

# The relevance between several air pollutants and the covid spread

Qiaochu Xing

King's College London, London, UK

qiaochuxing@kcl.ac.uk

**Abstract.** COVID-19 has become a global public health problem. This study aimed to explore the relationship between ambient air pollutants and the infection caused by the novel coronavirus. Daily confirmed cases, of air pollution concentration, in 43 cities were obtained from January 23, 2020, to February 20, 2020, in China. I applied a generalized logit model to investigate the associations of six air pollutants ( $PM_{2.5}$ ,  $PM_{10}$ ,  $SO_2$ ,  $CO$ ,  $NO_2$ , and  $O_3$ ) with COVID-19 confirmed cases. After generating the model and checking the p-value, I found that  $PM_{2.5}$ ,  $PM_{10}$ , and  $O_3$  are correlated with the increase in covid cases.

**Keywords:** Air pollutants, Covid, linear regression, p-value, hypothesis testing, correlation.

## 1. Introduction

With the development of society, the industry's growth will produce more greenhouse gas emissions.  $CO_2$ ,  $SO_2$ ,  $NO_2$ .... Some of the harmful gas will cause air pollution and it affects our health a lot. Exposure to air pollution can lead to a wide range of diseases, including stroke, respiratory disease, and high blood pressure, that raise the chances to get COVID-19 infection. The essay is aim to analyze the air quality data for different chemical gas emissions in 43 cities in China, to explore how air pollution connect with the covid spread and what can we do to predict and control the speed of spread of covid

## 2. Literature review

The studies pointed out that air pollutants are risk factors for respiratory infection to make pathogens more invasive to humans and affecting the body's immunity to make people more susceptible to pathogens (Cai et al., 2007; Horne et al., 2018; Xie et al., 2019; Xu

et al. 2016). The study also shows that air pollution with 1  $\mu g/m^3$  more  $PM_{2.5}$  concentrations will have 9.4 more COVID-19 cases, 3 more hospital admissions, and 2.3 more deaths. (Environmental and Resource Economics, August 4, 2020) As WHO reported, people with respiratory problems are more susceptible to COVID-19. The previous paper has suggested that  $NO_2$  concentration was positively associated with the transmission ability of COVID-19. (Yarongen, 2021). The previous report also found annual nitrogen dioxide exposure (a pollutant that comes from tailpipe emissions) to be associated with COVID-19 incidence and mortality in Los Angeles County neighborhoods while adjusting for numerous confounders, with an 8.7 ppb increase in  $NO_2$  to be associated with a 35–60% increase in mortality rate (Environment International, August 2021).

But there are not only  $NO_2$  can do a negative effect on the respiratory system, it still contains other air pollutants such as  $CO$  and  $PM_{2.5}$ ...

To provide useful implications for controlling and preventing this novel coronavirus disease, I will explore the correlation between air quality data and daily confirmed COVID-19 cases in 43 cities in China and detect the coefficient of different pollutants to explain how different air pollutants affect the increased cases of COVID-19.

### 3. Methodology

#### 3.1. Data collection:

This study included 43 cities (4 municipalities and 39 prefecture-level cities) According to the National Health Commission, 74,928 COVID-19 confirmed cases have been identified in the whole of China as of February 20, 2020. Our studied cities covered 60% of confirmed cases. I focused my analysis on these 43 cities because of the limitation of the obtained meteorological data and the air pollution data.

From January 23, 2020, to February 20, 2020, the daily data of newly confirmed cases in each city came from the report released by the official website of the local health committee. I took January 23, 2020 (Wuhan city closure day) as the starting point of the study period, to reduce the inclusion of imported cases from Wuhan as much as possible. Data ends on February 20, 2020, and COVID policy has already started, which is a factor that affects the speed of transmission

Air pollution data were collected from an online platform (<https://www.aqistudy.cn>) and six air pollutants were measured, including particles with diameters  $\leq 2.5 \mu\text{m}$  (PM2.5), particles with diameters  $\leq 10 \mu\text{m}$  (PM10), sulfur dioxide (SO<sub>2</sub>), carbon monoxide (CO), nitrogen dioxide (NO<sub>2</sub>), and ozone (O<sub>3</sub>), and sulfur trioxide(SO<sub>3</sub>).

#### 3.2. Mathematical analysis:

Previous studies have shown that the health effects of air pollution can last for several days, and generalized linear models are an effective way to examine the effects of meteorological factors and air pollution on health outcomes (Lin et al., 2018; Myung et al., 2019; Xie et al., 2019; Yang et al., 2020).

Thus, in this study, we used the GLM with a Normal distribution family to estimate the associations between the moving average concentrations of air pollutants and daily COVID-19 confirmed cases. Specifically, we examined the effects of six air pollutants by generating the linear model and checking the covariance between different air pollutants and daily increases. (Using R studio) The logit model was defined as follows: ( $\log y = a + bx_1 + cx_2 + dx_3 + ex_4 + fx_5 + gx_6 + e$ )

Where  $y$  represents the “daily increase in covid cases”, ‘ $x_1$ ’ represents the data of PM2.5, ‘ $x_2$ ’ represents the data of PM10, ‘ $x_3$ ’ represents the data of SO<sub>2</sub>, ‘ $x_4$ ’ represents the data of NO<sub>2</sub>, ‘ $x_5$ ’ represents the data of O<sub>3</sub>, ‘ $x_6$ ’ represents the data of CO, and the ‘ $e$ ’ represents the random error for the linear model)

First I just took the simple linear model ( and it generates by the method of least square method ( is a standard method in regression analysis to approximate the solution by minimizing the sum of the squares of the residuals) and we can get the result of coefficient of different independent variables to show how different air pollutants affect the daily increase.

Coefficients:

|                 | Estimate  | Std. Error | Pr(> t ) |
|-----------------|-----------|------------|----------|
| (Intercept)     | 4.274881  | 14.936138  | 0.7758   |
| PM2.5           | 0.374754  | 0.317231   | 0.2424   |
| PM10            | -0.221713 | 0.302077   | 0.4660   |
| SO <sub>2</sub> | -0.003093 | 1.824432   | 0.9987   |
| NO <sub>2</sub> | -0.509751 | 0.254710   | 0.0501   |
| O <sub>3</sub>  | 0.171065  | 0.099711   | 0.0917   |
| CO              | 6.240636  | 15.380861  | 0.6865   |

From the individual t-test, the P value is too large for all air pollutants so it didn’t show that the air pollutants are important factors that will increase the rate of covid infection. It didn’t fix our assumption I. So, we need to get a new model.

Then we try to remove some outliers ( the data with large standard error) and regenerate the model. We can use box cox transformation( create a monotonic transformation of data using power functions.

It is a data transformation technique used to stabilize the variability, make the data more normal distribution-like, and improve the validity of measures )

And pick the log of the independent factor and then generate the model as ( $\log y = a + bx_1 + cx_2 + dx_3 + ex_4 + fx_5 + gx_6 + e$ )

#### 4. Result

|                 | (coefficient Estimate) | Std. Error | Pr(> t ) |
|-----------------|------------------------|------------|----------|
| (Intercept)     | 1.319187               | 1.328284   | 0.017*   |
| PM2.5           | 0.24506                | 0.028212   | 0.0024** |
| PM10            | 0.21288                | 0.026864   | 0.00617* |
| SO <sub>2</sub> | 0.001753               | 0.162248   | 0.976    |
| NO <sub>2</sub> | 0.027002               | 0.022652   | 0.23     |
| O <sub>3</sub>  | 0.06602                | 0.008867   | 0.0135*  |
| CO              | 1.306784               | 1.367834   | 0.067    |

This paper uses a generalized linear model to explore the relationship between several air pollutants and daily confirmed cases of COVID-19.

We found that PM<sub>2.5</sub>, PM<sub>10</sub>, and O<sub>3</sub> were significantly positively associated with confirmed cases of COVID-19, Each unit increase of PM<sub>2.5</sub>, PM<sub>10</sub>, and O<sub>3</sub> will cause an increase in daily covid cases while SO<sub>2</sub>, NO<sub>2</sub>, CO seems to have no relationship with the number of daily confirmed cases( individual t-test shows that it is easy to accept the null hypothesis which is the coefficient of SO<sub>2</sub>, NO<sub>2</sub>, CO is 0).

These findings could provide evidence that air pollution is an important factor in COVID-19 infection. What's more, a previous report has shown that exposure to SO<sub>2</sub>, CO, and NO<sub>2</sub> was harmful to our health and increased the risk of respiratory disease (Chen et al.

2007) However, our results are different from previous studies since we observed no correlation between SO<sub>2</sub>, CO, NO<sub>2</sub> and COVID-19 confirmed cases. The virucidal property of SO<sub>2</sub> may be a possible reason (Berendt et al., 1971, Berendt et al., 1972), and additional research is needed to determine whether the SO<sub>2</sub>, NO<sub>2</sub>, CO will affect the result in biography and chemistry aspects.

#### 5. Discussion

Our research has certain guiding significance for the prevention and control of new coronary pneumonia

First, governments and the public should pay more attention to regions with high concentrations of PM<sub>2.5</sub>, PM<sub>10</sub>, and O<sub>3</sub>, since these regions may suffer a more serious COVID-19 epidemic. In other words, reducing air pollutants could be a useful way to control COVID-19 infections. In addition, it is worth noting that SO<sub>2</sub> is inversely associated with confirmed COVID-19 cases, and further laboratory studies are needed to elucidate the underlying mechanism.

In addition, the transmission of respiratory viruses is known to occur through contact (directly or indirectly through contaminants) or through contaminated droplets emitted by coughing, sneezing, breathing, and speaking of an infected person, both of which are associated with interpersonal contact, social distancing, and population movement. What's more, the air pollutants we focus on are indicators of traffic-related air pollutants, and the association between them and daily increase data for COVID-19 can be explained by the relationship between virus transmission and population movements. Further investigation is, of course, necessary to provide more details and explain such mechanisms.

## 6. Further works

Our study has several limitations. First, I just concentrated on the correlation between air pollutants data and COVID-19 confirmed cases and not the causal effect of air pollution on COVID-19 infection. Secondly, the findings were not globally representative since cities of other countries were not included in this study, which may cause the error as China has a great population base. Thirdly, from the statistical aspects, I ignore the collinearity among different air pollutants and it will affect the result while generating the model. Future studies are needed to overcome these limitations.

## Reference list

- [1] Zhu et al (Association between short-term exposure to air pollution and COVID-19 infection: Evidence from China)
- [2] R.F. Berendt, E.L. Dorsey, H.J. Hearn Virucidal properties of light and SO<sub>2</sub> II. Effect of a low gas concentration on aerosolized virus
- [3] Q.-C. Cai, J. Lu, Q.-F. Xu, Q. Guo, D.-Z. Xu, Q.-W. Sun, H. Yang, G.-M. Zhao, Q.-W. Jiang Influence of meteorological factors and air pollution on the outbreak of severe acute respiratory syndrome
- [4] C. Sohrabi, Z. Alsafi, N. O'Neill, M. Khan, A. Kerwan, A. Al-Jabir, C. Iosifidis, R. Agha, World Health Organization declares global emergency: a review of the 2019 novel coronavirus (COVID-19)
- [5] J. Xie, J. Teng, Y. Fan, R. Xie, A. Shen The short-term effects of air pollutants on hospitalizations for respiratory disease in Hefei
- [6] Zhu et al. Association between short-term exposure to air pollution and COVID-19 infection: Evidence from China
- [7] Health Effects Institute, State of global air 2019: A special report on global exposure to air pollution and its disease burden
- [8] Ciencewicki, I. Jaspers, Air pollution and respiratory viral infection.