Use of high resolution Aerosol Optical Depth for identification of fine particulates and local level pollution sources in ambient environment of Delhi

Dissertation

SUBMITTED TO

CENTRAL UNIVERSITY OF HARYANA, MAHENDERGARH, HARYANA

FOR THE AWARD OF THE DEGREE OF MASTER OF SCIENCE

IN STATISTICS

INTERNAL SUPERVISOR:
Dr. Anoop Kumar
Assistant Professor

EXTERNAL SUPERVISOR:
Er . Sunayana
Senior Scientist

SUBMITTED BY: Kailash

DEPARTMENT OF STATISTICS,
CENTRAL UNIVERSITY OF HARYANA,
MAHENDERGARH, HARYANA - 123031

JUNE 2024

DEPARTMENT OF STATISTICS

CENTRAL UNIVERSITY OF HARYANA MAHENDERGARH

Declaration

I hereby declare that the work presented in this Dissertation entitled, "**Use of high resolution Aerosol Optical Depth for identification of fine particulates and local level pollution sources in ambient environment of Delhi**", for the partial fulfilment of the requirements of the award of the degree **Master of Science in Statistics** and submitted to the Department of Statistics, Central University of Haryana, is an authentic record of my work carried out in IV semester under the supervision of **Dr. Anoop Kumar**, Assistant Professor, Department of Statistics, Central University of Haryana.

I have not submitted the matter presented in this project report for the award of any other degree/diploma of this or any other University/Institute.

Kailash
Roll No: 221026
Signature of the candidate

This is to certify that the above statement made by the candidate is correct to the best of my knowledge.

Dr. Anoop Kumar
Supervisor

The M.Sc. dissertation Viva-Voce Examination of Mr. **Kailash**, Roll No. **221026** has been held on June 2024

Head of the Department

सीएसआईआर
राष्ट्रीय पर्यावरण अभयांत्रिकी
अनुसंधान संस्थान
दिल्ली क्षेत्रिय प्रयोगशाला
सी.एस.आई.आर., आर एण्ड
डी_ सेंटर, ए - 93-94, फेस -
1 नारायणा औध्योगक क्षेत्र
नई दिल्ली - 110028

प्रेषक: वरिष्ठ वैज्ञानिक
From`: Senior Scientist

CSIR
National
Environmental
Engineering
Research
Institute

Delhi Zonal Centre
CSIR R&D Centre
A - 93-94, Phase - 1
Naraina Industrial Area
New Delhi – 110 028

Date: 31/05/2024

CERTIFICATE

This is to certify that dissertation entitled, "**Use of high resolution AOD for identification of fine particulates and local level pollution sources in ambient environment of Delhi**" which is submitted by **Mr. Kailash** Roll Number 221026 in partial fulfilment of the requirement for the Degree of Master of Science in Statistics, Central University of Haryana, Mahendergarh is a record of the work carried out by the candidate under my supervision at CSIR-NEERI, Zonal Centre, New Delhi during 13th February, 2024 to 19th May 2024.

सुनयना / Sunayana
वैज्ञानिक / Scientist
सीएसआईआर–नीरी क्षेत्रीय केंद्र / CSIR-NEERI Zonal Centre
ए-93-94, नारायणा औध्योगिक क्षेत्र / A-93-94, Naraina Indl. Area
नई दिल्ली–110028 / New Delhi-110028

Signature and stamp
External Supervisor/Guide

Acknowledgments

Firstly, I want to show my heartfelt gratitude to everyone who has contributed in completion of my dissertation. I am thankful to my internal supervisor Dr. Anoop Kumar, Assistant Professor at Central University of Haryana, department of statistics, for his continuous and invaluable support of my project work and research. He has been an incredible guide, offering invaluable support and mentorship throughout my journey. His expertise and guidance played a important role in completing this dissertation. I could not have imagined having a better advisor and mentor for my dissertation.

I am also thanful to Dr. Ranjan Kumar Sahoo sir, Head of department of statistics at Central University of Haryana for providing me this opportunity to undertake my dissertation. His support throughout this journey is invaluable. Besides, i would like to thank the rest of my statistics faculty: Dr. Ravinder Singh, Dr. Pavitra Kumari for their questions and insightful comments.

I also want to thank my fellow classmates, all my family members and all my friends for their great support and always encourage me throughout this journey. Their love and support towards me throughout this period of dissertation encourage me a lot. Lastly, i would like to appreciate all the individuals, institutions and organizations whose directly or indirectly contributions have been vital to the completion of this dissertation.

Date:
Place:

Kailash
Roll No: 221026

Certificate regarding Originality of Research Work
(Dissertation)

Name of Student**: Kailash**

Title of the M.Sc. Dissertation: **"Use of high resolution Aerosol Optical Depth for identification of fine particulates and local level pollution sources in ambient environment of Delhi**

Name of research supervisor: **Dr. Anoop Kumar**

Designation & address of supervisor: **Assistant Professor, Department of Statistics**

I/ We hereby declare that, the above mentioned research work is original & it doesn't contain any plagiarized content. The similarity index of this research work is **8%.** The plagiarism check report by supervisor's account of the thesis is attached herewith.

Declaration by Student

Further, we declare that, the contents of this M.Sc. Dissertation submitted to, The Department of Statistics, Central University of Haryana for award of M.Sc. Statistics Degree are original findings & Research work.

i) If any plagiarism is found in this dissertation I/we shall be solely responsible for it.
ii) This work has not been submitted for the award of any other Degree / Diploma in any other University / Institutes.

Signature of the Student

Verified by

Signature of the Supervisor

Date: Place:

Forwarded by the HoD :
Department Of Statistics
Central University of Haryana

Abstract

Air pollution is the main problem for many cities around the world. India's mostly big cities are currently facing the pollution crisis. Air quality management has a critical challenge in cities like Delhi facing concerns due to high population, industrial activities, vehicular emissions and geographical factors of landlocked from all sides increses the air pollution lavels. To moniter the air quality pollution patterns, high resolution MODIS (Moderate resolution imaging spectroradiometer) products are being used. This dissertation investigates the use of integrating high resolution MODIS products into ground lavel concentrations of Delhi. MODIS provides a large range of datasets on atmospheric pollution factors such as aerosols optical depth, particulate matter concentrations, land temperature and various other atmospheric particles. The study aims to establising the relation between the Aerosol Optical Depth (AOD) values with the ground values concentrations and study the impact of stubble burning in neighbouring states on black carbon concentration. The study use methods for integrating MODIS data with existing ground based black carbon data and satellite observations of AOD to create a spatiotemporal dataset of air quality indicators. Machine learning algorithms and various other statistical models are used to predict ground based Carbonaceous aerosols pollutant concentrations, identify pollution sources and impact of meteorological factors on air quality.

Contents

List of Figures

List of Tables

Chapter 1

Introduction and Objective

1.1 Introduction

1.1.1 Overview

Air quality indicates the degree to which the air is suitable for human beings or the environment conditions. Good air quality refers the air which is free from all the harmful substances or these harmful substances is in low concentration. Air pollution means the pollutants particle into the air that affects human health and the environment. Poor air quality and human health are closely related to each other. According to WHO, air pollution is the single largest environmental health risk factor that killing up to 12% people across the world. Air pollution is the biggest problem in all countries, whether it is developed or developing. There are many health problems arising due to bad air quality such as mild allergic reactions, heart diseases, respiratory diseases and even premature death. In addition, its effects on human health the air pollution can also harm ecosystem, damage vegetation and contribute to change in climatic factor. Urbanization, industrialization, forest fires, stubble burning and the number of vehicles are increasing very rapidly, causing common air pollution. In India, almost all big cities face problems due to pollution. The management of air quality in advanced cities is a major problem, and cities are facing progressive deterioration in ambient air quality attributable to increasing pollution. Air pollution refers to the presence of chemical, physical, or biological contaminants in the atmosphere which are harmful to human health as well as the environment. Atmosphere includes various gases in the air including 78% nitrogen, 21% oxygen, and 1% other gases like hydrogen, methane, carbon-dioxide, carbon-monoxide, noble gases (like helium, neon, argon, krypton, xenon), etc. These gases are present in balanced concentrations in the atmosphere. If the concentration of these gases exceeds, it leads to pollution. The sources of air pollution can be natural (such as wind-blown dust, forest fires, volcanoes) as well as human-related (such as power plants, industries, agricultural activities, and vehicles). Air pollution being the topmost environmental killer in the world that kills over 17 billions of people worldwide.

Air pollution generally takes place due to the combustion of fossil fuels. Gases, oil and coal are used to produce energy and release CO, NO into the air at a high level. Industrial activities mainly emits a lot of pollutants in the air and the key pollutants that are emitted from industries as particulate matter, NO2, SO2, O3 and CO. Wildfires, stubble burning are also the major contribution in increased $PM_{2.5}$ in the atmosphere which combines with other harmful gases and create smog, makes the air hazy to breathe. People find it difficult to breathe. Transportation system, openly burning of garbage waste and other substances, construction of road and buildings and demolition, all the agricultural activities, use of chemical products and synthetic products and many more to say, these are the major sources of causing air pollution. In past a number of tragedy happened due to air pollution for example Bhopal gas tragedy in 1984, great smog of London in 1952, Kuwait oil fires in 1991, Donora smog in the US in 1948 etc. These are some major air pollution disasters of all time in the world led to many casualties and impacted the environment.

India is among the top polluted country in the world. As India is a large populated under developing country due to which the pollution is a major concern for them. Air pollution in India is approximately killed about more than 1 million people every year. India has the world's highest mortality rate due to the diseases like respiratory diseases, cardiovascular diseases and asthma. In Delhi the capital city of India, poor air quality damages the lungs of millions children. Delhi's pollution crisis is also caused by the production of agriculture waste that are burnt to clean the fields by the farmers which results in smog production in air and other harmful particles are produced and make the air polluted. This happened mainly in crops production session of the year as Delhi is closely linked to the cities of Haryana and UP which already have very high Air quality index. A study conducted in 2016 to measured the sources of pollutants like aerosols optical depth, other particulate matter concentrations and average levels of various type of air pollution in Delhi of $PM_{2.5}$ pollutants, 40% came from the dust on the road, 20% to vehicles fuel burning emissions and rest from domestic fuel burning and various other industrial productions waste. Sources of PM10 pollutants, 60% came from dust on the road, 10% from concrete buildings and roads, 10% from industrial productions waste and 10% from vehicles fuel emissions and rest from industries, construction sites.

Children's population are generally more effective to the negative impacts of air pollution as their organs are in developing state which means that their lungs are so sensitive to these harmful pollutants. They mostly spend time outside the home due to which they are more exposed to its effects. Particulate matters caused the most premature death and various health related problems as per data. Air pollution also affects the lifestyle of human being as you can see there a difference between the lifestyle of rural and urban peoples. Mainly cities are affected more as compare to other area because cities are more developed, industrialized, overcrowded and many of the persons have their personal vehicles. Cities have all the big reasons of air pollution. In order to reduce the air pollution many strategies and technologies are available morever many countries have their certain law for air pollution. For some pollutants such as black carbon concentration increased because of large number of vehicles. Air pollution also the reason in changes of economic growth of the country. The problem is even worst in developing countries. Using of public transport instead of personal vehicles is a good choice in controlling air pollution but it totally depends on people's prospective. Delhi government's odd-even scheme also an example of controlling air pollution caused due to vehicles.

1.1.2 Objectives

The main objective of this study is to understand the effect of stubble burning in neighbouring states on AOD values and Black Carbon concentration. The objectives of this study are as-

1. Analysis of AOD values at 1 km resolution from MODIS for Delhi in different months.
2. Analysis of black carbon surface mass concentration from MERRA-2 and variation due to fire in neighbouring states during stubble burning period.
3. Developing a statistical model for prediction of ground level BC concentration

1.1.3 Study Area

Our region of interest for studying air pollution effects is Delhi which is the capital city of India located in the northern central region. It is among one of the most populated city and polluted city in the world. The capital city is divided into two sections Old Delhi and New Delhi. Delhi lies within the coordinates of 27.6581°N latitude to 29.113°N latitude and 76.2781°E longitude to 78.0249°E longitude covers the total area of about 1484 square kilometers. Delhi consists of 11 districts named as: North Delhi, North-West Delhi, West Delhi, South-West Delhi, New Delhi, South-East Delhi, East Delhi, Central Delhi, North-East Delhi and Shahdara. Total population of Delhi is about 18.7 millions as per census done in 2011. The region is known for its commercial activities, economic activities, a big transportation network, and dense population. Delhi is very small as compare to its population and a mix of land area uses like residential areas, industrial zones, agricultural land, commercial areas, schools, colleges etc. These all activities makes the capital region condition worst to live due to excessive pollution. Vehicles are the major source of pollution but heavy deforestation for development is the main problem that Delhi faces. Delhi airport is the largest airport in India with a network of large number of flights. These airplanes and jets emit large amount of fumes in the sky of Delhi.

Delhi is a landlocked city, shares its border with Haryana and Utter Pardesh. The districts of Haryana and UP that share their border with Delhi are part of the National Capital Region known as NCR. Delhi landlocked between heavily industrial Districts of Haryana and UP. These industries release heavy fumes and different types of chemical into the air that travels into city. The agricultural activities like stubble burning in neighbouring states turned Delhi into a gas chamber city. Rajasthan is possible the cause of large dust that enter into the city through air. There is no sea near Delhi thus the pollution remains there. Stubble burning and other fire activities happened in or around Delhi also affects its environment.

Further in our study, we established a relationship between the pollutants like BC MERRA-2 and number of fire counts in Delhi and its neighbouring states during stubble burning period to conclude the effects. Our study area is mainly for Delhi region but we also study the sub region named as IITM Janakpuri, which is the part of capital city Delhi. The landscape around the region of IITM, Janakpuri, Delhi includes green belts, undeveloped areas, vacant land, commercial areas, institutional areas, various types of utilities and a large number of recreational areas. Generally, Janakpuri is known for its green belt, forest areas, wide roads and a numbers of parks and gardens. These wide roads can't able to reduce the traffic congestion. Still like other parts of Delhi, IITM, Janakpuri also faces air pollution challenges. Vehicular emission is one the main factor in releasing the air pollutants along with construction and industrial activities.

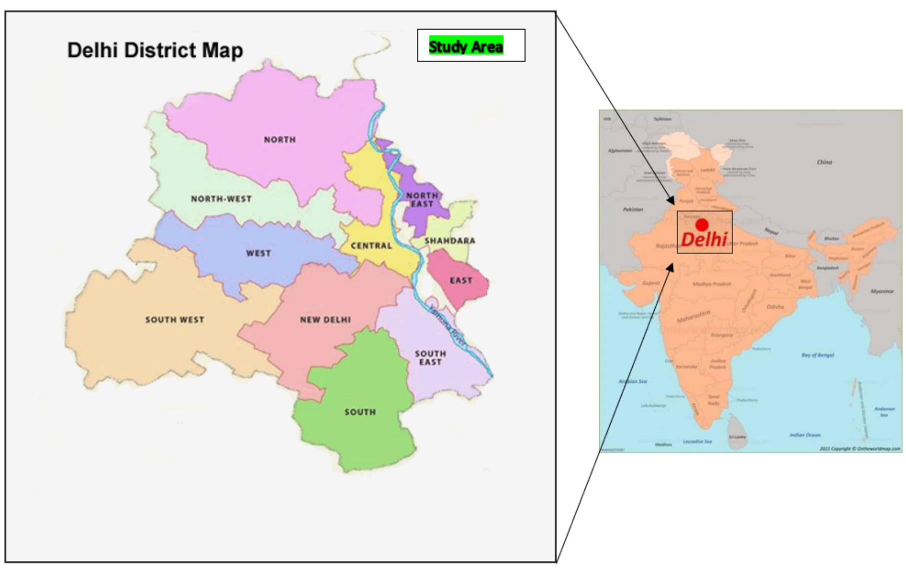

Figure 1.1: Study Area

Chapter 2

Review of Literature

2.1 Review of Literature

Air quality is important factor of environment and our health. Pollutants like particulate matter, black carbon, aerosol values having a huge impact on the human health, environment and climatic factors. In recent year, various satellite remote sensing programme or technologies developed and various studies has been done in hope finding the impact of different air pollutants on human health and environmental factors. These literatures study provide a overview of the methods in accessing air quality and focus on the measurements of particulate matter pollution using satellite remote sensing techniques.

- The paper published by Soni et al. (2014) provides us information about the analysis of aerosol satellite values distribution and its trends over the region of Gangatic Himalayan by using long term MODIS observation and ARIMA model. This paper highlights the impacts of aerosol values on atmosphere, health and climate change. It shows the increasing trend of aerosols over Asia due to various factors as biomass burning, industrialization and rapid increase of population in Asia. The main focus of the author is on Indo Gagantic plain which is located in northern India, a highly polluted land on the Earth. Various studies conducted in the Indo Gagantic plane for the increase in the level of satellite aerosols. These studies conducted in the Himalayas of central and western Indian as well as in Nepal to understand the importance of aerosol characteristics in these regions. The authors use time series model as a tool in analyzing climatic fluctuation on ARIMA model. They discuss about use of ARIMA model to predict the future trends and variance in aerosol.

  Previous study on that region basically focused on the air pollutants as NO2, SO2, CO and O3. They used MODIS datasets at 1 km high resolutions monthly AOD at 550 nm during period of March 2000 to May 2012 and level 3 monthly data was produced from Giovanni. ARIMA model used in predicting the future trends of AOD at 550nm. By following these steps, seasonal mean values of MODIS AOD at 550nm are summarized and largest values of seasonal AOD's are observed during the mansoon period due to high biofuel and biomass burning. During the pre mansoon period, the trends was tends to be seen decreased. During winter, the lowest values of AOD at 550nm was recorded due to minimization of transportation dust. By summarizing the paper, we found that the past AOD values have the influence on the future AOD values as we predict the values using the statistical methods.

- The article published by Majumdar et al. (2023) explore the spatial pattern and temporal pattern of biomass burning and BC concentration during the summer seasons from 2015 to 2021 in India. First of all, they study about the spatial distribution of biomass burning which were concentrated mostly in Northern, Central, North-Eastern and Western India with April having the most extensive concentration and there's a decrease in May-June period. The various parts of India received air mass movement from these fire clusters regions. As a result, black carbon surface mass concentration were highest in April and decreased in May-June corresponding with the decrease in fire incidents. Fire data of active biomass fire were retrieved by using the visible infrared imaging radiometer suite method. The region of interest India's fire data were retrieved during the period of 2015-2021. HYSPLIT is a model which used to find the air trajectories of different part of India. This helps us in understanding the movement of air mass that carrying Biomass burning of smoke and other air pollutants that harms the environment and healths. Now to analysis the black carbon surface mass concentrations, they used the MERRA-2 analysis data of aerosols. Data retrieved from the MERRA-2 analysis diagnostics over the period of 2015 to 2021 was first converted into the monthly data. Finally, the monthly average black carbon surface mass concentration data were studied over India for April-June summer period. This study examined the spatial pattern and temporal variation in biomass burning data over India's landmass using the MERRA-2 product for analysis. The study identifies several factors contributing to fires in summer season such as wheat stubble burning in Eastern, Northern and Central region of India, cultivation of crops, Forest fires mainly in the region like Chhattisgarh, Madhya Pradesh, Odisha, Maharashtra and many more. These all factors causes the burning of biomass and in this paper, the result comes out to be as impact of biomass burning on air quality and impacts on other pollutants like black carbon, particulate matter, CO and SO2. The author highlights the importance of the planetary layer in influencing spreading and dilution of the air pollutants particles. The literature provides a detailed analysis of the factors affecting the summer period's biomass burning and its impact on quality of air and pollution level, the need for further study and policy to reform this environmental issue.

- The paper by Sun et al. (2017) for the temporal difference and spatial variation in the satellite data values of aerosols in Beijing-Hubai, China with 1 km AOD MODIS product. Beijing-Hubai is one of the most polluted region of the China. This study comprises the use of high resolution data of aerosols optical depth values that are taken from MODIS datasets. This study highlights the importance of spatial difference and temporal variations in the values of AOD concentration and increase of particulate matter pollutants in Hubai regions. This study builts upon the previous studies in which focused on only aerosol monitoring and analysis using various techniques. It highlight the problem of taking ground measurement in spatial distribution of aerosols and focused on the need for continuous monitoring over large areas. The conclusing study of this paper is the monthly AOD was detected in the region from the year of 2013 and 2014 for which get the high values from February to June and get a low AOD values from September to January. This shows a increasing trends in the summer and spring season. In spring season, stubble burning, sand dust weather occurs which increase the AOD values. In winter season, snowfall and rainfall decrease the AOD values.

  For the spatial distribution of the region, we found that northern suburban area has the lowest aerosols optical depth values and southern has the maximum aerosols optical depth values. The main reasons behind the spatial distribution of aerosols optical depth values is the urbanization and sub-urbanization in region of study site. The AOD values for Beijing Hubai site with high spatial variation and temporal resolution was taken from the MODIS LIB data from the year of 2013 and 2014. This paper conclude that the seasonal variation of AOD values taken as temporal distribution was obtained by similar way spatial distribution region AOD values retrieved which means that AOD values in Hubai, China was affected by the climate conditions and the land form. This paper highlighting the importance of spatial and temporal resolution of aerosol data (AOD values) for environmental research and management. Many studies were considering on the retrieval process of AOD values and spatial, temporal distributed values from the MODIS data. By summarizing, we found that there is a relation between the spatial, temporal resolved distribution and the AOD values.

- The atmospheric article published by Sathe et al. (2019) on Black carbon emission and its impact from burning of biomass fuel and other anthropogenic sources in New South Wales, Sydney, Australia provide us extensive review related to the surface black carbon emission arises from the burning of biomass. This study highlights the main sources of BC emission as biomass burning occurs due to burns and bushfires, anthropogenic activities includes vehicular emissions and domestic fuel burning. Black carbon is a important component of pollutants particle like particulate matter $PM_{2.5}$ and a by-product of incomplete burning of organic materials. It causes a severe health problem like respiratory, cardiovascular etc. and more dangerous for those who already have these types of problem currently. Morever it also have impacts on climate changes by causing greenhouse effects. These all the things make this component to monitor and study to mitigate its emission. For collecting the data, NSW department installed 9 aethalometers in Sydney from 2014 to 2019 to measure black carbon level in the air and provide a dataset for analyzing spatial and temporal resolution of black carbon concentrations. Along with this data, trajectory modeling tools like HYSPLIT used to trace the sources of dispersion of black carbon and known as hybrid single particle lagrangian integrated trajectory, helps us to determining the sources of high black carbon concentration by studing the air particle trajectories and also distinguishing the difference between local and distant sources. To compare the satellite based data and ground base data, MERRA-2 data of reanalysis used to generate high resolution mapping of black carbon emission from both of the sources. This analysis reveals seasonal patterns in black carbon. In winter, BC level increased due to domestic wood burning and dispersion of it due to atmosphere. In summer, BC lavel increased due to the increase in the number of bushfires. In Sydney, there are less domestic wood burning during summer but more bushfires cases detected. Generally, urban area are more resoponsible for anthropogenic bc emissions. Also, the regression analysis between the BC and $PM_{2.5}$ shows that this relationship varies from season to season and by location. By integrating the advanced modeling tools and datasets, this study provides us the valuable information about the spatial variation and temporal dynamic structure of black carbon pollutants. This relation is used to mitigate the BC emission and the health problem regarding it. The health impact of black carbon on human being is highlighted due to its hazardeous results and high risk than the effects of pollutants matter like $PM_{2.5}$.

- The paper published by Srivastava et al. (2014) provide a a satellite view of aerosol characterstics over Delhi from and MODIS multiple angled imaging spectroradiometer (MSIR). This study includes the information about the aerosol optical depth, aerosol vertical distribution and their impact on the climate change or various environmental factors. The authors provide imformation that both MODIS and MISR products capture the trends, seasonal variation of AOD over region of interest Delhi. The seasonal variation presents in the aerosols optical depth values from both the satellite mission is the significant finding of the study. This study helps us in understanding the difference in seasonal variation in the values of aerosols optical depth and the difference in the values during the pre mansoon and post mansoon period or seasons. The authors also highlights the presence of very fine dust particles that dominants the aerosol values during all the seasons throughout the year. The study highlights the importance of considering aerosol composition and vertical distribution in estimating aerosol accurately. The authors discuss about the inclusion of vertical distribution for the climate modeling and representation of aerosol retrieving values.

Srivastava highlights the inclusion of dust that transported dessert area in pre mansoon season along with some anthropogenic activities which makes the air polluted. Such mixing process generally used in the estimating the DRF. In other seasons, local dust made the air polluted but not much as compared to the pre mansoon season. The seasonal cycle of variation was captured by both the satellite mission named as MODIS and MISR. Aerosol vertical level over the Delhi region has a strong seasonal variation where aerosol layer are below 2 km during post mansoon season and expand upto 6 km in the other seasons. This elavated aerosol layer over Delhi throughout the year is mainly due to the availability of dust particles in the air. Seasonal variation in day and night is same during a season but the aerosol layer generally at lower scale height during night as compare to day. Non-spherical dust particles are found to be more dominating the aerosol layer throughout the year. It provides a valuable insights of aerosol process over the Delhi and highlights the role between natural and anthropogenic aerosols and their influence on climate process.

- The paper published by Kumar et al. (2007) study the relationship between aerosol optical depth values that are driven from the satellite data and fine particulate matter that are obtained from ground level in region of Delhi. The main aim of this study is to establishing the relationship between the aerosols optical depth which is satellite based air quality and the particulate matter that obtained on the ground or Earth's surface in ROI Delhi region. Its aim to find whether the satellite based aerosols optical depth values can predict the ground based data at a high resolution of spatial and temporal distribution. The satellite based AOD data was collected from Terra MODIS. Ground was was computed using the 113 study area sites that are made in Delhi region to collect the ground based particulate matter that affacts the air quality. Then, the relationship was established between them by using the various statistical models which includes regression analysis, time series trends analysis etc. There is a significance difference between the data collection as we know that the satellite data collected from Terra MODIS at different spatial and temporal resolutions such as (1 km , 5 km, 10 km pixels). Since the particulate matter data are collected at points locations from the 113 sites made to collect air quality data. The main process of this study is to integrates these datasets by using different types of approaches to access the association between them at different geographic scales.

  This analysis reveals a significant positive association between AOD values and the particulate matter values. There are few similarities between these two values of air pollutants, this indicates that satellite based AOD values can be used as a indicator for the ground based particulate matter values. This relation is affected by various factors such as wather conditions, seasonality, vegetations, water bodies etc. leading some association or variability between them. Since, sources of air pollution very regionally can influence the AOD values and the relationship to the particulate matter values. This type of relationship between satellite based aerosols optical depth values and particulate matter can be used to predicting air quality for back years and allowing us for the examination of temporal variation and spatial distribution of air pollution and its impact on public health, environment and change in climatic conditions. This literature helps us in understanding how we can use remote sensing satellite data for estimation of ground data values like particulate matters in the regions with limited sources of monitering infrastructure. It highlights the importance of integration two different datasets and established relationship by using statistical methods. Further, the main aim for future study is by using these statistical relations to predict the health exposure due to air pollution.

- The research article by Lin et al. (2016) investigates the use of high resolution dat of aerosols from the satellite mission MODIS for the study of urban air quality studies. Its focusing on establishing relation between aerosol optical depth and ground level $PM_{2.5}$ concentrations. Air pollution is generally due to the concentration of AOD values and $PM_{2.5}$ values in air causes health risks. Various satellite missions for high resolution data was launched to get a better quality of data such as MODIS , MAIAC algorithm etc. The multi-angle impletation of atmospheric correction known as MAIAC algorithm was developed to provide high resolution data of aerosols optical depth values such as a finer resolution of 1 km as compared to the MODIS data. This algorithm provides better accuracy over the urban areas by removing the contamination. MAIAC data has a better correlation with ground level particulate matter values or concentrations as compare to MODIS data known as MOD04 because its temporal resolution of data for region is 10 km while MAIAC has a finer resolution of 1 km. Therefore, in estimating a better result, mixed effects model was used to predict the ground based $PM_{2.5}$ concentration based on the MAIAC aerosols optical depth values and MODIS aerosols optical depth values. These relationships was used to predict $PM_{2.5}$ ground level concentrations in the study area. The use of cross validation was a good decision in predicting $PM_{2.5}$ concentrations. This method can apply to other regions by adjustiing the variability in AOD $PM_{2.5}$ relation.

  They investigate the relation between aerosols optical depth values and $PM_{2.5}$ daily average values from the region sites for time period of 2002 to

2008 and made a direct comparison between the MAIAC and MOD04. By using the same data and applying regression analysis to find the seasonality in the data. The coefficients of determination for both MAIAC and MOD04 are almost same which shows as a good alternative of $PM_{2.5}$. Previous studies has shown that the relation between satellite based aerosols optical depth values and $PM_{2.5}$ are seasonally or by location but in this study, they used MAIAC and MOD04 to study the same effects on the $PM_{2.5}$ values. This shows that the spatial resolution of aerosols optical depth values affects the correlation factor between the satellite based aerosols optical depth values and the ground lavel based particulate matter $PM_{2.5}$ values. A comparison was done between MAIAC 1 km and MOD04 10 km AOD values which results as MAIAC 1km AOD provide a better correlation with the ground $PM_{2.5}$ values.

- The research paper published by Nair et al. (2020) highlights the complex issues of air pollutions in the region of Delhi, India where the air quality reach at its alarming states. The paper starts by highlighting the impacts of air pollution on human health, climate change and environment due to the urbanization, transportation issues, industrialization and many more. Delhi is among the most popular cities in world morever the capital city of India suffers with huge amount of air pollution surpassing the limits set by World Health Organisation on the daily basis. The main points highlighting in this paper which contributes in Delhi's poor air quality includes vehicles emissions, industrial activities and stubble burning in Delhi landlocked states. There's a lot of negative impact of this air pollution on human health like, respiratory problems, cardiovascular problems, and cancer among the population. The various steps has been taken by the Delhi's government to reduce air pollution but these are insufficient against the rising pollution lavels and demands for standard of livings.

  Stubble burning is a significant factor that affect the air pollution and its concentration on air. Stubble burning is done by the neighbouring states of Delhi as Punjab and Haryana because it is landlocked by them. The main aim of this study is to understand the impact of stubble burning emissions on air quality of Delhi using the various methods such as satellite based aerosols optical depth data, daily ground based covering of pollutants concentration and their analysis. This study analysis the fire radiative power, ground based pollutants concentration and the spatial variations and temporal distribution of satellite aerosols optical depth values with high resolution. This study finds that the high aerosols optical depth values during time period of stubble burning from Punjab and Haryana region. The smoke created from stubble burning is over the heights 2 km above from the Earth surface, hence the ground based data is neglected in this and stubble burning directly affects the satellite aerosol optical depth values. This study also concludes that the stubble burning particles staginated above 2 km range, hence the affect of stubble burning on human may be negligible.

- According to the article of Chen et al. (2021) on the estimation of spatial variations and temporal pattern of air pollutants depending on high resolution data of MAIAC AOD 1 Km in Hubai, China region. This article highlights the limitations of ground monitoring centre or sites which are unable to provide the regular data of monitoring. The author used a Modified support vector regression (MSVR) model to estimates particulate matter concentrations in Hubai, China by combining the MAIAC AOD, ground data measurements of particulate matter and meteorological data. The use of MSVR model is to capture the relationship between AOD, particulate matter and other auxiliary variables. MAIAC provide high resolution AOD products and other datasets. This datasets enables accurate estimation of particulate matter concentrations. There are a lot of effects of meteorological data such as temperature and humidity because air pollutants are absorbed by these auxiliary variables and affect the result. To obtain the trend of these variables, MSVR model used. Before estimating the particulate matter values, they finds the relationship between aerosols optical depth values and $PM_{2.5}$ concentration. By changing the different location at time, both the variables have shown similar trends. Ground data values are predicted by using the MSVR statistical model. Temporal and seasonal pattern of $PM_{2.5}$ concentrations are calculating by taking the data from different seasons and applying the ARIMA model. The ground data values are generally low in summer, high in winter due to the smoggy weather and during the spring and autumn season, the lowest values were recorded. To check the performance of MSVR model, both SVR and MSVR are applied to the same datasets and then compare their value of R squares or goodness of fit value. The accuracy for predicting the $PM_{2.5}$ values using MAIAC AOD values and meteorological data was comes out to be inconsistant as the values changes according to season. Evaluating all this experiment by the authors, they finds that the MSVR model performed better results in estimating the high resolution $PM_{2.5}$ data. The article provides information about the spatio-temporal variation and distribution of $PM_{2.5}$ concentration in Hubai, China and identify the regions with higher pollution level. The author also discuss about the factors that contribute in increasing AOD concentration such as dense populated areas and industrial activities.

- The article by He et al. (2017) gives better understanding and comparison of MODIS aerosol optical depth at 3 km values and 10 km values over China region. The main target of performing these measurements is focusing on spatial and temporal variations. MODIS provides globally aersol values with high temporal resolution. The 3 km aerosols optical depth product is expected to provide better result of local air pollution as compared to 10 km aerosols optical depth product. This study used aerosol values of Terra MODIS data during March 2000 to July 2015. AERONET ground based data of air pollution are taken from 18 sites across China during the same time period for the validation. Both the 3 km aerosols optical depth product and 10 km aerosols optical depth product shows high correlated values with the AERONET data, as R-value is greater than 0.9 for both the cases. But the spatial variation of the aerosols optical depth products shows both the product perform different when location changes. The 3 km aerosols optical depth product perform better as compared to 10 km product in developed areas of Southern China region. In northern region, 3 km aerosols optical depth products perform worse. In desert region of northwestern China, 10 km aerosols optical depth products has higher R values with the ground data. For temporal variation, both product perform worse during summer and spring due to the cloud cover. There is a difference in daily mean AOD values when compare the day to day variation in the values of 3 km aerosols optical depth products and 10 km aerosol products. Seasonal comparisons of both the products reveal that 3 km aerosols optical depth values are generally higher than the 10 km aerosols optical depth values more particularly in summer and spring. Atmospheric clouds and some surface activities are important factors that affects the accuracy of the MODIS to retrieve AOD values. 3 km aerosols optical depth values provide finer resolution of data and more information about local aerosols to study but its accuracy is slightly low as compared to 10 km product. This study reveals that the 3 km AOD products provide more information about the aerosols values of air pollution but the accuracy of 10 km product is good as compare to it. Both the products have good correlation coefficient with the ground based data to compare with aerosols values. The main conclusion of this study is, 3 km high resolution data product is more beneficial in highly populated areas and high polluted areas.

- The article by Vohra et al. (2021) examines the satellite based instruments for monitoring the air quality measurements in 4 main cities as London and Bermingham in the United Kingdom, Delhi and Kanpur in India. This research takes different type of data from satellite instruments which includes OMI for NO2, IASI for NH3 and MODIS for AOD as a substituent of particulate matter $PM_{2.5}$. These satellite datasets are used to predict the ground based concentration of aerosols. Aerosol optical depth value that are extracted from MODIS have no good correlation with $PM_{2.5}$ but consistent over the AERONET ground AOD values for both the sites in UK and India with a value of R greater than 0.8. According to the MODIS data, there is a decline in the $PM_{2.5}$ level in London since 2009.

Correlations between AOD and $PM_{2.5}$ are weak which might be due to the variability and difference in decomposition of aerosol levels. Both cities of India: Delhi and Kanpur has increasing trends in $PM_{2.5}$ which indicates that irrespective of regular efforts in controlling air pollution there is no such improvement. Kanpur is one of the most polluted city across the world according to the list given by WHO in 2018. The main reason of decreasing trends in UK and incresing in India is control on the vehicular emissions. This study collects the satellite based data of aerosols optical depth values from MODIS and ground-based data of the aerosol concentration. This study applies statistical methods including Theil-san estimator to determine the saesonality and trends in the following data. Satellite based data are compared with the ground based measurements in UK and India to check the efficiency of satellite observations in predicting ground based measurement or establishing statistical relationship. The observed trends and seasonality in these 4 cities highlights the upcoming challenges in controlling air pollution in future specially in cities like Kanpur and Delhi which have very high values. The conclusion of this article is that the satellite observations provides a overview about long term trends in urban air quality management. Regular efforts taken in improving air quality shows the benefits in London and Bermingham but there needs to improve the strategies in cities like Delhi and Kanpur as the trends for air pollution always tends to increasing.

- According to the paper published by Taneja, et al. (2020) about time series analysis of aerosols optical depth over Delhi region by using the Box-Jenkins ARIMA model approach. This study focuses on applying stochastic model, Box-Jenkins ARIMA model in order to predict the future trends of aerosols optical depth properties over Delhi, India knowns for its high pollution levels. Delhi's high levels of pollution from both natural and anthropogenic sources needs detailed studies like this for effective air quality management. Terra MODIS satellite data used to retrieved aerosol optical depth values at 550 nm. The aerosols optical depth values at 550 nm, monthly mean was taken from January 2004 to December 2014 period. Seasonal ARIMA model was used in SPSS statistical analysis software. The autocorrelation analysis shows the presence of seasonal pattern in AOD time series which indicates the AOD values peak during some certain months every year. The best fit model was used to predict the aerosols optical depth values for the year 2014 by using R square value, root mean square value, mean absolute error and many more. Residuals from the fitted models were analyzed to ensure that there was no significant autocorrelations between them. This step was used to check the adequacy of the models. Ljung-Box Q test was applied to check about the independency of residuals to ensures that the chosen models captured all significant information that is necessary from the data. Accurate prediction of the AOD values is crucial for climate modeling and public health planning. A cross-validation approach was used and the dataset was split into training and testing to predict the future values of aerosols optical depth products. Hence, this statistical model is first step towards finding the future values of aerosols optical depth, it can be applied later to other climatic purposes and future trend analysis of other products. This study conclude that the ARIMA modeling technique is very effective in predicting the future values of aerosol optical depth and its trends. This model can helps us in understanding and managing the impacts of aerosols on human health and climate changes in urban area like Delhi, India. The successful application of SARIMA models in this study highlights the potential of satellite data in providing accurate forecasts and helps in understanding the climatic factors and public health.

Chapter 3

Methodology

3.1 Methods

The methodology used for study the air quality for different region of Delhi (mainly focused on IITM JanakPuri) and carbonaceous aerosol concentration for the ROI were estimated using the satellite based data of aerosols optical depth and black carbon concentration.

3.1.1 Data Collection

We collected the satellite data that used in our study from various sources, including LAADS DAAC (Level-1 and Atmosphere Archive and Distribution System Distributed Active Archive Center) and Giovanni.

About LAADS DAAC:

LAADS DAAC website serves a platform for accessing the Earth observation data collected by various NASA's instruments. Its aim to provide easy access to Earth observation data for scientific research, decision-making processes and monitering the environmental. The website give us imformation about Earth observation data derived from various satellite missions including MODIS named as Moderate Resolution Imaging Spectroradiometer, VIIRS named as Visible Infrared Imaging Radiometer Suite, ASTER named as Advanced Spaceborne Thermal Emission and Reflection Radiometer etc. This data covers a diverse range of Earth scientific disciplines including atmospheric, land surface and oceanic observations. We can search data collections by specific parameters such as location, time and satellite instrument or use spatial and temporal filters to refine the search.

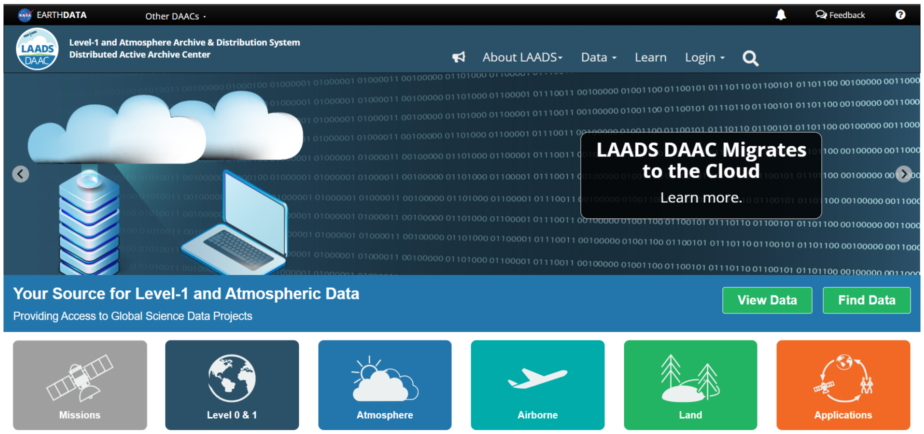

Figure 3.1: LAADS DAAC Portal

Accessing LAADS DAAC data:

We can access the LAADS DAAC data by following these steps:

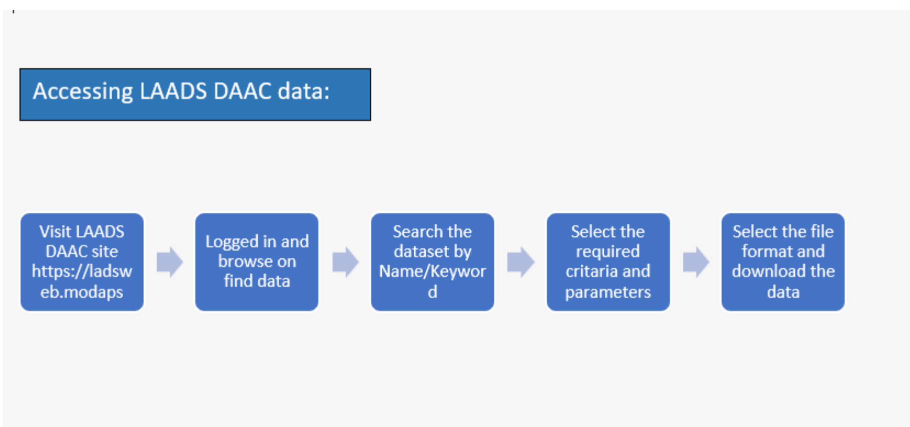

Figure 3.2: Accessing LAADS DAAC data

1. We visit the LAADS DAAC website by searching it on search engine and register for an account. Here is the URL for Giovanni: `https://ladsweb.modaps.eosdis.nasa.gov/`.
2. Once logged in, Browse on Find Data where we get the portal for veiw and downloading the data.
3. We search our data for study by dataset name/ keyword, date range, geographic location and day-night boundries. By following these steps we get to select the dataset files for downloading process.

4. After locating the desired dataset, we select dataset to view more detailed information about the data product, including metadata, file formats, and download options.

5. By following all these prompt, we get the download link.

About GIOVANNI:

Giovanni is an application developed by NASA's Goddard Earth Sciences Data and Information Services Center. It provides a online simple interface for accessing, visualizing and analyzing Earth science data. It offer access to a wide range of Earth science data derived from various NASA satellite missions. These data covers disciplines such as atmospheric, land surface, oceanic parameters and climate variables. One of the key feature of Giovanni is its ability to generate visualizations and perform basic data analysis without the need of software or programming skills. This makes it a valuable resource for scientists, researchers, and educators.

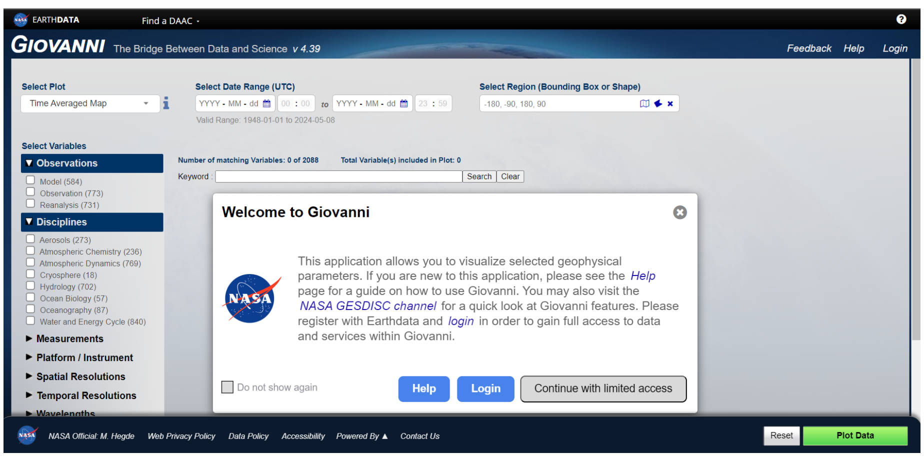

Figure 3.3: GIOVANNI Portal

Accessing GIOVANNI data:

We can access the GIOVANNI data by following these steps:

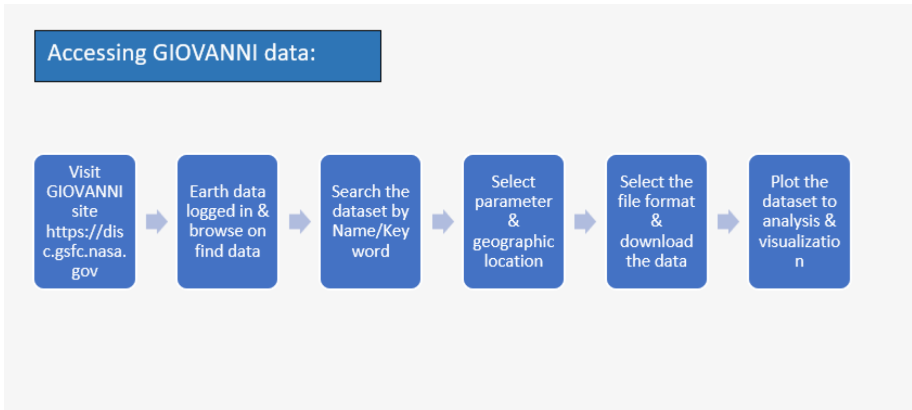

Figure 3.4: Accessing GIOVANNI data

1. We visit the GIOVANNI website by searching it on search engine and register for an Earth data login account. Here is the URL for Giovanni: `https://disc.gsfc.nasa.gov/information/tools?title=Giovanni`.

2. Once logged in, we select the parameters, date range, geographic location and plot type for the study data.

3. Then, we select the dataset from all the available datasets.

4. After following all these steps we get the plot for our dataset to analysis and visualize it.

3.1.2 Data processing

Data processing involves various steps to converting the downloaded data in a format for which the analysis is suitable. Here is the detailed explanation of the data processing steps that we have taken:

Processing of AOD values at 1 km resolution, ground level BC concentration and black carbon surface mass concentration from MERRA-2

1. Data obtaining: We have download the required data of study for 6 months from the LAADS DAAC website by giving the appropriate format command. for each day, we got a .hdf4 extension file for the 6 months data. we can't be able to view these types of file simply in our laptop or computer because this type of extension does not support by them. So we take the help of the programming language python by writing the command in Jupyter notebook.

2. Data Exploration: First of all, to find that what kind of data these hdf files store, we viewed the sds (Scientific Data Sets) of the files by giving

the command on all the files in python notebook. This process helped us to identify the parameters relevant to our study.

```
from pyhdf import SD
hdf_file_path = 'c:/Users/kaila/Downloads/MCD19A2.A2022315.h24v06.061.2023012193614.hdf'
try:
    hdf = SD.SD(hdf_file_path)
except Exception as e:
    print('Unable to open file: \n' + hdf_file_path + '\nSkipping...')
    print('Error:', e)
    exit()
datasets = hdf.datasets()
print('\nThis is a MODIS file. Here is a list of 1 km SDS in your file:\n')
for i, item in enumerate(datasets.items()):
    key, value = item[0], item[1]
    spatres = value[0]
    if any("1km" in name for name in spatres):
        print('{0}: {1}'.format(i + 1, key))
print('\nAll valid files in the directory have been processed')
```

Figure 3.5: Scientific Data Sets

3. Parameter Selection: Based on our study, we selected specific parameters which we needs in our further study because the files size are too large and this process made it easy to handle.

4. Conversion to CSV: We convert all these hdf files into csv files one by one manually by selecting the required parameters because if we select all the parameters then the converted csv file size should be very large and difficult to read. We apply this process of converting hdf file to csv by selecting the parameters manually because we have total 182 files and each file size is very large after converting. If we apply the same process to all these hdf files then the converted data size is too large and can't be handle by pc.

5. Location Filtering: Now, we have the data in csv format but the location of these data is not in proper coordinates so we apply filter on the latitude and longitude parameters by using the coordinates of Delhi (region of interest). we save all these files after applying same location filter to the latitude and longitude parameters of all the csv files and save.

```
A=A[(A['Latitude'] > 27.6581) & (A['Latitude'] < 29.113)]
A=A[(A['Longitude'] > 76.2781) & (A['Longitude'] < 78.0249)];A
```

Figure 3.6: Delhi coordinates

6. Subregion Filtering: Further filtering was applied to focus on a specific subregion of interest within Delhi, such as IITM Janakpuri.

```
import os
import pandas as pd
input_directory = 'c:/Users/kaila/OneDrive/Desktop/october2022/'
output_directory = 'c:/Users/kaila/OneDrive/Desktop/new data/'
latitude_min = 28.5962
latitude_max = 28.688
longitude_min = 77.1202
longitude_max = 77.230
for filename in os.listdir(input_directory):
    if filename.endswith('.csv'):
        filepath = os.path.join(input_directory, filename)
        df = pd.read_csv(filepath)
        filtered_df = df[(df['Latitude'] > latitude_min) & (df['Latitude'] < latitude_max) &
                         (df['Longitude'] > longitude_min) & (df['Longitude'] < longitude_max)]
        output_filepath = os.path.join(output_directory, filename)
        filtered_df.to_csv(output_filepath, index=False)
        print(f"Filtered data saved to: {output_filepath}")
```

Figure 3.7: IITM Delhi coordinates

7. Data Aggregation: We have the data for 6 months in 182 files for each day. It is difficult to analysis this data as each files had large number of values. So, we take the average of the values of satellite based data named as aerosols optical depth at 470 nm which is one of the parameter among all. We take the average for each days and make a new csv file by putting the average values for each day between 1 oct 2022 to 31 mar 2023. Similarly, we found the ground data of black carbon concentration for Delhi region and averaged it for each day at specific times (10 am, 11 am, 12 pm, and 1 pm IST) to match the timing of satellite data. Ater that, we download the data of black carbon concentration by selecting the required fields and product named as black carbon surface mass concentration (M2T1NXAER v5.12.4) on Giovanni. Giovanni allow us to download the data for single day, so we download the data for the timing of satellite data manually and take the average of the data for each day.

8. Combining Datasets: By combining the daily average ground data of black carbon concentration with the daily average satellite data for product Aerosol Optical Depth involving MERRA-2 average black carbon concentration data from Giovanni. Now, the daywise data of ground black carbon corbon concentration, satellite data of aerosols optical depth at 470 nm and average black carbon concentration of MERRA-2 that we have download from giovanni.

9. Data Cleaning: After combining the datasets, import the final csv file into Python and excluded the missing values from the dataset.

IMPORT FINAL CSV FILE

```
import pandas as pd
A=pd.read_csv("c:/Users/kaila/OneDrive/Desktop/FINAL_CSVFILE_OF_ALL_DATA.csv");A
```

	Unnamed: 0	DATE	AVERAGE OPTICAL_DEPTH_047	carbonaceous_aerosols_470(kg/m^3)	average_merra2(kg/m^3)_(M2T1NXAER v5.12.4)
0	0	01-10-2022	0.188398	NaN	3.500000e-09
1	1	02-10-2022	0.242980	NaN	3.610000e-09
2	2	03-10-2022	0.222058	NaN	1.860000e-09
3	3	04-10-2022	0.373552	9.040000e-09	2.670000e-09
4	4	05-10-2022	0.278813	1.410000e-08	2.780000e-09
...	...	...	...	...	...
177	177	27-03-2023	0.454712	4.360000e-09	1.250000e-09
178	178	28-03-2023	0.548524	4.620000e-09	1.280000e-09
179	179	29-03-2023	0.580424	8.430000e-09	2.120000e-09
180	180	30-03-2023	NaN	8.550000e-09	2.420000e-09
181	181	31-03-2023	0.422143	2.950000e-09	1.450000e-09

Figure 3.8: importing final csv file of all datasets

3.1.3 Calculation of Scaling Factors for variables

Scaling factors were calculated for two key variables by using a base variable to normalize the data and facilitate meaningful comparisons between two key variables. The steps involved in calculating the scaling factors are as follows:

- The two variables selected for scaling were Ground data of Carbonaceous aerosols concentration and BC MERRA-2 concentration.
- For each variable, the mean and standard deviation were computed from the cleaned dataset.
- Each value of the two key variables was scaled by dividing both the variables with each value of satellite data of Aerosols optical depth at 470nm wavelength as it is a unitless quantity, so the transformation is same for both.
- Now, linear regression model was fitted between the two scaled variables and calculate R-squared value by using train and split test to check the efficiency of data in predicting the values of ground data of carbonaceous aersols concentration.

SCALING FACTOR FOR GROUND DATA AND BC MERRA-2 DATA

```
import pandas as pd
A=pd.read_csv("c:/Users/kaila/OneDrive/Desktop/FINAL_CSVFILE_OF_ALL_DATA.csv");A
A=A.dropna();A
A['scaled_bc_mass_conc'] = A['average_merra2(kg/m^3)_(M2T1NXAER v5.12.4)'] / A['AVERAGE OPTICAL_DEPTH_047'];A
A['scaled_carbonaceous_aerosols_conc'] = A['carbonaceous_aerosols_470(kg/m^3)'] / A['AVERAGE OPTICAL_DEPTH_047'];A
```

amed: 0	DATE	AVERAGE OPTICAL_DEPTH_047	carbonaceous_aerosols_470(kg/m^3)	average_merra2(kg/m^3)_(M2T1NXAER v5.12.4)	scaled_bc_mass_conc	scaled_carbonaceous_aerosols_conc
3	04-10-2022	0.373552	9.040000e-09	2.670000e-09	7.147607e-09	2.420014e-08
4	05-10-2022	0.278813	1.410000e-08	2.780000e-09	9.970823e-09	5.057144e-08
5	06-10-2022	0.320570	1.490000e-09	1.670000e-09	5.209466e-09	4.647967e-09
6	07-10-2022	0.334944	1.720000e-09	3.850000e-09	1.149446e-08	5.135187e-09
7	08-10-2022	0.603907	2.200000e-09	5.890000e-09	9.753161e-09	3.642946e-09

Figure 3.9: Calculation of scaling factor of variables

This transformation ensured that the variables were on a comparable scale to facilitating further analysis.

Processing of fire data in neighbouring states during stubble burning period

1. Data obtaining: We have download the fire count data for Delhi region and in neighbouring states during stubble burning period from satellite data of MODIS by using coordinates 28.1634°N latitude to 31.4721°N latitude and 74.9787°E longitude to 78.6146°E longitude in a csv format.

2. Location Filtering: This data also contains some extra coordinates values because satellite took the data for a tile and direct filtering from the website can lead to some extra points. So, we again use the same coordinates filter by processing this data into our python programming notebook.

```
A=A[(A['latitude'] > 28.1634) & (A['latitude'] < 31.4721)]
A=A[(A['longitude'] > 74.9787) & (A['longitude'] < 78.6146)];A
```

Figure 3.10: fire data coordinates

3. Data Aggregation: We took the average of amount of fire count per day during the period 1 October 2022 to 31 December 2022 and save this data into a csv formet file.

4. Combining Datasets: We combined the fire count data with the data that we have already processed for a period of 1 October 2022 to 31

December 2022 as we are analysing the effect of fire count on black carbon concentration during stubble burning period.

5. Data Cleaning: After combining the datasets, we imported the final csv file into Python and excluded the missing values from the dataset.

```
import pandas as pd
A=pd.read_csv("c:/Users/kaila/OneDrive/Desktop/fire_count_final data.csv");A
```

	Unnamed: 0	DATE	AVERAGE OPTICAL_DEPTH_047	GROUND DATA AVERAGE(ng/m^3)	carbonaceous_aerosols_470(kg/m^3)	average_merra2(kg/m^3)_(M2T1NXAER v5.12.4)	fire_count
0	0	01-10-2022	0.188398	NaN	NaN	3.500000e-09	1.0
1	1	02-10-2022	0.242980	NaN	NaN	3.610000e-09	NaN
2	2	03-10-2022	0.222058	NaN	NaN	1.860000e-09	12.0
3	3	04-10-2022	0.373552	9039.717262	9.040000e-09	2.670000e-09	NaN
4	4	05-10-2022	0.278813	14072.437800	1.410000e-08	2.780000e-09	23.0
...	...	...	...	...	...	...	...

Figure 3.11: Importing fire data with BC MERRA-2 concentration

By following these steps, we transformed satellite, MERRA-2 and ground data into a dataset that suitable for further analysis.

Chapter 4

Statistical Models and Approach

4.1 Statistical Models and Approach

In the study conducted, various statistical approachs and models are used to analyze the 6 months daily average data for Delhi. The analysis includes visualizations, trend analysis, correlation studies, and predictive modeling to understand and predict air quality parameters for future references.

4.1.1 Discriptive Statistics

- Descriptive statistics are used to summarize and describe the main features of data in a quantitative manner.
- These include the mean, median, and mode. The mean provides the average value of the dataset, the median gives the middle value when data points are arranged in order, and the mode represents the most frequently occurring value.
- Standard deviation, variance, and range are key metrics about the dataset.
- It provide a way to present large amounts of data in a more understandable form, often using graphical representations like histograms, box plots, and scatter plots.
- These statistics help in identify patterns and trends in the data, such as seasonal variations which can be critical for further analysis and modeling.

```python
import pandas as pd
data = pd.read_csv("C:/Users/kaila/OneDrive/Desktop/FINAL_CSVFILE_OF_ALL_DATA.csv")
data = data.dropna()
data['DATE'] = pd.to_datetime(data['DATE'], format='%d-%m-%Y')
data = data.sort_values(by='DATE')
data['Month'] = data['DATE'].dt.month
months = [10, 11, 12, 1, 2, 3]
month_names = {10: 'Oct', 11: 'Nov', 12: 'Dec', 1: 'Jan', 2: 'Feb', 3: 'Mar'}
monthly_stats = data.groupby(data['Month'])['AVERAGE OPTICAL_DEPTH_047']\
.agg(['mean', 'std', 'min', 'max', 'var'])
monthly_stats.index = monthly_stats.index.map(month_names)
print(monthly_stats)
```

Figure 4.1: Discriptive Statistics test

4.1.2 Data Visualization

- Bar Graphs: Bar graphs are a simple but powerful tool to display the data in categorical form. In this study, bar graphs were used to represent monthly variations in the daily average data of satellite based AOD. This visualization helps in understanding the variability and distribution of the data across different months by comparing the heights of the bars as well as identify the months with the highest and lowest values.

- Line Graphs: Line graphs were used to illustrate trends of two variables BC MERRA-2 and fire count over time. The x-axis represented dates and y-axis represents BC MERRA-2 concentration values on left, per day fire count on right allowing for the observation of variations and trends over the 3 month period. Line graphs are particularly useful for time series data as they clearly show how a variable changes over time. By plotting two variables on the same graph, it is possible to assess their relationship and identify any trends or patterns presents in the data.

- Scatter Plots: It was used to analysis the relationship among two variables. This approach helps in visualizing the correlation between the two variables. Each point on the scatter plot represents an observation, with the x-coordinate corresponding to one variable and the y-coordinate corresponding to another. The pattern of the points can indicate the type and strength of the relationship, such as positive, negative, or no correlation. Scatter plots are essential in identifying outliers and understanding the linear or non-linear relationships between variables.

```
import matplotlib.pyplot as plt
x= A['scaled_bc_mass_conc']
y = A['scaled_carbonaceous_aerosols_conc']
bo = 2.92906098
b1 = 1.7092255585058988e-09
regression_line = bo * x + b1
plt.figure(figsize=(10, 8))
plt.scatter(x, y, color='red', label='Data')
plt.plot(x, regression_line, color='blue', label='Regression Line')
```

Figure 4.2: regression test

4.1.3 Trend Analysis

- Seasonal Mann-Kendall Test: To finds the trends over the 6 months time period data, the Seasonal Mann-Kendall test was applied in this study. The non-parametric test provides the presence of a monotonic trend which means the trend is either increasing or decreasing in time series data. This test is useful for data having seasonality and serial dependence. It does not assume any specific distribution for the data, making it durable against outliers and missing values. The test provides a p-value which helps us in determining the statistical significance of observed trend value. A statistically significant p-value (always less than 0.05) indicates a trend in the data.

```
import pymannkendall as mk
aod_series = data['AVERAGE OPTICAL_DEPTH_047']
result = mk.seasonal_test(aod_series, period=12)
print("Seasonal Mann-Kendall Test p-value:", result.p)
print("Trend:", result.trend)
slope, intercept = np.polyfit(np.arange(len(aod_series)), aod_series, 1)
trend_line = intercept + slope * np.arange(len(aod_series))
```

Figure 4.3: Seasonal Mann-Kandall test

4.1.4 Correlation Analysis

- Pearson Correlation Coefficient: Pearson Correlation Coefficient: To understand the relationship between two variables BC MERRA-2 and fire count, the Pearson correlation coefficient was calculated. The coefficient finds the strength between the variables and direction of the linear relationship between two continuous variables over time. The range is in between -1 and 1, where 1 indicates a perfect positive, -1 indicates a perfect negative and 0 indicates no correlation. Pearson correlation coefficient is useful for identifying how two variables are closely related to each other and changes in one variable are associated with changes in another. In this study, it helped in understanding the relationship between these two air quality parameters.

```
data = {
    "average_merra2(kg/m^3)_(M2T1NXAER v5.12.4)": average_merra2,
    "fire_count": fire_count
}
df = pd.DataFrame(data)
summary_stats = df.describe()
print(summary_stats)
correlation = df.corr()
print("\nCorrelation Matrix:")
print(correlation)
```

Figure 4.4: correlation test

4.1.5 Train-Test Split

The whole dataset comprising six months of daily average data for Delhi was divided into two parts as a training set and a testing set. The training set comprising a majority of the data (70%) was utilized to train the predictive models. The testing set comprising the remaining data (30%) was used to evaluate the performance of the trained models. This approach helps to mitigate overfitting and provides a more realistic estimation of the model.

R^2 BY TRAIN AND SPLIT TEST

```
x= A['scaled_bc_mass_conc'].array.reshape(-1, 1)
y = A['scaled_carbonaceous_aerosols_conc']
```

```
from sklearn.model_selection import train_test_split
from sklearn.metrics import r2_score,mean_absolute_error,mean_squared_error
xtrain,xtest,ytrain,ytest=train_test_split(x,y,test_size=0.2,random_state=14)
from sklearn.linear_model import LinearRegression
import numpy as np
lm=LinearRegression()
model=lm.fit(xtrain,ytrain)
y_pred=model.predict(xtest)
bo=model.coef_
b1=model.intercept_
r_squared = r2_score(ytest, y_pred)
n = len(ytest) + len(ytrain)
p = xtrain.shape[1]
adjusted_r_squared = 1 - ((n - 1) / (n - p - 1)) * (1 - r_squared)
mse = mean_squared_error(ytest, y_pred)
mae = mean_absolute_error(ytest, y_pred)
rmse = np.sqrt(mse)
tr_pred=model.predict(xtrain)
ts_pred=model.predict(xtest)
train_err=mean_absolute_error(ytrain,tr_pred)
test_err=mean_absolute_error(ytest,ts_pred)
```

Figure 4.5: Train-test split

Model Evaluation Metrics

To calculate the performance of predictive model, several evaluation metrics were used. These metrics include mean squared error (MSE) and R-squared (R^2). MSE provides a measure of the average squared difference between the predicted and observed values and R^2 indicates the proportion of variance in the dependent variable explained by the independent variables.

4.1.6 Predictive Modeling

- Linear Regression: Linear regression model was used in predicting the future values of a variable based on the past values of data. This method helps us in creating the relationship between a dependent variable and one or more independent variables by fitting a linear equation to the observed data. The slope of this fitted line representing the rate at which changes occure of the dependent variable with respect to the independent variable, while the intercept line of the scatter plot represents the value of the dependent variable when the independent variable is assumed to be

0. Linear regression model is valuable for making the predictions of the values of dependent variable and understanding the relationship between variables. In this study, it was used to predict future ground based data of scaled Carbonaceous aerosols with scaled MERRA-2 BC concentration based on past observations of data which provides insight into potential future trends.

By using these statistical models and approaches, this study effectively analysis the variations, trends, and correlations in the data. The predictive models helped in forecasting the future values of ground data of BC CA's which contributing to a better understanding of air quality in Delhi. These methods collectively provided a comprehensive analysis and allowing for decision-making and interventions to improve air quality.

Chapter 5

Results and Discussion

5.1 Results and Discussion

In this study, high resolution aerosols optical depth product from Level-1 and Atmosphere Archive & Distribution System Distributed Active Archive Center (LAADS-DAAC) with 1 km resolution were processed. Aerosols in ambient atmosphere are represented through AOD and it is difficult to measure particulate matter directly from satellite. MAIAC algorithm integrates both MODIS Aqua and MODIS Terra data for obtaining AOD at 1 km resolution. AOD at higher resolution helps in identifying the probable source of pollution at local level and availability of satellite pass time can help in mapping the time and source of pollution together. In this study, high resolution AOD at 1 km was derived from October, 2022 to March 2023. MERRA-2 reanalysis data was used to develop a linear model for prediction of Black Carbon fraction of Carbonaceous aerosols (CA's) and effect of fires in nearby areas on BC concentration over satellite data was also assessed and machine learning models for prediction of high resolution AOD were developed.

Month	Mean	Std Dev	Min	Max	Variance
Oct	0.464670	0.136007	0.236952	0.742714	0.018498
Nov	0.633985	0.379150	0.152484	1.675275	0.143754
Dec	0.484711	0.190789	0.205020	0.939163	0.036400
Jan	0.508079	0.227798	0.242599	1.084452	0.051892
Feb	0.631612	0.318096	0.357554	1.763067	0.101185
Mar	0.549679	0.158806	0.370635	0.993224	0.025219

Table 5.1: Monthly Statistics

5.1.1 Monthly variation of Aerosol optical depth (AOD) at 470nm

AOD at 470 nm represents blue wavelength and this is sensitive to smaller particles (fine mode aerosols) such as smoke, urban pollution and secondary organic aerosols. AOD at 470 nm better correlates with pollution from vehicles, industrial activities and biomass burning. AOD at 470 nm is effective for monitoring of urban and industrial pollution where fine particles are predominant. Important for assessing human health impacts related to very fine particulate matter ($PM_{2.5}$). For assessing the combined effect of vehicles, industrial activity and biomass burning AOD at 470nm is analyzed and discussed.

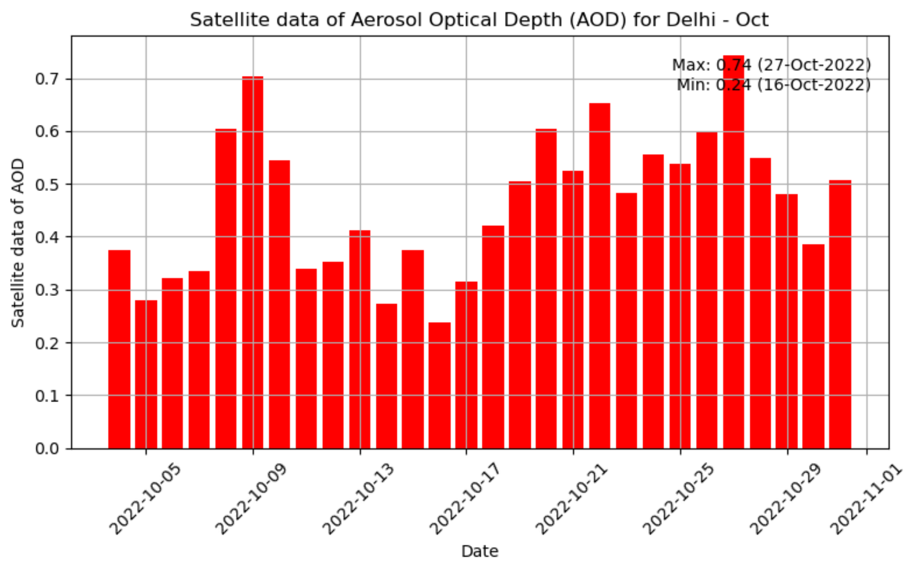

Figure 5.1: AOD values from MODIS for Delhi in october

From fig. 5.1, it is clear that AOD in October compared to starting of month in October remains at higher level by the end of month also. The mean value for AOD for October is 0.46±0.14. AOD after 27th October started rising and reached to maximum 0.74.

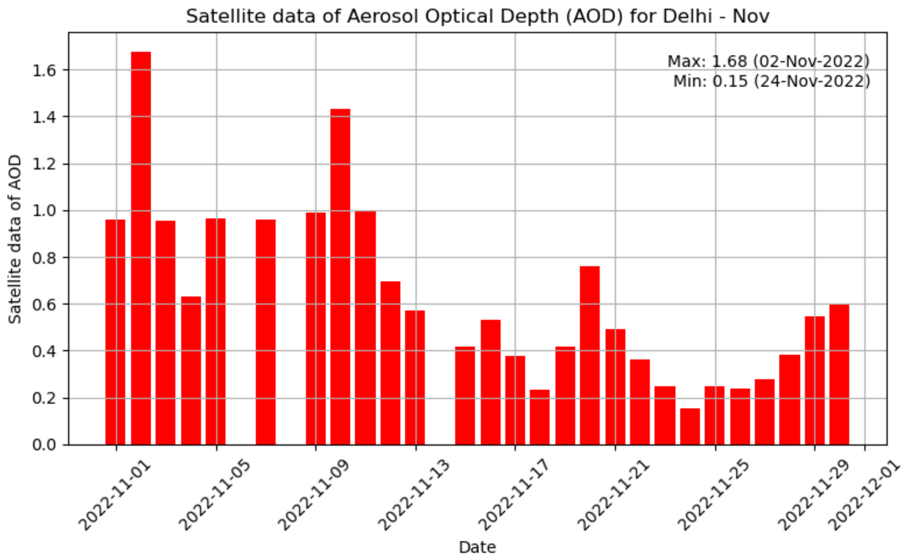

Figure 5.2: AOD values from MODIS for Delhi in november

From fig. 5.2, November is the month when winter sets in, there is moisture in atmosphere, and it can be seen that the AOD values remained higher than 0.8 from starting of November and AOD remained greater than 1.0 till 10 November indicating that the source of pollution is because of vehicles and biomass burning. The mean value for AOD for November is 0.63±0.38. The mean AOD value for November is greater than that of October and maximum AOD on 2nd November with 1.68 was observed.

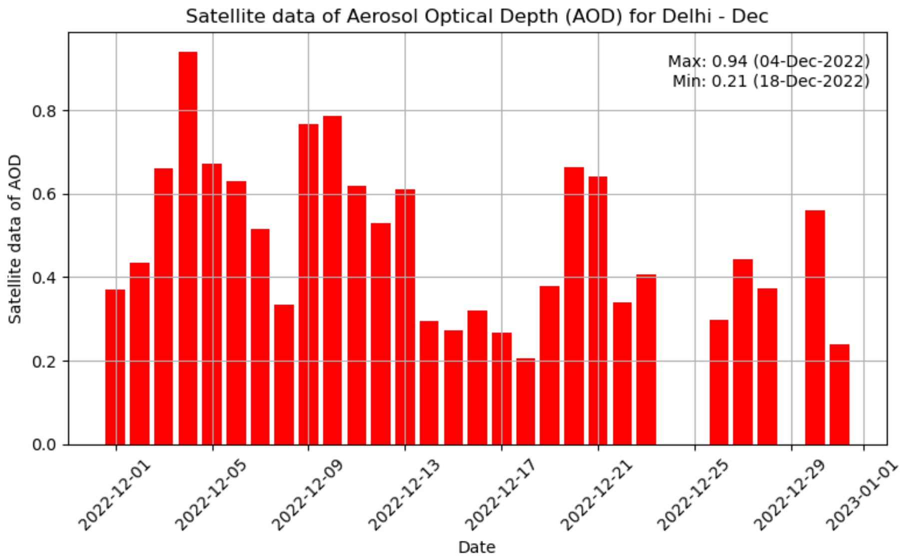

Figure 5.3: AOD values from MODIS for Delhi in december

From fig. 5.3, it is clear that dust AOD in December varied from high 0.94 (start of December) to low 0.22 (end of December). The mean AOD for December is 0.48±0.19. The mean AOD value for December is lesser than that of November. In December, peaks in AOD reaches maximum upto 0.94 and mean value less than 0.5 indicates that influence of biomass burning is not there and the pollution levels are because of industrial and vehicular activities.

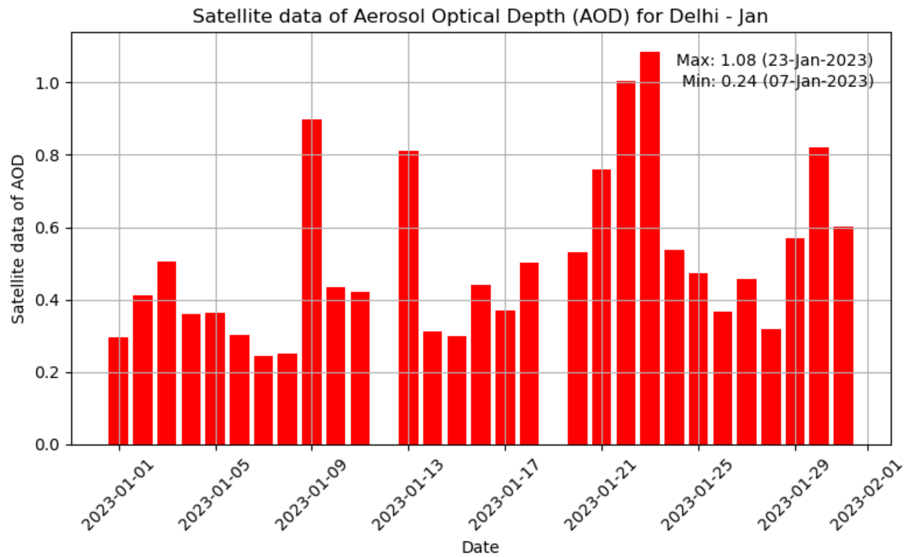

Figure 5.4: AOD values from MODIS for Delhi in january

From fig. 5.4, January is extreme winter month in Delhi and also the AQI is not in good category. The mean value for AOD for January is 0.51 ± 0.22. The mean AOD value for January is greater than that of December indicating that after mid of January, AOD levels in atmosphere starts increasing and at peak on 23 January with 1.08. The higher AOD values greater than 0.7 indicates vehicular, residential and industrial emissions dominance in ambient atmosphere.

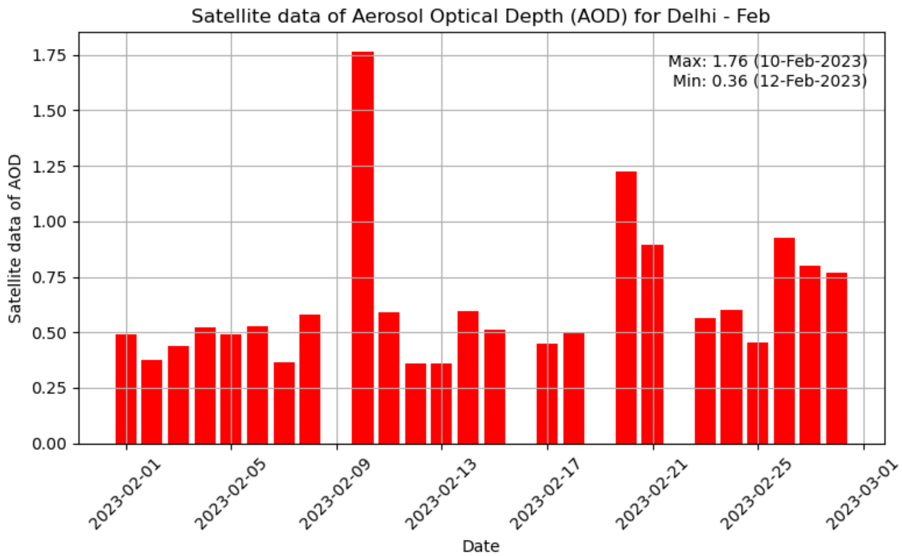

Figure 5.5: AOD values from MODIS for Delhi in february

From fig. 5.5, it is clear that AOD values started increasing in February and remained higher than mean AOD value in January. The mean value for AOD for February is 0.63±0.31. Based on mean AOD for February it can be said that it is a common scenario for Delhi and sources of pollution mainly would be vehicles, industrial activities and some open burning.

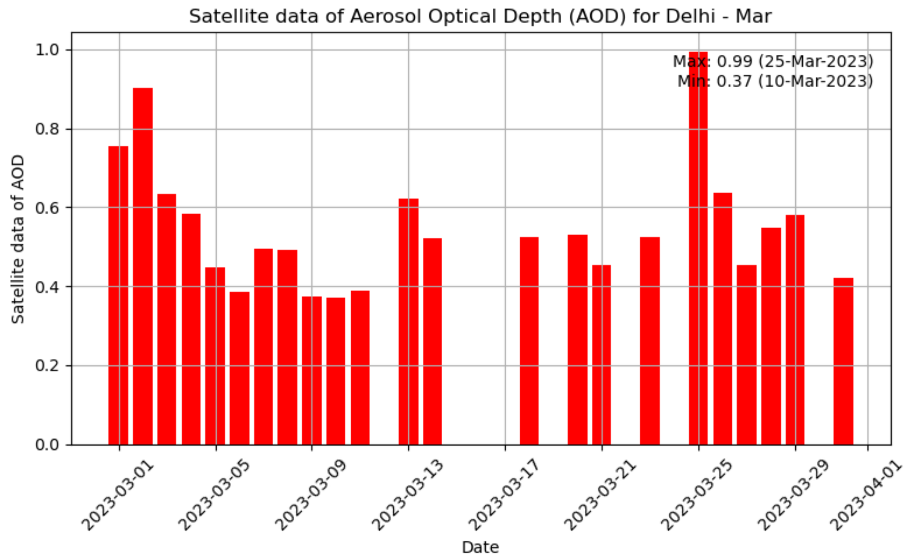

Figure 5.6: AOD values from MODIS for Delhi in march

From fig. 5.6, In March the AOD values increased when compared to February. The mean value for AOD for March is 0.55 ± 0.16 indicating the rise in various pollution activities such as vehicular emissions and industrial activities.

The monthly analysis of satellite based AOD values at 470 nm on daily basis from October 2022 to March 2023 indicates that October has least AOD values followed by December and maximum mean AOD values was found for November and February. The range of AOD values between 0.46 to 0.63 represents higher AOD values and this scenario is common in Delhi and there are days with extreme pollution event but these are regular AOD variations that are observed. AOD at 470 nm is generally higher compared to longer wavelengths if the atmosphere contains a significant amount of small particles. Thus, this indicating the presence of high amount of finer particles presents in atmosphere and which are more dangerous for human health and it is deduced that in month of November and February residents need to be more careful for their practice of using mask in exposed environment.

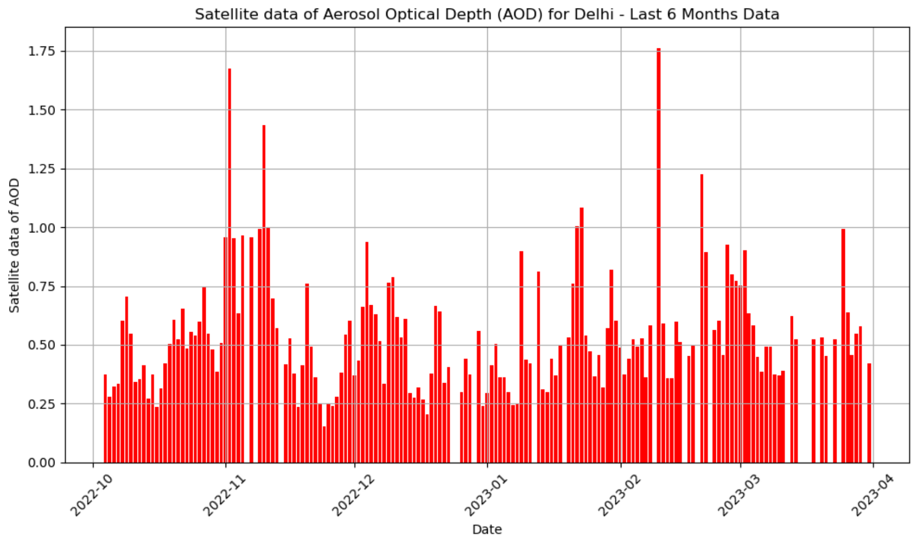

Figure 5.7: AOD values from MODIS for Delhi during oct-2022 to mar-2023

Trend analysis of AOD

The trend analysis graph from fig. 5.8, shows that there is a positively increased pattern during the winter period which concludes the increased fine particulates in ambient environment . It is conducted to understand the past variations and predict the future possibility using the analysis.

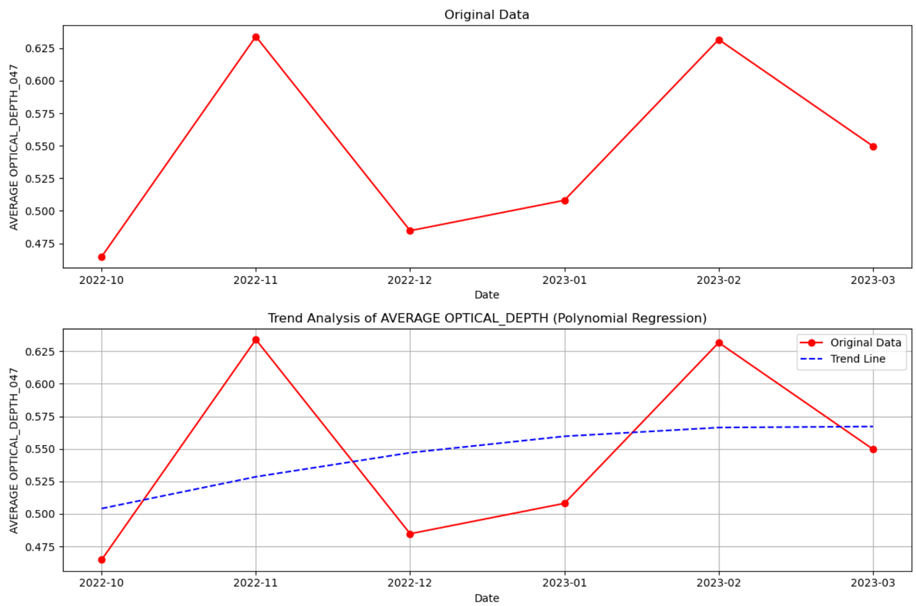

Figure 5.8: Trend analysis of satellite based AOD values at 470nm

Seasonal Mann-Kendall Test for Trend analysis

This test of trend analysis is done to assesses whether there is a monotonic trend which means the trend is either increasing or decreasing in the time series data. The p-value associated with this test is (0.047) which below the value of significance level of 0.05. This indicates that the correlation observed is statistically significant. If the p-value is less than 0.05, then it can be concluded that there is a statistically significant increasing trendpresent in the aerosols optical depth product (470nm) over time. The result of hypothesis (h) is TRUE implies that trend detected by the this test is statistically significant.

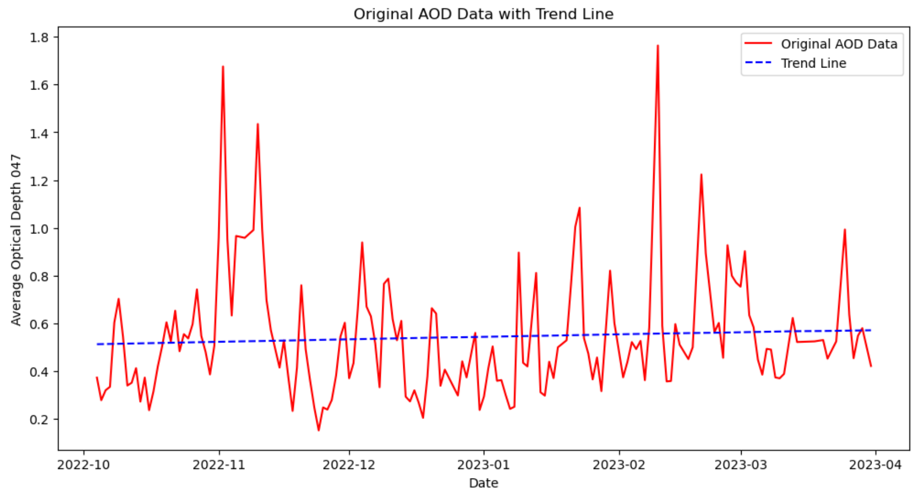

Figure 5.9: Seasonal Mann-Kandall test for trend analysis

5.1.2 Analysis of Black Carbon Surface Mass Concentration from MERRA-2

Black carbon (BC) is the type of very fine particulate matter ($PM_{2.5}$) that produced from incomplete combustion of fossil fuels, biofuel, and biomass burning. It is a primary component of soot and is known for its ability to absorb light and heat, making it a significant contributor to global warming and air pollution. Black carbon, due to its strong light-absorbing characteristics, significantly contributes to AOD. While AOD includes contributions from all types of aerosols (scattering and absorbing), BC's ability to absorb sunlight makes it a key component in the overall optical depth, particularly in the visible spectrum, including 470 nm. BC has a stronger impact on AOD at shorter wavelengths (e.g., 470 nm) because of its higher absorption efficiency for blue light. This makes AOD measurements at 470 nm particularly useful for detecting BC. For Delhi, during October-November there is a sight where blackness in ambient air is environment. The BC surface mass concentration was derived from MERRA-2 data with 0.5° × 0.625° resolution with region of interest set for Delhi.

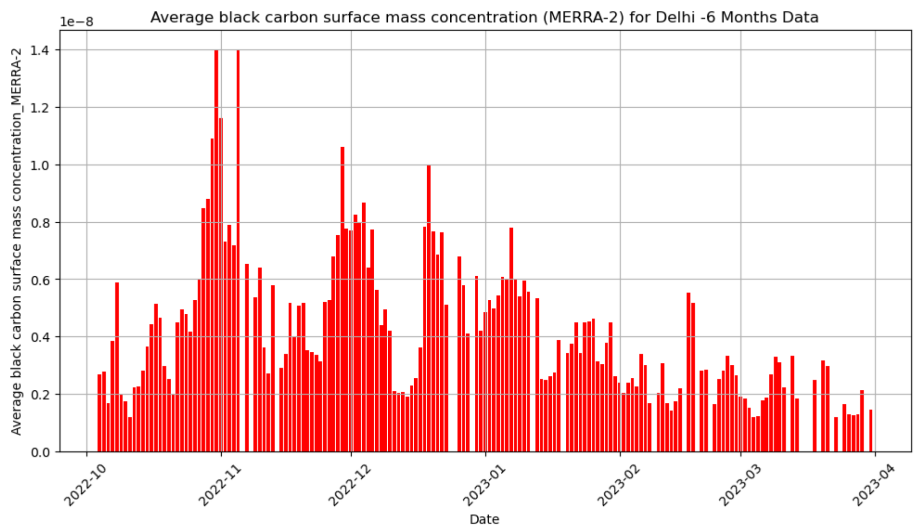

Figure 5.10: Analysis of Black Carbon Surface Mass Concentration from MERRA-2

The monthly analysis of BC on daily basis showed that mean BC mass concentration for Oct-Dec was 5.27 × 10-9 kg/m3 and for Jan-Mar was 3.14 × 10-9 kg/m3. This shows that Oct-Dec has higher BC concentration in atmosphere and this correlates well with several studies done for Delhi for this period. The higher concentration of BC as attributed by other studies also is due to burning activities in nearby areas, festival celebrations and vehicle exhaust emissions. The use of BC from MERRA-2 is helpful in identifying the fractions of BC in particulate matter and identifying the sources that are creating blackness in ambient air. This is experienced by residents of Delhi as blackening of masks and clothes and more prominent white clothes is observed during October-November and people spending sufficient time in open atmosphere experience more effect of blackening of air.

Effect of fire count on MERRA-2 Black Carbon (BC) product during stubble burning period

During October and November, fires in neighboring states of Delhi, particularly in Punjab and Haryana, significantly contribute to air pollution in the region. This period is marked by widespread of agricultural waste burning which commonly known as stubble burning, where farmers set fire to the leftover paddy straw after harvesting rice to quickly clear the fields for the next crop, typically wheat. Due to the short window between the rice harvest and the sowing of the next crop, farmers resort to burning as the quickest and most cost-effective method of clearing fields. The burning results in the presence of large amount of fine particulate matter and black carbon in ambient atmosphere. To analyze the impact of fire counts over black carbon. Black carbon surface mass concentration from MERRA-2 were analyzed for time period of 3 months (October-December).

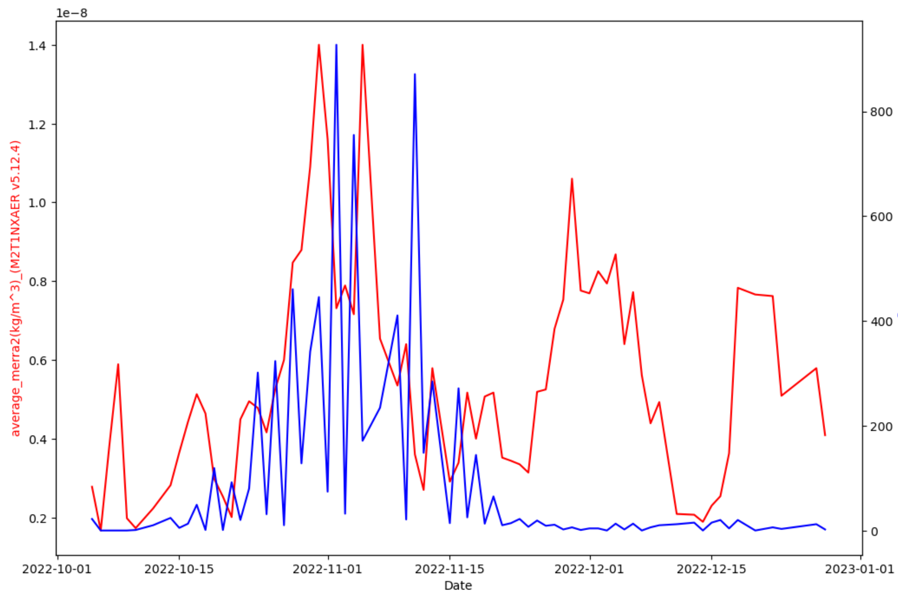

Figure 5.11: BC concentration variation with fire count

From Fig. 5.11, it can be seen that BC mass concentration matches well with fire count from October to December in 2022. It can be seen that BC and fire count are following similar pattern from 15th October to 15th November which is the period of festivals when firecrackers are burned and are source of BC in ambient environment and stubble burning also contributes to BC concentration. This shows that higher BC concentrations are due to fire counts and the source of BC in ambient is stubble burning and firecrackers. After 15th November fire count starts reducing and so does BC. December marks high concentration of BC and with almost no fire count showing the traffic and other sources contributing to BC rather than farm fires. This overlay can help in identifying the source contribution of BC. The below results provide interesting

insights into the relationship between black carbon surface mass concentration of MERRA-2 and fire count data during period of Oct 2002 to Dec 2022.

Table 5.2: Summary Statistics for Black Carbon Surface Mass Concentration and Fire Count

Statistic	*average_merra2(kg/m³)_M2T1NXAER v5.12.4*	*fire_count*
Count	8.000000×10^{1}	80.000000
Mean	5.229500×10^{-9}	92.837500
Std Dev	2.850550×10^{-9}	186.946783
Min	1.200000×10^{-9}	1.000000
25%	2.962500×10^{-9}	4.750000
50%	4.855000×10^{-9}	14.000000
75%	6.882500×10^{-9}	76.500000
Max	1.400000×10^{-8}	927.000000

Summary Statistics:

1. The dataset consists of 80 observations for both the variables.

2. On average, the black carbon surface mass concentration is around 5.23×10^{-9} kg/m^3 and standard deviation of approximately 2.85×10^{-9} kg/m^3.

3. The fire count varies with a mean of approximately 92.84 and a standard deviation of 186.95.

4. The minimum observed black carbon surface mass concentration is 1.20×10^{-9} kg/m^3 and the minimum fire count is 1. This indicates that there are instances in the data with both low and high concentrations and fire occurrences.

5. The maximum observed black carbon surface mass concentration is 1.40×10^{-8} kg/m^3 and the maximum fire count is 927.

Table 5.3: Correlation Matrix

	BC MERRA-2 CONC.	*fire_count*
BC MERRA-2 CONC.	1.000000	0.292651
fire_count	0.292651	1.000000

Correlation Analysis:

1. The correlation coefficient between black carbon surface mass concentration and daily fire count is approximately 0.29. This indicates a positive but relatively a weak correlation between these two variables.

2. The positive correlation indicates that as the fire count increases, there tends to slight increase in BC MERRA-2 concentration. Moreever, the correlation is not strong so other factors may also influence BC MERRA-2 concentration.

3. Other factors such as industrial emissions, high traffic density and meteorological conditions may also play significant roles in increasing black carbon surface mass concentration.

Overall, these results provide some valuable insights into the dynamics of black carbon surface mass concentration of MERRA-2 and daily fire occurrence and highlighting the need for comprehensive strategies to decreasing air pollution and manage fire risks effectively.

5.1.3 Linear Regression model

To find the fraction of BC in Carbonaceous aerosol, BC surface mass concentration obtained from MERRA-2 and ground data was used to develop a regression model. from fig. 5.12, this model estimates the ground level concentration of CA's BC fraction. The coefficient of correlation was about 0.88 indicating that developed equation is closely predicting ground level BC concentration. The development of prediction model is important because it reduces the need of repeated analysis.

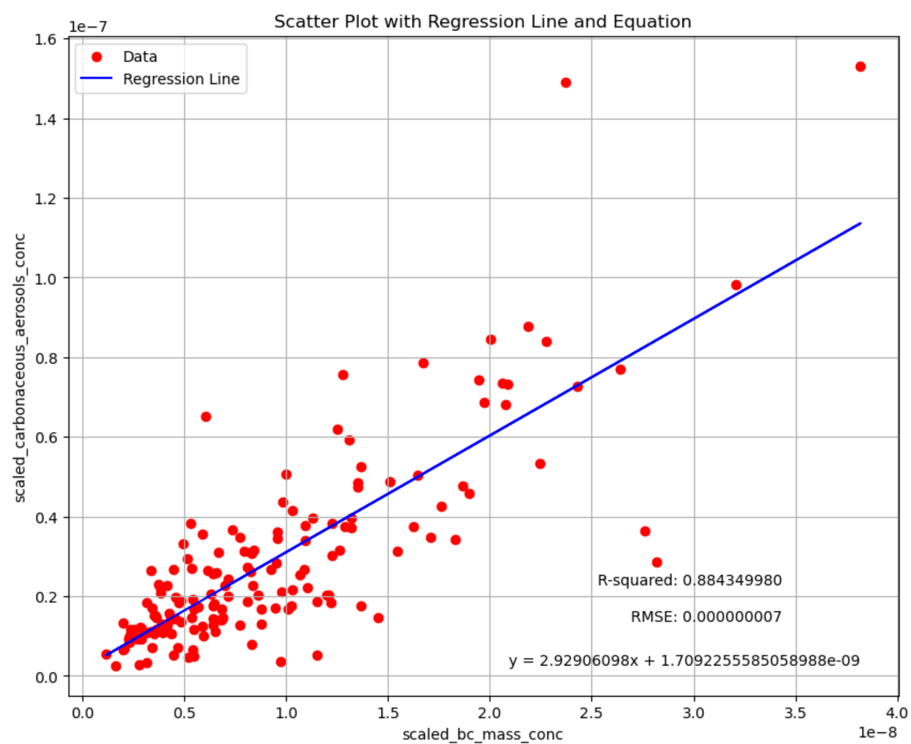

Figure 5.12: Regression model for prediction of ground level BC concentration

x (Independent Variable) represents the scaled black carbon surface mass

concentration
y (dependent Variable) represents the scaled carbonaceous aerosols concentration or ground data concentration.
70% was used for training and 30% for testing. So, if we have satellite and merra-2 values we can predict ground values.

Chapter 6

Conclusion

6.1 Conclusion

In this study, high resolution aerosols optical depth (AOD) from Level-1 and Atmosphere Archive & Distribution System Distributed Active Archive Center (LAADS-DAAC) with 1 km resolution were processed. Aerosols in ambient atmosphere are represented through AOD and it is difficult to measure particulate matter directly from satellite. In this study, high resolution AOD at 1 km was derived from October, 2022 to March 2023. MERRA-2 reanalysis data was used to develop a linear model for prediction of Black Carbon fraction of Carbonaceous aerosols (CA's) and effect of fires in nearby areas on BC concentration over satellite data.

AOD at 470 nm represents blue wavelength and this is sensitive to smaller particles (fine mode aerosols) such as smoke, urban pollution and secondary organic aerosols. AOD at 470 nm better correlates with pollution from vehicles, industrial activities and biomass burning. AOD at 470 nm is effective for monitoring of urban and industrial pollution where fine particles are predominant.

The monthly analysis of satellite based AOD values at 470 nm on daily basis from October 2022 to March 2023 indicates that October has least AOD values followed by December and maximum mean AOD values was found for November and February. The range of AOD values between 0.46 to 0.63 represents higher AOD values and this scenario is common in Delhi and there are days with extreme pollution event but these are regular AOD variations that are observed. AOD at 470 nm is generally higher compared to longer wavelengths if the atmosphere contains a significant amount of small particles. Thus, this indicates the presence of higher amount of finer particles in atmosphere and which are more dangerous for human health and it is deduced that in month of November and February residents need to be more careful for their practice of using mask in exposed environment. Analyzing AOD values, scientists and policymakers can infer the types and intensities of pollution sources, aiding in effective environmental management and health protection strategies.

The seasonal Mann-Kendall test of trend analysis over AOD is done to assess whether there is a monotonic trend which means the trend is either increasing or decreasing in the time series data. The p-value associated with the test (0.047) is below the significance level of 0.05. This indicates that the correlation factor observed is statistically significant. Therefore, it can be concluded that there is a statistically significant increasing trend in the aerosols optical depth values at 470nm over time. The result of the hypothesis (h) is TRUE implies that the trend detected by this test is statistically significant.

Black carbon (BC) is a type of very fine particulate matter ($PM_{2.5}$) produced from incomplete combustion of fossil fuels, biofuel, and biomass burning. It is a primary component of soot and is known for its ability to absorb light and heat, making it a significant contributor to global warming and air pollution.

For Delhi, during October-November there is a sight where blackness in ambient air is environment. The BC surface mass concentration was derived from MERRA-2 data with $0.5^{\circ} \times 0.625^{\circ}$ resolution with region of interest set for Delhi. The monthly analysis of BC on daily basis showed that mean BC mass concentration for Oct-Dec was $5.27 \times$ 10-9 kg/m3 and for Jan-Mar was $3.14 \times$ 10-9 kg/m3. This shows that Oct-Dec has higher BC concentration in atmosphere and this correlates well with several studies done for Delhi for this period. The higher concentration of BC as attributed by other studies also is due to burning activities in nearby areas, festival celebrations and vehicle exhaust emissions. The use of BC from MERRA-2 is helpful in identifying the fractions of BC in particulate matter and identifying the sources that are creating blackness in ambient air. This is experienced by residents of Delhi as blackening of masks and clothes and more prominent white clothes is observed during October-November and people spending sufficient time in open atmosphere experience more effect of blackening of air.

During October and November, fires in neighboring states of Delhi, particularly Punjab and Haryana, significantly contribute to air pollution in the region. This period is marked by widespread agricultural burning, commonly known as stubble burning, where farmers set fire to the leftover paddy straw after harvesting rice to quickly clear the fields for the next crop, typically wheat.

Black carbon surface mass concentration from MERRA-2 were analyzed for time period of 3 months (October-December). It can be seen that BC and fire count are following similar pattern from 15th October to 15th November which is the period of festivals when firecrackers are burned and are source of BC in ambient environment and stubble burning also contributes to BC concentration. This shows that higher BC concentrations are due to fire counts and the source of BC in ambient is stubble burning and firecrackers. After

15th November fire count starts reducing and so does BC. December marks high concentration of BC and with almost no fire count showing the traffic and other sources contributing to BC rather than farm fires. This overlay can help in identifying the source contribution of BC. Thus, from this study, it can be concluded that for understanding the distribution of fine particles in ambient atmosphere AOD at 470nm has to be used. The high resolution of AOD values makes it possible to be integrated in local area management plan of city for pollution monitoring, pattern detection and analysis. The overlay of fire count and BC showed that it is possible to identify the source contribution of BC during different months. The integration of machine learning models for prediction of AOD is a future for air quality management.

References

Chen, N., Yang, M., Du, W., and Huang, M. (2021). Pm2. 5 estimation and spatial-temporal pattern analysis based on the modified support vector regression model and the 1 km resolution maiac aod in hubei, china. *ISPRS international journal of geo-information*, 10(1):31.

Duc, H. N., Shingles, K., White, S., Salter, D., Chang, L. T.-C., Gunashanhar, G., Riley, M., Trieu, T., Dutt, U., Azzi, M., et al. (2020). Spatial-temporal pattern of black carbon (bc) emission from biomass burning and anthropogenic sources in new south wales and the greater metropolitan region of sydney, australia. *Atmosphere*, 11(6):570.

He, Q., Zhang, M., Huang, B., and Tong, X. (2017). Modis 3 km and 10 km aerosol optical depth for china: Evaluation and comparison. *Atmospheric environment*, 153:150–162.

Kumar, N., Chu, A., and Foster, A. (2007). An empirical relationship between pm2. 5 and aerosol optical depth in delhi metropolitan. *Atmospheric Environment*, 41(21):4492–4503.

Lin, C., Li, Y., Lau, A. K., Deng, X., Tim, K., Fung, J. C., Li, C., Li, Z., Lu, X., Zhang, X., et al. (2016). Estimation of long-term population exposure to pm2. 5 for dense urban areas using 1-km modis data. *Remote sensing of environment*, 179:13–22.

Majumdar, D. (2023). Spatial distribution and temporal variation of biomass burning and surface black carbon concentrations during summer (2015–2021) in india. *Air Quality, Atmosphere & Health*, 16(3):459–476.

Nair, M., Bherwani, H., Kumar, S., Gulia, S., Goyal, S., and Kumar, R. (2020). Assessment of contribution of agricultural residue burning on air quality of delhi using remote sensing and modelling tools. *Atmospheric Environment*, 230:117504.

Sathe, Y., Kulkarni, S., Gupta, P., Kaginalkar, A., Islam, S., and Gargava, P. (2019). Application of moderate resolution imaging spectroradiometer (modis) aerosol optical depth (aod) and weather research forecasting (wrf) model meteorological data for assessment of fine particulate matter (pm2. 5) over india. *Atmospheric Pollution Research*, 10(2):418–434.

Soni, K., Kapoor, S., Parmar, K. S., and Kaskaoutis, D. G. (2014). Statistical analysis of aerosols over the gangetic–himalayan region using arima model based on long-term modis observations. *Atmospheric Research*, 149:174–192.

Srivastava, P., Dey, S., Agarwal, P., and Basil, G. (2014). Aerosol characteristics over delhi national capital region: a satellite view. *International Journal of Remote Sensing*, 35(13):5036–5052.

Sun, L., Li, R.-B., Tian, X.-P., Wei, J., et al. (2017). Analysis of the temporal and spatial variation of aerosols in the beijing-tianjin-hebei region with a 1 km aod product. *Aerosol and Air Quality Research*, 17(3):923–935.

Vohra, K., Marais, E. A., Suckra, S., Kramer, L., Bloss, W. J., Sahu, R., Gaur, A., Tripathi, S. N., Van Damme, M., Clarisse, L., et al. (2021). Long-term trends in air quality in major cities in the uk and india. *Atmos. Chem. Phys*, 21:6275–6296.

Soni et al. (2014); Majumdar (2023); Sun et al. (2017); Srivastava et al. (2014); Kumar et al. (2007); Lin et al. (2016); Nair et al. (2020); Chen et al. (2021); Vohra et al. (2021); Duc et al. (2020); He et al. (2017); Sathe et al. (2019)

UNIVERSITY DISSERTATION Kailash-1.pdf

ORIGINALITY REPORT

8%	5%	5%	3%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	www.science.gov Internet Source	<1%
2	www.mdpi.com Internet Source	<1%
3	Submitted to American College of Education Student Paper	<1%
4	Moorthy Nair, Hemant Bherwani, Suman Kumar, Sunil Gulia, Sanjeev Goyal, Rakesh Kumar. "Assessment of contribution of agricultural residue burning on air quality of Delhi using remote sensing and modelling tools", Atmospheric Environment, 2020 Publication	<1%
5	ladsweb.modaps.eosdis.nasa.gov Internet Source	<1%
6	Ashwin Rachha, Mohammed Seyam. "Explainable AI In Education : Current Trends, Challenges, And Opportunities", SoutheastCon 2023, 2023 Publication	<1%