

# Positive correlation between long term emission of several air pollutants and COVID-19 deaths in Sweden

Lars Helander\*

December 5, 2020

## Abstract

Several recent studies have found troubling links between air pollution and both incidence and mortality of COVID-19, the pandemic disease caused by the virus SARS-CoV-2. Here, we investigate whether such a link can be found also in Sweden, a country with low population density and a relatively good air quality in general, with low background levels of important pollutants such as PM<sub>2.5</sub> and NO<sub>2</sub>. The investigation is carried out by relating normalized emission levels of several air pollutants to normalized COVID-19 deaths at the municipality level, after applying a sieve function using an empirically determined threshold value to filter out noise. We find a fairly strong correlation for PM<sub>2.5</sub>, PM<sub>10</sub> and SO<sub>2</sub>, and a moderate one for NO<sub>x</sub>. We find no correlation neither for CO, nor (as expected) for CO<sub>2</sub>. Our results are statistically significant and the calculations are simple and easily verifiable. Since the study considers only emission levels of air pollutants and not measurements of air quality, climatic and meteorological factors (such as average wind speeds) can trivially be ruled out as confounders. Finally, we also show that although there are small positive correlations between population density and COVID-19 deaths in the studied municipalities (which are for the most part rural and non densely populated) they are either weak or not statistically significant.

---

\*The author can be contacted at [lahe@outlook.com](mailto:lahe@outlook.com).

# 1 Introduction

Air pollution is a global health concern and numerous studies have found links between air pollution and both respiratory and chronic diseases [1, 2, 3, 4]. Since the start of the COVID-19 pandemic, several studies have found links between air pollution and COVID-19 incidence and mortality [5, 6, 7, 8, 9]. The present study aims at determining whether there is a link between air pollution and COVID-19 deaths at the municipality level in Sweden. We propose that Sweden is a suitable candidate for this kind of ecological study, due to its long tradition of high quality statistics coupled with the fact that the COVID-19 pandemic has not been met here with severe governmental mitigation measures such as lockdown or school closures or other coercive measures such as obligations to wear face masks. Also, regional responses to the pandemic have not varied much throughout the country, thus encouraging the disease to run its natural course and reach all corners of the country.

The method of the present study, which will be detailed in Section 2, consists in a first step of relating normalized emission levels to normalized COVID-19 deaths at the municipality level. Normalization refers here to relating the figures to the number of inhabitants of each municipality. A positive correlation thus means that municipalities with relatively high emissions and few inhabitants will account for relatively more COVID-19 deaths than those with lower emissions or more inhabitants. In Sweden, municipalities with high emissions and few inhabitants are traditionally called “bruksorter” (mill towns), and are centered around one or several industries (paper mill, steel mill, factory, mine, etc.) playing a major role in the lives of the inhabitants. The industries are required by law to report yearly emission levels of hazardous substances to a public register.

In a second step, we introduce a threshold value  $t_q$  for each pollutant  $q$ . Municipalities with normalized emissions of  $q$  below the threshold value  $t_q$  are excluded from the calculation of the correlation coefficient. The rationale for using a threshold is to remove noise, as COVID-19 death figures in municipalities with low emissions are likely to be driven by other factors than those emissions. Municipalities with high emissions and many inhabitants (such as big cities) may thus fall under the threshold value  $t_q$  even though the emissions very well may have an impact on COVID-19 mortality. However, excluding those cases should not lower the relevance of the study’s results.

## 2 Method

Emission [10], census [11] and population density [11] data at the municipality level were retrieved directly from Statistics Sweden (SCB), the Swedish government statistics agency with a history dating back to before 1749.

COVID-19 deaths registered up to November 30, 2020 [12] at the municipality level were retrieved directly from the Swedish National Board of Health and Welfare (Socialstyrelsen), another government agency. No data for municipalities with 1-3 COVID-19 deaths is given in [12] due to privacy concerns (in this case an 'X' is given in the table instead). For those municipalities, we assume for simplicity that there are 2 COVID-19 deaths.

Now, let  $q$  denote the pollutant under study, with  $q \in \{\text{PM}_{2.5}, \text{PM}_{10}, \text{SO}_2, \text{NO}_x, \text{CO}, \text{CO}_2\}$ . Let  $M$  denote the set of municipalities in Sweden ( $|M| = 290$ ). For each  $m \in M$ , let  $a_m^q$  denote the average emission of  $q$  in  $m$  for the years 2008 to 2018 (inclusive), let  $p_m$  denote the number of inhabitants of  $m$  for the year 2019 and let  $c_m$  denote the number of COVID-19 deaths in  $m$  up to November 30, 2020.

The normalized emissions  $e_m^q$  of  $q$  in  $m$  is calculated according to:

$$e_m^q = a_m^q \cdot 10000/p_m \quad (1)$$

The normalized COVID-19 deaths  $d_m$  in  $m$  is calculated according to:

$$d_m = c_m \cdot 100000/p_m \quad (2)$$

The following threshold values  $t_q$  for each type of pollutant has been identified empirically using the available data. (No correlation has been discernible for CO or CO<sub>2</sub> for any threshold):

<i>Pollutant</i> ( $q$ )	<i>Threshold</i> ( $t_q$ )	<i>Unit</i>
PM2.5	60	tonnes per 10000 inhabitants
PM10	100	tonnes per 10000 inhabitants
SO <sub>2</sub>	70	tonnes per 10000 inhabitants
NO <sub>x</sub>	220	tonnes per 10000 inhabitants
CO	0	tonnes per 10000 inhabitants
CO <sub>2</sub>	0	1000 tonnes per 10000 inhabitants

Table 1: Table of empirically determined threshold values for each pollutant.

We now introduce the sets  $S_q$  per pollutant, representing our samples, with the following definition (using the sieve function  $e_m^q > t_q$  to filter out noise):

$$S_q = \{(d_m, e_m^q) \mid m \in M \wedge e_m^q > t_q\} \quad (3)$$

For each pollutant  $q$  we compute Pearson’s correlation coefficient  $r$  and  $p$ -value of our sample  $S_q$  and a 95% confidence interval in the standard way using [13].

To determine any possible relationship between population density and COVID-19 deaths, we let  $\rho_m$  denote the number of inhabitants per square kilometer in the municipality  $m$  and we define the set  $P$  and sets  $P_q$  for each pollutant  $q$  as follows:

$$P = \{(d_m, \rho_m) \mid m \in M\} \quad (4)$$

$$P_q = \{(d_m, \rho_m) \mid m \in M \wedge e_m^q > t_q\} \quad (5)$$

As a comparison, we also define the set  $P_{t_{\rho 200}}$  below, which corresponds to the set of municipalities with a population density above 200 inhabitants per square kilometer. The threshold of 200 has been empirically determined, in the same way as the thresholds for the pollutants.

$$P_{t_{\rho 200}} = \{(d_m, \rho_m) \mid m \in M \wedge \rho_m > 200\} \quad (6)$$

We then compute Pearson’s correlation coefficient and 95% confidence interval in the same way as before for all these sets (using [13]).

### 3 Results

Below in Table 2 follows the main results of our study. For scatter plots, please refer to Appendix A.

<i>Pollutant</i>	<i># Samples</i>	<i>Pearson</i>	<i>95% confidence interval</i>
PM2.5	32	$r = 0.57, p = 0.0006$	[0.28, 0.77]
PM10	26	$r = 0.60, p = 0.0013$	[0.27, 0.80]
SO <sub>2</sub>	27	$r = 0.59, p = 0.0011$	[0.28, 0.79]
NO <sub>x</sub>	63	$r = 0.40, p = 0.0012$	[0.17, 0.59]
CO	290	<i>No correlation</i> ( $r = -0.11$ )	–
CO <sub>2</sub>	290	<i>No correlation</i> ( $r = 0.00$ )	–

Table 2: Pearson’s correlation coefficient and 95% confidence interval for the set  $S_q$  for each pollutant  $q$ , representing the relationship between emission levels of  $q$  and COVID-19 deaths, as defined in Section 2.

In Table 3 below, we present the correlations between population density and COVID-19 deaths in the different sets of municipalities.

<i>Population density</i>	<i># Samples</i>	<i>Pearson</i>	<i>95% confidence interval</i>
$P_{\text{PM2.5}}$	32	$r = 0.24, p = 0.18$	[−0.12, 0.55]
$P_{\text{PM10}}$	26	$r = 0.22, p = 0.28$	[−0.18, 0.56]
$P_{\text{SO}_2}$	27	$r = 0.10, p = 0.63$	[−0.29, 0.46]
$P_{\text{NO}_x}$	63	$r = 0.28, p = 0.02$	[0.04, 0.50]
$P$	290	$r = 0.22, p = 0.0001$	[0.11, 0.33]
$P_{t_{\rho 200}}$	28	$r = 0.46, p = 0.01$	[0.12, 0.70]

Table 3: Pearson’s correlation coefficient and 95% confidence interval for the set  $P$  and sets  $P_q$  for each pollutant  $q$ , representing the relationship between population density and COVID-19 deaths, as defined in Section 2.  $P_{t_{\rho 200}}$  corresponds to the set of municipalities with a population density above 200 inhabitants per square kilometer.

## 4 Conclusions

After applying a sieve function to filter our noise, we have proved statistically significant correlations between the air pollutants PM<sub>2.5</sub>, PM<sub>10</sub>, SO<sub>2</sub> and NO<sub>x</sub> and COVID-19 deaths at the municipality level in Sweden. The correlations for the three first substances are fairly strong, and for the last the correlation is more moderate. Our approach is new in the sense that we consider only emission levels of pollutants and not measurements of air quality, which means that we can easily exclude climate and weather (such as wind speeds) as confounding factors. We have also shown that the correlations between population density and COVID-19 deaths in the studied municipalities are either weak or not statistically significant, which means that the observed correlations with air pollution are not due to population density. When considering the country as a whole, there is a weak, but positive, correlation between population density and COVID-19 deaths. If we restrict ourselves to the municipalities with a population density above 200 inhabitants per square kilometer, the correlation becomes more pronounced.

**Limitations:** A limitation to this ecological study is that we cannot quantify to what extent toxic substances emitted in a municipality are actually inhaled by its inhabitants and thus to what extent they play a relevant role in the observed correlations. As a future improvement, we would like to incorporate meteorological factors at the municipality level, in particular average wind velocities, into our calculations as a means to modelize the speed of dispersal of the polluting substances. Also, we would like to remind the reader that we have only proved a correlation between data in statistical tables, not any causal relationship between air pollution and COVID-19 deaths. Nevertheless, we hope that this novel approach can be of value to the scientific community and to policymakers, as another small piece in the puzzle of the relationship between air pollution and disease burden in our societies.

## References

- [1] Kim D, Chen Z, Zhou LF, Huang SX. *Air pollutants and early origins of respiratory diseases.*, Chronic Dis Transl Med. 2018 Jun 7;4(2):75-94. doi: 10.1016/j.cdtm.2018.03.003. PMID: 29988883; PMCID: PMC6033955.
- [2] Romieu I, Samet JM, Smith KR, Bruce N. *Outdoor air pollution and acute respiratory infections among children in developing countries.* J Occup Environ Med. 2002 Jul;44(7):640-9. doi: 10.1097/00043764-200207000-00010. PMID: 12134528.
- [3] Chen H, Goldberg MS. *The effects of outdoor air pollution on chronic illnesses.*, McGill J Med. 2009 Jan;12(1):58-64. PMID: 19753290; PMCID: PMC2687917.
- [4] Chau, TT., Wang, KY. *An association between air pollution and daily most frequently visits of eighteen outpatient diseases in an industrial city.* Sci Rep 10, 2321 (2020). <https://doi.org/10.1038/s41598-020-58721-0>
- [5] Wu X, Nethery RC, Sabath BM, Braun D, Dominici F. *Exposure to air pollution and COVID-19 mortality in the United States: A nationwide cross-sectional study.* medRxiv [Preprint]., 2020 Apr 7:2020.04.05.20054502. doi: 10.1101/2020.04.05.20054502. Update in: Sci Adv. 2020 Nov 4;6(45): PMID: 32511651; PMCID: PMC7277007.
- [6] Coker, E.S., Cavalli, L., Fabrizi, E. et al. *The Effects of Air Pollution on COVID-19 Related Mortality in Northern Italy.* Environ Resource Econ 76, 611634 (2020). <https://doi.org/10.1007/s10640-020-00486-1>
- [7] Cole MA, Ozgen C, Strobl E. *Air Pollution Exposure and Covid-19 in Dutch Municipalities.* Environ Resour Econ (Dordr). 2020 Aug 4:1-30. doi: 10.1007/s10640-020-00491-4. Epub ahead of print. PMID: 32836849; PMCID: PMC7399597.
- [8] Andrea Pozzer, Francesca Dominici, Andy Haines, Christian Witt, Thomas Mnzl, Jos Lelieveld. *Regional and global contributions of air pollution to risk of death from COVID-19*, Cardiovascular Research, Volume 116, Issue 14, 1 December 2020, Pages 22472253,

- [9] Bo Pieter, Johannes Andree. *Incidence of COVID-19 and Connections with Air Pollution Exposure: Evidence from the Netherlands*, medRxiv 2020.04.27.20081562; doi: <https://doi.org/10.1101/2020.04.27.20081562>
- [10] SCB, *Air emissions by region (municipality, LAU2) and subject. Year 2008 - 2018*, [http://www.statistikdatabasen.scb.se/pxweb/en/ssd/START\\_\\_MI\\_\\_MI1301\\_\\_MI1301B/UtslappKommun/](http://www.statistikdatabasen.scb.se/pxweb/en/ssd/START__MI__MI1301__MI1301B/UtslappKommun/).
- [11] SCB, *Population density per sq. km, population and land area by region and sex. Year 1991 - 2019*, [https://www.statistikdatabasen.scb.se/pxweb/en/ssd/START\\_\\_BE\\_\\_BE0101\\_\\_BE0101C/BefArealTathetKon/](https://www.statistikdatabasen.scb.se/pxweb/en/ssd/START__BE__BE0101__BE0101C/BefArealTathetKon/).
- [12] Socialstyrelsen, *Statistics on number of COVID-19 deaths*, <https://www.socialstyrelsen.se/globalassets/1-globalt/covid-19-statistik/statistik-over-antal-avlidna-i-covid-19/statistik-covid19-avlidna.xlsx>
- [13] Zhiya Zuo, *Calculate Pearson Correlation Confidence Interval in Python*, <https://zhiyzuo.github.io/Pearson-Correlation-CI-in-Python/> March 31, 2018.



## A Scatter Plots

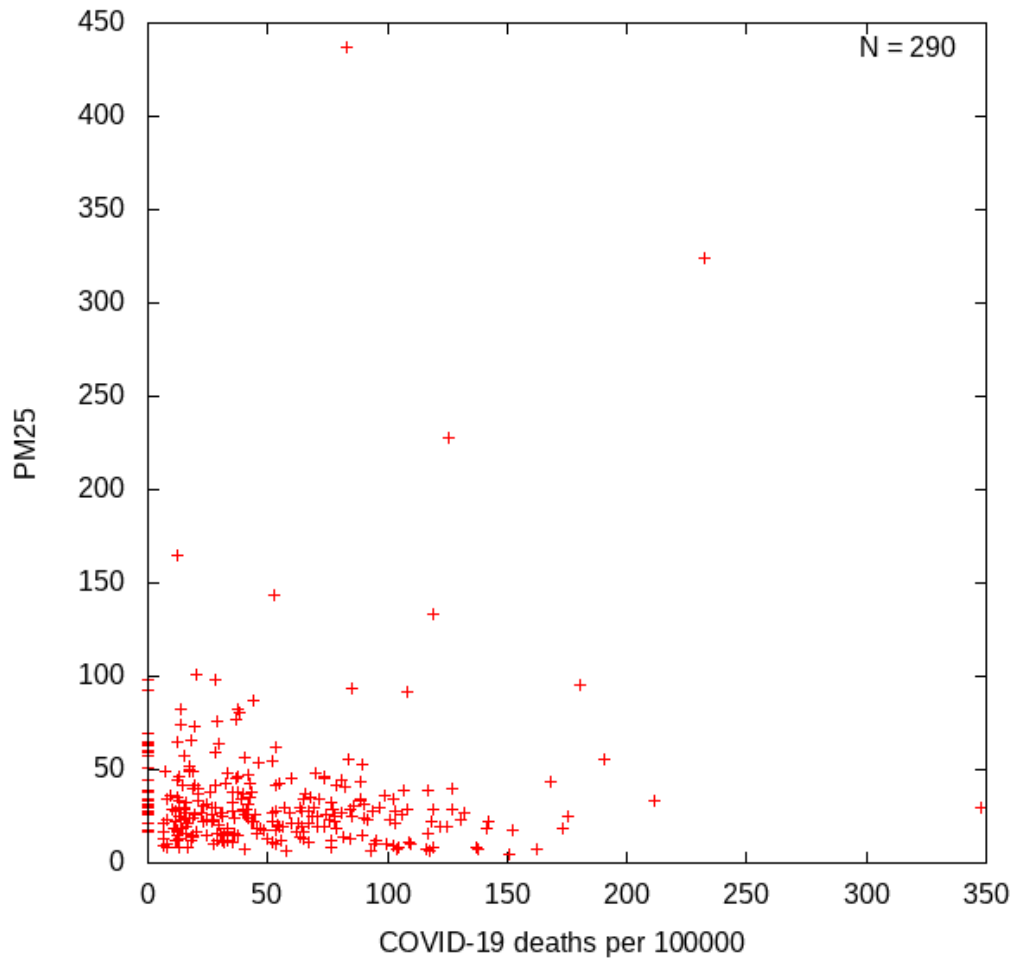


Figure 1: Emissions of PM2.5 (tonnes per 10000 inhabitants) related to COVID-19 deaths. One point per municipality.

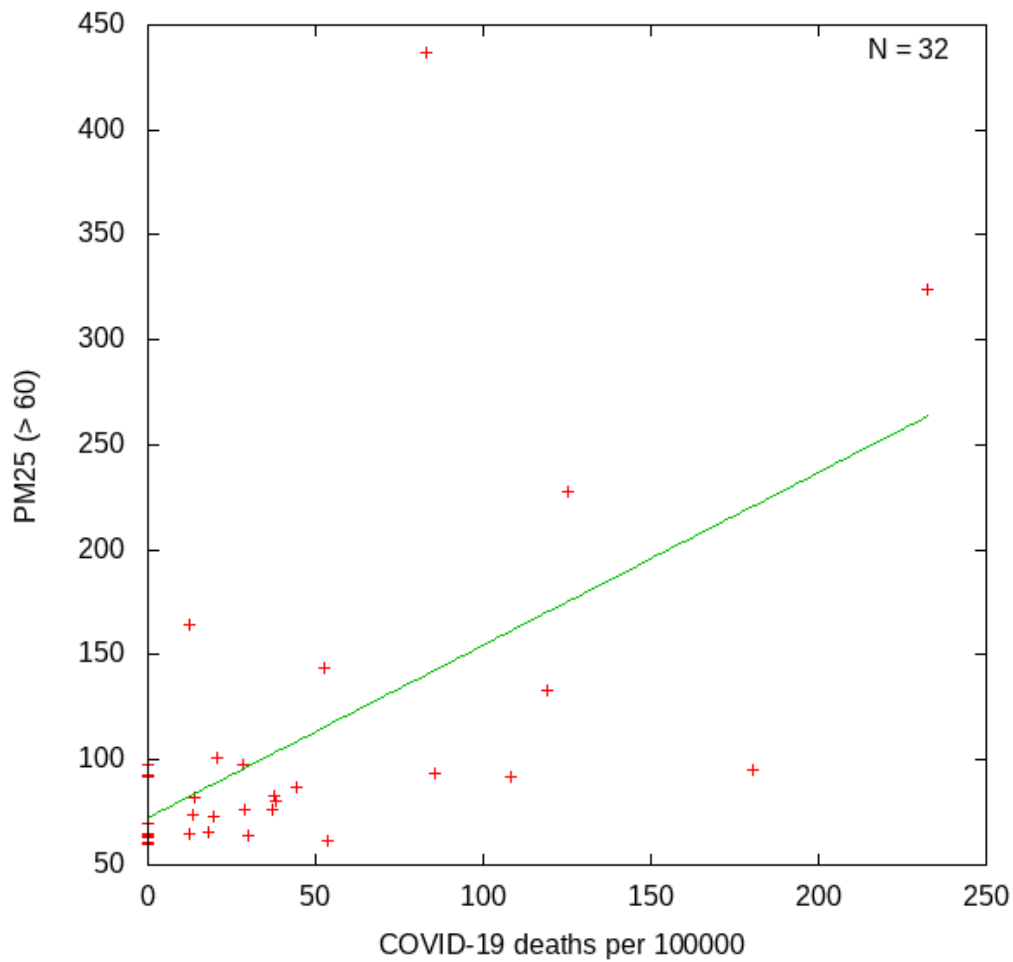


Figure 2: Emissions of PM2.5 (tonnes per 10000 inhabitants) related to COVID-19 deaths. One point per municipality with emissions  $> t_{PM2.5}$ . The included municipalities are Arjeplog, Askersund, Bengtsfors, Dorotea, Gotland, Grums, Gällivare, Härjedalen, Kalix, Karlshamn, Kiruna, Kramfors, Kristinehamn, Lindesberg, Lysekil, Mönsterås, Norsjö, Oxelösund, Piteå, Ragunda, Robertsfors, Rättvik, Sorsele, Storuman, Tjörn, Vilhelmina, Vindeln, Ydre, Åsele, Älvkarleby, Örnsköldsvik, Överkalix.

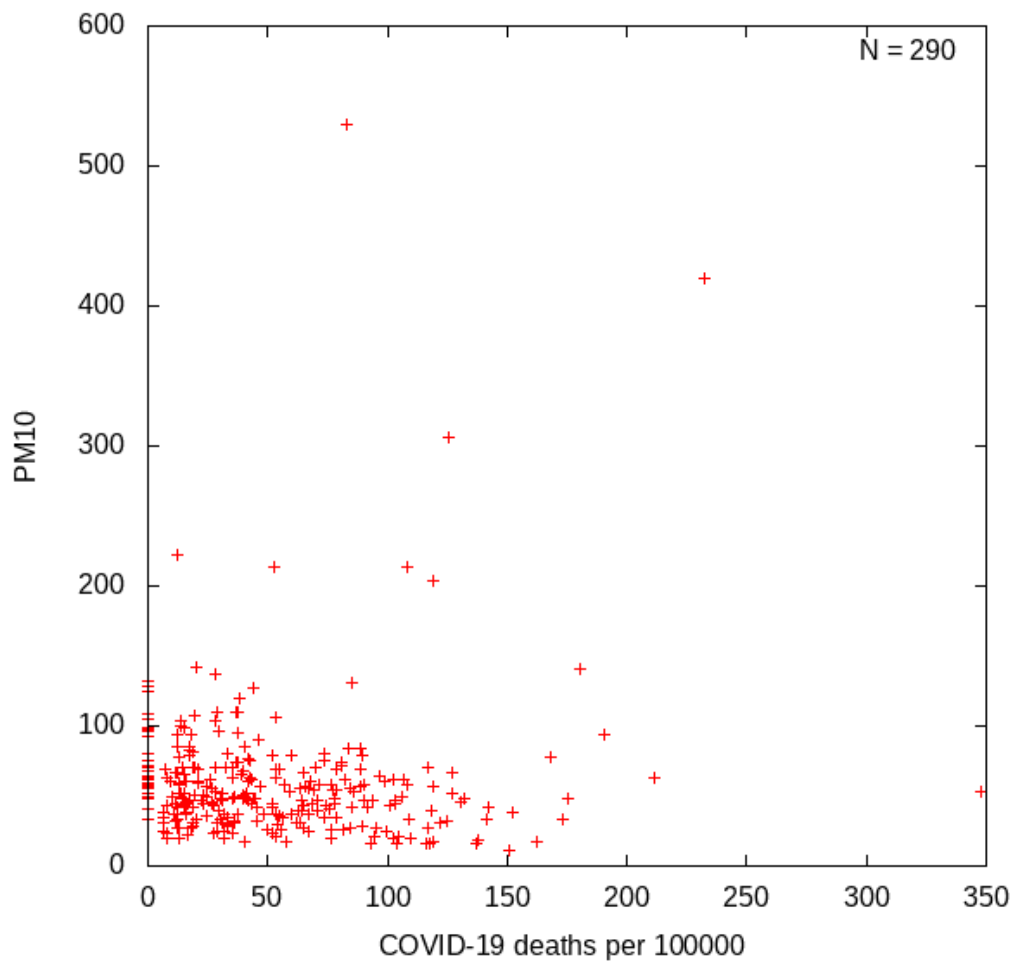


Figure 3: Emissions of PM10 (tonnes per 10000 inhabitants) related to COVID-19 deaths. One point per municipality.

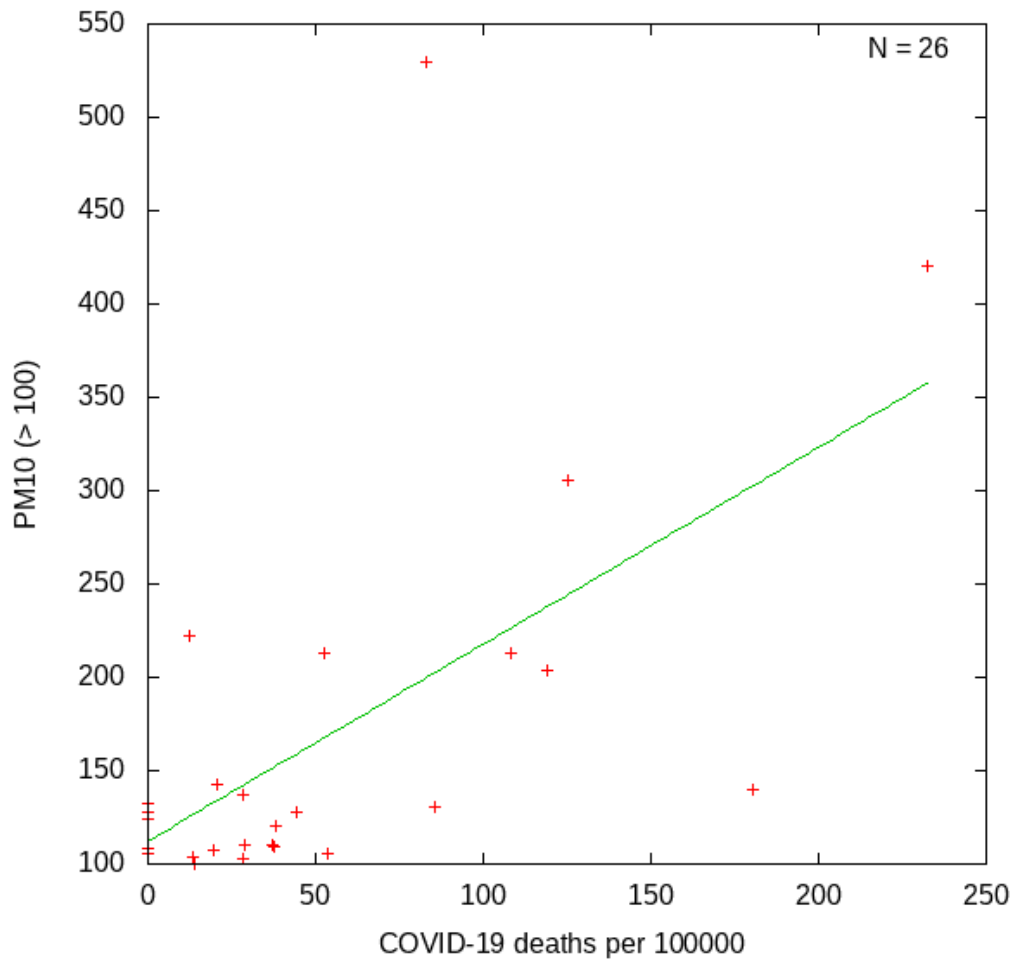


Figure 4: Emissions of PM10 (tonnes per 10000 inhabitants) related to COVID-19 deaths. One point per municipality with emissions  $> t_{PM10}$ . The included municipalities are Askersund, Bengtsfors, Berg, Dorotea, Gotland, Grums, Gällivare, Kalix, Kiruna, Kramfors, Kristinehamn, Lindesberg, Lysekil, Mönsterås, Norsjö, Oxelösund, Piteå, Ragunda, Sorsele, Storuman, Tjörn, Vilhelmina, Vindeln, Ydre, Älvkarleby, Örnsköldsvik.

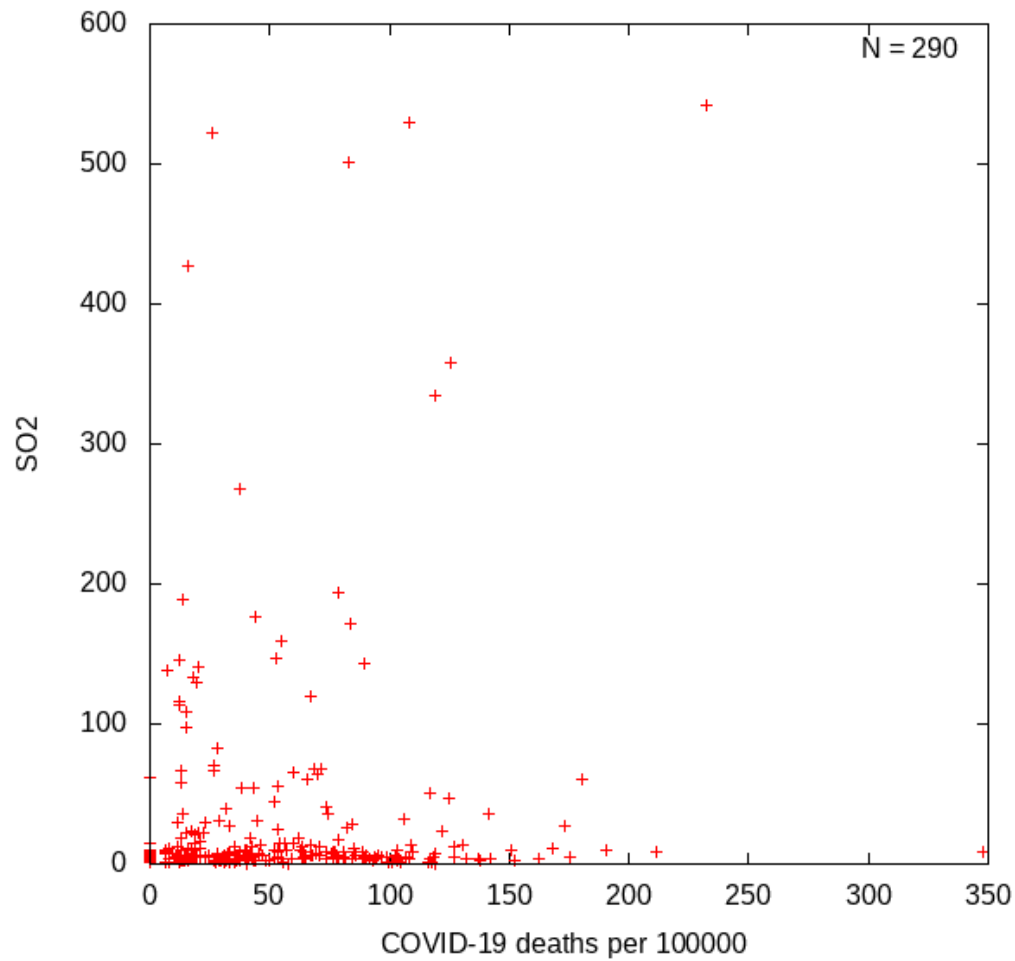


Figure 5: Emissions of SO<sub>2</sub> (tonnes per 10000 inhabitants) related to COVID-19 deaths. One point per municipality.

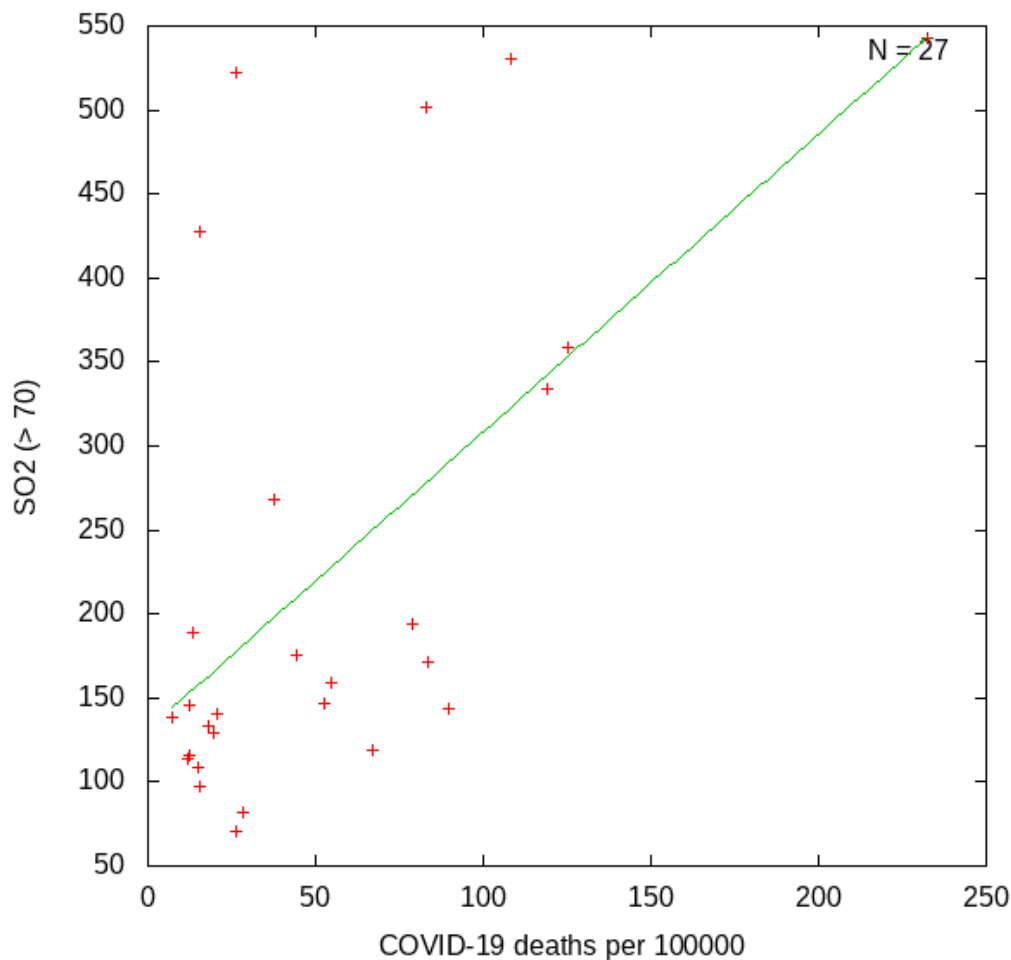


Figure 6: Emissions of  $\text{SO}_2$  (tonnes per 10000 inhabitants) related to COVID-19 deaths. One point per municipality with emissions  $> t_{\text{SO}_2}$ . The included municipalities are Askersund, Bengtsfors, Bromölla, Gotland, Grums, Gällivare, Göteborg, Götene, Hammarö, Kalix, Karlshamn, Kiruna, Lidköping, Luleå, Mönsterås, Oxelösund, Piteå, Rättvik, Skellefteå, Stenungsund, Söderhamn, Timrå, Tjörn, Tranemo, Trelleborg, Älvkarleby, Örnsköldsvik.

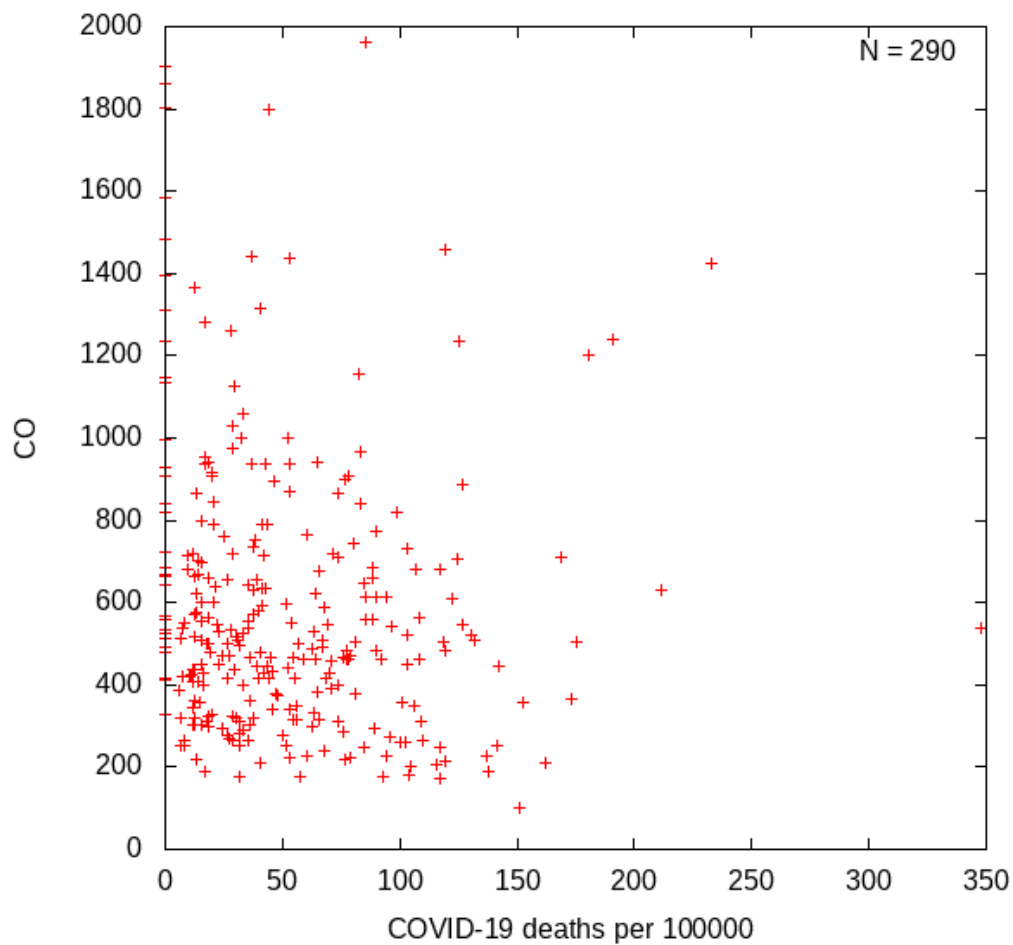


Figure 7: Emissions of CO (tonnes per 10000 inhabitants) related to COVID-19 deaths. One point per municipality. *No correlation found.*

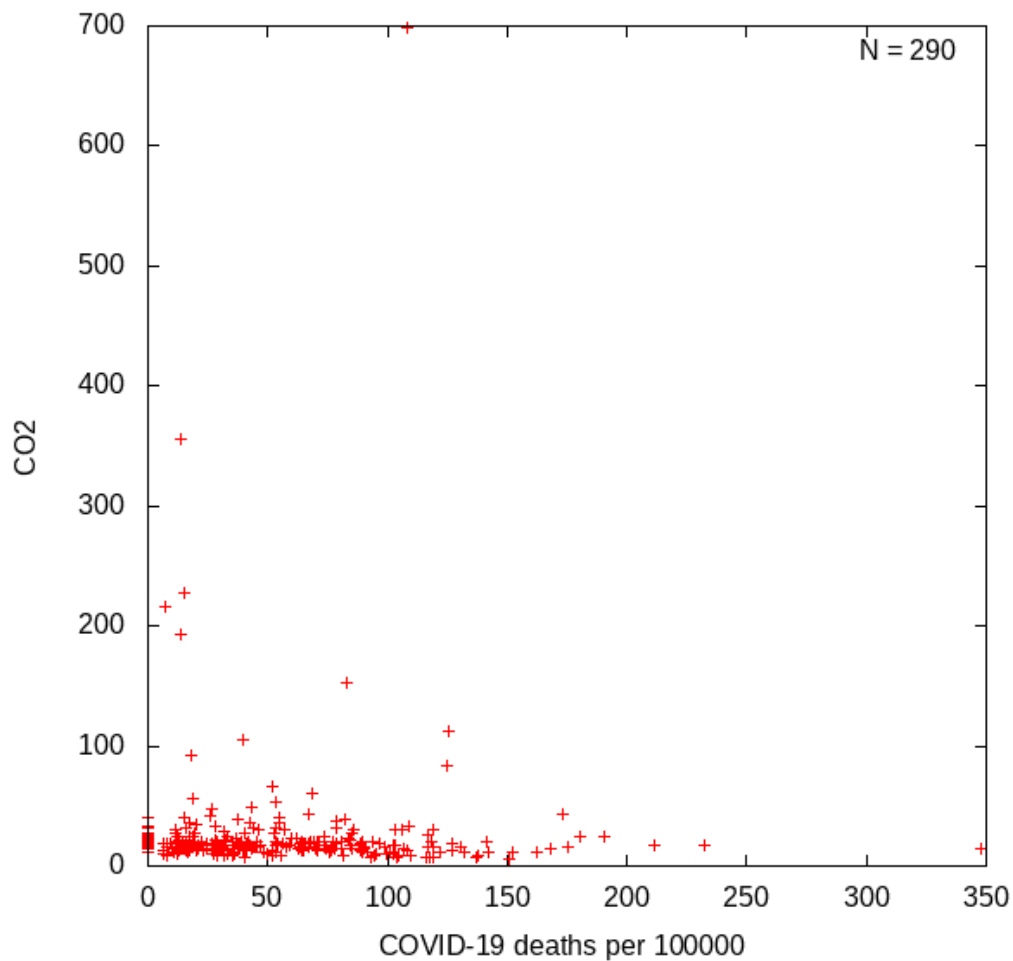


Figure 8: Emissions of CO<sub>2</sub> (1000 tonnes per 10000 inhabitants) related to COVID-19 deaths. One point per municipality. *No correlation found.*



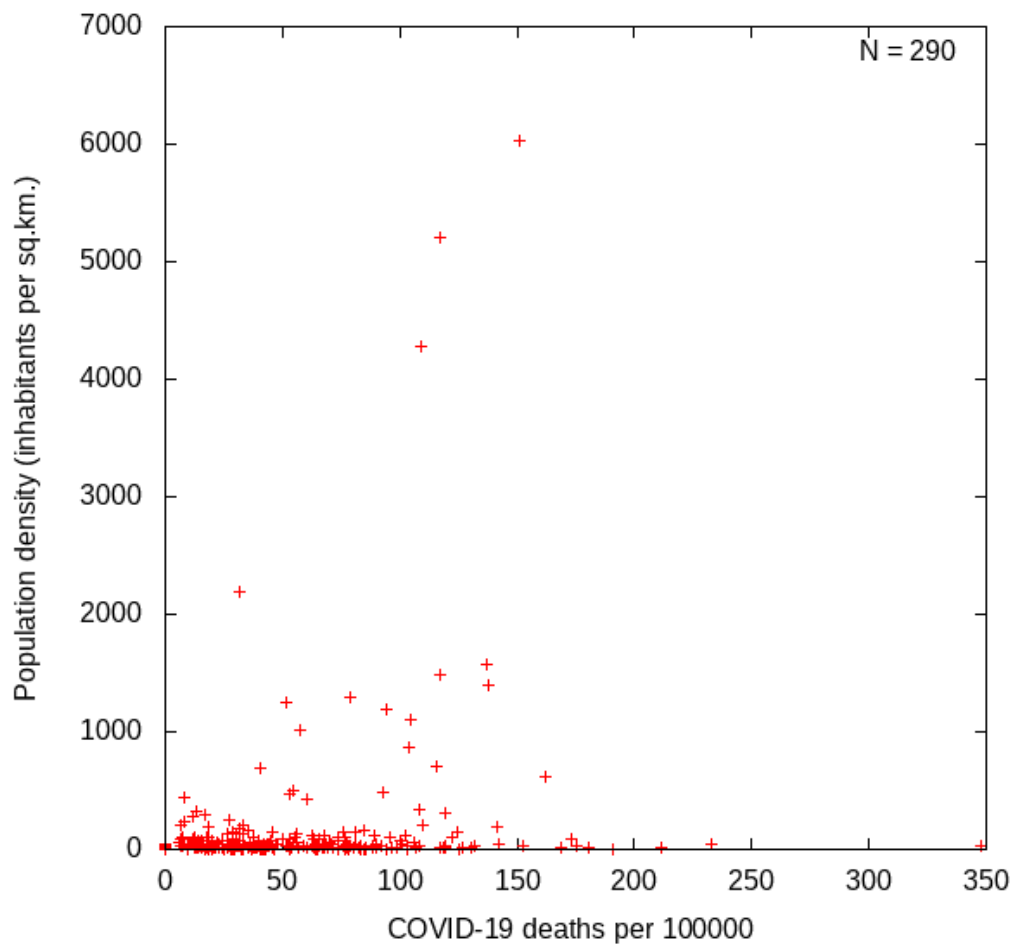


Figure 9: Population density (inhabitants per square kilometer) related to COVID-19 deaths. One point per municipality.