Air Pollutants as a significant predictor of SARS-CoV-2 transmission in the National Capital Region (Delhi NCR)

A DISSERTATION SUBMITTED IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE AWARD OF DEGREE

OF

MASTER OF TECHNOLOGY

IN

ENVIRONMENTAL ENGINEERING

Submitted by

GAURAV YADAV

2K20/ENE/04

Under the supervision of

DR. GEETA SINGH

DEPARTMENT OF ENVIRONMENTAL ENGINEERING

DELHI TECHNOLOGICAL UNIVERSITY

(Formerly Delhi College of Engineering)

Bawana Road, Delhi-110042

MAY 2022

CANDIDATE'S DECLARATION

I, Gaurav Yadav, Roll No. 2K20/ENE/04 student of MTech (Environmental Engineering), hereby declare that the project Dissertation titled "Air Pollutants as a significant predictor of SARS-CoV-2 transmission in National Capital Region (Delhi NCR)" which is submitted by me to the Department of Environmental Engineering, Delhi Technological University, Delhi in partial fulfilment of the requirement for the award of the degree of Master of Technology is original and not copied from any source without proper citation. This work has not previously formed the basis for the award of any Degree, Diploma Associateship, Fellowship or other similar title or recognition.

Place: Delhi Gaurav Yadav

Date: MAY 2022

CERTIFICATE

I hereby certify that the Project Dissertation titled "Air Pollutants as a significant predictor of SARS-CoV-2 transmission in National Capital Region (Delhi NCR)" which is submitted by **Gaurav Yadav, Roll No. 2K20/ENE/04**, Department of Environmental Engineering, Delhi Technological University, Delhi in partial fulfilment of the requirement for the award of the degree of Master of Technology is a record of the project work carried out by the student under my supervision. To the best of my knowledge this work has not been submitted in part or full for any Degree or Diploma to this University or elsewhere.

Place: Delhi **Dr. Geeta Singh**

Date **SUPERVISOR**

ACKNOWLEDGEMENT

I want to express my deepest gratitude to my supervisor Dr. Geeta Singh, Assistant Professor, Department of Environmental Engineering, Delhi Technological University, New Delhi, for her guidance, help, useful suggestions and supervision without which this report could not have been possible in showing a proper direction while carrying out the project. I also must acknowledge the unconditional freedom to think, plan, execute and express, that I was given in every step of my project work while keeping faith and confidence in my capabilities.

GAURAV YADAV

2K20/ENE/04

ABSTRACT

The continuous rise in the cases of the COVID-19 virus has created a challenging situation for the public health system of India. The cities with higher populations and poor air quality have been hard hit by coronavirus and fatality rates when the positivity rate increased during each wave. This study aims to examine the relation to whether the criteria air pollutants have promoted the spread of COVID-19 in Delhi NCR by regularizing the issue of multicollinearity among air pollutants.

It was hypothesized that criteria air pollutants would positively predict the spread of the coronavirus during the first wave (September 01, 2020 – December 11, 2020), the second wave (April 01, 2021 – May 22, 2021), and the third wave (January 01, 2022 – January 30, 2022) in the 25 most polluted districts of Delhi NCR. Elastic net-applied regularization was used in model exploration and coefficient estimation using EVIEWS 12 for testing of hypothesis. It was found that throughout each wave, $PM_{2.5}$ was among the most significant predictor of COVID-19 transmission in all districts of Delhi NCR.

This research explains the link between air pollutants and daily new corona case instances or how they individually contribute to the virus's transmission. Furthermore, exposure to areas with poor air quality may enhance the susceptibility and negatively impact the diagnosis of patients suffering from coronavirus.

Table of Content

List of Figures

List of Tables

CHAPTER 1

INTRODUCTION

Wuhan, a city, located in China, has become the control centre of a new coronavirus (SARS-CoV-2), which has given rise to a new disease named COVID-19 the February month of 2020 by WHO. (N. Zhu et al., 2020). The Indian government reported around 42,802,505 corona cases and 5,11,262 fatalities as of February 19, 2022 (www.mohfw.gov.in). In most Indian cities, due to high population density and poor air quality, COVID-19 infection is very high. Out of the 100 most polluted cities, 46 are from India (The 2020 World Air Quality Report). In India, Delhi NCR is distinguished as the region with the most polluted districts concerning air pollution. As to the Indian Government, the influence of coronavirus infection in all these districts is powerful during each wave (www.mohfw.gov.in).

Emerging epidemic and experimental data now suggest the involvement of air pollution in COVID-19-related effects. The existing literature on many respiratory infections indicates that air pollution may contribute to COVID-19-related morbidity and mortality. In many scientific studies, it has been observed that the transmission of COVID-19 is related to air pollution (Biktasheva, 2020). Many studies are available in which it has been examined that the combined impact of meteorological variables and air pollution will further lead to health problems in humans away lead to mortality (Ravindra et al., 2019). Chronic exposure to air pollution contributes to adverse health outcomes associated with many chronic diseases, most of which overlap with risk factors for severe COVID-19. The receptor for SARS-CoV-2 angiotensin-converting enzyme 2 (ACE2) is expressed in numerous organs explaining the wide range of symptoms associated with COVID-19 (Brandt & Mersha, 2021). As shown in Figure 1, coronavirus and air pollutants impact the same organ. The COVID-19 course demonstrates a vital role in the presence of comorbidities, which significantly increases the patient's risk profile. It includes many incurable diseases such as heart and body diseases and air pollution that hurts long-term and severe short-term exposure to air pollution. This may explain why exposure to air pollution can indirectly place people in the more powerful and dangerous forms of COVID-19, as suggested by epidemiological studies (Bourdrel et al., 2021) In, Figure 2, Figure 3, and Figure 4 from the collected dataset of this study it can easily be observed that during three waves the fluctuations in corona cases are influenced by some pollutants in the NCR region.

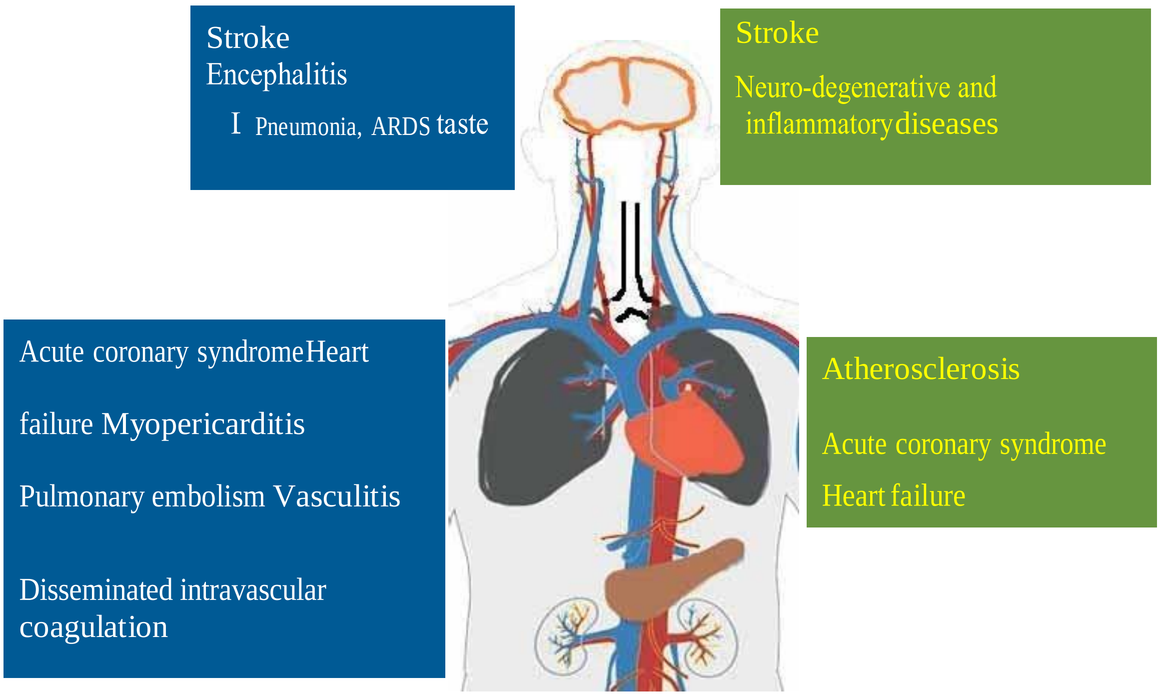


Figure 1. Target organs and the primary diseases that COVID-19 disease 2019 (blue) and air pollution (green) (Bourdrel et al., 2021)

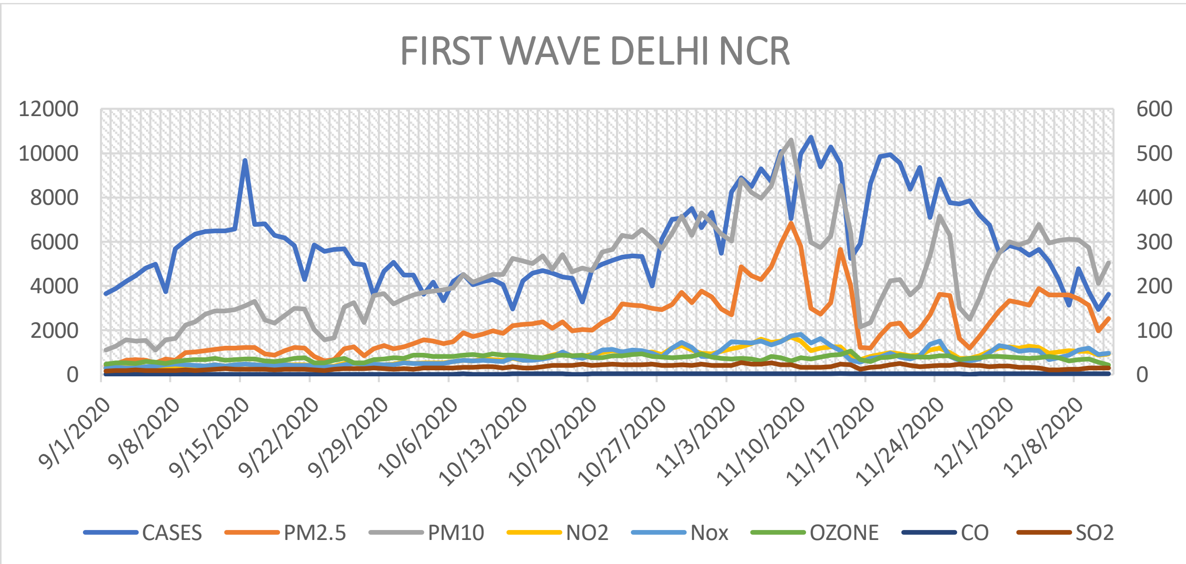


Figure 2. Daily COVID-19 cases and daily 24-hour mean concentration of air pollutants during the first wave

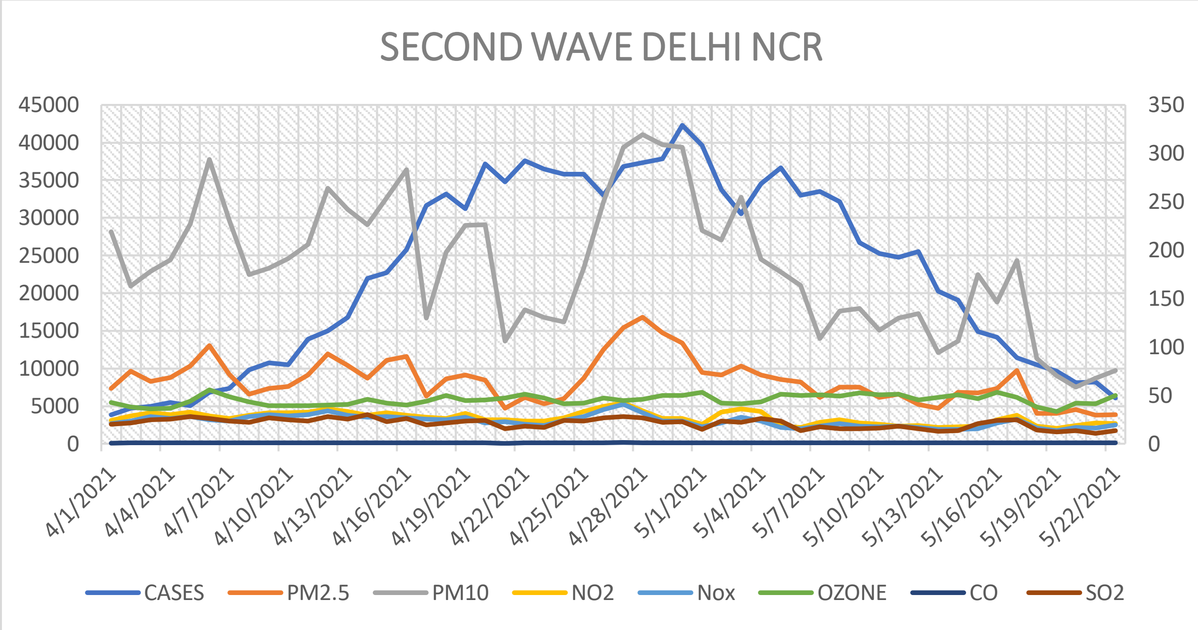


Figure 3. Daily COVID-19 cases and daily 24-hour mean concentration of air pollutants during the second wave

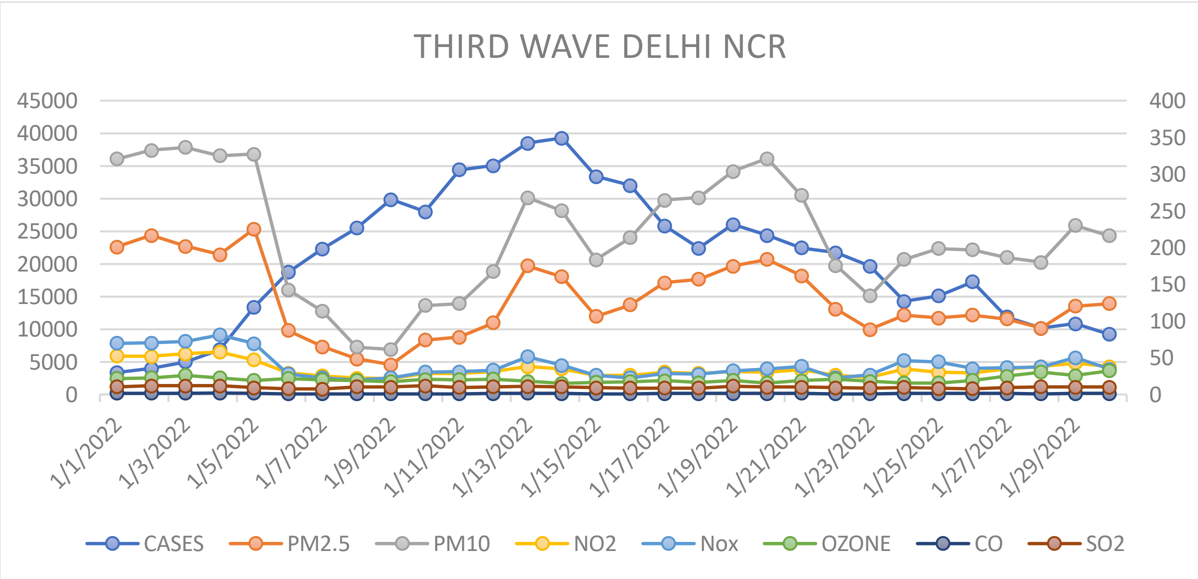


Figure 4. Daily COVID-19 cases and daily 24-hour mean concentration of air pollutants during the third wave

The previous studies which were finding relationship among SARS-CoV-2 and criteria air pollutants in India, and primarily focusing on Delhi were i.e., (Dutta & Dutta, 2021) used Spearman rank correlation analysis but found no positive association among air pollutants and the spread of coronavirus cases throughout two waves (April 2020 and April 2021), only O_3 shown an positive association with SARS-CoV-2 mortality during each wave; (Kolluru et al., 2021) studied five Indian megacities (including Delhi) in which a positive Pearson connection among $PM_{2.5}$, PM_{10}, and CO with daily new confirmed cases was reported; (Meo et al., 2022) by using Poisson regression analysis concluded that in India's largest metropolitan cities CO, O_3, and NO_2 were positively associated to COVID-19 daily cases and deaths in Delhi from March 02, 2020 to March 15, 2021; and (Wen et al., 2022) concluded that $PM_{2.5}$, increased SARS-CoV-2 cases by 0.439% in 10 of India's worst-affected states from March 9, 2020 to September 20, 2020. The available scientific literature worldwide spotlights that the presence of pollutants in the air is one of the essential sources of COVID-19-related transmissibility and deaths (Ayoub Meo et al., 2021).

So far, the studies conducted in the Delhi region primarily examined correlations at the bivariate level, which has the drawback of not considering the presence or effect of other independent variables between the two being investigated. Whereas the studies done at the multivariate level have excluded the issue of multicollinearity among the predictors. Air pollutants showed a positive correlation at the bivariate level, which further negatively impacted the analysis and severely limited the coefficient estimation.

1.1 Objective of study

To address these constraints, this study was undertaken to assess the part of air pollutants in COVID-19 transmission over National Capital Region (NCR) during three different study periods. It was hypothesized that all the criteria air pollutants (PM_{10}, $PM_{2.5}$, NO_2, NO_X, O_3, CO, SO_2) would positively predict COVID-19 propagation in Delhi NCR. To test this hypothesis, elastic net regression analysis was used because it is a model with minimal crucial independent variables; additionally, it punishes the fitting of unwanted independent variables. However, none of the studies that used elastic-net have been published to find an association between air pollution and COVID-19 transmission in India.

Many statisticians proved that elastic-net produces accurate results regarding the lucidity and precision of regression models (Edouard Grave, 2011). Elastic-net regression provides the most uncompromising model with perfect prediction when predicting a dependent variable with numerous predictors (Kim et al., 2016). Through cross-validation, it was also discovered that model founded by elastic-net regression predicts accurate results out of the bounds of data required for regression analysis (McNeish, 2015).

CHAPTER 2

LITERATURE REVIEW

A century after the Spanish Influenza Pandemic, a study reported that cities with more coal had tens of thousands of deaths from flu outbreaks compared with cities using less coal, taking into account factors such as socioeconomic status and primary health care (Clay et al., 2018; Pope et al., 2004). Prolonged exposure to air pollution is associated with increased influenza deaths in the USA (Pope et al., 2004). In 2003, a Chinese study on SARS-CoV-1 showed that people living in highly polluted areas were twice as likely to die due to SARS compared to people living in less polluted areas (Cui et al., 2003).

In 2015, a study conducted in 195 countries concluded that air pollution is a significant factor in the burden of lower respiratory infections (Troeger et al., 2017). As shown by the flu, exposure to air pollution increases the severity of respiratory infections caused by bacteria (Ciencewicki & Jaspers, 2007). In addition, several studies had shown that increased air pollution concentration was associated with an increase in respiratory infections among children and adults, especially when viral infection was accompanied by a temporary increase in exposure to air pollution. An increase in $PM_{2.5}$ concentrations has been associated with increased viral infections, namely influenza, respiratory syncytial virus (RSV), and measles. In these studies, the concentration of $PM_{2.5}$ was associated with several new cases of respiratory infection and a delay period of several days (Chen et al., 2017).

2.1 Studies Between Air pollutants and Covid-19:

Various studies have reported a link between exposure to air pollution and COVID-19 disease and mortality worldwide which are as follows: -

2.1.1 Particulate matters ($PM_{2.5\&10}$) and COVID-19

$PM_{2.5}$ can invade deeply into the lungs and deposit into alveoli. (Liu et al., 2021) done a study in 9 countries using discontinuous linear regression between January 21 to May 20, 2020, and found that PM_{10} plays a more substantial role in accelerating the spread of COVID-19 infection in China, England, Germany, and France.

Another study conducted in Milan of, Italy, between January 01 to April 30, 2020, by (Zoran et al., 2020b) found that daily maximum $PM_{2.5}$ and PM_{10} were positively associated with new COVID-19 cases.

Using a linear regression model with the dataset of Wuhan and Xiaogan (China) between January 26 to February 29, 2020 (Li et al., 2020) found that $PM_{2.5}$ was prominently correlated with COVID-19 incidence. (Jiang & Xu, 2021) did a study in the origin place of coronavirus Wuhan between January 25 to April 07, 2020, using Pearson's and Poisson's regression models, found that $PM_{2.5}$ was positively associated (relative risk [RR] = 1.079, 95% CI 1.071-1.086, $P < 0.01$) with COVID-19 deaths and PM_{10} was inversely related to COVID-19 deaths.

In 337 prefecture-level cities of China, using Spearman's rank correlation analysis and multiple linear regression, it was observed by (Q. Wang et al., 2021) that $PM_{2.5}$ and PM_{10} were positively correlated with newly confirmed findings of COVID-19 cases. Another study by (Pei et al., 2021) in 325 cities of China using geographically weighted regression up to May 27, 2020, concluded that $PM_{2.5}$ and PM_{10} had significantly positive impacts on COVID-19. Similarly, (Y. Zhu et al., 2020) conducted a study in 120 cities in China using a Generalized additive model between January 23 to February 29, 2020, and concluded 10 mg/m^3 increase in $PM_{2.5}$ was positively associated with 2.24% (95% CI: 1.02-3.46) increase in the daily counts of confirmed cases; 10 mg/m3 increase in PM_{10} was positively associated with 1.76% (95% CI: 0.89-2.63) increase in the daily counts of confirmed cases. (PM_{10}, $PM_{2.5}$, NO_2, NO_X, O_3, CO, SO_2)

Using Bayesian hierarchical models (Konstantinoudis et al., 2021) executed, a study in England up to June 30, 2020. It was observed that every 1 μg/m^3 increase in $PM_{2.5}$ was associated with a 1.4% (95% CI: −2.1%-5.1%) increase in COVID-19 mortality risk. (Mele & Magazzino, 2021) using correlation analysis in the dataset of Italy's $PM_{2.5}$ concentration and Daily new COVID-19 cases found that $PM_{2.5}$ was positively associated with the total number of COVID-19 cases.

Further, (Frontera et al., 2020) conducted a new large study in which he took a large dataset of 47 regional European Capitals and 107 major Italian cities for a time period of February 10 to April 10, 2020, and used a Binary classifier based on an artificial neural network to analyze the large collected dataset observed that $PM_{2.5}$ and PM_{10} were positively associated with the number of COVID-19 cases. (Setti et al., 2020) also conducted a study at a bivariate level in 110 Italian provinces from February 07 to March 15, 2020. He was the first to observe that the average number of exceedances of the PM_{10} daily limit value was positively associated with the number of COVID-19 cases in each province.

In 63 cities in China (B. Wang et al., 2020) used generalized additive models (GAM) with a quasi-Poisson's distribution on the dataset from January 01 to March 02, 2020, and found that a 10µg/m^3 increase in the concentration of PM_{10} and $PM_{2.5}$ were positively associated with the confirmed cases of COVID-19. The estimated strongest RRs (both at lag 7) were 1.05 (95% CI: 1.04-1.07) and 1.06 (95% CI: 1.04-1.07), respectively.

(Travaglio et al., 2021) by using the dataset of the UK biobank for 2018-2019 and by using generalized linear models, negative binomial regression analysis Concluded that an increase of 1m^3 in the long-term average of $PM_{2.5}$ was associated with a 12% increase in COVID-19 cases and a one-unit increase in PM_{10} was associated with approximately 8% more COVID-19 issues in the UK biobank.

A large study conducted worldwide by (Pozzer et al., 2020) up to June 2020 using the Global atmospheric chemistry general circulation model (EMAC) found that $PM_{2.5}$ contributed 15% (95%CI: 7%-33%) to COVID-19 mortality worldwide. A time-series analysis study was done in the origin place of coronavirus Wuhan between January 19 to March 15, 2020, and it is observed that $PM_{2.5}$ and PM_{10} were positively associated with the case fatality rate of COVID-19 (CFR). (Yao et al., 2020)

(Magazzino et al., 2020) conducted a study in Paris, Lyon, and Marseille using Artificial Neural Networks (ANNs) experiment Machine Learning (ML) methodology for data analysis found that $PM_{2.5}$ and PM_{10} directly correlated with COVID-19 fatality. In the Northern region of Italy (Coker et al., 2020), by using Negative binomial regression for January 01 to April 30, 2020, concluded that a one-unit increase in $PM_{2.5}$ concentration (µg/m^3) was associated with a 9%(95% CI: 6%-12%) increase in COVID-19 related mortality.

In 20 districts in Lima (Peru) (Vasquez-Apestegui et al., 2021) conducted, an ecological study using linear regression found that higher $PM_{2.5}$ levels were associated with a higher number of cases and deaths of COVID-19. In the USA (Hendryx & Luo, 2020), using Mixed model linear multiple regression analyses, Greater diesel particulate matter (DPM) was significantly associated with COVID-19 prevalence and mortality rates. (Lembo et al., 2021) in 33 European countries, Pearson's correlation analysis found that $PM_{2.5}$ was positively correlated with positive COVID-19 cases and deaths.

(Jiang et al., 2020) again conducted a study in 3 regions of China (Wuhan, Xiaogan, Huanggang) from January 25 to February 29, 2020, using Multivariate Poisson's regression analysed that $PM_{2.5}$ was positively associated with daily COVID-19 incidence in Wuhan (1.036, 95% CI: 1.032-1.039), Xiaogan (1.059, 95% CI: 1.046-1.072), and Huanggang (1.144, 95% CI: 1.12-1.169), and PM_{10} was negatively associated with daily COVID-19 incidence in Wuhan (0.964, 95% CI: 0.961-0.967), Xiaogan (0.961, 95% CI: 0.95-0.972), and Huanggang (0.915, 95% CI: 0.896-0.934).

(Wu et al., n.d.) using mixed binomial methods in 3000 counties of the USA concluded that a 1 mg/m3 increase in $PM_{2.5}$ was positively associated with an 8% increase in the COVID-19 death rate (95% CI: 2%-15%). In 71 provinces of Italy, a study was conducted by (Fattorini & Regoli, 2020) on April 27, 2020, and found that $PM_{2.5}$ and PM_{10} were favourable for the spread of virulence SARS-CoV-2.

(Bontempi, 2020) done study at a bivariate level using correlation analysis in 14 cities of Italy between February 10 to March 27, 2020, and observed no evidence of correlations between the presence of high quantities of PM_{10} and COVID-19 cases. Facilitate transmission of SARS-CoV-2 virus droplets and PM in indoor environments (Amoatey et al., 2020) in middle eastern countries.

By using new machine learning techniques for statistical analysis (Magazzino et al., 2021), New York observed that $PM_{2.5}$ accelerated COVID-19 death. (Liang et al., 2020) between January 22 to April 29, 2020, a study in 3122 counties of the USA with Zero-inflated negative binomial models showed no association between $PM_{2.5}$ and COVID-19. Similarly, (Adhikari & Yin, 2020) in Queens and New York, between March 01 to April 20, 2020, using a Negative binomial regression model, observed that a one-unit increase in the moving average of $PM_{2.5}$ ($\mu g/m^3$) was associated with a 33.11% (95% CI: 31.04-35.22) decrease in the daily new COVID-19 cases.

2.1.2 NO_2 and COVID-19

(Li et al., 2020) using linear regression in Wuhan and Xiaogan found that in both these cities of China, NO_2 was prominently correlated with COVID-19 incidence. During a similar period (Jiang et al., 2020) conducted a study using Multivariate Poisson's regression in one more additional city in China (hanging) and observed that NO_2 was positively correlated with daily

COVID-19 incidence in Wuhan (1.056, 95% CI: 1.053-1.059) and Xiaogan (1.115, 95% CI:1.095-1.136).

(Yao et al., 2021) in 11 Hubei cities, by using multiple linear regression, residual analysis, principal component analysis, and meta-analysis method from January 01 to February 08, 2020, observed that NO_2 concentration (with a 12-day time lag) was positively related to transmission ability (primary reproductive number) of the 11 Hubei cities (except Xining City).

In 3122 US counties from January 22 to April 29, 2020, using Zero-inflated negative binomial models (Liang et al., 2020) observed that per interquartile range (IQR) increase in NO_2 (4.6 ppb) was associated with an increase in COVID-19 case-fatality rate (7.1%, 95% CI: 1.2%-13.4%) and mortality rate (11.2%, 95% CI: 3.4%-19.5%), respectively.

(Travaglio et al., 2021) by using Generalized linear models, negative binomial regression analysis in England during 2018-2019 observed that NO_2 and NO were positively associated with COVID-19 infectivity, with an odds ratio of approximately 1.03 for both the single-year and multiyear model. In 66 administrative regions in Italy, Spain, France, and Germany from Jan to Feb 2020 (Ogen, 2020) concluded that NO_2 was positively correlated with COVID-19 fatality cases. Out of the 4443 fatality cases, 3487 (78%) were in five regions (have the highest NO_2).

In 29 provinces of China from January 21 to April 03, 2020 (Lin et al., 2020), by using the Chain-binomial model, correlation analysis found that NO_2 was inversely correlated to the primary reproductive ratio of COVID-19. In England, up to June 30, 2020 (Konstantinoudis et al., 2021), Bayesian hierarchical models observed that Every 1 μg/m3 increase in NO_2 was associated with a 0.5% (95% CI: −0.2%-1.2%) increase in COVID-19 mortality risk. In Milan of, Italy, from January 01 to April 30, 2020, using time series analysis (Zoran et al., 2020a), ground-level NO_2 was inversely correlated with COVID-19 infections.

In 9 countries using Discontinuous linear regression from January 21 to May 20, 2020 (Liu et al., 2021), observed that the aggravating effect of NO_2 on COVID-19 infection appears in Canada and France. In 33 European countries (Lembo et al., 2021) conducted a study at the bivariate level and found that NO_2 was positively correlated with positive COVID-19 cases and deaths. (Mele et al., 2021) using machine learning techniques in 3 major French cities from March 18 to April 27, 2020, found that NO_2 levels contribute to COVID-19 deaths and exist threshold values. (Magazine et al., 2021) also, machine learning experiments found that NO_2 accelerated COVID-19 deaths in 3 French cities.

Using a general additive model (Y. Zhu et al., 2020) concluded that in 120 cities in China, every 10 mg/m3 increase of NO_2 was associated with a 6.94% (95% CI: 2.38-11.51) increase in the daily counts of confirmed COVID-19 cases. (Suez et al., 2020) in Catalonia, Spain observed that NO_2 was significantly correlated with COVID-19 incidence, mortality, and lethality rates using Spearman's nonparametric correlation.

In 71 Italian provinces (Fattorini & Regoli, 2020) found that NO_2 was significantly correlated with cases of COVID-19. In 18 Indian states (Chakraborty et al., 2020) by using Pearson's correlation coefficient and regression analysis for data analysis, found that NO_2 showed a strong positive correlation between the absolute number of COVID-19 deaths ($r = 0.79$, $P < 0.05$) and case fatality rate ($r = 0.74$, $P < 0.05$).

(Filippini et al., 2020) with the help of a Multivariable restricted cubic spline regression model, concluded that in 28 northern provinces of Italy, NO_2 was significantly correlated with SARS-CoV-2 infection prevalence rate.

2.1.3 Surface Level Ozone (O_3) and COVID-19

(Liu et al., 2021) in 9 countries (China, Japan, Korea, Canada, America, Russia, England, Germany, and France) from January 21 to May 20, 2020, using Discontinuous linear regression for data analysis observed that O_3 presents a more pronounced positive effect on COVID-19 infection in more countries (such as Japan, Canada, America, Russia, France, etc.).

By using a general additive model for the dataset of 120 cities in China (Y. Zhu et al., 2020) found that with per 10 mg/m3 increase in O_3 was associated with a 4.76% (95% CI: 1.99-7.52) increase in the daily counts of confirmed cases, respectively.

In 47 regional capitals and 107 major cities in Europe, a study was conducted by (Fronza et al., 2020)using an artificial neural network for data analysis and observed that O_3 was negatively associated with the number of COVID-19 cases per million ($r = -0.44$). A study in the UK (Travaglio et al., 2021) using generalized linear models and negative binomial regression analyses concluded that O_3 was significantly associated with COVID-19 deaths and cases at the sub-regional level.

(Jiang et al., 2020) found that O_3 was negatively associated with daily COVID-19 incidence in Wuhan (0.99, 95% CI: 0.989-0.991) and Xiaogan (0.991, 95% CI: 0.989-0.993) and positively associated with daily COVID-19 incidence in Huanggang (1.016, 95% CI:

1.012-1.02) using Multivariate Poisson's regression in 3 Chinese cities.

In 3122 counties of the USA, a study was conducted using Zero-inflated negative binomial models (Liang et al., 2020) observed no significant associations between O_3 and COVID-19 cases.

In New York (Adhikari & Yin, 2020) conducted a study using negative binomial regression mode and concluded that a one-unit increase in O_3 was associated with a 10.51% (95% CI: 7.47-13.63) increase in the daily new COVID-19 cases. (Zoran et al., 2020b) Milan, Italy, conducted a time series analysis study and found that COVID-19 infections showed a positive correlation with ground-level O_3.

2.1.4 Carbon Monoxide (CO) and COVID-19

(Liu et al., 2021) while studying for CO dataset in 9 countries and using discontinuous linear regression for data analysis observed that CO will increase the propagation speed of COVID-19 infection, which is significant in COVID-19 disease, which is significant in Korea and China, respectively.

(Jiang & Xu, 2021) while studying found that CO was inversely associated with COVID-19 deaths from January 25 to April 07, 2020, in Wuhan, where this pandemic has originated.

(Q. Wang et al., 2021) in 337 prefecture-level cities in China, CO was positively correlated with newly confirmed cases using bivariate analysis. Similarly, (Pei et al., 2021) in China observed that CO harmed COVID-19 deaths up to May 27, 2020.

(Jiang et al., 2020) when studying for CO concluded that CO was positively correlated with the daily incidence in Wuhan (1.932, 95% CI: 1.763-2.118); but negatively correlated with the daily incidence in Xiaogan (0.041, 95%CI: 0.026-0.066) and Huanggang (0.032, 95%CI: 0.017-0.063). Similarly, (Lin et al., 2020), when studying in China, found that CO was positively correlated with the primary reproductive ratio of COVID-19.

2.1.5 Sulphur Di Oxide (SO_2) and COVID-19

SO_2 increased the propagation speed of COVID-19 infection, which is significant in Korea and China, respectively (Liu et al., 2021). SO_2 was inversely associated with COVID-19 deaths (Jiang & Xu, 2021). SO_2 was positively correlated with newly confirmed cases (Q.

Wang et al., 2021).

10 μg/m^3 increase of SO_2 was associated with a 7.79% decrease (95% CI: −14.57 to −1.01) in COVID-19 confirmed cases (Y. Zhu et al., 2020). SO_2 was positively correlated with positive COVID-19 cases and deaths (Lembo et al., 2021). SO_2 was not associated with daily COVID-19 incidence (Jiang et al., 2020).

In addition, to examine the link between air pollution and COVID-19 prevalence, several studies have analyzed the effect of air pollution on COVID-19 estimates. Although these studies answer a different question and may have their drawbacks, examining the prediction of COVID-19 is less affected by infection potential. It thus leads to more minor legal threats such as undiagnosed confusion. To remove the issue of multicollinearity present, this study has been conducted.

CHAPTER 3
MATERIALS AND METHODOLOGY

3.1 Study area

The study analysed the 25 most polluted districts in the Delhi NCR region. These districts lie under Delhi, Haryana, and Uttar Pradesh and are portrayed in two different Figure 5, and Figure 6. They have emerged as important hubs for the commercial, industrial, medical, and educational sectors, attracting people from all over the country. All these districts are struggling to handle the problem of air pollution and related health hazards due to the increase in population. In all these districts, the effect of COVID-19 has been highly observed.

3.1.1 Delhi

As shown in the Figure 5, all districts of Delhi (28.7041° N, 77.1025° E) with 38 CAAQMS MONITORING STATIONS have been selected for this study. The 24-hourly air quality and COVID-19 cases data have been collected from all these monitoring sites.

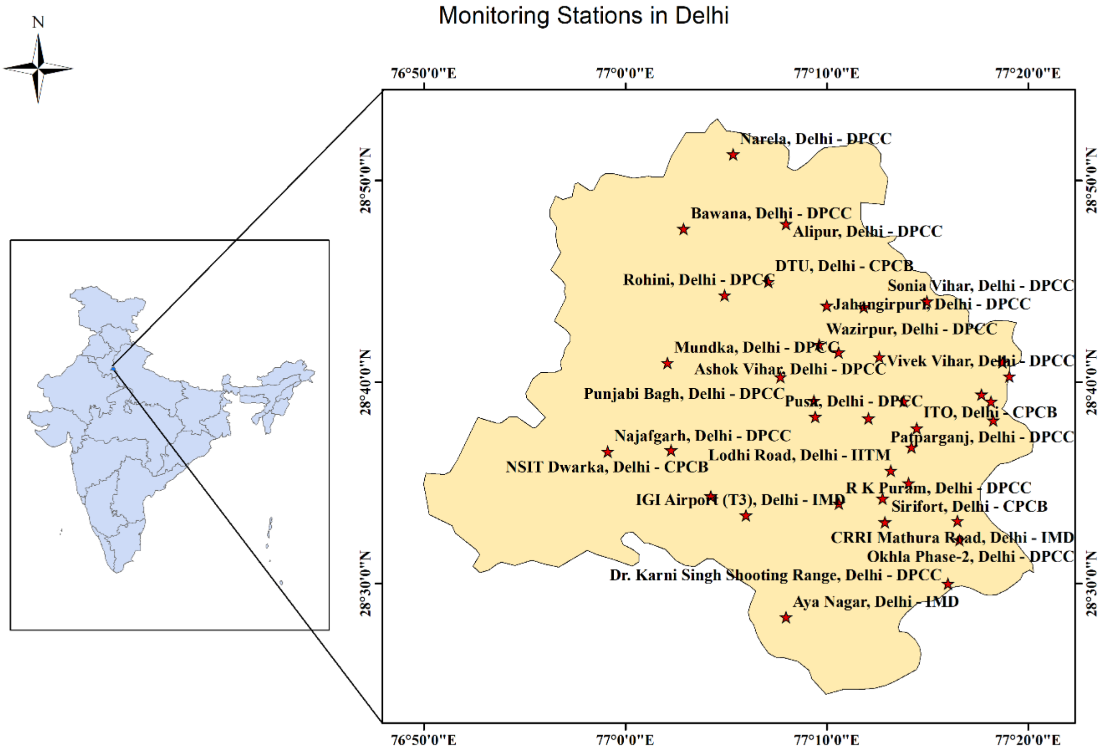


Figure 5. The study area (Delhi)

3.1.2 Haryana and Uttar Pradesh

Fifteen districts from Haryana (29.0588° N, 76.0856° E) and Uttar Pradesh (26.8467° N, 80.9462° E) with 23 CAAQMS MONITORING STATIONS have been selected for this study. All monitoring sites whose data has been used for this study have been portrayed in the Figure 6 using ArcGIS.

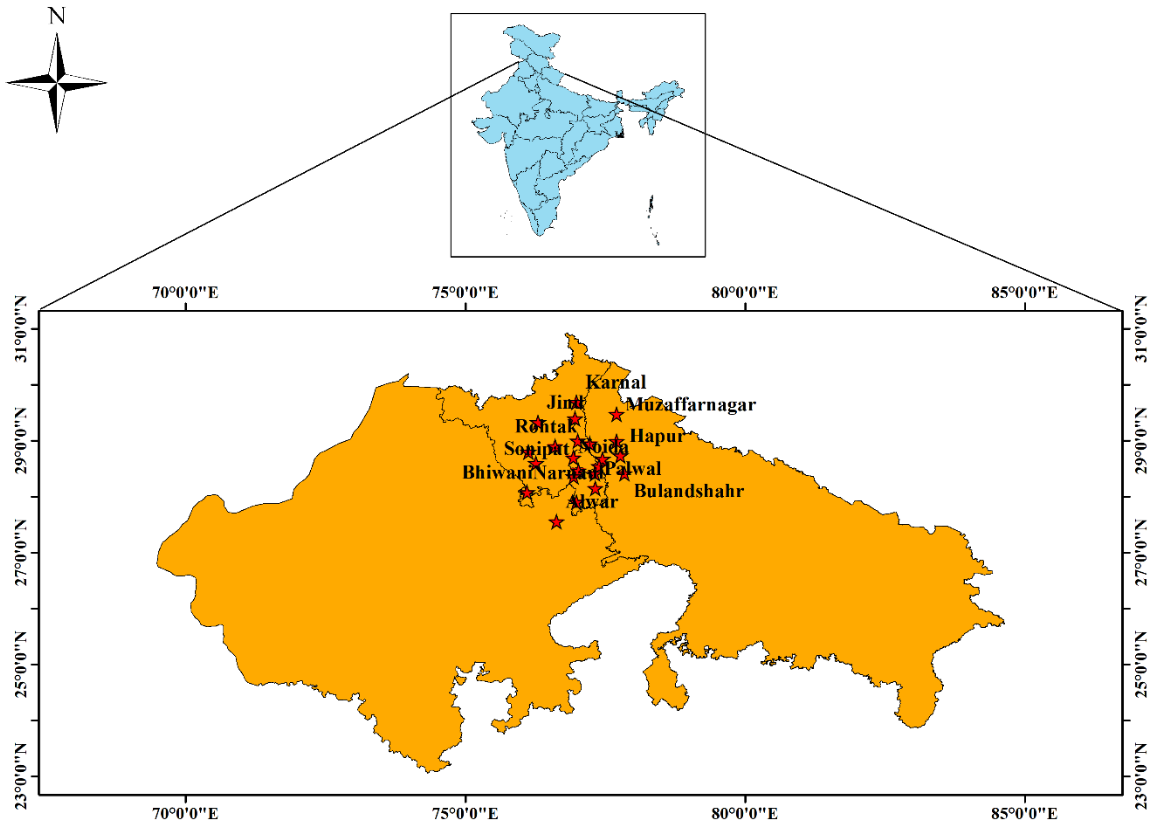


Figure 6. The study area of NCR

3.2 Data sources

For three study periods, i.e., (September 1, 2020 – December 11, 2020), (April 1, 2021 – May 22, 2021), and (January 1, 2022 – January 30, 2022), the total daily new COVID-19 cases were collected for all districts from (http://health.delhigovt.nic.in/), (http://nhmharyana.gov.in), (http://rajswasthya.nic.in), and (http://dgmhup.gov.in). The daily 24-hour mean concentration data of criteria air pollutants was collected from (https://app.cpcbccr.com/ccr) operated by CPCB (https://cpcb.nic.in). *Table 1* shows the average quality of the air and the sum of corona cases during each study period.

Table 1. The average and the total value of datasets

Variables	First wave	Second wave	Third Wave
Total Cases (Numbers)	612368	1184479	620572
Average $PM_{2.5}$ in µg/m^3	113.74	65.31	129.39
Average PM_{10} in µg/m^3	230.44	183.26	216.06
Average NO_2 in µg/m^3	40.35	26.52	34.29
Average NOx in ppb	40.07	23.32	39.53
Average O_3 in µg/m^3	36.84	45.13	20.37
Average CO in mg/m^3	1.2	0.8	1.28
Average SO_2 in µg/m^3	16.42	21	9.81

3.3 Data analysis

The acquired data set was analysed individually at three levels, univariate, bivariate, and multivariate. Every aspect influencing the impact of the ecological condition on human health is studied. Using three methods, primarily the trend in air quality during three waves was observed using ArcGIS (legend). Secondly, Pearson's correlations were calculated to determine the association among predictors; due to the presence of high association among pollutants. Stepwise standard multiple linear regression was calculated to determine the coefficient of determination (R^2), which will further help us determine VIF and tolerance level among multivariate predictors. In the last step, elastic net regression analysis will be used to predict a better model. All the methods used in this study are briefly described below sections.

3.3.1 Mapping of air quality in Delhi NCR using ArcGIS

The air maps were created using ArcGIS version 10.5, a cutting-edge GIS program, at all detected locations. Combining GIS with air models allows for the automatic generation of air data models from digital geographic data. The air data is collected, stored, managed, and controlled using a GIS database management system. The interpolation techniques available in GIS are used to create air pollutants concentration contours.

3.3.2 Pearson Correlation Coefficients: -

It is calculated by using the formulae stated in Equation (1): -

$$r = \frac{n\,(\sum XY) - (\sum X)\,(\sum Y)}{\sqrt{[n \sum X^2 - (\sum X)^2] * [n\,\sum Y^2 - (\sum Y)^2]}} \tag{1}$$

Wherein r is the Pearson correlation coefficient; n is the number of observations, and X and Y are the first and second variables in the context. This test was conducted to check the higher values of correlation coefficients among the selected predictors (air pollutants) for this study, which will further indicate the possibility of multicollinearity (correlation coefficients near 0.8 indicate the presence of collinearity)(Chatterjee, 2013).

3.3.3 Testing for multicollinearity

The variance inflation factor (VIF) is utilized to calculate how much variance of the predicted regression coefficient is inflated when the predictors are correlated. VIF is determined as stated in Equation (2): -

$$VIF = \frac{1}{1 - R^2} = \frac{1}{Tolerance} \tag{2}$$

Tolerance is the inverse of VIF, which means that the lower the tolerance, the more variables are prone to be multicollinear. Furthermore, VIF =1 shows that independent variables are not linked to each other. If VIF is around 1 to 5, it means that variables are relatively related to each other; when VIF is approximately 5 to 10, it means that variables are highly linked; if VIF ≥ 5 to 10. It means that there is multicollinearity among independent variables in the regression model, and VIF > 10 specifies that regression coefficients are not accurately estimated due to the presence of multicollinearity.

3.3.4 Elastic Net Regression

The hypothesis tests were conducted to find whether the criteria air pollutants carry an important influence on the spread of the corona virus in the 25 most polluted districts lying under Delhi NCR when the cases were spreading at a high rate. The daily new SARS-CoV-2 cases in all 25 districts of Delhi NCR were used as a dependent variable, and 24-hour mean air quality over all communities was taken as predictors to find how much variance each parameter was created on the spread of new daily corona cases. All the statistical analyses of elastic net regression were performed using EVIEWS 12 (www.eviews.com). Due to the presence of high multicollinearity (VIF > 10) between most of the predictor's elastic net-applied regularisation has been used in model exploration and coefficient estimation (Zou & Hastie, 2005) as shown in Equation (3), where α is the mixing parameter between ridge (α = 0) and lasso (α = 1).

$$L_{\text{enet}}\left(\hat{\beta}\right) = \frac{\sum_{i-1}^{n}\left(y_i - x_i'\hat{\beta}\right)^2}{2n} + \lambda\left(\frac{1-\alpha}{2}\sum_{j=1}^{m}\hat{\beta}_j^2 + \alpha\sum_{j=1}^{m}\left|\hat{\beta}_j\right|\right) \tag{3}$$

By minimizing regression coefficients to zero, the L1-norm (lasso) of penalty creates a sparse model. The L2-norm (ridge) of penalty removes the constraint on the number of selected variables, stimulates grouping, and stabilizes the L1 regularization route (Kuang et al., 2015). Elastic net reduces the regression coefficients by combining the L1-norm (lasso) and L2-norm (ridge) penalties; α = 0.5 was utilized, a midpoint among L1 and L2. Additionally, to reduce the threat of overfitting, elastic net regression was done with K-Fold cross-validation to measure the mean square error. The complete dataset was randomly split into ten folds. One block of split datasets (90% of the entire dataset) was used to calculate coefficients for every test.

Furthermore, the predictive performance of our prediction model was calculated with the remaining data block (10% of the entire dataset). For each test, a separate set of coefficients quantity was calculated while changing λ. Lambda (λ) at minimum error is used to calculate the most straightforward model with high prediction accuracy and regularized beta coefficients. To estimate an accurate algorithm OLS with 500 maximum iterations and 0.0001 convergence was used.

CHAPTER 4
RESULTS AND DISCUSSIONS

4.1 Air mapping of Air Pollutants

In most regions selected for study, a lockdown is imposed during each wave, but even after that, the air quality condition in most areas is deplorable. In the below sections, the trends of each pollutant during all three waves have been portrayed

4.1.1 Air mapping of $PM_{2.5}$ during each wave

The Figure 9 shows air quality mapping of $PM_{2.5}$ in the NCR region during the first COVID-19 wave (September 01 to December 11, 2020). $PM_{2.5}$ average concentration varies from 38.6 µg/m^3 to 308.11 µg/m^3 for the first wave and further shows that most districts of Delhi, Haryana, and Uttar Pradesh lying under the NCR region are facing the problem of bad air quality even after the lockdown. The monitoring site which recorded the worst average $PM_{2.5}$ concentration during the first wave was RK Puram ranging from 278.1 µg/m^3 to 308.11 µg/m^3. During the first wave, the monitoring stations with good average air quality were Bhiwani, Palwal, Alwar, and Hapur, ranging from 38.6 µg/m^3 to 68.5 µg/m^3.

The Figure 7 shows the air quality mapping of $PM_{2.5}$ in the NCR region during the second COVID-19 wave (April 01 to May 22, 2021). $PM_{2.5}$ average concentration varies from 20.5 µg/m^3 to 105.2 µg/m^3 for the second wave. The monitoring site which recorded the worst average $PM_{2.5}$ concentration during the second wave was Bawana ranging from 95.8 µg/m^3 to 105.2 µg/m^3. During the second wave, the monitoring stations with good average air quality were Bhiwani and Hapur, ranging from 20.5 µg/m^3 to 29.9 µg/m^3.

The Figure 8 shows the air quality mapping of $PM_{2.5}$ in the NCR region during the third COVID-19 wave (January 01 to January 31, 2022). $PM_{2.5}$ average concentration varies from 19.3 µg/m^3 to 222.7 µg/m^3 for the third wave. The monitoring site which recorded the worst average $PM_{2.5}$ concentration during the third wave was Bahadurgarh ranging from 204.5 µg/m^3 to 227.7 µg/m^3. During the third wave, the monitoring stations with good average air quality were Mandikhera, Palwal, and Rohtak, ranging from 19.3 µg/m^3 to 42.4 µg/m^3.

4.1.2 Air mapping of PM_{10} during each wave

The Figure 12 shows air quality mapping of PM_{10} in the NCR region during the first COVID-19 wave (September 01 to December 11, 2020). PM_{10} average concentration varies from 87.7 µg/m^3 to 361.9 µg/m^3 for the first wave. The monitoring sites which recorded the worst average PM_{10} concentration during the first wave were Rohini and Mundka, ranging from 330.8 µg/m^3 to 361.9 µg/m^3. During the first wave, the monitoring stations with good average air quality were Panipat and Alwar, ranging from 87.7 µg/m^3 to 112.9 µg/m^3.

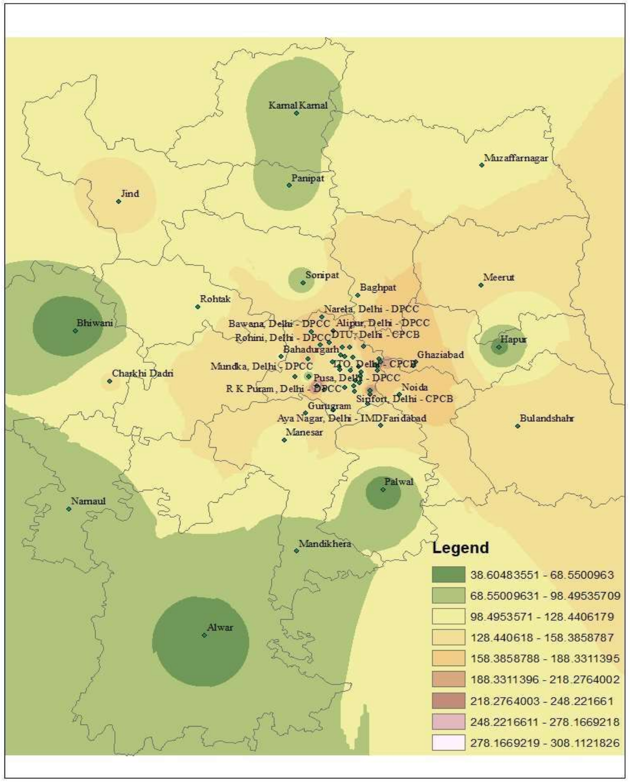


Figure 7. Air Mapping of $PM_{2.5}$ in Delhi NCR during first wave

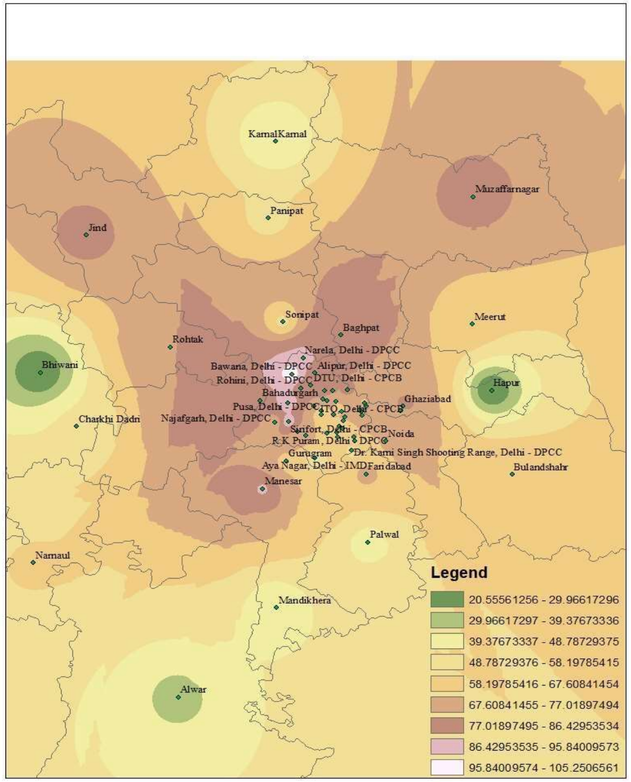


Figure 8. Air mapping of $PM_{2.5}$ in Delhi NCR during second wave

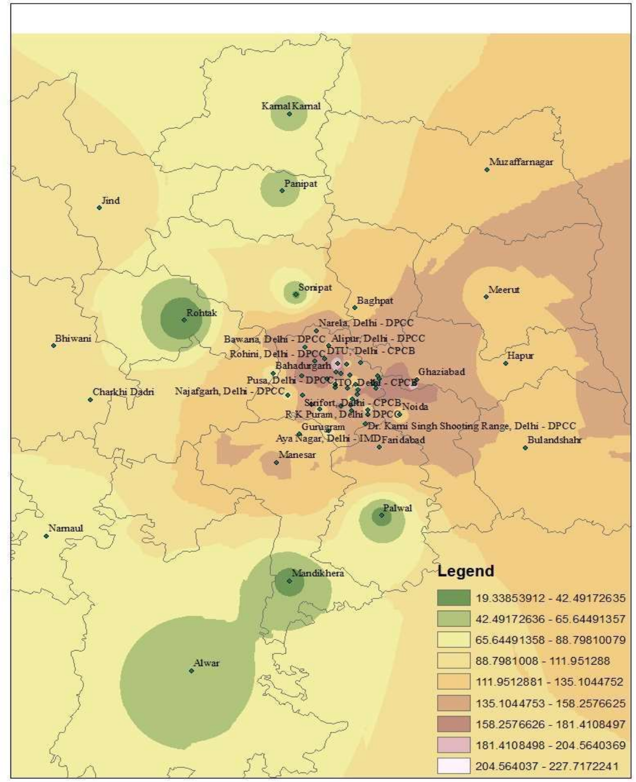


Figure 9. Air mapping of $PM_{2.5}$ in Delhi NCR during third wave third wave

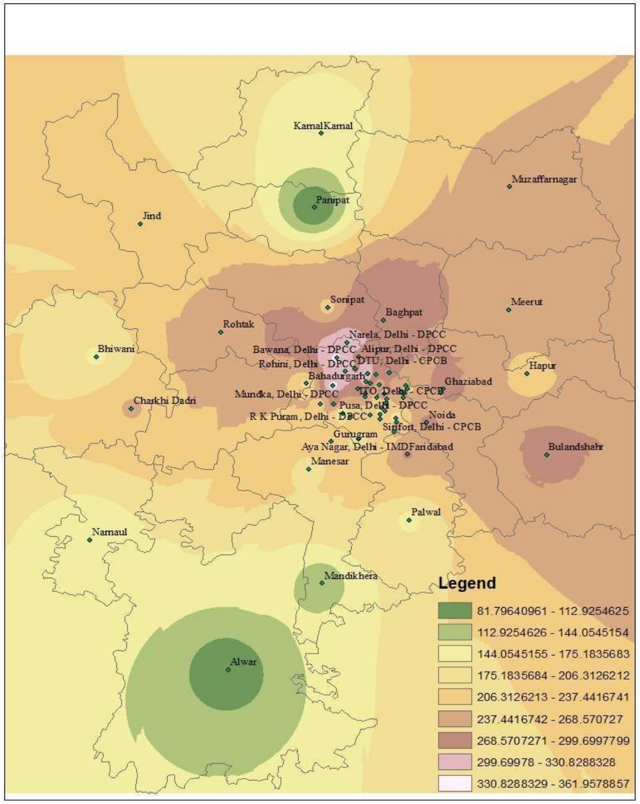


Figure 10. Air Mapping of PM_{10} in Delhi NCR during first wave

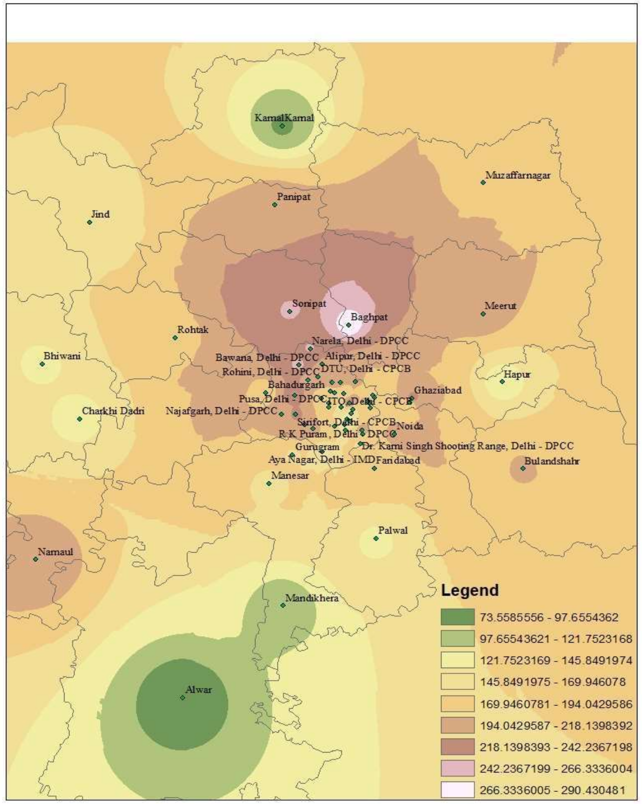


Figure 11. Air mapping of PM_{10} in Delhi NCR during second wave

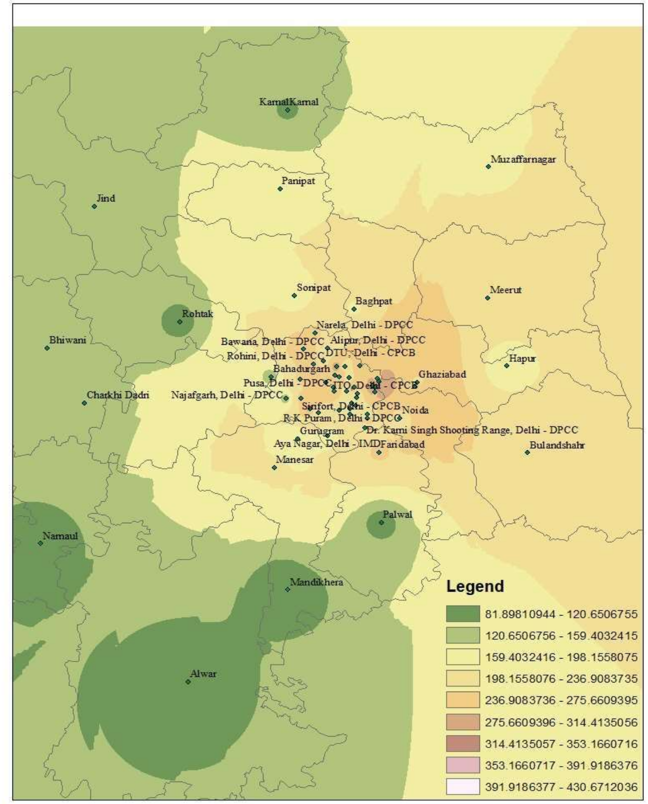


Figure 12. Air mapping of PM_{10} in Delhi NCR during third wave third wave

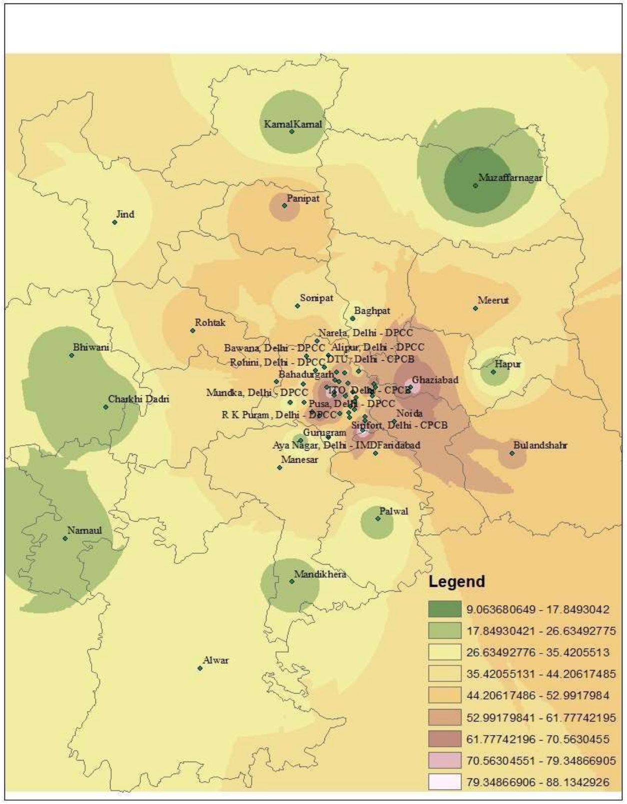


Figure 13. Air Mapping of NO_2 in Delhi NCR during first wave

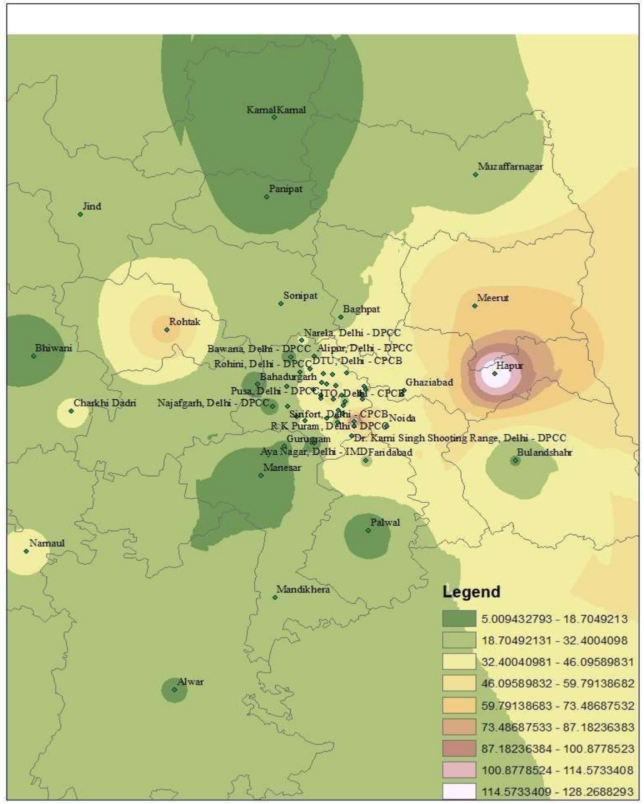


Figure 14. Air mapping of NO_2 in Delhi NCR during second wave

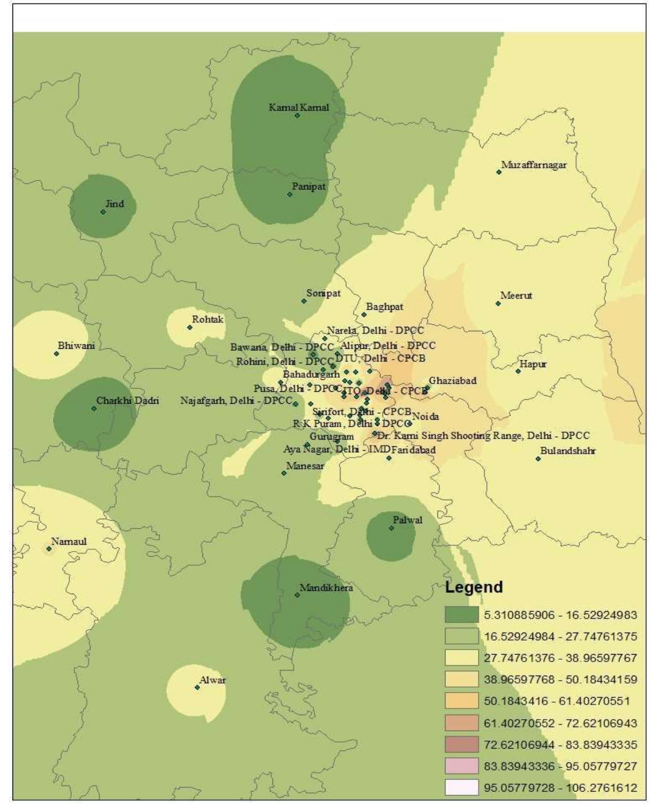


Figure 15. Air mapping of NO_2 in Delhi NCR during third wave third wave

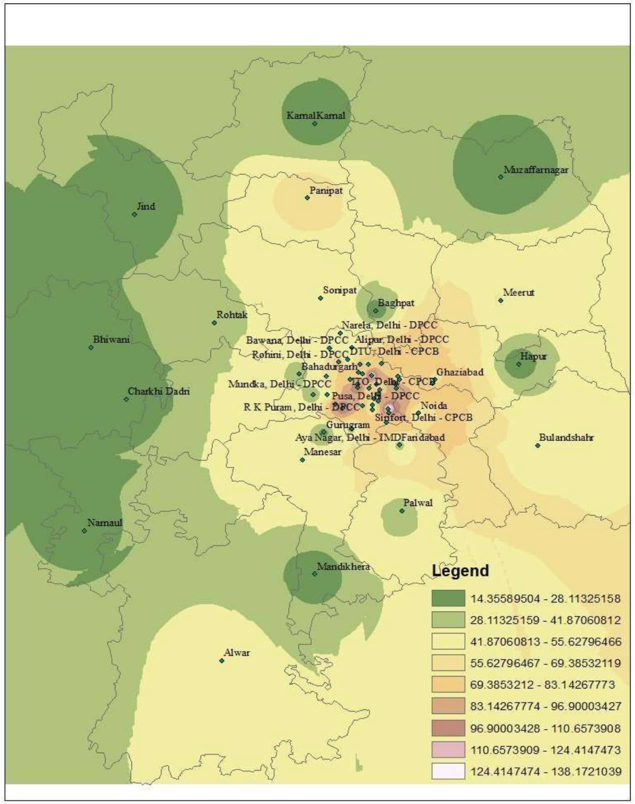


Figure 16. Air Mapping of NOx in Delhi NCR during first wave

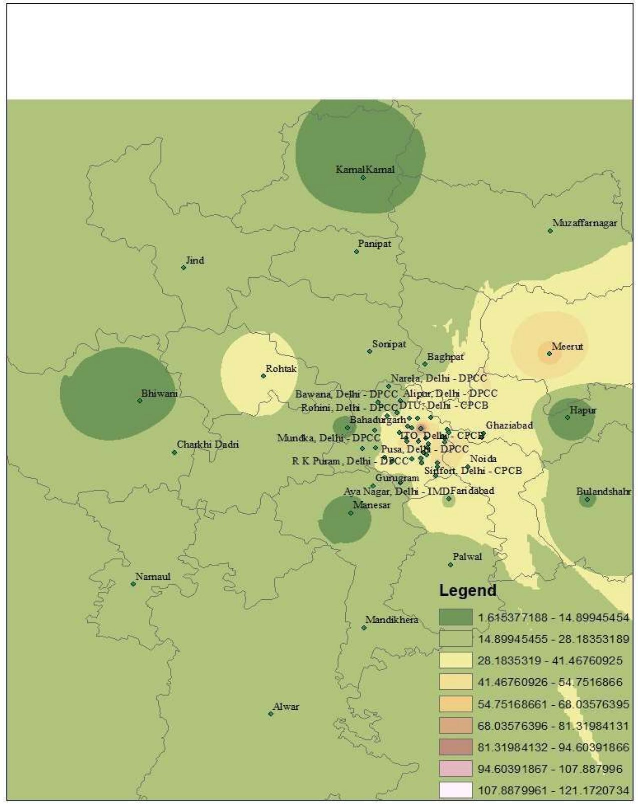


Figure 17. Air mapping of NOx in Delhi NCR during second wave

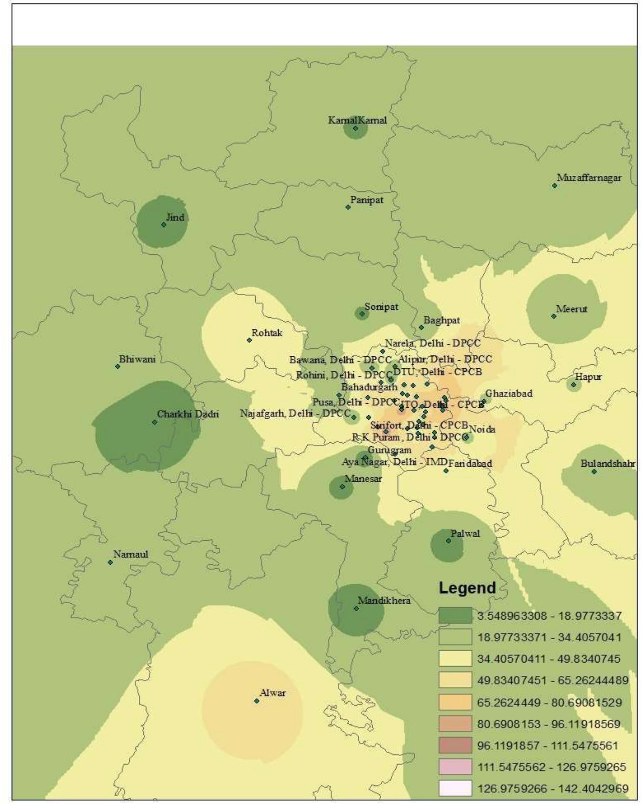


Figure 18. Air mapping of NOx in Delhi NCR during third wave third wave

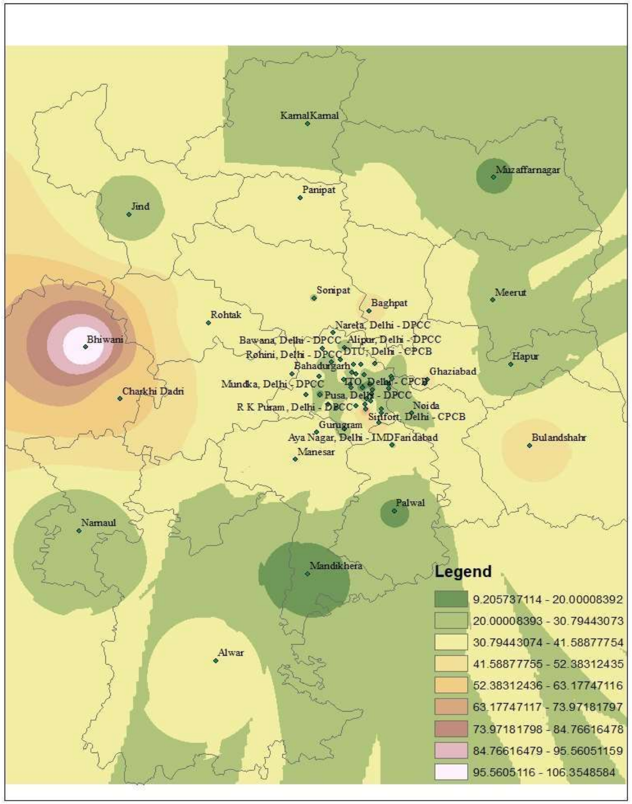


Figure 19. Air Mapping of O_3 in Delhi NCR during first wave

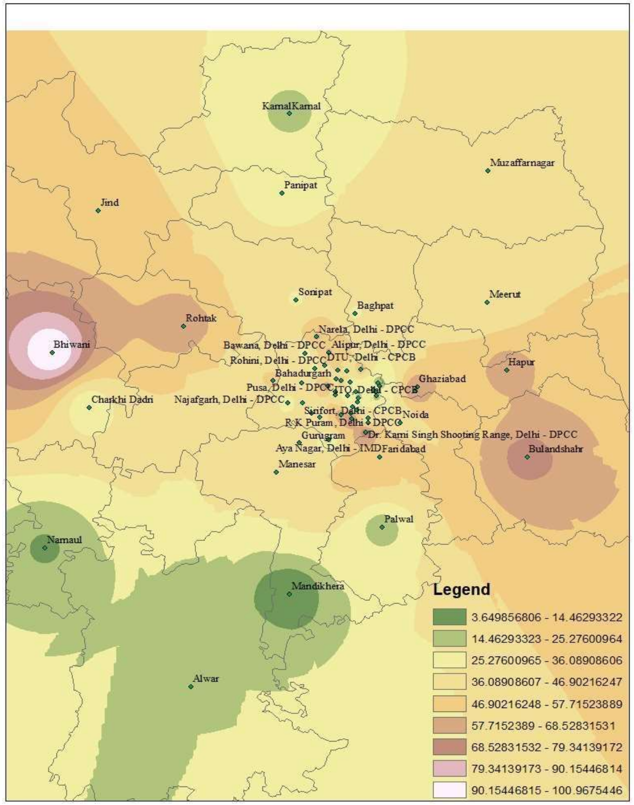


Figure 20. Air mapping of O_3 in Delhi NCR during second wave

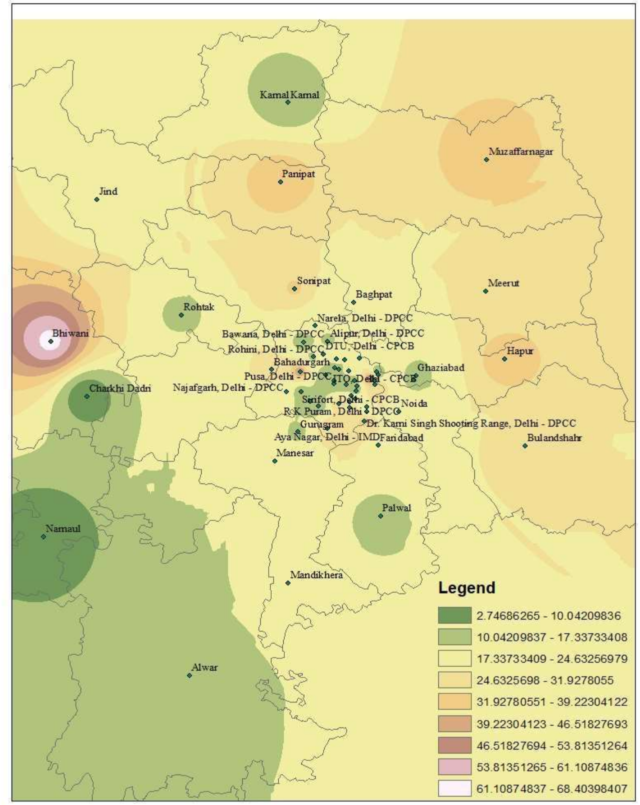


Figure 21. Air mapping of O_3 in Delhi NCR during third wave third wave

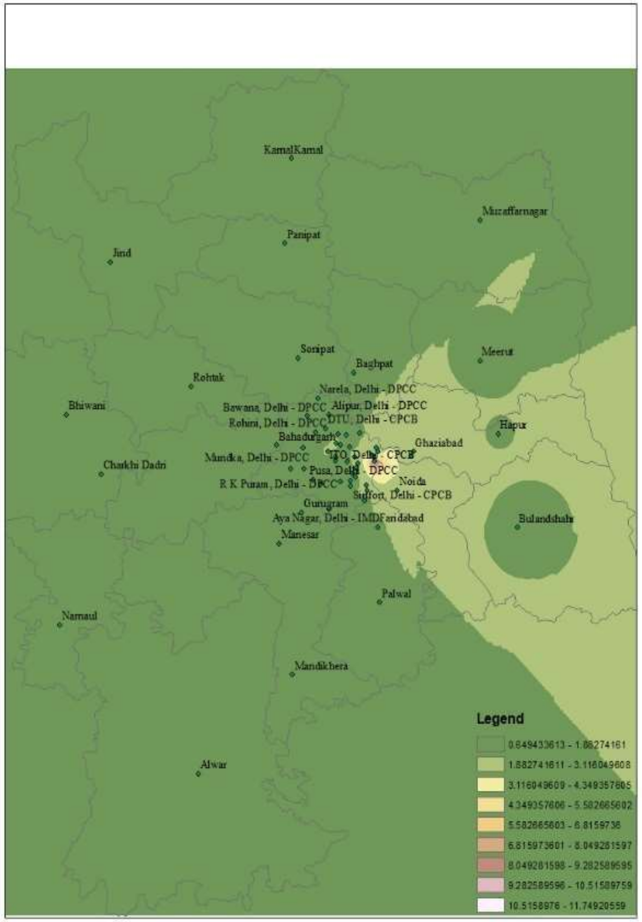


Figure 22. Air Mapping of CO in Delhi NCR during first wave

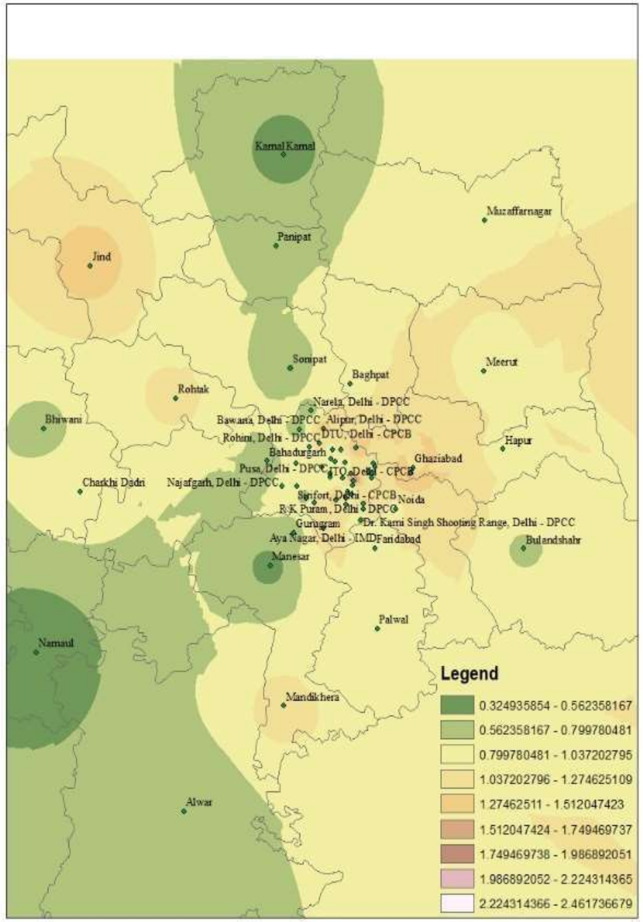


Figure 23. Air mapping of CO in Delhi NCR during second wave

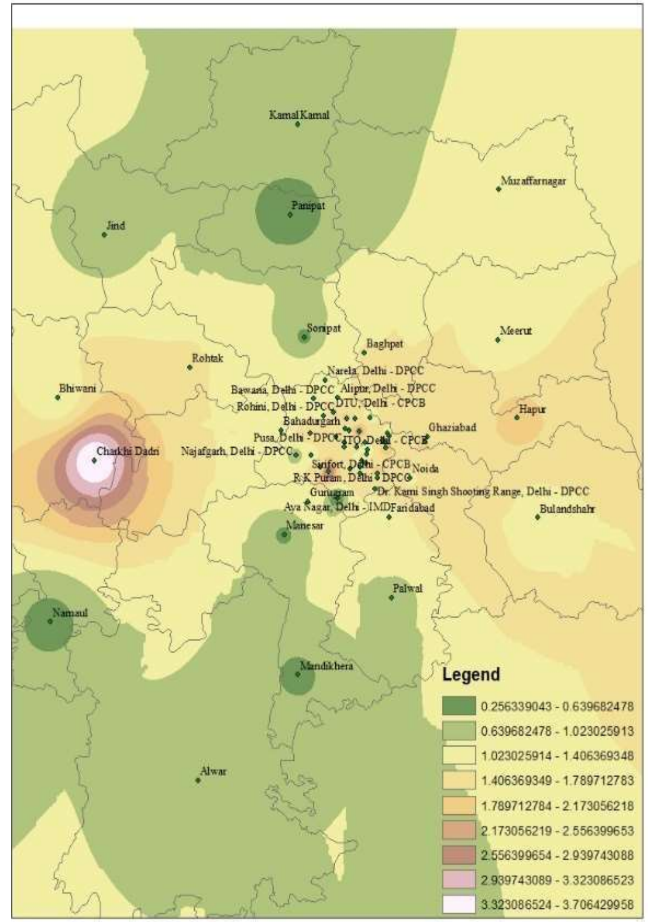


Figure 24. Air mapping of CO in Delhi NCR during third wave third wave

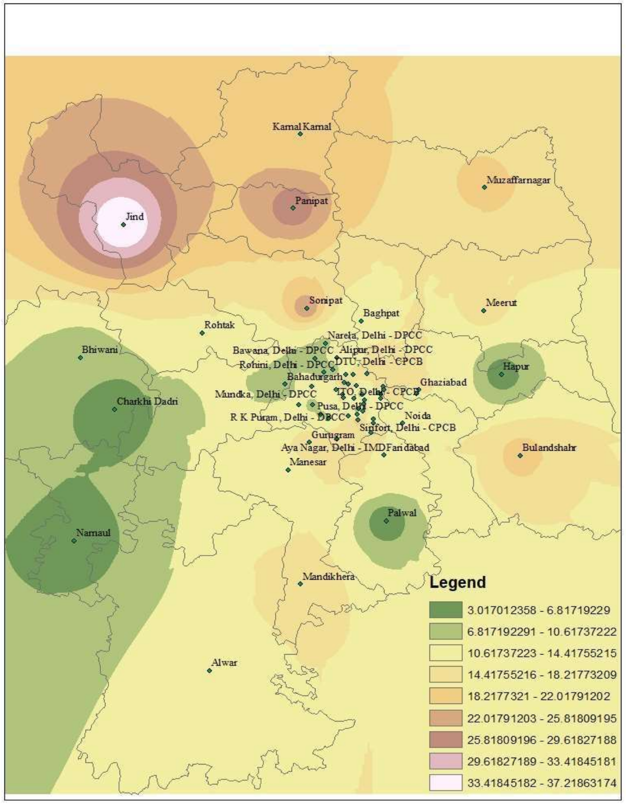


Figure 22. Air Mapping of SO_2 in Delhi NCR during first wave

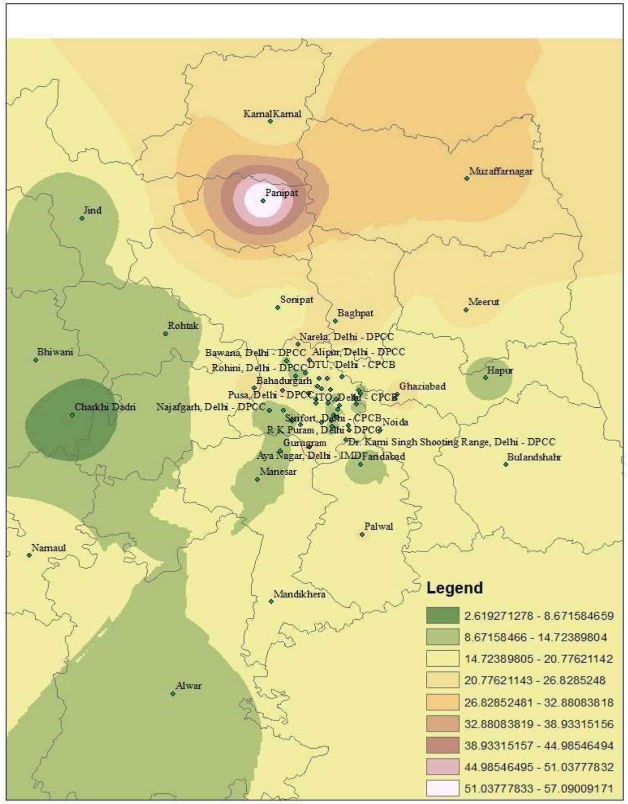


Figure 23. Air mapping of SO_2 in Delhi NCR during second wave

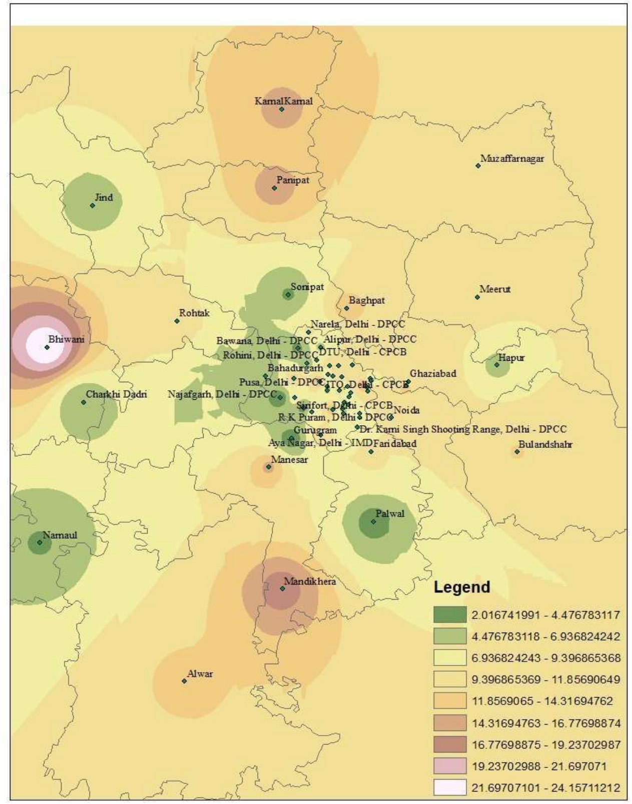


Figure 24. Air mapping of SO_2 in Delhi NCR during third wave third wave

The Figure 10 shows the air quality mapping of PM_{10} in the NCR region during the second COVID-19 wave (April 01 to May 22, 2021). PM_{10} average concentration varies from 73.5 $\mu g/m^3$ to 290 $\mu g/m^3$ for the second wave. The monitoring site which recorded the worst average $PM_{2.5}$ concentration during the second wave was Baghpat ranging from 266.3 $\mu g/m^3$ to 290 $\mu g/m^3$. During the second wave, the monitoring stations with good average air quality were Karnal and Alwar, ranging from 73.5 $\mu g/m^3$ to 97.6 $\mu g/m^3$.

The Figure 11 shows the air quality mapping of PM_{10} in the NCR region during the third COVID-19 wave (January 01 to January 31, 2022). PM_{10} average concentration varies from 81.8 $\mu g/m^3$ to 430.6 $\mu g/m^3$ for the third wave. The monitoring site which recorded the worst average PM_{10} concentration during the third wave was ITO ranging from 391.6 $\mu g/m^3$ to 430.6 $\mu g/m^3$. During the third wave, the monitoring stations with good average air quality were Narnaul, Alwar, Mandikhera, Palwal, and Rohtak, ranging from 81.8 $\mu g/m^3$ to 120.6 $\mu g/m^3$.

4.1.3 Air mapping of NO_2 during each wave

The Figure 14 shows the air quality mapping of NO_2 in the NCR region during the first COVID-19 wave (September 01 to December 11, 2020). NO_2 average concentration varies from 9 $\mu g/m^3$ to 88.1 $\mu g/m^3$ for the first wave. The monitoring sites which recorded the worst average NO_2 concentration during the first wave were ITO and Sirifort, ranging from 79.3 $\mu g/m^3$ to 88.1 $\mu g/m^3$. During the first wave, the monitoring stations with good average air quality were Muzzaffarnagar ranging from 9 $\mu g/m^3$ to 17.8 $\mu g/m^3$.

The Figure 13 shows the air quality mapping of NO_2 in the NCR region during the second COVID-19 wave (April 01 to May 22, 2021). NO_2 average concentration varies from 5 $\mu g/m^3$ to 128.2 $\mu g/m^3$ for the second wave. The monitoring site which recorded the worst average NO_2 concentration during the second wave was Hapur ranging from 114.5 $\mu g/m^3$ to 128.2 $\mu g/m^3$. The monitoring stations with good average air quality during the second wave were Bhiwani, Panipat, Palwal, Bahadurgarh, Najafgarh, Karnal, and Bulandshahr, ranging from 5 $\mu g/m^3$ to 18.7 $\mu g/m^3$.

The Figure 15 shows the air quality mapping of NO_2 in the NCR region during the third COVID-19 wave (January 01 to January 31, 2022). NO_2 average concentration varies from 5.3 $\mu g/m^3$ to 106.2 $\mu g/m^3$ for the third wave. The monitoring sites that recorded the worst average NO2 concentration during the third wave were ITO and Dr. Karni Singh, shooting from 95 $\mu g/m^3$ to 106.2 $\mu g/m^3$. During the third wave, the monitoring stations with good average air quality were Bawana, Charkhi Dadri, Jind, Mandikhera, Panipat, and Karnal, ranging from 5.3 $\mu g/m^3$ to 16.5 $\mu g/m^3$.

4.1.4 Air mapping of NO_X during each wave

The Figure 17 shows the air quality mapping of NO_X in the NCR region during the first COVID-19 wave (September 01 to December 11, 2020). NO_X average concentration varies from 14.3 ppb to 138.1 ppb for the first wave. The monitoring site which recorded the worst average NO_X concentration during the first wave was Sirifort ranging from 124.4 ppb to 138.1

ppb. During the first wave, the monitoring stations with good average air quality were Bhiwani, Muzzaffarnagar, Jind, and Mundka, ranging from 14.3 ppb to 28.1 ppb.

The Figure 18 shows the air quality mapping of NO_X in the NCR region during the second COVID-19 wave (April 01 to May 22, 2021). NO_X average concentration varies from 1.6 ppb to 121.1 ppb for the second wave. The monitoring site which recorded the worst average NO_X concentration during the second wave was ITO ranging from 107.8 ppb to 121.1 ppb. The monitoring stations with good average air quality during the second wave were Bhiwani, Hapur, Panipat, Manesar, Bahadurgarh, Karnal, and Bulandshahr, ranging from 1.6 ppb to 14.8 ppb.

The Figure 16 shows the air quality mapping of NO_X in the NCR region during the third COVID-19 wave (January 01 to January 31, 2022). NO_X average concentration varies from 3.5 ppb to 142.4 ppb for the third wave. The monitoring sites which recorded the worst average NO_X concentration during the third wave were ITO and Sirifort shooting, ranging from 126.9 ppb to 142.4 ppb. The monitoring stations with good average air quality during the third wave were Bawana, Charkhi Dadri, Jind, Mandikhera, Palwal, Aya Nagar, Manesar, and Karnal, ranging from 3.5 ppb to 18.9 ppb.

4.1.5 Air mapping of O_3 during each wave

The Figure 21 shows the air quality mapping of O_3 in the NCR region during the first COVID-19 wave (September 01 to December 11, 2020). O_3 average concentration varies from 9.2 µg/m^3 to 106.3 µg/m^3 for the first wave. The monitoring site which recorded the worst average O_3 concentration during the first wave was Bhiwani ranging from 95.5 µg/m^3 to 106.3 µg/m^3. During the first wave, the monitoring stations with good average air quality were Palwal, Mandikhera, and Muzzaffarnagar, ranging from 9.2 µg/m^3 to 20 µg/m^3.

The Figure 19 shows the air quality mapping of O_3 in the NCR region during the second COVID-19 wave (April 01 to May 22, 2021). O_3 average concentration varies from 3.6 µg/m^3 to 100.9 µg/m^3 for the second wave. The monitoring site which recorded the worst average O_3 concentration during the second wave was Bhiwani ranging from 90.1 µg/m^3 to 100.9 µg/m^3. During the second wave, the monitoring stations with good average air quality were Narnaul and Mandikhera, ranging from 3.6 µg/m^3 to 14.4 µg/m^3.

The Figure 20 shows the air quality mapping of O_3 in the NCR region during the third COVID-19 wave (January 01 to January 31, 2022). O_3 average concentration varies from 2.7 µg/m^3 to 68.4 µg/m^3 for the third wave. The monitoring site which recorded the worst average O_3 concentration during the third wave was Bhiwani shooting, ranging from 61.1 µg/m^3 to 68.4 µg/m^3. During the third wave, the monitoring stations with good average air quality were Charkhi Dadri and Narnaul, ranging from 2.7 µg/m^3 to 10 µg/m^3.

4.1.6 Air mapping of CO during each wave

The Figure 24 shows air quality mapping of CO in the NCR region during the first COVID-19 wave (September 01 to December 11, 2020). CO average concentration varies from 0.64 mg/m^3 to 1.8 mg/m^3 for the first wave. The monitoring site which recorded the worst average CO concentration during the first wave was ITO ranging from 8 mg/m^3 to 9.2 mg/m^3. During the first wave, the monitoring stations with good average air quality were all the area apart from ITO, ranging from 0.64 mg/m^3 to 1.8 mg/m^3.

The Figure 23 shows the air quality mapping of CO in the NCR region during the second COVID-19 wave (April 01 to May 22, 2021). For the second wave, the CO average concentration varies from 0.3 mg/m3 to 2.4 mg/m3. The monitoring site which recorded the worst average CO concentration during the second wave was ITO ranging from 2.2 mg/m^3 to 2.4 µg/m^3. During the second wave, the monitoring stations with good average air quality were Manesar, Narnaul, and Karnal, ranging from 0.3 mg/m^3 to 0.5 mg/m^3.

The Figure 22 shows the air quality mapping of CO in the NCR region during the third COVID-19 wave (January 01 to January 31, 2022). CO average concentration varies from 0.2 mg/m^3 to 3.7 mg/m^3 for the third wave. The monitoring site which recorded the worst average CO concentration during the third wave was Charkhi Dadri shooting, ranging from 3.3 mg/m^3 to 3.7 mg/m^3. During the third wave, the monitoring stations with good average air quality were Panipat, Sonipat, Narnaul, Gurugram, Manesar, and Mandikhera, ranging from 0.2 mg/m^3 to 0.6 mg/m^3.

4.1.7 Air mapping of SO_2 during each wave

The Figure 27 shows the air quality mapping of SO_2 in the NCR region during the first COVID-19 wave (September 01 to December 11, 2020). SO_2 average concentration varies from 3 µg/m^3 to 37.2 µg/m^3 for the first wave. The monitoring site which recorded the worst average SO_2 concentration during the first wave was Jind ranging from 124.4 µg/m^3 to 138.1 µg/m^3. During the first wave, the monitoring stations with good average air quality were Charkhi Dadri, Hapur, Narnaul, and Palwal, ranging from 3 µg/m^3 to 6.8 µg/m^3.

The Figure 26 shows the air quality mapping of SO_2 in the NCR region during the second COVID-19 wave (April 01 to May 22, 2021). SO_2 average concentration varies from 2.6 µg/m^3 to 57 µg/m^3 for the second wave. The monitoring site which recorded the worst average SO_2 concentration during the second wave was Panipat ranging from 51 µg/m^3 to 57 µg/m^3. During the second wave, the monitoring station with good average air quality was Charkhi Dadri ranging from 1.6 µg/m^3 to 14.8 µg/m^3.

The Figure 25 shows the air quality mapping of SO_2 in the NCR region during the third COVID-19 wave (January 01 to January 31, 2022). SO_2 average concentration varies from 2 µg/m^3 to 24.1 µg/m^3 for the third wave. The monitoring site which recorded the worst average SO_2 concentration during the third wave was Bhiwani ranging from 21.6 µg/m^3 to 24.1 µg/m^3. During the third wave, the monitoring stations with good average air quality were Bawana, Sonipat, Najafgarh, Palwal, Aya Nagar, Manesar, and Narnaul, ranging from 2 µg/m^3 to 4.4 µg/m^3.

4.2 Pearson's Correlation Coefficients

This method helps in determining the collinearity among predictors. (*Table 2*, *Table 3*, and *Table 4*) are showing the correlation analysis between new daily corona cases and 24-mean air quality. During each study period over Delhi, NCR results in a moderately significant positive correlation (r) for the first and second wave, whereas for the third wave, a weak association has been seen in most cases at $p < 0.001$.

Table 2 Pearson Correlation Matrix of the first wave

	CASES	*$PM_{2.5}$*	*PM_{10}*	*NO_2*	*NOx*	*OZONE*	*CO*	*SO_2*
CASES	1.000							
$PM_{2.5}$	0.398	1.000						
PM_{10}	0.338	0.973	1.000					
NO_2	0.463	0.914	0.911	1.000				
NOx	0.489	0.847	0.861	0.951	1.000			
OZONE	0.091	0.414	0.500	0.374	0.306	1.000		
CO	0.453	0.919	0.888	0.925	0.904	0.332	1.000	
SO_2	0.383	0.662	0.728	0.739	0.677	0.609	0.597	1.000

Table 3 Pearson Correlation Matrix of the second wave

	CASES	*$PM_{2.5}$*	*PM_{10}*	*NO_2*	*NOx*	*OZONE*	*CO*	*SO_2*
CASES	1.000							
$PM_{2.5}$	0.272	1.000						
PM_{10}	0.233	0.939	1.000					
NO_2	0.052	0.673	0.683	1.000				
NOx	0.028	0.684	0.689	0.951	1.000			
OZONE	0.405	0.081	0.029	-0.349	-0.310	1.000		
CO	0.128	0.705	0.527	0.512	0.540	0.062	1.000	
SO_2	0.008	0.698	0.759	0.811	0.803	-0.197	0.455	1.000

Table 4 Pearson Correlation Matrix of the third wave

	CASES	*$PM_{2.5}$*	*PM_{10}*	*NO_2*	*NOx*	*OZONE*	*CO*	*SO_2*
CASES	1							
$PM_{2.5}$	0.345	1.000						
PM_{10}	-0.387	0.978	1.000					
NO_2	-0.672	0.756	0.755	1.000				
NOx	-0.622	0.748	0.727	0.960	1.000			
OZONE	-0.585	0.017	0.068	0.455	0.259	1.000		
CO	0.479	0.897	0.886	0.877	0.908	0.104	1.000	
SO_2	0.150	0.387	0.365	0.560	0.513	0.310	0.459	1.000

The results that appeared after running Pearson correlation analysis in the collected dataset are also represented in the graphical format (Figure 28, Figure 29, and Figure 30) for each wave and individually discussed in below sections:

1. $PM_{2.5}$ and COVID-19 Cases
   In Delhi, NCR during the first, second, and third wave $PM_{2.5}$ has shown moderate positive, weak positive, and moderate positive association with daily new COVID-19 cases, respectively.
2. PM_{10} and COVID-19 Cases
   In Delhi, NCR, during the first, second, and third waves, PM_{10} has shown moderate positive, weak positive, and moderate negative association with daily new COVID-19 cases.

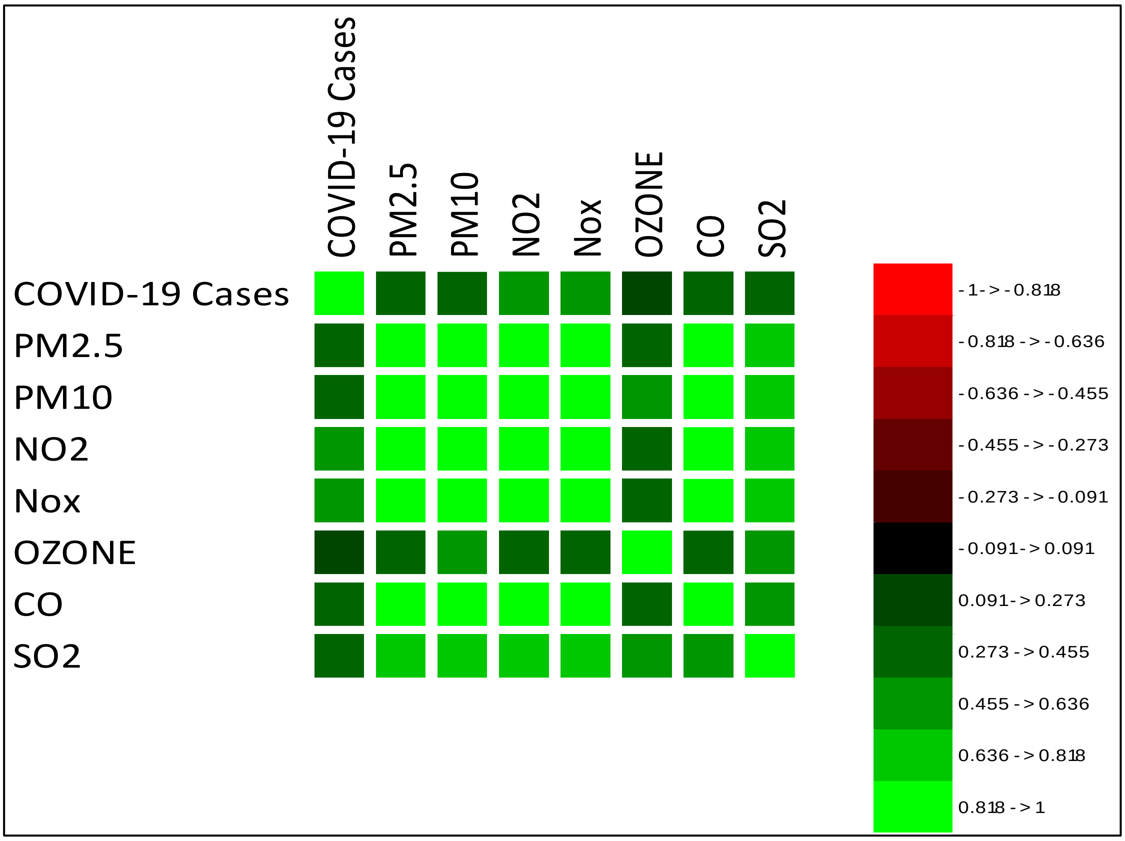


Figure 28. Pearson correlation among COVID-19 and air pollutants during the first wave

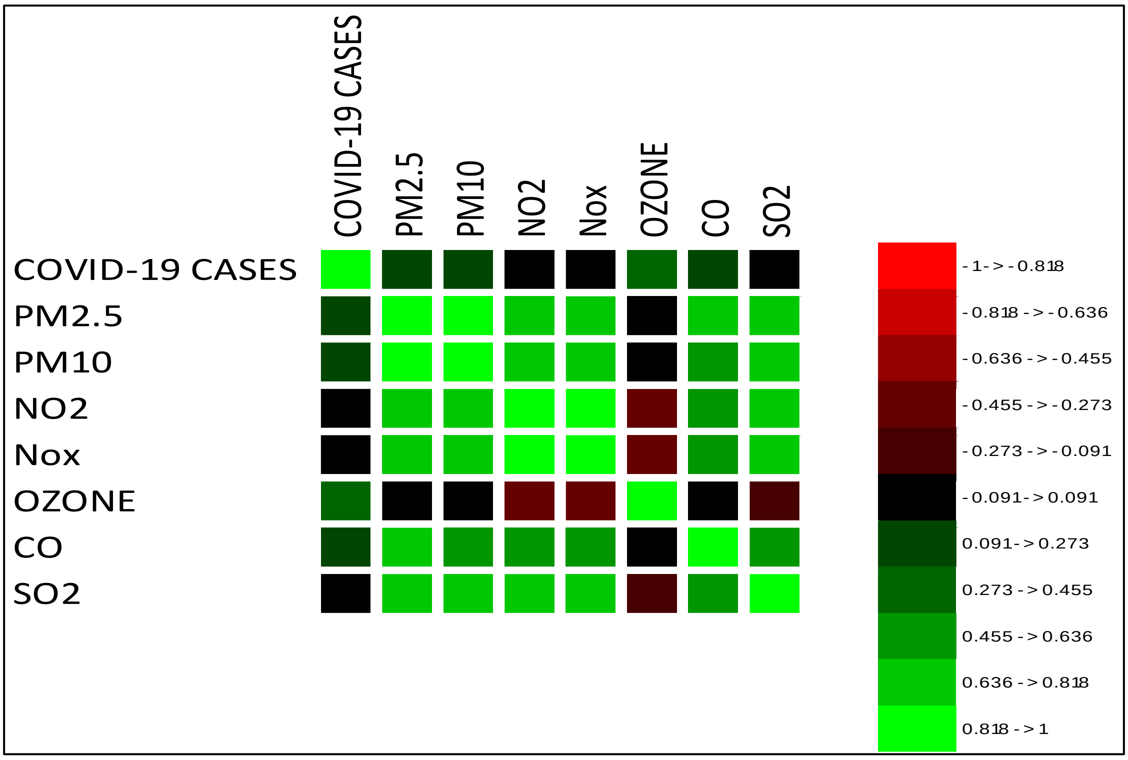


Figure 29. Pearson correlation among COVID-19 and air pollutants during the second wave

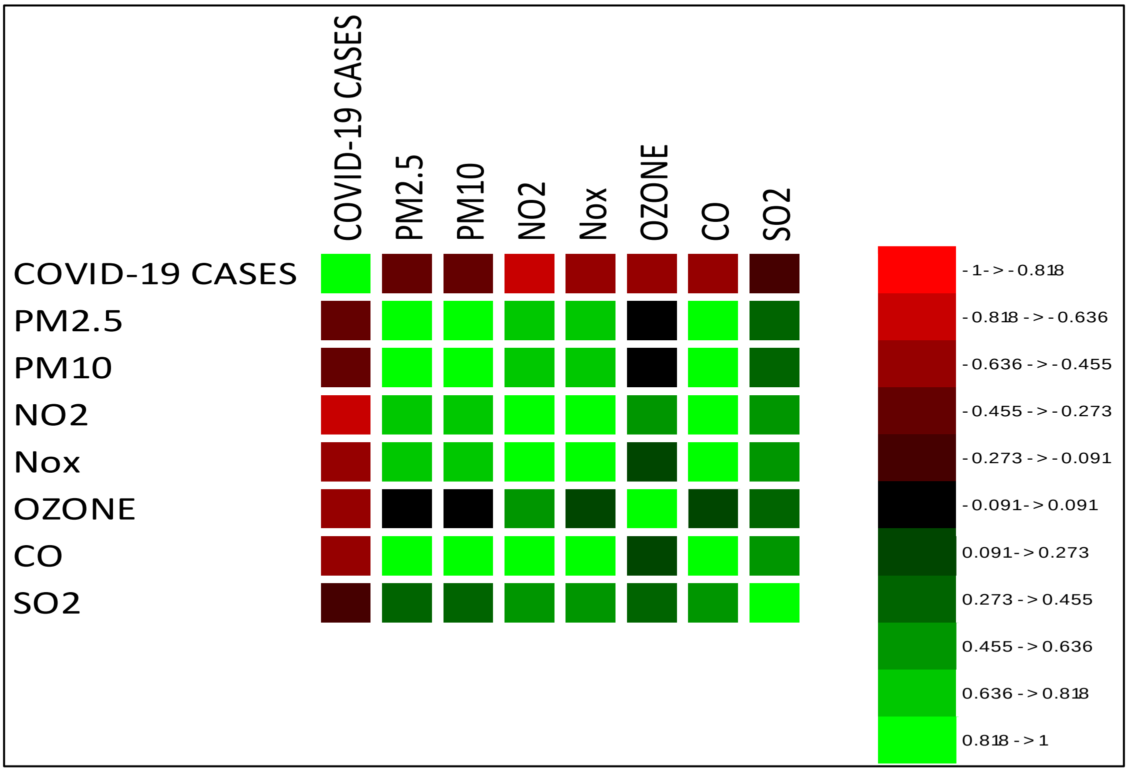


Figure 30. Pearson correlation among COVID-19 and air pollutants during the third wave

3. NO_2 and COVID-19 Cases

   In Delhi, NCR during the first, second, and third wave NO_2 has shown moderate positive, weak positive, and strong negative association with daily new COVID-19 cases, respectively.
4. NOx and COVID-19 cases
   In Delhi, NCR during the first, second, and third wave NO_X has shown moderate positive, weak positive, and strong negative association with daily new COVID-19 cases, respectively.
5. Surface ozone (O_3) and COVID-19 cases
   In Delhi, NCR during the first, second, and third wave O_3 has shown weak positive, moderately positive, and strong negative association with daily new COVID-19 cases, respectively.
6. Carbon Monoxide (CO) and COVID-19 cases
   In Delhi NCR, during the first, second, and third waves, CO has shown moderate positive, weak positive, and moderate positive association with daily new COVID-19 cases, respectively.
7. SO_2 and COVID-19 cases
   In Delhi NCR, during the first, second, and third waves, CO has shown moderate positive, weak positive, and weakly positive association with daily new COVID-19 cases, respectively.
   From the above results, it is observed that ($PM_{2.5}$, CO, and SO_2) were among the pollutants which were showing a positive association with daily new coronavirus cases ($p < 0.001$). Additionally, (PM_{10}, NO_2, NO_X, and O_3) were among the pollutants, which has shown a positive association with the spread of the virus during two out of three waves at bivariate level analysis ($p < 0.001$). From the above tables, it can be seen that the air pollutants (predictors) are showing a high correlation among them ($r > 0.8$), which indicates the presence of collinearity among variables. In the next step, the dataset will be analysed at a multivariate level using a multiple linear regression model to draw out some strong evidence for the existence of collinearity among variables.

4.3 Variance Inflation Factor (VIF)

The stepwise procedure of multiple linear regression was performed to find collinearity among predictors. Results appeared after analysing the dataset using the "least square (NLS and ARMA)" regression model for the dataset of each wave in Delhi NCR are as follows:

1. In Delhi NCR, during first wave the result shows that 40.53% of the variance in cases can be accounted for by the air pollutants, collectively, (*F-statistic*) = 9.153, $p < 0.000001$. Looking at the unique individual contributions of the predictors, the results show that $PM_{2.5}$ (β= 50.91, t= 3.627, $p < 0.0005$), NO_X (β= 93.7, t= 2,99, $p < 0.0035$), and SO_2 (β= 178.7, t= 2.81, $p < 0.006$) were among the most significant ($p < 0.05$) air pollutants that influenced the incidence of SARS-CoV-2 in Delhi NCR.

2. In Delhi NCR, during the first wave, the result shows that 28.1% of the variance in cases can be accounted for by the air pollutants, collectively (*F-statistic)* = 2.46, $p < 0.03$. Looking at the unique individual contributions of the predictors, the results show that O_3 (β= 998.71, t= 2.71, $p < 0.009$*)* was among the most significant ($p < 0.05$) air pollutants that influenced the incidence of SARS-CoV-2 in Delhi NCR.
3. In Delhi NCR, during the first wave, the result shows that the air pollutants can account for 68.96% of the variance in cases, collectively, (*F-statistic)* = 6.98, $p < 0.0001$. Looking at the unique individual contributions of the predictors, the results show that SO_2 (β= 259.99, t= 2.09, $p < 0.04$*)* was among the most significant ($p < 0.05$) air pollutants that influenced the incidence of SARS-CoV-2 in Delhi NCR.

Table 5 shows that 36.1 %, 16.7 %, and 59.1 % variance in COVID-19 cases can be accounted for by the pollutants in NCR during the first, second, and third wave, respectively, at $p < 0.001$. Furthermore, $PM_{2.5}$, PM_{10}, NO_2, NOx, and CO were the predictors whose VIF value is more significant than ten and tolerance level is less than 0.10 during each wave, which specifies that the pollutants are highly correlated with each other. The high values of VIF corresponding to the variables show that there is a problem with collinearity. From all these findings, it can be said that the association can't be determined using standard regression models because they will not be able to predict the correct coefficients. In the next section, to deal with this issue, elastic net regularization has been applied to find the best output for our collected dataset and to predict better results.

Table 5 Multicollinearity statistics (p <0.001)

	First Wave (R^2 = 0.361)		Second Wave (R^2 = 0.167)		Third Wave (R^2 = 0.591)	
Variables	Tolerance	VIF	Tolerance	VIF	Tolerance	VIF
$PM_{2.5}$	0.028	35.19	0.054	18.51	0.034	29.51
PM_{10}	0.031	31.75	0.066	15.25	0.032	31.07
NO_2	0.043	23.1	0.084	11.93	0.015	64.74
NOx	0.7	14.31	0.089	11.22	0.02	48.96
O_3	0.484	2.06	0.693	1.44	0.202	4.96
CO	0.083	12.04	0.323	3.09	0.059	16.81
SO_2	0.27	3.71	0.249	4.01	0.647	1.54

4.4 Association among air pollutants and SARS-CoV-2 daily cases using elastic net

In the previous section, a high correlation was found among most of the predictors (collinearity). To remove this discrepancy and predict a better model, elastic net regularization has been used in model exploration and coefficient estimation. The results that appeared after analysing the dataset are mentioned below lines:

4.4.1 During First Wave

In Delhi NCR, the result shows that air pollutants can account for 38.5% of the variance in cases, collectively shown in the Figure 31 of model output. As shown in Figure 32 Lambda (λ) at minimum error, RMSE (Root mean square error) and MAE (Mean absolute error) came out as 13.7, 1549.7, and 1279.7, respectively, after testing and training the model with K-Fold cross-validation with (K=10 folds) (Figure 33). In the Figure 34, this model's actual, fitted, and residual lines are portrayed in a graph that indicates the justifications for model output. Looking at the unique individual contributions of the predictors (pollutants), the results show that the air pollutants which influenced the occurrence of SARS-CoV-2 daily cases in Delhi NCR are portrayed in Equation (4)

$$New\ Cases_{First\ Wave} = 4240.13 + 40.28 * PM_{2.5} + 68.58 * NO_X + 16.72 * O_3 + 82.52 * CO + 85.64 * SO_2 - 3.9 * NO_2 - 32.6 PM_{10} \quad (5)$$

Furthermore, the Figure 35 displays the equation derivatives of new covid cases during the first wave. It is observed from the output of the first wave dataset that for every 1 μg/m^3 increase in $PM_{2.5}$, O_3, CO, and SO_2, COVID-19 cases will increase by 40.28, 16.7, 82.5, and 85.6, respectively, and when there will be 1 μg/m^3 decrease in PM_{10}, and NO_2, COVID-19 cases will decrease by 32.6, and 3.9 provided that at all the other parameter remain unchanged at a particular time. Additionally, with every one ppb increase in NO_X COVID-19, cases will increase by 68. Hence, during the first wave, mainly every pollutant has shown a significant association with the spread of the virus in the NCR region.

Equation: UNTITLED Workfile: DELHI NCR FIRST WAVE...

View | Proc | Object | Print | Name | Freeze | Estimate | Forecast | Stats | Resids

Dependent Variable: CASES
Method: Elastic Net Regularization
Date: 05/28/22 Time: 15:19
Sample: 9/01/2020 12/11/2020
Included observations: 102
Penalty type: Elastic Net (alpha = 0.5)
Lambda at minimum error: 13.7
Regressor transformation: None
Cross-validation method: K-Fold (number of folds = 10), rng=kn, seed=1574489245
Selection measure: Mean Squared Error

	(minimum)	(+ 1 SE)	(+ 2 SE)
Lambda	13.7	355.5	2285
Variable	**Coefficients**		
C	4240.130	4799.343	4722.495
PM2_5	40.28699	17.15347	6.602102
PM10	-32.66565	-9.813955	0.812498
NO2	-3.990125	13.27180	3.323408
NOX	68.58684	22.01408	5.117417
OZONE	16.72402	0.000000	0.000000
CO	82.52030	0.000000	0.000000
SO2	85.64206	5.900027	0.228754
d.f.	7	5	5
L1 Norm	4570.546	4867.497	4738.579
R-squared	0.387350	0.271950	0.168983

Figure 31. Model output summary of the first wave

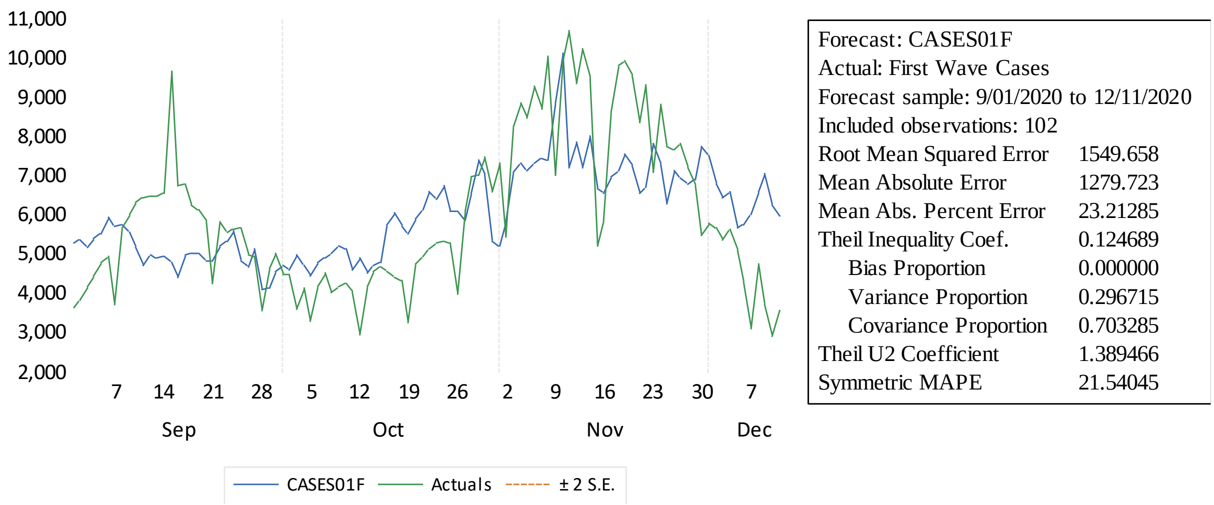


Figure 32. Results after analysing the dataset for the first wave

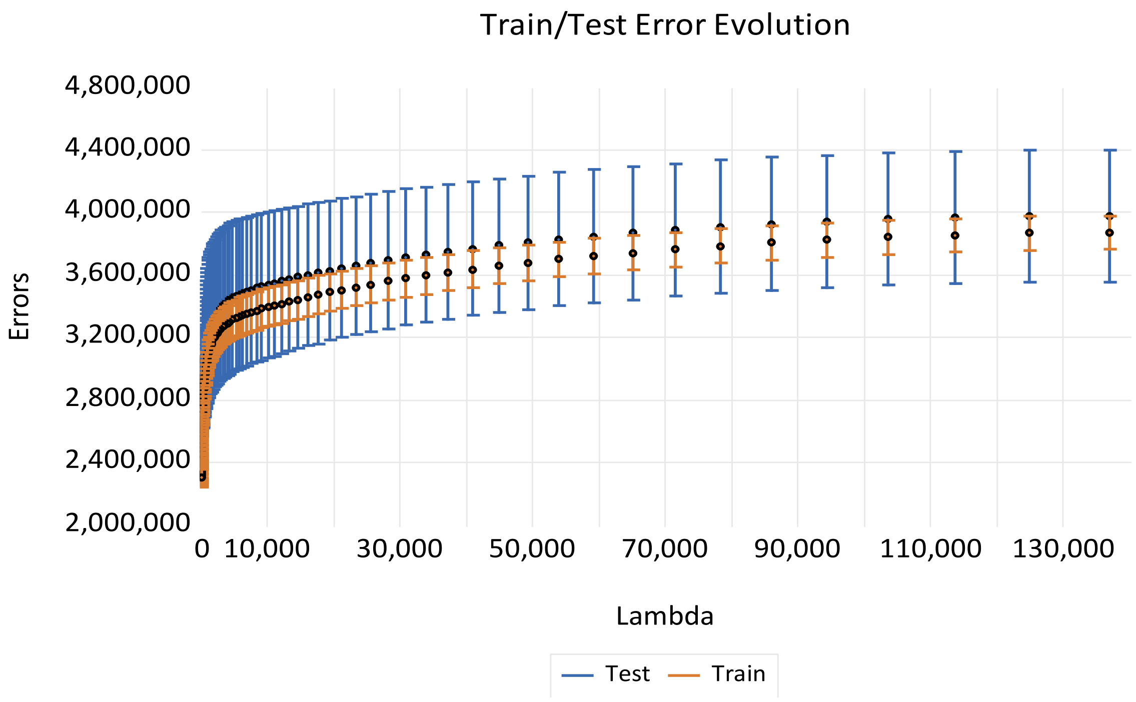


Figure 33. Training and testing of the dataset with K-Fold cross-validation for the first wave

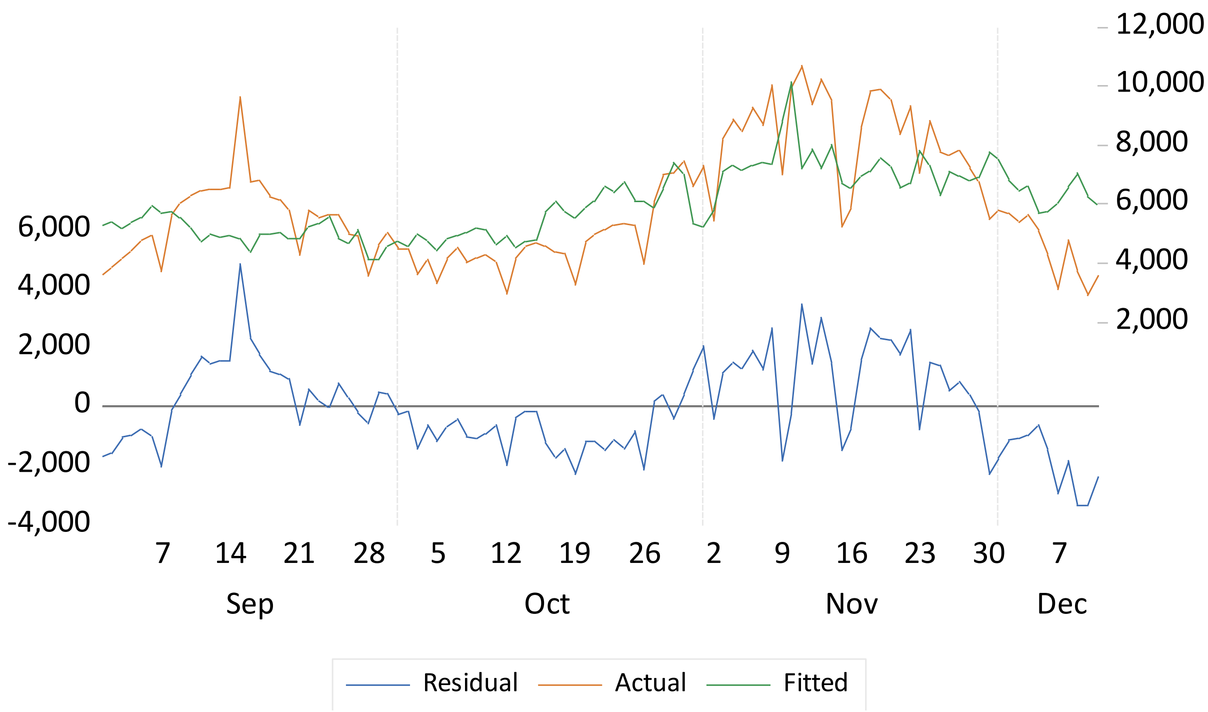


Figure 34. Residual, Actual and Fitted lines of regression model for first wave dataset

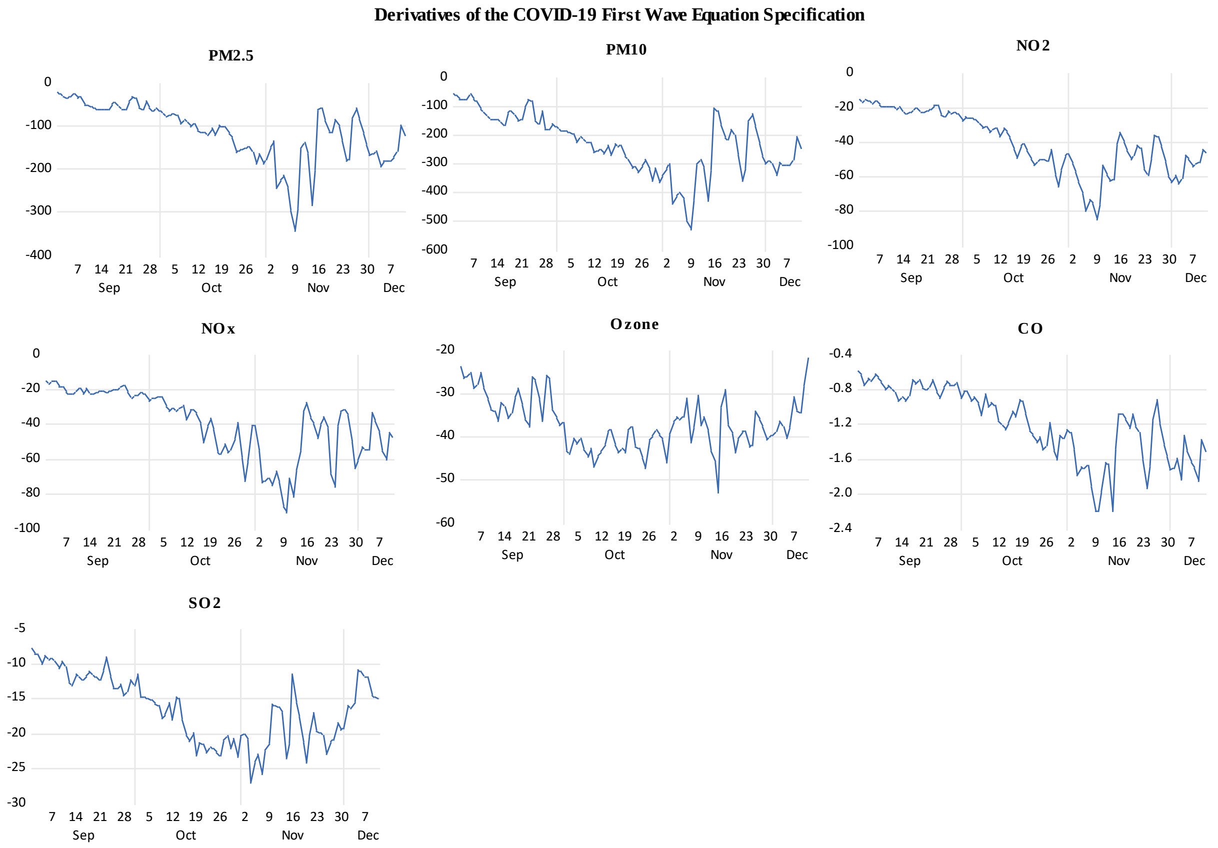


Figure 35. Equation derivatives for the first wave

4.4.2 During the Second wave

In Delhi NCR, the result shows that air pollutants can account for 19.28% of the variance in cases, collectively shown in the Figure 36 of model output. As shown in Figure 37 Lambda (λ) at minimum error, RMSE (Root mean square error) and MAE (Mean absolute error) came out as 61.4, 10851.9, and 9406.1, respectively, after testing and training the model with K-Fold cross-validation with (K=5 folds) (Figure 38). In the Figure 39, this model's actual, fitted, and residual lines are portrayed in a graph. Looking at the unique individual contributions of the predictors (pollutants), the results show that the air pollutants which influenced the occurrence of SARS-CoV-2 daily cases in Delhi NCR are portrayed in Equation (5)

$$New\ Case_{Second\ Wave} = 359.4 + 136.5 * PM_{2.5} + 382.3 * O_3 + 13.8 * PM_{10} - 2.2 * CO - 20.2 * NO_2 - 94.8 * NO_X - 168.4 * SO_2 \quad (5)$$

Furthermore, the Figure 40 displays the equation derivatives of new covid cases during the first wave. It is observed from the output of the second wave dataset that for every 1 $\mu g/m^3$ increase in $PM_{2.5}$, PM_{10}, and O_3, COVID-19 cases will increase by 136.5, 13.8, and 382.3, respectively. When there is a 1 $\mu g/m^3$ decrease in SO_2, CO, and NO_2, COVID-19 cases will decrease by 168.4, 2.2, and 20.2, provided that all the other parameters remain unchanged at a particular time. Additionally, with every one ppb increase in NO_X COVID-19, cases will increase by 94.8. Hence, during the second wave, mainly every pollutant has shown a significant association with the spread of the virus in the NCR region.

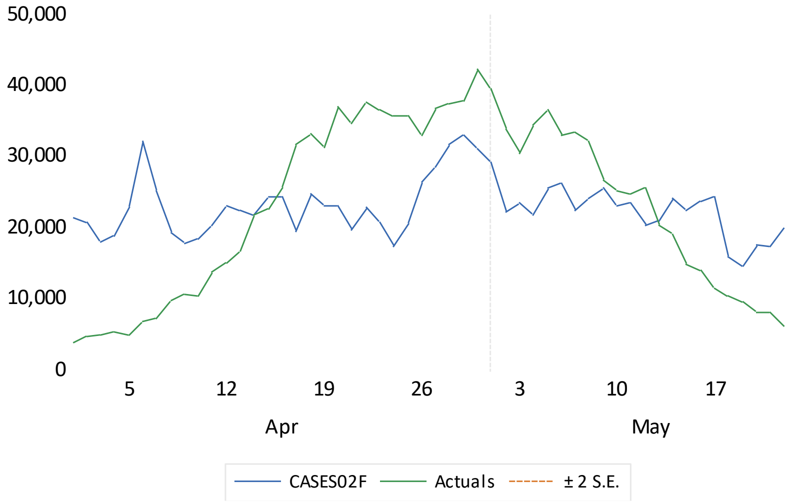
Equation: UNTITLED Workfile: DELHI NCR SECOND WAVE::Un...

View | Proc | Object | Print | Name | Freeze | Estimate | Forecast | Stats | Resids

Dependent Variable: CASES
Method: Elastic Net Regularization
Date: 05/28/22 Time: 14:36
Sample: 4/01/2021 5/22/2021
Included observations: 52
Penalty type: Elastic Net (alpha = 0.5)
Lambda at minimum error: 61.38
Regressor transformation: None
Cross-validation method: K-Fold (number of folds = 5), rng=kn, seed=1527321226
Selection measure: Mean Squared Error

	(minimum)	(+ 1 SE)	(+ 2 SE)
Lambda	61.38	3.2e+05	3.2e+05
Variable		Coefficients	
C	359.3781	22742.02	22742.02
PM2_5	136.4681	0.000000	0.000000
PM10	13.84291	0.198775	0.198775
NO2	-20.18806	0.000000	0.000000
NOX	-94.83591	0.000000	0.000000
OZONE	382.3332	0.000000	0.000000
CO	-2.162392	0.000000	0.000000
SO2	-168.4494	0.000000	0.000000
d.f.	7	1	1
L1 Norm	1177.658	22742.21	22742.21
R-squared	0.192728	0.000517	0.000517

Figure 36. Model output summary of the second wave



Forecast: CASES02F	
Actual: CASES SECOND WAVE	
Forecast sample: 4/01/2021 to 5/22/2021	
Included observations: 52	
Root Mean Squared Error	10927.38
Mean Absolute Error	9533.441
Mean Abs. Percent Error	80.97360
Theil Inequality Coef.	0.223240
Bias Proportion	0.000000
Variance Proportion	0.557778
Covariance Proportion	0.442222
Theil U2 Coefficient	9.216985
Symmetric MAPE	48.68036

Figure 37. Results after analysing the dataset for the second wave

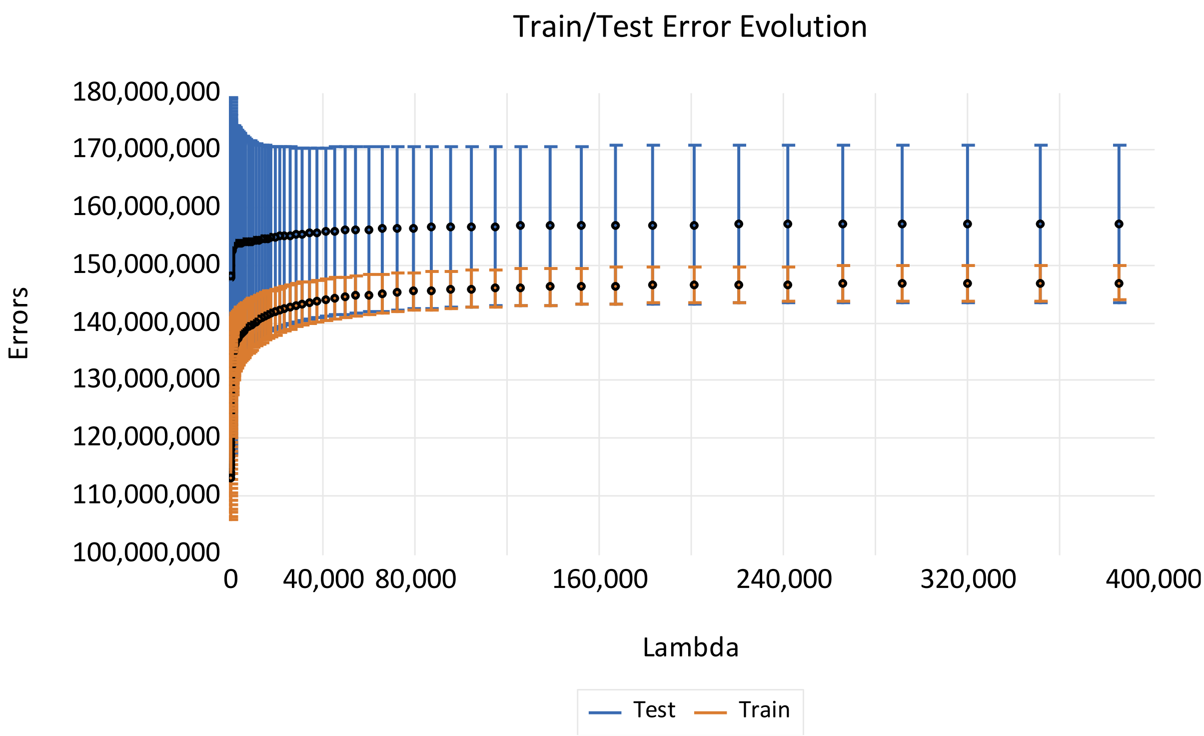


Figure 38. Training and testing of the dataset with K-Fold cross-validation for the second wave

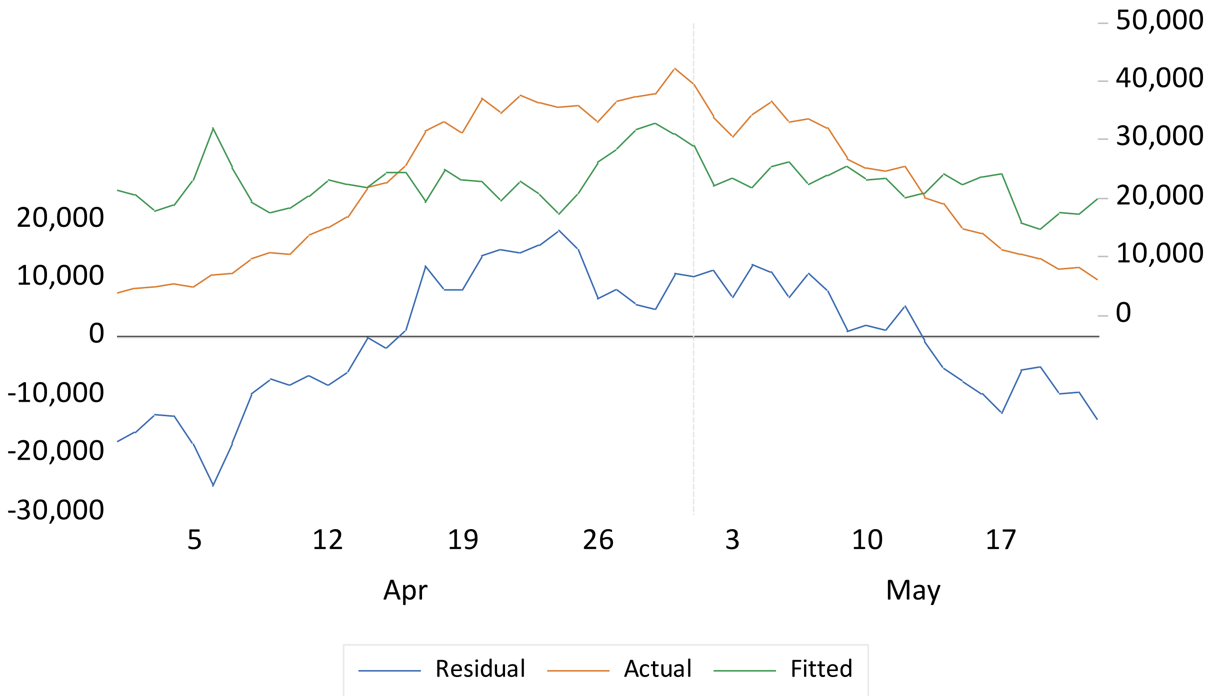


Figure 39. Residual, Actual and Fitted lines of regression model for second wave dataset

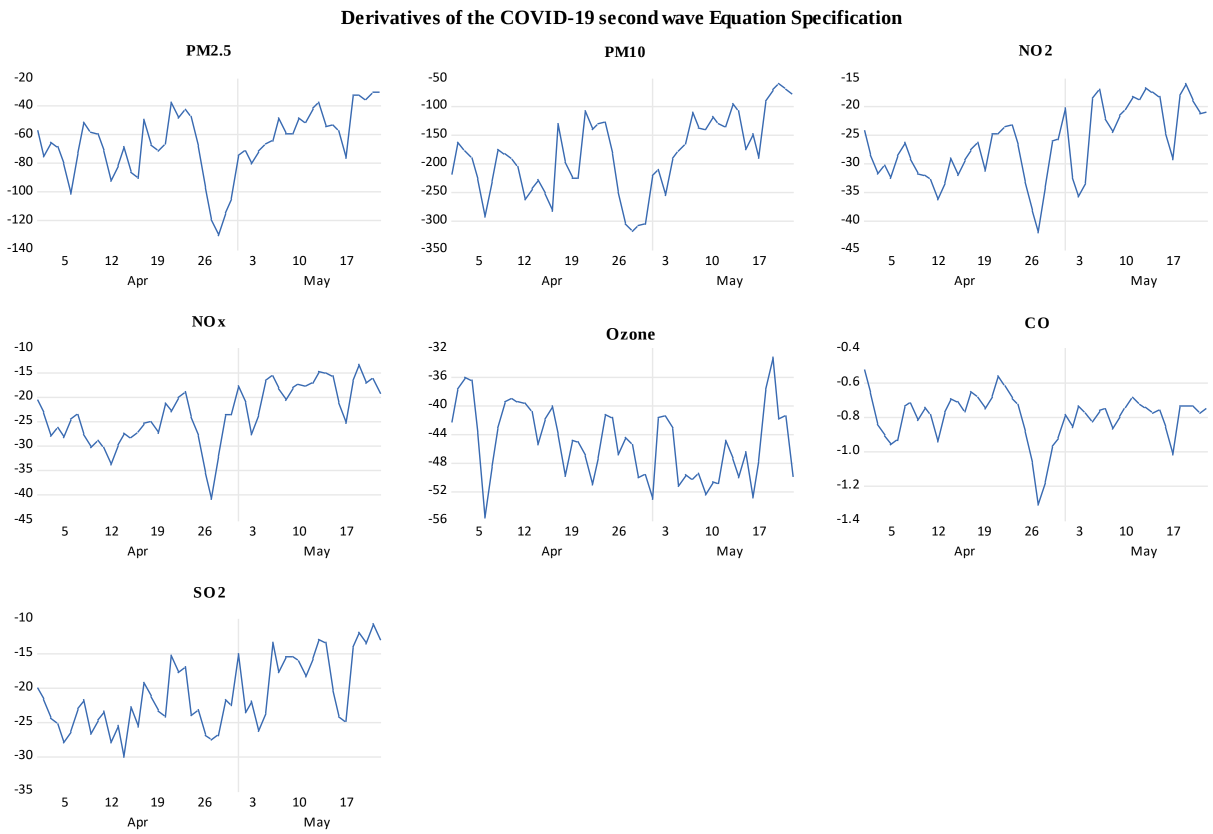


Figure 40. Equation derivatives for the second wave

4.4.3 During Third Wave

In Delhi NCR, the result shows that the air pollutants collectively account for 55.32% of the variance in cases as shown in Figure 41. As shown in Figure 42 Lambda (λ) at minimum error, RMSE (Root mean square error), and MAE (Mean absolute error) came out as 63.1, 6859.2, and 38.34, respectively, after testing and training the model with K-Fold cross-validation with (K=10 folds) (Figure 43). In the Figure 44, this model's actual, fitted, and residual lines are portrayed in a graph. Looking at the unique individual contributions of the predictors (pollutants), the results show that the air pollutants which influenced the occurrence of SARS-CoV-2 daily cases in Delhi NCR are portrayed in Equation (6)

$$New\ Case_{Third\ Wave} = 42102.3 + 164.8 * PM_{2.5} + 69.5 * SO_2 + 20.1 * CO - 86.8 * PM_{10} - 229.1 * NO_2 - 272.5 * NO_X - 297.5 * O_3 \quad (6)$$

Furthermore, the Figure 45 displays the equation derivatives of new covid cases during the first wave. It is observed from the output of the third-wave dataset that for every 1 μg/m^3 increase in $PM_{2.5}$, CO, and SO_2, COVID-19 cases will increase by 164.8, 20.1, and 69.5, respectively. When there is a 1 μg/m^3 decrease in PM_{10}, O_3, and NO_2, COVID-19 cases will decrease by 86.8, 297.5, and 229.1, provided that the other parameters remain unchanged at a particular time. Additionally, with every one ppb increase in NO_X COVID-19, cases will increase by 272.5. Hence, during the third wave, mainly every pollutant has shown a significant association with the spread of the virus in the NCR region.

Equation: UNTITLED Workfile: DELHI NCR THIRD WAVE::Untit...

View | Proc | Object | Print | Name | Freeze | Estimate | Forecast | Stats | Resids

Dependent Variable: CASES
Method: Elastic Net Regularization
Date: 05/28/22 Time: 14:59
Sample: 1/01/2022 1/30/2022
Included observations: 30
Penalty type: Elastic Net (alpha = 0.5)
Lambda at minimum error: 63.12
Regressor transformation: None
Cross-validation method: K-Fold (number of folds = 5), rng=kn, seed=159469012
Selection measure: Mean Squared Error

	(minimum)	(+ 1 SE)	(+ 2 SE)
Lambda	63.12	232.2	536.4
Variable	Coefficients		
C	42102.27	37386.24	34704.90
PM2_5	164.7669	86.41279	37.94661
PM10	-86.81093	-56.49266	-41.71044
NO2	-229.0758	-149.2330	-93.09820
NOX	-272.4715	-202.4848	-135.1388
OZONE	-297.4851	-130.6609	-68.89080
CO	20.07310	-0.049959	0.000000
SO2	69.49292	11.06112	2.173465
d.f.	7	7	6
L1 Norm	43242.45	38022.63	35083.85
R-squared	0.553228	0.457909	0.364531

Figure 41. Model output summary of the third wave

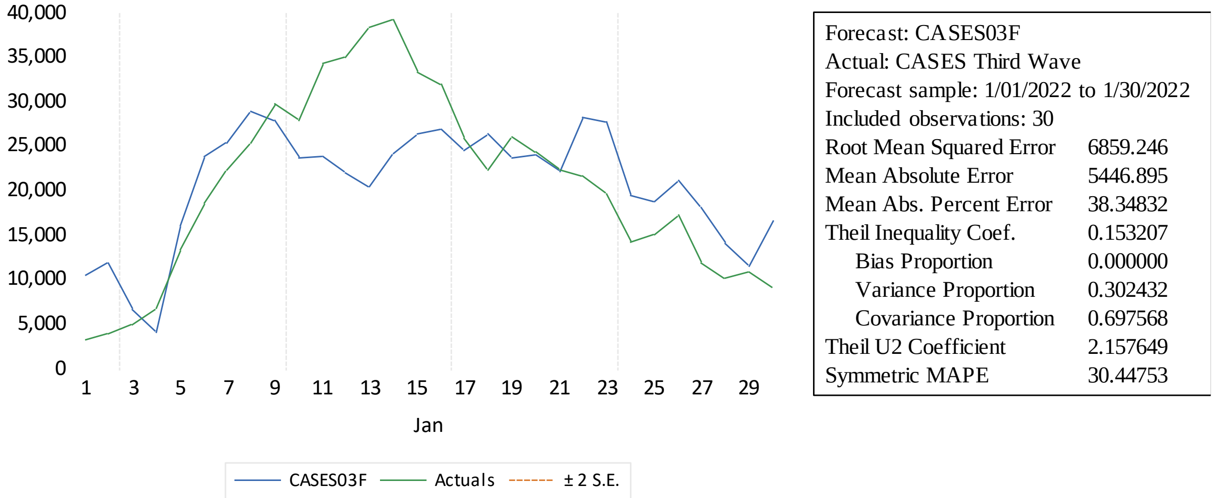


Figure 42. Results after analysing the dataset for the third wave

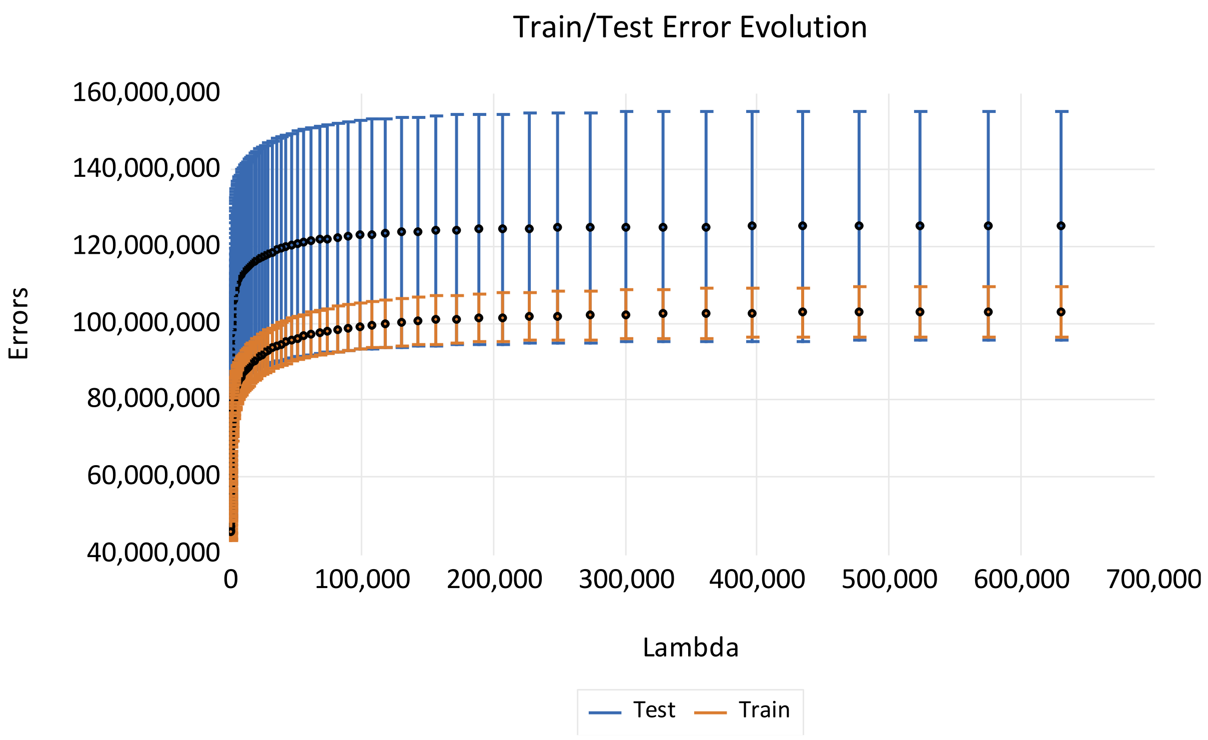


Figure 43. Training and testing of the dataset with K-Fold cross-validation for the third wave

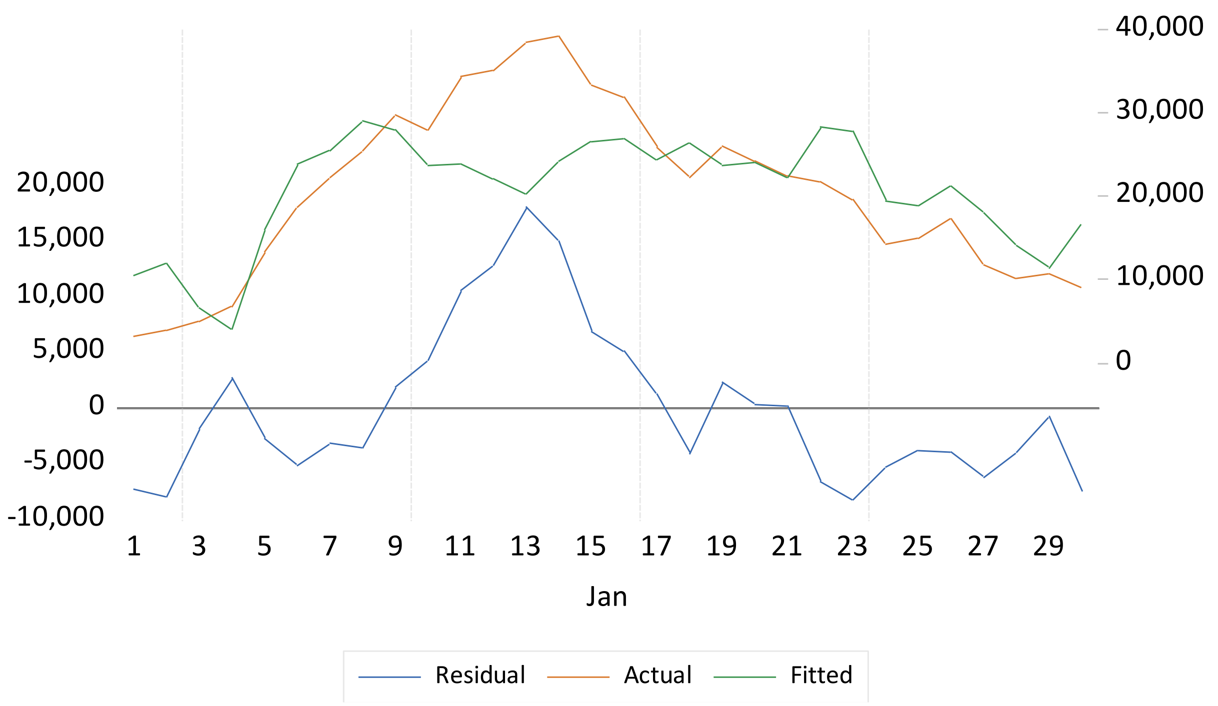


Figure 44. Residual, Actual and Fitted lines of regression model for the third-wave dataset

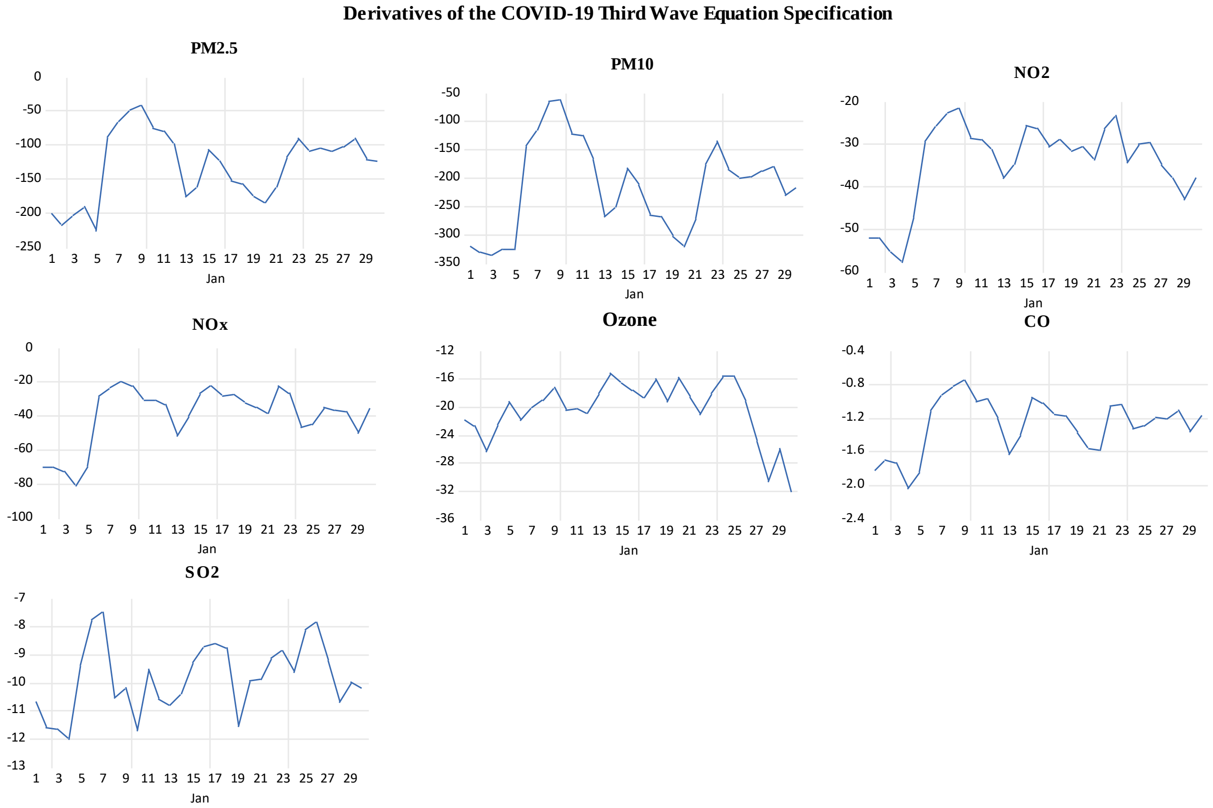


Figure 45. Equation derivatives for the third wave

4.5 Association among air pollution and COVID-19

4.5.1 Revelation to air pollutants when the spread rate of SARS-COV-2 is high

Findings appeared after running the best available method (as shown in Figure 46, and *Table 6*) to resolve the issue of collinearity among predictors and establish a relationship between daily new SARS-COV-2 cases and air pollutants using elastic net regression.

This study indicates that PM2.5 was the primary pollutant showing high correlation with daily new cases. Its influence on corona virus is also increasing during each consecutive study period. Which is making it a severe cause of concern for citizens, as it is seen that the virus is unstoppable even after taking the vaccine.

Additionally, O_3, CO, and SO_2 were among the pollutants which have shown a significant positive association with cases during two out of three study periods, while NOx and PM_{10} were the pollutants that have shown positive association only during one out of three study periods. Finally, nitrogen dioxide was the only pollutant showing negative association during each study period. Findings of Study show that most of the contaminants present in our atmosphere have a significant positive impact on the propagation of the SARS-COV-2 virus in Delhi NCR.

Few previous studies are available worldwide (countries or cities with the problem of poor air quality and COVID-19 infection). The association is examined at the multivariate level for different pollutants. Below sections show a comparison between the findings of this study with the results of previous studies, and the effects of the combination of pollutants and corona virus on human health will be discussed.

Table 6. Elastic net regularization output ($\alpha = 0.5$)

Variables	First Wave	Second Wave	Third Wave
Intercept	4237.7	359.4	42102.3
$PM_{2.5}$	39.6	136.5	164.8
PM_{10}	-32.1	13.8	-86.8
NO_2	-1.2	-20.2	-229.1
NOx	67.0	-94.8	-272.5
O_3	16.9	382.3	-297.5
CO	69.5	-2.2	20.1
SO_2	80.2	-168.4	69.5
Lambda (λ) at minimum error	13.7	61.4	63.1
L 1 Norm	4544.3	1177.6	43242.5
Adjusted R^2	0.385	0.1928	0.5532

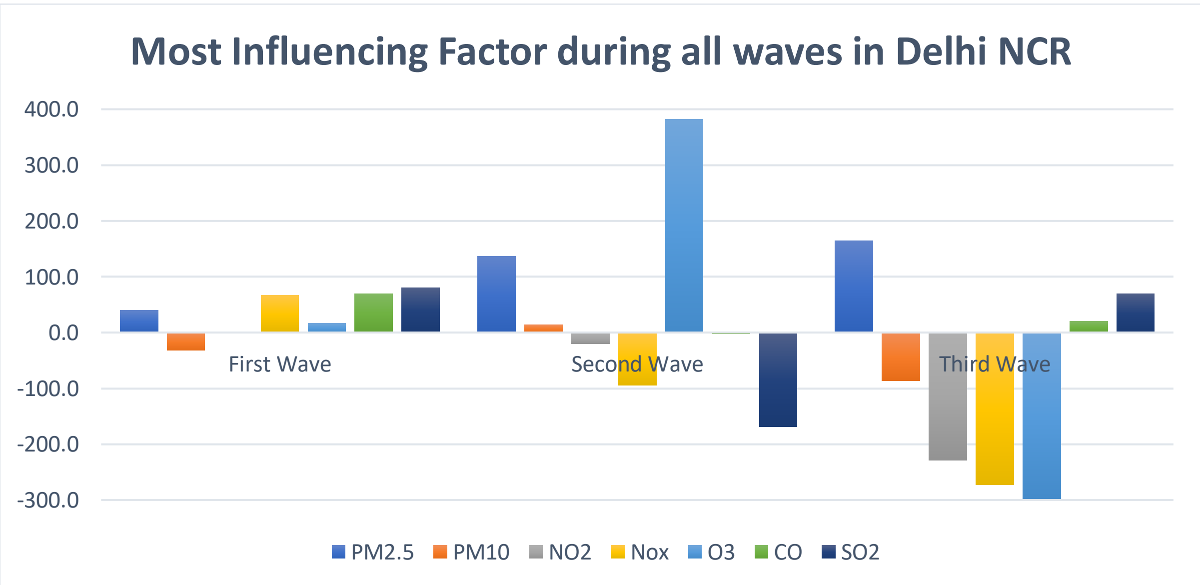


Figure 46. Most influencing predictor during each wave

4.5.2 COVID-19 and Particulate matters ($PM_{2.5}$ & PM_{10})

The influence of particulate matter less than 2.5 and 10 microns on SARS-COV-2 has been linked with oxidative stress, immunological dysregulation, cytotoxicity of polycyclic aromatic hydrocarbons, malfunctioning surfactants, inflammatory effects, metabolic pathways, and ACE-2 (it works as a receptor for COVID-19). $PM_{2.5}$ is the only main pollutant which has consistently shown a strong and positive relationship with SARS-COV-2 daily cases during the first ($\beta = 40.2$, $R^2 = 38.5\%$, $\lambda_{\text{min error}} = 13.7$), second ($\beta = 136.5$, $R^2 = 19.28\%$, $\lambda_{\text{min error}} = 61.4$), and third ($\beta = 164.8$, $R^2 = 55.32\%$, $\lambda_{\text{min error}} = 63.1$) waves. There are some previous studies available worldwide that claim a positive relationship between $PM_{2.5}$ and daily new corona cases, i.e., (Adhikari & Yin, 2020; Fattorini & Regoli, 2020; Frontera et al., 2020; Jiang et al., 2020; Jiang & Xu, 2021; Konstantinoudis et al., 2021; Li et al., 2020; Pei et al., 2021; Travaglio et al., 2021; Vasquez-Apestegui et al., 2021; B. Wang et al., 2020; Q. Wang et al., 2021; Y. Zhu et al., 2020),clearly indicating that $PM_{2.5}$ acts as a carrier for the virus. Air pollution (PM2.5) can contribute to the COVID-19 pandemic in two ways. First, by making people more sensitive to COVID-19 infection by increasing their susceptibility to chronic diseases and putting COVID-19 infected persons at immediate risk, if not death. Second, because COVID-19 can be transferred by microscopic particles or mixed with ultrafine aerosols, the risk of exposure to it is increased, and this situation is also explained in Figure 47 (Annesi-Maesano et al., 2021).

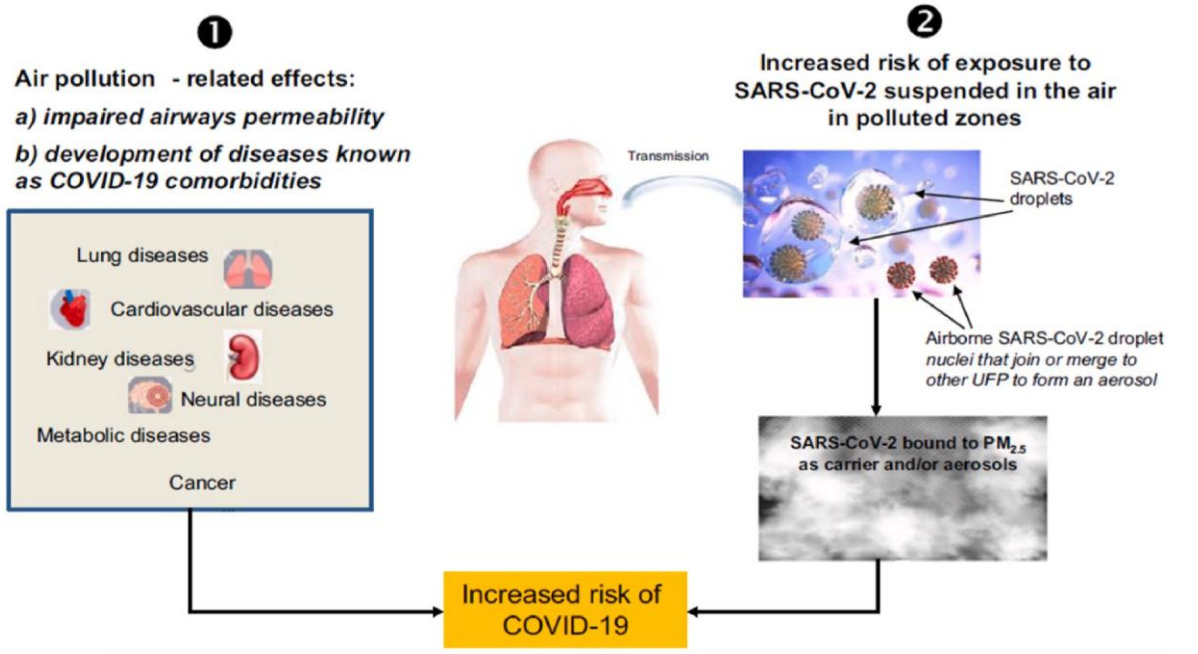


Figure 47. Air pollution and COVID-19 transmission. (Annesi-Maesano et al., 2021)

On the other hand, the findings for PM_{10} exhibited major contradictions since it showed a positive connection with corona virus instances during the second (β = 13.8, R^2= 19.28%, $\lambda_{\text{min error}}$ = 61.4) wave; earlier studies in other nations showed a positive association, i.e., (Liu et al., 2021; Pei et al., 2021; Setti et al., 2020; Travaglio et al., 2021; B. Wang et al., 2020; Q. Wang et al., 2021; Y. Zhu et al., 2020) establish a role of PM_{10} in virus spread. However, the model a revealed negative correlation for the first (β = -32.6, R^2= 38.5%, $\lambda_{\text{min error}}$ = 13.7) and third (β = -86.8, R^2= 55.32%, $\lambda_{\text{min error}}$ = 63.1) wave, which is also seen by (Bontempi, 2020), he gave an observation that city with serve PM_{10} pollution are negatively associated with cases.

4.5.3 COVID-19 and Nitrogen Dioxide (NO_2)

The effects of nitrogen dioxide on COVID-19 have been correlated with monocyte enrichment, increased pneumonic epithelial permeability, inflammatory impacts, immunological dysregulation, and lipid pathways. In this study NO_2 shown negative association among SARS-COV-2 daily cases during the first (β = -3.9, R^2= 38.5%, $\lambda_{\text{min error}}$ = 13.7), second (β = -20.2, R^2= 19.28%, $\lambda_{\text{min error}}$ = 61.4), and third (β = -229.2, R^2= 55.32%, $\lambda_{\text{min error}}$ = 63.1) wave, which indicate that there is an inverse relationship among both, which is also has been studied by other researchers (Lin et al., 2020; Zoran et al., 2020a). Previous studies in the Delhi region also found a negative connection with NO_2, but epidemiological studies reveal that nitrogen dioxide increases our respiratory system's vulnerability to coronavirus infection (Mele & Magazzino, 2021).

4.5.4 COVID-19 and Ground-level Ozone (O_3)

The effects of ground-level ozone on COVID-19 have been because it can ameliorate inflammation and pain because of antiparasitic, bactericidal, and virucidal properties (Fernández-Cuadros et al., 2020). Additionally, O_3, with the help of its metabolites, can regulate our immune structure by administering the release of cytokines (Alberto, 2011). As a result of the presence of the host immune system, it can enhance the development of bacterial activity with the help of ozone (Babior et al., 2003). O_3 shown a positive connection with new daily cases during first (β = 16.72, R^2= 38.5%, $\lambda_{\text{min error}}$ = 13.7) and second (β = 382.3, R^2=

19.28%, $\lambda_{\text{min error}}$ = 61.4) wave, also seen in other studies, i.e., (Adhikari & Yin, 2020; Liu et al., 2021; Travaglio et al., 2021; Y. Zhu et al., 2020; Zoran et al., 2020a). However, during the third wave, ozone showed a negative correlation with corona cases (β = -297.5, R^2= 55.32%, $\lambda_{\text{min error}}$ = 63.1), which also has been observed by several researchers, i.e., (Fronza et al., 2020; Jiang et al., 2020; Liang et al., 2020). (Nassan et al., 2021), claimed that ozone does not induce metabolite changes, but its provisional exposure is associated with the SAM metabolism cysteine, a member of taurine, cysteine, and methionine.

4.5.5 COVID-19 and Carbon Monoxide (CO)

The effects of carbon monoxide on COVID-19 have been correlated. It produces anti-inflammatory effects, vasodilation, and the high concentrations of CO may enhance COVID-19 spread because of the damaged alveolar-capillary unit. Positive associations among carbon monoxide and corona cases were observed in this study during the first (β = 82.52, R^2= 38.5%, $\lambda_{\text{min error}}$ = 13.7) and third (β = 20.1, R^2= 55.32%, $\lambda_{\text{min error}}$ = 63.1) wave, indicating that CO has helped in increasing corona virus spread during both waves, i.e., (Jiang et al., 2020; Lin et al., 2020; Liu et al., 2021; Q. Wang et al., 2021) also studied the same. However, a minor negative relationship was observed during the second wave (β = -2.2, R^2= 19.28%, $\lambda_{\text{min error}}$ = 61.4), which is also seen in another research, i.e., (Jiang and Xu, 2021; Pei et al., 2021).

4.5.6 COVID-19 and Sulfur Dioxide (SO_2)

The effects of sulfur dioxide on COVID-19 have been correlated, i.e., as the concentrations of sulfur dioxide rise, it starts to damage our respiratory tract and increase the susceptibility of a host in our body. In this study, a positive relationship was discovered between SO_2 and corona cases during the first (β = 85.64, R^2= 38.5%, $\lambda_{\text{min error}}$ = 13.7) and third (β = 69.5, R^2= 55.32%, $\lambda_{\text{min error}}$ = 63.1) waves, indicating that SO_2 increased cases in the NCR region during both waves, which is also seen in other researches, i.e., (Lembo et al., 2021; Liu et al., 2021; Q. Wang et al., 2021; Y. Zhu et al., 2020). However, during the second wave (β = -168.4, R^2= 19.28%, $\lambda_{\text{min error}}$ = 61.4), it has shown a negative relationship, also studied by Jiang et al., 2020; Jiang and Xu, 2021.

4.5.7 COVID-19 and NOx

The effects of oxides of nitrogen on COVID-19 have been correlated, as it can cause inflammation of our airway system at a high level when its concentration is high in our surroundings. In this study, there was a positive correlation between NOx and cases only during the first (β = 68.58, R^2= 38.5%, $\lambda_{\text{min error}}$ = 13.7) wave, while in the second (β = -94.8, R^2= 19.28%, $\lambda_{\text{min error}}$ = 61.4) and third (β = -272.5, R^2= 19.28%, $\lambda_{\text{min error}}$ = 61.4) waves, there was an inverse effect concerning the daily COVID -19 cases was found which indicates that NOx was playing a role in the transmission of coronavirus in some circumstances.

CHAPTER 5
CONCLUSION

Beginning from 7 criteria air pollutants and using the best available method to remove multicollinearity among predictors, the results were obtained in which $PM_{2.5}$ pollutant was a strong predictor of daily SARS-COV-2 transmission in Delhi NCR. Furthermore, results also reveal that the (O_3), (CO), and (SO_2) showed a significant positive association in the spread of daily new corona cases. Additionally, PM_{10}, $NO_{2,}$ and NOx have a significantly less or inverse relationship with the spread of the virus.

From these findings, it can be said that air pollutants have played a significant favourable influence on the rate of multiplication of SARS-COV-2 daily cases in the regions where pollution levels are high during each study period. The Indian Government has to make effective policies to control the emission of pollutants in the NCR region, which will further help minimize the influence of climate change and air pollution on current and future pandemics.

REFERENCES

Adhikari, A., & Yin, J. (2020). Short-Term Effects of Ambient Ozone, PM2.5, and Meteorological Factors on COVID-19 Confirmed Cases and Deaths in Queens, New York. *International Journal of Environmental Research and Public Health*, *17*(11), 4047. https://doi.org/10.3390/ijerph17114047

Alberto, P. O. (2011). *Ozone the One and Only Drug* (pp. 143–146). https://doi.org/10.1007/978-3-211-99370-5_22

Amoatey, P., Omidvarborna, H., Baawain, M. S., & Al-Mamun, A. (2020). Impact of building ventilation systems and habitual indoor incense burning on SARS-CoV-2 virus transmissions in Middle Eastern countries. *Science of The Total Environment*, *733*, 139356. https://doi.org/10.1016/j.scitotenv.2020.139356

Annesi-Maesano, I., Maesano, C. N., D'Amato, M., & D'Amato, G. (2021). Pros and cons for the role of air pollution on COVID-19 development. *Allergy*, *76*(8), 2647–2649. https://doi.org/10.1111/all.14818

Ayoub Meo, S., Adnan Abukhalaf, A., Sami, W., & Hoang, T. D. (2021). Effect of environmental pollution PM2.5, carbon monoxide, and ozone on the incidence and mortality due to SARS-CoV-2 infection in London, United Kingdom. *Journal of King Saud University - Science*, *33*(3), 101373. https://doi.org/10.1016/j.jksus.2021.101373

Babior, B. M., Takeuchi, C., Ruedi, J., Gutierrez, A., & Wentworth, P. (2003). Investigating antibody-catalyzed ozone generation by human neutrophils. *Proceedings of the National Academy of Sciences*, *100*(6), 3031–3034. https://doi.org/10.1073/pnas.0530251100

Biktasheva, I. v. (2020). Role of a habitat's air humidity in Covid-19 mortality. *Science of The Total Environment*, *736*, 138763. https://doi.org/10.1016/j.scitotenv.2020.138763

Bontempi, E. (2020). First data analysis about possible COVID-19 virus airborne diffusion due to air particulate matter (PM): The case of Lombardy (Italy). *Environmental Research*, *186*, 109639. https://doi.org/10.1016/j.envres.2020.109639

Bourdrel, T., Annesi-Maesano, I., Alahmad, B., Maesano, C. N., & Bind, M.-A. (2021). The impact of outdoor air pollution on COVID-19: a review of evidence from *in vitro* , animal, and human studies. *European Respiratory Review*, *30*(159), 200242. https://doi.org/10.1183/16000617.0242-2020

Brandt, E. B., & Mersha, T. B. (2021). Environmental Determinants of Coronavirus Disease 2019 (COVID-19). *Current Allergy and Asthma Reports*, *21*(3), 15. https://doi.org/10.1007/s11882-021-00993-1

Chakraborty, P., Jayachandran, S., Padalkar, P., Sitlhou, L., Chakraborty, S., Kar, R., Bhaumik, S., & Srivastava, M. (2020). Exposure to Nitrogen Dioxide (NO2) from

Vehicular Emission Could Increase the COVID-19 Pandemic Fatality in India: A Perspective. *Bulletin of Environmental Contamination and Toxicology*, *105*(2), 198–204. https://doi.org/10.1007/s00128-020-02937-3

Chatterjee, S. and J. S. S. (2013). *Handbook of regression analysis*. John Wiley & Sons.

Chen, G., Zhang, W., Li, S., Williams, G., Liu, C., Morgan, G. G., Jaakkola, J. J. K., & Guo, Y. (2017). Is short-term exposure to ambient fine particles associated with measles incidence in China? A multi-city study. *Environmental Research*, *156*, 306–311. https://doi.org/10.1016/j.envres.2017.03.046

Ciencewicki, J., & Jaspers, I. (2007). Air Pollution and Respiratory Viral Infection. *Inhalation Toxicology*, *19*(14), 1135–1146. https://doi.org/10.1080/08958370701665434

Clay, K., Lewis, J., & Severnini, E. (2018). Pollution, Infectious Disease, and Mortality: Evidence from the 1918 Spanish Influenza Pandemic. *The Journal of Economic History*, *78*(4), 1179–1209. https://doi.org/10.1017/S002205071800058X

Coker, E. S., Cavalli, L., Fabrizi, E., Guastella, G., Lippo, E., Parisi, M. L., Pontarollo, N., Rizzati, M., Varacca, A., & Vergalli, S. (2020). The Effects of Air Pollution on COVID-19 Related Mortality in Northern Italy. *Environmental and Resource Economics*, *76*(4), 611–634. https://doi.org/10.1007/s10640-020-00486-1

Cui, Y., Zhang, Z.-F., Froines, J., Zhao, J., Wang, H., Yu, S.-Z., & Detels, R. (2003). *Environmental Health: A Global Access Science Source Air pollution and case fatality of SARS in the People's Republic of China: an ecologic study*. http://www.ehjournal.net/content/2/1/15

Dutta, A., & Dutta, G. (2021). Association of air pollution and meteorological variables with the two waves of COVID-19 pandemic in Delhi: A critical analysis. *Heliyon*, *7*(11), e08468. https://doi.org/10.1016/j.heliyon.2021.e08468

Edouard Grave, G. R. O. F. B. (2011). Trace Lasso: a trace norm regularization for correlated designs. *Advances in Neural Information Processing Systems*, *24*.

Fattorini, D., & Regoli, F. (2020). Role of the chronic air pollution levels in the Covid-19 outbreak risk in Italy. *Environmental Pollution*, *264*, 114732. https://doi.org/10.1016/j.envpol.2020.114732

Fernández-Cuadros, M. E., Albaladejo-Florín, M. J., Peña-Lora, D., Álava-Rabasa, S., & Pérez-Moro, O. S. (2020). Ozone (O3) and SARS-CoV-2: Physiological Bases and Their Therapeutic Possibilities According to COVID-19 Evolutionary Stage. *SN Comprehensive Clinical Medicine*, *2*(8), 1094–1102. https://doi.org/10.1007/s42399-020-00328-7

Filippini, T., Rothman, K. J., Goffi, A., Ferrari, F., Maffeis, G., Orsini, N., & Vinceti, M. (2020). Satellite-detected tropospheric nitrogen dioxide and spread of SARS-CoV-2 infection in Northern Italy. *Science of The Total Environment*, *739*, 140278. https://doi.org/10.1016/j.scitotenv.2020.140278

Frontera, A., Cianfanelli, L., Vlachos, K., Landoni, G., & Cremona, G. (2020). Severe air pollution links to higher mortality in COVID-19 patients: The "double-hit" hypothesis. *Journal of Infection*, *81*(2), 255–259. https://doi.org/10.1016/j.jinf.2020.05.031

Fronza, R., Lusic, M., Schmidt, M., & Lucic, B. (2020). Spatial–Temporal Variations in Atmospheric Factors Contribute to SARS-CoV-2 Outbreak. *Viruses*, *12*(6), 588. https://doi.org/10.3390/v12060588

Hendryx, M., & Luo, J. (2020). COVID-19 prevalence and fatality rates in association with air pollution emission concentrations and emission sources. *Environmental Pollution*, *265*, 115126. https://doi.org/10.1016/j.envpol.2020.115126

Jiang, Y., Wu, X.-J., & Guan, Y.-J. (2020). Effect of ambient air pollutants and meteorological variables on COVID-19 incidence. *Infection Control & Hospital Epidemiology*, *41*(9), 1011–1015. https://doi.org/10.1017/ice.2020.222

Jiang, Y., & Xu, J. (2021). The association between COVID-19 deaths and short-term ambient air pollution/meteorological condition exposure: a retrospective study from Wuhan, China. *Air Quality, Atmosphere & Health*, *14*(1), 1–5. https://doi.org/10.1007/s11869-020-00906-7

Kim, M.-H., Banerjee, S., Park, S. M., & Pathak, J. (2016). Improving risk prediction for depression via Elastic Net regression - Results from Korea National Health Insurance Services Data. *AMIA ... Annual Symposium Proceedings. AMIA Symposium*, *2016*, 1860–1869.

Kolluru, S. S. R., Patra, A. K., Nazneen, & Shiva Nagendra, S. M. (2021). Association of air pollution and meteorological variables with COVID-19 incidence: Evidence from five megacities in India. *Environmental Research*, *195*, 110854. https://doi.org/10.1016/j.envres.2021.110854

Konstantinoudis, G., Padellini, T., Bennett, J., Davies, B., Ezzati, M., & Blangiardo, M. (2021). Long-term exposure to air-pollution and COVID-19 mortality in England: A hierarchical spatial analysis. *Environment International*, *146*, 106316. https://doi.org/10.1016/j.envint.2020.106316

Kuang, D., Choo, J., & Park, H. (2015). Nonnegative Matrix Factorization for Interactive Topic Modeling and Document Clustering. In *Partitional Clustering Algorithms* (pp. 215–243). Springer International Publishing. https://doi.org/10.1007/978-3-319-09259-1_7

Lembo, R., Landoni, G., Cianfanelli, L., & Frontera, A. (2021). Air pollutants and SARS-CoV-2 in 33 European countries. *Acta Bio-Medica : Atenei Parmensis*, *92*(1), e2021166. https://doi.org/10.23750/abm.v92i1.11155

Li, H., Xu, X.-L., Dai, D.-W., Huang, Z.-Y., Ma, Z., & Guan, Y.-J. (2020). Air pollution and temperature are associated with increased COVID-19 incidence: A time series study. *International Journal of Infectious Diseases*, *97*, 278–282. https://doi.org/10.1016/j.ijid.2020.05.076

Liang, D., Shi, L., Zhao, J., Liu, P., Sarnat, J. A., Gao, S., Schwartz, J., Liu, Y., Ebelt, S. T., Scovronick, N., & Chang, H. H. (2020). Urban Air Pollution May Enhance COVID-19 Case-Fatality and Mortality Rates in the United States. *The Innovation*, *1*(3), 100047. https://doi.org/10.1016/j.xinn.2020.100047

Lin, S., Wei, D., Sun, Y., Chen, K., Yang, L., Liu, B., Huang, Q., Paoliello, M. M. B., Li, H., & Wu, S. (2020). Region-specific air pollutants and meteorological parameters influence COVID-19: A study from mainland China. *Ecotoxicology and Environmental Safety*, *204*, 111035. https://doi.org/10.1016/j.ecoenv.2020.111035

Liu, Q., Xu, S., & Lu, X. (2021). Association between air pollution and COVID-19 infection: evidence from data at national and municipal levels. *Environmental Science and Pollution Research*, *28*(28), 37231–37243. https://doi.org/10.1007/s11356-021-13319-5

Magazzino, C., Mele, M., & Sarkodie, S. A. (2021). The nexus between COVID-19 deaths, air pollution and economic growth in New York state: Evidence from Deep Machine Learning. *Journal of Environmental Management*, *286*, 112241. https://doi.org/10.1016/j.jenvman.2021.112241

Magazzino, C., Mele, M., & Schneider, N. (2020). The relationship between air pollution and COVID-19-related deaths: An application to three French cities. *Applied Energy*, *279*, 115835. https://doi.org/10.1016/j.apenergy.2020.115835

McNeish, D. M. (2015). Using Lasso for Predictor Selection and to Assuage Overfitting: A Method Long Overlooked in Behavioral Sciences. *Multivariate Behavioral Research*, *50*(5), 471–484. https://doi.org/10.1080/00273171.2015.1036965

Mele, M., & Magazzino, C. (2021). Pollution, economic growth, and COVID-19 deaths in India: a machine learning evidence. *Environmental Science and Pollution Research*, *28*(3), 2669–2677. https://doi.org/10.1007/s11356-020-10689-0

Mele, M., Magazzino, C., Schneider, N., & Strezov, V. (2021). NO2 levels as a contributing factor to COVID-19 deaths: The first empirical estimate of threshold values. *Environmental Research*, *194*, 110663. https://doi.org/10.1016/j.envres.2020.110663

Meo, S. A., Ahmed Alqahtani, S., Saad binmeather, F., Abdulrhman AlRasheed, R., Mohammed Aljedaie, G., & Mohammed Albarrak, R. (2022). Effect of environmental pollutants PM2.5, CO, O3 and NO2, on the incidence and mortality of SARS-COV-2 in largest metropolitan cities, Delhi, Mumbai and Kolkata, India. *Journal of King Saud University - Science*, *34*(1), 101687. https://doi.org/10.1016/J.JKSUS.2021.101687

Nassan, F. L., Kelly, R. S., Kosheleva, A., Koutrakis, P., Vokonas, P. S., Lasky-Su, J. A., & Schwartz, J. D. (2021). Metabolomic signatures of the long-term exposure to air pollution and temperature. *Environmental Health*, *20*(1), 3. https://doi.org/10.1186/s12940-020-00683-x

Ogen, Y. (2020). Assessing nitrogen dioxide (NO2) levels as a contributing factor to coronavirus (COVID-19) fatality. *Science of The Total Environment*, *726*, 138605. https://doi.org/10.1016/j.scitotenv.2020.138605

Pei, L., Wang, X., Guo, B., Guo, H., & Yu, Y. (2021). Do air pollutants as well as meteorological factors impact Corona Virus Disease 2019 (COVID-19)? Evidence from China based on the geographical perspective. *Environmental Science and Pollution Research*, *28*(27), 35584–35596. https://doi.org/10.1007/s11356-021-12934-6

Pope, C. A., Burnett, R. T., Thurston, G. D., Thun, M. J., Calle, E. E., Krewski, D., & Godleski, J. J. (2004). Cardiovascular Mortality and Long-Term Exposure to Particulate Air Pollution. *Circulation*, *109*(1), 71–77. https://doi.org/10.1161/01.CIR.0000108927.80044.7F

Ravindra, K., Rattan, P., Mor, S., & Aggarwal, A. N. (2019). Generalized additive models: Building evidence of air pollution, climate change and human health. *Environment International*, *132*, 104987. https://doi.org/10.1016/j.envint.2019.104987

Saez, M., Tobias, A., & Barceló, M. A. (2020). Effects of long-term exposure to air pollutants on the spatial spread of COVID-19 in Catalonia, Spain. *Environmental Research*, *191*, 110177. https://doi.org/10.1016/j.envres.2020.110177

Setti, L., Passarini, F., de Gennaro, G., Barbieri, P., Licen, S., Perrone, M. G., Piazzalunga, A., Borelli, M., Palmisani, J., di Gilio, A., Rizzo, E., Colao, A., Piscitelli, P., & Miani, A. (2020). Potential role of particulate matter in the spreading of COVID-19 in Northern Italy: first observational study based on initial epidemic diffusion. *BMJ Open*, *10*(9), e039338. https://doi.org/10.1136/bmjopen-2020-039338

Travaglio, M., Yu, Y., Popovic, R., Selley, L., Leal, N. S., & Martins, L. M. (2021). Links between air pollution and COVID-19 in England. *Environmental Pollution*, *268*, 115859. https://doi.org/10.1016/j.envpol.2020.115859

Troeger, C., Forouzanfar, M., Rao, P. C., Khalil, I., Brown, A., Swartz, S., Fullman, N., Mosser, J., Thompson, R. L., Reiner, R. C., Abajobir, A., Alam, N., Alemayohu, M. A., Amare, A. T., Antonio, C. A., Asayesh, H., Avokpaho, E., Barac, A., Beshir, M. A., … Mokdad, A. H. (2017). Estimates of the global, regional, and national morbidity, mortality, and aetiologies of lower respiratory tract infections in 195 countries: a systematic analysis for the Global Burden of Disease Study 2015. *The Lancet Infectious Diseases*, *17*(11), 1133–1161. https://doi.org/10.1016/S1473-3099(17)30396-1

Vasquez-Apestegui, B. v., Parras-Garrido, E., Tapia, V., Paz-Aparicio, V. M., Rojas, J. P., Sanchez-Ccoyllo, O. R., & Gonzales, G. F. (2021). Association between air pollution in Lima and the high incidence of COVID-19: findings from a post hoc analysis. *BMC Public Health*, *21*(1), 1161. https://doi.org/10.1186/s12889-021-11232-7

Wang, B., Liu, J., Li, Y., Fu, S., Xu, X., Li, L., Zhou, J., Liu, X., He, X., Yan, J., Shi, Y., Niu, J., Yang, Y., Li, Y., Luo, B., & Zhang, K. (2020). Airborne particulate matter, population mobility and COVID-19: a multi-city study in China. *BMC Public Health*, *20*(1), 1585. https://doi.org/10.1186/s12889-020-09669-3

Wang, Q., Dong, W., Yang, K., Ren, Z., Huang, D., Zhang, P., & Wang, J. (2021). Temporal and spatial analysis of COVID-19 transmission in China and its influencing factors. *International Journal of Infectious Diseases*, *105*, 675–685. https://doi.org/10.1016/j.ijid.2021.03.014

Wen, C., Akram, R., Irfan, M., Iqbal, W., Dagar, V., Acevedo-Duqued, Á., & Saydaliev, H. B. (2022). The asymmetric nexus between air pollution and COVID-19: Evidence from a non-linear panel autoregressive distributed lag model. *Environmental Research*, *209*, 112848. https://doi.org/10.1016/j.envres.2022.112848

Wu, X., Nethery, R. C., Sabath, B., Braun, D., Dominici, F., & James, C. (n.d.). *Exposure to air pollution and COVID-19 mortality in the United States: A nationwide cross-sectional study*. https://doi.org/10.1101/2020.04.05.20054502

Yao, Y., Pan, J., Liu, Z., Meng, X., Wang, W., Kan, H., & Wang, W. (2020). Temporal association between particulate matter pollution and case fatality rate of COVID-19 in Wuhan. *Environmental Research*, *189*, 109941. https://doi.org/10.1016/j.envres.2020.109941

Yao, Y., Pan, J., Liu, Z., Meng, X., Wang, W., Kan, H., & Wang, W. (2021). Ambient nitrogen dioxide pollution and spreadability of COVID-19 in Chinese cities. *Ecotoxicology and Environmental Safety*, *208*, 111421. https://doi.org/10.1016/j.ecoenv.2020.111421

Zhu, N., Zhang, D., Wang, W., Li, X., Yang, B., Song, J., Zhao, X., Huang, B., Shi, W., Lu, R., Niu, P., Zhan, F., Ma, X., Wang, D., Xu, W., Wu, G., Gao, G. F., & Tan, W. (2020). A Novel Coronavirus from Patients with Pneumonia in China, 2019. *New England Journal of Medicine*, *382*(8), 727–733. https://doi.org/10.1056/NEJMoa2001017

Zhu, Y., Xie, J., Huang, F., & Cao, L. (2020). Association between short-term exposure to air pollution and COVID-19 infection: Evidence from China. *Science of The Total Environment*, *727*, 138704. https://doi.org/10.1016/j.scitotenv.2020.138704

Zoran, M. A., Savastru, R. S., Savastru, D. M., & Tautan, M. N. (2020a). Assessing the relationship between ground levels of ozone (O3) and nitrogen dioxide (NO2) with coronavirus (COVID-19) in Milan, Italy. *Science of The Total Environment*, *740*, 140005. https://doi.org/10.1016/j.scitotenv.2020.140005

Zoran, M. A., Savastru, R. S., Savastru, D. M., & Tautan, M. N. (2020b). Assessing the relationship between surface levels of PM2.5 and PM10 particulate matter impact on COVID-19 in Milan, Italy. *Science of The Total Environment*, *738*, 139825. https://doi.org/10.1016/j.scitotenv.2020.139825

Zou, H., & Hastie, T. (2005). Regularization and variable selection via the elastic net. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, *67*(2), 301–320. https://doi.org/10.1111/j.1467-9868.2005.00503.x