Air pollution impacts of COVID-19–related containment measures

Guillaume P. Chossière¹, Haofeng Xu¹, Yash Dixit¹, Stewart Isaacs¹, Sebastian D. Eastham¹,², Florian Allroggen¹,², Raymond L. Speth¹,², Steven R. H. Barrett¹,²,³*

Responses to the COVID-19 outbreak resulted in one of the largest short-term decreases in anthropogenic emissions in modern history. To date, there has been no comprehensive assessment of the impact of lockdowns on air quality and human health. Using global satellite observations and ground measurements from 36 countries in Europe, North America, and East Asia, we find that lockdowns led to reductions in NO2 concentrations globally, resulting in ~32,000 avoided premature mortalities, including ~21,000 in China. However, we do not find corresponding reductions in PM2.5 and ozone globally. Using satellite measurements, we show that the disconnect between NO2 and ozone changes stems from local chemical regimes. The COVID-related lockdowns demonstrate the need for targeted air quality policies to reduce the global burden of air pollution, especially related to secondary pollutants.

INTRODUCTION

In an effort to reduce the spread of coronavirus disease 2019 (COVID-19), governments around the world imposed restrictions on both social life and economic activity. By the end of March 2020, 76% of the global population lived in countries with stay-at-home orders; workplace closures were in place across countries that collectively generated 99% of global gross domestic product (GDP) in 2018; and 92% of the global population lived in countries with school closures (1, 2). The result was an unprecedented reduction in economic activity, as measured by industrial production [−27% year-on-year (YOY) in April 2020 in the Euro area (3); −15% YOY in April 2020 for the United States (4)] and by production of services [−12% YOY in March 2020 in China (5)]. Furthermore, mobility declined abruptly both for long-distance air travel [−94% YOY in global air transport revenue passenger kilometers in April 2020 (6)] and for surface transportation [−74% in public transit in major cities (7) and 40 to 80% reductions in car usage (8)].

These changes in economic activity reduced energy use and anthropogenic emissions (9, 10). Le Quéré et al. (11) found that global daily anthropogenic CO2 emissions declined by 17% in early April 2020 compared to average year-2019 levels, largely due to reductions in surface transportation emissions (−36%) in daily CO2 emissions in early 2020 compared to year-2019 mean levels. In addition, satellite data suggested reductions in emissions of air pollutants such as nitrogen oxides (NOx) [e.g., up to 40% reduction in observed nitrogen dioxide (NO2) concentrations in China (10)].

Reductions in emissions of primary pollutants such as NO2 are expected to also yield reductions in ambient concentrations of secondary pollutants such as fine particulate matter (PM2.5) and ozone. Investigations into this effect have shown inconsistent results. In Europe, the European Space Agency (ESA) reported that, of these pollutants, only NO2 concentrations substantially decreased (12). Wang et al. (13) used air quality modeling in China and found decreases in PM2.5 in January and February 2020. Using monitor data for January and March 2020, Shi and Brasseur (14) found decreases in NO2 and PM2.5 in China. Elsewhere, Chauhan and Singh (15) reported decreases in PM2.5 in nine cities around the world between January and March 2020, while Berman and Ebisu (16) found statistically significant decreases in NO2 and PM2.5 in the United States between January and April. However, these studies were focused on the regional effects of lockdown measures on pollution concentrations. They were not able to compare the varying impacts of lockdown measures on air quality across countries on a consistent basis.

The epidemiological literature has established that changes in concentrations of ozone and PM2.5 affect mortality rates (17–22). In addition, several studies have shown that increased exposure to NO2 can itself result in negative health outcomes, independent of local concentrations of other pollutants (23–25). The relationships between exposure to pollution and mortality identified by these studies are routinely used to estimate the negative health impacts of human activity (26–29). For example, long-term exposure to ozone and PM2.5 resulting from human activity is estimated to cause ~3.8 million premature mortalities each year (30). For COVID-19–related lockdowns, Giani et al. (31) estimated that ~24,000 premature deaths were avoided in China because of reductions in PM2.5 concentrations in February and March and ~2200 in Europe between February and May. In the medium term, Liu et al. (32) suggested that between 99,000 and 147,000 premature mortalities in 76 countries could be avoided due to the lockdown measures between January and July. Chen et al. (33) estimated that the lockdown measures in China resulted in ~8900 avoided mortalities from NO2 reductions and ~3200 from PM2.5 reductions from 23 January to 14 March. However, most existing studies neglected confounding factors such as interannual trends and seasonal variation. In countries that have seen rapid declines in emissions over the past years, such as China (34), ignoring these factors may lead to an overestimate of the air quality benefits of lockdown measures. In particular, it is unclear whether variations between the pollution and health outcomes in each country are the result of regional atmospheric conditions, lockdown stringency, or the method used to estimate what a “no-COVID” 2020 would have looked like. In addition, existing studies largely focus on local changes and, while valuable, do not allow a global comparison of changes in emissions and impacts across countries.

¹Laboratory for Aviation and the Environment, Department of Aeronautics and Astronautics, Massachusetts Institute of Technology, Cambridge, MA, USA. ²Joint Program on the Science and Policy of Global Change, Massachusetts Institute of Technology, Cambridge, MA, USA. ³Department of Mechanical Engineering, Seoul National University, Seoul, South Korea.
*Corresponding author. Email: sbarrett@mit.edu
This study produces the first consistent, global estimate of the short-term air quality and human health impacts of the lockdown measures implemented in response to the COVID-19 pandemic. We develop an approach using air quality monitoring data to evaluate the short-term impact of COVID-related lockdowns on population exposure to air pollution and suggest possible explanations for the observed trends. Our approach consists of three steps: We first test the impact of the stringency of lockdown restrictions on regional air quality; we then model a counterfactual case in the absence of lockdowns and calculate resulting changes in air pollutant levels; and last, we quantify the change in air pollution–related premature mortality in each region.

Impacts of lockdown stringency on regional air quality

In the first step, we test statistically whether lockdown restrictions correlate with a change in regional levels of air pollutants (NO2, PM2.5, and ozone). To capture geographical differences within large countries, the regions considered in this study are defined by the first-level administrative boundaries (i.e., states, prefectures, and provinces) of the United States, China, South Korea, Japan, France, Italy, Spain, Germany, and the United Kingdom. Other regions are included at the country level (see Materials and Methods). We collect hourly air quality measurement data from ground stations in 36 countries from 1 January 2016 to 6 July 2020 (35–39) and tropospheric NO2 column measurements from the ESA SENTINEL-5 satellite (40) from 1 May 2018 to 17 June 2020. The satellite-based dataset allows us to include countries where ground monitor data are not available and provides a second set of results in regions where ground monitor data are available. Surface-level NO2 concentrations are estimated from column measurements using two different methods for robustness (see Materials and Methods). This dataset of air quality measurements is combined with population density data provided by the Center for International Earth Science Information Network (41) to quantify population-weighted average exposure in each region before and during the lockdown restrictions.

These air quality measurements are combined with the indicator of the stringency of lockdown measures developed by Hale et al. (1), which measures the stringency of the lockdown restrictions in each country on a scale from 0 to 100. The stringency index is set to 0 before the implementation of lockdown restrictions. For the United States, where the lockdown restrictions were implemented at the state level, we develop a similar, state-level stringency index (see Materials and Methods). We test the hypothesis that lockdown policies (as measured by the stringency index) led to reduced pollution exposure in each region using an auto-regressive, integrated, moving average (ARIMA) time-series model with the stringency index as an exogenous variable. The model is fitted to the full (pre- and post-restrictions) dataset to calculate, for each region, the coefficient associated with the stringency index along with its associated P value and 95% confidence interval (CI).

Changes in regional air quality associated during the lockdown period

In the second step, we train and validate a second ARIMA time-series model on pre-lockdown measurements and use it to estimate regional air pollution in a hypothetical case with no lockdown restrictions. This second model is applied to calculate counterfactual concentrations of NO2, PM2.5, and ozone in each region, i.e., the concentrations of pollutants that would have been expected had there been no lockdown restrictions. This model captures expected seasonal and weekly patterns, thereby producing estimates of counterfactual levels of air pollution (see Materials and Methods and the Supplementary Materials for validation). To estimate differences in pollutant levels between the areas covered by ground monitors and the rest of the region, we interpolate measured values at monitor locations to cover the entire region. Similar to previous studies (42–44), the spatial interpolation is performed using ordinary kriging.

The difference between the results from this second ARIMA model (counterfactual) and air quality measurements during the lockdown restrictions allows us to estimate the impact of lockdown restrictions on regional air quality. Monte Carlo simulation is used to quantify the uncertainty associated with the prediction of the counterfactual in each region. On the basis of the comparison of the changes in regional air quality among the different pollutants (NO2, PM2.5, and ozone), we offer initial insights into the atmospheric mechanisms at play and their relevance to air quality policy.

Change in air pollution–related premature mortality

In the third step, concentration-response functions (CRFs) adapted from the epidemiological literature (22, 45, 46) are used to relate the difference between actual and counterfactual population exposure to the resulting health impacts from short-term exposure to NO2, PM2.5, and ozone between the first day of application of lockdown measures in each region and the last day of study (6 July 2020). Uncertainty in the parameters of the CRFs is included in the Monte Carlo simulation.

RESULTS AND DISCUSSION

This section is organized as follows: We start by presenting the results of our first and second steps regarding the impacts of lockdown stringency on air pollution in each region. We then present our premature mortalities estimates (third step) in each region.

Regional air quality during the lockdown period

This section first presents the impacts of lockdown stringency on air pollution in each region (first step) and the changes in air pollution compared to the counterfactual (second step) for each pollutant (NO2, PM2.5, and ozone) individually, followed by a multispecies analysis.

Decreases in NO2 exposure obtained from ground monitor data coincide with the tightening of the COVID–19 lockdowns as measured by the stringency index in 213 of the 252 regions considered, accounting for 93% of the population in the countries considered (see Fig. 1 and table S1 for the full results by region). These impacts are statistically significant at the 5% level in 120 of these regions, accounting for 70% of the population. This means that, in these regions, the null hypothesis, that there were no changes in levels of ambient air pollution due to the increase in the stringency index, can be rejected with 95% confidence.

In light of these results, we calculate the average difference between the observed NO2 exposure and the modeled counterfactual during the lockdown period in each region (second step of our analysis, using the second ARIMA model). The reductions are largest in China, where the population-weighted changes in NO2 concentration average −16 μg m−3 (95% CI, −26 to −7.2 μg m−3), or a −53% (95% CI, −83 to −23%) change compared to the counterfactual case. The average NO2 concentrations between the start of the local lockdown

and 6 July 2020 change by −25% (95% CI, −46 to −5.0%) in South Korea, −24% (−47 to −0.96%) in Europe, −8.6% (95% CI, −25 to −7.0%) in Japan, and −4.3% (95% CI, −14 to +5%) in the United States. These results are confirmed with satellite-based data: We find that 222 of the 252 regions analyzed with the ground monitor method show decreases in NO2 levels in response to tighter lockdown measures (including 179 that have decreases with all three methods). In addition, the same result is found for 92 of 222 countries not covered by ground monitor data. The largest changes in NO2 levels in countries not covered by monitor data are in Peru (−59%; 95% CI, −90 to −14%), Lebanon (−35%; 95% CI, −44 to −25%), and Singapore (−23%; 95% CI, −74 to 28%) (fig. S8 and table S1).

In 163 of 252 regions, covering 69% of the population, ground monitor data suggest that PM2.5 concentrations decreased with higher lockdown stringency. The effect is statistically significant at the 5% level in 77 of these regions based on our statistical test. In China, the resulting changes in PM2.5 concentrations amount to −36% (95% CI, −69 to −2.4%) compared to the values expected in the counterfactual. In Japan, the relative change in PM2.5 levels during the period of study is −17% (95% CI, −24 to −10%), while South Korea experienced a −16% (95% CI, −42 to −9.7%) change in PM2.5. In contrast, in Europe and the United States, we find that ambient levels of PM2.5 did not change with tighter lockdown restrictions: We find a +8.5% (95% CI, −85 to +88%) change in PM2.5 in Europe during the period of study, and a +1.4% (95% CI, −21 to +24%) change in the United States. Only 7 of 106 regions in Europe (2% of the European population) and four states in the United States (11% of the U.S. population) have statistically significant changes in PM2.5 concentration at the 5% level (fig. S3 and table S1).

For ozone, 146 of the 252 regions globally—accounting for 46% of the population analyzed—show decreases in concentration in response to the lockdown measures. In 45 of these regions, decreases in population-weighted average ozone concentrations are statistically significant at the 5% level. Changes are largest in South Korea (−28%; 95% CI, −43 to −13% compared to the counterfactual) and Japan (−5.4%; 95% CI, −12 to +1.0% change). Results in the United States, China, and Europe show no statistically significant change in ozone: the United States has a −0.82% (95% CI, −5.6 to +4.0%) change in ozone concentrations, and China and Europe have changes of +0.45% (95% CI, −25 to +24%) and −3.1% (95% CI, −17 to +9.4%), respectively.

We analyze the correlation between changes in secondary air pollutants (ozone and PM2.5) and NO2 using ground monitoring data to provide further insight into where trends align between species and where these trends break. We find that 140 of 252 regions (and 65% of the population) have decreases in PM2.5 and NO2 after the lockdown measures are implemented (see fig. S12) and 123 in both NO2 and ozone (43% of the population). Eighty-three of
252 regions (and 32% of the population), 48 of which are in Japan and South Korea, have decreases in all three pollutants attributable to lockdown measures. Among the 38 regions where the changes in both NO2 and PM2.5 are statistically significant at the 5% level, the ratio of the change in PM2.5 to the change in NO2 varies by a factor of 7. However, 88 regions have decreases in NO2 and increases in ozone. This potentially suggests a counterintuitive atmospheric response to changes in NO2 in some regions.

We explore possible pathways explaining the local increases in ozone by collecting additional, satellite-based data to characterize the chemical regime of each region based on pre-shutdown conditions using the ratio between satellite-derived column measurements of formaldehyde (HCHO) and NO2 (see Materials and Methods; Fig. 2A). Globally, we find that 94% of the regions under consideration are in a "transition" regime with HCHO-to-NO2 ratios between 1 and 4 (47–50). Past research has found that, in the transition regime, decreases in NO2 can result in either an increase or a decrease in ozone. We find that the probability of having a reduction in ozone in a given region as a result of an NO2 decrease varies approximately linearly with HCHO/NO2 (Fig. 2B), suggesting a well-characterized gradual transition. These results are consistent with a negligible net ozone benefit of the NO2 reductions globally and provide, for each region, a first estimate of the changes in volatile organic compounds (VOCs) that would be needed for NO2 reductions to translate into decreases in ozone (figs. S5 and S6). This is the most general characterization to date of the relationship of satellite-observed HCHO and NO2 columns with surface ozone. Details about each region and region-specific ozone isopleths are presented in figs. S5 and S6.

**Change in air pollution–related premature mortality**

Aggregating across pollutants based on monitor data, we find that changes in pollutant concentrations associated with COVID lockdowns up to 6 July 2020 resulted in −95,000 (95% CI, −160,000 to −29,000) premature mortalities from NO2, PM2.5, and ozone exposure during the lockdown period. Statistically significant changes in air pollution at the 5% level contribute to −67,000 (95% CI, −115,000 to +19,000) of the total premature mortalities. For comparison, the global death toll of COVID-19 as of 6 July 2020 was 544,000 (51). Changes in NO2 exposure account for −31,000 (95% CI, 65,000 to −2300) premature mortalities, while changes in PM2.5 exposure account for −64,000 premature mortalities (95% CI, −95,000 to −32,000) and changes in ozone exposure for 30 additional premature mortalities globally (95% CI, −190 to 800). China accounts for 79% of the estimated decrease in total premature mortalities, including 65% of the reductions in NO2 mortalities, and 85% of the reductions in PM2.5 mortalities. For Europe, we estimate a total change of −6600 premature mortalities (95% CI, −13,000 to +380) from reduced exposure to NO2 and −6100 premature mortalities (95% CI, −9100 to +3100) from changes in PM2.5. If different CRFs are used (see Materials and Methods), we find average mortality estimates of −43,000 (95% CI, −62,000 to −25,000) for PM2.5 and −51,000 (95% CI, −62,000 to −40,000) for NO2 (see Materials and Methods). When using cause-specific CRFs, we find that changes in ambient levels of NO2 resulted in −2800 (95% CI, −11,000 to +6000) premature mortalities from respiratory diseases and −10,000 (95% CI, −41,000 to +21,000) from cardiovascular diseases. Ambient changes in PM2.5 are found to result in −10,000 (95% CI, −14,000 to −7000) premature mortalities from respiratory diseases and −38,000 (95% CI, −50,000 to −25,000) from cardiovascular diseases. Overall, using cause-specific CRFs results in a total of −60,000 (95% CI, −116,000 to −4200) premature mortalities, compared to −95,000 with the all-cause CRFs, with 75% of avoided premature mortalities being from cardiovascular diseases and 25% from respiratory diseases.

Reductions in air pollution–related premature mortalities per capita represent less than 2.8% (95% CI, 0.56 to 5.1%) of the deaths per capita from COVID-19 in the United States (51) and 6.4% (95% CI, 1.0 to 12%) in Europe, but are between 2.3 and 16 times greater than the reported number of COVID-19 deaths per capita in Asia (Fig. 3).

In East Asia, where lockdown measures yielded a reduction in both primary and secondary air pollution, reductions in premature mortalities due to COVID-19–related lockdown measures correlate with the average stringency over the period since lockdown measures were implemented in that region. The Pearson correlation coefficient (r²) between the average stringency over the period of study and the avoided air quality–related premature mortalities
population exposure to air pollution (53, 54), the COVID-19 lockdowns highlight the need for targeted air quality policies to maximize public health benefits at a reasonable economic cost. In particular, accounting for the prevailing chemical regime is key to reducing ozone levels, and our results provide a worst-case bound and the probability of an ozone benefit from regional NO2 reductions, which are available globally. Recent examples of successful air quality policies include the 2013 Air Pollution Action Plan in China (34, 55) or the Regional Greenhouse Gas Initiative in the United States (56). In addition, the observations collected during this shutdown offer an unprecedented opportunity to further improve our understanding of atmospheric chemistry, and from there to better predict the effect of future, structured measures to improve air quality globally.

**MATERIALS AND METHODS**

**Impacts of lockdown stringency on regional air quality**

**Ground monitor measurements**

We retrieve hourly ground measurements of PM2.5, ozone, and NO2 from 1 January 2016 to 7 July 2020 from the national environmental monitoring agencies of the United States, the People's Republic of China, Japan, the Republic of Korea, and the European Union (35–39). Retrieved values are filtered to exclude negative values as well as invalid measurements where applicable and processed to produce daily average PM2.5 and NO2 concentrations and daily maximum 8-hour ozone concentration (MDA8) at each monitor location. The regions covered are defined by the first-level administrative boundaries (i.e., states, prefectures, and constituent countries) of the United States, China, South Korea, Japan, France, Italy, Spain, Germany, and the United Kingdom. In addition, we include other member countries of the European Union, Norway, and Switzerland at the national level.

From daily pollutant concentrations at each monitor location, we compute the daily population-weighted average concentration in each region and for each pollutant. The population count is taken from the Center for International Earth Science Information Network (CIESIN) at a resolution of 1/24° (0.04°, about 5 km). In cases where several monitors belong to the same grid cell, we average their measurements. To test the statistical significance of changes in concentration due to lockdown measures for each pollutant and region, we perform the spatial integration of monitor data only at monitor locations. This avoids the introduction of interpolated values in the signal. The value of the average concentration, \( C_{r,d,\text{poll}} \), of a pollutant, poll, in region, \( r \), for day, \( d \), is computed with the following formula

\[
C_{r,d,\text{poll}} = \frac{\sum_{i,j \in J \times I} p_{i,j} c_{i,j,d,\text{poll}}}{\sum_{i,j \in J \times I} p_{i,j}}
\]

where \( I \times J \) is the set of indices within region \( r \) where the population grid contains a monitor, \( p_{i,j} \) is the population at location \( i,j \), and \( c_{i,j,d,\text{poll}} \) is the daily concentration of the pollutant derived from hourly measurements at monitor location \( i,j \).

**Satellite measurements**

We use high-resolution satellite data from the Tropospheric Monitoring Instrument (TROPOMI) onboard the ESA's Sentinel-5 Precursor (S5P) satellite (launched October 2017, data available from April 2018) to obtain global measurements of tropospheric NO2 vertical column densities (VCDs). The satellite has a Sun-synchronous orbit and provides daily retrievals of NO2 VCDs over the globe. We use the TROPOMI data to estimate regional average NO2 concentrations by first averaging daily VCDs from the same monitor over region \( r \) and then dividing by the area of the region.

**Impacts of lockdown stringency on regional air quality**

**Ground monitor measurements**

We retrieve hourly ground measurements of PM2.5, ozone, and NO2 from 1 January 2016 to 7 July 2020 from the national environmental monitoring agencies of the United States, the People's Republic of China, Japan, the Republic of Korea, and the European Union (35–39). Retrieved values are filtered to exclude negative values as well as invalid measurements where applicable and processed to produce daily average PM2.5 and NO2 concentrations and daily maximum 8-hour ozone concentration (MDA8) at each monitor location. The regions covered are defined by the first-level administrative boundaries (i.e., states, prefectures, and constituent countries) of the United States, China, South Korea, Japan, France, Italy, Spain, Germany, and the United Kingdom. In addition, we include other member countries of the European Union, Norway, and Switzerland at the national level.

From daily pollutant concentrations at each monitor location, we compute the daily population-weighted average concentration in each region and for each pollutant. The population count is taken from the Center for International Earth Science Information Network (CIESIN) at a resolution of 1/24° (0.04°, about 5 km). In cases where several monitors belong to the same grid cell, we average their measurements. To test the statistical significance of changes in concentration due to lockdown measures for each pollutant and region, we perform the spatial integration of monitor data only at monitor locations. This avoids the introduction of interpolated values in the signal. The value of the average concentration, \( C_{r,d,\text{poll}} \), of a pollutant, poll, in region, \( r \), for day, \( d \), is computed with the following formula

\[
C_{r,d,\text{poll}} = \frac{\sum_{i,j \in J \times I} p_{i,j} c_{i,j,d,\text{poll}}}{\sum_{i,j \in J \times I} p_{i,j}}
\]

where \( I \times J \) is the set of indices within region \( r \) where the population grid contains a monitor, \( p_{i,j} \) is the population at location \( i,j \), and \( c_{i,j,d,\text{poll}} \) is the daily concentration of the pollutant derived from hourly measurements at monitor location \( i,j \).

**Satellite measurements**

We use high-resolution satellite data from the Tropospheric Monitoring Instrument (TROPOMI) onboard the ESA’s Sentinel-5 Precursor (S5P) satellite (launched October 2017, data available from April 2018) to obtain global measurements of tropospheric NO2 vertical column densities (VCDs). The satellite has a Sun-synchronous orbit and provides daily retrievals of NO2 VCDs over the globe. We use the TROPOMI data to estimate regional average NO2 concentrations by first averaging daily VCDs from the same monitor over region \( r \) and then dividing by the area of the region.
or with a local overpass time of around 13:30 with daily global full-surface coverage (57). The TROPOMI nadir-viewing short-wave spectrometer has a spatial resolution of 3.5 km × 7 km (3.5 km × 5.5 km since August 2019) in the ultraviolet–near-infrared spectrum (405 to 465 nm), which is used for NO₂ measurements. Its resolution is higher than that of similar previous instruments such as OMI and GOME-2. TROPOMI’s ability to resolve individual emissions sources and plumes has been demonstrated in regional validation studies of the Canadian Oil Sands (58), Helsinki (59), and South Korea (60), which verified TROPOMI’s performance against ground-based spectroscopy, ground-based monitors, and integrated air quality models.

We use the level 2 product (“SSP_L2__NO2”) of tropospheric NO₂ VCD. TROPOMI’s NO₂ retrieval method, developed by the Royal Netherlands Meteorological Institute, uses a differential optical absorption spectroscopy method to determine the NO₂ slant column density, which is then converted to a tropospheric VCD using a data-assimilated chemistry transport model (TM5-MP). A quality band (“qa_value”) is provided for each pixel, ranging from 0 (poor) to 1 (best); we use a pixel selection criterion of qa_value > 0.75, which removes scenes covered by snow/ice, errors, and problematic retrievals. Satellite data are regridded to a 0.25° × 0.25° grid by averaging measurements within a grid cell. The average daily coverage between 30 April 2018 and 17 June 2020 is 67.9% by population (daily min. 8.5%, max. 80.8%) and 56.0% by area (daily min. 23.8%, max. 61.4%).

To model the relationship between tropospheric VCD, \( C^\text{trop} \), and ground-level concentration, \( C^g \), we use two independent linear models for robustness: The first one makes use of the ground monitor measurements presented earlier, and the second one makes use of modeled data.

In the first model, the relationship between these two quantities is as follows

\[
C^g_{x,y,day_k} = K^0_{x,y} + K^1_{x,y} C^\text{trop}_{x,y,day_k}
\]

(2)

where \( K^0 \) and \( K^1 \) are empirically derived constants that we find for each grid cell. We derive \( K^0 \) and \( K^1 \) for grid cells where ground monitor measurements are available (mean 2587 grid cells, daily min. 922, daily max. 2882) by performing a least squares regression between historical ground monitor measurements and satellite measurements from the years 2018 and 2019 (2020 data are not included). We constrain the slope \( K^1 \) to be nonnegative; where a negative value of \( K^1 \) results from the regression, it is set to zero and \( K^0 \) is the mean of the historical ground monitor measurements. Maps of \( K^0 \) and \( K^1 \) as well as the coefficient of determination of the regression can be found in fig. S1. To obtain a relationship between tropospheric VCD and ground-level concentration for locations where monitor data are not available, inverse distance weighting is used to interpolate values of \( K^1 \) and \( K^2 \) between monitor locations.

In addition to the previous approach, we also use the GEOS Composition Forecast (GEOS-CF) system as an alternate method to estimate ground NO₂ from satellite data of tropospheric NO₂ VCD. The GEOS-CF system combines the GEOS weather analysis and forecasting system with the state-of-the-science GEOS-Chem chemistry module (61–63) to provide detailed chemical forecast of a wide range of air pollutants including ozone, carbon monoxide, nitrogen oxides, and fine particulate matter (PM₂.₅). Specifically, we use data of NO₂ concentrations from 2019 at a spatial resolution of 0.25° × 0.25°. The chemical forecast provides both the VCD and the ground concentration at the satellite local overpass time of 13:30, allowing us to estimate a relationship between the two. This method does not rely on ground monitor measurements and provides an independent assessment of population exposure. It also has the advantage of providing global coverage without the need for interpolation and in areas where ground monitor data are not available.

We use a linear model (similar to that of the satellite-monitor method) to estimate the ground concentration, \( C^g \), from tropospheric VCD

\[
C^g_{x,y,day_k} = K^0_{x,y,month_l} + K^3_{x,y,month_l} C^\text{trop}_{x,y,day_k}
\]

(3)

where \( K^2 \) and \( K^3 \) are empirically derived constants that we find for each grid cell and for each month of the year and \( C^\text{trop} \) is the satellite NO₂ VCD measurement. \( K^2 \) and \( K^3 \) are determined by least squares regression between historical data of ground concentration from the chemical forecast and tropospheric VCD from the chemical forecast for each month. The monthly regression allows this model to capture intra-annual variations in the relationship between tropospheric VCD and ground concentration. Maps of \( K^2 \) and \( K^3 \) as well as the coefficient of determination of the regression for the month of April can be found in fig. S2.

### Stringency index

To measure the effect of lockdown measures on changes in economic activity and pollutant emissions in each country, we use the indicator of the stringency of lockdown measures developed by Hale et al. (1). This index ranges from 0 to 100 tracks governments’ responses to the COVID-19 crisis in 150 countries over time, at a daily resolution, and takes into account several measures that affect economic activity. For each date, Hale et al. (1) aggregate country-specific data regarding containment and closure measures (namely, school closings, workplace closings, cancellation of public events, restrictions on gathering size, closing of public transport, stay-at-home requirements, restrictions on internal movement, and restrictions on international travel). These dimensions are each coded onto a numeric scale and then normalized and averaged to produce a single value by country and by day. We use this index to represent numerically the different lockdown measures that were implement ed in China, South Korea, Japan, and every country in the European Union along with Switzerland, Norway, and Iceland.

In the United States, because lockdown measures were decided at the state level (64), we develop a similar lockdown index at the state level by tracking school closures, declaration of a state of emergency, closing of nonessential businesses, stay-at-home orders, restrictions on internal movement at the state level, and restriction on international travel at the federal level over time. We then apply the same method as Hale et al. (1), standardize, and average the index for each day and each state.

### Significance testing

To relate any changes in observed PM₂.₅, ozone, and NO₂ concentrations to COVID-19–related lockdown measures, we perform a regression of daily, population-weighted, average concentration in each region obtained with the methods described above against the stringency index described above for each region under study. This regression features ARIMA errors and is implemented using the Python package pmdarima (65). The degree of differencing \( d \) and the order \((p,q)\) of the ARIMA error are chosen to minimize the Akaike information criterion (AIC) (66). We model the time dependence...
of regional daily population-weighted concentration $y_i$ with Fourier terms, following Hyndman (67). This captures the seasonality of regional exposure to air pollution. Overall, the regression for a given region can be expressed as

$$y_i = a + \sum_{k=1}^{K} \left( \alpha_k \sin \left( \frac{2\pi kt}{m} \right) + \beta_k \cos \left( \frac{2\pi kt}{m} \right) \right) + \gamma S_i + N_i \tag{4}$$

where $S_i$ is the value of the shutdown index in the region at time $t$ and $N_i$ the ARIMA(p,d,q) term verifies the relationship

$$(1 - \phi_1 B - \ldots - \phi_p B^p) N_i = (1 + \theta_1 B + \ldots + \theta_q B^q) \epsilon_t \tag{5}$$

with $B$ the backshift operator defined as follows

$$\forall n \in \mathbb{N}, B^n N_i = N_{i-n} \tag{6}$$

and $\epsilon_t$ a white noise series. The seasonality factor $m$ is set to 365 and $K$ is set to 100 to capture not only annual variations but also higher-frequency periodic variations. Although the ARIMA model does not lend itself to a simple decomposition between seasonal and trend component, our formulation represents seasonality with the Fourier terms, while the auto-regressive part of the model (the combination of $N$ terms) captures the influence of previous days' air pollution level on the current level and the moving average part (the $\epsilon$ terms) accounts for past prediction errors.

For any region and pollutant where the $P$ value of the $\gamma$ coefficient is below 0.05, we conclude that the correlation between the imposed lockdown measures and the change in population exposure due to lockdown measures in each region is prepared for regression by removing any negative values, 0 values, and values above 500 $\mu$g m$^{-3}$ and linearly interpolating missing values along the time dimension. Leading and trailing missing values are excluded altogether. Furthermore, any time series with more than 10% missing values before interpolation is excluded from further analysis.

An ARIMA model is then trained on the time series described above and truncated on the last day before lockdown measures were implemented. Similar to the first part of this study, we forecast daily exposure using historical data and previous forecasted values to generate predictions. Overall, the regional average concentration of a pollutant is given by

$$y_i = a + \sum_{k=1}^{K} \left( \alpha_k \sin \left( \frac{2\pi kt}{m} \right) + \beta_k \cos \left( \frac{2\pi kt}{m} \right) \right) + N_i \tag{7}$$

where

$$N_i = \phi_1 N_{i-1} + \phi_2 N_{i-2} + \ldots + \phi_p N_{i-p} + \epsilon_t + \theta_1 \epsilon_{t-1} + \ldots + \theta_q \epsilon_{t-q} \tag{8}$$

Similar to the significance testing step, we set $m$ to 365 and $K$ to 100 to account for medium- and long-term temporal effects. The ARIMA term $N_i$ captures the short-term variations, and the order of the auto-regressive and moving average parts, $p$ and $q$, respectively, are chosen to minimize the AIC. Last, $\epsilon_t$ is a white noise series whose squares the fitting step minimizes.

To validate the model, we train the predictive model for the counterfactual in each region with data from 1 January 2016 through the day of 2019 corresponding to the last day before lockdown measures were in place in the region in 2020 and validated over the corresponding period of 2019 (first day of lockdown through 6 July for the ground monitor method and 17 June for the satellite methods). We validate the predicted monthly average concentration of each pollutant in the region against the measured values for 2019 using the coefficient $r^2$ and the average relative root mean square error. Validation results by region are included in table S1.

### Calculating changes in air pollution relative to a counterfactual scenario

#### Counterfactual pollutant levels.

To predict regional population-weighted, average concentrations in the counterfactual case (i.e., in the absence of lockdown measures) between the first day the lockdown measures were in place and the last day of study (6 July 2020), we generate a time series of daily populations’ concentrations between 1 January 2016 until the last day before lockdown measures were implemented by integrating spatially interpolated monitor-derived values over each region. The spatial interpolation makes use of ordinary kriging [following previous studies (42–44)]. On the basis of results from a grid search cross-validation comparing ordinary and universal kriging with linear, spherical, Gaussian, and power variogram models, we use ordinary kriging with a spherical variogram model in all regions. The parameters are selected for each country using 10-fold cross-validation leaving 10% of the monitor-derived values out of the training set. Interpolated daily concentrations are then weighted by population and averaged regionally to produce the time series. Therefore, the value of the average concentration in any given region on any given day is calculated from Eq. 1, with $J \times I$ the set of all indices within the region. To evaluate the sensitivity of the mortality results to the spatial interpolation, we replicate the generation of the regional time series using the 2.5th and 97.5th percentiles of the spatially interpolated values and find that our overall mortality results vary by 3%.

Last, the time series of the average concentration in each region is prepared for regression by removing any negative values, 0 values,

$$\text{Validation results by region are included in table S1.}$$

### Calculating changes in pollutant levels relative to the counterfactual scenario

For each region, we predict the mean value of the daily population-weighted average concentration, as well as the 95% CI. The daily difference between actual and counterfactual concentrations (and the associated 95% CIs) is used to calculate the changes in pollutant levels relative to the counterfactual scenario. This difference is also used as input to the concentration response functions used to derive the air quality–related health impacts of the lockdown measures. The uncertainty associated with the prediction of the counterfactual is included in a Monte Carlo simulation that also includes uncertainty in the spatial interpolation and uncertainty in the premature mortality calculation (see below).

Estimated changes in air pollutant concentrations in each region lead to changes in air quality–related premature mortalities. On the basis of the comparison between actual regional average concentration of PM$_{2.5}$, NO$_x$, and ozone derived using the methods described in the sections below, we calculate daily changes in population exposure due to lockdown measures in each region. Population density and count are taken from the Center for International Earth Science Information Network (41) at a resolution of 0.04° (approximately 5 km).

### Change in air pollution–related premature mortality

We calculate the response of daily all-cause mortality to changes in population exposure to PM$_{2.5}$, NO$_x$, and ozone by adapting CRFs from the
epidemiological literature. For each of these species, we adopt a log-linear CRF relating increases in the risk of all-cause mortality to changes in population exposure. For NO$_2$, we use the results of He et al. (45) who find a 0.57% (95% CI, −0.04 to 1.18%) increase in all-cause mortality per 10 µg m$^{-3}$ increase in the average NO$_2$ concentration over the past 7 days. We also quantify premature mortality from cardiovascular and respiratory mortality using the estimates from He et al. (45) [0.25% (95% CI, −0.66 to 1.18%) and 0.45% (−0.96 to 1.89%) increases in cardiovascular and respiratory mortality, respectively, per 10 µg m$^{-3}$ increase in the average NO$_2$ concentration over the past 7 days]. For PM$_{2.5}$, Atkinson et al. (46) report a 1.04% increase in all-cause mortality (95% CI, 0.52 to 1.56%) for every 10 µg m$^{-3}$ increase in same-day PM$_{2.5}$. In a similar fashion to the NO$_2$ case, we quantify cardiovascular and respiratory premature mortality as well [0.84% (95% CI, 0.41 to 1.28%) and 1.51% (95% CI, 1.01 to 2.01%) increases in cardiovascular and respiratory mortality, respectively, per 10 µg m$^{-3}$ increase in same-day PM$_{2.5}$]. We use the same CRFs for PM$_{2.5}$ and NO$_2$ in all regions under study.

For ozone, we use country-specific results from Vicedo-Cabrera et al. (22) where applicable (for European countries that are not included in their results, we apply the European average; for other countries with no information, we apply the global average). Changes in relative risk for all-cause mortality range from 0.066% (95% CI, −0.08 to 0.19%) in Spain to 0.35% (95% CI, 0.24 to 0.46%) in the United Kingdom per 10 µg m$^{-3}$ increase in same-day maximum 8-hour ozone.

We also perform a CRF sensitivity analysis for each species under study using parameters derived from Chen et al. (23) for NO$_2$, who find a 0.9% (95% CI, 0.7 to 1.1%) increase in all-cause mortality per 10 µg m$^{-3}$ in 2-day NO$_2$, from Bell et al. (68) for O$_3$, who report a 0.52% (95% CI, 0.27 to 0.77%) increase in daily all-cause mortality per 10 ppb (parts per billion) increase in the previous week’s ozone, and from Brook et al. (69), who find a 0.7% (95% CI, 0.4 to 1.0%) increase in all-cause mortality per 10 µg m$^{-3}$ increase in same-day PM$_{2.5}$.

Average annual incidence rates of all-cause mortality as well as low and high estimates (interpreted as the 95% CI and fitted with a normal distribution) are taken from the World Health Organization Global Burden of Disease 2017 study (70) and divided by 365 to produce daily incidence of all-cause mortality. We take incidence rates at the finest resolution available for each country. As a result, incidence rates are region specific in the United States and Japan, and country-specific elsewhere. Overall, the change in daily mortality $M$ is computed using

$$\Delta M = p_{\text{aff}} I_0 \Delta I_0$$

where $\Delta I$ is the change in incidence rate due to the lockdown measures, $p_{\text{aff}}$ is the affected population, and $I_0$ is the incidence rate at the baseline (counterfactual) concentration of pollutant. By definition, the last ratio can be expressed as

$$\Delta I_0 = \frac{RR_{\text{actual}} - RR_0}{RR_0}$$

where $RR_0$ is the relative risk at the counterfactual concentration of pollutant and $RR_{\text{actual}}$ is the relative risk at the actual (measured) concentration of pollutant. Furthermore, a Taylor expansion of the above numerator yields

$$RR_{\text{actual}} - RR_0 = \Delta \chi \left( \frac{\partial RR}{\partial \chi} \right)_{\chi=0} + O(\chi^2)$$

where $\chi$ is the concentration of pollutant.

Given that we assume that the relative risk has a log-linear shape

$$RR = \exp(\beta \chi)$$

and neglecting second- and higher-order terms

$$\frac{RR_{\text{actual}} - RR_0}{RR_0} = \beta \Delta \chi$$

which is calculated from the parameters reported in the corresponding epidemiological study (22, 46, 47). Combining Eqs. 9 to 11, we obtain

$$\Delta M = p_{\text{aff}} I_0 \beta \Delta \chi$$

Daily estimates of premature mortalities due to changes in pollutants’ concentration in each region between the first day of lockdown measures and the last day of study (7 July 2020 in the case of monitor data) are then added to produce our final mortality estimates. Uncertainty in the CRF’s $\beta$ parameter is estimated from the reported CI for the relative risk in the original epidemiological study and propagated to our results using a Monte Carlo simulation with 10,000 samples. Total calculated changes in mortality in each region are compared to the number of deaths due to COVID-19 as reported by the Johns Hopkins University COVID-19 Data Repository (51).

Characterization of the ozone regime

Satellite-derived ozone isopleths

To explain the observed changes in ozone concentrations and the differences that we find between regions, we derive daily tropospheric column measurements of HCHO above monitor locations from the TROPOMI instrument for dates from 1 May 2018 to 17 June 2020, in addition to NO$_2$ column measurements.

For each region, we aggregate these measurements into 15 NO$_2$ and 15 HCHO bins and we average the corresponding monitor-derived maximum daily 8-hour ozone. We use these bins to construct the isopleths presented in fig. S5. Satellite OMI and, more recently, TROPOMI products have been used extensively in the literature (47–50) to predict changes in surface-level ozone concentrations in Europe, the United States, and East Asia. Previous studies found that HCHO/NO$_2$ ratios below 1 are indicative of a NO$_2$-saturated regime, where decreases in HCHO and NO$_2$ levels may result in increases in ozone. HCHO/NO$_2$ ratios above 2 (48) or, depending on the region, 4 (47, 50) are indicative of a NO$_2$-limited regime where decreases in NO$_2$ and HCHO levels lead to decreases in ozone. Ratios in between are characteristic of a transition regime.
In addition, we estimate, for each region, the average columns HCHO and NO₂ between the first day of lockdown in the region of interest and 17 June 2020 (Fig. 2). We also calculate the average satellite-derived column HCHO and NO₂ during the same period in 2018 and 2019. The difference between the 2018–2019 average and 2020 is indicative of the expected ozone change and helps explain the results described in the main paper (fig. S5).

SUPPLEMENTARY MATERIALS
Supplementary material for this article is available at http://advances.sciencemag.org/cgi/content/full/7/21/eabe1178/DC1

REFERENCES AND NOTES
4. Industrial Production Index, FRED, Federal Reserve Bank of St. Louis (2020); https://fred.stlouisfed.org/series/INDPRO.
10. COVID-19: Nitrogen dioxide over China. esa.int (2020); https://www.esa.int/Applications/Observing_the_Earth/Copernicus/Sentinel-SP/Covid19-nitrogen_dioxide_over_China.


64. Forecasting with long seasonal periods | R. J. Hyndman (2010); https://robjhyndman. com/hyndsight/longseasonality/.


Acknowledgments
Funding: This publication was partially supported by U.S. EPA grant RD-835872-01. Its contents are solely the responsibility of the grantee and do not necessarily represent the official views of the U.S. EPA. Furthermore, U.S. EPA does not endorse the purchase of any commercial products or services mentioned in the publication. We also gratefully acknowledge support of this research from the VoLo Foundation.

Author contributions:
G.P.C., H.Y.D., and S.I. contributed data curation, formal analysis, investigation, methodology, software development, and validation for this study. G.P.C. and S.I. contributed to the drafting of the manuscript, with the detailed structure collectively developed and the bulk of the first draft completed by G.P.C. with input from other authors.

Competing interests: The authors declare that they have no competing interests.

Data and materials availability: All the code files necessary to reproduce the results of this study are available at https://zenodo.org/record/4553547 (DOI: 10.5281/zenodo.4553547). Full results are available at https://doi.org/10.5281/zenodo.4571643.

Submitted 31 July 2020
Accepted 2 April 2021
Published 21 May 2021
10.1126/sciadv.abe1178

Air pollution impacts of COVID-19–related containment measures
Guillaume P. Chossière, Hao Feng Xu, Yash Dixit, Stewart Isaacs, Sebastian D. Eastham, Florian Allroggen, Raymond L. Speth and Steven R. H. Barrett

Sci Adv 7 (21), eabe1178.
DOI: 10.1126/sciadv.eabe1178