 5(12), 21319-21323.