5-4-2022

Spatiotemporal Analysis of COVID-19 Infection and Air Quality in India

Grace Yu

Georgia State University

Spatiotemporal Analysis of COVID-19 Infection and Air Quality in India

by

Grace Yu

Under the Direction of Yi Jiang PHD

A Thesis Submitted in Partial Fulfillment of the Requirements for the Degree of

Master of Science

in the College of Arts and Sciences

Georgia State University

2022

ABSTRACT

COVID-19 infects the respiratory tract leading to lung damage. Air pollutants such as PM 2.5 is one of the main causes and aggravating factors of many respiratory diseases. A known COVID-19 and air pollution "hotspot" is India. India reported a devastating number of COVID-19 cases in early 2020. As the country went into lockdown, the air quality improved significantly, providing a rare opportunity to study correlation between COVID-19 cases and air quality. The spatial autocorrelation analysis between the regions for air quality and COVID-19 cases revealed no significant clustering within the regions. Cross-correlation in time series and regression analysis established a positive correlation between PM2.5 emissions and COVID-19 deaths with a time lag of 20-60 days. Spatiotemporal correlation reveals that there is a positive correlation of PM 2.5 and COVID-19 deaths with time lag of 30-50 days and 300 kilometers.

INDEX WORDS: Spatial Analysis, COVID-19, Air Quality, India, Autocorrelation, Time Series

Spatiotemporal Analysis Of COVID-19 Infection and Air Quality in India

by

Grace Yu

Committee Chair: Yi Jiang

Committee: Gengsheng Qin

Jun Kong

Electronic Version Approved:

Office of Graduate Services

College of Arts and Sciences

Georgia State University

May 2022

DEDICATION

While writing this thesis, I overcame some personal difficulties through faith. I dedicate this thesis to God. I also dedicate this thesis to my family and friends, who had been patient and kind towards me during difficult times.

ACKNOWLEDGEMENTS

I would like to express my sincere gratitude to my advisor, Dr. Yi Jiang. Her patience and guidance supported me to complete this thesis and our lab meetings were always insightful. I would also like to thank my thesis committee, Dr. Gengsheng Qin and Dr. Jun Kong on their expertise and insights. I am very grateful for all the wonderful teachings I have received from the mathematics and statistics department during my studies at Georgia State University.

TABLE OF CONTENTS

1 INTRODUCTION

1.1 COVID-19 and Air Quality Relevance

In December 2019, a respiratory outbreak caused by SARS-CoV-2 virus was identified for the first time in Wuhan district of China. Since then, the contagious disease rapidly spread to all parts of the world, leading to WHO's declaration of a COVID-19 pandemic in March 2020. It strongly affected the population dense parts of the world such as cities in Europe, North America, and Asia. Some locations can be named COVID-19 "hotspots", where the population is dense hence allowing for greater spread. These hotspots are spatial clusters described as areas of elevated incidence, higher transmission risk, or higher probability of disease emergence.

SARS-CoV-2 viruses infect the respiratory tract and lead to lasting lung damage. At the same time air pollution is a cause and aggravating factor of many respiratory diseases. Fine particulate matter, or PM 2.5, is one the causative factors of human non-accidental death. PM 2.5 refers to tiny particles or droplets in the air that are 2.5 microns or less in width (Liu et al, 2021). It is known to cause asthma, respiratory inflammation, and cancers. Cars, exhausts, power plants, smoking, cooking, heaters, and fireplaces are many examples of causes of PM 2.5 (Sahu et al, 2021).

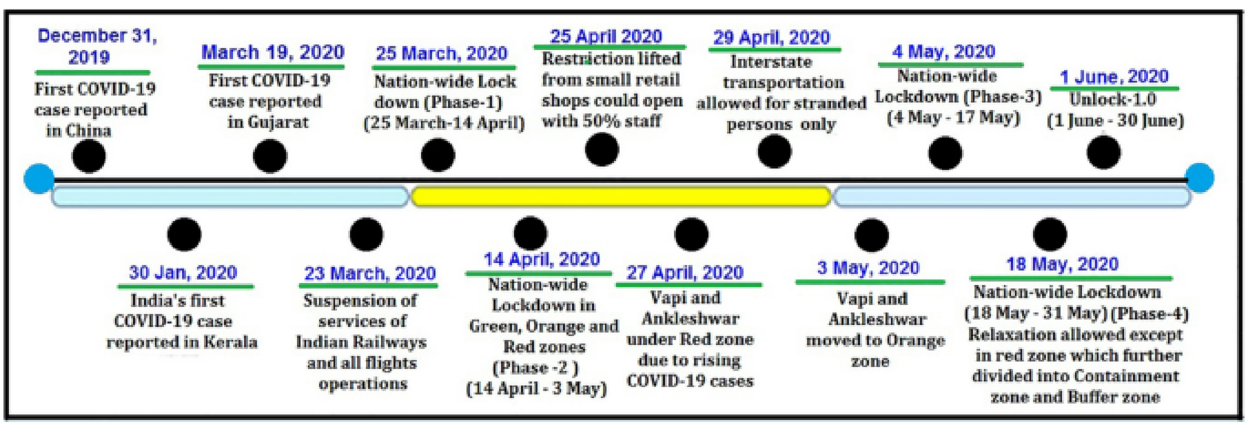
Long-term air quality data has been found to correlate with COVID-19 cases in the Italian Provinces which suggest that chronic atmospheric pollution may favor COVID-19 spread. Citizens of Varese in Italy were linked to regional datasets for COVID-19 case ascertainment and annual exposure to outdoor concentration of air quality (Veronesi et al., 2022). Long-term exposure to high levels of air pollutants, especially PM2.5 increased the incidence of COVID-19 was concluded using Poisson regression models (Veronesi et al., 2022). Various other research has been published to suggest a relationship between air pollution and climate change with COVID-19 cases (Veronesi et al., 2022).

1.2 COVID-19 in India

A known COVID-19 hotspot is the country of India. Ever since the first confirmed COVID-19 cases on January 30, 2020, India reported a devastating number of COVID-19 cases in early 2020 (Ghosh et al, 2020). To stop the spread of the viral disease, India declared a nation-wide lockdown. As the country went into lockdown, the air quality improved drastically, providing a rare opportunity to study correlation between COVID-19 infection and polluted air quality. The lockdown policies timeline for India is shown in Table 1.1 which shows that the beginning of lockdown started on March 25, 2020.

Table 1.1 Lockdown Policy Dates for India

Source: Ministry of Home Affairs, Govt. of India



1.3 Air Quality in India

India is one of the world's worst countries in air pollution. Many of India's cities are polluted due to emissions from vehicles, industry, brick kilns, coal-based power plants, and crop residue burning (Singh et al. 2004, Prasad et al. 2006, Venkataraman et al. 2018). Poor air quality has continuously increased throughout the years in India. Table 1.2 shows the Environmental Protection Agency (EPA)'s categories of PM 2.5 levels (Du & Varde, 2016). According to IQ Air, the average PM 2.5 index level for India is 83.2 $\mu g/m^3$ which is in the "Unhealthy" category. India's toxic air is known to kill more than one million people each year and increases risk for heart disease, diabetes, and respiratory disease (Chatterjee P., 2019).

Table 1.2 EPA PM 2.5 Levels by Category

AQI Category	Index Values	Previous Breakpoints (1999 AQI) ($\mu g/m^3$, 24-hour average)	Revised Breakpoints ($\mu g/m^3$, 24-hour average)
Good	0 - 50	0.0 - 15.0	0.0 – 12.0
Moderate	51 - 100	>15.0 - 40	12.1 – 35.4
Unhealthy for Sensitive Groups	101 – 150	>40 – 65	35.5 – 55.4
Unhealthy	151 – 200	> 65 – 150	55.5 – 150.4
Very Unhealthy	201 – 300	> 150 – 250	150.5 – 250.4
Hazardous	301 – 400	> 250 – 350	250.5 – 350.4
	401 – 500	> 350 – 500	350.5 – 500

1.4 COVID-19 and Air Quality Relationship in India

Due to the COVID-19's lockdown policies in early 2020, there were reports that India's air quality improved during those months due to the reduction in industrial and transportation emissions (Sahu et al, 2021). As the country went into lockdown, it provided a rare opportunity to study the correlation between COVID infection and polluted air quality. The purpose of this thesis is to perform a spatial and temporal analysis of air quality and COVID-19 cases and mortality in India.

2 LITERATURE REVIEW

2.1 Spatial Autocorrelation

Spatial Autocorrelation describes the presence of systematic spatial variation in a variable and positive spatial autocorrelation is the tendency for areas or sites that are close together to have similar values (Haining, 2001). It can be indexed or quantified by including an autoregressive parameter in a regression model or filtered from variables. It is the fundamental concept in spatial analysis (Getis, 1992).

2.1.1 Moran's I statistic

Global Moran's I statistic is a tool that measures spatial autocorrelation based on locations and feature values at the same time. Given features and associated attribute, it would evaluate whether there is a pattern that is clustered, dispersed, or random. Moran's I Index value and a z-score and p-value evaluates the significance of the Index. The calculation for Moran's I statistic is as below (Getis, 1992):

$$I = \frac{n}{S_o}\frac{\sum_{i=1}^{n}\sum_{j=1}^{n} w_{i,j} z_i z_j}{\sum_{i=1}^{n} z_i^2}$$

Where z_i is the deviation of an attribute for feature I from its mean $(x_i - \bar{X}), w_{i,j}$ is the spatial weight between feature i and j, n is equal to the total number of features, and S_o is the aggregate of all the spatial weights:

$$S_o = \sum_{i=1}^{n}\sum\nolimits_{j=1}^{n} w_{i,j}$$

The z_I score for the statistic is computed as:

$$Z_I = \frac{I - E[I]}{\sqrt{V[I]}}$$

Global Moran's I index must always be interpreted within the context of its null hypothesis. The null hypothesis states that the attribute is randomly distributed among the study area. Only if the p-value is statistically significant, the null hypothesis can be rejected.

2.2 Linear Regression

Linear regression is a statistical method examining the linear relationship between quantitative variables. The model equation can be expressed as per equation,

$$y = \alpha + \beta x_i$$

The x_i is the explanatory variable and y is the dependent variable. The slope of the line is β, and α is the intercept.

$$r_{xy} = \frac{\overline{xy} - \bar{x}\bar{y}}{\sqrt{(\overline{x^2} - x^2)(\overline{y^2} - y^2)}}$$

The coefficient of determination, or R squared, is equal to r_{xy}^2 when the model is linear with a single independent variable.

2.3 Cross-correlation Time Series

Given two time series x_t and y_t, we can delay x_t by T samples and then calculate the cross-covariance between the pair of signals

$$\sigma_{xy}(T) = \frac{1}{N-1}(x_{t-T} - \mu_x)(y_t - \mu_y)$$

where μ_x and μ_y are the means of each time series and there are N samples in each. The function $\sigma_{xy}(T)$ is the cross-covariance function. The cross-correlation is a normalized version

$$r_{xy}(T) = \frac{\sigma_{xy}(T)}{\sqrt{\sigma_{xx}(0)\sigma_{yy}(0)}}$$

Where we note that $\sigma_{xx}(0) = \sigma_x^2$ and $\sigma_{yy}(0) = \sigma_y^2$ are the variances of each signal

$$r_{xy}(0) = \frac{\sigma_{xy}}{\sigma_x \sigma_y}$$

2.4 Spatiotemporal Correlation

Spatiotemporal is when data is collected across space and time which describes a particular location at a specific period of time. Spatial correlation relates to the existence of a functional relationship between "what happens at one point in space and what happens elsewhere" Spatiotemporal correlation compares spatial patterns observed at multiple times. Existing literature aims to investigate the spatial and temporal characteristics and uses a spatiotemporal (ST) kriging tool to examine the ST variables across space and time (Raja et al. 2017). Kriging was originally used for spatial data only, but now allows for the incorporation of temporal information by considering temporal data as a co-kriged variable (Aydin, 2020). Much of the spatiotemporal analysis is conducted through data visualization. One example of past research is the observation of seasonality of drought in certain locations in Ethiopia (Moges, 2019). The amount of rainfall was observed in respect to years and locations of rainfall stations. Other example includes the application of spatiotemporal computing in the mining industry (Chaulya, 2016). Environmental modeling uses spatiotemporal analysis to represent processes that occur in the real world in space and time.

3 METHODS

The data points are in 3 different time frames: before lockdown (Jan 30 - Mar 24, 2020), during lockdown (Mar 25 – May 31, 2020), and after lockdown (Jun 1- Jun 30, 2020) in the spatial autocorrelation dataset. Daily PM2.5 levels were collected from *Air Quality Open Data Platform* in which 14 distinct cities were usable as air quality stations. Number of 'unhealthy' days was calculated by using the PM2.5 levels of the cities within before/during/after date ranges. To utilize PM2.5 levels for all states, an estimated PM 2.5 gross yearly emission level was found using previous literature (Sahu et al., 2021). Daily and cumulative COVID-19 confirmed and deceased case numbers were collected from *COVID19India.org* and mortality rate was calculated.

3.1 Analysis of Moran's I statistic for PM2.5 levels in India

The data points are in 3 different date ranges: before lockdown (Dec 31-Mar 24, 2020), during lockdown (Mar 25 – May 31, 2020), and after lockdown (Jun 1-Jun 30, 2020). Using the *sp, rgdal spedp,* and *rgeos* libraries in R, spatial data analysis was conducted. A map of before, during, and after lockdown of deceased number of COVID-19 cases by state was created to run a global spatial autocorrelation for a shapefile. A global spatial autocorrelation is run by defining neighbors which will give a single measure to represent the area. The single measurement, or Moran's I statistic is a correlation score between -1 and 1, and like a correlation coefficient, 1 represents a perfect positive spatial autocorrelation, 0 identifies that the data is randomly distributed, and -1 represents a negative spatial autocorrelation. This method was repeated for the mortality rate by state to observe the Moran's I statistic. Because there are only 14 usable cities with air quality stations, a spatial autocorrelation cannot be found, but a visual map was plotted to observe the change of PM2.5 levels during before/during/after lockdown dates.

3.2 Linear Regression of COVID-19 Cases and Air Quality Variables

Linear regression model was used in R to find the relationship between COVID-19 cases and estimated PM 2.5 emission. The model was evaluated using p-value and R-squared values to test for significance of relationship between the PM2.5 emission levels and COVID-19 cases. The methods were repeated for confirmed cases vs. PM 2.5 levels, deceased cases vs. PM 2.5 levels, mortality rate vs. PM 2.5 levels.

3.3 Time Series Cross-Correlation

Two relatively high level PM2.5 cities were selected and two relatively low-level PM 2.5 cities were selected for time series cross-correlation. Because India's overall air quality difference is miniscule, the selection process used the spatial maps to visualize India's PM2.5 improvement during the period of lockdown. From the visual map, the states with improvements and no improvements in air quality can be identified. The high-level PM 2.5 states were selected to be Delhi and West Bengal, and low-level PM 2.5 states were Tamil Nadu and Andhra Pradesh. The daily deceased cases and PM 2.5 levels were each plotted on a time series graph. Using the R libraries, *feasts, tsibble, lubridate,* the cross-correlation function between daily PM 2.5 levels and deceased COVID cases separated by *k* days was found. The method was repeated for each city to observe the correlation between high-level PM2.5 cities and low-level PM 2.5 cities against COVID-19 cases.

3.4 Spatiotemporal Correlation Analysis

Correlation is found using the R function, cor(), with uses the Pearson correlation formula:

$$r = \frac{\sum(x - m_x)(y - m_y)}{\sqrt{\sum(x - m_x)^2 \sum(y - m_y)^2}}$$

$$m_x \; and \; m_y \; are \; the \; means \; of \; x \; and \; y \; variables$$

3.4.1 Temporal Correlation

First, a temporal correlation is like the time series cross-correlation in that the observed correlation describes the relationship between the two variables with regard to the time lag between variables. Correlation of time lag is found between PM 2.5 levels and number of deceased cases. Using the formula, PM 2.5 as x and deaths at different time lag as y, the correlation are shown in a plot.

3.4.2 Spatial Correlation

Spatial correlation observes the relationship of PM 2.5 and number of deceased cases with regard to the distance from city A to city B. Distance is calculated using the longitude and latitude of the cities and finding the distance between cities in kilometers. Using the formula, PM 2.5 as x and deaths at different distance as y, the correlation are shown in a plot.

3.4.3 Spatiotemporal Correlation

Spatiotemporal correlation shows the relationship between PM 2.5 and number of deceased cases in combination of distance and time lag. Like the spatial and temporal correlation models, the distance is found between two points and a time lag is factored in. Using the formula, PM 2.5 as x and deaths at different time lag and distance as y, the correlation are shown in a plot.

4 RESULTS

4.1 Spatial Autocorrelation

4.1.1 Clustering (Hotspot) of Deceased Number of Cases/Mortality Rate

Before lockdown, there are no clusters observed for deceased cases and mortality as shown by the Moran's I statistic value and p-value in Table 4.1.1. At a state level, it can be stated that there is a random pattern between the states of deceased cases and mortality rate. The state with the highest mortality level shown in Figure 4.1.1 is Karnataka which includes the city Bengaluru.

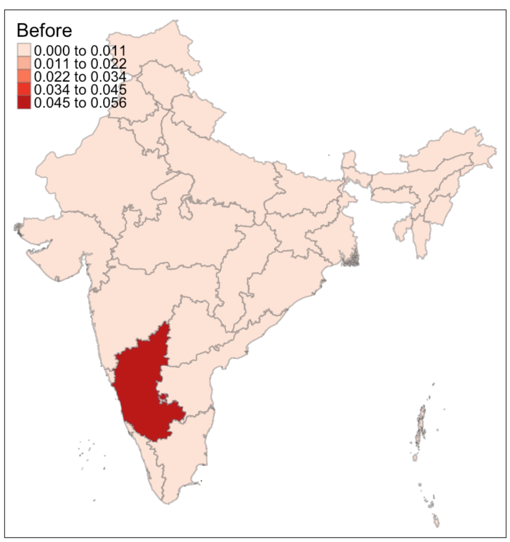


Figure 4.1.1 Mortality Rate by state of India for Before Lockdown

During lockdown, there are also no clusters observed for deceased cases and mortality as shown by the Moran's I statistic value and p-value in Table 4.1.2. At a state level, it can be stated that there is a random pattern between the states of deceased cases and mortality rate. During lockdown, high mortality rates can be seen in western states such as Maharashtra, Gujarat, Madhya Pradesh, and Telangana.

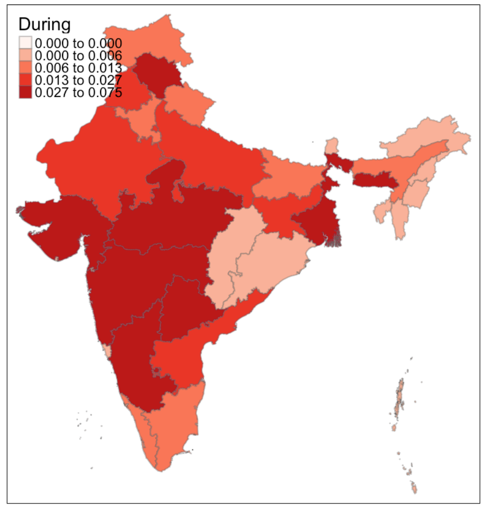


Figure 4.1.2 Mortality Rate by state of India for During Lockdown

After lockdown, there are also no clusters observed for deceased cases and mortality as shown by the Moran's I statistic value and p-value in Table 4.1.3. At a state level, it can be stated that there is a random pattern between the states of deceased cases and mortality rate. After lockdown, northern states which are known for high pollution due to industrial emissions have higher mortality rates.

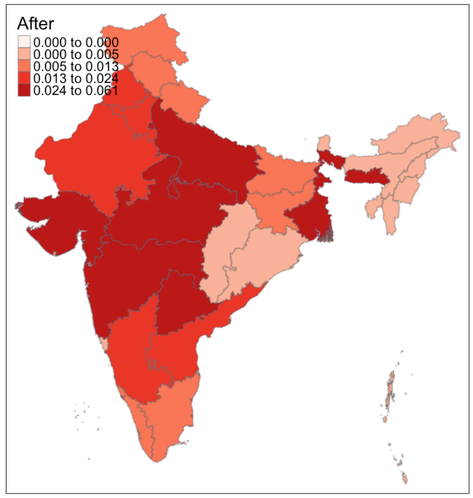


Figure 4.1.3 Mortality Rate by state of India for After Lockdown

Table 4.1.1 Moran's I Statistic Results for Deceased Cases and Mortality Rate

		Moran's I statistic	P-value	Expectation	Variance	Pattern
Before Lockdown (Jan 30 - Mar 24, 2020)	Deceased	-0.0441	0.7716	-0.0286	0.0004	Random
	Mortality Rate	-0.0441	0.7716	-0.0286	0.0004	Random
During (Mar 25 - May 31, 2020)	Deceased	0.0727	0.1213	-0.0286	0.0075	Random
	Mortality Rate	-0.0999	0.7214	-0.0286	0.0148	Random
After (Jun 1 - Jun 30, 2020)	Deceased	-0.0044	0.3756	-0.0286	0.0058	Random
	Mortality Rate	0.0330	0.3081	-0.0286	0.0151	Random

4.1.2 Air Quality improvement during lockdown

Air quality ratio of before, during, and after lockdown periods were calculated using the recorded daily PM 2.5 levels before lockdown can be observed in Figure 4.1.4 which shows that all states with active air quality stations reported high number of days of unhealthy ($<55.5\mu g/m^3$) PM 2.5 levels.

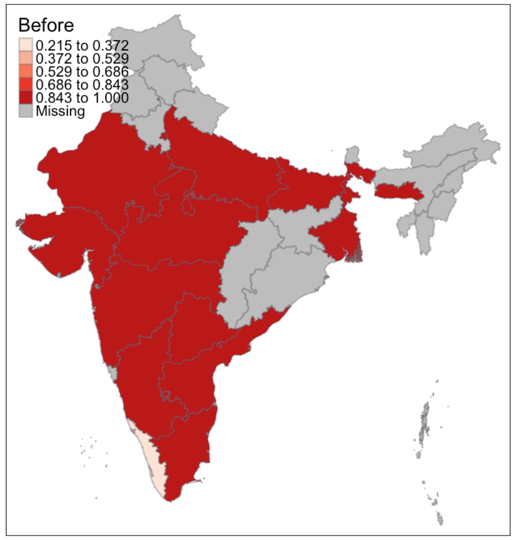


Figure 4.1.4 Ratio of Unhealthy PM2.5 days by state of India for Before Lockdown

In figure 4.1.5, it can be observed that the ratio of the number of unhealthy PM 2.5 levels improved in many states during the lockdown period. It can be stated that there was an improvement in air quality in states like Karnataka, Maharashtra, Tamil Nadu, and Andhra Pradesh.

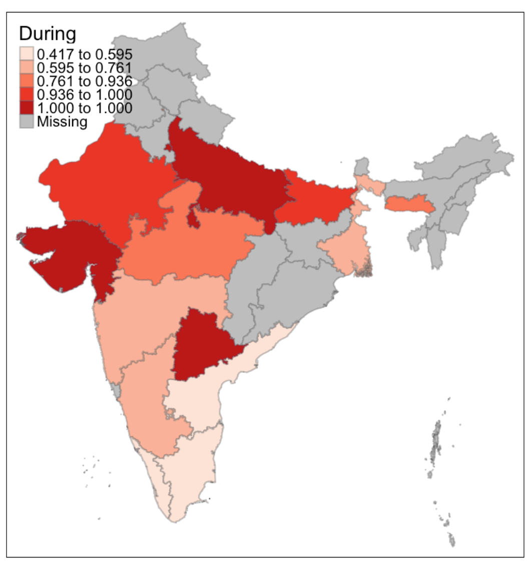


Figure 4.1.5 Ratio of Unhealthy PM2.5 days by state of India for During Lockdown

After the lockdown, the state's PM 2.5 level unhealthy days ratio increases in northern states as industrial and transportation activities resume. The change in ratios can be seen in figure 4.1.6. Rajasthan is an example of a state that had a lower PM2.5 levels during the lockdown, but again increased after the lockdown had ended.

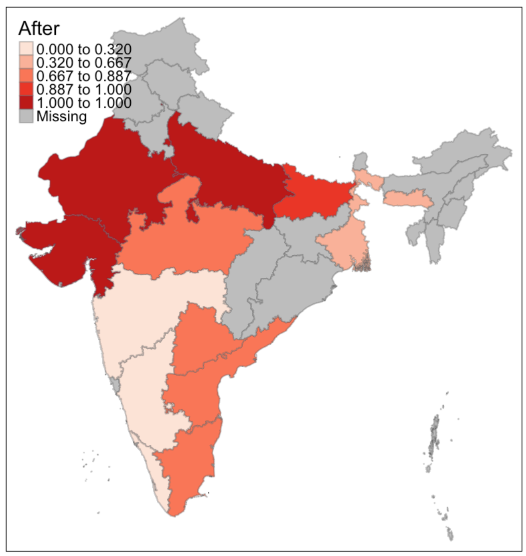
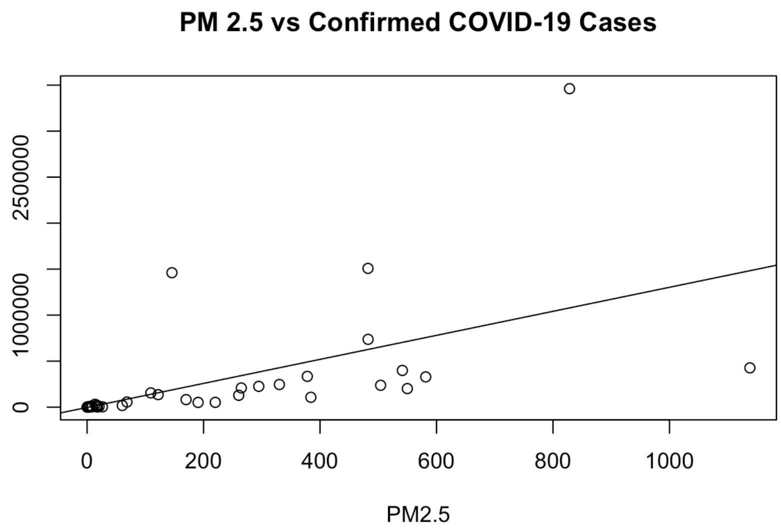


Figure 4.1.6. Ratio of Unhealthy PM2.5 days by state of India for After Lockdown

4.2 Linear Regression of COVID-19 Cases and PM 2.5 Emission Levels

4.2.1 Confirmed cases vs. PM 2.5 Emission

A significant weak positive correlation is shown between the number of confirmed cases and PM 2.5 estimated gross yearly levels can be observed. The R-squared value for the regression is .275 and p-value is .000609. We can significantly say with 95% confidence that there is a positive relationship with PM 2.5 emission levels and confirmed cases.

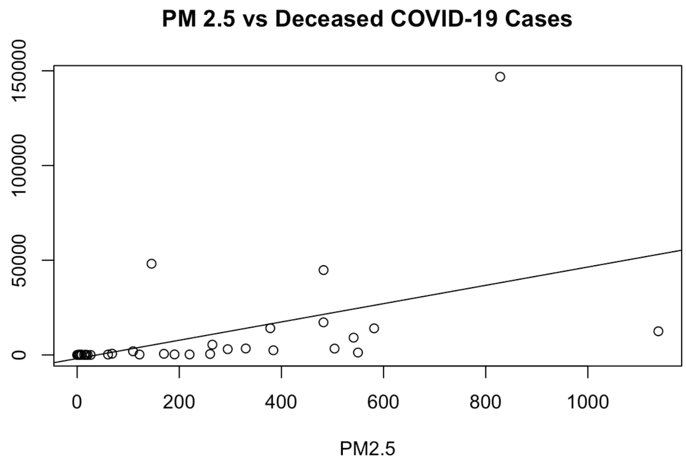


R^2	$p - value$
.275	.000609

Figure 4.2.1 Mortality Rate by state of India for During Lockdown

4.2.2 Deceased cases vs. PM 2.5 Emission

A significant weak positive correlation between the number of deceased cases and PM 2.5 estimated gross yearly levels can be observed. The R-squared value for the regression is .227 and p-value is .00195. We can significantly say with 95% confidence that there is a positive relationship with PM 2.5 emission levels and deceased cases.

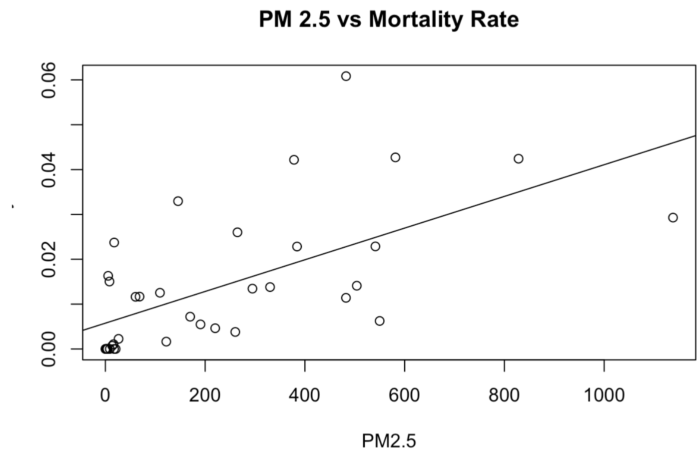


R^2	$p - value$
.227	.00195

Figure 4.2.2 Mortality Rate by state of India for During Lockdown

4.2.3 Mortality Rate vs. PM 2.5 Emission

A significant weak positive correlation between mortality rate and PM 2.5 estimated gross yearly levels can be observed. The R-squared value for the regression is .368 and p-value is .0000519. We can significantly say with 95% confidence that there is a positive relationship with PM 2.5 emission levels and mortality rate. Increased in PM 2.5 emission levels would result in higher mortality rate in states.

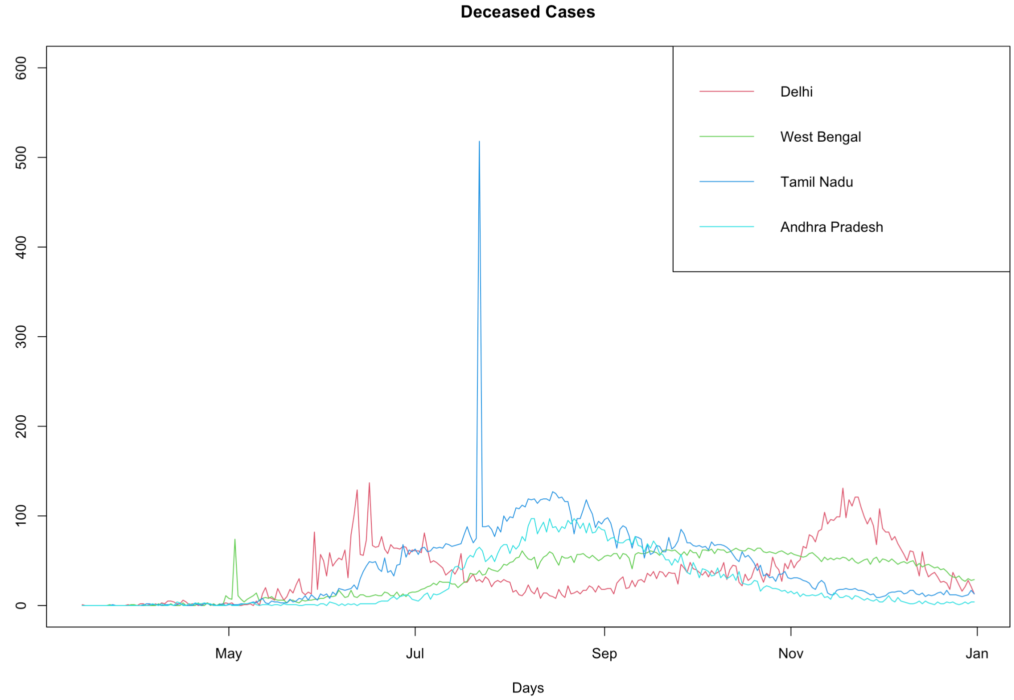


R^2	$p-value$
.368	.0000519

Figure 4.2.3 Mortality Rate by state of India for During Lockdown

4.3 Time Series Cross-correlation

To observe the correlation between PM 2.5 and number of deaths, a total of four cities were selected to explore the cross correlation. Daily number of deceased cases and PM 2.5 levels were plotted, then a cross correlation function was used to determine how the variables are dependent on each other. The four cities time series plot of daily PM 2.5 levels and daily deceased counts can be found on figure 4.3. After lockdown, the number of deceased increases suddenly with steeper peaks.



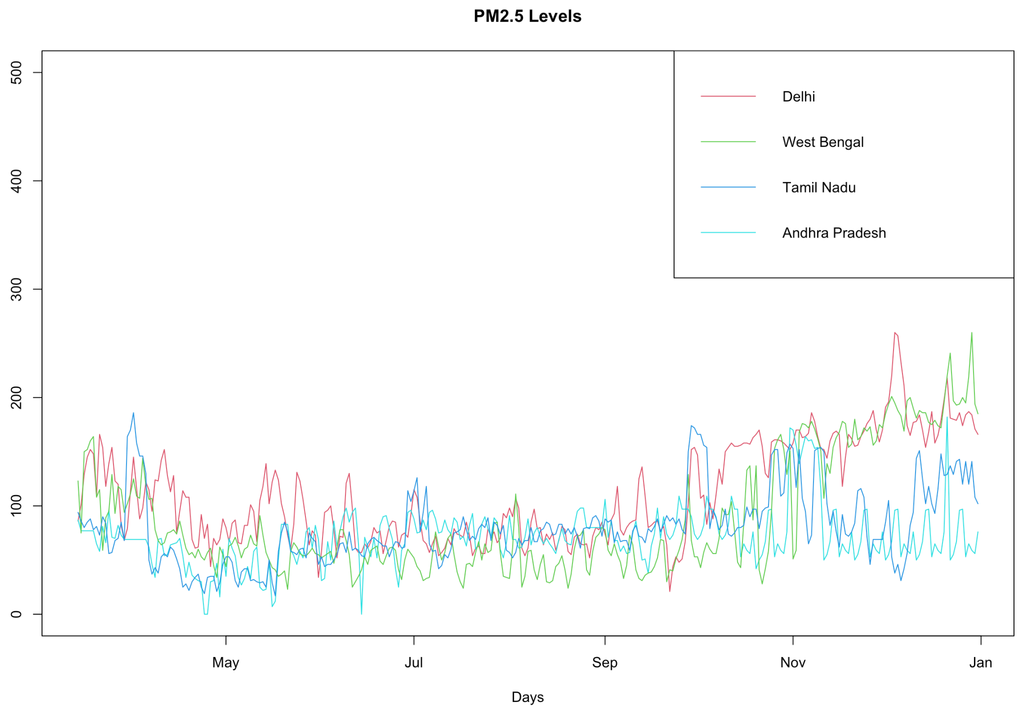


Figure 4.3 Time Series Plot of Daily Deceased Cases and Daily PM2.5 Levels for Four Cities

4.3.1 Cross Correlation for Delhi

Delhi had a peak of deceased cases around June and then again in December 2020. It can be seen in the PM 2.5 graph that during lockdown months, the PM 2.5 level slightly decreased. We can observe from figure 4.3.1 that the most dominant cross correlation occurs around h=20. The positive correlation value indicates that the above value of PM 2.5 has a peak positive relationship with the number of deceased cases 20 days after.

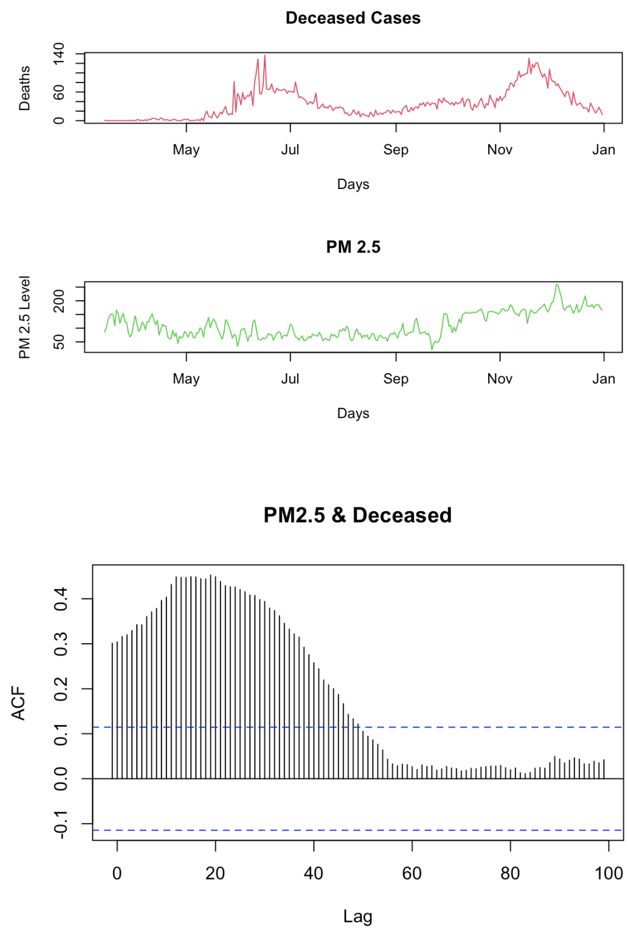


Figure 4.3.1 Correlation between daily Covid-19 Deaths and PM2.5 for Delhi in 2020

4.3.2 Cross Correlation for West Bengal

West Bengal had a sharp peak of deaths in May 2020, and a broad peak from July to December of 2020. PM 2.5 level decreased during lockdown months. Figure 4.3.2 shows a positive correlation between PM 2.4 and daily Covid-19 deaths with a peak correlation at 60 days after.

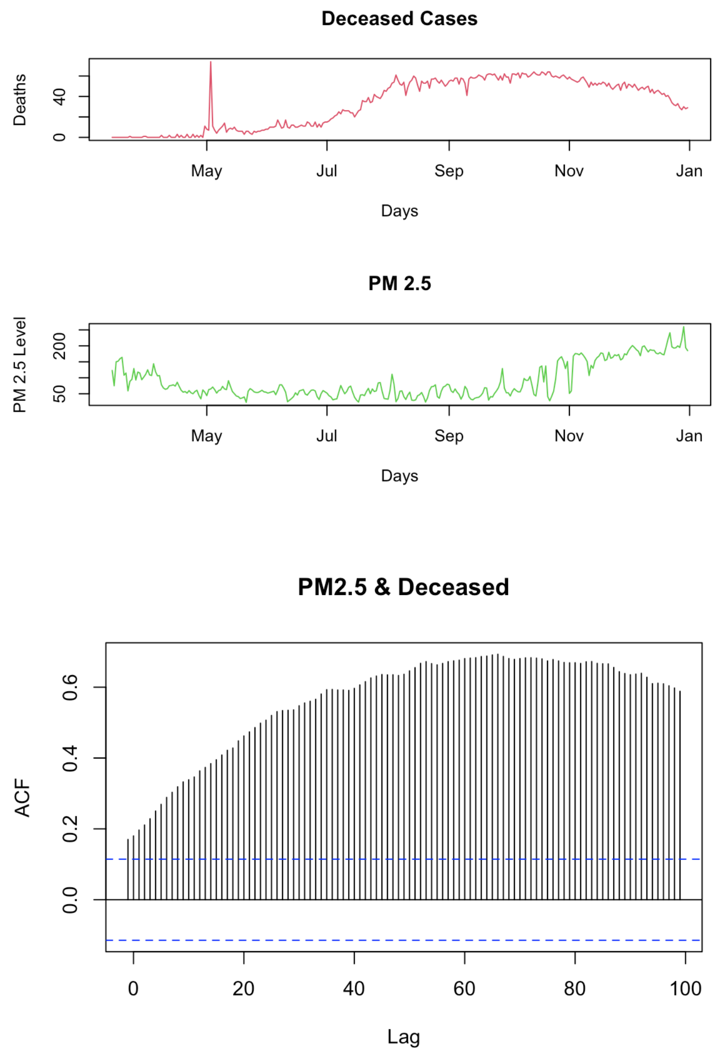


Figure 4.3.2 Correlation between daily Covid-19 Deaths and PM2.5 for West Bengal in 2020

4.3.3 Cross Correlation of Tamil Nadu

Tamil Nadu had a peak of deceased cases around the end of July 2020. It can be seen in the PM 2.5 graph that during lockdown months, the PM 2.5 level slightly decreased. We can observe from figure 4.3.3 that the most dominant cross correlation occurs around h=75. The positive correlation value indicates that above the value of PM 2.5 has a positive relationship with the number of deceased cases 75 days after.

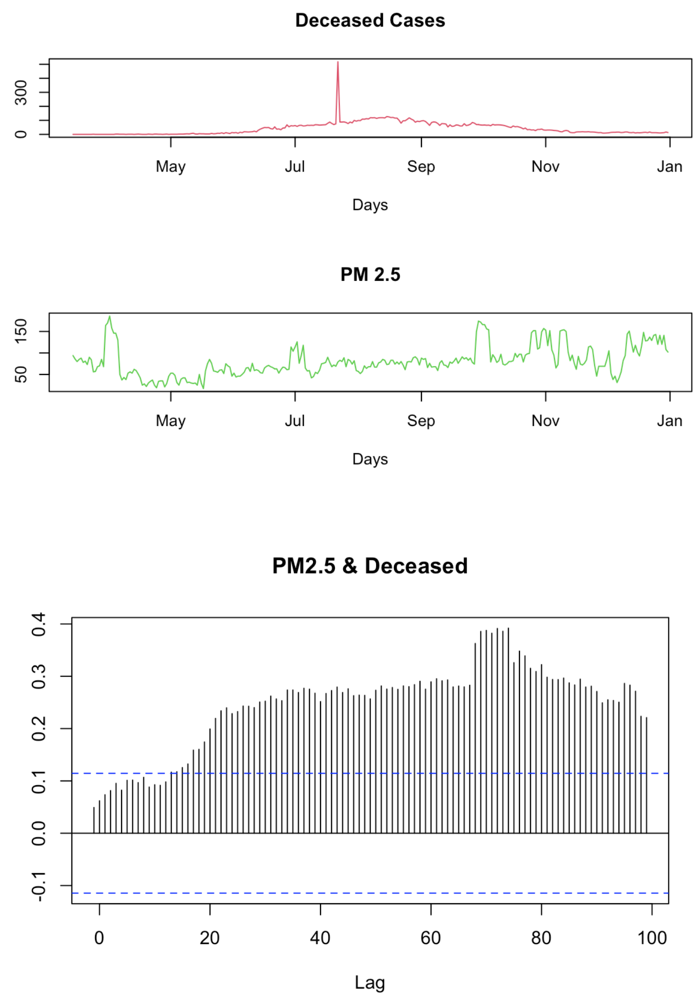
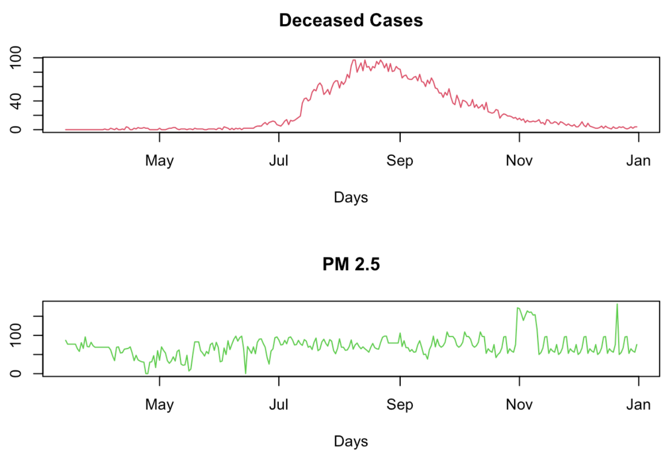


Figure 4.3.3 Correlation between daily Covid-19 Deaths and PM 2.5 for Tamil Nadu in 2020

4.3.4 Cross Correlation of Andhra Pradesh

Andhra Pradesh COVID-19 daily deaths steadily rose until the end of August then decreased afterwards. The PM 2.5 level for the state is relatively lower than the other cities. Figure 4.3.4 shows that the most dominant cross correlation occurs around h=50. The positive correlation value indicates that PM 2.5 has a peak correlation with the number of deceased cases 50 days after.



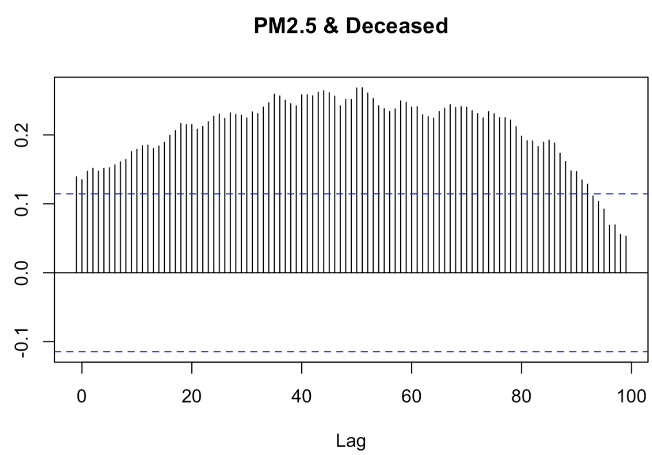


Figure 4.3.4 Correlation between daily Covid-19 Deaths and PM2.5 for Andhra Pradesh in 2020

4.4 Spatiotemporal Analysis

4.4.1 Temporal Correlation

A peak can be seen around 50 days in figure 4.4.1 which signifies that there is a positive relationship between PM2.5 and COVID-19 deceased cases. An elevated PM 2.5 level is correlated to an elevation of COVID-19 deceased cases 50 days later. Another peak around 250 days can be also observed, which may represent a long-term exposure effect of bad air quality on COVID-19 cases. It could also be a numerical artifact as the total length of time analyzed is 292 days.

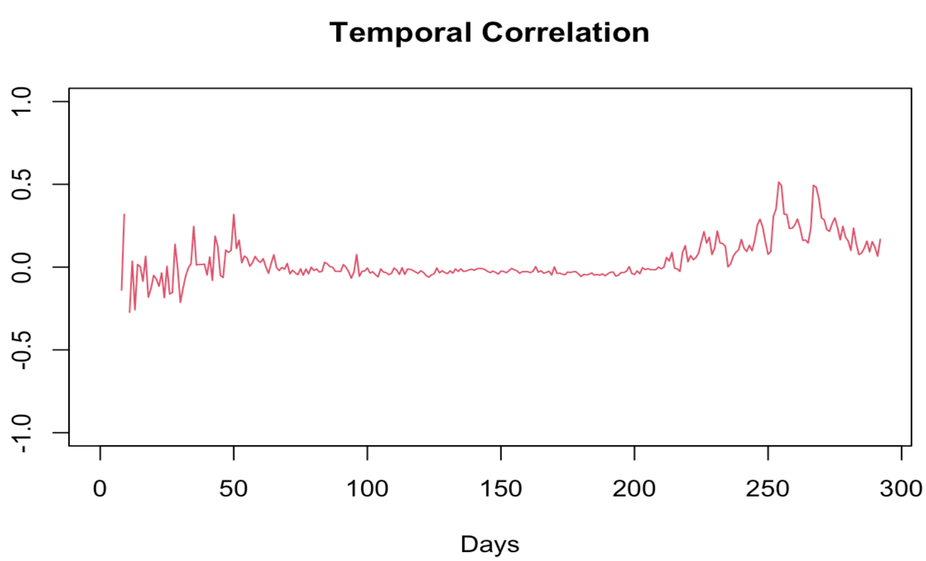


Figure 4.4.1 Temporal Correlation of PM2.5 and Covid-19 Cases in regard to Time Lag

4.4.2 Spatial Correlation

No visual spatial correlation can be seen on figure 4.4.2 which states that there is no relationship between PM 2.5 levels and deceased cases according to distance.

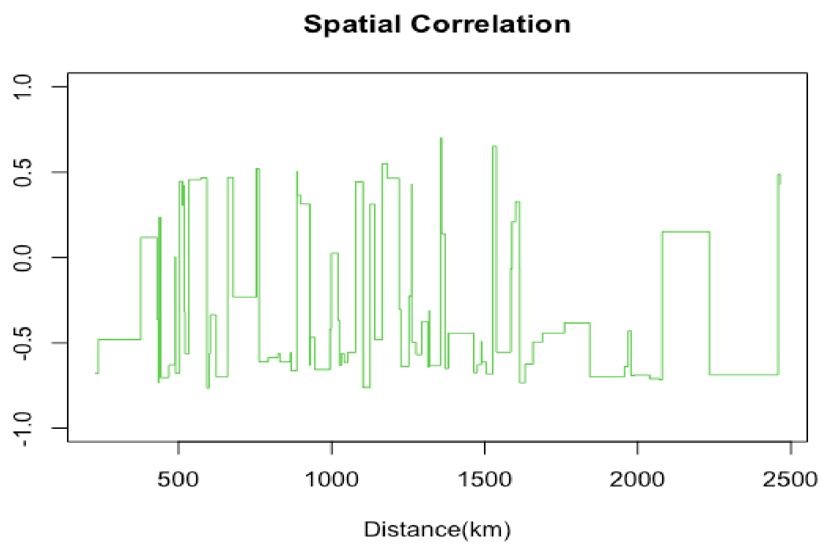


Figure 4.4.2 Spatial Correlation of PM2.5 and Covid-19 Deaths in regard to distance

4.4.3 Spatiotemporal Correlation

Spatiotemporal correlation results show that there is a high correlation between time lag of 15 to 50 days and distance of around 300 kilometers. We can interpret the correlation value that there is a positive relationship between PM 2.5 levels and number of deceased cases after 15-50 days and within 300 kilometers of distance. Figure 4.4.3 shows a heat map of the spatiotemporal correlation. A 3D surface plot and a scatter 3D plot are also shown in figure 4.4.5 and figure 4.4.6 to visualize the spatiotemporal correlation.

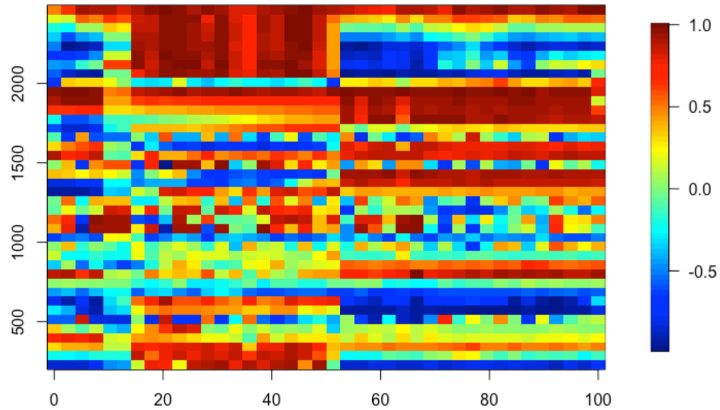


Figure 4.4.3 Heat Map of Correlation of PM 2.5 and Covid-19 Deaths in regard to Time and Distance

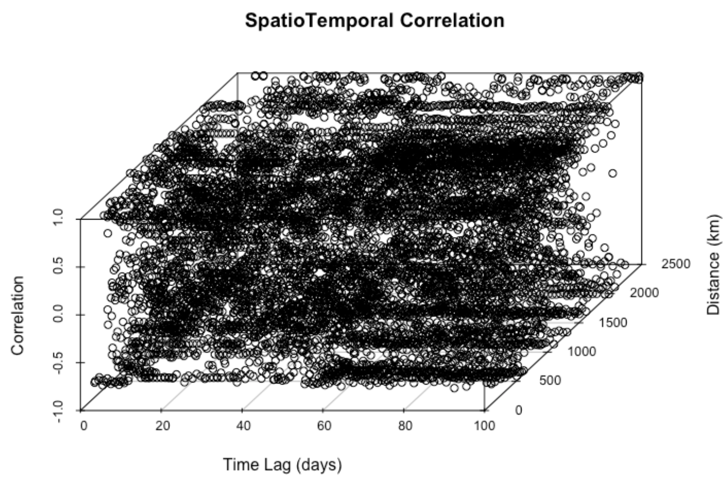


Figure 4.4.5 3D Scatter plot of Correlation of PM 2.5 and Covid-19 Deaths in regard to Time and Distance

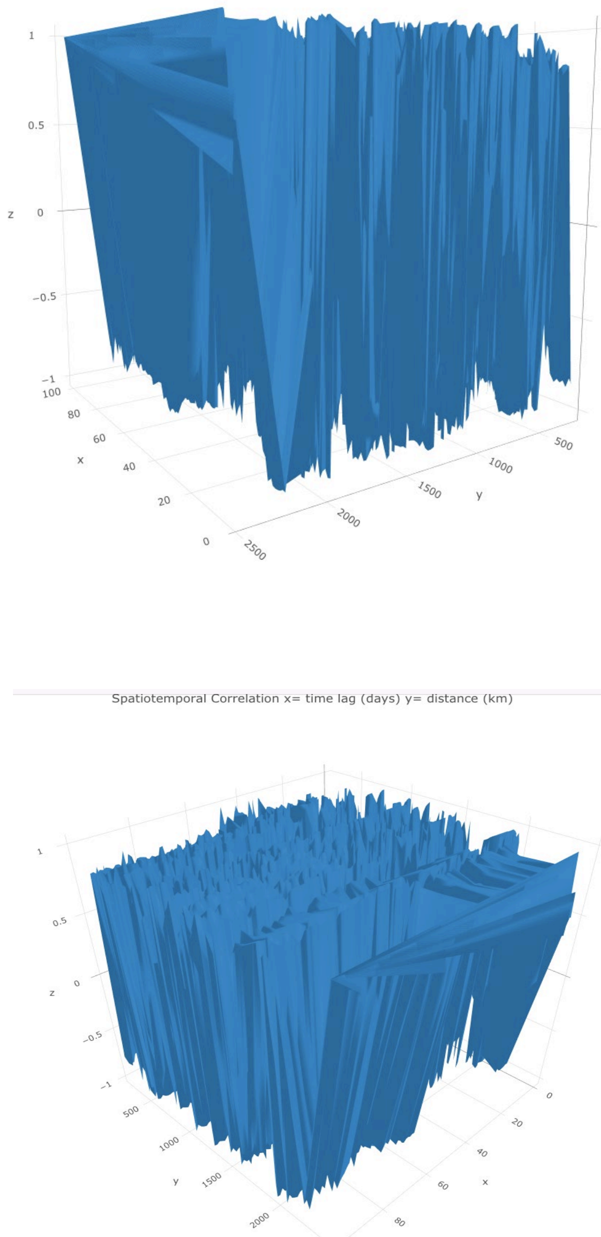


Figure 4.4.6 Surface Plot of Correlation of PM 2.5 and Covid-19 Deaths in regard to Time and Distance

5 CONCLUSION

According to the various methods used, it can be concluded that there is a positive relationship between unhealthy PM 2.5 levels and deceased number of COVID-19 cases. People who are likely to have been exposed to unhealthy PM 2.5 levels are more likely to be susceptible to COVID-19 infection and at a higher chance of mortality. Linear Regression shows a weak positive relationship between state's PM2.5 emission levels and COVID-19 cases. COVID-19's confirmed cases, deceased cases, and mortality rate increases as PM2.5 gross yearly emission levels increases. Hot spot analysis using Moran's I statistic shows us that at state level, there is no cluster formation of COVID-19 mortality rates, COVID-19 number of deceased cases, and PM2.5 unhealthy levels. Cross-correlation through time series shows that there is a positive relationship between PM 2.5 and deceased number of cases 20-60 days after. Spatiotemporal analysis shows there is a high correlation of PM 2.5 and number of deceased cases when time lag is around 30-50 days and within 300 kilometers. This suggests that the effect of an increase in PM 2.5 levels will have an elevated effect on COVID-19 cases after 30 days and will affect the surrounding area up to 300 kilometers.

Future research should be dedicated to exploring not only PM 2.5 levels as an aggravating factor but other variants such as PM 10 and other harmful particles in air pollution. A factor that should be taken into consideration is the short-term and long-term consequences of unhealthy air quality. Short-term consequences include coughing and respiratory distress while long-term exposure would be respiratory diseases and death. This research focuses on the deceased cases but does not take into consideration the other long-term consequences of extended presence in bad air quality.

REFERENCES

Aydin, O., Raja, N.B. Spatial-temporal analysis of precipitation characteristics in Artvin, Turkey. Theor Appl Climatol 142, 729–741 (2020). https://doi.org/10.1007/s00704-020-03346-6

Coronavirus in India: Latest Map and case count. Coronavirus Outbreak in India. (n.d.). Retrieved February 13, 2022, from https://www.covid19india.org/

Chatterjee P. (2019). Indian air pollution: loaded dice. The Lancet. Planetary health, 3(12), e500–e501. https://doi.org/10.1016/S2542-5196(19)30247-5

Du, Xu & Varde, Aparna. (2016). Mining PM2.5 and traffic conditions for air quality. 33-38. 10.1109/IACS.2016.7476082.

S.K. Chaulya, G.M. Prasad, in Sensing and Monitoring Technologies for Mines and Hazardous Areas, 2016

Getis, Arthur, and J. K. Ord. "The Analysis of Spatial Association by Use of Distance Statistics." Geographical Analysis 24, no. 3. 1992.

Ghosh, A., Nundy, S., & Mallick, T. K. (2020). How India is dealing with COVID-19 pandemic. Sensors international, 1, 100021. https://doi.org/10.1016/j.sintl.2020.100021

Lansley, G and Cheshire, J (2016). *An Introduction to Spatial Data Analysis and Visualisation in R.* CDRC Learning Resources. Online: https://data.cdrc.ac.uk/tutorial/an-introduction-to-spatial-data-analysis-and-visualisation-in-r

Liu, S., et al. (2021). Long-term exposure to low-level air pollution and incidence of asthma: the ELAPSE project. The European respiratory journal, 57(6), 2003099. https://doi.org/10.1183/13993003.03099-2020

Raja NB, Aydin O, Türkoğlu N, Çiçek I (2017) Space-time kriging of precipitation variability in Turkey for the period 1976–2010. Theor Appl Climat 129:293–304

S.K. Chaulya, G.M. Prasad, in *Sensing and Monitoring Technologies for Mines and Hazardous Areas*, 2016

Sahu, S. K., Mangaraj, P., Beig, G., Tyagi, B., Tikle, S., & Vinoj, V. (2021). Establishing a link between fine particulate matter (PM2.5) zones and COVID -19 over India based on anthropogenic emission sources and air quality data. Urban climate, 38, 100883. https://doi.org/10.1016/j.uclim.2021.100883

Singh RP, Dey S, Tripathi SN, Tare V, Holben B (2004) Variability of aerosol parameters over Kanpur, northern India. J Geophys Res-Atmos 109(D23)

The World Air Quality Index project. (n.d.). Air pollution in world: Real-time Air Quality Index Visual Map. aqicn.org. Retrieved February 13, 2022, from https://aqicn.org/map/world

Veronesi G, De Matteis S, Calori G, et alLong-term exposure to air pollution and COVID-19 incidence: a prospective study of residents in the city of Varese, Northern ItalyOccupational and Environmental Medicine 2022;79:192-199.