Spatio-temporal modeling of air pollutant concentrations in Germany using machine learning

Vigneshkumar Balamurugan¹, Jia Chen¹, Adrian Wenzel¹, and Frank N. Keutsch²,³

¹Environmental Sensing and Modeling, Technical University of Munich (TUM), Munich, Germany.
²School of Engineering and Applied Science, Harvard University, Cambridge, MA, USA.
³Department of Chemistry and Chemical Biology, Harvard University, Cambridge, MA, USA.

Correspondence: Vigneshkumar Balamurugan (vigneshkumar.balamurugan@tum.de), Jia Chen (jia.chen@tum.de)

Abstract. Machine learning (ML) models are becoming a meaningful tool for modeling air pollutant concentrations. ML models are capable of learning and modeling complex non-linear interactions between variables, and they require less computational effort than chemical transport models (CTMs). In this study, we used gradient boosted tree (GBT) and multi-layer perceptron (MLP; neural network) algorithms to model near-surface nitrogen dioxide (NO₂) and ozone (O₃) concentrations over Germany at 0.1 degree spatial resolution and daily intervals.

We trained the ML models using TROPOMI satellite column measurements combined with information on emission sources, air pollutant precursors and meteorology as feature variables. We found that the trained GBT model for NO₂ and O₃ explained a major portion of the observed concentrations (R² = 0.68-0.88, RMSE = 4.77-8.67 µg m⁻³ and R² = 0.74-0.92, RMSE = 8.53-13.2 µg m⁻³, respectively). The trained MLP model performed worse than the trained GBT model for both NO₂ and O₃ (R² = 0.46-0.82 and R² = 0.42-0.9, respectively).

Our NO₂ GBT model outperforms the CAMS model, a data-assimilated CTM, but slightly under-performs for O₃. However, our NO₂ and O₃ ML models require less computational effort than CTM. Therefore, we can analyze people’s exposure to near-surface NO₂ and O₃ with significantly less effort. During the study period (2018-04-30 and 2021-07-01), it was found that around 36% of people lived in locations where the WHO NO₂ limit was exceeded for more than 25% of the days, while 90% of the population resided in areas where the WHO O₃ limit was surpassed for over 25% of days. Although metropolitan areas had high NO₂ concentrations, rural areas, particularly in southern Germany, had high O₃ concentrations.

Furthermore, our ML models can be used to evaluate the effectiveness of mitigation policies. Near-surface NO₂ and O₃ concentrations changes during the 2020 COVID-19 lockdown period over Germany were indeed reproduced by the GBT model, with meteorology-accounted for near-surface NO₂ significantly decreased (by 23±5.3%) and meteorology-accounted for near-surface O₃ slightly increased (by 1±4.6%) over ten major German metropolitan areas, compared to 2019. Finally, our O₃ GBT model is highly transferable to other countries, at least to neighboring countries and locations where no measurements are available (R² = 0.87-0.94), whereas our NO₂ GBT model is moderately transferable (R² = 0.32-0.64).
1 Introduction

Air pollution is a major threat to human health and impacts ecosystems (Bell et al., 2011; Lelieveld et al., 2015; Zhang et al., 2019; Xie et al., 2019). Based on the source, air pollutants are classified as primary (directly emitted from anthropogenic/natural sources) or secondary (formed through complex atmospheric chemical reactions). Near-surface nitrogen oxide (\(\text{NO}_X = \text{NO} + \text{NO}_2\)) is a primary air pollutant emitted largely by fossil-fuel-consuming sectors such as vehicles, industries, power plants, etc., but there are also natural sources such as lightning, soil emissions, and biomass burning. Near-surface ozone (\(\text{O}_3\)) is a secondary air pollutant produced solely by the photolysis of \(\text{NO}_2\) (nitrogen dioxide) in the presence of sunlight (Crutzen, 1988; Council et al., 1992).

Tropospheric \(\text{NO}_X\) and \(\text{O}_3\) are chemically strongly coupled via complex atmospheric chemical reactions (Jacob, 1999). The majority of \(\text{NO}_X\), from primary sources such as fossil-fuel combustion, is emitted in the form of nitric oxide (NO), which rapidly converts to \(\text{NO}_2\) by reacting with \(\text{O}_3\). In turn, \(\text{O}_3\) and NO are generated again by photolysis of \(\text{NO}_2\), forming a null cycle. Therefore, the amount of sunlight present and the total concentration of \(\text{NO}_X\) determine ozone production via this \(\text{NO}_X\) null cycle. However, the oxidation of volatile organic compounds (VOCs) via the hydroxyl (OH) radical results in the formation of hydro-peroxy radicals (\(\text{HO}_2\)) and organic-peroxy radicals (\(\text{RO}_2\)), which can alter the \(\text{NO}/\text{NO}_2\) ratio. The presence of hydroxyl radical initiates the VOC oxidation process, followed by the formation of hydro- and organic peroxy radicals, which convert the NO to \(\text{NO}_2\), which can form additional \(\text{O}_3\), as well as converting \(\text{HO}_2\) back to OH thus forming a catalytic cycle (\(\text{HO}_X\) catalytic cycle). However, ozone production is non-linear in relation to its precursors (\(\text{NO}_X\) and VOC) due to termination reactions that occur within the catalytic cycle (Lin et al., 1988; Nussbaumer and Cohen, 2020; Pusede and Cohen, 2012; Pusede et al., 2014). To that end, the response of ozone production is categorized into three regimes: \(\text{NO}_X\)-saturated (high \(\text{NO}_X\) with low VOC), \(\text{NO}_X\)-limited (low \(\text{NO}_X\) with high VOC), and transitional (Sillman et al., 1990; Sillman, 1999). In the \(\text{NO}_X\)-saturated regime (typically urban areas), ozone production is inversely proportional to \(\text{NO}_X\) concentration, whereas ozone production is directly proportional to VOC concentration. However, in \(\text{NO}_X\)-limited regimes (typically rural areas), ozone production is directly proportional to \(\text{NO}_X\) concentration, whereas VOCs have little effect on ozone production. This complex ozone production vs. precursor emission response is also evident in real-time observations, such as urban weekend ozone levels being higher than weekday levels (Sicard et al., 2020) and high ozone levels during public holidays and national shutdowns (e.g., the COVID-19 lockdown) due to low \(\text{NO}_X\) emissions (Balamurugan et al., 2021, 2022).

Chemical transport models (CTMs) are commonly used to study air pollution and its drivers (Hu et al., 2016; Lou et al., 2015), but these models are dependent on emissions as represented in emission inventories (Pisoni et al., 2018). Emission inventories are typically developed using the bottom-up method, based on data such as economic activity, fuel consumption, traffic density, etc (McDuffie et al., 2020; Osses et al., 2022). However, bottom-up emission inventories can be highly uncertain due to inaccuracies in the data used in the bottom-up method, especially from unaccounted sources (Chen et al., 2020; Crippa et al., 2019; Trombetti et al., 2018). Because of the significant computational effort and storage space requirements, CTMs often perform at coarse spatial resolution, making it unable to solve fine transport and chemical mechanisms, particularly over complex topography (Singh et al., 2021). Machine learning (ML) models have been shown to be an effective complement to
these computationally expensive CTMs (Vlasenko et al., 2021). The performance of machine learning models for modeling air pollutants is promising (Balamurugan et al., 2022a; Cheng et al., 2022; Lee et al., 2020; Li et al., 2022; Liang et al., 2020; Liu et al., 2022; Zaini et al., 2022; Zhao et al., 2023). Meteorological variables such as solar radiation and temperature have been shown to be important parameters in near-surface ozone modeling using machine learning (Diao et al., 2021; Hu et al., 2021). Meteorological conditions influence the concentration of O$_3$ both directly and indirectly. Solar UV radiation is responsible for the photolysis of O$_3$ precursors (NO$_2$ and VOCs). Temperature directly influences the photochemical reaction rate. Furthermore, meteorology influences biogenic and fuel-leak-related VOC emissions (exponentially proportional to temperature), which account for a significant portion of total VOC emissions (Guenther et al., 1993). In addition to meteorology, when emission source information is included, ML models predict near-surface NO$_2$ very well (Ghahremanloo et al., 2021; De Hoogh et al., 2019).

In-situ air quality measurements are sparse and concentrated primarily in urban areas. Recent advancements in satellite remote sensing allow us to analyze urban and non-urban air quality with adequate spatial and temporal coverage; however, they typically only measure the total or tropospheric column of specific air quality species, making it difficult to interpret people’s exposure to near-surface air pollutants concentration. Therefore, in this study, we trained two ML models for near-surface NO$_2$ and O$_3$ concentrations over Germany using available information on proxies for near-surface air pollutants (satellite column measurements) and emission sources, precursors of air pollutants, as well as meteorology. Many recent studies, similar to ours, have attempted to model near-surface NO$_2$ and O$_3$ concentrations at national/regional spans (De Hoogh et al., 2019; Kang et al., 2021; Li et al., 2020; Zhu et al., 2022); there are, however, very few attempts over Germany. To the best of the authors’ knowledge, only one study (Chan et al., 2021) used TROPOMI satellite NO$_2$ tropospheric column measurements and other auxiliary information (e.g., meteorology) to model near-surface NO$_2$ concentrations over Germany using a MLP model. Furthermore, previous studies have focused on a single pollutant (e.g., NO$_2$), whereas in this study, we model and analyze the spatio-temporal variations in both NO$_2$ and O$_3$, which are chemically strongly coupled. In terms of anthropogenic emissions, we also evaluate the ML model performance of NO$_2$ and O$_3$ during the 2020 COVID-19 lockdown period, which serves as a natural experiment period with significantly lower primary anthropogenic emissions (Gensheimer et al., 2021).

2 Study region, Data sets, Model, and Method

All data sets used in this study, as well as their spatial and temporal resolutions, are summarized in Table 1.

2.1 Study region and near-surface NO$_2$ and O$_3$ measurements

We focused on the spatial domain of 5-15°E and 47-55.5°N, particularly over Germany. Near-surface NO$_2$ and O$_3$ data from measurement stations across Germany were used in this study. However, not all measuring stations collect data on both pollutants; there are less stations measuring O$_3$ than those measuring NO$_2$. There were also temporal gaps in the measurement data. Therefore, we only considered stations that had more than 80% data coverage during the study period. In the end, we...
considered 321 stations for modeling NO$_2$ and 256 stations for modeling O$_3$. The selected measurement stations are located throughout the entire country and are situated in high-traffic, industrial, and background locations (Fig. 1 & Table A1).

2.2 Predictor variables of ML model

Predictor variables or input features for the ML models include satellite column measurements of air pollutants, meteorology and auxiliary data containing information on the area of interest.

2.2.1 Satellite column measurements

Tropospheric column NO$_2$, total column O$_3$, and tropospheric column HCHO data are used, which are level-2 retrieval products from TROPOMI (TROPOspheric Monitoring Instrument) aboard the Sentinel-5P satellite. Sentinel-5P overpasses the study area between 13:00 and 14:00 local standard time. The spatial resolution of TROPOMI data is 7*3.5 km (increased to 5.5*3.5 km after August 6, 2019). We applied the data quality filtering described in the product manual to each data product (S5P (2022b) for NO$_2$, S5P (2022c) for O$_3$, and S5P (2022a) for HCHO). Tropospheric column NO$_2$ is used in the NO$_2$ ML model because it can be considered as proxy for near-surface NO$_2$. Since NO$_2$ is the precursor for O$_3$, we also included the tropospheric column NO$_2$ in the O$_3$ ML model. Because formaldehyde (HCHO) is an intermediate gas-product of VOC oxidation, it can be used as a proxy for VOC-oxidation (Jin et al., 2017). Therefore, we included tropospheric column HCHO in the O$_3$ model. We also considered the “TROPOMI FNR” (ratio of “TROPOMI HCHO” and “TROPOMI NO$_2$”) in the O$_3$ ML model, which in previous studies has been shown to be a useful indicator of ozone production regime (Jin et al., 2020; Wang et al., 2021). We included total column O$_3$ in the O$_3$ ML model by considering total column O$_3$ as a proxy for near-surface O$_3$. 

Figure 1. Locations of near-surface NO$_2$ (a) and O$_3$ (b) measurement stations considered in this study. The color bar depicts the mean of near-surface NO$_2$ and O$_3$ for each measurement station during the study period.
<table>
<thead>
<tr>
<th>Data source</th>
<th>Data</th>
<th>Temporal resolution</th>
<th>Spatial resolution</th>
</tr>
</thead>
<tbody>
<tr>
<td>Governmental in situ measurements</td>
<td>Near-surface NO\textsubscript{2} and O\textsubscript{3}</td>
<td>1 hr</td>
<td>-</td>
</tr>
<tr>
<td>TROPOMI satellite measurements</td>
<td>Tropospheric column NO\textsubscript{2}, total column O\textsubscript{3} and tropospheric column HCHO</td>
<td>Daily</td>
<td>7 km<em>3.5 km (5.5 km</em>3.5 km, after 6 August 2019)</td>
</tr>
<tr>
<td>ERA5 (ECMWF reanalysis)</td>
<td>Temperature, relative humidity, wind speed, wind direction, downwind UV solar radiation at surface, boundary layer height, surface pressure and temperature of air at 2 m above the surface</td>
<td>1 hr</td>
<td>0.25*0.25-degree</td>
</tr>
<tr>
<td>U.S. Geological Survey</td>
<td>Surface elevation</td>
<td>-</td>
<td>1*1-km</td>
</tr>
<tr>
<td>GRIP global roads database</td>
<td>Road density</td>
<td>-</td>
<td>8*8-km</td>
</tr>
<tr>
<td>CAMS European air quality reanalyses</td>
<td>Near-surface NO\textsubscript{2} and O\textsubscript{3}</td>
<td>1 hr</td>
<td>0.1*0.1-degree</td>
</tr>
<tr>
<td>GEOS-Chem chemical transport model</td>
<td>Near-surface NO\textsubscript{2} and O\textsubscript{3}</td>
<td>1 hr</td>
<td>0.5*0.625-degree</td>
</tr>
</tbody>
</table>

### 2.2.2 Vegetation index

Normalized difference vegetation index (NDVI) and enhanced vegetation index (EVI) data were obtained from MODIS (Moderate Resolution Imaging Spectroradiometer) measurements aboard the Terra and Aqua satellites. We used the “MOD13A2 (16-day 1-km) VI” data set, which contains NDVI and EVI data at 1 km spatial resolution and 16 day temporal resolution. To generate daily intervals, the NDVI and EVI data were linearly interpolated. We considered these vegetation indexes in the O\textsubscript{3} ML model because vegetation contributes a considerable amount of VOCs. We also considered these vegetation indexes in the NO\textsubscript{2} ML model as a supplementary information to check whether changes in vegetation cover has any implications on NO\textsubscript{2} concentration changes.

### 2.2.3 Meteorology

Meteorology has both direct and indirect effects (e.g., dispersion, photochemical reactions) on pollutant concentrations. Meteorological variables such as temperature (T), relative humidity (RH), wind speed (WS), and wind direction (WD) were obtained from the ERA-5 reanalysis product. These variables were derived from the lowest model level (1000 hPa) of the “ERA-5 hourly data on pressure levels” data set. Downward UV solar radiation at the surface (DUV), boundary layer height (BLH), surface pressure (SP) and temperature of the air at 2 m above the surface (T2m) were derived from the “ERA-5 hourly data on single levels” data set. These meteorological data have a spatial resolution of 0.25 degree and a temporal resolution of one hour. In both the NO\textsubscript{2} and O\textsubscript{3} ML models, we took all meteorology variables into account.

### 2.2.4 Proxy for NO\textsubscript{X} emission source

Because vehicle (transport sector) emissions are a significant source of NO\textsubscript{X} emissions, considering a proxy for vehicle emissions is crucial. Therefore, we used road density as a proxy for the source of NO\textsubscript{X} emissions. We are aware that traffic volume...
or density would be the ideal proxy, but data on traffic volume or density on a national/regional span is not available. The road density (RD) data was obtained from the GRIP global roads database, with a spatial resolution of 8 km.

2.2.5 Additional features

Additional supplementary data such as surface elevation (E) was obtained from the U.S. Geological Survey (USGS), with a spatial resolution of 1 km. Surface elevation was taken into account because it influences the tropospheric/total column value of measurements. We also considered “DOW” (day of the week), and “season” (season of the year) information in both the NO\textsubscript{2} and O\textsubscript{3} models since both NO\textsubscript{2} and O\textsubscript{3} have distinct weekly and seasonal cycles. Because NO\textsubscript{2} is an important precursor to O\textsubscript{3}, in addition to “TROPOMI NO\textsubscript{2}”, we also included “Near-surface NO\textsubscript{2}” modeled from NO\textsubscript{2} ML model as a feature variable in the O\textsubscript{3} ML model.

2.3 Study period and data pre-processing

The study period was chosen to be between 2018-04-30 and 2021-07-01, which corresponds to the availability of TROPOMI data retrievals with the same processing version. Despite the fact that satellites pass over the study area between 13:00 and 14:00 local standard time, we found that the satellite data represents the daily mean of air pollutants well. Therefore, we considered the daily 24-hr mean for near-surface NO\textsubscript{2} and the daily maximum 8-hour mean (i.e. the mean of the 8 highest hourly values during a day) for near-surface O\textsubscript{3} as our variables of interest (dependent variables to model), as these are commonly used metrics in air quality research (Hoffmann et al., 2021).

Because each data set has a different spatial and temporal resolution, we re-sampled all of the data to the same spatial (0.1°×0.1 degree) and temporal (daily) resolution. The 0.1 degree (≈ 10 km) resolution was chosen because it corresponds to the resolution of the main features such as road density (spatial resolution of 8 km), TROPOMI satellite measurements (spatial
Table 2. Evaluation metrics of our GBT model in different testing strategies.

<table>
<thead>
<tr>
<th></th>
<th>Random (1-fold)</th>
<th>Random (5-fold)</th>
<th>Time-leave-out (5-fold)</th>
<th>Location-leave-out (5-fold)</th>
</tr>
</thead>
<tbody>
<tr>
<td>NO\textsubscript{2}</td>
<td>R\textsuperscript{2}</td>
<td>0.88</td>
<td>0.89±0.002</td>
<td>0.74±0.07</td>
</tr>
<tr>
<td>GBT model</td>
<td>RMSE (µg m\textsuperscript{-3})</td>
<td>4.77</td>
<td>4.65±0.034</td>
<td>6.77±0.7</td>
</tr>
<tr>
<td>O\textsubscript{3}</td>
<td>R\textsuperscript{2}</td>
<td>0.92</td>
<td>0.92±0.001</td>
<td>0.74±0.09</td>
</tr>
<tr>
<td>GBT model</td>
<td>RMSE (µg m\textsuperscript{-3})</td>
<td>8.53</td>
<td>9.36±0.068</td>
<td>13.2±1.1</td>
</tr>
</tbody>
</table>

resolution of 7*3.5 km), and concurrent high-resolution (0.1 degree) reanalysis air quality datasets from CAMS (Copernicus Atmosphere Monitoring Service). We computed the daily 24-hr mean for near-surface NO\textsubscript{2} and the daily maximum 8-hr mean for near-surface O\textsubscript{3} for each in-situ measurement station and then calculated the mean of all stations that fell within 0.1 degree grid. The mean of surface elevation, NDVI, EVI, TROPOMI (NO\textsubscript{2}, HCHO, O\textsubscript{3}), and road density for each day were then calculated for the corresponding 0.1 degree grids. The surface elevation and road density were assumed to be constant during the study period. The ERA-5 meteorology product was resampled to 0.1 degree resolution using the nearest-neighbor method and the 24-hr mean was computed.

2.4 Machine learning model and evaluation strategies

We primarily used the gradient boosted tree (GBT) machine learning algorithm, XGBoost (Chen and Guestrin, 2016), to model near-surface NO\textsubscript{2} and O\textsubscript{3} concentrations. The GBT algorithm is a gradient-boosted decision tree-based algorithm that is expected to outperform deep neural network-based algorithms for structured data (Lundberg et al., 2020). Furthermore, tree-based models are more interpretable and require less time to train than deep neural network algorithms. However, for comparison, we also used the multi-layer perceptron (MLP; neural network) algorithm (Gardner and Dorling, 1998). The GBT and MLP algorithms were implemented using "scikit-learn", a Python module (https://scikit-learn.org/stable/). When training the MLP model, we normalized the discrete feature variables between 0 and 1. The corresponding predictor variables and data flow for the NO\textsubscript{2} and O\textsubscript{3} ML model is shown in Fig. 2.

To evaluate the ML model, we used the R\textsuperscript{2} (coefficient of determination) and RMSE (root-mean-square error) metrics. We split the available data into training (70% of the data) and testing (the remaining 30%). The training data set was used to iteratively vary the hyper-parameters (combinations) and select the best set of hyper-parameters using a 5-fold CV (cross-validation). The hyper parameters used in this study are shown in Table A2 and Table A3. We also evaluated the ML model using three different 5-fold CV testing strategies (random 5-fold CV, time-leave-out 5-fold CV, and location-leave-out 5-fold CV) with 100% of the data (Meyer et al., 2018). In the random 5-fold CV testing strategy, the data was randomly split into five parts, four of which were used for training and one for testing. This procedure was repeated until all five parts had been used as test. The mean (and standard deviation) of R\textsuperscript{2} and RMSE from the 5-fold CV were then computed. In the time-leave-out 5-fold CV testing strategy, the 5-fold CV procedure was the same, but the data was split based on time period (by date;
Figure 3. Comparison between ground-truth and GBT-simulated near-surface NO$_2$ (a) and O$_3$ (b). Feature importance (top 10) calculated based on SHAP (SHapley Additive exPlanations) values for NO$_2$ (c) and O$_3$ (d) GBT model. RD: Road Density, BLH: Boundary Layer Height, E: Surface Elevation, T: Temperature, DOW: Day of the week, RH: Relative Humidity, T2m: Temperature at 2 meter height, DUV: Downwind UV radiation, WS: Wind speed, WD: Wind Direction.
from the start of study period to the end of study period). Similarly, in the location-leave-out 5-fold CV testing strategy, the data was split based on location (by latitude). Figure A1 shows the first one-fold step in a 5-fold CV for time-leave-out and location-leave-out testing strategies. To interpret the importance of feature variables in the fitted model, we use SHAP (SHapley Additive exPlanations) values. The SHAP method (https://christophm.github.io/interpretable-ml-book/shap.html) is the most commonly used method for interpreting ML model output, which calculates the contribution of each feature variable to the final prediction. Thus, higher SHAP values indicate greater feature importance.

2.5 CAMS model data

We obtained near-surface NO$_2$ and O$_3$ reanalysis data from CAMS in order to compare the performance of our ML model to that of the chemical transport model. This data set is based on a data-assimilation technique that combines real-time measurements with an ensemble of eleven air quality models to provide air quality data with high spatial resolution (0.1 degree) and 1 hr temporal resolution over Europe; however, it is only available for three years in the rolling archive. We used data from 2019-07-17 to 2020-01-31. We did not use data after 2020-01-31 due to COVID-19 lockdown restrictions, which limited many anthropogenic emission activities, and CAMS had not adjusted the emission inventory for changes in emissions. Furthermore, because NO$_2$ has a shorter lifetime, the effect of assimilated observations is minimal, and the CAMS reanalysis NO$_2$ product mostly reflects emissions prescribed in the inventory (Inness et al., 2015).

2.6 GEOS-Chem model data

In this study, GEOS-Chem (GC) chemical transport model simulations were used to disentangle the meteorology contribution when estimating the influence of COVID-19 lockdown restrictions on air pollutant concentration changes. The GC simulations over the study area were obtained with a spatial resolution of 0.5 × 0.625 degree and 1-hr temporal resolution for the 2020 strict COVID-19 lockdown period (March 21 to May 31) and the same period in 2019. Identical anthropogenic emissions from the 2014 CEDS inventory were used for both 2020 and 2019, but with the corresponding meteorology, natural, and fire emissions in the respective years. Therefore, the difference in GC-simulated species (X) concentrations between 2020 and 2019 results from changes in meteorology, natural, and fire emissions between 2020 and 2019 (GC X$_{2020-2019}$); here, X refers to either NO$_2$ or O$_3$. Then, we subtracted the GC X$_{2020-2019}$ from the observed near-surface X$_{2020-2019}$ to estimate the changes in concentrations of species X due to changes in anthropogenic emissions in the 2020 lockdown period (refer to studies Balamurugan et al. (2021); Qu et al. (2021) for the detailed description of the method).

3 Results

3.1 ML model evaluation and feature importance

The trained GBT model with 70% of the data for NO$_2$ reproduced the observed NO$_2$ concentration well in the test case, with an $R^2$ of 0.88 and RMSE of 4.77 µg m$^{-3}$ (Fig. 3(a) and Table 2). The random 5-fold CV results were in the same
Figure 4. Top: Comparison between ground-truth near-surface NO$_2$ and CAMS reanalysis near-surface NO$_2$ (a) and O$_3$ (b) for the period between 17-07-2019 and 31-01-2020. Bottom: Comparison between ground-truth near-surface NO$_2$ and GBT-simulated near-surface NO$_2$ (c) and O$_3$ (d) for the period between 17-07-2019 and 31-01-2020. The dotted line represents a 1:1 line, while the solid line represents a linear fit.

The other two testing strategies (time-leave-out 5-fold CV and location-leave-out 5-fold CV) showed slightly worse agreement (Table 2), indicating that different validation strategies should be performed to interpret the ML model capability. Otherwise, it may result in an overoptimistic view of ML models (Meyer et al., 2018). Furthermore, the worse agreement in the location-leave-out 5-fold CV testing strategy suggests that there is less confidence in modeling the near-surface NO$_2$ over new locations that the GBT model has not been trained on before. However, these results outperformed the MLP model trained by another study (Chan et al. (2021); R = 0.8 and RMSE = 6.32 µg m$^{-3}$ obtained for the testing strategy of random split of 90% of data used for training and 10% of data used for testing) for near-surface NO$_2$ over Germany. Feature importance, based on the SHAP values, indicates that road density is the most important feature in the fitted model for NO$_2$ (Fig. 3(c)), because traffic is the main source of near-surface NO$_X$ in urban areas. The next most important features were TROPOMI NO$_2$, boundary layer height, and elevation. Because the majority of NO$_X$ sources are present at the surface, tropospheric column NO$_2$ data plays an important role in explaining near-surface NO$_2$.
Near-surface NO\textsubscript{2} typically has a negative correlation with boundary layer height, as increasing BLH disperses more and vice versa (Balamurugan et al., 2021). Therefore, BLH is one of the most important features. It is unexpected that elevation was an important feature. The cause could be that the surface elevation varies greatly across Germany, influencing the total tropospheric column of NO\textsubscript{2} and thus serving as a link between the tropospheric column of NO\textsubscript{2} and near-surface NO\textsubscript{2}. A previous study (Chan et al., 2021) also found that elevation was an important feature in the fitted MLP model for near-surface NO\textsubscript{2} over Germany.

The GBT model trained with 70% of the data for O\textsubscript{3} also well represented the observed O\textsubscript{3} concentrations in the test case, with an $R^2$ of 0.92 and RMSE of 8.53 $\mu$g m\textsuperscript{-3} (Fig. 3(b)). Similar to the NO\textsubscript{2} GBT model findings, time-leave-out 5-fold CV and location-leave-out 5-fold CV testing strategies showed less agreement than the random 5-fold CV testing strategy (Table 2). In comparison to our NO\textsubscript{2} GBT model, our O\textsubscript{3} GBT model demonstrated greater confidence in modeling near-surface O\textsubscript{3} over locations the model was not trained on. According to SHAP values, the five most important features were DUV, T, RH, BLH, and season, with DUV having the greatest influence (Fig. 3(d)). Because ozone is formed in the atmosphere from the photolysis of NO\textsubscript{2}, DUV plays a significant role in the fitted model that explains near-surface O\textsubscript{3}. Temperature is the second most important feature, which is also not surprising as it drives biogenic VOC emissions (an important precursor to O\textsubscript{3}). Previous studies also show similar findings (Diao et al., 2021; Hu et al., 2021). GBT-modeled near-surface NO\textsubscript{2} was the sixth most important feature in the fitted model, according to the SHAP values, and it was also more important than TROPOMI NO\textsubscript{2}.

Figure A2 shows the results obtained from the MLP model. Both the NO\textsubscript{2} and O\textsubscript{3} MLP models performed worse than the NO\textsubscript{2} and O\textsubscript{3} GBT models, respectively (Table A4 vs. Table 2). In particular, MLP model findings showed low agreement in time-leave-out 5-fold CV and location-leave-out 5-fold CV testing strategies. This supports previous studies (Heaton, 2020; Lundberg et al., 2020) showing MLP model is unlikely to outperform tree-based models for tabular data. Because the GBT model outperforms the MLP model, we only considered the GBT model results in the following.

### 3.2 GBT model performance compared to CAMS

To evaluate how well our GBT model performs compared to CAMS, we compared the high-resolution near-surface NO\textsubscript{2} and O\textsubscript{3} reanalysis data from CAMS with observations, and GBT-simulated near-surface NO\textsubscript{2} and O\textsubscript{3} with observations, for the period between 2019-07-17 and 2020-01-31, i.e., CAMS comparison period, (Fig. 4). Please note this time period was not used for training the GBT model for this comparison. Our NO\textsubscript{2} GBT model reproduced the observed near-surface NO\textsubscript{2} concentrations well during this comparison period, with an $R^2$ of 0.82 and RMSE of 5.76 $\mu$g m\textsuperscript{-3}, while CAMS NO\textsubscript{2} reanalysis showed poor representation ($R^2 = 0.37$ and RMSE = 14.96 $\mu$g m\textsuperscript{-3}). However, CAMS O\textsubscript{3} reanalysis agreed slightly better with observed concentrations ($R^2 = 0.93$ and RMSE of 9.2 $\mu$g m\textsuperscript{-3}) compared to our O\textsubscript{3} GBT model ($R^2 = 0.85$ and RMSE = 13 $\mu$g m\textsuperscript{-3}). It should be noted that CAMS model reanalyses were based on data assimilation techniques. Therefore, the CAMS models are expected to outperform our GBT models. However, our NO\textsubscript{2} GBT model outperforms CAMS, possibly because the effect of data assimilation is minimal in the CAMS reanalysis product due to the short NO\textsubscript{2} lifetime.
3.3 Spatio-temporal changes in near-surface NO₂ and O₃ over the study domain

After the discussed model evaluation, we trained the GBT model using 100% of the data and modeled the near-surface NO₂ and O₃ concentrations over the study domain at 0.1 degree resolution and daily (24-hr mean for NO₂ and 8-hr maximum mean for O₃) intervals. The averaged GBT-modeled near-surface NO₂ concentrations over the study domain during the study period are shown in Fig. 5(a). The spatial variability of near-surface NO₂ correlates with Germany’s population density, and the main hotspots correspond to Germany’s major metropolitan areas (Figure A3). The study domain’s main hotspot is western Germany (North Rhine-Westphalia; a federal state of Germany), Germany’s industrial heartland. The number of days (%) that exceeded the 2021 WHO NO₂ limit (24-hr mean > 25 µg m⁻³) over major metropolitan areas in Germany was more than 50%, with western Germany exceeding the WHO NO₂ limit on more than 80% of the days during the study period (Fig. 7). Around 36% of people live in locations where more than 25% of days exceed the WHO NO₂ limit during the study period (Fig. 8). The GBT-simulated near-surface O₃ showed distinct spatial variability compared to NO₂, with high O₃ concentrations over southern Germany and low O₃ concentrations over northern Germany (Fig. 6). This could be due to the fact that O₃ is a secondary pollutant that is primarily driven by photochemical reactions influenced by meteorology; DUV and temperature values, which were the most influencing factors for photochemical reactions and accordingly the most important features fitted in the O₃ GBT model, were higher in southern Germany than northern Germany (Figure A4). During the study period, more
than 50% of days in southern Germany exceeded the 2021 WHO O₃ limit (maximum 8-hr mean > 100 µg m⁻³). Nearly 90% of people live in locations where more than 25% of days exceed the WHO O₃ limit (Fig. 8). Another interesting fact is that southern metropolitan areas and high NOₓ regions have less days that exceeded the WHO O₃ limit than southern rural regions (Fig. 7). It is a well-known fact that rural regions have higher ozone levels than urban regions (Malashock et al., 2022). It could be because NO is a significant O₃ scavenger in higher NOₓ (NO₂ is a proxy for NOₓ) regions or due to being in a NOₓ saturated regime. Furthermore, it is due to the fact that rural regions being the downwind locations of emission plume and are the primary source of biogenic VOC emissions (Zong et al., 2018).

The GBT-simulated near-surface NO₂ showed seasonal variations, as expected, with higher values in the winter season (Fig. 5). This is because of high-residential heating demand and favorable meteorology (e.g., a low boundary layer height) for pollutant accumulation and less NO₂ photolysis due to low solar radiation in the winter. The near-surface NO₂ hotspots were the same in all seasons, as seen in the overall study period average. In contrast, near-surface O₃ showed strong seasonal variations, with high values in the spring and summer due to high solar radiation (Fig. 6). It is worth noting that, as seen in the overall study period average, O₃ values in southern Germany were significantly higher in spring and summer than in northern Germany. Because near-surface O₃ is mainly driven by meteorology (DUV and temperature, which drive photochemical reactions and precursor emissions), the spatial and temporal variability is attributed to changes in meteorology. We also compared the spatial variability of GBT-simulated near-surface NO₂ and O₃ to the CAMS reanalysis product for the period
3.4 Influence of COVID-19 lockdown restrictions on near-surface NO$_2$ and O$_3$ changes

Due to the COVID-19 out-break, many nations, including Germany, announced a lockdown in the spring of 2020. During that time period, various anthropogenic emission activities were restricted, affecting particularly traffic-related emissions. To estimate the influence the lockdown restrictions on air pollutant concentration changes, we compared the GBT-simulated 2020 lockdown concentration with the same period in 2019. The 2020 lockdown period measurements were not used for GBT model training in this comparison. This can also be regarded as the critical performance evaluation of the GBT model.

When comparing different time periods, it is crucial to account for meteorological effects when estimating the impact of anthropogenic emission reductions (i.e., lockdown effects) on changes in air pollutant concentrations. Therefore, as described in the method section, we used GC simulations to exclude the meteorology contribution from GBT-simulated concentrations. After disentangling the meteorology contribution, it is noticeable that high near-surface NO$_2$ levels decreased primarily over
the previously observed hotspots (Fig. 9). The near-surface O$_3$ increased over western Germany while decreasing elsewhere, particularly over low NO$_X$ regions. We already observed that western Germany was a NO$_X$ hotspot, possibly a NO$_X$ saturated regime, so a reduction in NO$_X$ increases ozone. Also, we could see that changes in near-surface O$_3$ were either negligible or slightly increased over metropolitan areas. The meteorology-accounted for mean lockdown near-surface NO$_2$ decreased by about 23 ($\pm 5.3$)%, while meteorology-accounted for mean lockdown near-surface O$_3$ increased by 1 ($\pm 4.6$)%, over ten major metropolitan areas (Berlin, Bremen, Cologne, Dresden, Düsseldorf, Frankfurt, Hamburg, Hanover, Munich, and Stuttgart), compared to 2019. It increased by about 9% in the Cologne and Düsseldorf metropolitan areas (located in western Germany) and slightly increased or decreased (between -3 and +2%) in other metropolitan areas, compared to 2019. This finding is consistent with other studies that found a decrease in meteorology-accounted for lockdown near-surface NO$_2$ and the small increase in lockdown near-surface O$_3$ over German metropolitan areas compared to 2019 using in-situ measurements (Balamurugan et al., 2021, 2022b). We also evaluated our GBT model’s ability to represent different emission scenarios by comparing weekends and weekdays; typically, anthropogenic NO$_X$ emissions on weekends are lower than on weekdays due to reduced vehicle transportation. Our GBT model was also able to distinguish between the weekend and weekday emission scenarios; weekend near-surface NO$_2$ was lower than weekday near-surface NO$_2$, and, as expected, there were no or only slight changes in weekend near-surface O$_3$ compared to weekdays, with slight increases particularly over metropolitan areas (Figure A7).

3.5 Transferability of our GBT model

Our study domain also covered parts of other European countries. However, we trained our GBT model using data from German measurement stations only. Therefore, comparing our trained GBT model simulations with measurements in other countries demonstrates how well our GBT model models near-surface NO$_2$ and O$_3$ concentrations in neighboring parts of the world; similar to the location-leave-out testing strategy. We chose five major cities (Salzburg, Prague, Strasbourg, Liège,
and Groningen) in different European countries covered by our study domain and compared their measured NO$_2$ and O$_3$ concentrations with GBT modeled NO$_2$ and O$_3$ concentrations (Fig. 10 & Table A5).

Our trained NO$_2$ GBT model based on German measurement stations explained 32-64% ($R^2$ ranges between 0.32 and 0.64, and RMSE ranges between 9.76 and 13 $\mu$g m$^{-3}$) of near-surface NO$_2$ measured in five metropolitan areas located outside of Germany, while O$_3$ GBT model simulations agreed well with observations ($R^2$ ranges between 0.87 and 0.94, and RMSE ranges between 9.55 and 14.32 $\mu$g m$^{-3}$). Since near-surface O$_3$ is mainly driven by meteorology, the O$_3$ GBT model trained using German measurement stations explains a large portion of near-surface O$_3$ in other locations. The worse agreement between NO$_2$ GBT model predictions and NO$_2$ observations in other European countries suggests that information is lacking in the NO$_2$ GBT model for better representation of other locations, similar to location-leave-out 5-fold CV, which also showed low agreement for the NO$_2$ GBT model when modeling new locations (Table 2). Differences in vehicle fleet composition and emission standards across different countries/locations would have an impact on our NO$_2$ GBT model predictions when applied to other countries/locations. In future work, other features/proxies besides road density could be considered to represent traffic emission.

### 4 Conclusion

This study simulated near-surface NO$_2$ and O$_3$ concentrations using an ML model over Germany at 0.1 degree resolution and daily intervals. The ML model was used to link satellite column measurements (proxies for near-surface air pollutants), meteorology and proxies of emission source information to near-surface NO$_2$ and O$_3$ concentrations. The ML models are
extremely effective at learning the complex non-linear relationships between variables. Therefore, in this study, we explored the capabilities of ML models in the spatio-temporal prediction of air pollutants. In addition, we investigated three aspects of the ML model: 1. how well our ML model performs compared to the chemical transport model, 2. how well our ML model can be used to assess the effectiveness of mitigation initiatives; and 3. how well our ML model can be transferred to locations where measurements are unavailable.

Four different testing strategies were performed to evaluate the ML model’s spatio-temporal prediction: 1. Random split of data (70% for training and 30% for testing), 2. Random 5-fold CV, 3. Time-leave-out 5-fold CV, and 4. Location-leave-out 5-fold CV. The gradient boosted tree (GBT) model trained for NO\textsubscript{2} explained about 68-88% of observed NO\textsubscript{2} concentrations in Germany, with RMSE of 4.77-8.67 µg m\textsuperscript{-3}, whereas the GBT model trained for O\textsubscript{3} performed even better, with an R\textsuperscript{2} of 0.74-0.92 and RMSE of 8.53-13.2 µg m\textsuperscript{-3}. The evaluation metrics of the GBT model for different testing strategies differed significantly. The location-leave-out 5-fold CV testing strategy showed poor agreement for the NO\textsubscript{2} GBT model, whereas the time-leave-out 5-fold CV testing strategy showed poor agreement for the O\textsubscript{3} GBT model. This points out the importance of performing different testing strategies to interpret the true capability of the ML model. The road NO\textsubscript{X} emission source proxy (road density) and TROPOMI tropospheric column NO\textsubscript{2} were the most important features in the fitted NO\textsubscript{2} GBT model. However, for O\textsubscript{3}, the most important features were downward UV radiation at the surface and temperature. The multi-layer perceptron (MLP) model trained for both NO\textsubscript{2} and O\textsubscript{3} performed worse than the GBT model.

We also showed that our NO\textsubscript{2} GBT model outperforms the CAMS model, while slightly under-performing for near-surface O\textsubscript{3}. The CAMS model reanalysis data set uses real-time observations with an ensemble of eleven air-quality models through data assimilation techniques, which are expected to be more computationally expensive than our GBT model. Therefore, the
spatio-temporal variability of near-surface NO$_2$ and O$_3$ concentrations and human exposure at locations where no measurements are available can be studied with lower computational effort when using our GBT model. Near-surface NO$_2$ hotspots were found over German metropolitan areas, particularly western Germany. The near-surface NO$_2$ hotspots locations did not change with the seasons but had high values in the winter. However, near-surface O$_3$ showed high seasonal variability, with high values in the spring and summer and no definite hotspots. Overall, southern Germany experiences higher ozone levels than northern Germany due to higher downward UV radiation and temperatures in southern Germany compared to northern Germany. Even though metropolitan areas were the NO$_2$ hotspots, rural regions, particularly in southern Germany, had higher O$_3$ concentrations than metropolitan areas. It is because rural areas are dominated by meteorology-driven biogenic VOC emissions and are generally situated downwind of the emission plume. About 36% of people live in locations where WHO NO$_2$ limit exceeds more than 25% of days during the study period. Meanwhile, 90% of the people lives in areas where the WHO O$_3$ limit is exceeded for more than 25% of days.

Our study also demonstrated the GBT model’s capability to assess the efficacy of mitigation strategies. For example, our GBT model reproduced the observations that, during the 2020 COVID-19 lockdown period, meteorology-accounted for near-surface NO$_2$ was significantly reduced, while meteorology-accounted for near-surface O$_3$ was slightly increased or decreased over metropolitan and industrial areas over Germany, compared to 2019. These findings agreed with those of other studies that used in-situ measurements.

Our GBT ML model’s transferability is assessed by comparing simulations from our GBT model trained with measurements in Germany to measurements in other European countries. Our NO$_2$ GBT model showed moderate agreement with observations from other countries ($R^2$ ranges between 0.32 and 0.64, and RMSE ranges between 9.76 and 13 $\mu g$ m$^{-3}$), implying a lack of information in the GBT model when modeling near-surface NO$_2$ over other countries, which may have different vehicle fleet composition and emissions standards. However, our O$_3$ GBT model performed well ($R^2$ ranges between 0.87 and 0.94, and RMSE ranges between 9.55 and 14.32 $\mu g$ m$^{-3}$), indicating that our O$_3$ GBT model can be used to model the O$_3$ concentrations in other countries, at least in neighboring European countries.

---

**Code and data availability.** The various data sets and code used to conduct this study will be made available on GitHub following publication.
Appendix A

Table A1. Different type of stations (%) considered in this study (based on locations specified by the European Environment Agency).

<table>
<thead>
<tr>
<th></th>
<th>Traffic</th>
<th>Industrial</th>
<th>Background</th>
</tr>
</thead>
<tbody>
<tr>
<td>Near-surface NO$_2$</td>
<td>37.1%</td>
<td>5.3%</td>
<td>57.6%</td>
</tr>
<tr>
<td>Near-surface O$_3$</td>
<td>2.7%</td>
<td>5.8%</td>
<td>91.4%</td>
</tr>
</tbody>
</table>

Table A2. The hyperparameters of the GBT model for each pollutant used in the study.

<table>
<thead>
<tr>
<th>Hyper parameter</th>
<th>NO$_2$ model</th>
<th>O$_3$ model</th>
</tr>
</thead>
<tbody>
<tr>
<td>Max_depth</td>
<td>10</td>
<td>10</td>
</tr>
<tr>
<td>Learning_rate</td>
<td>0.3</td>
<td>0.3</td>
</tr>
<tr>
<td>reg_lambda</td>
<td>12</td>
<td>4</td>
</tr>
<tr>
<td>reg_alpha</td>
<td>18</td>
<td>26</td>
</tr>
<tr>
<td>gamma</td>
<td>20</td>
<td>8</td>
</tr>
<tr>
<td>min_child_weight</td>
<td>16</td>
<td>8</td>
</tr>
<tr>
<td>n_estimators</td>
<td>2500</td>
<td>2500</td>
</tr>
</tbody>
</table>

Table A3. The hyperparameters of the MLP model for each pollutant used in the study.

<table>
<thead>
<tr>
<th>Hyper parameter</th>
<th>NO$_2$ model</th>
<th>O$_3$ model</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hidden_layers (neurons in each layer)</td>
<td>3 (200,100,50)</td>
<td>4 (350,150,75,37)</td>
</tr>
<tr>
<td>activation</td>
<td>tanh</td>
<td>tanh</td>
</tr>
<tr>
<td>alpha</td>
<td>0.04</td>
<td>0.1</td>
</tr>
<tr>
<td>learning_rate</td>
<td>adaptive</td>
<td>adaptive</td>
</tr>
<tr>
<td>solver</td>
<td>sgd</td>
<td>lbfgs</td>
</tr>
<tr>
<td>Max_iter</td>
<td>2000</td>
<td>1500</td>
</tr>
</tbody>
</table>
Table A4. Evaluation metrics of our MLP model in different testing strategies.

<table>
<thead>
<tr>
<th></th>
<th>Random (70%/30%)</th>
<th>Random (5-fold)</th>
<th>Time-leave-out (5-fold)</th>
<th>Location-leave-out (5-fold)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>NO₂</strong></td>
<td><strong>R²</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MLP model</td>
<td>0.79</td>
<td>0.82±0.006</td>
<td>0.54±0.29</td>
<td>0.46±0.25</td>
</tr>
<tr>
<td><strong>RMSE (µg m⁻³)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MLP model</td>
<td>4.77</td>
<td>5.9±0.11</td>
<td>8.6±1.76</td>
<td>13.2±1.07</td>
</tr>
<tr>
<td><strong>O₃</strong></td>
<td><strong>R²</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MLP model</td>
<td>0.83</td>
<td>0.9±0.001</td>
<td>0.42±0.37</td>
<td>0.71±0.13</td>
</tr>
<tr>
<td><strong>RMSE (µg m⁻³)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MLP model</td>
<td>12.15</td>
<td>9.6±0.027</td>
<td>20.1±7.3</td>
<td>14.9±3.2</td>
</tr>
</tbody>
</table>

Table A5. Metropolitan areas in other European cities considered for the evaluation of GBT model. The evaluation metrics (comparison between GBT simulations and in-situ measurements) for NO₂ and O₃ shown in last two columns for each city.

<table>
<thead>
<tr>
<th>Metropolitan area (country)</th>
<th>Coordinates</th>
<th>R² and RMSE (µg m⁻³) for NO₂</th>
<th>R² and RMSE (µg m⁻³) for O₃</th>
</tr>
</thead>
<tbody>
<tr>
<td>Salzburg (Austria)</td>
<td>47.80° N, 13.05° E</td>
<td>0.32 and 12.52</td>
<td>0.87 and 12.43</td>
</tr>
<tr>
<td>Prague (Czech Republic)</td>
<td>50.07° N, 14.43° E</td>
<td>0.43 and 10.05</td>
<td>0.79 and 14.32</td>
</tr>
<tr>
<td>Strasbourg (France)</td>
<td>48.57° N, 7.75° E</td>
<td>0.47 and 13</td>
<td>0.94 and 9.55</td>
</tr>
<tr>
<td>Liège (Belgium)</td>
<td>50.63° N, 5.56° E</td>
<td>0.64 and 11.9</td>
<td>0.88 and 12.04</td>
</tr>
<tr>
<td>Groningen (Netherlands)</td>
<td>53.21° N, 6.56° E</td>
<td>0.34 and 9.76</td>
<td>0.87 and 11.33</td>
</tr>
</tbody>
</table>

Figure A1. A first one-fold step in 5-fold CV is illustrated for time-leave-out (a) and location-leave-out (b) testing strategies. In time-leave-out 5-fold CV, the data was divided into 5 parts based on time period (date-wise), with four parts used for training and one part tested. This process is repeated until each part (a total of 5) has been tested. Similarly, in location-leave-out 5-fold CV, the data was divided into 5 parts based on location (latitude), with four parts used for training and one part tested. This process is repeated until each part (a total of 5) has been tested.
Figure A2. Comparison between ground-truth and MLP-simulated near-surface NO$_2$ (a) and O$_3$ (b). The dotted line represents a 1:1 line, while the solid line represents a linear fit.

Figure A3. Population density for the year 2020 (a) and the locations of major German metropolitan areas (b).
Figure A4. Averaged “Downward UV radiation at the surface” (a) and “Temperature” (b) over the study domain during the study period.

Figure A5. Averaged GBT-simulated near-surface NO$_2$ concentrations (a) and CAMS reanalysis near-surface NO$_2$ concentrations (b) over the study domain for the period between 2019-07-17 and 2020-01-31.
Figure A6. Averaged GBT-simulated near-surface O$_3$ concentrations (a) and CAMS reanalysis near-surface O$_3$ concentrations (b) over the study domain for the period between 2019-07-17 and 2020-01-31.

Figure A7. The difference in GBT-simulated near-surface NO$_2$ (a) and O$_3$ (b) concentrations between weekend and weekday during the study period.
Author contributions. VB, JC and FNK conceived the study and designed the concept. VB obtained all of the data, performed the modelling work and analysed the results. VB and AW developed the methodology. JC and FNK acquired the funding and supervised the work. VB wrote the manuscript. JC, AW and FNK reviewed and edited the manuscript.

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

Acknowledgements. This research has been funded by the Institute for Advanced Study, Technical University of Munich (grant no. 291763). The authors thank the European Environment Agency, the Copernicus Services, the GES DISC data archive and the United States Geological Survey for providing free access to the various data sets used in this study.
References


Balamurugan, V., Chen, J., Qu, Z., Bi, X., and Keutsch, F. N.: Secondary PM 2.5 decreases significantly less than NO 2 emission reductions during COVID lockdown in Germany, Atmospheric Chemistry and Physics, 22, 7105–7129, 2022b.


