Estimation of NO$_2$ emission strengths over Riyadh and Madrid from space from a combination of wind-assigned anomalies and machine learning technique

Qiansi Tu$^{1,2}$, Frank Hase$^2$, Zihan Chen$^3$, Matthias Schneider$^2$, Omaira García$^4$, Farahnaz Khosrawi$^2$, Thomas Blumenstock$^2$, Fang Liu$^5$, Kai Qin$^6$, Song Lin$^1$, Hongyan Jiang$^{1,7}$, Dianjun Fang$^{1,8}$

$^1$Tongji University, School of Mechanical Engineering, Shanghai, China
$^2$Karlsruhe Institute of Technology (KIT), Institute of Meteorology and Climate Research (IMK-ASF), Karlsruhe, Germany
$^3$Karlsruhe Institute of Technology (KIT), Department of Informatics, Karlsruhe, Germany
$^4$Izaña Atmospheric Research Centre (IARC), Meteorological State Agency of Spain (AEMet), Tenerife, Spain
$^5$Beijing Chehejia Automobile Technology Co., Ltd., Beijing, China
$^6$China University of Mining and Technology, School of Environment and Spatial Informatics, Jiangsu, China
$^7$Jiangsu University of Science and Technology, School of Mechatronic and Power Engineering, Zhenjiang, China
$^8$Qingdao Sino-German Institute of Intelligent Technologies, Qingdao, China

Correspondence to: Matthias Schneider (matthias.schneider@kit.edu), Dianjun Fang (fang@tongji.edu.cn)

Abstract. Nitrogen dioxide (NO$_2$) air pollution provides valuable information for quantifying NOx emissions and exposures. This study presents a comprehensive method to estimate average tropospheric NO$_2$ emission strengths derived from three-year (April 2018 – March 2021) TROPOMI observations by combining a wind-assigned anomaly approach and a Machine Learning (ML) method, the so-called Gradient Descent. This combined approach is firstly applied to the Saudi Arabian capital city Riyadh, as a test site, and yields a total emission rate of 1.04×10$^{26}$ molec./s. The ML-trained anomalies fit very well with the wind-assigned anomalies with an $R^2$ value of 1.0 and a slope of 0.99. Hotspots of NO$_2$ emissions are apparent at several sites where the cement plant and power plants are located and over areas along the highways. Using the same approach, an emission rate of 1.80×10$^{25}$ molec./s is estimated in the Madrid metropolitan area, Spain. Both the estimate and spatial pattern are comparable to the CAMS inventory.

Weekly variations of NO$_2$ emission are highly related to anthropogenic activities, such as the transport sector. The NO$_2$ emissions were reduced by 24% at weekends in Riyadh, and high reductions are found near the city center and the areas along the highway. An average weekend reduction estimate of 30% in Madrid is found. The regions with dominant sources are located in the east of Madrid, where the residential areas and the Madrid-Barajas airport are located. Additionally, the NO$_2$ emissions decreased by 21% in March-June 2020 compared to the same period in 2019 induced by the COVID-19 lockdowns in Riyadh. A much higher reduction (60%) is estimated for Madrid where a very strict lockdown policy was implemented. The high emission strengths during lockdown only persist in the residential areas and cover smaller areas during weekdays than at weekends. The spatial patterns of NO$_2$ emission strengths during lockdown are similar to those observed at weekends in both
cities. Though our analysis is limited to two cities as testing examples, the method has proved to provide reliable and consistent results. Therefore, it is expected to be suitable for other trace gases and other target regions.

1 Introduction

Nitrogen oxides (NOx, also known as oxides of nitrogen) are a group of highly reactive trace gases (NO and NO2). NOx are toxic to human health and play a key role in tropospheric chemistry by catalyzing tropospheric O3 formation and acting as aerosol precursors, and this tropospheric O3 is a secondary pollutant that is also harmful to human health (IPCC, 2021). The emission of NOx is dominated by human activities and is mostly related to fossil fuel or biomass combustion (Goldberg et al., 2019). The major anthropogenic source in Europe is road transport (39%), followed by another four sectors with similar shares: energy production and distribution (14%), commercial, institutional, and households (13%), energy use in industry (11%) and agriculture (11%) (EEA, 2021). The near-surface abundance of NOx has generally increased with urbanization and industrialization (IPCC, 2021; Barré et al., 2021). Additionally, due to its short atmospheric lifetime (1 – 12 h) (Beirle et al., 2011; Stavrakou et al., 2013), NOx concentrations are highly variable and strongly correlated with local emission sources (Goldberg et al., 2019). Thus, NO2 observations can be considered as an excellent indicator to NOx emissions. The accurate knowledge of spatial and temporal distribution of NO2 atmospheric abundances, for this reason, is critical.

Space missions succeed in delivering well-resolved maps of tropospheric NO2 columns, from the early Global Ozone Monitoring Experiment (GOME) (Burrows et al., 1999), to the widely used Ozone Monitoring Instrument (OMI) (Boersma et al., 2007; He et al., 2021), to the latest TROPOspheric Monitoring Instrument (TROPOMI) (Veeckind et al., 2012). Among them, TROPOspheric Monitoring Instrument (TROPOMI) onboard Sentinel-5 Precursor (S-5P) since October 2017 has an outstanding importance. It is a push broom grating spectrometer, and measures direct and reflected sunlight in ultraviolet, visible, near-infrared, and shortwave infrared bands (Veeckind et al., 2012). TROPOMI offers daily coverage of data with an unprecedented spatial resolution of 3.5 × 7 km2 (3.5 × 5.5 km2 since August 2019) and a high signal-to-noise ratio (Copernicus Sentinel-5P, 2018; van Geffen et al., 2021). The TROPOMI NO2 data have been used for a variety of studies to estimate the NOx lifetime and emissions. For example, Lorente et al. (2019) has demonstrated that the strength and distribution of NOx emissions from Paris can be directly determined from the TROPOMI NO2 measurements. Beirle et al. (2019) mapped the NOx emissions on high spatial resolution based on the continuity equation and quantified urban pollution from Riyadh, Saudi Arabia (8.5 kg/s over 250 × 250 km²). A top-down NOx emission estimate approach was developed by Goldberg et al. (2019) and it reported that three megacities (New York City, Chicago, and Toronto) in North America emitted 3.9 – 5.3 kg/s NOx. Liu et al. (2020) demonstrated a 48% drop in the tropospheric NO2 column densities in China during the COVID-19 lockdown. The reductions of NO2 emission across the European urban areas resulting from the lockdown were studied by Barré et al. (2021) and -23% changes on average were obtained based on TROPOMI NO2 observations.
TROPOMI is unique due to its very high spatial- and temporal-resolution, which provides a large amount of data despite a planned mission lifetime of only about four years. This huge data set offers a possibility for its exploitation by the quickly developed artificial intelligence – Machine Learning (ML) techniques. For example, the application of ML to assess the NO\textsubscript{2} pollution changes during the COVID19 lockdown (Petetin et al., 2020; Keller et al., 2021; Barré et al., 2021; Chan et al., 2021). However, most studies focus on changes in NO\textsubscript{2} column abundances. The accurate amount and spatial pattern of deduced emission strengths are also important and can help air quality policy development.

In this study, the Gradient Descent (GD) approach in ML incorporating the wind-assigned method (Tu et al., 2022a, 2022b) is used to train the TROPOMI tropospheric NO\textsubscript{2} columns to estimate the NO\textsubscript{2} emission strengths of two (mega)cities: Riyadh (Saudi Arabia) and Madrid (Spain). The paper is organized as follows. Sect. 2 presents the data set and the combined method (wind-assigned and ML methods). The approach will be first applied to the Saudi Arabian capital city Riyadh for its evaluation and then applied to Madrid, followed by the discussion of the differences on weekdays and at weekends, and the changes before and during the COVID-19 lockdown period (Sect. 3). Conclusions are given in Sect. 4.

## 2 Data and Methodology

### 2.1 TROPOMI tropospheric NO\textsubscript{2} columns and wind data

We use operational Copernicus TROPOMI tropospheric NO\textsubscript{2} level 2 product (Copernicus Sentinel-5P (processed by ESA), 2021) from April 2018 to March 2021 over Saudi Arabian capital city of Riyadh and another (mega)city in Europe, such as Madrid, Spain. The quality flag (qa_value) is recommended to be >0.75, with which data are restricted to cloud-free (cloud radiance fraction < 0.5), and snow-ice-free observations (van Geffen et al., 2021). There are nearly 910,000 in Riyadh and 580,000 measurements in Madrid of good quality over three years. These observations are then binned for this study on a regular 0.1° × 0.1° grid, provided the number of observations is larger than 5 at each grid.

We use the horizontal wind information from the ERA5, which is the fifth-generation climate reanalysis produced by the European Centre for Medium-Range Weather Forecasts (ECMWF) at a spatial resolution of 0.25° × 0.25° (Copernicus Climate Change Service, 2017). Wind near the surface is significantly influenced by the local topography and we, therefore, use ERA5 at a higher level ~330 m above ground, referring to Figure A-1. This empirical choice is based on Tu et al. (2022b), in which the best fit between the estimated emission strength and the inventory for methane emission in the Upper Silesian Coal Basin (Poland) is obtained by using the ERA5 wind at ~330 m.

### 2.2 Wind-assigned and ML methods

The averaged distribution of emitted NO\textsubscript{2} over a long-term period can be approximated by an evenly distributed cone-shape plume, which is prescribed by wind speed and direction, and source strength with consideration of its temporal decay:
\[ \Delta \text{NO}_2(x,y,t) = \frac{\varepsilon}{v \cdot d(x,y)} \cdot \alpha \cdot \exp \left( \frac{-t}{\tau} \right) \]  
Eq. (1)

where \( \varepsilon \) is the emission strength and has an initialized value of \( 1\times10^{26} \) molec./s. The study area is binned on a regular \( 0.1^\circ \times 0.1^\circ \) grid and emission rates at each grid are assumed to be constant during the study period. \( \alpha \) is the angle of the emission cone and has an empirical value of \( 60^\circ \) (Tu et al., 2022a). \( d \) and \( \tau \) are the distance and transport time between the downwind location and \( \text{NO}_2 \) emission source, respectively. \( v \) is the wind speed from ERA5 and \( \tau \) is the lifetime/decay time for \( \text{NO}_2 \) and empirical values based on Beirle et al. (2019, 2011), i.e., 4 hours for Riyadh and 7 hours for Madrid, are used in this study. The daily plumes (\( \Delta \text{NO}_2 \)) from the individual emission source are computed based on Eq. (1) and then are super-positioned to have a total daily plume. The ERA5 model wind is divided into two opposite wind regimes based on the predominant wind regimes in each site (i.e., S: 90°-270° and N: the rest for Riyadh; SW: 135°-315° and NE: the rest for Madrid, see Figure A-1). A temporally averaged \( \Delta \text{NO}_2 \) plume is obtained for each wind regime and the difference between the two plumes generates the wind-assigned anomalies (more details in Tu et al., 2022a, 2022b).

The study area has \( x \times y (=N) \) grids. Each grid cell is considered as an independent point source at position \( (s_{\text{lat}}, s_{\text{lon}}) \), which yields a map of wind-assigned anomalies \( c_{s_{\text{lat}},s_{\text{lon}}} \). The wind information is assumed to be constant at each time over the study area in this study. The modeled wind-assigned anomalies derived from the point source located at the center grid \( (\text{lat}_{0}, \text{lon}_{0}) \) is considered as a parent map (see Figure A-2a):

\[ c_{s_{\text{lat}_{0}},s_{\text{lon}_{0}}} = (p_{\text{lat}_{0},\text{lon}_{1}} \cdots p_{\text{lat}_{0},\text{lon}_{j}} \cdots p_{\text{lat}_{0},\text{lon}_{y}}) \]  
Eq. (2)

The anomalies derived from other point source is identical to the parent anomalies, and value at each grid depends on the relative location to the parent one (see Figure A-2b):

\[ c_{s_{\text{lat}},s_{\text{lon}}} = (p_{\text{lat}_{0}-\text{lat}_{1},\text{lon}_{1}-\text{lon}_{0}} \cdots p_{\text{lat}_{0}-\text{lat}_{j},\text{lon}_{j}-\text{lon}_{0}} \cdots p_{\text{lat}_{0}-\text{lat}_{y},\text{lon}_{y}-\text{lon}_{0}}) \]  
Eq. (3)

These maps of wind-assigned anomalies at each grid are the inputs for the further step, which needs to be reformatted. The locations of the grids are renumbered in the sequence of latitude and longitude values from west to east and from north to south. The first grid at \( (\text{lat}_1, \text{lon}_1) \) locates in the far northwest and the last grid \( (\text{lat}_y, \text{lon}_y) \) locates in the far southeast. Therefore, each map of wind-assigned anomalies is converted to a new column vector \( c_k = (a_{k,1} \cdots a_{k,N})^T \), i.e., \( a_{k,k} \) represents the wind-assigned anomalies at k\textsuperscript{th} grid cell derived from point sources at k\textsuperscript{th} grid cell. The N grids generate N vectors to construct an \( N \times N \) matrix:

\[ M = (c_1 \cdots c_N) = \begin{pmatrix} a_{1,1} & \cdots & a_{1,N} \\ \vdots & \ddots & \vdots \\ a_{N,1} & \cdots & a_{N,N} \end{pmatrix} \]  
Eq. (4)

The estimated emission rate is a column vector \( w = (w_1 \cdots w_N)^T \). Since the emission rates cannot be negative, we use \( \log(w_k) \) as a proxy of the \( w_k \). The final result is then the exponent of the \( \log(w_k) \) and scaled by the initial \( \varepsilon \) of \( 1\times10^{26} \) molec./s. Then the modeled-calculated map \( m \) of the wind-assigned anomalies can be written as:
The wind-assigned anomaly method is also applied to the TROPOMI tropospheric NO2 column, yielding to a true map \( y = (y_1 \cdots y_N)^T \).

To estimate the emission strengths accurately, the modeled map \( m \) should approximate the true map \( y \). This problem is then converted to find the best \( w \) which results in the minimum value of the difference between \( y \) and \( m \), i.e., the cost function:

\[
L(y, m) = \frac{1}{N} \sum_{i=1}^{N} (y_i - m_i)^2 = \frac{1}{N} \sum_{i=1}^{N} (y_i - (a_{1,i} \cdots a_{N,i}) \times (w_1 \cdots w_N)^T)^2
\]

Eq. (6)

In our approach, the above equation can be considered as solving a linear system with constraints over the coefficients. In the ML framework, the popular GD algorithm can be a simple yet effective solution to find the coefficients. These coefficients can satisfy the approximation and the constraints at the same time, by formulating some of the constraints into the loss function that needs to be optimized. The main idea of GD is to find the partial derivatives of all coefficients in the system with respect to the loss function and use the local (gradient) information to reach the solution closer to the true state, which minimizes the approximation loss. In practice, this is implemented in an iterative process in which the data are sampled for the required gradients. However, there is only one single “data point” (one column vector) in our problem formulation. For each iteration (Eq. (7)), the new weight \( (w_{t+1}) \) is equal to the old weight \( (w_t) \) minus the gradient times the learning rate \( \eta \) (or so-called step size).

\[
w_{t+1} = w_t + \Delta w_t = w_t - \eta \times \sum_{i=1}^{N} \frac{\delta L}{\delta w_t}
\]

Eq. (7)

The selected areas in this study are highly isolated from the neighboring sources and thus, the emission rates at the edge can be assumed to be zero. However, the initialized constrain of them can increase the final biases. Therefore, we use a larger study area with \((n+1) \times (m+1)\) grids as the input data and remove the outmost ring to the target area of \( n \times m \) grids.

When applying GD for complicated systems with many parameters, there are many variations of GD that rely not only on the gradients but also introduce additional temporal information, i.e., the accumulation of gradients over time known as “momentum” to help GD converge faster and more reliably. Among those algorithms we decided to use Adaptive Moment Estimation (ADAM), because it is characterized by efficiency, little cost requirement (Kingma and Ba, 2015) compared to second-order methods such as BFGS (Broyden–Fletcher–Goldfarb–Shanno), and for our problem, it can slightly outperform other GD variations, such as the original gradient descent (GD) with momentum or Adadelta/Adagrad. In addition, it has been documented that it is superior by employing the cumulative first-order and second-order moments and, thus, become the de-facto method in the current deep learning scene when dealing with large data and parameters (Kingma and Ba, 2015).

It is notable that, in practice, GD or its variations are implemented under the name “Stochastic Gradient Descent” or “Mini-batch Gradient Descent” because using all datasets for one GD iteration is infeasible. The name “Stochastic” suggests having
the model parameters to be randomly generated and the dataset is randomly sampled into mini-batches (a small subset of the dataset) and the gradients are accumulated over the mini-batch solely for each update. This practice is not necessary with our problem.

3 Results and Discussion

3.1 Approach test for NO\textsubscript{2} emissions in Riyadh

Riyadh was chosen as the test site because this city with arid climate, has high NOx emissions due to the high population density (~4,300 residents/km\textsuperscript{2}; https://worldpopulationreview.com/world-cities/riyadh-population, last access: 29 March 2022) and it has punctual strong NOx sources close to the metropolitan area, such as cement plant and power plants. Moreover, Riyadh is remote from other sources, and has favorable weather conditions with low cloud cover and high surface albedo (Beirle et al., 2019; Rey-Pommier et al., 2022).

Figure 1 illustrates the averaged wind-assigned plumes derived from the TROPOMI tropospheric NO\textsubscript{2} and ML method over the analyzed period (April 2018 - March 2021). The ML-modelled plumes agree excellently with the satellite’s results (true map). A stronger plume is observed in the south of Riyadh, as the wind is more from the north (Figure A-1). The good correlation between these two maps is also presented in the one-to-one figure with an $R^2$ value of 1.0 and a slope value of 0.99 (Figure 1c). The estimated emission strengths based on the ML model show a very similar spatial pattern to the results in Beirle et al. (2019) (Fig. 2). Hotspots of NO\textsubscript{2} emissions are apparent at several sites where the cement plant and power plants are located and over areas along the highways (Figure 1d). These power plants have capacities larger than 1 GW and use crude oil and partly natural gas as fossil fuels (Beirle et al., 2019). The total emission rate is about $1.04\times10^{26}$ molec./s. Our estimate is slightly higher than the Beirle’s result ($8.3\times10^{25}$ molec./s from December 2017 to October 2018), in which wind fields from the ECMWF operational analysis at about 450 m above the ground were used (Beirle et al., 2019). The difference might be due to the different study periods and methods used. The wind speeds and directions are similar at lower levels (1000, 975, and 950 hPa) (Figure A-1), therefore, it is expected that wind at these levels introduces minor impacts on the estimates.

Figure 1: Wind-assigned plumes derived from (a) TROPOMI NO\textsubscript{2} and (b) ML method, (c) correlation plot between (a) and (b) for each grid, and (d) estimated emission rates in Riyadh, Saudi Arabia. Data in (a), (b), and (d) are oversampled to 0.1° ×0.1° resolution.
In (d) number in the figure’s title presents the total emission rate; triangle symbols and right-triangle symbol represent power plants and cement plant, respectively; grey lines represent the highways (data derived from www.openstreetmap.org © OpenStreetMap and www.mapcruzzin.com, last access: 11 April 2022).

3.2 NO$_2$ emission in Madrid

As a (mega)city in Europe, Madrid in Spain is another target in this study. The population of the Madrid metropolitan area is estimated to be about 6.7 million and nearly half of the residents live in Madrid city, resulting in a population density of ~5,400 residents/km$^2$ (https://worldpopulationreview.com/world-cities/madrid-population, last access: 29 March 2022). Figure 2a and b display the wind-assigned anomalies derived from TROPOMI observations and ML method, showing clearly pronounced bipolar plumes which are symmetrical in the Madrid city center. The ML-trained anomalies have a very good agreement with the TROPOMI one with an $R^2$ value of 0.99 and a slope of 0.99 (Figure 2c). The spatial pattern of estimated emission strengths is shown in Figure 2d, which is similar to that of CAMS-REG-AP (Copernicus Atmospheric Monitoring Service regional anthropogenic emission inventory, https://eccad.aeris-data.fr/catalogue/, last access: 31 March 2022; Granier, et al., 2019; Kuenen, et al., 2021) (Figure 2e). The CAMS-REG-AP covers emissions from the UNECE-Europe for the main air pollutants (e.g., NOx, expressed as NO$_2$) with a spatial resolution of $0.05^\circ \times 0.1^\circ - 0.1^\circ \times 0.1^\circ$ in longitude and latitude on a yearly basis over Europe (Kuenen et al., 2014). The total emission rate over the whole study is about $1.80\times 10^{25}$ molec./s, close to the CAMS inventory value of $1.12\times 10^{25}$ molec./s. Our estimate is smaller than a previous estimate of $6.8\times 10^{25}$ molec./s derived from the Ozone Monitoring Instrument (OMI) data during 2005-2009 (Beirle et al., 2011). This discrepancy is partly due to the different periods, methods, and data sets used. Apart from that, it is important highlighting that in the last decades, considerable efforts have been made in promoting the control and regulation of air quality policies across Europe (EEA report, 2020). In this context, Madrid City Council launched the Air Quality and Climate Change Plan for the city of Madrid (Plan A) in 2017, aiming at reducing pollution and adapting to climate change (https://www.madrid.es/UnidadesDescentralizadas/Sostenibilidad/CalidadAire/Ficheros/PlanAire&CC_Eng.pdf, last access: January 21, 2022). With an expectation, these actions may help to decrease NO$_2$ concentrations by ~25% in the central area by 2020. The binned emission rates agree well between the CAMS inventory and the ML-trained results with an $R^2$ value of 0.67 and a slope of 1.16 (Figure 2f). The ML-trained results are higher than the inventory. This is probably related to the fact that TROPOMI measures real-time NO$_2$ emissions which are not fully considered in the CAMS inventory.

Based on the spatial pattern, the dominant NO$_2$ sources can be easily distinguished. High NO$_2$ emissions are found near the city center, while the highest emissions are slight to the east, south, and southwest where the residential areas are located. The northwest of Madrid is the natural space and the Guadarrama mountains range runs in the NE-SW direction. Therefore, no obvious NO$_2$ sources can be found in these mountain regions. The Madrid-Barajas airport, which is the main international airport in Spain and the second-largest airport in Europe, is near the northeast of the city center where the region shows high NO$_2$ emissions. This is because aircraft exhaust emissions are highly enriched in NO$_2$ during taxiing and taking off (Herndon et al., 2004) and the near-airport NO$_2$ concentrations are higher than the emissions from highways and busy roadways (Hudda et al., 2020). In addition, the orographic feature, i.e., the development of mountain breezes along the slope of the Guadarrama
range causes the accumulation of pollutants in the NE-SW axis (Querol et al., 2018). Significant plumes of NO$_2$ columns are observed for wind from narrow wind regimes covering NE$_{1/2}$ (0°-90°) and SW$_{1/2}$ (18°-270°) (Figure A-3(a)-(b)). NO$_2$ accumulates near the city center for NW$_{1/2}$ wind (270°-360°), and a much weaker plume is found for SE$_{1/2}$ (90°-180°) wind regimes due to fewer wind days and weaker wind speed (Figure A-3(c)-(d)).

NO$_2$ emission changes on weekdays and at weekends

NOx emission variations result in significant changes in the weekly cycle, which is an unequivocal sign of anthropogenic sources (Beirle et al., 2003).

The estimated emission rates for weekdays (Sunday to Thursday) and weekends (Friday and Saturday) in Riyadh are presented in Figure 3. It should be noted that the weekends in Saudi Arabia are Fridays and Saturdays. The lowest NO$_2$ column abundances are observed on Fridays, followed by the ones on Saturdays (Figure A-4). The NO$_2$ emissions were reduced by 24% at weekends, and high reductions were found near the city center and the areas along Highway 65. Highway 65 is a major north-south highway in central Saudi Arabia and runs in the southeast-northwest direction, connecting Riyadh to Al Majma’ah in the northwest and to Kharj in the southeast (Figure A-5).

Figure 2: (a)-(d): same figures as Figure 1, but for the Madrid area, Spain. (e): spatial distribution of CAMS-REG-AP inventory, (f) correlation of emission rates between ML and CAMS-REG-AP inventory.
Significant column declines are found in large cities, especially in Europe, at weekends (Stavrakou et al., 2020). The weekly cycle of NO$_2$ column abundances in the Madrid area is different from that in Riyadh, as the lowest amounts are on Sundays, the second days of weekends (Figure A-6). An outstanding difference becomes apparent as much higher NO$_2$ amounts are found on working days, especially in urban areas. These high emissions are mainly due to road transport, which is the largest NOx contributor in Europe (Crippa et al., 2018) and emits up to 90% NO$_2$ in Madrid (Borge et al., 2014).

The ML-estimated emission strengths are presented in Figure 4. High NO$_2$ emission sources on weekdays are evenly distributed around the city center (Figure 4a). However, for weekends, the northeastern regions, far away from the city center, are the main sources, and no obvious sources are observed in the southwestern regions (Figure 4b). The total NO$_2$ emission strength in the urban area (dashed rectangles) during weekends ($7.24 \times 10^{24}$ molec./s) are smaller than those observed during weekdays ($9.85 \times 10^{24}$ molec./s) by about 26%. This result is similar to the result observed in another European city – Helsinki, where the weekly variability of traffic-related emissions was reduced by 30% at weekends (Ialongo et al. 2020). By subtracting weekends’ emissions from the ones of weekdays (Figure 4c), we found that the dominant NO$_2$ sources are in east-to-northeast and south-to-southwest regions, where the residential areas and working places are mainly located (Figure A-7). The orographic feature further causes the accumulation of NO$_2$ in these regions (see Section 3.2). The wind-assigned anomalies and correlation plots are presented in Figure A-8. Note that slightly higher scattering in the results at weekends is mostly due to fewer data points.
3.4 COVID-19 lockdown effect

The current global pandemic caused by coronavirus disease (COVID-19) largely impacts human life and the economic situation. To minimize the spread of the COVID-19 SARS-CoV-2 virus, countries around the world have enforced lockdown measures. Recent studies have reported decreasing NOx concentrations in the atmosphere due to lockdown, and additional reductions with more stringent lockdown, such as in Spain (Barré et al., 2021; Sun et al., 2021; Liu et al., 2021; Virghileanu et al., 2020; Keller et al., 2021; Bawens et al., 2020; Fan et al., 2020; Huang and Sun, 2020). The time series of TROPOMI tropospheric NO2 columns displays an obvious decrease since the lockdown started in early 2020 (Figure A-9). The NO2 amounts reach the lowest values in April 2020 and in the meanwhile are gradually back to normal levels as in previous years.

We analyze the same seasonal period in 2019 (before lockdown, March – June 2019) and in 2020 (during the lockdown, March – June 2020) for Riyadh and Madrid.

Figure 5 presents the spatial distribution of estimates before and during the lockdown in Riyadh. NO2 emissions decreased by 21% from $1.24 \times 10^{26}$ molec./s before lockdown to $9.79 \times 10^{26}$ molec./s during lockdown. The spatial distribution of estimates during lockdown is similar to that at weekend, when significant decreases are observed along Highway 65 and emissions are generally reduced in the city center and in the areas where the cement plant and power plants are located.
The NO$_2$ emission estimate in the urban area is about $1.04 \times 10^{25}$ molec./s before lockdown and it decreases by 60% to $4.04 \times 10^{24}$ molec./s during the lockdown period (Figure 6). This result fits well with the recent studies (Baldasano, 2020; Barré et al., 2021; Guevara et al., 2021). European Environment Agency (EEA) also reported a 56% - 72% reduction in NO$_2$ concentrations in Madrid based on in situ monitoring data (EEA report, 2020). Even compared to the emission at weekends, the lockdown emission was reduced by 43%. The regions with high NO$_2$ emissions are constrained only in the east of Madrid, where there are residential areas. Note that the lockdown spatial pattern reproduces that observed at weekends during the whole period (Figure 4b), corroborating that NO$_2$ emissions are highly related to transportation. Civil aviation was also restricted during the lockdown and thus less NO$_2$ emission strength is observed close to the airport. The reduction is larger than that in Riyadh as Madrid was under a very strict lockdown policy.
4. Conclusions

The paper proposes a combination of wind-assigned anomalies and Machine Learning (ML) methods to estimate the average tropospheric NO$_2$ emission strengths and its spatial pattern derived from TROPOMI observations from April 2018 to March 2021. The ADAM algorithm, as one of the Gradient Descent algorithms in ML is chosen because of its high efficiency and little cost requirement.

Riyadh is first used as a test studying site due to its high population density, remote from other sources, and favorite weather conditions, which allow for the high availability of space-based observations. Consistent wind-assigned plumes are found based on the TROPOMI measurements, so as the ML-trained plumes. A very good correlation between them is obtained with an R$^2$ value of 1.0 and a slope of 0.99. The spatial pattern of the estimated emission strengths agrees very well with the results from Beirle et al. (2019) as well. Several NO$_2$ emission hotspots, associated with the cement plant and power plants, are discernible. The total emission rate over the whole area is about $1.04 \times 10^{26}$ molec./s, which is higher than the previous study ($8.5 \times 10^{25}$ molec./s, Beirle et al., 2019). This difference might be due to the different study period and methods. These results suggest that our combined method works properly and is reliable.

We extended this method to the (mega)city of Madrid, Spain. The averaged NO$_2$ emission estimates are $1.80 \times 10^{25}$ molec./s in total and the dominant emitting area is around the city center, especially in the north-to-northeast and south-to-southeast regions. The region with the international Madrid-Barajas Airport in the northeast is also distinguished with high emission rates, as aircraft exhaust emissions are highly enriched in NO$_2$ during taxiing and taking off (Herndon et al., 2004). The orographic feature also causes NO$_2$ accumulation in the NE-SW regions, along the Guadarrama mountains range.

NO$_2$ emission is highly related to transportation and thus NO$_2$ emission changes between weekdays and weekends are investigated as well. Different weekly cycles of NO$_2$ are observed in Riyadh and Madrid. The lowest NO$_2$ column abundances are observed on Fridays, followed by the ones on Saturdays in Riyadh. The NO$_2$ emissions were reduced by 24% at weekends, and high reductions are found near the city center and the areas along Highway 65. Regions in the west and southwest of Madrid are not main NO$_2$ emitting areas at weekends but are on weekdays, indicating that many working places are located in the southwest. The estimates are $9.85 \times 10^{24}$ molec./s on weekdays and $7.24 \times 10^{24}$ molec./s at weekends in the urban area (70 km $\times$ 70 km$^2$). This 26% reduction in NO$_2$ emission is mainly due to commuting from home to the city center and working places.

Many studies have demonstrated that the lockdown policy response to the COVID-19 pandemic reduces NO$_2$ emissions (Barré et al., 2021; Sun et al., 2021; Liu et al., 2021; Virghileanu et al., 2020; Keller et al., 2021; Bauwens et al., 2020; Fan et al., 2020; Huang and Sun, 2020). Countries like Spain imposed a very stringent lockdown since March 2020. An average reduction of 60% in NO$_2$ emissions is observed during lockdown (March – June 2020) compared to the period of March – June 2019. The regions with dominant NO$_2$ emissions during lockdown are limited in the east of Madrid where there are residential areas.
Reduced NO$_2$ emissions (27\%) were observed in Riyadh, especially near the city center. This reduction is much smaller than that in Madrid, as the latter was under a very strict lockdown regulation.

Our easy-to-apply method has successfully probed its consistency and reliability in two contrasting examples (Riyadh and Madrid). But, it is can be applied to other key gases such as carbon dioxide or methane, and in other regions. Meanwhile, the powerful ML framework allows related analysis to be addressed, such as the estimation of the lifetime of NO$_2$.

---

**Data availability.** The TROPOMI data set is publicly available from https://scihub.copernicus.eu/ (last access: 18 January 2022; ESA, 2020). The access and use of any Copernicus Sentinel data available through the Copernicus Open Access Hub are governed by the legal notice on the use of Copernicus Sentinel Data and Service Information, which is given here: https://sentinels.copernicus.eu/documents/247904/690755/Sentinel_Data_Legal_Notice (last access: 18 January 2022; European Commission, 2020).

**Author contributions.** Qiansi Tu, Frank Hase, Zihan Chen, and Matthias Schneider developed the research question. Qiansi Tu wrote the manuscript and performed the data analysis with input from Frank Hase, Zihan Chen, Matthias Schneider, Omaira García, and Farahnaz Khosrawi. All authors discussed the results and contributed to the final manuscript.

**Competing interests.** The authors declare that they have no conflict of interest.

**Acknowledgements.** We acknowledge Quan Ngoc Pham from Interactive Systems Lab, Karlsruhe Institute of Technology for the useful technical support in the ML model. We would like to thank Emissions of atmospheric Compounds and Compilation of Ancillary Data (ECCAD) for providing CAMS-REG-AP inventory data. Thanks should also go to the TROPOMI team for making NO$_2$ data publicly available. We also acknowledge the project of Joint R&D and Talents Program funded by the Qingdao Sino-German Institute of Intelligent Technologies (kh0100020213319) and the project of Transnational Interoperability Rules and Solution Patterns in Collaborative Production Networks based on IDS and GAIA-X funded by Ministry of Science and Technology, PRC (SQ2021YFE010470).
Appendix

Figure A-1: Windrose of ERA5 model wind at 975 hPa (~330 m) in the daytime during TROPOMI overpasses, and wind speed with respect to pressure levels in Riyadh ((a)-(b)), and in Madrid ((c)-(d)).

Figure A-2: Examples of wind-assigned plume for the point source at 24.65°N, 46.85°E and at 25.05°N, 46.85°E in Riyadh.
Figure A- 3: TROPOMI tropospheric NO\textsubscript{2} column for narrow wind regimes covering (a) NE (0º-90º), (b) SW (18º-270º), (c) NW (270º-360º), and (d) SE (90º-180º), respectively.

Figure A- 4: TROPOMI tropospheric NO\textsubscript{2} column during week in Riyadh. Number in the figures’ title represents the average column abundances over the area.
Figure A-5: Map for Riyadh. Area in the white rectangle represents the study area. © Google Maps

Figure A-6: same as Figure A-4, but for the Madrid area.
Figure A-7: Map of the study area (up) © Google Maps and zoom version (bottom, https://land.copernicus.eu/local/urban-atlas/urban-atlas-2018, last access: 25 April 2022) for Madrid.
NO$_2$ plume on weekdays, Madrid

(a) TROPOMI tropospheric NO$_2$

(b) Machine learning NO$_2$

NO$_2$ plume at weekends, Madrid

(c) TROPOMI tropospheric NO$_2$

(d) Machine learning NO$_2$

Correlation of wind-assigned anomalies

(e) weekdays

(f) weekends
Figure A-8: Wind-assigned plumes derived from TROPOMI observations (a-b), ML method (c-d), and their correlation plots (e-f) on weekdays (left) and at weekends (right) in Madrid.

Figure A-9: Time series of TROPOMI tropospheric NO$_2$ columns in terms of daily, 10-day and monthly mean in (a) Riyadh and (b) Madrid. Areas marked with lavender and orange colors are the study periods in 2019 and 2020, respectively.
References


