Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.
The impact of the COVID-19 pandemic on air pollution: A global assessment using machine learning techniques

Jasper S. Wijnands a, b, *, Kerry A. Nice a, Sachith Seneviratne a, Jason Thompson a, Mark Stevenson a, c, d

a Transport, Health and Urban Design Research Lab, Melbourne School of Design, The University of Melbourne, Parkville VIC 3010, Australia
b Royal Netherlands Meteorological Institute (KNMI), 3731 GA De Bilt, The Netherlands
c Faculty of Engineering and Information Technology, The University of Melbourne, Parkville VIC 3010, Australia
d Melbourne School of Population and Global Health, The University of Melbourne, Parkville VIC 3010, Australia

A R T I C L E   I N F O

Keywords:
COVID-19
Nitrogen dioxide
Particulate matter
Ozone
Mobility
Decision trees

A B S T R A C T

In response to the COVID-19 pandemic, most countries implemented public health ordinances that resulted in restricted mobility and a resultant change in air quality. This has provided an opportunity to quantify the extent to which carbon-based transport and industrial activity affect air quality. However, quantification of these complex effects has proven to be difficult, depending on the stringency of restrictions, country-specific emission source profiles, long-term trends and meteorological effects on atmospheric chemistry, emission levels and in-flow from nearby countries. In this study, confounding factors were disentangled for a direct comparison of pandemic-related reductions in absolute pollution levels, globally. The non-linear relationships between atmospheric processes and daily ground-level NO₂, PM₁₀, PM₂.₅ and O₃ measurements were captured in city- and pollutant-specific XGBoost models for over 700 cities, adjusting for weather, seasonality and trends. City-level modelling allowed adaptation to the distinct topography, urban morphology, climate and atmospheric conditions for each city, individually, as the weather variables that were most predictive varied across cities. Pollution forecasts for 2020 in absence of a pandemic were generated based on weather and formed an ensemble for country-level pollution reductions. Findings were robust to modelling assumptions and consistent with various published case studies. NO₂ reduced most in China, Europe, and India, following severe government restrictions as part of the initial lockdowns. Reductions were highly correlated with changes in mobility levels, especially trips to transit stations, workplaces, retail and recreation venues. Further, NO₂ did not fully revert to pre-pandemic levels in 2020. Ambient PM₂.₅ pollution, which has severe adverse health consequences, reduced most in China and India. Since positive health effects could be offset to some extent by prolonged exposure to indoor pollution, alternative transport initiatives could prove to be an important pathway towards better health outcomes in these countries. Increased O₃ levels during initial lockdowns have been documented widely. However, our analyses also found a subsequent reduction in O₃ for many countries below what was expected based on meteorological conditions during summer months (e.g., China, United Kingdom, France, Germany, Poland, Turkey). The effects in periods with high O₃ levels are especially important for the development of effective mitigation strategies to improve health outcomes.

1. Introduction

At the end of 2019, a new, highly infectious and deadly coronavirus was detected, transmitted via human-to-human contact (Riou and Althaus, 2020). Since then, the rapid spread of various variants of SARS-CoV-2 has put unprecedented pressure on economies and healthcare systems, resulting in significant morbidity and a death toll of several million people (World Health Organization, 2021). To slow disease transmission, many countries implemented public health ordinances that included an array of measures such as mask wearing, quarantining of positive cases and reduced population mobility across and between cities. Even before travel and work restrictions were enacted and the World Health Organization (WHO) declared a pandemic on 11 March 2020, mobility dropped dramatically (Google, 2020). By April 2020, more than half of the world’s population had reduced
their travel by more than 50% (Forster et al., 2020; Le Quéré et al., 2020). Heavy restrictions resulted in deserted cities, empty roads and clear skies in cities frequently blanketed in air pollution (e.g., He et al., 2020). The enforced reductions in mobility and industrial activity provide a natural experiment to explore the effects of these activities on air pollution. For example, it provides an opportunity to investigate the effects on air quality of decarbonising the road transport system as part of climate change mitigation.

1.1. Quantifying the impact of COVID-19 lockdowns

During the first seven months following the start of the pandemic, over 200 papers were accepted for publication that explored changes in air pollution levels (Gkatzelis et al., 2021). Not all pollutants were affected equally, due to the different sources of pollution. For example, secondary pollutants such as \(O_3\) form through chemical reactions of \(NO_2\) and volatile organic compounds under the influence of solar radiation (Seinfeld and Pandis, 2016). \(NO_2\) is mainly emitted through fuel combustion from road transport and fossil fuel power plants. \(PM_{2.5}\) is also emitted in these sectors, but mostly from diesel motor vehicles (e.g., commercial heavy-duty diesel trucks) and coal-fired power stations. It is also generated in other sectors, for example, through industrial activity and coal-fired winter heating (He et al., 2020). Thunis et al. (2018) allocated \(PM_{2.5}\) emissions in Europe to agriculture (23%), industry (20%), natural (19%), transport (14%) and residential (13%) sources. In Scotland, \(PM_{2.5}\) pollution depends mostly on natural and non-traffic sources (Dobson and Semple, 2020). Although road transport strongly declined during lockdown periods, some locations were mainly exposed to pollution sources that were affected to a limited extent by the pandemic (e.g., agriculture, natural sources). Hence, besides the stringency of restrictions, the mix of emission sources in a specific country influences the reduction in air pollution that might be expected.

1.1.1. Nitrogen dioxide

Studies used either ground-level measurements or satellite remote sensing to quantify the reduction in \(NO_2\) during the initial lockdown period (i.e., March/April 2020 for most countries). In the United States, reported \(NO_2\) reductions were 20% compared to the corresponding period in 2010–2019 (Bekbalat et al., 2021) and 21.6% compared to 2019 (Goldberg et al., 2020). Other studies found average reductions of 1.3 parts per billion (reported as −27% compared to 2015–2019, before de-trending) (Archer et al., 2020) to 4.8 parts per billion (reported as −25.5% compared to 2017–2019) (Berman and Ebisu, 2020), with the largest reductions in urban counties. Sharma et al. (2020) found an average \(NO_2\) reduction of 18% compared to 2017–2019 in 22 cities across India, while Mahato et al. (2020) reported a 53% reduction compared to the pre-lockdown period in Delhi. In Wuhan, reported reductions were 53.3% compared to the pre-lockdown period (Lian et al., 2020) and 57% compared to 2017–2019 (Sicard et al., 2020). Sicard et al. (2020) also found a 53% \(NO_2\) reduction in selected European cities with high pollution levels (i.e., Nice, Rome, Valencia and Turin), while Bauwens et al. (2020) reported reductions of 2% compared to 2019 in Western Europe and 40% in China. Reductions varied substantially between cities (Goldberg et al., 2020). Conerton et al. (2020) found \(NO_2\) reductions of 24% compared to 2015–2019 in New York City, 25% in São Paulo, 38% in Los Angeles, and 39% in Paris.

\(NO_2\) reductions were short-lived and returned towards normal ranges after the lockdowns ended. For example, Dentener et al. (2020) observed reductions in the average \(NO_2\) tropospheric column of over 50% during March (compared to 2019) in a selection of major Asian cities. However, over the three-month period from March to May 2020, reductions were much lower at 15%–20%. Similarly, the observed three-month reductions were 20% in Germany and the Benelux, 15% in Italy, 10%–15% in North America, Spain, France, the United Kingdom, Poland and Czech Republic, and 8% in Romania.

1.1.2. Particulate matter

For particulate matter, results were mixed with large reductions in some countries and no significant effects in others. Substantial \(PM_{2.5}\) reductions of 36% compared to 2017–2019 (Sicard et al., 2020) and 36.9% compared to the pre-lockdown period (Lian et al., 2020) were reported for Wuhan. He et al. (2020) found \(PM_{2.5}\) reductions of 21.1 and 7.1 \(\mu g/m^3\) compared to 2019 in Chinese cities with and without formal lockdowns, respectively. Effects were larger in colder, richer, and more industrialised cities. Similarly, Giani et al. (2020) found an average reduction of 14.5 \(\mu g/m^3\) in population-weighted \(PM_{2.5}\) across China, compared to 2016–2019. In the areas most affected by COVID-19, two-month average reductions up to 40 \(\mu g/m^3\) were observed. In India, Sharma et al. (2020) found an average reduction in \(PM_{2.5}\) of 43% compared to 2017–2019 across 22 cities, while Mahato et al. (2020) reported a reduction of 39% compared to 2019 in Delhi.

In contrast, the average \(PM_{2.5}\) reduction in Europe was found to be only 2.2 \(\mu g/m^3\) compared to 2016–2019 (Giani et al., 2020). Reductions in Scotland were very limited (i.e., within 1 \(\mu g/m^3\) compared to 2017 and 2018), even though motor vehicle journeys reduced by 65% (Dobson and Semple, 2020). Further, Sicard et al. (2020) reported an average reduction of 4% compared to 2017–2019 in four cities with high pollution levels in Southern Europe. In the United States, a slight increase in \(PM_{2.5}\) was observed of 0.28 \(\mu g/m^3\) compared to April 2015–2019 (Archer et al., 2020). Bekbalat et al. (2021) found a 10% increase in \(PM_{2.5}\) compared to 2010–2019 and concluded this was within the normal range of variability.

1.1.3. Ozone

In contrast to the reductions observed for other pollutants, many studies found increased \(O_3\) during lockdown periods. For example, \(O_3\) increases of 17% compared to 2017–2019 were reported for both Europe (Sicard et al., 2020) and India (Sharma et al., 2020). Sicard et al. (2020) reported a 36% increase compared to 2017–2019 in Wuhan, while Lian et al. (2020) found an increase of 116.6% compared to the pre-lockdown period. Increased \(O_3\) levels were also reported for São Paulo (30% compared to 2015–2019), Paris (12%) and New York City (7%), while a decrease of 10% was observed for Los Angeles (Conerton et al., 2020). Many studies discussed atmospheric chemistry for \(O_3\) formation as the potential cause of observed ozone increases. \(O_3\) titration occurs particularly during winter months if the \(NO_2\) level is high, reducing the \(O_3\) level (Stillman, 1999). Hence, reduced \(NO_2\) during the initial lockdowns could lead to higher \(O_3\) pollution.

1.2. Limitations and opportunities

Many papers present case studies for a specific country (e.g., China, He et al. (2020), Scotland, Dobson and Semple (2020)) or a selection of cities (e.g., Connerton et al., 2020; Sicard et al., 2020). However, very few studies have performed city-level analyses at a global scale. Global studies have the advantage that they allow for a consistent comparison between countries. One example of a global analysis is the study by Venter et al. (2020), finding substantial variations in country-level temporal and spatial pollution anomalies for \(NO_2\), \(PM_{2.5}\) and \(O_3\) that remain unexplained. Although not considering cities individually, they compared pollution changes across 34 different countries.

Multiple confounding factors complicate the attribution of air pollution to pandemic-related changes. For example, long-term trends in air pollution can be observed for some pollutants, as government policies shift towards renewable energy sources. Several studies did not account for long-term trends in pollution levels, which could lead to incorrectly attributing a decrease in air pollution to actions taken during the pandemic. Further, various studies compared 2020 measurements to equivalent periods in previous years (e.g., 2019, or 2017–2019) without adjustments for weather effects (e.g., Bauwens et al., 2020; Berman and Ebisu, 2020; Sicard et al., 2020). This approach can have limitations if
the weather in 2020 differs from atmospheric conditions in the equivalent historical period (e.g., more rainfall or higher temperatures). Goldberg et al. (2020) found that weather variations between years can cause fluctuations in monthly NO₂ levels of approximately 15%. Further, Wang et al. (2020) found that benefits of emission reductions in China during the lockdown period were sometimes overwhelmed by adverse meteorology, leading to severe air pollution events. These examples indicate that, in absence of weather adjustments, pollution estimates need to be averaged over extended time periods to reduce weather influences; an approach that is taken by many studies (as mentioned above). The importance of weather adjustments is frequently acknowledged, but not often addressed. Specifically, Gkatzelis et al. (2021) found that two thirds of studies that quantified changes in air pollution did not correct for weather effects.

Studies that did adjust for weather influences used a variety of approaches, including difference-in-difference methods (e.g., Navinnya et al., 2020), generalised additive models (e.g., Ropkins and Tate, 2021) and atmospheric chemical transport models (e.g., Zhao et al., 2020). Further, several studies have investigated the use of machine learning approaches. For example, Grange and Carslaw (2019) investigated the normalisation of time series with respect to weather effects using a random forest model, obtaining time series under average weather conditions. The use of non-parametric machine learning methods for meteorological normalisation has considerable advantages over parametric alternatives such as regression analysis, as atmospheric processes are complex, non-linear, and variables frequently exhibit high multicollinearity (Grange and Carslaw, 2019). Further, Petetin et al. (2020) quantitatively illustrated the benefits of using a machine learning approach to incorporate meteorological variability, as opposed to directly comparing measurements in previous years to 2020 levels. Importantly, machine learning is especially beneficial for accurately predicting pollution at finer spatial and temporal scales (Petetin et al., 2020). Therefore, for a robust comparison to daily city-level mobility data (i.e., fine spatial and temporal scale), accurate weather-normalised pollution estimates are a prerequisite.

With respect to the selection of a particular machine learning approach, Ma et al. (2020a), ZamaniJoharestani et al. (2019) and Ren et al. (2020) all found that XGBoost (Chen and Guestrin, 2016) performed better for air pollution modelling than other machine learning techniques such as random forest and support vector regression. Further, Ma et al. (2020b) showed that an XGBoost model significantly improved the PM_{2.5} forecasts of an operational atmospheric chemical-transport model at the Shanghai Meteorological Service. Our research should be seen in light of these developments. In this paper, we present a city-level analysis at a global scale, including adjustment for weather variables such as temperature, wind speed and precipitation. City- and pollution-specific modelling based on historical atmospheric data and annual trends, provides air pollution estimates for 2020 in absence of a pandemic (i.e., a counterfactual, ‘business as usual’ scenario). Various methods were explored to optimise forecast accuracy. The identified pollution anomalies were then compared to country-specific government policies intended to limit the spread of COVID-19. Hence, our research investigates air pollution both at the micro- and the macro-level. It explores how different countries obtained different results based on the stringency of restrictions and associated changes in city-level mobility patterns.

2. Materials and methods

2.1. Data

The following global data sources were selected to provide information on air pollution levels, weather, the severity of government restrictions, and mobility patterns of city residents. Data was checked for completeness, cleaned and processed using Java, Python and R.

2.1.1. Pollution

The World Air Quality Index project (AQICN, 2021) provides ground-level readings of pollutants for cities in 132 countries sourced from world-wide environmental protection agencies. Daily measurements of NO₂, PM_{10}, PM_{2.5} and O₃ were downloaded over the period 2015–2020 for approximately 900 cities (i.e., a subset of the 1692 largest cities in the world (United Nations, 2015) for which sufficient data was available). All measurements were a 24-h average of hourly readings, accounting for the local time zone. Since the data provider quoted pollution measurements in terms of AQI levels based on standards of the Environmental Protection Agency (Gilliam and Hall, 2016), measurements were converted back to their original unit for comparative analysis in this study. Our final dataset consisted of NO₂ and O₃ measurements in parts per billion (ppb) and PM_{10} and PM_{2.5} measurements in µg/m².

2.1.2. Weather

Meteorological data was obtained from the ERA5 reanalysis (Hersbach et al., 2020) of the European Centre for Medium-Range Weather Forecasts, providing a large range of atmospheric, land and oceanic climate variables. Importantly, the ERA5 reanalysis combines past observations with models to generate globally consistent time series, while limiting missing data. Hourly (solar radiation and precipitation only) and 8-hourly ERA5 data was downloaded at a 0.25° × 0.25° resolution via the Copernicus Climate Change Service Climate Data Store. Using each city’s latitude and longitude coordinates (United Nations, 2015), daily city-level information was extracted from the grid cell containing the location of the city centre. The following atmospheric and land-based variables were obtained for each 24-h period based on the local time zone:

- mean air temperature at 2 m above the surface (K);
- total net solar radiation at the surface (J m⁻²);
- total precipitation (m);
- mean wind speed (m s⁻¹);
- mean wind direction (degrees); and
- mean leaf area index of vegetation (m² m⁻²).

2.1.3. Stringency index

Governments adopted a broad variety of (initial) approaches to deal with the COVID-19 pandemic, ranging from trying to maintain business as usual (e.g., Brazil, Sweden), to imposing strict lockdowns and border closures in attempts to eliminate the virus (e.g., China, New Zealand). A globally consistent assessment of the stringency of these policy measures has been provided by the Oxford COVID-19 Government Response Tracker (Hale et al., 2021). This daily, country-level stringency index is based on a combination of containment and closure policies, providing an overall score of a country’s physical distancing policies between 0 (no restrictions) and 100 (most severe restrictions). Components of the stringency index include the severity of restrictions on gatherings, cancellation of public events, stay-at-home requirements, school, university, workplace and public transport closures, local, regional and international travel controls, and the presence of public information campaigns. The stringency index purely evaluates the strictness of government policies, without providing a qualitative assessment on the appropriateness of restrictions.

2.1.4. Mobility

Google’s COVID-19 Community Mobility Reports (Google, 2020) were downloaded to assess the mobility patterns of city residents. These reports are based on GPS location information of users who had turned on Location History on their mobile phone. Google assigns location information to places using six categories: retail and recreation, grocery and pharmacy, parks, transit stations, workplaces, and residences. The total number of visitors is recorded for each category, except for residences which is measured by average time spent. Reports provide
a daily percentage change, comparing actual mobility to levels before widespread disruptions due to the pandemic. Baseline mobility levels represent the median value for the corresponding category and day of week during the five-week period between January 3 and February 6, 2020. The Community Mobility Reports consist of county/province-level data to preserve privacy. In this research, county-level data was fully allocated to a single city when a city was located inside the county area (i.e., Clark County for Las Vegas), or combined when the city extended across multiple counties (i.e., Bronx, Kings, New York, Queens, and Richmond counties for New York City).

### 2.2. Modelling

As air pollution levels strongly depend on meteorological conditions, pollution forecasts should be adjusted for weather-related effects before analysis. In this study, the relationship between weather and air pollution was modelled based on historical data from January 2015 until December 2019. A separate model was calibrated (i.e., fitted) for each city, as city topography and the presence of pollution sources in or near the city have a strong influence on the impact of changes in weather. For example, southerly winds in Melbourne (Australia) transport clean marine air masses to the city, while the same southerly winds in continental Europe might bring in pollution from nearby cities. Further, four separate models were calibrated for pollutants NO\(_2\), PM\(_{10}\), PM\(_{2.5}\) and O\(_3\). To ensure a sufficient amount of ground-level pollution data is available for model calibration and analysis, cities were excluded if less than 365 training samples were available or more than 30 pollution measurements in 2020 were missing. This led to a sample of 720, 710, 751 and 707 cities for NO\(_2\), PM\(_{10}\), PM\(_{2.5}\) and O\(_3\), respectively. Fig. 1 provides a schematic overview of the modelling process for each city and pollutant, which is described in more detail in the following paragraphs.

City-specific model forecasts based on the actual weather in 2020 provided an indication of pollution levels if the COVID-19 pandemic had not occurred. Model forecasts for 2020 were then compared to observed air pollution levels in 2020 to obtain the forecast errors (i.e., residuals). Residuals of the model consist of (i) imperfections in the statistical model for the weather–pollution relationship, (ii) the impact of pandemic-related government restrictions and reduced mobility on air pollution, and (iii) other confounders such as bushfires. Therefore, it is important to model the weather–pollution relationship with high accuracy, reducing noise in subsequent analyses.

#### 2.2.1. XGBoost

Machine learning was used to model city-specific air pollution with weather-related variables. Specifically, this study used XGBoost, an eXtreme Gradient Boosting algorithm based on decision trees. XGBoost is an adaptation of gradient boosting machines (Friedman, 2001), which perform additive optimisation in functional space. This means that a sequence of decision trees is calibrated, where each decision tree aims to explain the residuals left by the previous tree. Compared to gradient boosting machines, XGBoost adds a regularisation constraint to the objective function to prevent overfitting and includes several methodological improvements to enhance scalability. Model calibration was performed on the University of Melbourne’s high-performance computing system (Lafayette et al., 2016), using the XGBoost implementation in R.

#### 2.2.2. Hyperparameter tuning

XGBoost has several hyperparameters that influence the learning process. For example, the shrinkage parameter \(\eta\) controls the learning rate of the algorithm, by scaling down the contribution of each decision tree to the pollution estimate. Generally, a low \(\eta\) leads to a more robust model by limiting the influence of a single decision tree, although resulting in longer calibration times (i.e., more trees are required). Further, constraints on the structure of each decision tree can limit model complexity. These constraints include the maximum depth of a tree \((\tau_d)\) and the minimum number of samples required in a single node of a tree \((\tau_s)\). The latter reduces complexity by preventing further splits when the data does not provide sufficient evidence (i.e., only a small number of samples is available at the node). Further, the calibration of each decision tree using a random sample of the data can improve model performance by reducing the correlation of subsequent trees. Sub-sampling can be controlled using hyperparameters for the ratio of training samples \((\rho_s)\) and the ratio of features \((\rho_f)\) that is used for model calibration.

In this study, hyperparameters were tuned using grid search. Initial experiments used a large grid for a small sample of models \((n = 45; 15\) cities and 3 pollutants). The initial grid, presented in Table 1, led to a total of 720 combinations of hyperparameter values. However, extensive experimentation on the high-performance computing cluster was not feasible using a grid search this large. Note that a single experiment involved calibrating 2888 XGBoost models, capturing over 700 cities and four different pollutants. Further, each of these model calibrations involved a grid search over the hyperparameter space, including a five-fold cross-validation for each possible combination of the hyperparameters and up to 1000 decision trees per sub-model.

### Table 1

| Hyperparameter | Initial set | Final set |
|----------------|-------------|-----------|
| \(\eta\)       | \((0.01, 0.05, 0.10)\) | \((0.01, 0.05)\) |
| \(\tau_d\)     | \((2, 4, 5, 6, 8, 10)\) | \((5, 6, 8)\) |
| \(\tau_s\)     | \((1, 3, 5, 7)\) | \((5, 7)\) |
| \(\rho_s\)     | \((0.5, 0.6, 0.7, 0.8, 1)\) | \((0.5, 0.6)\) |
| \(\rho_f\)     | \((0.8, 1)\) | \((1)\) |

Fig. 1. Flowchart describing the modelling process that was followed for each city and pollutant.
Experiments using the full grid aimed to narrow down the range of hyperparameter values that performed well on our dataset. Hyperparameter combinations were ranked based on lowest root mean squared error (RMSE) in cross-validation. The final set of hyperparameter values in Table 1 were frequently present in the top models of the sampled cities. Therefore, a grid search over these values, a total of 24 hyperparameter combinations, was used for the full-scale experiments in this study. In particular, a low learning rate ($\eta \in (0.01, 0.05)$) reduced overfitting by making the boosting process more conservative. Further, the minimum number of instances required in each node was increased from the default setting (i.e., 1) to improve generalisability ($\tau \in (5, 7)$). Sub-sampling was applied to samples, but not to features. That is, either 50% or 60% of observations was selected at random to calibrate each decision tree, while a tree could use all weather-related variables. Besides limiting overfitting, sub-sampling also reduced computation time for a single tree, aiding the processing of many cities. Strongly restricting the maximum depth of each tree did not lead to good overall accuracy. Hence, values for this hyperparameter were set to $d \in (5, 6, 8)$.

### 2.2.3. Feature selection

Various feature selection experiments explored the best approach to incorporate the meteorological variables in the XGBoost models, based on the mean absolute error (MAE) of out-of-sample forecasts for the first two weeks of January 2020. This indicated that some variables have increased prediction power at different time lags. It was explored whether to include the value on the day itself, day $t-1$, $t-2$, $t-3$, or the average over the last or preceding three days ($t-2$ to $t$, or $t-3$ to $t-1$), taking into account interactions with the other features. Adding additional features capturing weather during the preceding three days improved model accuracy the most. Specifically, the total precipitation, mean temperature and mean solar radiation over the preceding three days were added as features. Wind speed and direction were most predictive when measured on the day itself. Since the leaf area index did not fluctuate substantially from day to day, only the measurement on day $t$ was used. Finally, a time variable was included to capture any annual trends in air pollution levels, as government policies sometimes resulted in a negative trend over the 2015–2019 period (e.g., cleaner fuels, gradual introduction of electric vehicles, etc.). An overview of the final selection of variables is provided in Table 2.

Some variables that have potential to further improve pollution forecasts have not been included in the modelling on purpose. For example, a ‘day of week’ variable can explain weekly patterns in air pollution caused by mobility fluctuations (i.e., limited travel during weekends). However, since mobility will be explored later on, models did not incorporate any mobility-related variables even if this could have led to higher accuracy. The only features used in the XGBoost models were atmospheric, vegetation-related and temporal variables, capturing weather, seasonality and any annual trends in air pollution levels.

### 2.2.4. Model forecasts and assessment

Final XGBoost models were calibrated using the full training set, containing the features in Table 2. Models used the city- and pollutant-specific hyperparameter set that resulted in the lowest RMSE in cross-validation. To improve explainability, variable importance in the final models was assessed through ranking. Specifically, tree-based models such as XGBoost allow for an investigation of the contribution of each feature to the model (i.e., gain). The gain quantifies how important a feature is in making a branch of a decision tree more pure, based on the sum of improvements in squared error over all internal nodes where the feature was chosen as a splitting variable (Hastie et al., 2009, p. 367–368). The averaged feature importance characteristics over all decision trees in one XGBoost model provide an overall score of feature importance. Feature importance was explored in various case studies to provide further insights into the modelling approach.

Each calibrated model was used to predict air pollution levels for the corresponding city and pollutant throughout 2020. These weather-based forecasts for 2020 do not incorporate any pandemic-related reductions in emissions. Hence, forecasts provide a counterfactual indication of pollution levels had a pandemic not occurred (i.e., the XGBoost models are agnostic of the pandemic). The differences between 2020 forecasts and the actual pollution measurements in 2020 are referred to as ‘unexplained pollution’, representing the meteorology-normalised reduction in air pollution. Importantly, a separate XGBoost model was calibrated for each city, independently. Therefore, city-specific air pollution models provide a multi-model ensemble forecast for the unexplained pollution in a country (i.e., for air pollution in large cities with a population over 300,000).

Finally, the unexplained NO\textsubscript{2} pollution was selected to explore associations with mobility patterns using the Google Mobility Reports. Motor-vehicle traffic is a major source of NO\textsubscript{2} pollution, while many emission sources/processes contribute to PM\textsubscript{10}, PM\textsubscript{2.5} and O\textsubscript{3} levels. Since the NO\textsubscript{2} pollution has been adjusted for weather influences, seasonality and annual trends, unexplained pollution estimates could theoretically be allocated to mobility levels on a day-to-day basis. However, the Google Mobility Reports used a different baseline for each day of the week, complicating the analysis of mobility data for consecutive days. Therefore, 7-day smoothing was applied to both time series for a more robust comparison. The Google Mobility Reports stratify mobility reductions based on trip purpose. A correlation analysis was performed to determine which type of trips were most highly associated with unexplained NO\textsubscript{2} levels. Correlations were computed between 15 February and 15 April 2020, capturing more regular mobility levels during the pre-lockdown period and subsequent reductions during the lockdown period. Analysis was performed at a country-level, showing how these relationships varied, globally.

### 3. Results

#### 3.1. Micro level

#### 3.1.1. Model fit

In total, 2888 XGBoost models for the weather-pollution relationship were calibrated, corresponding to 720, 710, 751 and 707 cities for NO\textsubscript{2}, PM\textsubscript{10}, PM\textsubscript{2.5} and O\textsubscript{3}, respectively. The average explained variation in training data, as measured by $R^2$, was 86.3%, 85.8%, 86.2% and 86.3% for each of these four pollutants. A wide variety of seasonal patterns were observed for different cities and pollutants, which were adequately captured by the XGBoost models. For example, Fig. 2 provides out-of-sample XGBoost model forecasts for 2020, showing NO\textsubscript{2} in Chongqing, PM\textsubscript{10} in Antwerpen and PM\textsubscript{2.5} in Ulanbaatar. The assessment of out-of-sample performance using 2020 measurements is complicated, because of the influences of the pandemic. However, results of further statistical model validation have been provided in Section 3.3.
3.1.2. Feature importance

Tables 3 and 4 show the feature importance statistics per pollutant, computed across all calibrated XGBoost models. For each feature in Table 3, the corresponding gains of all 2888 models were averaged to obtain the mean gain. Note that the gain is an indicator of relative feature importance (i.e., how important is a feature compared to the other features) and the gains across all features sum to 1. The three-day temperature average and the leaf area index were the most useful features for NO$_2$, PM$_{10}$ and PM$_{2.5}$, while the current solar radiation was most important for O$_3$. Variations in temperature and the amount of foliage throughout the year are both indicators of seasonal influences that affect air pollution levels. Urban vegetation also impacts pollution levels directly by removing a substantial amount of NO$_2$ and particulate matter from the air (Nowak et al., 2006). In contrast, temperature levels have a significant effect on the emission of pollutants into the atmosphere. For example, the increased pollution levels in Ulaanbaatar, Mongolia (see Fig. 2c) were mainly caused by the use of heating stoves and heat only boilers during winter (The World Bank, 2009). The XGBoost models were in line with these observations, as $T$ was the most important feature to predict air pollution in Ulaanbaatar (i.e., $T$ gain = 0.695 for PM$_{2.5}$). High temperatures can also result in increased pollution levels through atmospheric chemistry (von Schneidemesser et al., 2015) or increased energy consumption such as through the use of air conditioning (Davis and Gertler, 2015). For example, Fig. 2b shows increased PM$_{10}$ levels during the early August 2020 heat wave in Belgium, captured accurately by the XGBoost model.

The top features per pollutant were not always the best-performing features, as considerable variation in feature importance was observed across cities (see Table 4). For example, wind direction was the most important feature for several coastal cities (e.g., Liverpool, Nantong, Jakarta). In these cities, clean ocean air can substantially reduce pollution levels, while unfavourable wind conditions transport pollution from other cities or nearby power plants into the city. This assertion is supported by air pollution, wind speed and wind direction observations during 2015–2019. Bivariate polar plots of these observations (created using the polarplot/openair R package by Carslaw and Ropkins, 2012) illustrate how air pollution levels in these cities varied based on wind direction and speed. In Liverpool (Fig. 3a), wind directions between East-northeast (ENE) and South-southeast (SSE) transport particulate matter from nearby cities Manchester, Birmingham and London, and continental Europe into the city (Graham et al., 2020). Further, wind directions from the ocean transport clean maritime air masses, substantially reducing air pollution in Liverpool and Nantong (Fig. 3a–b).

Wind speed was a slightly more informative feature in NO$_2$ models than in PM$_{10}$, PM$_{2.5}$ and O$_3$ models. In contrast, the total precipitation over the past three days ($P$) was more useful for predicting PM$_{10}$ and PM$_{2.5}$ than NO$_2$ and O$_3$. This finding is consistent with the larger washout effect of precipitation for particulate matter compared to NO$_2$ (Yoo et al., 2014). Due to many null observations, $P$ and $P$ are not a differentiating feature in cities with a mostly dry climate (e.g., PM$_{2.5}$ gain = 0.023, 0.019) and (0.024, 0.010) for PM$_{10}$ and PM$_{2.5}$ in the semi-arid climate of Aguascalientes, Mexico). However, the precipitation variables rank highly for some cities with a sufficient amount of rainfall. For example, $P$ was the most influential predictor for PM$_{2.5}$ in Kathmandu (Nepal), which receives a substantial amount of rain annually while the local topography limits the ability of wind to clear the area of pollution (see Fig. 3c). Overall, the results of the feature importance analysis support our approach to model air pollution separately for each city and pollutant. This allowed the XGBoost models to adapt to the distinct urban features, climate and atmospheric conditions for each city, individually.

3.2. Macro level

After exploring micro-level effects in Sections 3.1.1 and 3.1.2, this section will focus on macro-level effects. Figs. 4–6 show the results of the ensemble forecasts for the unexplained pollution in several countries, using the XGBoost models of all corresponding cities. The figures
present the absolute reduction in NO\(_2\), PM\(_{2.5}\) and O\(_3\), stated in ppb, \(\mu g/m^3\) and ppb, respectively. For brevity, PM\(_{10}\) has not been included, as country-level results showed similar patterns for unexplained PM\(_{2.5}\) and PM\(_{10}\) (even though PM\(_{2.5}\) and PM\(_{10}\) models were calibrated independently). The unexplained pollution is the difference between 2020 measurements and the forecasts of the XGBoost models using weather, vegetation and temporal variables. Hence, these outcomes are agnostic of the pandemic and the associated restrictions on mobility. However, the unexplained pollution matched the timing of the restrictions in different countries very well. In particular, the unexplained pollution in China peaked between late January and early February, consistent with the lockdown during this period (see first column of Fig. 4). In other countries, the peaks were observed around early April. In larger countries such as China and India, regional lockdowns led to larger confidence intervals, as restriction levels varied across cities (He et al., 2020). In smaller countries restrictions generally applied more uniformly, resulting in narrower confidence intervals. As the XGBoost models did not incorporate mobility-related variables, many of the unexplained pollution charts still showed 7-day temporal patterns. This is especially visible for NO\(_2\) in the United States and several European cities, showing 52 peaks of reduced pollution corresponding to weekends in 2020. The same patterns can be observed for the mobility time series, for example, capturing different effects of the pandemic on workplace travel during weekdays and weekends.

Besides differences in timing, the amplitude of the pollution reductions also varied substantially between countries. For NO\(_2\), the largest reductions were observed in China, India and Europe. Further, the amplitude of NO\(_2\) reductions was in line with the magnitude of observed reductions in mobility for the corresponding country. Large reductions in PM\(_{2.5}\) were mainly observed for China and India. Note that in absolute terms only slight reductions in PM\(_{2.5}\) were apparent for some countries. However, these reductions were still substantial in relative terms, as the regular ranges of PM\(_{2.5}\) levels were already significantly lower than in China and India to start with. Since negative health effects are computed based on absolute levels, less benefits may be obtained in these countries by interventions targeting reductions in mobility, or alternative transport initiatives.

Most existing studies have investigated air pollution during the initial lockdown periods in March and April 2020. Our study found that after the initial lockdowns, NO\(_2\) and PM\(_{2.5}\) emissions did not fully return to pre-pandemic levels. Reduced mobility during most of 2020 led to a sustained reduction in air pollution, although not to the same extent as during the initial lockdowns. For O\(_3\), many studies reported an increase during the initial lockdowns (see Section 1.1.3), consistent with our results. However, from June/July to September, our analyses found a subsequent reduction in O\(_3\) for many countries below what was expected based on meteorological conditions (e.g., China, United Kingdom, France, Germany, Poland, Turkey). This period corresponds to summer in the Northern Hemisphere, when O\(_3\) levels are normally elevated. Note that photochemical production of O\(_3\) at mid and high latitudes in the Northern Hemisphere is low in winter due to reduced sunlight and temperatures (Dentener et al., 2020). Most studies reported increases in O\(_3\) during March and April, from already low levels. However, the implications in periods with high O\(_3\) levels are especially important with respect to health consequences. This should be investigated further in future research.

Figs. 4–6 also illustrate the effects of different baselines for each day of the week in the mobility dataset. This resulted in positive peaks during weekends, corresponding to lower mobility reductions with respect to already low baseline levels (e.g., see workplace travel). In contrast, the unexplained NO\(_2\) pollution estimates showed negative peaks during weekends. Table 5 reports correlations after 7-day smoothing between the NO\(_2\) and mobility time series from mid-February to mid-April, linking mobility patterns for each country to the unexplained pollution levels. Correlations were reported for all countries where at least one city was modelled, noting that one city with complete measurements resulted in \(n = 61\) data points. For various countries, the total unexplained NO\(_2\) pollution was highly correlated with changes in mobility levels. Exposures for each trip type were different per country (e.g., travel to workplaces may constitute the majority of trips), indicating that policy interventions in some categories may be more promising to explore than others. Trips to transit stations, workplaces, retail and recreation venues generally had a higher correlation with unexplained pollution than trips to groceries and pharmacies. Further, park visits varied significantly between countries. In most of the Nordic countries (Denmark, Finland, Norway) the correlations between park usage and unexplained pollution were opposite to the rest of the world, presumably because the baseline levels for park visits during winter were very low due to the cold weather conditions in these countries.

3.3. Robustness of results

To assess the impact of modelling assumptions used in this research, various alternative approaches were explored. Specifically, the weather–pollution relationship was modelled using linear regression, generalised linear models (GLM) using various link functions, and random forests. The goodness-of-fit on training data, averaged across all cities and pollutants, varied from linear regression (\(R^2 = 36.7\%\)), Gamma GLM with inverse link function (\(R^2 = 40.6\%\)), random forest (\(R^2 = 50.3\%\)), to XGBoost (\(R^2 = 86.2\%\)). As the goodness-of-fit of each model was different, Table 6 shows the variation of model fit across cities. Further, a time series modelling approach without weather variables was explored. As the time series forecasting approach differed substantially from our preferred XGBoost method described in Section 2.2, it is described here in more detail. First, one time series was created per pollutant and city based on pollution measurements during 2015–2019. Missing values were imputed, or the time series was disregarded if there were too many missing values. Each remaining time
Fig. 4. Unexplained NO$_2$ (ppb) and PM$_{10}$ (μg/m$^3$) across cities, stringency of COVID-19 restrictions, and mobility patterns for China, United States and Italy. Shading indicates 50, 80 and 90% confidence intervals. Google mobility data was not available for China; US park visits between mid-August and September had data quality issues.
Fig. 5. As Fig. 4, but for Spain, France and Germany. Park usage in France and Germany exceeds +100% during summer months, compared to the January baseline.
Fig. 6. As Fig. 4, but for India, Japan and United Kingdom.
series was decomposed into seasonal, trend and remainder components using Loess (STL) (Cleveland et al., 1990). The trend and remainder components were then recombined and used to calibrate an ARIMA model. A 2020 forecast was computed as the prediction of the ARIMA model plus the seasonal component obtained earlier. This resulted in a severe outbreak of COVID-19 during 2020 and residents in Santiago were subject to an extended 143-day lockdown period from March to mid-August. The various methods were consistent in predicting substantially higher NO\textsubscript{2} levels for the contextual scenario, based on the observed weather. Finally, the large seasonal PM\textsubscript{10} fluctuations in Pune (India) were captured accurately, including pollution spikes in November 2020, although forecasts also showed some variations between methods.

Tables 7 and 8 provide a statistical assessment of out-of-sample performance. The accuracy of pollution forecasts for the first two weeks of NO\textsubscript{2} during these periods, all methods forecasted substantially higher levels as part of very severe lockdown restrictions. Based on the weather of Ürümqi were locked inside their homes from 18 July to 31 August initial lockdown in China during February and early March, residents Ürümqi experienced multiple COVID-related lockdowns. Besides the captured the seasonal pollution patterns of the different cities. Notably, 10 and Santiago (Chile), and PM\textsubscript{10} in Pune (India), all methods correctly in Ürümqi (China) pollution normally than measured. In Santiago (Chile), NO\textsubscript{2} was observed to mid-August. The various methods were consistent in predicting a severe outbreak of COVID-19 during 2020 and residents in Santiago were subject to an extended 143-day lockdown period from March to mid-August. The various methods were consistent in predicting substantially higher NO\textsubscript{2} levels for the contextual scenario, based on the observed weather. Finally, the large seasonal PM\textsubscript{10} fluctuations in Pune (India) were captured accurately, including pollution spikes in November 2020, although forecasts also showed some variations between methods.

Tables 7 and 8 provide a statistical assessment of out-of-sample performance. The accuracy of pollution forecasts for the first two weeks
of January 2020 was computed based on MAE and RMSE across all cities. The effects of the pandemic on mobility levels were not as widespread in early January 2020. Hence, forecasting errors during this period can, to a large extent, be attributed to the quality of the statistical models, except for instances where increased pollution was not related to weather effects (e.g., the large-scale bushfires in Australia in January 2020, or reduced travel during the holiday period). Results show that the XGBoost models obtained the best out-of-sample performance, except for PM$_{2.5}$ and O$_3$, where the random forest approach led to a slightly lower MAE and RMSE.

Importantly, when aggregating results to country level, the alternative approaches led to similar results, but with an increased level of noise (see Fig. 8). The XGBoost method presented in this study had the highest prediction accuracy (based on training and out-of-sample data) and hence, provides the most accurate assessment of the impact of the pandemic on pollution levels. It also led to smaller confidence intervals compared to a time series or linear modelling approach. Replication of the main results using alternative methodologies provides further confidence in the robustness of the presented results. Deep learning methods have not been investigated in this study, but could potentially further improve accuracy.

Some XGBoost models had a lower goodness-of-fit on training data than the 86% average $R^2$, for example, due to remaining data quality issues or external events impacting pollution levels during 2015–2019. To assess the sensitivity of the results to model quality, various experiments were performed with different exclusion criteria for XGBoost models with a reduced quality of fit. For example, experiments were conducted that excluded XGBoost models with an $R^2$ on training data below 50%. This resulted in the exclusion of only 21 out of 2888 models (i.e., 0.7%) and had limited effects on the overall results. Finally, various experiments were performed using air pollution measurements obtained via satellite remote sensing, as opposed to ground-level measurements. Ground-level measurements were preferred, leading to improved correlations of the unexplained pollution with the mobility time series.

4. Discussion

The attribution of changes in air pollution levels for primary and secondary pollutants due to COVID-19 is complex. For example, the impact on primary pollutants is affected by the source profile of emissions applicable in a specific country. Government restrictions during the pandemic generally impacted transport emissions heavily and industrial activity to a lesser extent. In contrast, other sources such as agriculture were largely unaffected, while residential emissions mostly increased due to stay-at-home orders. Attribution is further complicated by government policies in response to climate change, extreme events such as large-scale bushfires, and weather conditions impacting atmospheric chemistry, emission levels, and the in-flow of particulate matter from nearby countries. Our study attempted to disentangle some of these confounding factors to directly compare absolute reductions in pollution levels due to the COVID-19 pandemic, globally.

4.1. Consistency with other publications

Daily meteorology-normalised pollution estimates were obtained for a wide variety of countries. As the temporal and spatial coverage is a superset of most other studies, results can easily be compared to those reported elsewhere. Table 9 provides a comparison to studies that investigated a specific country, region, or city. Our estimates are a slice of the full results for the same location/region and time period as reported by the respective study. Note that differences are expected due to variations in methodology, confounding factors that are adjusted for (if any), as well as the historical baseline period for reporting the change. Table 9 is presented for reference purposes and not intended to be an exhaustive comparison to all studies published in this domain.

Overall, results of our global analysis are broadly consistent with the patterns found in published region-specific case studies. For example, Archer et al. (2020) and Berman and Ebisu (2020) both investigated air pollution anomalies in the United States. Archer et al. (2020) compared the average NO$_2$ levels in April to the corresponding period in 2015–2019, observing a mean reduction of 2.02 ppb and reductions up to 5 ppb for some locations. Adjustments for the annual trend (but not for weather effects) led to a final estimate of --1.3 ppb. Berman and Ebisu (2020) found NO$_2$ reductions of 4.8 ppb in daily 1-h maximum NO$_2$, between 13 March and 21 April, compared to a baseline of 2017–2019, without adjustments for trends or weather effects. Our study found reductions in between these studies: a 2.5 ppb weather- and trend-adjusted reduction in 24-h average NO$_2$ between 1 April and 30 April, and 2.9 ppb between 13 March and 21 April.
country. Alternative data sources include Apple Maps Mobility Trends line in the same country, the measurements are consistent within a

However, as mobility is reported as the percentage change from base-

accuracy of place categorisation may vary from country to country.

reductions in pollution when also including rural areas. For the Google

comparatively high in urban locations, other studies may find lower

study to provide country-level pollution estimates. Since pollution is

major cities with a population over 300,000 were investigated in this

ensemble forecast, resulting in larger observed variations. Further,

did not have as many XGBoost models available for a multi-model

and 90% confidence intervals.

Sensitivity analysis, showing unexplained NO$_3$ (ppb) across cities in China (top row) and Italy (bottom row) for different modelling approaches. Shading indicates 50, 80 and 90% confidence intervals.

Fig. 8.

Comparison of results to other studies that quantified air pollution changes during lockdowns in 2020.

| Study               | Region            | Time period     | Pollutant | Reported estimate | Our estimate |
|---------------------|-------------------|-----------------|-----------|-------------------|--------------|
| Adams (2020)        | Ontario, Canada   | 22-Mar-25-Apr   | NO$_3$    | −2 ppb            | −1.9 ppb     |
| Adams (2020)        | Ontario, Canada   | 22-Mar-25-Apr   | PM$_{2.5}$| 0 μg/m$^3$        | −0.6 μg/m$^3$|
| Adams (2020)        | Ontario, Canada   | 22-Mar-25-Apr   | O$_3$     | −1 ppb            | 1.1 ppb      |
| Archer et al. (2020)| United States     | 1-Apr-30-Apr    | NO$_3$    | −2.02, −1.3 ppb$^a$| −2.5 ppb     |
| Archer et al. (2020)| United States     | 1-Apr-30-Apr    | PM$_{2.5}$| 0.05, 0.28 μg/m$^{3\alpha}$ | −0.1 μg/m$^3$|
| Berman and Ebisu (2020)| United States    | 8-Jan-12-Mar    | NO$_3$    | −1.17 ppb         | −0.6 ppb     |
| Berman and Ebisu (2020)| United States    | 13-Mar-21-Apr   | NO$_3$    | −4.76 ppb         | −2.9 ppb     |
| Berman and Ebisu (2020)| United States    | 8-Jan-12-Mar    | PM$_{2.5}$| −0.29 μg/m$^3$    | −0.6 μg/m$^3$|
| Berman and Ebisu (2020)| United States    | 13-Mar-21-Apr   | PM$_{2.5}$| −0.28 μg/m$^3$    | −0.1 μg/m$^3$|
| Connerton et al. (2020)| Los Angeles, USA | 1-Mar-31-Mar    | PM$_{2.5}$| −2.99 μg/m$^3$    | −1.9 μg/m$^3$|
| Connerton et al. (2020)| New York, USA    | 1-Mar-31-Mar    | PM$_{2.5}$| −2.03 μg/m$^3$    | −4.5 μg/m$^3$|
| Connerton et al. (2020)| Paris, France    | 1-Mar-31-Mar    | PM$_{2.5}$| −2.56 μg/m$^3$    | −2.4 μg/m$^3$|
| Connerton et al. (2020)| São Paulo, Brazil| 1-Mar-31-Mar    | PM$_{2.5}$| −0.54 μg/m$^3$    | −0.3 μg/m$^3$|
| Jia et al. (2020)   | Memphis, USA     | 25-Mar-4-May    | PM$_{2.5}$| 0.3 μg/m$^3$       | 0.1 μg/m$^3$ |
| Jia et al. (2020)   | Memphis, USA     | 25-Mar-4-May    | O$_3$     | −1.9 ppb          | −1.6 ppb     |
| Ordóñez et al. (2020)| Europe            | 15-Mar-30-Apr   | NO$_3$    | −9.2, −13.1 μg/m$^{1\beta}$ | −10.7 μg/m$^3$|
| Ordóñez et al. (2020)| Europe            | 15-Mar-30-Apr   | O$_3$     | 6.2, 0.1 μg/m$^{1\beta}$ | 4.5 μg/m$^3$ |
| Petetin et al. (2020)| Spain             | 14-Mar-29-Mar   | NO$_3$    | −3.4, −5.6 ppb$^c$ | −5.9 ppb     |
| Petetin et al. (2020)| Spain             | 30-Mar-9-Apr    | NO$_3$    | −5.2, −7.4 ppb$^c$ | −6.8 ppb     |
| Petetin et al. (2020)| Spain             | 10-Apr-23-Apr   | NO$_3$    | −4.3, −6.8 ppb$^c$ | −6.1 ppb     |
| Ropkins and Tate (2021)| United Kingdom   | 10-Mar-10-Apr   | NO$_3$    | −4.16, −7.58 μg/m$^{3\gamma}$ | −8.1 μg/m$^3$|
| Ropkins and Tate (2021)| United Kingdom   | 10-Mar-10-Apr   | PM$_{2.5}$| 4.79, 5 μg/m$^{3\gamma}$ | −1.7 μg/m$^3$|
| Ropkins and Tate (2021)| United Kingdom   | 10-Mar-10-Apr   | O$_3$     | 6.96, 7.39 μg/m$^{3\gamma}$ | 5.8 μg/m$^3$ |
| Tanner-Gruener et al. (2020)| Pittsburgh, USA| 14-Mar-30-Apr | PM$_{2.5}$| −2.8 μg/m$^3$     | −1.7 μg/m$^3$|
| Venter et al. (2020) | China             | 24-Jan-15-May   | PM$_{2.5}$| −16 μg/m$^3$      | −9.2 μg/m$^3$|
| Venter et al. (2020) | India             | 29-Feb-15-May   | PM$_{2.5}$| −15 μg/m$^3$      | −15.6 μg/m$^3$|
| Zheng et al. (2020) | Wuhan, China      | 23-Jan-22-Feb   | PM$_{2.5}$| −24.8 μg/m$^3$    | −28.9 μg/m$^3$|

*Comparison to 2015–2019 and trend-adjusted change, respectively.

*Comparison to 2015–2019 and meteorology-adjusted change, respectively.

*Meteorology-normalised changes at background and traffic sites, respectively.

*Break-point/segment analyses for urban background and traffic sites, respectively.

4.2. Limitations

Our study has a few limitations. For example, smaller countries did not have as many XGBoost models available for a multi-model ensemble forecast, resulting in larger observed variations. Further, major cities with a population over 300,000 were investigated in this study to provide country-level pollution estimates. Since pollution is comparatively high in urban locations, other studies may find lower reductions in pollution when also including rural areas. For the Google COVID-19 Community Mobility Reports, it should be noted that the accuracy of place categorisation may vary from country to country. However, as mobility is reported as the percentage change from baseline in the same country, the measurements are consistent within a country. Alternative data sources include Apple Maps Mobility Trends Reports (Apple, 2020), providing daily percent changes in Apple Maps routing requests from baseline levels at January 13, 2020. However, a limitation of this data source is that travellers mainly require routing information for new routes and the full extent of daily mobility on routes that are well-known to a traveller may not be captured.

4.3. Impact

It is well known that exposure to air pollution can lead to adverse health outcomes, such as respiratory and cardiovascular illnesses. Every year, ambient air pollution results in an estimated 4.2 million and 254,000 premature deaths from PM$_{2.5}$ and O$_3$ exposure, respectively (Cohen et al., 2017). Indoor air pollution leads to an additional 3.8 million deaths per year, caused by the use of solid fuels and kerosene for heating and cooking (World Health Organization, 2018). Globally, over half of deaths related to PM$_{2.5}$ exposure occur in China.
and India. PM$_{2.5}$ emissions related to residential energy use have the largest impact on mortality in these countries (Lelieveld et al., 2015). Our study shows the largest reductions in ambient particulate matter during the pandemic occurred in China and India, where adverse health consequences of air pollution are most prevalent. Other studies have shown that mobility restrictions also influence the intraday patterns of air pollutants. For example, mobility restrictions could eliminate the morning rush-hour peak in PM$_{2.5}$, reducing acute exposures in high-traffic environments (Tanzer-Gruener et al., 2020). However, He et al. (2020) found that PM$_{2.5}$ levels in China during lockdown periods were still four times higher than the WHO air quality guidelines for annual mean PM$_{2.5}$ of 10 µg/m$^3$ (World Health Organization, 2006). Lockdown periods also resulted in elevated exposure to indoor particulate matter. For example, residents in India spent over 30% more time at home during the initial lockdown (see Fig. 6). Although targeting mobility reductions in these countries can be effective in reducing exposure to ambient air pollution, the unintended consequences of increased exposure through indoor pollution should be carefully considered in any interventions.

With respect to O$_3$, reduced titration with NOx and volatile organic compounds was discussed in many studies as an explanation for increased O$_3$ levels. This process is the dominant mechanism affecting O$_3$ levels in winter (Yang et al., 2019). The majority of studies have investigated pollution levels during the initial lockdown, which occurred during late winter and early spring in the Northern Hemisphere. Our study also observed increased O$_3$ in various countries during these initial lockdowns. However, a subsequent reduction in O$_3$ was observed during summer months in the Northern Hemisphere (i.e., below levels expected based on meteorological conditions), which was outside the period of investigation of many other studies. As these reductions occurred when O$_3$ levels typically peak, this period is important to investigate with respect to the health consequences of altered O$_3$ levels. Hence, the formation of O$_3$ is an important research question to investigate further, especially in relation to the development of effective mitigation strategies.

5. Conclusions

The COVID-19 pandemic has prompted extraordinary measures to limit disease transmission, restricting mobility of citizens around the world. These restrictions also influenced air pollution, although quantification has proven to be complex. The aim of our study was to disentangle confounding factors in a city-level analysis at a global scale. Non-parametric machine learning methods were used to model the non-linear relationship between atmospheric processes and air pollution, adjusting for weather, seasonality, trends and city characteristics such as topography. Daily pollution levels throughout 2020 were predicted in a counterfactual scenario, agnostic of the pandemic, and compared to the observed air pollution. Finally, city-level predictions were assessed for each country, providing a country-specific multi-model ensemble forecast for the meteorology-normalised pollution reductions. It was shown that results were broadly consistent with those reported elsewhere. Hence, one of the contributions of this study is showing how case studies could be scaled up to a global level.

The unexplained pollution matched the timing of government restrictions in different countries very well, based on data from the Oxford COVID-19 Government Response Tracker. The largest reductions in NO$_x$ were observed during the initial lockdowns in China, Europe and India. Further, pollution did not fully revert to pre-pandemic levels throughout 2020. For particulate matter, the largest absolute reductions occurred in China and India. In many countries, increased O$_3$ levels were observed during the initial lockdowns, which has been documented widely (see Section 1.1.3). Reduced titration of O$_3$ with NOx and volatile organic compounds was generally postulated as the main driver; this process is the dominant mechanism affecting O$_3$ levels in winter. However, during the summer period from June/July to September, our analyses found a subsequent reduction in O$_3$ for many countries below what was expected based on meteorological conditions (e.g., China, United Kingdom, France, Germany, Poland, Turkey). This was outside the period of investigation of many other studies, but important for the development of effective mitigation strategies with respect to air pollution consequences. The adverse health consequences of air pollution are most prevalent in China and India. Although targeting mobility reductions in these countries can be effective in reducing exposure to ambient air pollution, the unintended consequences of increased exposure through indoor PM$_{2.5}$ should be carefully considered in any intervention. Note that in many countries, NOx reductions were highly correlated with changes in mobility levels, especially trips to transit stations, workplaces, retail and recreation venues. Therefore, alternative transport initiatives for these type of trips, such as electric vehicles, bicycles and ride sharing, could be an important pathway towards improved health outcomes.

CRediT authorship contribution statement

Jasper S. Wijnands: Conceptualisation, Methodology, Formal analysis, Writing - original draft, Writing - review & editing, Funding acquisition. Kerry A. Nice: Conceptualisation, Data curation, Writing - review & editing, Funding acquisition. Sachith Seneviratne: Methodology, Writing - review & editing. Jason Thompson: Writing - review & editing. Mark Stevenson: Writing - review & funding acquisition.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This work was supported by the Melbourne Energy Institute. J.T. is supported by an Australian Research Council Discovery Early Career Research Award [grant number DE180101411]. M.S. is supported by a National Health and Medical Research Council (Australia) Fellowship [grant number APP1136250]. The funding sources had no involvement in study design, analysis, writing, and the decision to submit the article for publication. The authors would like to acknowledge the three anonymous reviewers for their valuable feedback, which helped improve the quality of the original manuscript.

References

Adams, M.D., 2020. Air pollution in Ontario, Canada during the COVID-19 state of emergency. Sci. Total Environ. 742, 140516. http://dx.doi.org/10.1016/j.scitotenv.2020.140516.

Apple, 2020. Apple maps mobility trends reports. URL: https://covid19.apple.com/mobility.

AQICN, 2021. The world air quality index project. URL: https://aqicn.org (accessed 18 February 2021).

Archer, C.L., Cervone, G., Golbazi, M., Al Fahel, N., Hulquist, C., 2020. Changes in air quality and human mobility in the USA during the COVID-19 pandemic. Bull. Atmos. Sci. Technol. 1 (3), 491–514. http://dx.doi.org/10.1007/s42865-020-00019-9.

Bauwens, M., Coppenolle, S., Stavracos, T., Müller, J.-F., van Gent, J., Eskes, H., Levelt, P.F., van der A, R., Veefkind, J.P., Vlietinck, J., Yu, H., Zehner, C., 2020. Impact of coronavirus outbreak on NO2 pollution assessed using TROPOMI and OMI observations. Geophys. Res. Lett. 47 (11), e2020GL087978. http://dx.doi.org/10.1029/2020GL087978.

Beldibar, B., Apte, J.S., Müller, D.B., Robinson, A.L., Wells, K.C., Presto, A.A., Marshall, J.D., 2021. Changes in criteria air pollution levels in the US before, during, and after COVID-19 stay-at-home orders: evidence from regulatory monitors. Sci. Total Environ. 769, 144693. http://dx.doi.org/10.1016/j.scitotenv.2020.144693.

Berman, J.D., Ediss, K., 2020. Changes in U.S. air pollution during the COVID-19 pandemic. Sci. Total Environ. 739, 139864. http://dx.doi.org/10.1016/j.scitotenv.2020.139864.
United Nations, 2015. World Urbanization Prospects: the 2014 Revision. Technical Report ST/ESA/SER.A/366, Department of Economic and Social Affairs, Population Division, New York, NY.

Venter, Z.S., Aunan, K., Chowdhury, S., Lelieveld, J., 2020. COVID-19 lockdowns cause global air pollution declines. Proc. Natl. Acad. Sci. 117 (32), 18984–18990. http://dx.doi.org/10.1073/pnas.2006853117.

von Schneidemesser, E., Monks, P.S., Allan, J.D., Bruehwiler, L., Forster, P., Fowler, D., Lauer, A., Morgan, W.T., Paasonen, P., Righi, M., Sindelarova, K., Sutton, M.A., 2015. Chemistry and the linkages between air quality and climate change. Chem. Rev. 115 (10), 3856–3897. http://dx.doi.org/10.1021/acs.chemrev.5b00089.

Wang, P., Chen, K., Zhu, S., Wang, P., Zhang, H., 2020. Severe air pollution events not avoided by reduced anthropogenic activities during COVID-19 outbreak. Resour. Conserv. Recy. 158, 104814. http://dx.doi.org/10.1016/j.resconrec.2020.104814.

World Health Organization, 2018. Household air pollution and health. URL: https://www.who.int/news-room/fact-sheets/detail/household-air-pollution-and-health (accessed 27 July 2021).

World Health Organization, 2021. Coronavirus (COVID-19) dashboard. URL: https://covid19.who.int/ (accessed 20 October 2021).

Yang, J., Lia, J., Han, S., Yao, Q., Cai, Z., 2019. Study of the meteorological influence on ozone in urban areas and their use in assessing ozone trends in all seasons from 2009 to 2015 in Tianjin, China. Meteorol. Atmos. Phys. 131 (6), 1661–1675. http://dx.doi.org/10.1007/s00703-019-00664-x.

Yoo, J.-M., Lee, Y.-R., Kim, D., Jeong, M.-J., Stockwell, W.R., Kundi, P.K., Oh, S.-M., Shin, D.-B., Lee, S.-J., 2014. New indices for wet scavenging of air pollutants (O3, CO, NO2, SO2, and PM10) by summertime rain. Atmos. Environ. 82, 226–237. http://dx.doi.org/10.1016/j.atmosenv.2013.10.022.

Zamani Jokarastani, M., Cao, C., Ni, X., Bashir, B., Talebiesfandarani, S., 2019. PM2.5 prediction based on random forest, XGBoost, and deep learning using multisource remote sensing data. Atmosphere 10 (7), 373. http://dx.doi.org/10.3390/atmos10070373.

Zhao, Y., Zhang, K., Xu, X., Shen, H., Zhu, X., Zhang, Y., Hu, Y., Shen, G., 2020. Substantial changes in nitrogen dioxide and ozone after excluding meteorological impacts during the COVID-19 outbreak in mainland China. Environ. Sci. Technol. Lett. 7 (6), 402–408. http://dx.doi.org/10.1021/acs.estlett.0c00304.

Zheng, H., Kong, S., Chen, N., Yan, Y., Liu, D., Zhu, B., Xu, K., Cao, W., Ding, Q., Lan, B., Zhang, Z., Zheng, M., Fan, Z., Cheng, Y., Zheng, S., Yao, L., Bai, Y., Zhao, T., Qi, S., 2020. Significant changes in the chemical compositions and sources of PM2.5 in Wuhan since the city lockdown as COVID-19. Sci. Total Environ. 739, 140000. http://dx.doi.org/10.1016/j.scitotenv.2020.140000.