Assessing the impact of long-term exposure to nine outdoor air pollutants on spatial spread of Covid-19: evidence from 107 Italian provinces

Gaetano Perone (gaetano.perone@unibg.it)
University of Bergamo: Universita degli Studi di Bergamo https://orcid.org/0000-0002-0614-6727

Research Article

Keywords: air quality, COVID-19 transmission, environmental pollution, heavy metals, Italy, negative binomial regression, OLS, particulate matter, spatial analysis

Posted Date: September 21st, 2021

DOI: https://doi.org/10.21203/rs.3.rs-851961/v1

Abstract

The coronavirus (COVID-19) pandemic has dramatically changed every aspect of people's lives around the world over the past year and a half. Although the global vaccination campaign is progressing worldwide, new variants of COVID-19 have emerged, driving many countries into a fourth wave of COVID-19 contagion. This paper investigates the air quality in 107 Italian provinces in the period 2014–2019 and the association between long-term exposure to nine outdoor air pollutants and the prevalence of COVID-19 in the same areas. The methods used were negative binomial (NB) regressions, ordinary least squares (OLS) models, and spatial autoregressive models with autoregressive disturbances (SARAR). The air pollutants examined were common air pollutants (NO_2, O_3, $PM_{2.5}$, PM_{10}), polycyclic aromatic hydrocarbons (PAHs) (benzene and BaP), and heavy metals (As, Cd, and Ni). The results showed that i) common air pollutants were generally highly and positively correlated with density of large firms, energy and gas consumption, public transport, and the livestock sector; and ii) long-term exposure to NO_2, $PM_{2.5}$, PM_{10}, benzene, BaP, and Cd was positively and significantly correlated with the spread of COVID-19, even after controlling for cofactors and spatial effects. This outcome seems of interest and relevance because PAHs and heavy metals have not been considered at all in recent literature. It also seems to suggest the need for a national strategy to drive down air pollutant concentrations in order to cope better with possible future pandemics.

1. Introduction

The coronavirus disease of 2019 (COVID-19) is a severe acute respiratory syndrome that officially emerged for the first time in Wuhan, a city in the Hubei province of China, in December 2019. From the end of February 2020, the virus was rapidly spreading across the globe, dramatically changing every aspect of people's lives. As of 1 August 2021, the COVID-19 pandemic had affected 220 countries and territories, with almost 200 million confirmed cases and more than 4.2 million deaths (Worldometer 2021). At the time of writing, the virus has been mutating by generating new forms or variants of itself—the most important of which were first identified in the UK, South Africa, Brazil, and India (WHO 2021)—making the fight against the outbreak even more difficult. In fact, many countries which are approaching the third or even fourth wave of infections have had to reintroduce or extend their lockdowns and social distancing measures. The worst-hit countries include both advanced and developing ones, such as Argentina, Brazil, Colombia, France, India, Italy, Russia, Turkey, the UK, and the US.

In these circumstances, it has become crucial to identify the optimal containment and mitigation policies in order to prevent and manage the spread of the outbreak and prepare a plan to tackle the risk of future epidemics and pandemics. In the last year, a closer look has been taken at the potential adverse impact of air pollution on the spread dynamic and death toll of COVID-19. In fact, it is widely recognized that several air pollutants, such as benzo[a]pyrene (BaP), nitrogen dioxide (NO_2), ozone (O_3), particulate matter (PM), and sulfur dioxide (SO_2), can cause irritation, inflammation, and serious infections and diseases to the lungs and airways (WHO 2016; EEA 2019; Schraufnagel et al. 2019). This is a matter of great concern, considering that according to the latest EEA report (2020a, pp. 40, 42), the annual emissions of $PM_{2.5}$ and

PM_{10} in 2018 exceeded the limits set by the World Health Organization (WHO 2006) Air Quality Consultant (AQG) at 70% and 53% of the stations spread across European countries respectively.

In particular, the relationship between air pollution exposure and the spread of COVID-19 revealed that poor air quality may have favored COVID-19 transmissibility around the world (Cole et al. 2020; Hendryx and Luo 2020; Zhang et al. 2020).

This study may be of interest for two main reasons. First, as of 1 August 2021, Italy is one of the most affected countries worldwide, with 4,355,343 confirmed cases, that is, about 7.2% of the whole resident population. Second, although the literature has already established a positive and significant relationship between the two phenomena in Italy (Bontempi 2020; Coccia 2020; Comunian et al. 2020; Fattorini and Regoli 2020; Filippini et al. 2020; Lolli et al. 2020; Zoran et al. 2020; Collivignarelli et al. 2021; De Angelis et al. 2021), these studies may have suffered from some limitations: i) they mainly focused on a number of regions and provinces and referred to the early phase of the outbreak; ii) in many cases, they focused on the impact of short-term exposure to common air pollutants—NO_2, O_3, PM, and SO_2—on COVID-19 infections and deaths; iii) they did not consider other potentially dangerous air pollutants, such as polycyclic aromatic hydrocarbons (PAHs) and heavy metals; iv) they did not consider other important cofactors (except for De Angelis et al. 2021), such as demographic characteristics, weather conditions, population habits and structure, industrial centers, and hospital bed saturation; and v) finally, they did not explicitly consider the spatial dependency of COVID-19 infections, that is, the possibility that neighboring territories may have affected each other through the movement of people.

In this study, I try to partially fill this gap by jointly considering all these aspects. Thus, the goals of this study are the following: i) I investigate the general air quality in the Italian provinces in the period 2014–2019, trying to assess the main sources of outdoor air pollution and identifying the most polluted territories in the country; and ii) I use negative binomial (NB) regression models, an ordinary least squares (OLS) econometric approach, and spatial autoregressive models with autoregressive disturbances (SARAR) to assess the relationship between long-term exposure to nine air pollutants—NO_2, O_3, $PM_{2.5}$, PM_{10}, benzene, BaP, arsenic (As), cadmium (Cd), and nickel (Ni)—and cumulative COVID-19 cases and prevalence rates at the second peak of the outbreak.

The rest of the paper is organized as follows. Section 2 discusses the air quality in the Italian provinces; Sect. 3 discusses the related literature; Sect. 4 presents the data used in the empirical analysis; Sect. 5 discusses the empirical strategy; Sect. 6 presents and discusses the results; and finally, Sect. 7 provides some concluding considerations.

2. Environmental Pollution In The Italian Provinces

In this section, the main sources of nine air pollutants and the general quality of air in the 107 Italian provinces are investigated. As is well established by the European Environment Agency (EEA 2020b), industry processes, road transport, agricultural activities, waste management, energy production and distribution (especially from fossil sources), natural phenomena (i.e., volcanic eruptions, sandstorms, etc.),

public buildings, and households are the main causes of outdoor air pollution. For instance, exhaust emissions from vehicles and the abrasion of pneumatics and brakes can release benzene, Cd, carbon monoxide (CO_2), lead (Pb), mercury (Me), NO_2, $PM_{2.5}$, PM_{10}, and sulfur oxides (NO_x and SO_x) into the atmosphere (EEA 2016; De Donno et al. 2018), and favor chemical reactions that increase the likelihood of O_3 formation. Business activities, livestock buildings, and households are the major factors responsible for production of $PM_{2.5}$ (Lovarelli et al. 2020). Industrial activities burning fuels (coal, petroleum, wood, etc.), components of smoke cigarettes, forest fires, and vehicle exhaust emissions are the main causes of benzene and BaP (WHO 2010; EPA 2016).

Thus, in Table 1, I report the Pearson's correlation coefficient between the nine investigated air pollutants and six potential sources of environmental pollution in the period 2014–2019:[1] big firms with over 250 employees per square kilometer in the period 2014–2018; final consumption of energy and natural gas expressed as tons of oil equivalent per square kilometer in the period 2014–2019; number of vehicles used to transport goods and passengers (cars, motorcycles, and other vehicles) per square kilometer in the period 2015–2019; overall supply of local public transport (in the province's capital city) expressed as number of seats per inhabitants in the period 2014–2018; the production of cattle fodder from permanent grassland expressed as quintal per square kilometer in the period 2015–2019; and the number of livestock (bovines, buffalos, and pigs) per square kilometer in the period 2015–2019.[2] The results show that common air pollutants are positively and significantly correlated with big firms, energy and gas consumption, density of vehicles, public transport, cattle fodder, and livestock density. The highest correlation coefficients were exhibited by big firms, energy and gas consumption, and livestock density. Notably, NO_2, $O_{3\ (>120)}$, $O_{3\ (>180)}$, $PM_{2.5}$, and $PM_{10\ (>50)}$ showed correlation coefficients ranging from 0.66 to 0.75 for livestock density. This may have been partially caused by the ammonia (NH_3) generated in the urine and feces of cattle (Webb et al. 2005; Laubach et al. 2013), which contributes to the formation of two relevant (secondary) components of particulate matter, ammonium nitrate and ammonium sulphate (McCubbin et al. 2002). In fact, according to Greenpeace and the Italian Institute for Environmental Protection and Research (ISPRA) (Greenpeace 2020), animal husbandry was the second leading cause of air pollution in Italy in the period 1990–2018, accounting for 17% of all $PM_{2.5}$ formation.

Among PAHs, benzene is positively correlated with big firms, energy and gas consumption, vehicle density, and public transport at a 1% level of statistical significance. BaP is positively associated with cattle fodder production and livestock density at 5% and 1% levels of significance respectively. Generally, heavy metals were not significantly correlated with any of the sources of pollution considered.[3]

This is particularly worrying because according to the Air Quality Standards established by the European Commission, the legal threshold for key air pollutants was violated multiple times by most of the Italian provinces in the period 2014–2019 (Table 2).[4] Specifically, almost all provinces (106 out of 107) violated the PM_{10} limit of 50 $\mu g/m^3$ both in the short and long term, resulting in a national average of 25.15 violations per year. Notably, the legal thresholds for both measures of O_3 were also violated several times both in the short and long term, with a maximum of 95 provinces involved. Regarding the average concentrations of NO_2, $PM_{2.5}$, and PM_{10}, the violations were relatively fewer, respectively involving 15, 17,

and five provinces in the short term and 11, four, and no provinces in the long term. Among the PAHs, the legal limit for BaP was violated by 13 provinces in the short term and seven in the long term, while the legal threshold for benzene was never exceeded. No provinces registered violations for heavy metals, except Aosta and Terni, which exceeded the legal limit of Ni in the short term.

The situation becomes even worse when we consider the most restrictive legal thresholds set by the World Health Organization (WHO 2015). In this case, the legal thresholds for $PM_{2.5}$ and PM_{10} were violated respectively by 88 and 93 provinces in the short term and by 85 and 86 provinces in the long term (Table 3). Unlike EU law, the WHO has not established safe limits for the PAHs (benzene and BaP) and heavy metals (As and Ni) considered, except for Cd, which remains unchanged. This is not very surprising because according to the EEA (2021), Italian and Polish cities were the ones with the highest levels of $PM_{2.5}$ in the period 2019–2020, among 323 investigated localities. In fact, among Europe's 53 worst cities for $PM_{2.5}$ levels, 20 were in Italy.

In Table 4, I also calculate a synthetic environmental pollution index for the Italian provinces in the period 2014–2019, using data on NO_2, $O_{3\ (>120)}$, $PM_{2.5}$, and PM_{10}, for which there are sufficient observations. Specifically, the index is compiled by switching the data on each of the four air pollutants considered to fixed-base indexes (with average = 1), from whose arithmetic mean I achieve the final standardized index. Provinces are ranked from the most polluted to the cleanest.

The output shows that the top positions are all in Northern Italy. In particular, the 29 most polluted Italian provinces are all concentrated in the eight northern regions of Italy. Among them, the top six positions are held by provinces within Lombardy, that is, the Italian region which has been most severely hit by the COVID-19 outbreak.

On the contrary, the southern provinces generally hold the lowest positions in the ranking. In the bottom 20 positions of the ranking, 17 are southern provinces, only three provinces are in Central Italy (Macerata, Pistoia, and Viterbo), and none are in Northern Italy.

While the most polluted southern provinces are Naples and Chieti, they are in 29th and 41st place respectively. The results seem to reflect the deep historical gap in industrialization and development between the north and south of Italy (Malanima and Zamagni 2010; Bigoni et al. 2019).

An air pollution map for the average long-term concentrations or violations of each air pollutant in the Italian provinces is given in Fig. 1.

[1] The air pollutants are described in detail in Table 6 (Section 4).
[2] All the values are at the provincial level, except for data on the livestock numbers, which are available only at the regional level. Data were extracted from I.Stat (2021a), except for supply of local public transport and number of vehicles used to transport goods and passengers, which were extracted from ISTAT (2021b) and ISTAT (2021c) respectively.

[3] Although these correlations do not imply causation, they warn of the potentially dangerous effects of large firms, vehicles, energy and gas consumption, and livestock.
[4] The health-based standards for the concentrations of air pollutants are provided at https://ec.europa.eu/environment/air/quality/standards.htm (accessed 24 March 2021).

3. Literature Review

It is well established that air pollution exposure can adversely affect lung function. NO_2, O_3, $PM_{2.5}$, and PM_{10} can be risk factors for several respiratory diseases, such as asthma (Cadelis et al. 2014), bronchiectasis (Goeminne et al. 2018), chronic obstructive pulmonary disease (COPD) (Liang et al. 2019), invasive pneumococcal disease (IPD) (Johannson et al. 2018), lung cancer (Xing et al. 2019), and general respiratory infections (Zheng et al. 2017).[5] Similarly, exposure to heavy metals, especially to Cd, can contribute to oxidative stress and inflammation in the lungs (Rokadia and Agarwal 2013; Mo et al. 2019). Meanwhile, exposure to airborne PAHs can exacerbate respiratory infections and increase the risk of several non-malignant respiratory diseases associated with exposure to other air pollution, such as particulate matter (Låg et al. 2020).

Therefore, in the last year and a half, a large body of literature has focused its attention on the relationship between air quality and the propagation pattern of the COVID-19 pandemic. Bashir et al. (2020) used two non-parametric statistical techniques—Kendal and Spearman rank-order correlation coefficients—to investigate the association between seven air pollutants and COVID-19 cases and deaths in California. Specifically, they analyzed the concentrations of CO, NO_2, Pb, $PM_{2.5}$, PM_{10}, SO_2, and volatile organic compounds (VOC) from 4 March 2020 to 24 April 2020. They found that CO, NO_2, $PM_{2.5}$, PM_{10}, and SO_2 were significantly and positively correlated with COVID-19 cases and deaths, and the highest correlation coefficients were exhibited by NO_2 and $PM_{2.5}$.

Becchetti et al. (2020) utilized several statistical techniques, such as the difference-in-difference (DID) approach, fixed effects (FE) panel regression, ordinary least square (OLS) panel regression, and cross-sectional spatial autoregressive combined models (SAC), to investigate the role of three major air pollutants in the spread of COVID-19 in 96 Italian provinces from 24 February 2020 to 4 April 2020. They found that average concentrations of NO_2, $PM_{2.5}$, and PM_{10} (registered in 2017) were highly significant and positively associated both with COVID-19 mortality and infections. The results were also confirmed after controlling for several demographic, environmental, economic, and healthcare cofactors.

By using a mixed linear multiple regression approach, Hendryx and Luo (2020) analyzed the effect of long-term exposure to diesel particulate matter (DPM), O_3, and $PM_{2.5}$ in relation to COVID-19 susceptibility or outcomes in the US. Specifically, they investigated the cumulative confirmed cases as of 31 May 2020, finding that DPM alone was significantly and positively associated with COVID-19 prevalence, and robust enough against changes in the specifications. Although positive, the coefficient of $PM_{2.5}$ was not robust enough.

Cole et al. (2020) examined the link between confirmed COVID-19 cases, deaths, hospitalizations, and long-term exposure to three major air pollutants (O_3, $PM_{2.5}$, and SO_2) in 355 municipalities in the Netherlands. By using instrumental variable (IV) regressions, NB approaches, and SARAR models, they found that only the $PM_{2.5}$ coefficient was significant and robust against changes in the specifications. Specifically, for every 1 μg/m^3 increase in $PM_{2.5}$ concentrations, there was an increase of 9.4 cases, 2.3 deaths, and three hospitalizations.

By using a generalized additive model (GAM), Zhu et al. (2020) investigated the short-term relationship between several air pollutants and daily confirmed COVID-19 cases in 120 Chinese cities from 23 January 2020 to 29 February 2020. They found that 1-unit μg/m^3 increases in NO_2, O_3, $PM_{2.5}$, and PM_{10} were associated with 0.69%, 0.48%, 0.22%, and 0.18% increases respectively in daily confirmed COVID-19 cases. On the contrary, a 1-unit μg/m^3 increase in SO_2 was linked with a 0.78% decrease in daily confirmed COVID-19 cases.

Solimini et al. (2021) used negative binomial mixed effect models to investigate the association between long-term exposure to PM_{10} and $PM_{2.5}$ and COVID-19 cases in a large sample of countries. The data came from 63 countries, 730 regions, and five continents, and was updated on 30 May 2020. After adjusting the models for several regional and country covariates and spatial correlation, they found that 1-unit μg/m^3 increases in the $PM_{2.5}$ and PM_{10} concentrations were significantly correlated with increases of 0.81% and 1.15% respectively in the total number of confirmed COVID-19 cases in a 14-day window.

Table 5 summarizes 22 international studies on the relationship between environmental pollution and the spread of COVID-19 infections.

[5] For an extensive and comprehensive review of the related literature, see Bălă et al. (2021).

4. Data

In this section, I report the variables used in the empirical analysis. First, to avoid spurious correlations and mitigate the problem of omitted variables, I implement 18 covariates to account for geographical proximity, demographic characteristics, population habits and structure, industrial centers, hospital saturation, and weather conditions:

- four dummy variables to identify the provinces that border Austria, France, Slovenia, and Switzerland respectively;
- a dummy variable to identify the provinces that are also the regional capital;
- the distance between the provincial capital's center and the nearest airport with at least 50,000 passengers in the period from January to November 2020;
- the share of population aged 0–19 in each province, in 2020;
- the share of male population in each province, in 2020;
- the degree of urbanization of the population in each province;

- the average share of alcohol drinkers at regional level, in the period 2016–2019;
- the average share of obese individuals at regional level, in the period 2016–2019;
- the average share of smokers at regional level, in the period 2016–2019;
- the average deaths from chronic respiratory disease per 100,000 inhabitants in each province, in the period 2014–2019;
- the ratio between people who have been tested positive for COVID-19 and the average ordinary hospital beds in the period 2017–2018, in each province;
- the share of firms with 250 or more employees in each province, in the period 2014–2019;
- the average altitude of the capital of the province;
- the average annual days of rain in each province, in the period 2007–2018;
- the average annual temperature in each province, in the period 2008–2018.

Regarding the explanatory variables, I chose the following nine air pollutants, calculated – when data are available – for each Italian province:

- the average concentrations of NO_2, expressed in micrograms per cubic meter of air (μg/m^3), in the period 2014–2019;
- the average number of days in which Ozone exceeded the limit of 120 μg/m^3, in the period 2014–2019;
- the average number of hours in which Ozone exceeded the limit of 180 μg/m^3, in the period 2014–2018;
- the average concentrations of $PM_{2.5}$, expressed in μg/m^3, in the period 2014–2019;
- the average concentrations of PM_{10}, expressed in μg/m^3, in the period 2014–2019;
- average number of days in which PM_{10} exceeded the limit of 50 μg/m^3 in the period 2014–2018;
- the average concentrations of benzene, expressed in nanogram per cubic meter of air (ng/m^3), in the period 2014-2016;
- the average concentrations of BaP, expressed in μg/m^3, in the period 2014–2016;
- the average concentrations of As, expressed in ng/m^3, in the period 2014–2016;
- the average concentrations of Cd, expressed in ng/m3, in the period 2014– 2016;
- the average concentrations of Ni, expressed in ng/m^3, in the period 2014–2016.

As dependent variables, I use i) the number of cumulative confirmed COVID-19 cases on 30 November 2020, in each province; ii) and the proportion of the total resident population infected by COVID-19 on 30

November 2020 (or on 20 February 2021) in each province. All the independent and dependent variables are also described in detail in Table 6.[6]

[6] A summary of the main descriptive statistics is provided in Table A1 (Appendix A).

5. Empirical Strategy

The main goal of this paper is to estimate the relationship between long-term exposure to nine air pollutants and the spread of COVID-19 across 107 Italian provinces, using different econometric techniques. To measure the spread of COVID-19, I use both the absolute confirmed cases of the disease and its prevalence, expressed as a percentage of the population, as of 30 November 2020. This date was chosen by looking at the peak of the use of daily nasal swabs for testing COVID-19 at the second peak of the epidemic, which can be approximately dated to the end of November 2020. In fact, at that time, more than 200,000 swabs were consistently used daily (Sole 24 Ore 2021) and it is possible to hypothesize that they presented a reliable snapshot of reality.[7] In fact, in a cross-section analysis the differences across units are more important than the number of infections. This choice may mitigate the inevitable bias in detecting infected people, which was also probably raised in early 2021 due to the start of the nationwide COVID-19 vaccination campaign.

Regarding the empirical strategy, I use a negative binomial regression that fits well when the dependent variable is a count variable, such as the cumulative confirmed cases of COVID-19. The choice of a negative binomial approach instead of a standard Poisson regression is based on the evaluation of the likelihood-ratio (LR) test on the overdispersion parameter alpha. The negative binomial regression can be considered a generalization of Poisson regression that allows the conditional variance to exceed the conditional mean. To do this, the negative binomial approach considers an extra parameter that corrects the effects of the larger variance on the p-values (Cameron and Johansson 1997). To avoid biased results, I also include the size of the provincial population as an exposure variable. This is a pivotal point, because it allows me to standardize the cumulative confirmed cases, that is, convert each observation from a count variable into a rate. As result, I estimate the following basic equation:

$$\mathrm{Covid_i} = \beta_0 + \beta_1 \mathrm{D_i} + \beta_2 \mathrm{SD_i} + \beta_3 \mathrm{E_i} + \beta_4 \mathrm{C_i} + \beta_5 \mathrm{Pollutant_i} + \epsilon_i \quad [1]$$

Where i identifies each province, β_0 is a constant, $\mathrm{D_i}$ is a vector of dummy variables for identifying Italian provinces with international borders, $\mathrm{DE_i}$ is a vector of demographic and economic factors, $\mathrm{E_i}$ is a vector of epidemiological features, $\mathrm{C_i}$ is a vector of climate variables, $\mathrm{Pollutant_i}$ refers to the concentrations or violations of nine selected air pollutants (NO_2, O_3, $PM_{2.5}$, PM_{10}, benzene, BaP, As, Cd, and Ni), and ϵ_i is the error term.

As sensitivity checks, I modeled the prevalence of COVID-19 expressed in percentage points, using a standard ordinary least squares (OLS) approach and a spatial-autoregressive (SAR) framework. OLS can be considered the most widely used econometric technique for linear statistical models. It takes the same

form of equation [1], with the only exception of the dependent variable, which is the COVID-19 prevalence rate at the provincial level.

However, this procedure is not immune from issues, because from a theoretical point of view it is unlikely that neighboring provinces did not affect each other. In fact, the transmission within neighbor territories may have been reasonably affected by the movement of people, which is easier and faster across provinces' borders. The presence of spatial dependence in the dependent variable may lead to substantial bias in OLS models (Anselin 1998), resulting in inconsistent outcomes. Thus, I controlled for possible spatial effects in the dependent variable by following two sequential steps: i) I investigated the map of COVID-19 prevalence on 30 November 2020 to make sure that an eventual spatial pattern was visible; and ii) I calculated a common measure of spatial autocorrelation, the global Moran's I statistic (Moran 1948; Cliff and Ord 1973), to verify whether each infection had the same likelihood of occurring at any location. [8] Based on the evaluation of these metrics, I implemented a spatial-autoregressive model with autoregressive disturbances (SARAR). In particular, the model was estimated with a maximum likelihood (ML) approach instead of the more common generalized spatial two-stage least squares (GS2SLS) approach. This choice is justified by performing Cameron and Trivedi's (1990) decomposition of White's information matrix (IM) test over the hypothesis of normality and heteroscedasticity of the errors, which needs to be met to implement the ML estimator (Drukker et al. 2013, p. 236).

The equation estimated for the SARAR model was eventually obtained by adding a spatially lagged dependent variable and a spatial error term to the previous basic equation [1]. The spatially lagged dependent variable aimed to verify if and how much a given province was influenced by the prevalence of the neighbor provinces, while the spatial error term was meant to test whether an exogenous shock to one province may have caused a change in the prevalence of the neighbor provinces. The final equation takes the following form:

$$\mathrm{Covid}_i = \beta_0 + \beta_1 D_i + \beta_2 DE_i + \beta_3 E_i + \beta_4 C_4 + \beta_5 \mathrm{Pollutant}_i + \rho w_i \mathrm{Covid}_i + \lambda w_i \epsilon_i + \epsilon_i$$ [2]

Where i identifies each province, $\beta_0$is a constant, D_i is a vector of dummy variables for identifying Italian provinces with international borders, DE_i is a vector of demographic and economic factors, E_i is a vector of epidemiological features, $\mathrm{Pollutant}_i$ refers to the average concentrations or violations of nine selected air pollutants (NO_2, O_3, $PM_{2.5}$, PM_{10}, Benzene, BaP, As, Cd, and Ni), ρ_i is the spatially lagged dependent variable, λ_i is the spatial error term, w_i is an inverse-distance weighted matrix, and ϵ_i is the error term. The matrix was row standardized because: i) this allows for comparing spatial parameters that come from different models; and ii) since all the weights summed to 1, the fact that one feature may have two neighbors and another may have many more does not have a large effect on the results.

Finally, as a further sensitivity check, I used the data on COVID-19 prevalence rates from 20 February 2021, that is, exactly one year after the start of the COVID-19 outbreak in Italy. This aimed to test whether the relationship between major air pollutants and COVID-19 spread was maintained over time.

[7] Although the number of daily swabs was even higher during the third wave of the COVID-19 epidemic, I preferred not to use these data. In fact, the third peak of the epidemic occurred around 8 April 2021, when more than 14% of the Italian population had received at least one dose of a COVID-19 vaccine (Mathieu et al. 2021).

[8] I also computed the global Moran's I statistics for all the dependent variables.

6. Results And Discussion

6.1 Negative binomial regressions

In Table 7, I present the negative binomial model estimations for the 107 Italian provinces.[9] All 12 models were significant; in fact, the Fisher-Snedecor distribution assumed values far higher than the tabulated critical values at the 1% level of significance. The McFadden's (1974) pseudo-R-square is substantially homogenous across specifications and ranges between 0.1004 and 0.1235. Since pseudo-R-square values are usually much lower than those of the classic R-square, the results obtained can be considered a decent fit (McFadden 1978). Moreover, the likelihood-ratio (LR) chi-square test allows us to strongly reject the null hypothesis that the dispersion parameter alpha is equal to zero. Thus, the negative binomial approach is a better fit for the data than the Poisson regression.

Regarding control variables, the results showed that a border with Switzerland, the share of foreigners, population density, and altitude were significantly and positively correlated with cumulative confirmed COVID-19 cases on 30 November 2020. Conversely, distance from the nearest main airport, saturation of ordinary beds, and average temperature were significantly and negatively associated with total confirmed COVID-19 cases. With regard to air pollutants, NO_2, $O_{3(>120)}$, $PM_{2.5}$, PM_{10}, benzene and BaP showed a positive and statistically significant relationship with COVID-19 infections. However, if the first four air pollutants were verified at 1% levels of significance, benzene and BaP were verified at 5% and 10% levels of significance respectively. For the remainder, As, Cd, and Ni were not significant at all.

Since coefficients that come from negative binomial models cannot easily be interpreted, in Table 8 I computed the marginal effect for the air pollutants that were statistically significant. The most significant coefficients were $PM_{2.5}$ and PM_{10}, which were verified at 1% levels of significance, followed by NO_2, benzene, and BaP, which were verified at 5% levels of significance. Regarding primary pollutants, 1 $\mu g/m^3$ increases in $PM_{2.5}$, PM_{10}, and NO_2 were associated with average increases of 297.8, 282.9, and 117.2 COVID-19 infections respectively, while for PAHs, a 0.1 $\mu g/m^3$ increase in benzene and a 0.1 ng/m^3 increase in BaP were associated with average increments of 193.2 and 166.5 COVID-19 infections respectively.[10] Thus, among common air pollutants, $PM_{2.5}$ and PM_{10} seemed to have the most adverse effects on COVID-19 spread, while benzene was the most dangerous among the remaining pollutants.

6.2 OLS regression models

To strengthen the results, in Table 9, I estimated an OLS regression model for the prevalence of COVID-19 infections in the 107 Italian provinces. Since standard errors are usually biased in small samples, I corrected them for heteroscedasticity by applying the HC2 estimator proposed by MacKinnon and White (1985), which performs well even when sample size is not large (Davidson and MacKinnon 1993, p. 533). The Fisher-Snedecor distribution was highly significant and verified at a 1% level of significance for all the OLS models; therefore, the choice of the independent variables can be considered appropriate and justified. Moreover, the R-square was very high and ranged from 0.8165 to 0.8725, showing that all models were a good fit and explained a large fraction of the variability of COVID-19 prevalence. The variance inflation factors (VIF) were always less than the threshold of 5, suggesting that there were no severe multicollinearity issues (Rogerson 2001). The only exception was the coefficient of the temperature in model 1, which was carefully excluded by the other models.

The results are very similar to those obtained from the negative binomial regression models. Concerning the control variables, a border with Switzerland, foreigners, deaths from respiratory disease, population density, and altitude were significantly and positively correlated with COVID-19 prevalence; meanwhile, distance from the nearest airport, ordinary bed saturation, and temperature were significantly and negatively associated with infection rates. Notably, just borders with Switzerland were significant, while Austria, France, and Slovenia had no association with COVID-19 cases. This may be due to the flow of the 65,000 cross-border workers who reside in Italy and work in Switzerland, and who account for a total of 63.73% of all Italian cross-border commuters (European Commission 2021, pp. 184–185). The significance of foreigner population could be explained by foreigners' greater propensity to travel to their native countries, which could have increased the probability of meeting infected people.

The direction of the correlation between population density and COVID-19 cases is consistent with recent literature (Wong and Li 2020; Diao et al. 2021), suggesting the importance of keeping a safe physical distance from others to limit the spread of the outbreak. The positive significance of altitude, conversely, is in contrast with most of the recent literature (Cano-Pérez et al. 2020; Segovia-Juarez et al. 2020; Arias-Reyes et al. 2021; Fernandes et al. 2021). However, these studies mainly focused on Latin American countries, such as Colombia, Peru, and Brazil, which have cities with altitude differences of up to more than 3,000 meters. As shown by Table A1 (Appendix A), the difference between the most low-altitude city (Venice) and the most high-altitude city (L'Aquila) is just 1,168.2 meters, suggesting a lower isolation of the population. The positive effect of the number of deaths from respiratory diseases seems to stress the more significant vulnerability of people with comorbidities, who are more likely to get infected (Ejaz et al. 2020).

On the contrary, higher temperatures may have favored a reduction of COVID-19 transmission, and this result appears consistent with most of the recent literature (Sarkodie and Owusu 2020; Tobías and Molina 2020; Chen et al. 2021). The negative relationship between transmission and distance from the nearest airport seems to advocate the beneficial effect of travel restrictions. The adverse effect of the saturation of ordinary hospital beds may suggest that the shortage of beds forced patients to share the same rooms or areas. In this sense, the saturation of healthcare facilities may have favored the propagation of the virus.

Regarding air pollutants, $PM_{2.5}$, PM_{10}, and $PM_{10\ (>50)}$ were statistically significant at the 1% level, while NO_2, $O_{3(>120)}$, benzene, and BaP showed significance levels of 5%. Among common air pollutants, 10 μg/m^3 increases in the concentrations of NO_2, $PM_{2.5}$, and PM_{10} were associated respectively with average increments of 0.2%, 0.43%, and 0.45% of COVID-19 prevalence. Among PAHs, a 1-unit μg/m^3 increase in benzene and a 1-unit ng/m^3 in BaP were associated respectively with increases of 2.47% and 2.94% in nationwide COVID-19 prevalence. On the contrary, the concentrations of heavy metals were not significant.

6.3 Robustness checks: Spatial-autoregressive analysis

In Table 10, I present the results of the SARAR models regarding COVID-19 prevalence on 30 November 2020. The use of the SARAR approach is justified by the global Moran's I, which allowed me to reject the null hypothesis that data were randomly distributed both for the dependent and independent variables, with the only exception of distance from the nearest airport. Specifically, the global Moran's I was always positive and statistically significant at 1%, ranging from 0.032 to 0.329.[11]

Moreover, since the spatially lagged dependent variable is always highly significant, the SARAR approach is more appropriate than a classical OLS econometric technique. The use of an ML estimator was also justified by the use of Cameron and Trivedi's (1990) decomposition of IM test over the OLS models reported in Table A1 (Appendix A). All the tests confirmed the hypothesis that OLS errors were homoscedastic and reasonably close to a normal distribution, definitively advocating the ML approach (Drukker et al. 2013, p. 236).

Specifically, the outcomes show that the spatially lagged dependent variable was always positive and verified at a 1% level of significance, suggesting that neighboring provinces tended to display similar patterns in terms of the spread of COVID-19. A higher level of infection rate in one province seemed to increase prevalence in the neighboring provinces. This may largely be due to the fact that people usually move more easily to neighboring provinces, increasing the likelihood of meeting someone with COVID-19 and spreading the infection. On the contrary, the spatially autocorrelated error term was never significant, suggesting that there were no common-shock effects.

NO_2, O_3, $PM_{2.5}$, PM_{10}, benzene, and BaP remained positive and statistically significant despite the inclusion of the spillover effects, although the significance level of NO_2, O_3, and BaP switched from 5% to 10%. Notably, the coefficients of Cd became significant at the 5% level. In Table 11, I estimated the marginal effects for each air pollutant, consisting of a direct, indirect, and total effect. The direct effect refers to the impact of each air pollutant on COVID-19 spread in the same province, while the indirect effect is the impact on spread in other provinces. If among common air pollutants, the coefficients with the highest magnitude belonged to $PM_{2.5}$ and PM_{10}, among PAHs and heavy metals the highest values belonged to benzene and Cd respectively. Specifically, $PM_{2.5}$ and PM_{10} respectively exhibited direct effects of 0.032 and 0.039 and indirect effects of 0.2 and 0.27. Benzene and Cd respectively showed direct effects of 0.27 and 0.36 and indirect effects of 1.95 and 1.55. Thus, the indirect effect of air pollutants was much greater than the direct one.

As a further sensitivity check, in Table 12 I computed the SARAR models for the prevalence registered one year after the start of the outbreak, that is, on 20 February 2021. The results showed that air pollutants substantially maintained a positive significance, although the significance level had changed in some cases. Specifically, $O_{3\ (>120)}$ became not significant at all, benzene switched from 5% to 10%, and BaP and Cd increased in significance respectively from 10% to 1% and from 5% to 1%. Moreover, $PM_{2.5}$ and PM_{10} showed direct and indirect effects slightly lower than those on 30 November 2021 (Table 13). On the contrary, benzene and Cd showed direct and indirect effects that were definitively greater than those registered on 30 November 2020. Specifically, benzene and Cd respectively exhibited direct effects of 0.36 and 0.77 and indirect effects of 0.84 and 0.65. Thus, the results substantially show the persistence of the link between environmental pollution and the transmission of COVID-19, also suggesting the potentially dangerous effect of PAHs and heavy metals, such as benzene, BaP, and Cd.

[9] It is necessary to stress that data for all Italian territories were available only for NO_2, $O_{3\ (>120)}$, PM_{10}, and $PM_{10\ (>50)}$. Conversely, the observations for the remaining air pollutant concentrations or violations ranged from 60 to 98 territories.
[10] I chose 0.1 units for benzene and BaP because their legal threshold was comparatively much lower than that for common air pollutants.
[11] It should be noted that the Moran's index for COVID-19 prevalence is very low (0.032). However, since it is highly significant, the use of spatial analysis seems appropriate. The Moran's I statistics for the dependent and independent variables are reported in Table B1 (Appendix B).

Conclusions

In this article, I investigated the common sources of outdoor air pollution and the global air quality in the 107 Italian provinces in the period 2014–2019, and the link between long-term exposure to nine air pollutants in the same period and the spread of COVID-19 infections. The results showed that: i) common air pollutants (NO_2, O_3, $PM_{2.5}$, and PM_{10}) and PAHs (benzene and BaP) exhibited a positive and significant correlation with the presence of large firms, energy and gas consumption, vehicles, public transport, cattle fodder, and livestock; ii) the provinces located in the north of Italy were generally much more polluted than the southern ones; and iii) long-term exposure to air pollutants was positively correlated with the spread of COVID-19 infections across the Italian provinces.

The dangerous effect of the common air pollutants NO_2, O_3, $PM_{2.5}$ and PM_{10} was consistent with most of the recent literature (Bolaño-Ortiz et al. 2020; Cole et al. 2020; Vasquez-Apestegui et al. 2020; Zhu et al. 2020; Solimini et al. 2021). Moreover, this study found that as well as the common air pollutants, PAHs and heavy metals may also have played an important role in the spread of the COVID-19 outbreak. This outcome seems interesting and of relevance, given that these air pollutants have not been considered at all by recent scientific literature. Finally, the results seem to suggest the need for national strategies and economic policies that aim at reducing air pollutant concentrations in order to improve air quality levels (especially in Northern Italy) and to cope more effectively with similar unexpected pandemics in the future.

Declarations

Founding: The authors did not receive support from any organization for the submitted work.

Conflicts of interest/Competing interests: The author has no relevant financial or non-financial interests to disclose.

No Acknowledgment.

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.

Malanima, P., & Zamagni, V. (2010). 150 years of the Italian economy, 1861–2010. Journal of Modern Italian Studies, 15(1), 1-20.

Bigoni, M., Bortolotti, S., Casari, M., & Gambetta, D. (2019). At the root of the North–South cooperation gap in Italy: Preferences or beliefs?. The Economic Journal, 129(619), 1139-1152.

Anselin L (1988) Spatial econometrics: methods and models. Kluwer, Dordrecht

Arias-Reyes, C., Carvajal-Rodriguez, F., Poma-Machicao, L., Aliaga-Raduán, F., Marques, D. A., Zubieta-DeUrioste, N., ... & Soliz, J. (2021). Decreased incidence, virus transmission capacity, and severity of COVID-19 at altitude on the American continent. PloS one, 16(3), e0237294.

Associazione Italiana Gestori Aeroporti (ASSOAEROPORTI) (2021). https://assaeroporti.com/statistiche/. Accessed 13 January 2021.

Bălă, G. P., Râjnoveanu, R. M., Tudorache, E., Motișan, R., & Oancea, C. (2021). Air pollution exposure—the (in) visible risk factor for respiratory diseases. Environmental Science and Pollution Research, 1-14.[qui tutti I riferimenti utilizzati]

Bashir, M. F., Bilal, B. M., & Komal, B. (2020). Correlation between environmental pollution indicators and COVID-19 pandemic: A brief study in Californian context. Environmental Research, 109652.

Becchetti, L., Conzo, G., Conzo, P., & Salustri, F. (2020). Understanding the heterogeneity of adverse COVID-19 outcomes: the role of poor quality of air and lockdown decisions. Available at SSRN 3572548.

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, 109639.

Cameron, A. C., & Johansson, P. (1997). Count data regression using series expansions: with applications. *Journal of Applied Econometrics*, *12*(3), 203-223.

Cameron, A. C., and P. K. Trivedi (1990) The information matrix test and its applied alternative hypotheses. Working paper 372, University of California–Davis, Institute of Governmental Affairs.

Cano-Pérez, E., Torres-Pacheco, J., Fragozo-Ramos, M. C., García-Díaz, G., Montalvo-Varela, E., & Pozo-Palacios, J. C. (2020). Negative correlation between altitude and COVID-19 pandemic in Colombia: a preliminary report. The American Journal of Tropical Medicine and Hygiene, 103(6), 2347.

Chen, S., Prettner, K., Kuhn, M. et al. Climate and the spread of COVID-19. Sci Rep 11, 9042 (2021). https://doi.org/10.1038/s41598-021-87692-z

Coccia, M. (2020). Factors determining the diffusion of COVID-19 and suggested strategy to prevent future accelerated viral infectivity similar to COVID. Science of the Total Environment, 729, 138474.

Cole, M. A., Ozgen, C., & Strobl, E. (2020). Air pollution exposure and Covid-19 in Dutch municipalities. Environmental and Resource Economics, 76(4), 581-610.

Comunian, S., Dongo, D., Milani, C., & Palestini, P. (2020). Air pollution and COVID-19: the role of particulate matter in the spread and increase of COVID-19's morbidity and mortality. International journal of environmental research and public health, 17(12), 4487.

Collivignarelli, M. C., Abbà, A., Caccamo, F. M., Bertanza, G., Pedrazzani, R., Baldi, M., ... & Miino, M. C. (2021). Can particulate matter be identified as the primary cause of the rapid spread of CoViD-19 in some areas of Northern Italy?. Environmental Science and Pollution Research, 1-13.

Davidson, Russell & MacKinnon, James G., 1993. "Estimation and Inference in Econometrics," OUP Catalogue, Oxford University Press, New York.

De Angelis, E., Renzetti, S., Volta, M., Donato, F., Calza, S., Placidi, D., ... & Rota, M. (2021). COVID-19 incidence and mortality in Lombardy, Italy: an ecological study on the role of air pollution, meteorological factors, demographic and socioeconomic variables. Environmental research, 195, 110777.

De Donno, A., De Giorgi, M., Bagordo, F., Grassi, T., Idolo, A., Serio, F., ... & MAPEC_LIFE Study Group. (2018). Health risk associated with exposure to PM10 and benzene in three Italian towns. International journal of environmental research and public health, 15(8), 1672.

Delnevo, G., Mirri, S., & Roccetti, M. (2020). Particulate matter and COVID-19 disease diffusion in Emilia-Romagna (Italy). Already a cold case?. Computation, 8(2), 59.

Diao, Y., Kodera, S., Anzai, D., Gomez-Tames, J., Rashed, E. A., & Hirata, A. (2021). Influence of population density, temperature, and absolute humidity on spread and decay durations of COVID-19: A comparative study of scenarios in China, England, Germany, and Japan. One Health, 12, 100203.

Drukker, D. M., Prucha, I. R., & Raciborski, R. (2013). Maximum likelihood and generalized spatial two-stage least-squares estimators for a spatial-autoregressive model with spatial-autoregressive disturbances. The Stata Journal, 13(2), 221-241.

Ejaz, H., Alsrhani, A., Zafar, A., Javed, H., Junaid, K., Abdalla, A. E., ... & Younas, S. (2020). COVID-19 and comorbidities: Deleterious impact on infected patients. Journal of infection and public health, 13(12), 1833-1839.

European Commission. Air Quality Standards. https://ec.europa.eu/environment/air/quality/standards.htm

European Commission (2021), Annual report on intra-EU labour mobility 2020. https://op.europa.eu/en/publication-detail/-/publication/ab706f9b-74bf-11eb-9ac9-01aa75ed71a1/language-en.

European Environmental Agency (EEA) (2019). Healthy environment, healthy lives: how the environment influences health and well-being in Europe. https://www.eea.europa.eu/publications/healthy-environment-healthy-lives

European Environmental Agency (EEA) (2021). European city air quality viewer. January 17, 2021. https://www.eea.europa.eu/themes/air/urban-air-quality/european-city-air-quality-viewer

European Environmental Agency (EEA) (2020a). Air quality in Europe – 2020 report, EEA report NO. 09/2020. https://www.eea.europa.eu/publications/air-quality-in-europe-2020-report

European Environmental Agency (EEA) (2020b), Air pollution sources. https://www.eea.europa.eu/themes/air/air-pollution-sources-1

Eurostat (2013). Rural Development. “Overview Regions & cities”. https://ec.europa.eu/eurostat/web/rural-development/methodology.

Fattorini, D., & Regoli, F. (2020). Role of the chronic air pollution levels in the Covid-19 outbreak risk in Italy. Environmental Pollution, 114732.

Fernandes, J. S. C., da Silva, R. S., Silva, A. C., Villela, D. C., Mendonça, V. A., & Lacerda, A. C. R. (2021). Altitude conditions seem to determine the evolution of COVID-19 in Brazil. Scientific Reports, 11(1), 1-12.

Greenpace (2020). Covid-19, esposizione al particolato e allevamenti intensivi. https://www.greenpeace.org/static/planet4-italy-stateless/2020/04/184484ca-ricerca-particolato-def.pdf

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.

Italian Institute for Environmental Protection and Research (ISPRA) (2017). XIII Rapporto Qualità dell’ambiente urbano – Edizione 2017 (Chapter: Qualità dell’aria). https://www.isprambiente.gov.it/it/pubblicazioni/stato-dellambiente/xiii-rapporto-qualita-dell2019ambiente-urbano-edizione-2017.

Italian Institute for Environmental Protection and Research (ISPRA) (2016). XII Rapporto Qualità dell'ambiente urbano – Edizione 2016 (Chapter: Qualità dell'aria). https://www.isprambiente.gov.it/it/pubblicazioni/stato-dellambiente/xii-rapporto-qualita-dell2019ambiente-urbano-edizione-2016.

Italian Institute for Environmental Protection and Research (ISPRA) (2015). Qualità dell'ambiente urbano – XI Rapporto. Edizione 2015 (Chapter: Qualità dell'aria). https://www.isprambiente.gov.it/it/pubblicazioni/stato-dellambiente/qualita-dellambiente-urbano-xi-rapporto.-edizione-2015.

Italian Ministry of Health (2018). Posti letto per stabilimento ospedaliero e disciplina. http://www.dati.salute.gov.it/dati/dettaglioDataset.jsp?menu=dati&idPag=96. Accessed 20 January 2021.

Italian Ministry of Health (2020). COVID-19, Dati Province. https://github.com/pcm-dpc/COVID-19/tree/master/dati-province.

Italian National Institute of Health (ISS) (2021). Sorveglianza Passi. https://www.epicentro.iss.it/passi/.

Italian National Institute of Statistics (ISTAT) (2015). Qualità dell'ambiente urbano – Fattori di pressione. https://www.istat.it/it/archivio/173187.

Italian National Institute of Statistics (ISTAT) (2017). Ambiente Urbano. https://www.istat.it/it/archivio/207482.

Italian National Institute of Statistics (ISTAT) (2019). Ambiente Urbano. https://www.istat.it/it/archivio/236912

Italian National Institute of Statistics (ISTAT) (2020a). Temperatura e precipitazioni nelle città capoluogo di provincia. https://www.istat.it/it/archivio/242010.

Italian National Institute of Statistics (ISTAT) (2020b). Ambiente Urbano. https://www.istat.it/it/archivio/244648.

Italian National Institute of Statistics (ISTAT) (2021a). I.Stat database. http://dati.istat.it Accessed 1 March 2021.

Italian National Institute of Statistics (ISTAT) (2021b). Ambiente Urbano. 25 February 2021. https://www.istat.it/it/archivio/254037.

Italian National Institute of Statistics (ISTAT) (2021c). Ambiente Urbano. 17 June 2021.chttps://www.istat.it/it/archivio/258691

Italian National Institute of Statistics (ISTAT) (2021d). Principali Statistiche Geografiche sui Comuni. https://www.istat.it/it/archivio/156224

Låg, M., Øvrevik, J., Refsnes, M., & Holme, J. A. (2020). Potential role of polycyclic aromatic hydrocarbons in air pollution-induced non-malignant respiratory diseases. Respiratory Research, 21(1), 1-22.

Laubach, J., Taghizadeh-Toosi, A., Gibbs, S. J., Sherlock, R. R., Kelliher, F. M., & Grover, S. P. P. (2013). Ammonia emissions from cattle urine and dung excreted on pasture. Biogeosciences, 10(1), 327-338.

Legambiente (2020). Ecosistema Urbano. Rapporto sulle performance ambientali delle città 2020. https://www.legambiente.it/wp-content/uploads/2020/11/Ecosistema-Urbano-2020.pdf

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.

Lin, S., Wei, D., Sun, Y., Chen, K., Yang, L., Liu, B., ... & Wu, S. (2020). Region-specific air pollutants and meteorological parameters influence COVID-19: A study from mainland China. Ecotoxicology and environmental safety, 204, 111035.

Bolaño-Ortiz, T. R., Camargo-Caicedo, Y., Puliafito, S. E., Ruggeri, M. F., Bolaño-Diaz, S., Pascual-Flores, R., ... & Cereceda-Balic, F. (2020). Spread of SARS-CoV-2 through Latin America and the Caribbean region: a look from its economic conditions, climate and air pollution indicators. Environmental research, 191, 109938.

Lolli, S., Chen, Y. C., Wang, S. H., & Vivone, G. (2020). Impact of meteorological conditions and air pollution on COVID-19 pandemic transmission in Italy. Scientific reports, 10(1), 1-15.

Lovarelli, D., Conti, C., Finzi, A., Bacenetti, J., & Guarino, M. (2020). Describing the trend of ammonia, particulate matter and nitrogen oxides: The role of livestock activities in northern Italy during Covid-19 quarantine. Environmental research, 191, 110048.

McCubbin, D. R., Apelberg, B. J., Roe, S., & Divita, F. (2002). Livestock ammonia management and particulate-related health benefits. Environmental Science and Technology, 36(6), 1141-1146.

McFadden D. (1978), Quantitative Methods for Analyzing Travel Behaviour of Individuals: Some Recent Developments, in Hensher D. and Stopher P. (eds.), Behavioural Travel Modelling, Croom Helm, 1978, pp. 279-318.

McFadden, D. (1974) “Conditional logit analysis of qualitative choice behavior.” Pp. 105-142 in P. Zarembka (ed.), Frontiers in Econometrics. Academic Press. http://eml.berkeley.edu/~mcfadden/travel.html

Ministero delle politiche agricole alimentari e forestali (Mipaaf). Osservatorio agroclimatico: statistiche provinciali. https://www.politicheagricole.it/flex/FixedPages/Common/miepfy700_provincie.php/L/IT

Mo, Y., Jiang, M., Zhang, Y., Wan, R., Li, J., Zhong, C. J., ... & Zhang, Q. (2019). Comparative mouse lung injury by nickel nanoparticles with differential surface modification. Journal of nanobiotechnology, 17(1), 1-18.

Moran, P. A. (1948). The interpretation of statistical maps. *Journal of the Royal Statistical Society: Series B (Methodological)*, *10*(2), 243-251.

Rogerson P. A. (2001), Statistical methods for geography. Sage: London.

Rokadia, H. K., & Agarwal, S. (2013). Serum heavy metals and obstructive lung disease: results from the National Health and Nutrition Examination Survey. Chest, 143(2), 388-397.

Sarkodie, S. A., & Owusu, P. A. (2020). Impact of meteorological factors on COVID-19 pandemic: Evidence from top 20 countries with confirmed cases. Environmental Research, 191, 110101.

Schraufnagel, D. E., Balmes, J. R., Cowl, C. T., De Matteis, S., Jung, S. H., Mortimer, K., ... & Wuebbles, D. J. (2019). Air pollution and noncommunicable diseases: A review by the Forum of International Respiratory Societies' Environmental Committee, Part 2: Air pollution and organ systems. Chest, 155(2), 417-426.

Segovia-Juarez, J., Castagnetto, J. M., & Gonzales, G. F. (2020). High altitude reduces infection rate of COVID-19 but not case-fatality rate. Respiratory Physiology & Neurobiology, 281, 103494.

Setti, L., Passarini, F., De Gennaro, G., Barbieri, P., Perrone, M. G., Borelli, M., ... & Miani, A. (2020). SARS-Cov-2RNA found on particulate matter of Bergamo in Northern Italy: first evidence. Environmental research, 188, 109754.

Sole 24 Ore (2021). Lab24: Tamponi giornalieri e contagiati. https://lab24.ilsole24ore.com/coronavirus/

Solimini, A., Filipponi, F., Fegatelli, D. A., Caputo, B., De Marco, C. M., Spagnoli, A., & Vestri, A. R. (2021). A global association between Covid-19 cases and airborne particulate matter at regional level. Scientific reports, 11(1), 1-7.

Tobías, A., & Molina, T. (2020). Is temperature reducing the transmission of COVID-19?. Environmental research, 186, 109553.

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.

United States Environmental Protection Agency (US EPA 2017). Toxicological Review of BaP. https://iris.epa.gov/static/pdfs/0136_summary.pdf

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), 1-13.

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), 1-13.

Mathieu, E., Ritchie, H., Ortiz-Ospina, E., Roser, M., Hasell, J., Appel, C., ... & Rodés-Guirao, L. (2021). A global database of COVID-19 vaccinations. Nature human behaviour, 1-7.

Webb, J., Menzi, H., Pain, B. F., Misselbrook, T. H., Dämmgen, U., Hendriks, H., & Döhler, H. (2005). Managing ammonia emissions from livestock production in Europe. Environmental pollution, 135(3), 399-406.

Wong, D. W., & Li, Y. (2020). Spreading of COVID-19: Density matters. Plos one, 15(12), e0242398.

World Health Organization (WHO) (2006). Air quality guidelines global update 2005: particulate matter, ozone, nitrogen dioxide and sulfur dioxide. Copenhagen: WHO Regional Office for Europe. https://apps.who.int/iris/handle/10665/107823

World Health Organization (WHO) (2010). Preventing Disease Through Healthy Environments. Exposure to benzene: a major public health concern. https://www.who.int/ipcs/features/benzene.pdf

World Health Organization (WHO) (2015). WHO Expert Consultation: Available evidence for the future update of the WHO Global Air Quality Guidelines (AQGs). https://www.euro.who.int/__data/assets/pdf_file/0013/301720/Evidence-future-update-AQGs-mtg-report-Bonn-sept-oct-15.pdf.

World Health Organization (WHO) (2016). Ambient air pollution: a global assessment of exposure and burden of disease. World Health Organization. https://apps.who.int/iris/handle/10665/250141

World Health Organization (WHO) (2021). Tracking SARS-COV-2 variants. https://www.who.int/en/activities/tracking-SARS-CoV-2-variants/.

Zhang, Z., Xue, T., & Jin, X. (2020). Effects of meteorological conditions and air pollution on COVID-19 transmission: Evidence from 219 Chinese cities. Science of the Total Environment, 741, 140244.

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.

Zoran, M. A., Savastru, R. S., Savastru, D. M., & Tautan, M. N. (2020). 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, 139.

Tables

Table 1. Pearson’s correlation coefficients between nine air pollutants and six potential sources of environmental pollution.

Air Pollutants	Large firms per km^2	Energy and natural gas consumption	Vehicles per km^2	Public transport	Cattle fodder	Livestock per km^2
Common air pollutants						
NO_2	0.5887***	0.6982***	0.5783***	0.5206***	0.2624***	0.6679***
$O_{3\ (>120)}$	0.5942***	0.4464***	0.1967*	0.2489**	0.4033***	0.7486***
$O_{3\ (>180)}$	0.3738***	0.3688***	0.1986*	0.1212	0.3334***	0.6605***
$PM_{2.5}$	0.5841***	0.517***	0.2963***	0.3373***	0.3096***	0.6732***
PM_{10}	0.4802***	0.4585***	0.3656***	0.3089***	0.2686***	0.4416***
$PM_{10\ (>50)}$	0.3922***	0.5247***	0.3348***	0.3347***	0.3497***	0.6557***
PHAs						
Benzene	0.2897***	0.4285***	0.5417***	0.3458***	-0.0006	0.0216
BaP	0.2152*	0.2114*	0.1057	0.1192	0.2357**	0.3373***
Heavy metals						
As	0.1001	0.1917	0.0998	0.1316	0.2362*	0.0807
Cd	0.0618	0.0651	0.0012	0.1523	-0.0557	-0.1798
Ni	-0.0106	0.216*	0.131	0.0197	0.2007	0.0689

Notes: p-value < 0.01***; p-value < 0.05**; p-value < 0.1*.

Table 2. Provinces that exceeded the EU legal threshold in the period 2014–2019.

Air Pollutants	EU legal threshold	National averages	Provinces with long-term violations	Provinces with at least 1-year violation
NO_2	40 µg/m^3	26.3279	11	15
O_3	>120 µg/m^3	28.2256	95	95
O_3	>180 µg/m^3	8.9064	58	58
$PM_{2.5}$	25 µg/m^3	15.3609	4	17
PM_{10}	40 µg/m^3	24.952	0	5
PM_{10}	>50 µg/m^3	25.1509	106	106
Benzene	5 µg/m^3	1.2069	0	0
BaP	1 ng/m^3	0.4363	7	13
As	6 ng/m^3	0.9559	0	0
Cd	5 ng/m^3	0.343	0	0
Ni	20 ng/m^3	3.6301	0	2

Source: https://ec.europa.eu/environment/air/quality/standards.htm.

Table 3. Provinces that exceeded the WHO AGQ threshold in the period 2014–2019.

Air Pollutants	WHO AQG threshold	National averages	Provinces with long-term violation	Province with at least 1-year violation
NO_2	40 µg/m^3	26.3279	11	15
$O_{3\ (8\ hours)}$	> 100 µg/m^3	28.2256*	95*	95*
$PM_{2.5}$	10 µg/m^3	15.3609	85	88
PM_{10}	20 µg/m^3	24.952	86	93
PM_{10}	> 50 µg/m^3	25.1509	106	106
Benzene	No safe level	1.2069	-	-
BaP	No safe level	0.4363	-	-
As	No safe level	0.9559	-	-
Cd	5 ng/m	0.343	0	0
Ni	No safe level	3.6301	-	-

Source: (WHO 2015).

Notes: *Due to a lack of data, these violations referred to the legal threshold limit of 120 µg/m^3.

Table 4. A synthetic environmental pollution index for the Italian provinces in the period 2014–2019.

Province	Index	Province	Index
1-Monza and Brianza	1.7639	55-Ascoli Piceno	0.8696
2-Bresca	1.6969	56-Rieti	0.8675
3-Milan	1.6681	57-Avellino	0.8572
4-Bergamo	1.6278	58-Caserta	0.8464
5-Lodi	1.6175	59-Bari	0.8408
6-Cremona	1.6097	60-Foggia	0.8315
7-Turin	1.5723	61-Livorno	0.8237
8-Pavia	1.5631	62-Pescara	0.8195
9-Mantua	1.5302	63-Perugia	0.8118
10-Alessandria	1.5261	64-Syracuse	0.8086
11-Piacenza	1.5229	65-Aosta	0.8058
12-Vicenza	1.512	66-Isernia	0.8012
13-Como	1.4881	67-Crotone	0.7981
14-Varese	1.4826	68-Teramo	0.7977
15-Genoa	1.4759	69-Pisa	0.7955
16-Venice	1.4727	70-Ancona	0.7936
17-Padua	1.451	71-Grosseto	0.7822
18-Modena	1.4293	72-Campobasso	0.7798
19-Verona	1.4267	73-Massa and Carrara	0.7749
20-Parma	1.4176	74-Benevento	0.7702
21-Treviso	1.3975	75-La Spezia	0.7599
22-Lecco	1.3836	76-Cosenza	0.7556
23-Reggio Emilia	1.3724	77-Siena	0.7352
24-Vercelli	1.30156	78-Cagliari	0.7232
25-Rovigo	1.2989	79-Latina	0.7228
26-Rimini	1.2815	80-Taranto	0.7076
27-Novara	1.2781	81-Savona	0.6869
28-Bologna	1.2741	82-Brindisi	0.6868
29-Ferrara	1.2291	83-Vibo Valentia	0.6826
30-Naples	1.2182	84-L'Aquila	0.6815
31-Frosinone	1.2172	85-Enna	0.6806
32-Trento	1.2082	86-Imperia	0.6743
33-Florence	1.2056	87-Salerno	0.6518
34-Terni	1.1402	88-Macerata	0.6335
35-Prato	1.0913	89-Barletta-Andria-T.	0.6168
36-Forlì-Cesena	1.0686	90-Viterbo	0.6166
37-Pordenone	1.0526	91-Catania	0.6118
38-Asti	1.0452	92-Lecce	0.574
39-Udine	1.0426	93-Pistoia	0.5671
40-Ravenna	1.0379	94-Potenza	0.5539
41-Chieti	1.0297	95-Reggio Calabria	0.5467
42-Cuneo	1.024	96-Ragusa	0.5457
43-Sondrio	0.988	97-Catanzaro	0.5352
44-Palermo	0.9861	98-Oristano	0.5269
45-Gorizia	0.9729	99-Caltanissetta	0.5237
46-Rome	0.9632	101-Messina	0.5131
47-Biella	0.9541	102-Agrigento	0.5065
48-Lucca	0.9168	103-Fermo	0.4961
49-Verbano-Cusio-O.	0.91	104-Sassari	0.496
50-Arezzo	0.9034	105-South Sardinia	0.4654
51-Pesaro e Urbino	0.892	105-Trapani	0.3985
52-Trieste	0.8747	106-Matera	0.3825
53-Belluno	0.8729	107-Nuoro	0.3741
54-Bolzano	0.8729		

Notes: the provinces are ranked from the most polluted to the cleanest. The 10 most polluted provinces are red colored, while the 10 cleanest provinces are green colored.

Table 5. 22 Selected studies on the relationship between air pollutants and the spread of COVID-19 across the world.

Author	Area	Air pollutants	Method	Results
Adhikari and Yin (2020)	New York (USA)	O_3, $PM_{2.5}$	OLS	O_3 (+), $PM_{2.5}$ (-)
Bashir et al. (2020)	California	CO, NO_2, Pb, $PM_{2.5}$, PM_{10}, SO_2, VOC	Kendall and Spearman correlation	CO (+), NO_2 (+), $PM_{2.5,}$ (+), PM_{10} (+), SO_2 (+)
Becchetti et al. (2020)	96 Italian provinces	NO_2, $PM_{2.5}$, PM_{10}	DID, OLS, SAC	NO_2 (+), $PM_{2.5,}$ (+), PM_{10} (+)
Bolaño-Ortiz et al. (2020)	10 big cities from Latin America and the Caribbean	NO_2, $PM_{2.5}$, PM_{10}	Spearman correlation	NO_2 (+), $PM_{2.5,}$ (+), PM_{10} (+)
Bontempi (2020)	Lombardy & Piedmont (Italy)	PM_{10}	Simple graphical analysis	The relation between PM_{10} and COVID-19 spread was not evident.
Coccia (2020)	55 Italian provinces	O_3 & PM_{10}	OLS, quadratic model	O_3 (+), PM_{10} (+)
Cole et al. (2020)	355 Dutch municipalities	$NO_{2,}$ $PM_{2.5}$, SO_2	IV, NB, SARAR	$PM_{2.5}$ (+), NO_2 (+) long-term
Delnevo et al. (2020)	Modena and Ravenna (Italy)	$PM_{2.5}$, PM_{10}	Granger causality	$PM_{2.5}$ (+), PM_{10} (+)
Fattorini and Regoli (2020)	71 Italian provinces	NO_2, O_3, $PM_{2.5}$, PM_{10}	Pearson correlation	NO_2 (+), O_3 (+), $PM_{2.5}$ (+), PM_{10} (+)
Filippini et al. (2020)	28 Italian provinces (Emilia-Romagna, Lombardy, Veneto)	NO_2	Multivariable RCS regression	NO_2 (+)
Hendryx and Luo (2020)	3,143 USA counties	DPM, O_3, $PM_{2.5}$	Mixed linear multiple regression	DPM (+), $PM_{2.5}$ (+)
Li et al. (2020)	Wuhan and XiaoGan (China)	AQI, CO, NO_2, $PM_{2.5}$, PM_{10}	Pearson correlation	AQI (+), $PM_{2.5}$ (+), and NO_2 (+)
Lin et al. (2020)	29 China provinces	AQI, CO, NO_2, O_3, $PM_{2.5}$, PM_{10}, SO_2	Pearson/Spearman correlation	CO (+), NO_2 (-)
Lolli et al. (2020)	Florence, Milan, Trento (Italy)	$PM_{2.5}$	Spearman/Kendall correlation	$PM_{2.5}$ (+)
Setti et al. (2020)	Bergamo (Italy)	PM_{10}	Molecular markers	Molecular marker genes (E, N, and RdRP) confirmed the presence of traces of SARS-CoV-2 RNA on PM_{10}.

Vasquez-Apestegui et al. (2020)	24 Districts of Metropolitan Lima (Peru)	$PM_{2.5}$	Pearson correlation	Long-term $PM_{2.5}$ (+)
Zhu et al. (2020)	120 Chinese cities	NO_2, O_3, $PM_{2.5}$, PM_{10}, SO_2	GAM	NO_2 (+), O_3 (+), $PM_{2.5}$ (+), PM_{10} (+), SO_2 (-)
Zhang et al. (2020)	2,019 Chinese cities	AQI	Spearman/Kendall correlation, OLS	AQI (+)
Zoran et al. (2020)	Milan (Italy)	AQI, $PM_{2.5}$, PM_{10}	Pearson correlation	AQI (+), $PM_{2.5}$ (+), PM_{10} (+)
De Angelis et al. (2021)	1,439 municipalities of Lombardy (Italy)	NO_2, $PM_{2.5}$, PM_{10}	$NBMM_s$	NO_2 (-), $PM_{2.5}$ (+), PM_{10} (+),
Solimini et al. (2021)	730 regions, 63 countries, 5 continents	$PM_{2.5}$, PM_{10}	$NBMM_s$	$PM_{2.5}$ (+), PM_{10} (+)
Travaglio et al. (2021)	England (sub-regional and individual data)	NO_2, NO_x, $PM_{2.5}$, PM_{10}	GLM, NB regression	Individual data: NO_2 (+), NO_x (+), $PM_{2.5}$ (+), PM_{10} (+). Sub-regional data: NO_2 (+), NO_x (+)

Notes: AIQ, air quality index; DID, difference-in-difference; FE, fixed effect; GAM, generalized additive model; GLM, generalized least model; IV, instrumental variables; NB, negative binomial; $NBMM_s$, negative binomial mixed effect; OLS, ordinary least square; RCS, restricted cubic spline; RdRP, RNA-dependent RNA polymerase gene; SAC, spatial autoregressive combined models.

Table 6. Definition of variables used in the empirical analysis.

Variables	Definitions	Sources
Dependent variables		
Confirmed cases	The number of COVID-19 cumulative cases in each province, on 30 November 2020.	Italian Ministry of Health (2021)
Prevalence rate	The ratio between people who have been tested positive for COVID-19 on 30 November 2020 (and on 20 February 2021), and total resident population on 1 January 2020.	I.Stat (2021a), Italian Ministry of Health (2021)
Independent variables		
AUT border	A dummy that takes 1 when the province borders Austria and 0 elsewhere.	Google Maps
FRA border		Google Maps
SLO border	A dummy that takes 1 when the province borders France and 0 elsewhere.	Google Maps
SWI border		Google Maps
Airport distance	A dummy that takes 1 when the province borders Slovenia and 0 elsewhere.	Assoaeroporti (2021),
Foreigners	A dummy that takes 1 when the province borders Switzerland and 0 elsewhere.	www.michelin.it I.Stat (2021a)
Aged 0-19	The distance in kilometers between the provincial capital's center and the nearest airport with at least 50,000 passengers over the period January–November 2020.	I.Stat (2021a)
Male	The foreign-born population measured as a percentage of the total resident population in each province on 1 January 2020.	I.Stat (2021a)
Population Density	The percentage of the resident population aged 0-19 in each province on 1 January 2020.	Eurostat (2013)
Urbanization	The percentage of resident population that is male on 1 January 2020.	
	The number of inhabitants per square kilometer of land area in each province, on 1 January 2020.	
	An ordinal index that ranks population of each province by urban-rural structure: predominantly rural (1), intermediate (2), and predominantly urban (3).	
Alcohol	The average percentage of alcohol drinkers in each region, in the period 2016–2019.	ISS (2021)
Obesity	The average percentage of obese individuals in each region, in period 2016–2019.	ISS (2021)
Smokers	The average percentage of smokers in each region, in the period 2016–2019.	ISS (2021)
Lung disease	The average deaths from chronic respiratory disease (per 100,000 inhabitants) in each province, in the period 2014–2019.	I.Stat (2021a)
Beds/infected	The ratio between people who have been tested positive for COVID-19 on 30 November 2020 and average ordinary hospital beds in the period 2017–2018.	Italian Ministry of Health (2018, 2020)
Large firms	The percentage of firms that employed 250 or more employees in each province, in the period 2014–2018.	I.Stat (2021a)
Capitals	A dummy that takes 1 when the province is also the regional capital and 0 elsewhere.	
Altitude	The average altitude of the provincial capital.	I.Stat (2021a)
Rainy days	The average annual days of rain in the provincial capital, in the period 2007–2018.	Istat (2021d)
Temperature	The average annual temperature of provincial capital, in the period 2009–2018.	Istat (2020a) Mipaaf (2021)
NO_2	The average concentrations of nitrogen dioxide, expressed in micrograms per cubic meter of air (μg/m^3), in the period 2014–2019.	Istat (2015, 2017, 2019),
$O_{3\ (>120)}$	The average number of days in which ozone exceeded the limit of 120 micrograms per cubic meter of air (μg/m^3), in the period 2014–2019.	Legambiente (2020)
$O_{3\ (>180)}$	The average number of hours in which ozone exceeded the limit of 180 micrograms per cubic meter of air (μg/m^3) in each province, in the period 2014–2018.	Istat (2015, 2017 2019), Legambiente (2020)
$PM_{2.5}$	The average concentrations of particulate matter less than 2.5 micrometers in diameter, expressed in micrograms per cubic meter of air (μg/m^3), in the period 2014–2019.	Istat (2015, 2017, 2019)
		Istat (2015, 2017, 2019), Legambiente (2020)

PM_{10}	The average concentrations of particulate matter less than 10 micrometers in diameter, expressed in micrograms per cubic meter of air ($\mu g/m^3$), in the period 2014–2019.	Istat (2015, 2017, 2019), Legambiente (2020)
PM_{10} (>50)	The average number of days in which PM_{10} exceeded the limit of 50 micrograms per cubic meter of air ($\mu g/m^3$), in the period 2014–2018.	Istat (2015, 2017, 2019)
Benzene	The average concentrations of benzene, expressed in nanogram per cubic meter of air ($\mu g/m^3$), in the period 2014-2016.	ISPRA (2015, 2016, 2017)
BaP	The average concentrations of benzo[a]pyrene, expressed in expressed in nanogram per cubic meter of air (ng/m^3), in the period 2014-2018.	Istat (2015, 2017, 2019)
As	The average concentrations of arsenic, expressed in nanogram per cubic meter of air (ng/m^3), in the period 2014–2016.	ISPRA (2015, 2016, 2017)
Cd	The average concentrations of cadmium, expressed in nanogram per cubic meter of air (ng/m^3), in the period 2014–2016.	ISPRA (2015, 2016, 2017)
Ni	The average concentrations of nickel, expressed in nanogram per cubic meter of air (ng/m^3), in the period 2014–2016.	ISPRA (2015, 2016, 2017)

Table 7 (part A). Results from negative binomial regressions on COVID-19 cumulative cases registered on 30 November 2020.

Variables	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
AUT border	-0.044	0.0751	0.1889	0.1373	0.0815	0.1366
	[0.1351]	[0.1357]	[0.1304]	[0.1354]	[0.1321]	[0.1291]
FRA border	-0.0842	0.058	0.1087	0.0803	0.0412	0.0618
	[0.1085]	[0.1064]	[0.1017]	[0.1041]	[0.1042]	[0.1007]
SLO border	-0.0899	-0.0131	-0.0882	-0.0991	0.0181	0.0405
	[0.1189]	[0.1262]	[0.1173]	[0.1472]	[0.1213]	[0.1198]
SWI border	0.1176	0.2767***	0.3266***	0.3048***	0.2415**	0.3161***
	[0.1026]	[0.0984]	[0.0931]	[0.0957]	[0.0966]	[0.0927]
Aged 0-19	0.0022	0.009	-0.0009	0.0017	0.0336	0.0129
	[0.0197]	[0.0205]	[0.0204]	[0.0208]	[0.0214]	[0.0194]
Airport dis.	-0.0014**	-0.0013***	-0.0014**	-0.0013**	-0.001	-0.0011*
	[0.0006]	[0.0006]	[0.0006]	[0.0006]	[0.0006]	[0.0006]
Foreigners	0.0226***	0.0229***	0.0255***	0.0281***	0.0165*	0.0221***
	[0.008]	[0.0084]	[0.0094]	[0.0095]	[0.0084]	[0.0079]
Male	-0.0235	-0.0352	-0.0451	-0.0382	-0.0989	-0.0839
	[0.0546]	[0.0573]	[0.0591]	[0.0632]	[0.0611]	[0.0548]
Pop. Density	0.0002***	0.0002**	0.0002***	0.0002***	0.0001**	0.0002***
	[0.0001]	[0.0001]	[0.0001]	[0.0001]	[0.0001]	[0.0001]
Urbanization	-0.0165	-0.0485	-0.0235	-0.0289	-0.0391	-0.0432
	[0.0372]	[0.0394]	[0.0384]	[0.0427]	[0.0397]	[0.0365]
Lung disease	0.0467	0.0517*	0.0322	0.0492	0.0564*	0.0504*
	[0.0293]	[0.0308]	[0.0301]	[0.0307]	[0.033]	[0.0291]
Beds/infected	-2.746***	-2.8647***	-2.5472***	-2.622***	-2.8946***	-2.7838***
	[0.2703]	[0.2763]	[0.2807]	[0.2978]	[0.2797]	[0.2639]
Large firms	-0.005	0.0029	0.0043	0.0071	-0.0014	-0.0019
	[0.0086]	[0.0085]	[0.0089]	[0.0094]	[0.0087]	[0.0082]
Altitude	0.0001	0.0004***	0.0003***	0.0003***	0.0004***	0.0005***
	[0.0001]	[0.0001]	[0.0001]	[0.0001]	[0.0001]	[0.0001]
Rainy days	-0.0029	-0.001	-0.0007	-0.0008	-0.0045	0.0001
	[0.0019]	[0.0019]	[0.0018]	[0.0019]	[0.0019]	[0.0018]
Temperature	-0.0625***					
	[0.0151]					
No_2		0.0078***				
		[0.0029]				
$O_{3\ (>120)}$			0.0035***			
			[0.0013]			
$O_{3\ (>180)}$				0.0016		
				[0.0014]		
$PM_{2.5}$					0.0188***	
					[0.0046]	
PM_{10}						0.0188***
						[0.0042]
Pseudo R	0.1047	0.1004	0.1094	0.1058	0.1038	0.1059
N	107	107	98	95	97	107
LR test (p-v.	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000

Notes: p-value < 0.01***; p-value < 0.05**; p-value < 0.1*. All models included a dummy for regional capitals and controls for alcohol drinkers, smokers, and obese individuals.

Table 7 (part B). Results from negative binomial regressions on COVID-19 cumulative cases registered on 30 November 2020.

Variables	Model 7	Model 8	Model 9	Model 10	Model 11	Model 12
AUT border	0.1507	-0.2965*	0.1905	-0.1308	-0.1756	-0.1895
	[0.131]	[0.1758]	[0.1231]	[0.177]	[0.1673]	[0.1653]
FRA border	0.0559	0.1432	0.1061	0.317***	0.3063**	0.347***
	[0.102]	[0.1166]	[0.1071]	[0.12]	[0.1198]	[0.1226]
SLO border	-0.0136	0.1221	-0.1391	0.1129	0.136	0.1478
	[0.1193]	[0.1596]	[0.1299]	[0.1574]	[0.1562]	[0.1558]
SWI border	0.2979***	0.2884***	0.2861***	0.4054***	0.4293***	0.4297***
	[0.0938]	[0.1065]	[0.0998]	[0.1009]	[0.1028]	[0.0994]
Aged 0-19	0.0138	0.0602***	-0.041	0.0061	0.0104	0.0169
	[0.0197]	[0.0216]	[0.0253]	[0.023]	[0.0223]	[0.0221]
Airport dis.	-0.001*	-0.0007	-0.0011*	0.0009	0.0008	0.0013*
	[0.0006]	[0.0007]	[0.0006]	[0.0007]	[0.0008]	[0.0007]
Foreigners	0.0215***	0.0237***	0.0252***	0.0224**	0.0216*	0.0244**
	[0.008]	[0.0081]	[0.0084]	[0.0112]	[0.0112]	[0.0113]
Male	-0.0915	-0.1438**	0.0589	-0.1971***	-0.1939***	-0.2135***
	[0.0563]	[0.0606]	[0.0712]	[0.0712]	[0.0731]	[0.0704]
Pop. Density	0.0001**	0.0002***	0.0002***	0.0001	0.0001	0.0001
	[0.0001]	[0.0001]	[0.0001]	[0.0001]	[0.0001]	[0.0001]
Urbanization	-0.0316	-0.0939**	0.027	-0.049	-0.0577	-0.0469
	[0.0369]	[0.0412]	[0.0431]	[0.0483]	[0.0482]	[0.0478]
Lung disease	0.0396	0.0474	0.0433	-0.0666*	-0.0651	-0.0752*
	[0.0294]	[0.03]	[0.0332]	[0.04]	[0.0417]	[0.0393]
Beds/infected	-2.9141***	-2.5812***	-2.5642***	-2.8808***	-2.8403***	-2.8695***
	[0.266]	[0.2687]	[0.2673]	[0.3665]	[0.3865]	[0.3625]
Large firms	-0.0026	0.0045	0.0154*	0.0264**	0.0274**	0.0224**
	[0.0084]	[0.0084]	[0.0091]	[0.0107]	[0.0113]	[0.0108]
Altitude	0.0004***	0.0003**	0.0004***	0.0001	0.0001	0.0001
	[0.0001]	[0.0001]	[0.0001]	[0.0001]	[0.0002]	[0.0001]
Rainy days	-0.0003	-0.0001	-0.0021	-0.0024	-0.0024	-0.003
	[0.0019]	[0.0019]	[0.0019]	[0.0023]	[0.0023]	[0.0023]
$PM_{10\ (>50)}$	0.0042***					
	[0.001]					
Benzene		0.1174**				
		[0.0455]				
BaP			0.1*			
			[0.0594]			
As				0.032		
				[0.0372]		
Cd					0.0348	
					[0.066]	
Ni						-0.0133
						[0.0097]
Pseudo R	0.1045	0.1122	0.1235	0.1168	0.1164	0.1176
N	107	88	73	60	60	60
LR test	0.000	0.0000	0.0000	0.0000	0.000	0.000

Notes: p-value < 0.01***; p-value < 0.05**; p-value < 0.1*. All models included a constant, a dummy for regional capitals, and controls for alcohol drinkers, smokers, and obese individuals. Due to collinearity, alcohol drinkers were excluded in models 10, 11, and 12.

Table 8. The average marginal effects get from negative binomial regressions.

	NO_2	$PM_{2.5}$	PM_{10}	Benzene	BaP
	1 µg/m^3 increase			0.1 µg/m^3 increase	0.1 ng/m^3 increase
Marginal effects	117.1943** [46.63]	297.76*** [72.45]	282.88*** [64.43]	193.15** [75.63]	166.53** [75.05]
95% Conf. interval	25.8 – 208.59	155.77 – 439.76	156.59 – 409.16	45.64 – 340.67	19.52 – 313.54

Notes: p-value < 0.01***; p-value < 0.05**.

Table 9 (part A). Results from OLS models on COVID-19 prevalence rate registered on 30 November 2020.

Variables	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
AUT border	0.3017	0.539	0.7522	0.6557	0.581	0.7155*
	[0.4664]	[0.4847]	[0.4926]	[0.5026]	[0.4186]	[0.4067]
FRA border	0.1064	0.3932	0.5217*	0.486	0.3407	0.3982
	[0.2841]	[0.2616]	[0.2991]	[0.3439]	[0.2808]	[0.2515]
SLO border	-0.3872*	-0.1974	-0.3704	-0.4566	-0.1599	-0.0957
	[0.2162]	[0.2535]	[0.302]	[0.3603]	[0.2717]	[0.2659]
SWI border	0.7822**	1.1121***	1.1985***	1.1975***	1.0315***	1.2232***
	[0.3487]	[0.2263]	[0.2286]	[0.2249]	[0.2076]	[0.2]
Aged 0-19	-0.0175	-0.004	-0.0188	-0.0174	0.0393	0.0012
	[0.0578]	[0.0644]	[0.0579]	[0.0638]	[0.067]	[0.059]
Airport dis.	-0.0036**	-0.0034**	-0.0032*	-0.0034*	-0.0026*	-0.003*
	[0.0015]	[0.0016]	[0.0016]	[0.0017]	[0.0015]	[0.0015]
Foreigners	0.0639***	0.0638***	0.0642**	0.0724***	0.0496**	0.0628***
	[0.0202]	[0.02]	[0.0246]	[0.0268]	[0.0216]	[0.0195]
Male	0.0753	0.0628	0.0068	-0.0106	-0.0401	-0.0365
	[0.1756]	[0.1944]	[0.2004]	[0.2259]	[0.2126]	[0.1685]
Pop. Density	0.0008***	0.0006***	0.0007***	0.0007***	0.0005***	0.0006***
	[0.0002]	[0.0002]	[0.0002]	[0.0001]	[0.0002]	[0.0002]
Urbanization	0.0217	-0.0469	-0.0139	-0.062	-0.0127	-0.0306
	[0.1087]	[0.1128]	[0.1202]	[0.1379]	[0.1167]	[0.106]
Lung disease	0.1931**	0.2103**	0.1539*	0.1837**	0.2378**	0.2096**
	[0.0853]	[0.0904]	[0.0901]	[0.0883]	[0.101]	[0.0885]
Beds/infected	-4.3731***	-4.5036***	-3.8405***	-4.1506***	-4.7116***	-4.3196***
	[0.9185]	[1.0157]	[0.7985]	[0.7401]	[0.9585]	[0.8715]
Large firms	0.0014	0.0155	0.0136	0.0277	0.0087	0.0039
	[0.0212]	[0.0177]	[0.0212]	[0.0234]	[0.0201]	[0.0168]
Altitude	0.0005*	0.001***	0.0008***	0.0008***	0.0011***	0.0012***
	[0.0003]	[0.0003]	[0.0003]	[0.0003]	[0.0003]	[0.0003]
Rainy days	-0.0101*	-0.0064	-0.0061	-0.0048	-0.0041	-0.0037
	[0.0056]	[0.0058]	[0.0048]	[0.0051]	[0.0056]	[0.005]
Temperature	-0.1283***					
	[0.0457]					
No_2		0.0196**				
		[0.0076]				
$O_{3\ (>120)}$			0.0097**			
			[0.0037]			
$O_{3\ (>180)}$				0.0053		
				[0.0051]		
$PM_{2.5}$					0.0431***	
					[0.012]	
PM_{10}						0.0453***
						[0.0104]
R-square	0.8229	0.8165	0.8429	0.8269	0.8257	0.8631
N	107	107	98	95	97	107
F-test	25.63***	24.59***	27.01***	23.45***	23.74***	28.04***
VIF (range)	1.37-6.02	1.32-3.77	1.35-3.97	1.36-3.8	1.37-3.89	1.32–3.77

Notes: p-value < 0.01***; p-value < 0.05**; p-value < 0.1*. Notes: p-value < 0.01***; p-value < 0.05**; p-value < 0.1*. All models include a constant, a dummy for regional capitals, and controls for alcohol drinkers, smokers, and obese individuals.

Table 9 (part B). Results from OLS models on COVID-19 prevalence rate registered on 30 November 2020.

Variables	Model 7	Model 8	Model 9	Model 10	Model 11	Model 12
AUT border	0.7357*	-0.3681	0.8067**	0.0279	-0.0147	-0.0507
	[0.4335]	[0.3731]	[0.3243]	[0.3936]	[0.3651]	[0.3895]
FRA border	0.3794	0.7743***	0.6809***	1.067***	1.0507***	1.0775***
	[0.2574]	[0.278]	[0.1974]	[0.223]	[0.2128]	[0.2286]
SLO border	-0.1899	-0.0645	-0.6107**	-0.0882	-0.0466	-0.0503
	[02652]	[0.31]	[0.2426]	[0.3032]	[0.2843]	[0.3142]
SWI border	1.1752***	1.1262***	1.0214***	1.3727***	1.4442***	1.3966***
	[0.2054]	[0.2489]	[0.3056]	[0.2329]	[0.2339]	[0.2342]
Aged 0-19	0.0107	0.0962	-0.1027	-0.0061	-0.003	0.0041
	[0.059]	[0.0584]	[0.0668]	[0.07]	[0.0712]	[0.0692]
Airport dis.	-0.0027*	-0.0016	-0.0023	0.0012	0.0007	0.0015
	[0.0016]	[0.0019]	[0.0015]	[0.0016]	[0.0018]	[0.0019]
Foreigners	0.0603***	0.0688***	0.0651***	0.0502*	0.049*	0.051*
	[0.0189]	[0.0206]	[0.0175]	[0.0253]	[0.0256]	[0.0266]
Male	-0.0837	-0.1741	0.1853	-0.2866	-0.2582	-0.2965
	[0.1808]	[0.1953]	[0.1755]	[0.2288]	[0.2374]	[0.2323]
Pop. Density	0.0006***	0.0007***	0.001***	0.0007*	0.0007**	0.0007**
	[0.0002]	[0.0002]	[0.0003]	[0.0003]	[0.0003]	[0.0003]
Urbanization	-0.0148	-0.1707	0.0488	-0.1106	-0.1331	-0.1137
	[0.1059]	[0.1217]	[0.1154]	[0.1538]	[0.1485]	[0.152]
Lung disease	0.1832**	0.1817**	0.2072**	-0.0206	0.0042	-0.0311
	[0.0866]	[0.0834]	[0.0807]	[0.1075]	[0.1136]	[0.1019]
Beds/infected	-4.6259***	-4.0403***	-4.0097***	-4.9935***	-4.7193***	-5.005***
	[0.8618]	[0.9]	[0.7222]	[1.3218]	[1.3171]	[1.2906]
Large firms	0.0000	0.0162	0.0295	0.0587*	0.0658*	0.0556*
	[0.0173]	[0.0194]	[0.0269]	[0.0324]	[0.0344]	[0.0324]
Altitude	0.0011***	0.0007*	0.001***	0.0003	0.0003	0.0003
	[0.0003]	[0.0004]	[0.0003]	[0.0003]	[0.0003]	[0.0003]
Rainy days	-0.0045	-0.0036	-0.0071	-0.0053	-0.0051	-0.0059
	[0.005]	[0.0055]	[0.0058]	[0.0058]	[0.0058]	[0.0058]
$PM_{10\ (>50)}$	0.0111***					
	[0.0025]					
Benzene		0.2468**				
		[0.1172]				
BaP			0.2937**			
			[0.1422]			
As				0.0472		
				[0.1011]		
Cd					0.1565	
					[0.2312]	
Ni						-0.0084
						[0.0177]
R-square	0.8361	0.8323	0.8725	0.8639	0.8658	0.8635
N	107	88	73	60	60	60
F-test	28.04***	22.59***	25.64***	20.71***	21.03***	20.65***
VIF (range)	1.32-3.79	1.49-4.53	1.53-4.3	1.64-4.12	1.68-4.5	1.62-4.32

Notes: p-value < 0.01***; p-value < 0.05**; p-value < 0.1*. All models included a constant, a dummy for regional capitals, and controls for alcohol drinkers, smokers, and obese individuals. Due to collinearity, alcohol drinkers were excluded in models 10, 11, and 12.

Table 10 (part A). Results from SARAR models on COVID-19 prevalence rate registered on 30 November 2020.

Variables	Model 1	Model 2	Model 3	Model 4	Model 5
ρ	0.9021***	0.8689***	0.8752***	0.88***	0.8855***
	[0.0941]	[0.117]	[0.1178]	[0.111]	[0.1078]
λ	-0.0465	-0.6803	0.2826	0.0361	0.1268
	[0.8365]	[0.9604]	[0.6648]	[0.755]	[0.7401]
NO_2	0.0124*				
	[0.0073]				
$O_{3\ (>120)}$		0.0053*			
		[0.0032]			
$PM_{2.5}$			0.0295**		
			[0.0119]		
PM_{10}				0.0366***	
				[0.0104]	
$PM_{10\ (>50)}$					0.0074***
					[0.0026]
Pseudo R^2	0.8326	0.8641	0.8271	0.8419	0.8346
N	107	98	97	107	107
Wald spatial terms	96.35***	55.46***	61.34***	66.34***	72.36***

Notes: p-value < 0.01***; p-value < 0.05**; p-value < 0.1*. All models included a constant and the following controls: dummies for regional capitals and national borders, population aged 0–19, distance from nearest airport, share of foreigners, share of male population, population density, degree of urbanization, deaths due to lung disease, alcohol drinkers, smokers, obese individuals, large firms, altitude, and rainy days.

Table 10 (part B). Results from SARAR models on COVID-19 prevalence rate registered on 30 November 2020.

Variables	Model 6	Model 7	Model 8	Model 9	Model 10
ρ	0.8882***	0.8129***	0.8462***	0.8257***	0.8712***
	[0.1072]	[0.1529]	[0.1339]	[0.1409]	[0.1158]
λ	0.2147	-1.5979	-1.3417	-1.6609	-1.2576
	[0.707]	[1.1238]	[1.1228]	[1.1024]	[1.1591]
Benzene	0.249**				
	[0.1086]				
BaP		0.2782*			
		[0.1453]			
As			0.046		
			[0.0892]		
Cd				0.3324**	
				[0.1432]	
Ni					0.0219
					[0.0249]
Pseudo R^2	0.8578	0.8819	0.9066	0.9126	0.9081
N	88	73	88	60	60
Wald spatial terms	78.62***	29.68***	40.06***	35.9***	56.63***

Notes: p-value < 0.01***; p-value < 0.05**; p-value < 0.1*. All models included a constant and the following controls: dummies for regional capitals and national borders, population aged 0–19, distance from nearest airport, share of foreigners, share of male population, population density, degree of urbanization, deaths due to lung disease, alcohol drinkers, smokers, obese individuals, large firms, altitude, and rainy days. Due to collinearity, alcohol drinkers were excluded in models 8, 9, and 10.

Table 11. Direct and indirect effects of air pollutants after fitting SARAR models (30 November 2020).

Air pollutants	Direct	Indirect	Total
NO_2	0.0135	0.1131	0.1266
$O_{3\ (>120)}$	0.0057	0.0351	0.0408
$PM_{2.5}$	0.0317	0.2044	0.2361
PM_{10}	0.0393	0.2661	0.3053
$PM_{10\ (>50)}$	0.008	0.0569	0.0648
Benzene	0.2724	1.954	2.2265
BaP	0.2948	1.1917	1.4865
As	0.0504	0.2488	0.2992
Cd	0.3598	1.5479	1.9077
Ni	0.0245	0.1459	0.1704

Table 12 (part A) Results from SARAR models on COVID-19 prevalence rate registered on 20 February 2021.

Variables	Model 1	Model 2	Model 3	Model 4	Model 5
ρ	0.8172***	0.7717***	0.7889***	0.7862***	0.7905***
	[0.1636]	[0.1963]	[0.1845]	[0.1841]	[0.1822]
λ	0.4782	0.264	-0.0368	0.3552	0.3727
	[0.449]	[0.5822]	[0.7313]	[0.5247]	[0.5161]
No_2	0.02*				
	[0.0114]				
$O_{3\ (>120)}$		0.0022			
		[0.0055]			
$PM_{2.5}$			0.0562***		
			[0.0192]		
PM_{10}				0.0453***	
				[0.0174]	
$PM_{10\ (>50)}$					0.0099**
					[0.0044]
Pseudo R^2	0.7777	0.7978	0.8012	0.7883	0.7846
N	107	98	97	107	107
Wald spatial terms	28.75***	17.12***	18.89***	20.18***	21***

Notes: p-value < 0.01***; p-value < 0.05**; p-value < 0.1*. All models included a constant and the following controls: dummies for regional capitals and national borders, population aged 0–19, distance from nearest airport, share of foreigners, share of male population, population density, degree of urbanization, deaths due to lung disease, alcohol drinkers, smokers, obese individuals, large firms, altitude, and rainy days.

Table 12 (part B). Results from SARAR models on COVID-19 prevalence rate registered on 20 February 2021.

Variables	Model 6	Model 7	Model 8	Model 9	Model 10
ρ	0.706***	0.7168***	0.5674*	0.4635	0.6498**
	[0.2307]	[0.2328]	[0.3366]	[0.3593]	[0.2868]
λ	0.1519	-0.3531	-0.3795	-0.7666	-0.5084
	[0.64]	[0.929]	[1.1857]	[1.2722]	[1.3028]
Benzene	0.3537*				
	[0.1818]				
BaP		0.9559***			
		[0.2826]			
As			0.1908		
			[0.1819]		
Cd				0.7641***	
				[0.32]	
Ni					-0.0166
					[0.0504]
Pseudo R^2	0.8038	0.8345	0.8141	0.8298	0.8124
N	88	73	60	60	60
Wald spatial terms	10.55***	9.48***	2.95	1.71	5.29*

Notes: p-value < 0.01***; p-value < 0.05**; p-value < 0.1*. All models include a constant and the following controls: dummies for regional capitals and national borders, population aged 0–19, distance from nearest airport, share of foreigners, share of male population, population density, degree of urbanization, deaths due to lung disease, alcohol drinkers, smokers, obese individuals, large firms, altitude, and rainy days.

Table 13. Direct and indirect effects of air pollutants after fitting SARAR models (20 February 2021).

Air pollutants	Direct	Indirect	Total
NO_2	0.0209	0.0885	0.1094
$O_{3\ (>120)}$	0.0023	0.0074	0.0097
$PM_{2.5}$	0.0584	0.2079	0.2664
PM_{10}	0.0469	0.1648	0.2117
$PM_{10\ (>50)}$	0.0103	0.037	0.0472
Benzene	0.3634	0.8395	1.2029
BaP	0.9878	2.3876	3.3754
As	0.1945	0.2465	0.441
Cd	0.7727	0.6515	1.4242
Ni	-0.0171	-0.0303	-0.0474

Figures

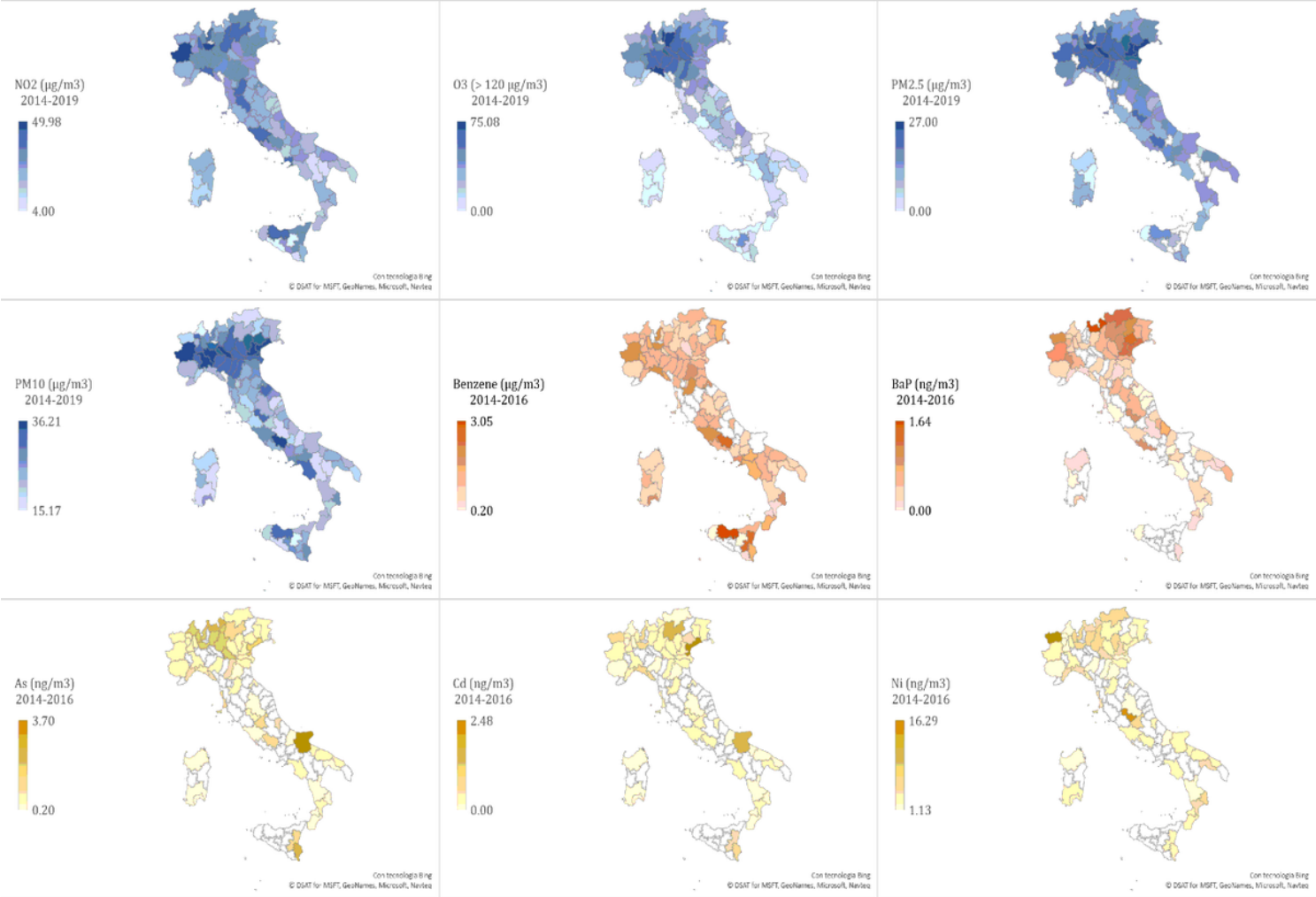

Figure 1

Average long-term outdoor concentrations (or violations) of NO2, O3, PM2.5, PM10, benzene, BaP, As, Cd, and Ni, in the 107 Italian provinces. Notes: when no data are available, the province is light grey colored.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- AppendixAandB.docx