

# Estimating Monthly Global Ground-Level NO<sub>2</sub> Concentrations Using Geographically Weighted Panel Regression

## Abstract

Predicting long-term ground-level nitrogen dioxide (NO<sub>2</sub>) is important globally to support environmental and public health research and to provide information to governments and society for air pollution control policies. The ozone monitoring instrument (OMI), onboard Aura Satellite, detects monthly global tropospheric column amounts (TrCA) of NO<sub>2</sub> molecules. However, the relationship between the ground-level NO<sub>2</sub> concentration and TrCA of NO<sub>2</sub> molecules remains elusive because NO<sub>2</sub> molecules in the air are not evenly distributed vertically. We use geographically weighted panel regression (GWPR) to examine the relationship between satellite-derived data, measured ground-level NO<sub>2</sub> concentrations, and several controlling meteorological variables from January 2015 to October 2021. The GWPR can analyze unbalanced panel data and capture the spatial variability of the relationship. Based on the GWPR estimation, 82 monthly global ground-level NO<sub>2</sub> concentrations are predicted from January 2015 to October 2021. The GWPR is reliable, as indicated by the 10-fold cross-validation. The accuracy of the raster prediction of global ground-level NO<sub>2</sub> from January 2015 to October 2021 is 69.61%. The coefficient of correlation, root mean square error and mean absolute error between globally predicted and measured ground-level NO<sub>2</sub> are 0.838, 7.84  $\mu\text{g}/\text{m}^3$  and 4.07

$\mu g/m^3$ , respectively, while the mean of globally measured ground-level  $NO_2$  is 19.47  $\mu g/m^3$ . Overall, this research provides critical basic data to environmental and public health science and valuable information for governments and societies to make more reasonable policies.

## Keywords

Monthly Ground-Level  $NO_2$  Concentration; GWPR model; spatial nonstationarity; OMI

## Introduction

The air pollutant nitrogen dioxide ( $NO_2$ ) is adversely related to various health outcomes (Lelieveld et al. 2015; Newell et al. 2017; Ogen 2020; Orellano et al. 2020). The relationship between all-cause mortality and  $NO_2$  concentration is positive (Brunekreef and Holgate 2002). Specifically, a 10  $\mu g/m^3$  increase in  $NO_2$  concentration in the short term is associated with a 0.72% increase in the relative risks of all-cause mortality (Orellano et al. 2020). The European Union and World Health Organization (WHO) recommend that the annual average exposure to ambient  $NO_2$  should be lower than 40  $\mu g/m^3$ . Additionally, other air pollutants, such as particulate matter and ozone, are associated with approximately 3.3 million premature deaths per year globally, and  $NO_2$  is one of the critical ozone precursors (Beckerman et al. 2008; Lelieveld et al. 2015). Evidence shows a link between  $NO_2$  and various acute and chronic diseases, especially decreased lung function and lung

cancer (Chiusolo et al. 2011; Hamra et al. 2015; Rice et al. 2013). Furthermore, several studies, while controlling for several other crucial factors, including age, gender, and restriction policy, indicate that in areas with high NO<sub>2</sub> concentrations, people are more easily infected by the coronavirus and find it more difficult to recover (Li and Managi 2022; Ogen 2020; Wu et al. 2020; Yao et al. 2021). In addition, air pollution affects human well-being directly (Li and Managi 2021b). For example, an increase of 10  $\mu\text{g}/\text{m}^3$  in annual average NO<sub>2</sub> is related to an approximately 50% decrease in life satisfaction in London (Mackerron and Mourato 2009). Thus, the ground-level NO<sub>2</sub> concentration affects human well-being and public health the most (Lelieveld et al. 2015; Li and Managi 2021b, 2022). However, satellite measurements focus mainly on the tropospheric column amounts (TrCA) and the total column amounts (ToCA) of NO<sub>2</sub> (**Figure 1**) (OMI Team 2012). The association of the ground-level NO<sub>2</sub> concentration with TrCA and ToCA remains unclear because the NO<sub>2</sub> molecules are not vertically evenly distributed in the air. Several recent studies have investigated the regional ground-level NO<sub>2</sub> concentrations based on satellite observations. For example, Kim et al. estimated the hourly near-surface NO<sub>2</sub> in the alpine domain from June 2018 to May 2020 (Kim et al. 2021); Li et al. detected the ground-level NO<sub>2</sub> concentrations in the Wuhan urban agglomeration during 2019 (Li and Wu 2021; Li et al. 2020); and Qin et al. probed the ground-level NO<sub>2</sub> over central-eastern China from May 2013 to April 2014 (Qin et al. 2017). However, there is a lack of ground-level NO<sub>2</sub> concentration measurements in low-income countries, which induces broad uncertainty in the global ground-level NO<sub>2</sub> distribution (Larkin et al. 2017). However, to help detect the adverse impacts of NO<sub>2</sub> on public health and human well-being, global time-series ground-level NO<sub>2</sub> data are strongly desired.

Figure 1 is located here.

Geographically weighted regression (GWR) is an advanced method that estimates air pollution while taking the spatial contexts into account (Bigdeli et al. 2021; Jiang et al. 2017). To examine the relationship between satellite-derived data and measured ground-level air pollution, a variety of models have been developed, such as land-use regression (Bechle et al. 2015; Larkin et al. 2017), chemical transport models (Geddes et al. 2016), linear regression model (Lin et al. 2019; Liu et al. 2017), GIS-based multisource and multibox models (Wang and Chen 2013), geographically and temporally weighted regression (GTWR) (Li et al. 2020; Qin et al. 2017), and machine learning methods (Kim et al. 2021). All these models show relatively high accuracy in the estimations. To further improve the accuracy, we take the spatial variability of the relationship between satellite-derived data and measured ground-level air pollution into account. The GWR can illustrate the spatial variability of the relationship according to the local regression technique and spatial weights (Fotheringham and Oshan 2016; Hu et al. 2013). It must be noted that here, the global models are not spatially global models but statistically global models. The statistically global models mainly investigate the conditional interactions between the dependent and independent variables on the complete dataset, while the statistically local models are based on a partial dataset taking each observing subject as the center (Beenstock and Felsenstein 2019).

Currently, GWR can only analyze cross-sectional data (data collected by observing many subjects at one period of time) or panel data (data collected by observing many subjects across time) by using pooled ordinary least squares (POLS). To incorporate the temporal effect into the analysis, there are two other advanced methods, namely, GTWR

89 (Fotheringham et al. 2015; Qin et al. 2017) and geographically weighted panel regression  
90 (GWPR) (Yu et al. 2021), to regress the balanced panel data (data collected by observing  
91 many subjects at all periods during the study period). GTWR is a spatially and temporally  
92 nonstationary model. Integrating spatial and temporal distance into spatiotemporal distance  
93 poses fresh unknowns, although Euclidean distance is still a powerful tool (Fotheringham  
94 et al. 2015; Li and Managi 2021a; Qin et al. 2017). Furthermore, the recent GTWR model  
95 requires that the input dataset is a balanced panel dataset (Gollini et al. 2015). Yu et al.’s  
96 GWPR improves GWR by using the least squares dummy variable estimator, but models  
97 with other effects, such as fixed effects and random effects, are under development (Yu et  
98 al. 2021). This method works when the panel data have only a handful of periods. Moreover,  
99 neither of these two methods can solve the problem with unbalanced panel data (except  
100 balanced panel data, all panel data are unbalanced) due to the issues in the spatially  
101 weighted matrix. However, this limitation in the GWPR model is induced by the current  
102 toolboxes in most statistical software rather than the GWPR algorithm itself, although none  
103 of the current toolboxes specifically addresses the GWPR. For example, the outstanding  
104 GWR package “GWmodel” in R can also execute GWPR in specific situations using  
105 pretransformed datasets. However, this algorithm builds spatially weighted matrix rows for  
106 each observation, which consumes a large volume of computer memory and might cause  
107 failure due to a lack of memory. Furthermore, when it calibrates the adaptive distance  
108 bandwidth in GWPR using unbalanced panel data, it might miscalculate the numbers of  
109 spatial neighbors of some observations because the observation numbers vary in each  
110 location in unbalanced panel data. In fact, most observed datasets are unbalanced, including  
111 our dataset. To analyze the panel data by GWPR and facilitate other researchers, we bundle

the GWPR algorithm into a new R package, “GWPR.light” (<https://cran.r-project.org/web/packages/GWPR.light/index.html>), and release it on CRAN (The Comprehensive R Archive Network). The “GWPR.light” requires less memory even for big data calculations, solves the issues of the numbers of spatial neighbors of observations, and can perform GWPR with fixed or random effects.

In this paper, first, based on a series of statistical tests and model comparisons, we choose the GWPR with fixed time-fixed effects. This GWPR model has excellent accuracy. Second, the GWPR analysis involves several auxiliary variables, including terrain atmospheric pressure, land surface temperature, normalized difference vegetation index (NDVI), precipitation, planetary boundary layer height (PBLH), and year dummy variables, to detect the relationship between the ground-level and TrCA of NO<sub>2</sub>. Third, we calibrate the optimal bandwidth, which is critical to building the spatially weighted matrix. Fourth, we estimate the spatially nonstationary coefficients using the GWPR model, then interpolate the coefficients into raster data. Finally, we provide the monthly global ground-level NO<sub>2</sub> concentration from January 2015 to October 2021.

## Materials and Methodology

### *Materials*

#### *Satellite-Derived NO<sub>2</sub> Estimation*

The satellite Ozone Monitoring Instrument (OMI) provides vertical column amounts of NO<sub>2</sub>. The OMI/Aura NO<sub>2</sub> Total and Tropospheric Column Daily L2 Global Gridded 0.25 degree × 0.25 degree V3 (OMNO2G) is the global daily grid dataset with a

0.25-degree resolution to gauge TrCA and ToCA of NO<sub>2</sub> (**Figure 1**). It is based on the observation of the NASA Earth Observing System (EOS) Aura satellite platform (Irie et al. 2012; Li and Wu 2021; Nickolay et al. 2019) ([https://disc.gsfc.nasa.gov/datasets/OMNO2G\\_003/summary](https://disc.gsfc.nasa.gov/datasets/OMNO2G_003/summary)). This satellite was launched on July 15, 2004, and began providing daily global TrCA, stratospheric column amounts and ToCA of NO<sub>2</sub> observations (unit: *molecules/cm<sup>2</sup>*) from October 1, 2004 (Nickolay et al. 2019; OMI Team 2012). The OMI-NO<sub>2</sub> datasets, especially OMNO2G, have been widely used in emission and pollution monitoring (Curier et al. 2014; Li and Wu 2021; Liu et al. 2017). Air pollution near the surface impacts human health and daily life most harmfully (Hamra et al. 2015; Rice et al. 2013). However, the ground-level NO<sub>2</sub> concentration is unavailable in many areas, particularly in rural areas and most developing countries. Therefore, the conversion or estimation from TrCA or ToCA to ground-level data becomes a critical issue. The EOS Aura orbit is a sun-synchronous polar orbit at an altitude of 705 km, and the satellite goes over each location between 12:00 and 15:00 local time (OMI Team 2012; Schoeberl et al. 2006). Due to the orbit altitude, the ToCA of NO<sub>2</sub> covers the NO<sub>2</sub> molecules in the troposphere, stratosphere, mesosphere, and thermosphere, while the TrCA of NO<sub>2</sub> takes only the troposphere into account. Because the troposphere is most affected by human activities and air pollution, it has the most substantial impact on human health (Lelieveld et al. 2015; Li and Managi 2022) and well-being (Li and Managi 2021b), and the TrCA of NO<sub>2</sub> is an ideal independent variable to predict the ground-level NO<sub>2</sub> concentration. However, owing to anthropogenic climate change, tropopause height is slowly rising (Meng et al. 2021). Although this continuous rise is relatively small in comparison to the tropopause height, whether it influences the estimation of the TrCA of

NO<sub>2</sub> remains elusive. To choose the satellite NO<sub>2</sub> variable, we put both the ToCA and TrCA of NO<sub>2</sub> in the analyses as the primary variables. Both the goodness of fit of the model taking the TrCA of NO<sub>2</sub> as the dependent variable ( $R^2$ : 74.45%) and the cross validation results are slightly better than those of the other model using the ToCA of NO<sub>2</sub> ( $R^2$ : 74.39%). Moreover, compared with the satellite orbit height, the tropopause height is closer to the surface, which might make the TrCA of NO<sub>2</sub> more accurate. Therefore, the TrCA of NO<sub>2</sub> is employed in the final analysis.

We convert the unit of the TrCA of NO<sub>2</sub> from *molecules/cm<sup>2</sup>* to micrograms per square meter ( $\mu\text{g}/\text{m}^2$ ), to make it understandable to the public as follows:

$$ATrCA = RTrCA/N_A \times MM_{NO_2} \times 10,000 \text{ cm}^2/\text{m}^2 \times 1,000,000 \mu\text{g}/\text{g} \quad (1)$$

where *ATrCA* is the TrCA of NO<sub>2</sub> after the unit conversion, whose unit is  $\mu\text{g}/\text{m}^2$ , *RTrCA* is the value extracted from the OMNO2G directly, whose unit is *molecules/cm<sup>2</sup>*,  $N_A$  is the Avogadro constant ( $6.022 \cdot 10^{23} \text{ mol}^{-1}$ ), and  $MM_{NO_2}$  is the molar mass of NO<sub>2</sub> (46.007 *g/mol*).

The temporal resolution of our analysis is monthly, although both the OMNO2G and ground-level measurements are seemingly daily. Due to the orbit height of the EOS Aura, the satellite revisit period is 16 days and the orbit period is 98.83 minutes (Schoeberl et al. 2006). The EOS Aura's telescope field of view is 2600 km on the ground. Its daily data are from the day-lit portion of the satellite's 14 orbits. For all these reasons, each location could be observed several times in a month rather than every day. In other words, some locations are unobserved in a few days. Furthermore, other auxiliary satellite data, such as land surface temperature and NDVI, are obtained from the EOS Terra and Aqua



satellites, whose satellite revisit periods are also 16 days. For these reasons, we convert the daily TrCA of NO<sub>2</sub> from the OMNO2G dataset into the monthly average TrCA of NO<sub>2</sub>.

### *Ground-Level NO<sub>2</sub> Measurements*

This dataset can be downloaded from the Air Quality Open Data Platform (<https://aqicn.org/data-platform/covid19/>). The data have been available in over 530 major cities since 2015. The data provider has converted to the unified unit, part per billion (PPB), in accordance with the requirement of the United States Environmental Protection Agency. Initially, the dataset included 539 cities, but nine were dropped because there were no more than five monthly records from January 2015 to October 2021. This study exploits the monthly average NO<sub>2</sub> concentration as the dependent variable, calculated from the daily data, to merge with the satellite data correctly. The Air Quality Open Data Platform also provides the locations of the cities with measurement points (<https://aqicn.org/data-platform/covid19/airquality-covid19-cities.json>) (**Figure 2, Supplementary Materials Table S1: City List**). In most cities, there are generally several measurement points. The dataset records the medians of ground-level NO<sub>2</sub> using the data from different measurement points in a specific city. The ground-level measured NO<sub>2</sub> concentration data are daily data. To connect them with the satellite data, we convert them into monthly average values. It must be underscored that we use all the daily observations to calculate the monthly mean of ground-level NO<sub>2</sub>. However, there is no significant difference between the means obtained from all the daily observations and using the observations sampled at OMI measurement dates. In particular, the correlation between the means in each measurement using all the daily observations and using the observations sampled at

OMI measurement dates is 0.980 (p value < 0.1%). The variation caused by the difference is only 4.17% of the total variation of the means using all daily observations. Therefore, the means of all the daily observations are valid. Additionally, OMNO2 is a Level-2 grid product. The missing values of the satellite measurement appear in the data swath gaps. Even though the satellite does not observe a specific city on a certain day, other grids close to that city might be detected. Thus, ignoring some ground-level measurements due to no satellite data might decrease the accuracy of predictions.

*Figure 2 is located here.*

The unit of monthly average ground-level measured NO<sub>2</sub> concentration data is PPB, while the unit of the monthly average TrCA of NO<sub>2</sub> is micrograms per square meter ( $\mu g/m^2$ ). To make the units of these two datasets consistent, we convert the unit of monthly average ground-level measured NO<sub>2</sub> concentration data from PPB to  $\mu g/m^3$  using the following equation:

$$GNO2 = \frac{P \times MW_{NO2} \times RGN02}{R \times T \times 1000} \quad (2)$$

where  $GNO2$  is the monthly average ground-level NO<sub>2</sub> concentration with unit  $\mu g/m^3$ ,  $RGN02$  is the NO<sub>2</sub> concentration before conversion, whose unit is PPB,  $P$  is the terrain atmospheric pressure (mmHg), which can be extracted from OMNO2G and converted into the monthly average values (OMNO2G data unit is hPa, and the conversion coefficient between mmHg and hPa is 0.750 mmHg/hPa),  $MW_{NO2}$  is 46.007,  $R$  is the ideal gas constant ( $62.4 L \cdot Torr \cdot K^{-1} \cdot mol^{-1}$ ), and  $T$  is the monthly average temperature from MODIS product, MOD11C3.

## 222 *Auxiliary Data*

223         Auxiliary data, including terrain atmospheric pressure, temperature, NDVI,  
224 precipitation, and PBLH, are taken as the control variables to improve the model's  
225 performance. A considerable number of previous studies have involved terrain atmospheric  
226 pressure in the analyses of NO<sub>2</sub> concentrations (Chi et al. 2021; Li et al. 2020). Terrain  
227 atmospheric pressure data are from OMNO2G, with a 0.25-arc-degree resolution.  
228 Although the temporal resolution of this dataset is daily, it has the same problem as the  
229 TrCA of the NO<sub>2</sub> dataset, namely, a long satellite revisit period of the Aura satellite.  
230 Therefore, we convert it into a monthly average value. Additionally, to make the  
231 calculation understandable, the unit of land surface temperature is converted into Celsius.

232         The temperature data are provided by the Moderate Resolution Imaging  
233 Spectroradiometer (MODIS, <https://modis.gsfc.nasa.gov/>). The products, MOD11C3 and  
234 MYD11C3, are based on the MODIS Terra and Aqua satellite observations and  
235 measurements, including the daytime and nighttime average temperatures, with a 0.05-arc-  
236 degree spatial resolution (<https://modis.gsfc.nasa.gov/data/dataproduct/mod11.php>). Due to  
237 the 16-day revisit period, there are one or two global raster data based on each satellite's  
238 monthly observation. In a month, we take the mean of all the raster data, involving the  
239 daytime and nighttime average temperature in a month, to obtain the monthly average  
240 temperature raster. We note that the spatial resolution of the monthly average temperature  
241 raster is still 0.05 arc degrees. To make it consistent with the monthly average TrCA of  
242 NO<sub>2</sub>, we reduce the resolution to 0.25 arc degrees by applying the average value.

243         The monthly NDVI data are acquired from the MODIS products, MOD13C2 and  
244 MYD13C2 (<https://modis.gsfc.nasa.gov/data/dataproduct/mod13.php>), with a 0.05-arc-

degree spatial resolution. We also use the average value to make its resolution 0.25 arc degrees. The NVDI value ranges from -1 to 1. Because the ground-level NO<sub>2</sub> data are from cities, the situation in rural areas is not carefully gauged. To enable the model to predict the environment in rural areas, we incorporate the NDVI into the analysis since the most notable difference between urban and rural areas is landscape and human activity density. Both the NDVI and land cover raster can depict landscape, but land cover data are discrete and yearly and hence are not applicable for this analysis. Therefore, we employ the NDVI to reduce the influence of the difference between urban and rural areas. Additionally, nighttime light is associated with transportation (Chang et al. 2019) and human activity (Zhao et al. 2019), which are related to air pollution emissions. To represent their impacts, we also put the nighttime light data into the initial model. However, this has little effect on the prediction accuracy. Hence, it is removed in the final model.

The precipitation data are obtained from NASA Global Land Data Assimilation System Version 2 (GLDAS) Noah Land Surface Model L4 monthly 0.25 x 0.25 degree V2.1 (GLDAS\_NOAH025\_M, [https://disc.gsfc.nasa.gov/datasets/GLDAS\\_NOAH025\\_M\\_2.1/summary](https://disc.gsfc.nasa.gov/datasets/GLDAS_NOAH025_M_2.1/summary)). The temporal and spatial resolutions are one month and 0.25 arc degrees, respectively. This dataset has some missing values because some regions are too close to the seas and oceans. To avoid the reduction of the data size in our calculation, we impute the missing values by averaging their closest surrounding grids. We estimate the dataset twice, that is, the original GLDAS\_NOAH025\_M dataset boundary extends approximately 0.5 arc degrees into the seas and oceans. The original unit of this dataset is  $kg/(m^2 \cdot s)$ , but most values are less

than  $0.0001 \text{ kg}/(\text{m}^2 \cdot \text{s})$ . To improve the readability and robustness of the calculation, we convert its unit to  $\text{kg}/(\text{m}^2 \cdot \text{h})$ .

The PBLH data are extracted from Monthly mean, Time-Averaged, Single-Level, Assimilation, Surface Flux Diagnostics V5.12.4 (M2TMNXFLX, [https://disc.gsfc.nasa.gov/datasets/M2TMNXFLX\\_5.12.4/summary](https://disc.gsfc.nasa.gov/datasets/M2TMNXFLX_5.12.4/summary)). The spatial resolution of this dataset is  $0.5^\circ \times 0.625^\circ$ . We employ the bilinear interpolation method to impute a new dataset with a 0.25-arc-degree resolution. The unit of PBLH is meters (m).

**Figure 3** demonstrates the variable selection process of this research. Simply put, we select a variable that meets one of the following three requirements. First, the variable is significantly related to the dependent variable and has low correlations with other variables in the model. Second, the variable can significantly improve the model accuracy. Third, previous studies prove that the variable is necessary for the analysis. In the model design stage, we consider many other variables, including humidity, nighttime light, wind speed, cloud pressure, aerosol optical depth, and ultraviolet aerosol index. However, none of these variables are effective due to either their poor contribution to improving the prediction accuracy or due to the multicollinearity between independent variables. For example, the correlation coefficient between humidity and precipitation significantly exceeds 0.5. Furthermore, the correlation coefficient between humidity and temperature is over 0.7, while the relationship between precipitation and temperature is not as strong, namely, 0.15. Therefore, we keep the precipitation and temperature in the analysis. Moreover, the GWPR results are spatial point data frames. To secure the coefficient raster data for the prediction, we have to employ interpolation methods, which would cause some slight errors. If the benefit of involving one more variable cannot offset the errors in its

interpolation, it should not be taken into account. In view of numerous tests, finally, we select the TrCA of NO<sub>2</sub>, terrain atmospheric pressure, temperature, NDVI, precipitation, and PBLH as independent variables.

*Figure 3 is located here.*

### *Descriptive Statistics*

We establish an unbalanced dataset based on the above data sources for modeling and prediction with timestamps and geographical locations from January 2015 to October 2021. The number of complete observations in this dataset is 29,686. The imbalance of data is expressed in both temporal and spatial extent. Every city has at least five months and at most 65 months of data, i.e., it is the imbalance in temporal extent. Moreover, the spatial distribution is also unbalanced. For instance, the smallest distance between the two closest cities is 0.103 arc degrees, while the largest distance is 24.710 arc degrees. As shown in **Figure 2**, most cities with the data are located in the Northern Hemisphere, while only a few are distributed in South America, Africa, and Australia, which increases the difficulty of the interpolation.

**Figure 4** summarizes the statistical distributions and descriptive statistics of the variables in the original dataset. The statistical distribution of the concentration of ground-level NO<sub>2</sub> is one peak and right-skewed because the skewness is 7.17. Due to the 4.17 skewness, the statistical distribution of the monthly TrCA of NO<sub>2</sub> is one right-skewed peak. The statistical distribution of terrain atmospheric pressure is one left-skewed peak, whose skewness is -2.37. The shape of the monthly average temperature distribution is one

slightly left-skewed peak because its skewness is -0.39. The statistical distribution of NDVI is one peak and slightly left-skewed (skewness: -0.34). The shape of its statistical distribution of precipitation is one right-skewed peak because its skewness is 2.25. The statistical distribution of the PBLH is one right-skewed peak (skewness: 1.79). According to the variable distribution shape illustrated in **Figure 4**, some variable distributions cannot be deemed normal distributions. Therefore, directly putting them into regression is unsuitable. Put another way, the POLS cannot be employed as the basic analysis model.

We choose instead the fixed effect model (FEM) as the basic algorithm based on a series of statistical tests. This will be described in the next section. In FEM, the data are transformed as follows:

$$\tilde{V}_{it} = V_{it} - \bar{V}_i \quad (3)$$

where  $\tilde{V}_{it}$  denotes the transformed dependent and independent variables in measurement  $i$  during month  $t$ ,  $V_{it}$  denotes the original dependent and independent variables in measurement  $i$  during month  $t$ , and  $\bar{V}_i$  denotes the mean values of each original variable in measurement  $i$ . **Figure 5** demonstrates the variable distributions in the transformed dataset. The variables in the transformed dataset are almost normally distributed because their skewnesses and kurtosises are close to 0 except for ground-level NO<sub>2</sub>, monthly TrCA of NO<sub>2</sub>, and precipitation. The distributions of those three variables in the transformed dataset are right-skewed and leptokurtic because their skewnesses are 11.74, 4.00, and 1.68, and their kurtoses are 556.92, 54.23, and 10.11, respectively. We find that the transformed dataset still has 17 extremely large numbers of ground-level NO<sub>2</sub>, which are more than ten standard deviations from the mean. They mainly appear in five cities, including Izmir (Turkey), Konya (Turkey), Khorramabad (Iran), Qom (Iran), and Puebla (Mexico). If we

ignore these extreme numbers, the skewness and kurtosis decrease significantly. It must be underscored that the extreme numbers in the transformed dataset mean that the variable in specific measurements fluctuated significantly, rather than that they are absolutely large. The variables, monthly TrCA of NO<sub>2</sub> and precipitation, show similar situations, although extreme values appear in different cities. Because the share of the extreme numbers is no more than 0.5%, their impacts on regression can be ignored.

*Figures 4 and 5 are located here.*

**Figures 6 and 7** show the linear trends and correlation coefficients between the concentration of ground-level NO<sub>2</sub> and other variables in the original and transformed datasets, respectively. In **Figures 6 and 7**,  $r$  and  $p$  represent Pearson's correlation coefficient and p value, respectively (Taylor 1990).  $r$  is calculated as follows:

$$r = \frac{\sum_{i=1}^n (X1_i - \bar{X1})(X2_i - \bar{X2})}{\sqrt{\sum_{i=1}^n (X1_i - \bar{X1})^2} \sqrt{\sum_{i=1}^n (X2_i - \bar{X2})^2}} \quad (4)$$

where  $r$  is the correlation coefficient,  $X1$  and  $X2$  are the aimed variables put into the test, and  $n$  is the data size. To estimate the significance of the correlation coefficient, we calculate the t test value ( $t$ ) at first:

$$t = \frac{r\sqrt{n-2}}{\sqrt{1-r^2}} \quad (5)$$

Here, the p value ( $p$ ) is calculated as follows:

$$p = 2 \times \Pr(T > t) \quad (6)$$

where  $T$  follows a t distribution with  $n - 2$  degrees of freedom, and  $\Pr(T > t)$  represents the probability of  $T > t$ . According to **Figures 6 and 7**, all correlation coefficients are significant at the 0.1% level.



Figures 6 and 7 are located here.

## **Methodology**

### *Spatially Stationary Global Model*

We first assume that the relationships between the ground-level concentration of NO<sub>2</sub> and other independent variables are spatially stationary (Brunsdon et al. 1998; Fotheringham et al. 2002). We apply the three typical global models, namely, the random effects model (REM), FEM, and POLS, to estimate the relationships among the variables. The only difference between these three global models is the time-fixed effects in the models. Therefore, substantial tests about the time-fixed effects are needed.

We select global models with different time-fixed effects, namely, REM, FEM, and POLS. To compare FEM with POLS, we apply the F test for individual effects. FEM is better than POLS since the test result is significant, and the null hypothesis that no time-fixed effects are needed is rejected (Breusch and Pagan 1980; Croissant and Millo 2008). Based on the Breusch-Pagan Lagrange Multiplier test for random effects, the significant result indicates that REM is appropriate because the null hypothesis that no time-fixed effects are needed is not accepted (Breusch and Pagan 1980). We also run the Hausman test to detect a reasonable model between FEM and REM (Kang 1985). The significant result indicates that FEM is preferred, as the null hypothesis that the preferred model is REM is rejected. Here, the FEM in our analysis is as follows:

$$GNO2_{it} = \beta X'_{it} + \alpha_i + \mu_{it} \quad (7)$$

where  $GN02_{it}$  denotes the monthly average ground-level concentration of  $NO_2$  ( $\mu g/m^3$ ) in measurement  $i$  during month  $t$ ,  $\mathbf{X}_{it}$  denotes a matrix of independent variables, including TrCA of  $NO_2$  ( $\mu g/m^2$ ), terrain atmospheric pressure (hPa), temperature ( $^{\circ}C$ ), NDVI, precipitation ( $kg/(m^2 \cdot h)$ ), PBLH (m), and year dummy variables in measurement  $i$  during month  $t$ ,  $\alpha_i$  denotes the time-fixed effects,  $\mu_{it}$  denotes an idiosyncratic error, and  $\beta$  is a vector of parameters to be estimated. Some ignored variables, such as air pollution policy in countries and economic status, among others, are changed year by year. We therefore add the year dummy variables to the model.

### *Spatially Nonstationary Local Model*

Unlike global linear regression models, including REM, FEM, and POLS, the GWPR allows the coefficients to vary spatially (Brunsdon et al. 2010; Fotheringham and Oshan 2016). Because the relationship between the dependent variable and the independent variables could vary by the spatial context, the estimation accuracy of the ground-level  $NO_2$  concentration, such as the goodness of fit of the models, dramatically improves. Basically, the GWPR divides the total sample into numerous subsamples according to the optimal bandwidth and spatially weighted matrix. The bandwidth in GWPR is the threshold distance to judge whether two objects have a spatial relationship (Beenstock and Felsenstein 2019; Brunsdon et al. 2010; Brunsdon et al. 1998). If two objects have a spatial relationship, they are considered to be neighbors of each other. Currently, two types of bandwidth are widely used, fixed distance bandwidth and adaptive distance bandwidth (Gollini et al. 2015). The fixed distance bandwidth is a fixed threshold distance, and if two spatial objects are within this distance, they are spatially related. The adaptive distance

bandwidth is the number of objects in a subsample because every object always has a certain number of neighbors. Hence, the threshold distance for each object might be different. To calibrate the optimal bandwidth in the GWPR, the mean square prediction error is the critical index. When the mean square prediction error is the smallest, the model with a certain bandwidth has the highest accuracy (Gollini et al. 2015). For a specific bandwidth ( $b$ ), the mean square prediction error is calculated as follows:

$$MSPE(b) = \frac{m \sum_j [y_j - \widehat{y_j(b)}]^2}{(m - p + 1)^2} \quad (8)$$

where  $MSPE(b)$  is the mean square prediction error based on a specific bandwidth  $b$ ,  $m$  is the data size,  $y_j$  is the dependent variable's value of the  $j$ th record,  $\widehat{y_j(b)}$  is the predicted value of the  $j$ th record based on bandwidth  $b$ , and  $p$  is the number of parameters in the analysis. It must be noted that bandwidth  $b$  can be either the fixed distance bandwidth or the adaptive distance bandwidth here. The GWPR is an improvement of GWR in that it considers time-fixed effects. In GWR, previous studies generally assume that **Equation 8** is a U-shaped function (Brunsdon et al. 1998; Fotheringham et al. 2002; Gollini et al. 2015). However, as those studies mentioned, there is no statistical or mathematical evidence for or against this assumption. The benefit from this assumption is the reduction of calculating time because the calculation could use the golden-section search to find the extremum to obtain optimal bandwidth (Gollini et al. 2015).

The step increment selection is applied to calibrate the optimal bandwidth due to the recently designed program that requires fewer computer resources. We calibrate the optimal fixed and adaptive distance bandwidths. In the fixed distance bandwidth selection, the selection extent is from 0.25 arc degrees to 20 arc degrees, and the examined bandwidth

increases by 0.25 arc degrees every time because the spatial resolution of all the grid data is 0.25 arc degrees. The optimal fixed distance bandwidth is 2.25 arc degrees (**Figure 8**). Since the spatial distribution of the cities with data is unbalanced and some points are far from others (**Figure 2**), many points would have no neighbor if using this bandwidth. Data islands severely destroy the stability of the GWPR because they have no relationship with other points in the dataset. Dropping the data islands and increasing the bandwidth would slash the accuracy of the analysis. Hence, we adopt the adaptive distance bandwidth. In the adaptive distance bandwidth selection, the selection extent is from 4 to 100, and the tested bandwidth increases by one every time. When the adaptive distance bandwidth is less than 4, some local regressions do not have enough degrees of freedom. According to the analysis, the optimal adaptive distance bandwidth is selected, which is 7. In other words, in each subsample, every locally weighted regression involves seven cities' data.

*Figure 8 is located here.*

The spatially weighted matrix, based on the calibrated bandwidth, is calculated as follows:

$$\mathbf{W}_i = \begin{cases} [1 - (\frac{\mathbf{d}}{b_i})^2]^2, d_k \in \mathbf{d} \text{ if } d_k \leq b_i \\ 0, d_k \notin \mathbf{d} \text{ if } d_k > b_i \end{cases} \quad (9)$$

where  $\mathbf{W}_i$  denotes a vector of spatial weights between city  $i$  and its neighbors,  $\mathbf{d}$  represents a vector of distances between city  $i$  and its neighbors,  $b_i$  is the threshold distance of city  $i$ 's neighbors, and  $d_k$  is the distance between city  $i$  and city  $k$ . It should be emphasized that  $b_i$  might vary city by city, as here we use the adaptive distance bandwidth. Every vector is a column of the spatially weighted matrix. According to the spatial weight vectors, the total sample is divided into 530 subsamples, taking every city as the center.

437 The GWPR estimates the coefficients in every location in light of the spatially  
 438 weighted matrix. In global model statistical tests, FEM is the preferred model, and in the  
 439 locally statistical test, most local regressions show significant time-fixed effects. Therefore,  
 440 the GWPR model also follows the FEM. The GWPR model is expressed as follows:

$$GNO2_{it} = \beta_i X'_{it} + \alpha_i + \epsilon_{it} \quad (10)$$

441 where  $\beta_i$  denotes a vector of parameters in the regression with the subsample taking city  $i$   
 442 as the center. In terms of the equation expression, the only difference between FEM  
 443 (**Equation 7**) and the GWPR based on FEM (**Equation 10**) is that  $\beta_i$  varies spatially. It  
 444 must be noted that the algorithm would change to the GWPR based on POLS if we directly  
 445 used the GWR method. According to cross validation, the accuracy and stability of the  
 446 GWPR based on POLS are lower than those of the GWPR based on FEM. Thus, the new  
 447 method, GWPR, is necessary for this analysis.

448 To estimate the coefficient in the GWPR, because the time-fixed effects ( $\alpha_i$ ) are  
 449 unknown (Croissant and Millo 2008), we need to expand **Equation 10** as follows:

$$GNO2_{it} - \hat{\theta}_i \overline{GNO2}_i = (1 - \hat{\theta}_i) \beta_{0i} + \beta_i (X_{it} - \hat{\theta}_i \overline{X}_i)' + (\epsilon_{it} - \hat{\theta}_i \overline{\epsilon}_i) \quad (11)$$

450 where  $\hat{\theta}_i$  denotes the share of the individual effects in the total effects,  $\overline{GNO2}_i$  denotes the  
 451 mean of the ground-level NO2 concentration in the dataset in city  $i$ ,  $\beta_{0i}$  is the estimated  
 452 intercept in city  $i$ ,  $\overline{X}_i$  denotes a vector of means of independent variables in city  $i$ , and  $\overline{\epsilon}_i$   
 453 denotes the mean of the error in city  $i$ . If the GWPR follows the data transformation of  
 454 FEM,  $\hat{\theta}_i$  is always equal to 1. Therefore,  $\beta_{0i}$  is removed, in the GWPR based on FEM. To  
 455 simplify **Equation 11**, we define the matrix of independent variables  $x_{it}$  and the dependent  
 456 variable  $gno2_{it}$  as follows:

$$\mathbf{x}_{it} = \mathbf{X}_{it} - \overline{\mathbf{X}}_t \quad (12)$$

$$gno2_{it} = GNO2_{it} - \overline{GNO2}_t \quad (13)$$

$$\sigma_{it} = \epsilon_{it} - \overline{\epsilon}_t \quad (14)$$

457 In light of **Equations 12-14**, the GWPR model is transformed as follows:

$$gno2_{it} = \beta_i \mathbf{x}'_{it} + \sigma_{it} \quad (15)$$

458 This algorithm becomes the typical GWR without the intercept. The coefficient estimation  
459 can be displayed as follows:

$$\beta_i = [\mathbf{x}_{it}^T \mathbf{W}_i \mathbf{x}_{it}]^{-1} \mathbf{x}_{it}^T \mathbf{W}_i gno2_{it} \quad (16)$$

460

#### 461 *Statistical Indicators*

462 Previous studies widely use several statistical indicators, including  $R^2$ , root mean  
463 square error (RMSE), mean absolute error (MAE),  $r$  between observed and predicted  
464 values, and regression coefficients between observed and predicted values (Kim et al. 2021;  
465 Li and Wu 2021; Liu et al. 2017). Hence we also apply these indicators to evaluate the  
466 estimation accuracy of the GWPR model.  $R^2$  is a critical statistical indicator describing the  
467 goodness of fit. A high  $R^2$  in the models means that the discrepancy between the predicted  
468  $NO_2$  concentration and the measured ground-level  $NO_2$  concentration is small. The  $R^2$   
469 calculation is as follows:

$$R^2 = 1 - \frac{\sum_{k=1}^n (MNO2_k - PNO2_k)^2}{\sum_{k=1}^n (MNO2_k - \overline{MNO2})^2} \quad (17)$$

470 where  $n$  represents the observation number in the total sample,  $MNO2_k$  represents the  $k$ th  
471 record of the measured ground-level  $NO_2$  concentration,  $PNO2_k$  represents the  $k$ th record  
472 of the predicted  $NO_2$  concentration, and  $\overline{MNO2}$  represents the mean of the measured

ground-level NO<sub>2</sub> concentration. It must be noted that this R<sup>2</sup> of the GWR or the GPWR is the global value because here the total sample is used. The RMSE is also calculated, indicating the differences between the predicted NO<sub>2</sub> concentration and the measured ground-level NO<sub>2</sub> concentration. The RMSE is sensitive to both systematic error and random error, shown as follows:

$$RMSE = \sqrt{\frac{1}{n} \sum_{k=1}^n (MNO2_k - PNO2_k)^2} \quad (18)$$

The MAE is calculated as follows:

$$MAE = \text{mean}(|MNO2_k - PNO2_k|) \quad (19)$$

In the analysis, the MAE should be low. Additionally, the regression coefficients between observed and predicted values are imputed as follows:

$$MNO2_k = \alpha + \beta PNO2_k + \delta_k \quad (20)$$

where  $\alpha$  is the intercept in the regression and the ideal value of  $\alpha$  is 0,  $\beta$  is the slope and its ideal value is 1, and  $\delta_k$  is a random error term.

#### *Spatial Interpolation of the Coefficients*

We use the ordinary kriging (OK) method, described in a previous study (Li and Heap 2011), to interpolate the coefficients of the GWPR result into raster data. Directly using the GWPR result is insufficient to predict and obtain the ground-level NO<sub>2</sub> concentration raster data because the result is a spatial point data frame and the coefficients vary spatially. There are several reasons for employing the OK method. First, kriging methods can obtain relatively high accuracy when the interpolation covers a large spatial scope (Bivand et al. 2008). Second, for each interpolation, it is valid that the variogram

model is the same on an even global scale. Third, the interpolated data are spatially isotropic. Fourth, we interpolate the local coefficients from the GWPR result, whose mean and variance are constant across the spatial field. The general form of the OK method is expressed (Pebesma 2004) as follows:

$$\hat{\beta}(s_0) = \sum_{l=1}^m \lambda_l \beta(s_l) \text{ with } \sum_{l=1}^m \lambda_l = 1 \quad (21)$$

where  $\hat{\beta}(s_0)$  represents the coefficient at unobserved location  $s_0$ ,  $\beta(s_l)$  represents the coefficient value at known location  $s_l$ ,  $\lambda_l$  is the spatial weight for the coefficient value at the  $l$ th location, and  $m$  is the number of known values. To obtain the optimal  $\lambda_l$ , there are two requirements in OK: unbiased and minimal variance of estimation. If  $\sum_{l=1}^m \lambda_l$  is equal to 1, then it is unbiased (Pebesma 2004). To minimize the estimation variance, we use the spherical semivariance mathematical model. Harnessing the OK method, the GWPR results would be estimated to several coefficient raster data with a 0.25-arc-degree spatial resolution. Additionally, the mean value raster data of variables are also interpolated by the OK method. In the prediction process, the data transformations following **Equations 12** and **13** are necessary. To check the reliability of the interpolation process, we apply leave-one-out cross validation. Simply stated, leave-one-out cross-validation requires that the number of folds equals the number of cities in our dataset. In the cross validation process, one specific city is selected as the test set, while all 529 other cities are used as the training set. This process will repeat 530 times until each city has been selected as the test set.

### *Predictions of Ground-level NO<sub>2</sub> Concentration and Change Trends*

The monthly global ground-level NO<sub>2</sub> concentration raster data are predicted based on the original raster data, coefficient rasters, and mean value rasters, following **Equation**



15. Furthermore, because the ground-level NO<sub>2</sub> concentration cannot be negative, any value less than zero is set to zero in the prediction grid data. The prediction process creates 82-month grid data. To detect the monthly change trends of each grid, the monthly NO<sub>2</sub> concentrations of each grid are regressed with their month order. The first month of our dataset is January 2015, so the month order of data in January 2015 is the 1<sup>st</sup>. The month order of October 2021 is 82<sup>nd</sup>. The monthly change trends are estimated as follows:

$$PNO2_{LT} = \alpha_L + \beta_L MO_T + \varphi_{LT} \quad (22)$$

where  $PNO2_{LT}$  is the predicted NO<sub>2</sub> concentration in grid  $L$  in the  $T$ th month,  $MO_T$  is the month order of  $PNO2_{LT}$ ,  $\alpha_L$  is the intercept in grid  $L$ ,  $\beta_L$  is the monthly NO<sub>2</sub> concentration change trend in grid  $L$ , and  $\varphi_{LT}$  is a random error term. It must be noted that the data from a certain grid should be subjected to the regression once if there are no less than 30 records. Additionally, every individual regression uses only the data from one grid.

## Results

### *Validation of the GWPR Results*

The overall accuracy of the GWPR based on FEM is 74.45%, the RMSE is 7.171  $\mu g/m^3$ , and the MAE is 3.48  $\mu g/m^3$ . The mean globally measured ground-level NO<sub>2</sub> is 19.47  $\mu g/m^3$ . **Figure 9** shows the relationship between the predicted and measured ground-level NO<sub>2</sub> concentrations in the GWPR, and all coefficients in **Figure 9** are significant. The coefficient of the correlation ( $r$ ) between the predicted and measured values is 0.863, following **Equation 4**, i.e., there is a strong correlation. Furthermore, the OLS regression between measured and predicted values also illustrates the significant

relationship since the slope is approximately 1.021 significantly, and the ideal value is 1. To confirm whether the goodness of fit of a specific year is apparently lower than that of other years, we calculate all the yearly statistical indicators. In **Table 1**, the yearly statistical indicators, including  $R^2$ , RMSE, correlation coefficient ( $r$ ), the slope and intercept of regressions between measured and predicted values, are listed. The lowest yearly  $R^2$  is 51.54% in 2021. The yearly RMSE and MAE peak in 2021 and 2018, respectively, which are 9.619 and  $4.071 \mu\text{g}/\text{m}^3$ . In September 2021, the measured ground-level  $\text{NO}_2$  concentration in Puebla, a Mexican city, suddenly increased to more than 20 times the usual level. In October 2021, the data of this city were missing. As we mentioned in the Materials section, this seemingly strange record also causes the high skewness of the measured ground-level  $\text{NO}_2$  concentration distribution. The available daily data in that city are close to its monthly average value in September 2021. Thus, we keep that record in the analysis. However, if we ignore this individual record, the  $R^2$  in 2021 will reach 78.10%. Comparisons of monthly time series illustrate that the GWPR captures the monthly temporal variability accurately (**Figure 10**).

*Figure 9 is located here.*

*Table 1 is located here.*

*Figure 10 is located here.*

**Table 2** reports the results of 10-fold cross validation. The total dataset is randomly divided into ten subsets. In every single cross validation, nine subsets are used to train the model. Then the reserved subset and the coefficients from the trained model are employed to predict the dependent variable. The performance of the GWPR based on FEM is stable in terms of goodness of prediction for the test subsets. The lowest prediction accuracy for

the test subset is 68.55%, and its RMSE, MEA, and coefficient ( $r$ ) of the correlation between the predicted and measured values are  $15.569 \mu\text{g}/\text{m}^3$ ,  $4.399 \mu\text{g}/\text{m}^3$ , and 0.627, respectively. The 10-fold cross validation results show the reliability of the GWPR model.

*Table 2 is located here.*

### ***Validation of the OK Interpolations***

**Table 3** illustrates the result of the leave-one-out cross validation of the coefficient and mean value interpolations. The worst performance of the OK method is from the interpolation of the mean value of measured ground-level  $\text{NO}_2$ , whose  $R^2$  is 34.50%, but its RMSE and MAE are not high, which are  $7.773$  and  $5.633 \mu\text{g}/\text{m}^3$ , respectively. For the low RMSE and MAE values, this interpolated grid dataset is still acceptable. There is a new indicator named on point  $R^2$  in **Table 3**. This indicator depicts how much the observed values change after interpolation. In the OK method, the spatial weight might not equal 1 when the distance between two points equals 0, i.e., they overlap. Hence, after OK interpolation, the predicted values from the imputed grid dataset might differ slightly from the observed values. The on-point  $R^2$  indicates that none of the interpolations lose the accuracy of more than 2%, and most of them only slash the accuracy by 0.1%.

*Table 3 is located here.*

### ***Prediction and Monthly Change Trends***

The accuracy of the predicted grid dataset is 69.61%, and the RMSE and MAE are  $7.82$  and  $4.07 \mu\text{g}/\text{m}^3$ , respectively. Due to the slight change in the interpolation process,

the accuracy of the final prediction is lower than that of the GWPR result. **Figure 11** demonstrates the relationship between the predicted and measured ground-level NO<sub>2</sub> concentrations in the predicted grid datasets, and all coefficients in **Figure 11** are significant. The  $r$  is 0.838, which is lower than the  $r$  in the GWPR result (0.859). According to the tiny difference of all statistical indicators between the GWPR result and the grid data prediction result, the final prediction data are reasonable.

*Figure 8 is located here.*

**Figure 12** presents the average monthly ground-level NO<sub>2</sub> concentration. The results show that the ground-level NO<sub>2</sub> concentration in Asia is relatively higher than that in other continents. Moreover, in western Asia, the ground-level NO<sub>2</sub> pollution is the most severe. In most regions of Iran, Iraq, and Turkmenistan, the average values even exceed 100  $\mu\text{g}/\text{m}^3$ . It is much higher than WHO's ambient NO<sub>2</sub> concentration limits, which is an annual mean of 40  $\mu\text{g}/\text{m}^3$  in link with previous studies (Kasparoglu et al. 2018; Yousefian et al. 2020). Moreover, some regions in northern China, India, Chile, Bolivia, South Africa, northern Italy, and the west coast of the United States are also severely polluted. This is consistent with previous research (Di et al. 2020; Kim et al. 2021; Zheng et al. 2019). Moreover, in **Figure 12**, we plot 110 main cities' average measured ground-level NO<sub>2</sub> concentrations from January 2015 to October 2021. The measured values are consistent with our predictions. The map of average monthly ground-level NO<sub>2</sub> concentrations without city points and monthly ground-level NO<sub>2</sub> concentrations from January 2015 to October 2021 are shown in the **Supplementary Materials (Figures S1 – S83)**. **Figure 13** shows the average monthly TrCA of the NO<sub>2</sub> concentration from January 2015 to October 2021. There are some differences between the two spatial distributions. As assumed, the

relationship between the TrCA of NO<sub>2</sub> and the ground-level NO<sub>2</sub> spatially varies. Thus, the same amounts of NO<sub>2</sub> in the troposphere might contribute disproportionately to the ground-level NO<sub>2</sub> concentration in the different places. **Figure 14** displays the monthly trends of the means of predicted ground-level NO<sub>2</sub> concentrations in all the grids and the cities with the measurements. Both the mean values of all the global grids and the cities with measurements show downward trends. The monthly change trends of each grid and the means of all the global grids and the cities with measurements are estimated following **Equation 20**. The slope of the regression between the average ground-level NO<sub>2</sub> concentration of the cities and month order is -0.097 (95% confidence interval: -0.140 - -0.053, p value < 0.1%), indicating an approximately 0.097 (95% CI: 0.053 – 0.140)  $\mu\text{g}/\text{m}^3$  decrease per month from January 2015 to October 2021. Moreover, the mean of all the grids decreases by approximately 0.066 (95% CI: 0.033 - 0.099)  $\mu\text{g}/\text{m}^3$  (p value < 0.1%) per month. **Figure 15** illustrates the monthly change trends of each grid. All the values shown in **Figure 15** should be significant at the 0.1 level. **Figure 15** demonstrates that the ground-level NO<sub>2</sub> concentration gradually decreases in most regions. However, in the Red Sea region, Arabian Peninsula, Persian Gulf, the ground-level NO<sub>2</sub> concentration is increasing, even though these are already the most severely polluted areas. In addition, the changes in most places in India and Mexico also show an upward trend. In some areas of California and Nevada in the United States too, the change trends are increasing.

*Figures 12-15 are located here.*

## Discussion

We employ the GWPR model to examine the association of satellite-derived data with measured ground-level NO<sub>2</sub> concentrations from January 2015 to October 2017 among 530 cities. The accuracies of the GWPR (74.45%) and raster prediction (69.61%) exceed most previous studies on regional ground-level NO<sub>2</sub> concentrations (Kim et al. 2021; Li et al. 2020; Qin et al. 2017). Our analysis provides the first example of the GWPR on an unbalanced panel dataset. Our study predicts 82 monthly global ground-level NO<sub>2</sub> concentrations with a 0.25-arc-degree spatial resolution from January 2015 to October 2021 (shown in the **Supplementary Materials**). Additionally, our analysis demonstrates the monthly change trends of each grid.

In terms of accuracy, our study has made some improvements in the ground-level NO<sub>2</sub> concentration estimation based on the satellite dataset. The accuracy of a study on ground-level NO<sub>2</sub> concentrations over Central-Eastern China employing GTWR is 0.60, which is lower than our 0.70 (Qin et al. 2017). Additionally, its MEA is  $9.28 \mu\text{g}/\text{m}^3$ , while our MEA is  $5.633 \mu\text{g}/\text{m}^3$ . A study over the Wuhan urban agglomeration applying space-time neural networks obtained 0.69 accuracy and an  $8.29\text{-}\mu\text{g}/\text{m}^3$  RMSE, while our RMSE was  $7.82 \mu\text{g}/\text{m}^3$  (Li et al. 2020). Additionally, a study over Switzerland and northern Italy using machine learning technology provides a 59%-accuracy prediction, and its MEA is  $7.69 \mu\text{g}/\text{m}^3$  (Kim et al. 2021). However, these studies excel at either spatial or temporal resolution.

The spatial distribution and concentrations of the predictions are compared with previous studies. According to a study by Di et al., in the metropolitan areas of Michigan,

Illinois, Wisconsin, New England, Colorado, Nevada and California, NO<sub>2</sub> pollution is at harmful levels, higher than 25 PPB (Di et al. 2020). Our prediction is in line with this study since the ground-level NO<sub>2</sub> in those places is predicted to be approximately 40  $\mu\text{g}/\text{m}^3$  (under an ambient pressure of 1 atmosphere and a temperature of 25 °C, 1 PPB NO<sub>2</sub> roughly equals 1.88  $\mu\text{g}/\text{m}^3$ ). Kasparoglu et al. proposed that in urban sites of the Marmara regression of Turkey, the NO<sub>2</sub> concentrations were higher than 75  $\mu\text{g}/\text{m}^3$  (Kasparoglu et al. 2018), consistent with our study. Yousefian et al. proposed that during 2012-2017, the NO<sub>2</sub> concentration in Tehran, the capital of Iran, was always 1.5-2.5 times higher than WHO limits (40  $\mu\text{g}/\text{m}^3$ ) (Yousefian et al. 2020), which was also the most polluted area in our research. Kim et al.'s research hints that in an alpine domain, the NO<sub>2</sub> in urban regions is always over 40  $\mu\text{g}/\text{m}^3$  (Kim et al. 2021), which is in keeping with our study. Additionally, Chi et al. estimated the annual average ground-level NO<sub>2</sub> concentration in the main regions of China over 2014-2020 (Chi et al. 2021). Its results are similar to our predictions. Moreover, because research on the ground-level NO<sub>2</sub> concentration in polar regions is rare, it is difficult to compare our study with previous studies. However, the low temperature and PBLH might account for the high NO<sub>2</sub> concentration in polar regions. Evidence shows that in winter, the NO<sub>2</sub> concentration is apparently higher than it is in summer (Fan et al. 2020; Shen et al. 2021; Yousefian et al. 2020). PBLH is negatively related to ground-level air pollution (Xiang et al. 2019). To summarize, our study is in line with the most recent studies on ground-level NO<sub>2</sub> concentrations.

It must be underscored that the spatial distributions of the ground-level NO<sub>2</sub> concentration and TrCA of NO<sub>2</sub> are different, as **Figures 12** and **13** show. Put another way, the impacts of the TrCA of NO<sub>2</sub> on the ground-level NO<sub>2</sub> concentration vary spatially.

**Figure 3.a** illustrates the discrepancy between the TrCA of NO<sub>2</sub> and the ground-level NO<sub>2</sub> concentration. For example, some points with no more than 1000  $\mu\text{g}/\text{m}^2$  TrCA of NO<sub>2</sub> have a more than 60  $\mu\text{g}/\text{m}^3$  ground-level NO<sub>2</sub> concentration. The potential reason for this is that most NO<sub>2</sub> in a certain column distributes near the surface, although the total amount of NO<sub>2</sub> in this column is small. Moreover, several other points with more than 20000  $\mu\text{g}/\text{m}^2$  TrCA of NO<sub>2</sub> have a less-than-30- $\mu\text{g}/\text{m}^3$  ground-level NO<sub>2</sub> concentration. In this case, there is a large amount of NO<sub>2</sub> molecules in a specific column, but most of them might be hundreds of meters away from the surface. Several geoengineering measures could drive air pollution to penetrate the planetary boundary layer into the troposphere to avoid clustering at the surface (Liu et al. 2021). Ground-level atmospheric monitoring equipment cannot detect these NO<sub>2</sub> molecules, even though emissions might be high there. Because the physical and chemical properties of the atmosphere affected by geographical and environmental factors are not the same everywhere, we employ GWPR to predict the ground-level NO<sub>2</sub> concentration rather than use other global models to estimate it directly. The spatially varied relationship between the TrCA of NO<sub>2</sub> and the ground-level NO<sub>2</sub> concentration is the main reason leading to the difference between **Figures 12** and **13**.

Our monthly grid data demonstrate that the temporal variation in the spatial distribution of the ground-level NO<sub>2</sub> concentration is evident (shown in **Supplementary Materials** and **Figure 10**). In December, January, and February, the NO<sub>2</sub> pollution over the Northern Hemisphere is relatively serious, while it is high in June, July, and August over the Southern Hemisphere. The possible reasons for this temporal variation are as follows: first, NO<sub>2</sub> can last longer, or it is difficult to be removed in a low-temperature environment (Yousefian et al. 2020); second, residential energy use for heating might



produce more NO<sub>2</sub> (Fan et al. 2020); and third, the low temperature changes other meteorological factors that eventually leads to NO<sub>2</sub> being unable to spread to other zones far from the surface (Fan et al. 2020; Shen et al. 2021). However, no matter what the main reason is, policy-makers should take this temporal variation into account when devising air pollution policies.

Some limitations remain in our study. First, the number of measurement points is limited, and their spatial distribution is unbalanced. Our dataset takes only 530 cities into account globally without considering rural areas. Especially in Africa and South America, there are few measurement points. Second, the spatial resolutions of one control variable, namely, PBLH, are not good enough. Third, some hidden variables are ignored. For instance, the research does not involve the impacts of air pollution policies and COVID-19, although we have used year dummy variables to reduce their effects. Fourth, the spatial and temporal resolutions are relatively low. Future studies should focus on the following topics. First, the GWPR model should be further optimized. For example, if the dataset contains data islands based on a certain bandwidth, whether the GWRP is still statistically reliable remains unknown. Second, more confounders, such as land cover and human activity, among others, need to be included in future analyses. Finally, the vertical distribution of NO<sub>2</sub> in air and its physical and chemical causes should be carefully explored. Third, the vertical distribution of air pollutants in the atmosphere in different places should be probed and analyzed.

## Conclusions

This study creates a GWPR model to detect the spatial variation in the relationship between the measured ground-level NO<sub>2</sub> concentration and satellite-derived data. It predicts the monthly ground-level NO<sub>2</sub> concentration based on the data from the OMI and other satellite platforms. The GWPR model exhibits outstanding accuracy (74.45%),  $r$  value (0.863), RMSE ( $7.171 \mu\text{g}/\text{m}^3$ ) and MAE ( $3.480 \mu\text{g}/\text{m}^3$ ) in the estimation and stability in the cross validation. The GWPR is a simple but effective model to estimate monthly global ground-level NO<sub>2</sub> concentrations at a 0.25-degree spatial resolution. The predictions provide critical basic data to environmental and public health science and valuable information for governments and societies to formulate more reasonable and efficient policies.

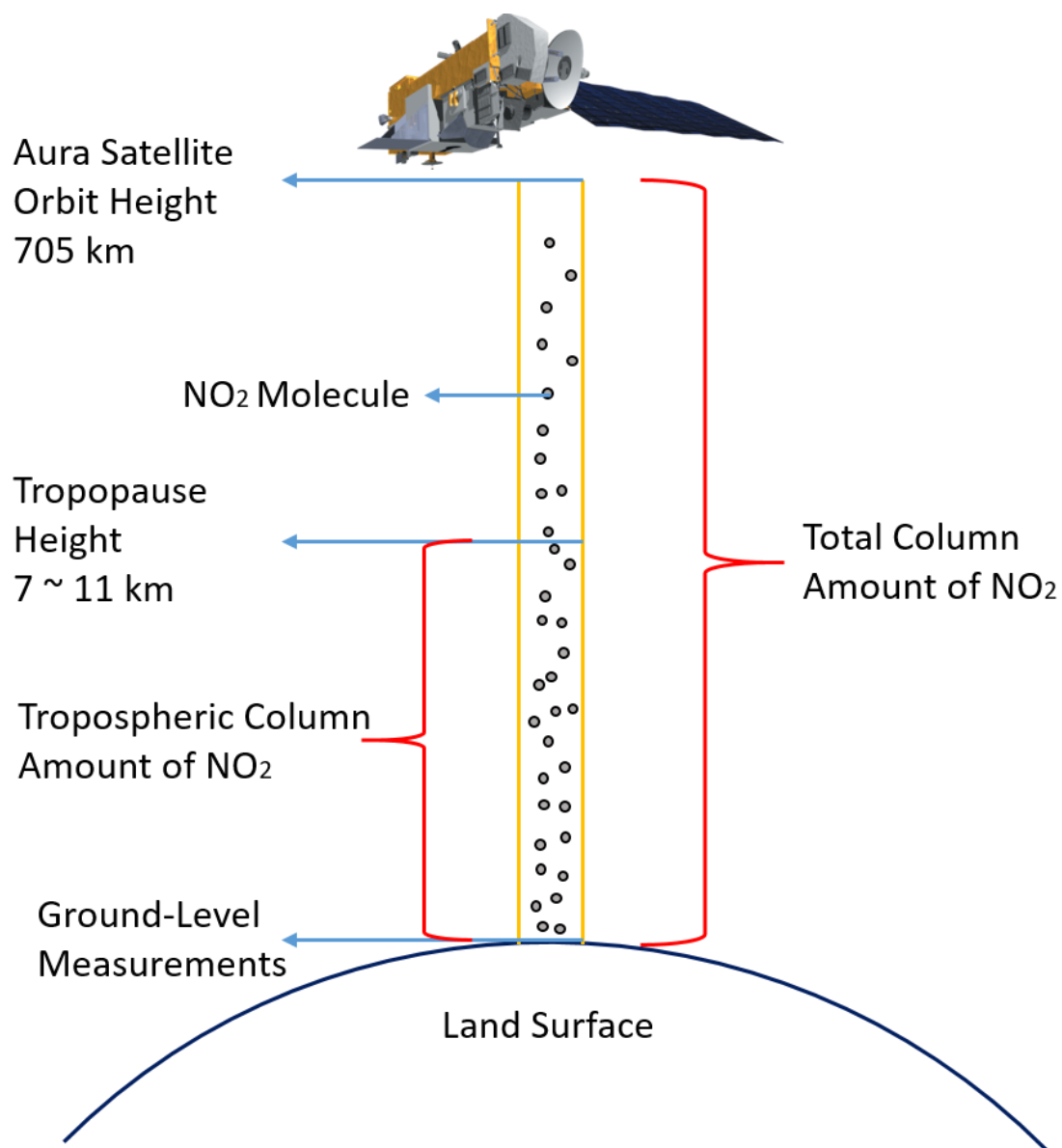
## Data Availability

All data sources used in the analyses, along with fully reproducible code, are publicly available at [https://github.com/MichaelChaoLi-cpu/Monthly\\_Global\\_Ground\\_Level\\_NO2](https://github.com/MichaelChaoLi-cpu/Monthly_Global_Ground_Level_NO2). The predicted data from January 2015 to October 2021 and the animation of ground-level NO<sub>2</sub> concentration could also be found in the abovementioned websites.

## **Acknowledgments**

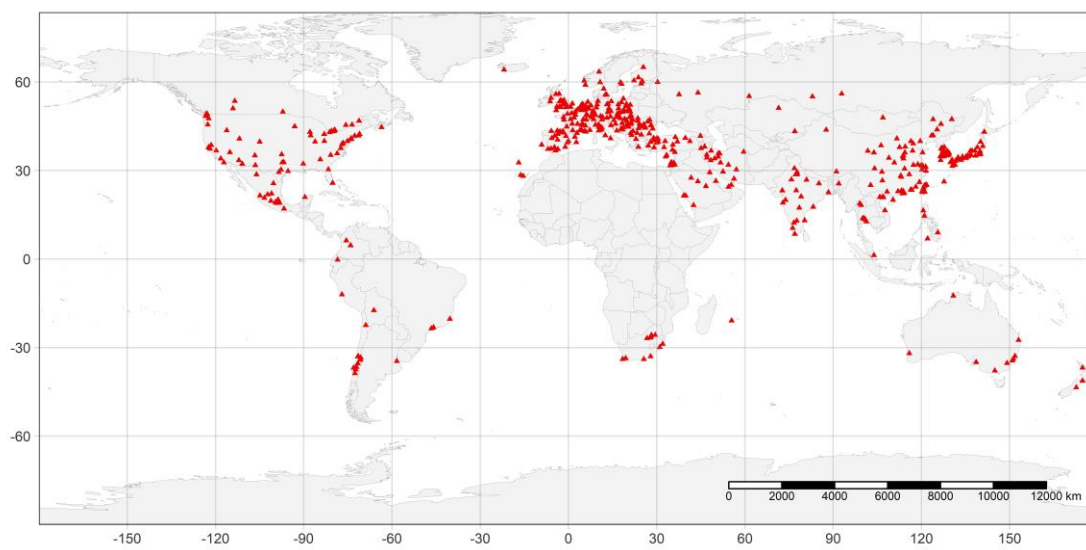
The authors gratefully acknowledge all the support, financial or otherwise, in enabling the work here to be carried out. This research was supported by the following funding agencies: JSPS KAKENHI (Grant No. JP20H00648), the Environment Research and Technology Development Fund of the Environmental Restoration and Conservation Agency of Japan (Grant No. JPMEERF20201001), and also JST SPRING (Grant No. JPMJSP2136).

740 **Figure**



741

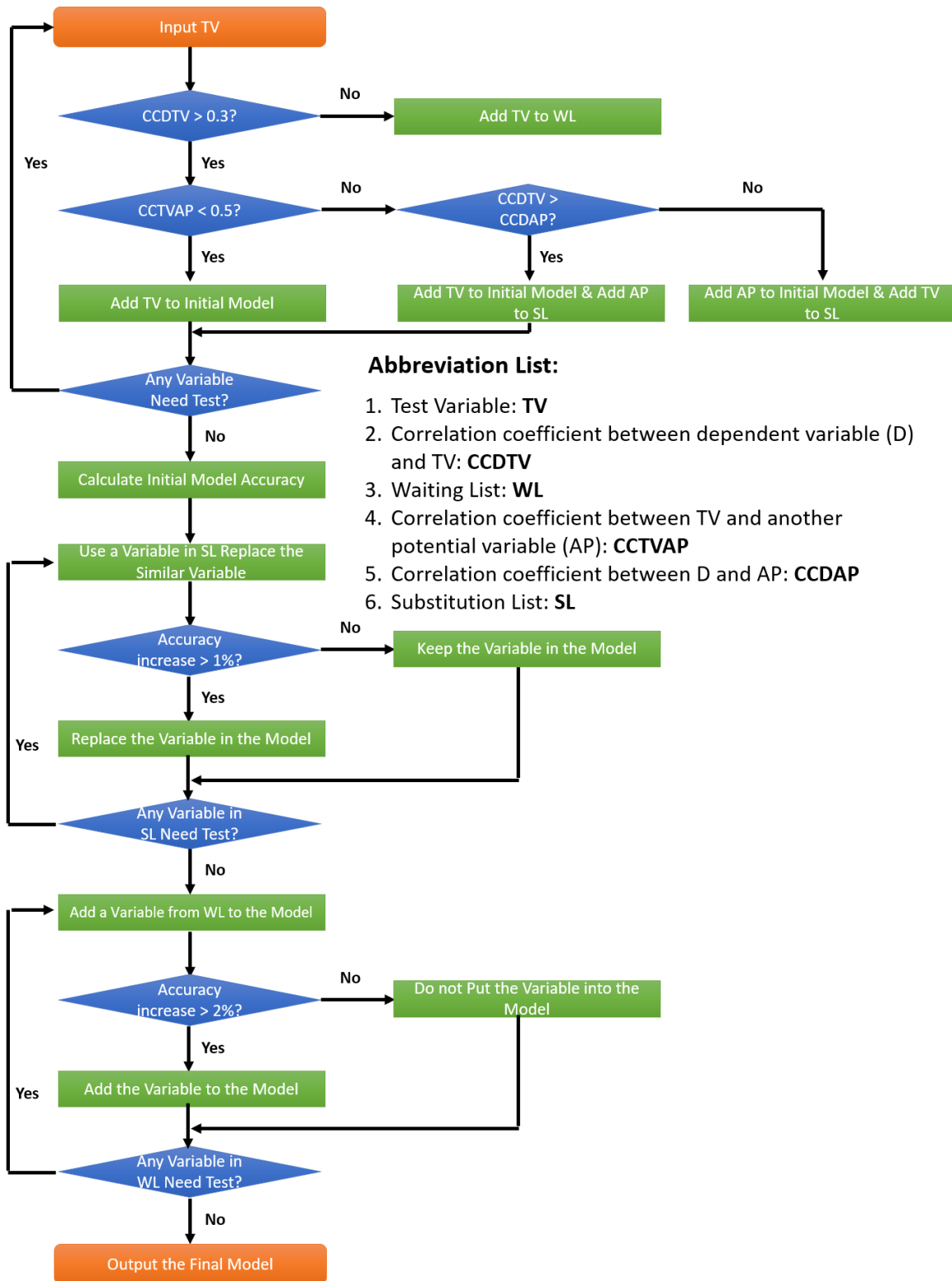
742 **Figure 1: Illustration of TrCA, ToCA, and Ground-Level Measurements**



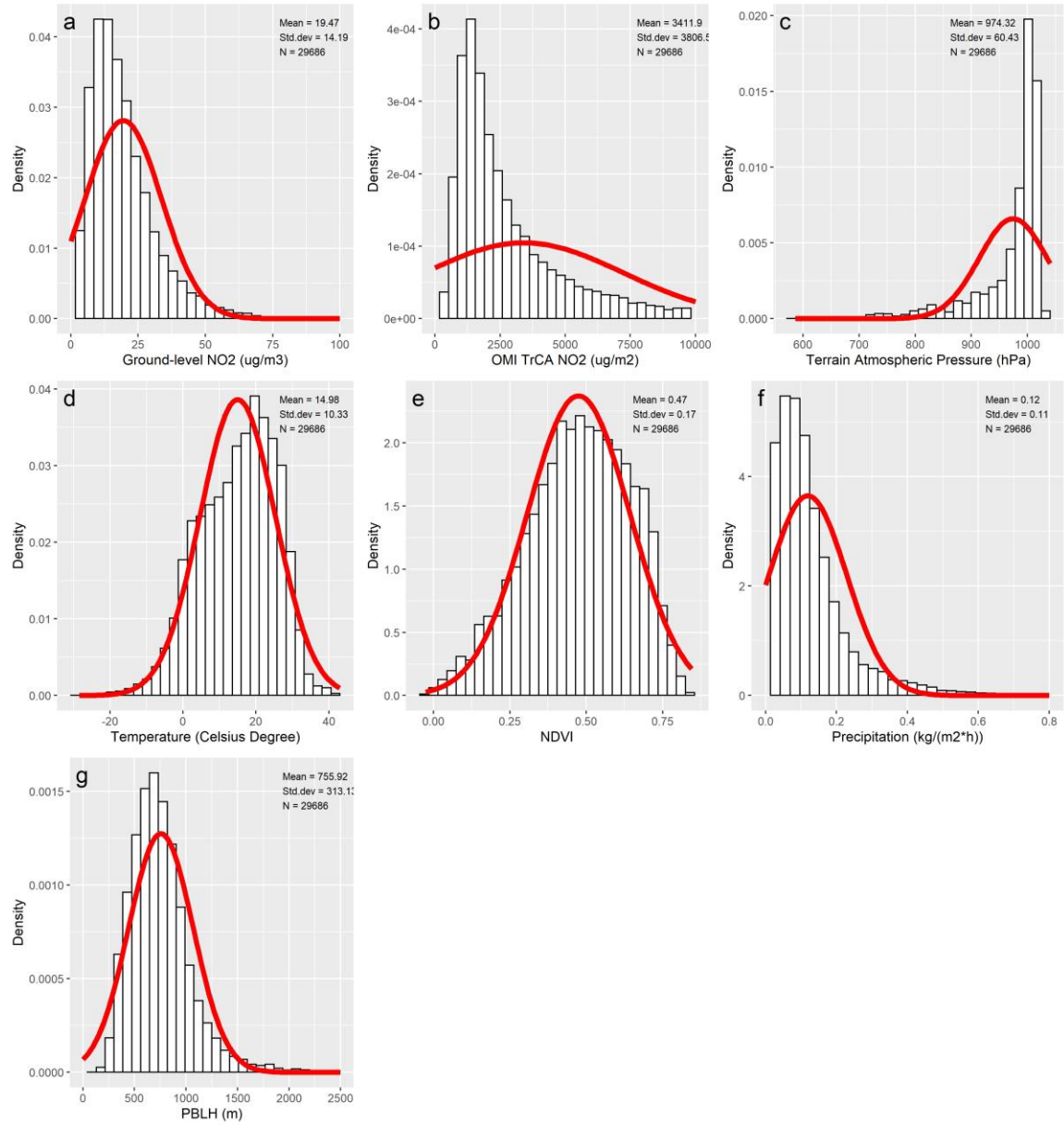
743

744

**Figure 2: Map of City Locations with the Measurement Points**

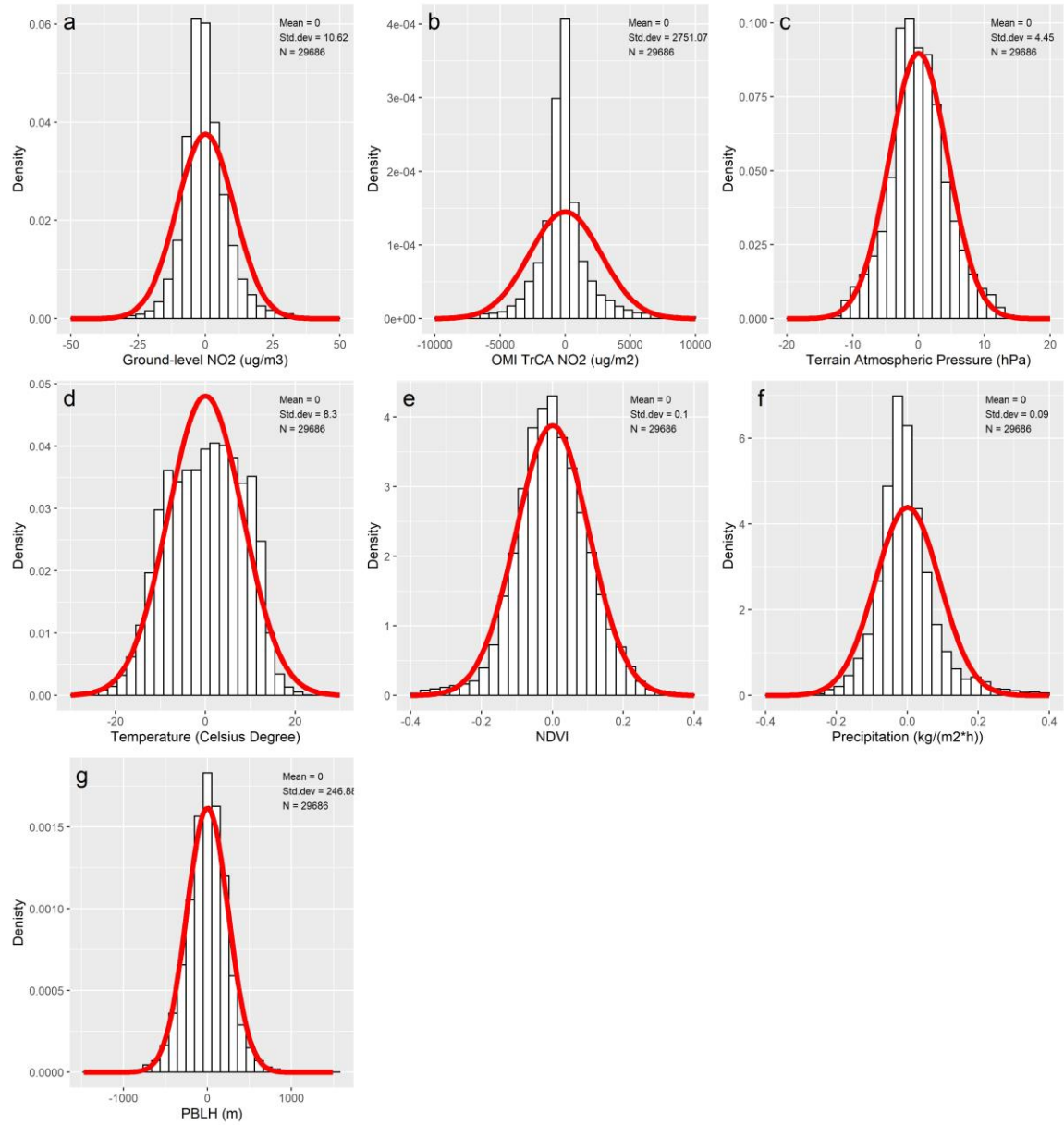


**Figure 3: Flow Chart of Variable Selection Process**



**Figure 4: Statistical Distributions and Descriptive Statistics of the Variables in the Original Dataset**

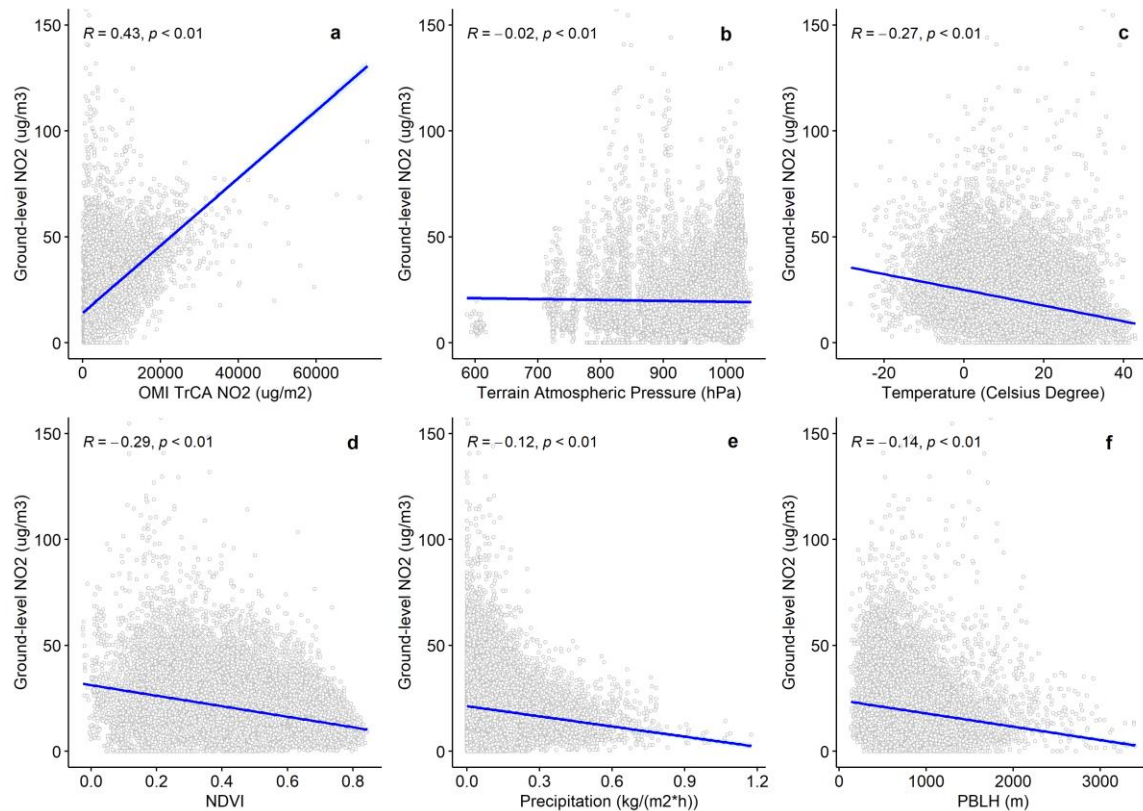
(The red line is the normal curve.)



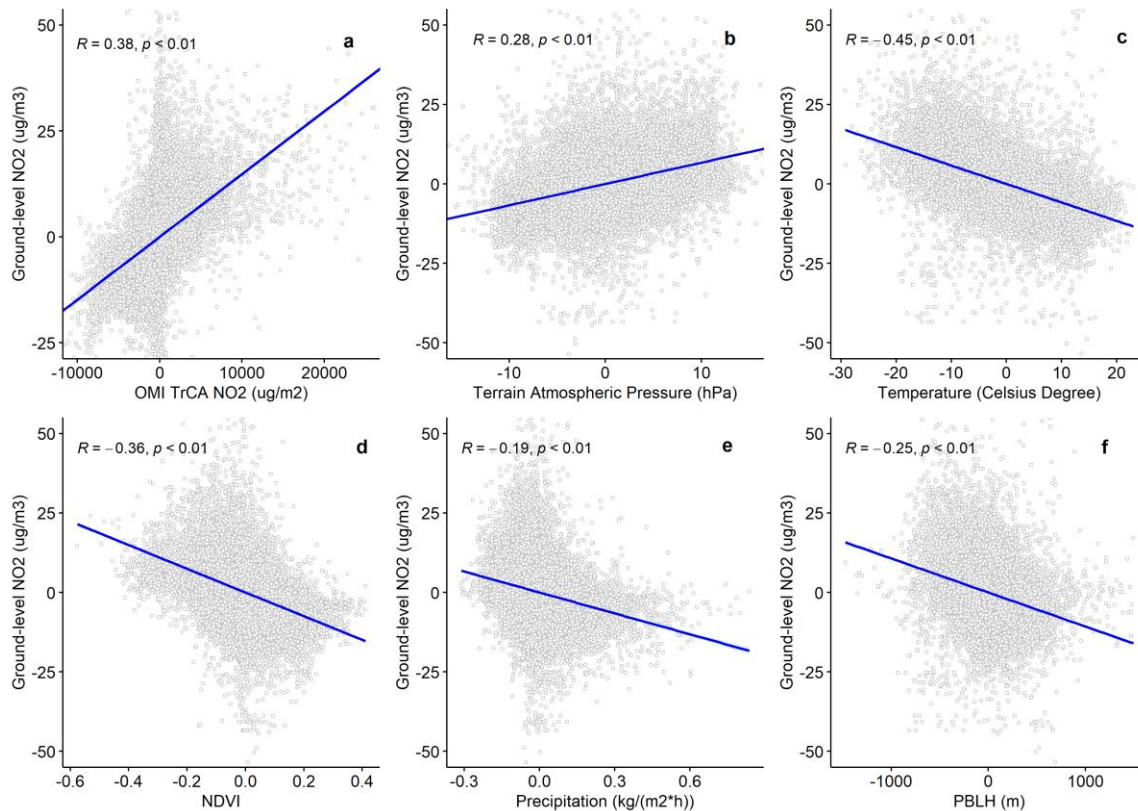
**Figure 5: Statistical Distributions and Descriptive Statistics of the Variables in the Transformed Dataset**

(The red line is the normal curve.)

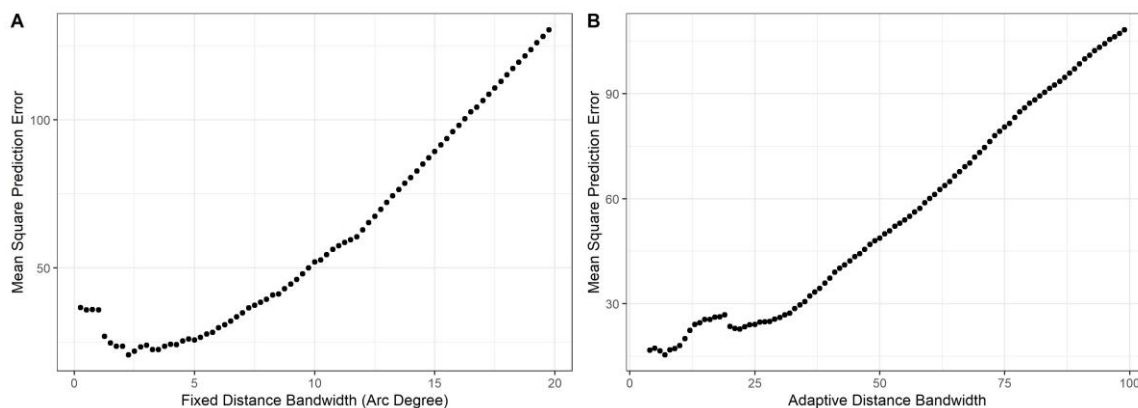




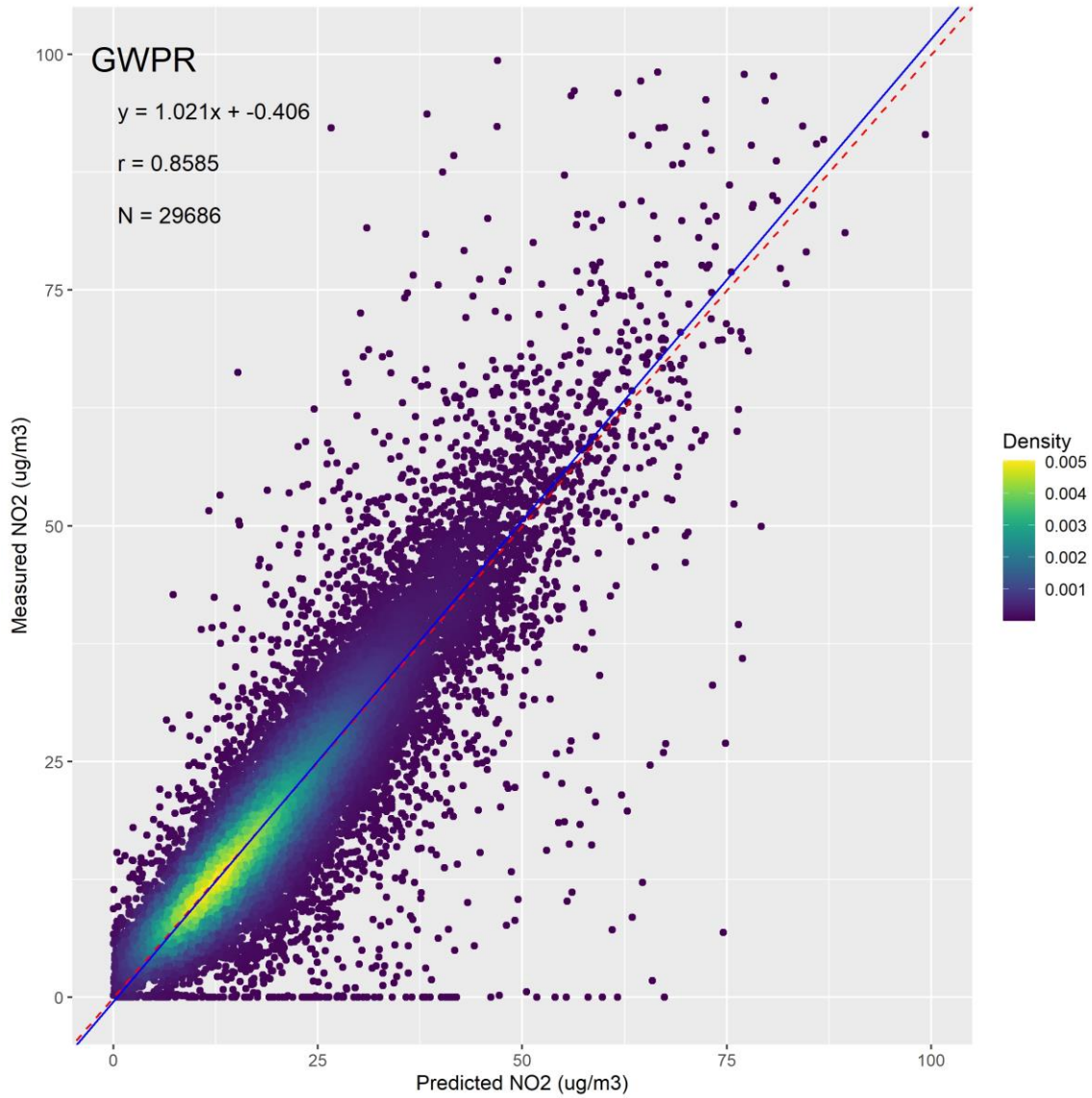
**Figure 6: Linear Trends and Correlation Coefficients between Variables in the Original Dataset**



**Figure 7: Linear Trends and Correlation Coefficients between Variables in the Transformed Dataset**

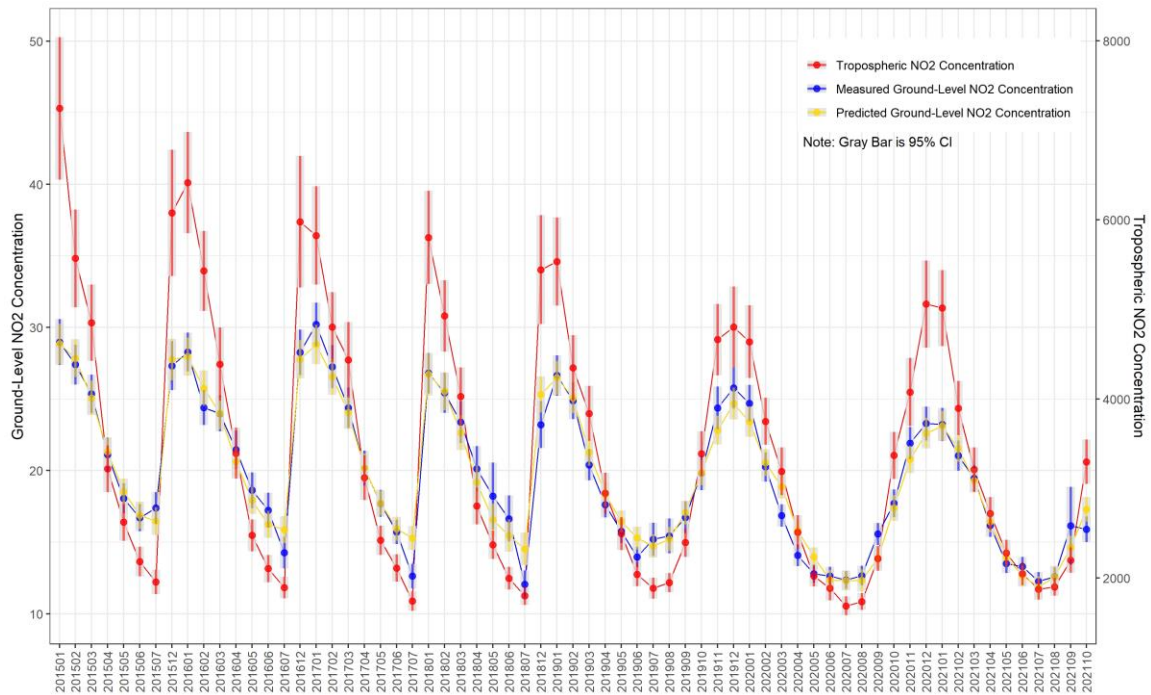


**Figure 8: Step Increment Bandwidth Selection**



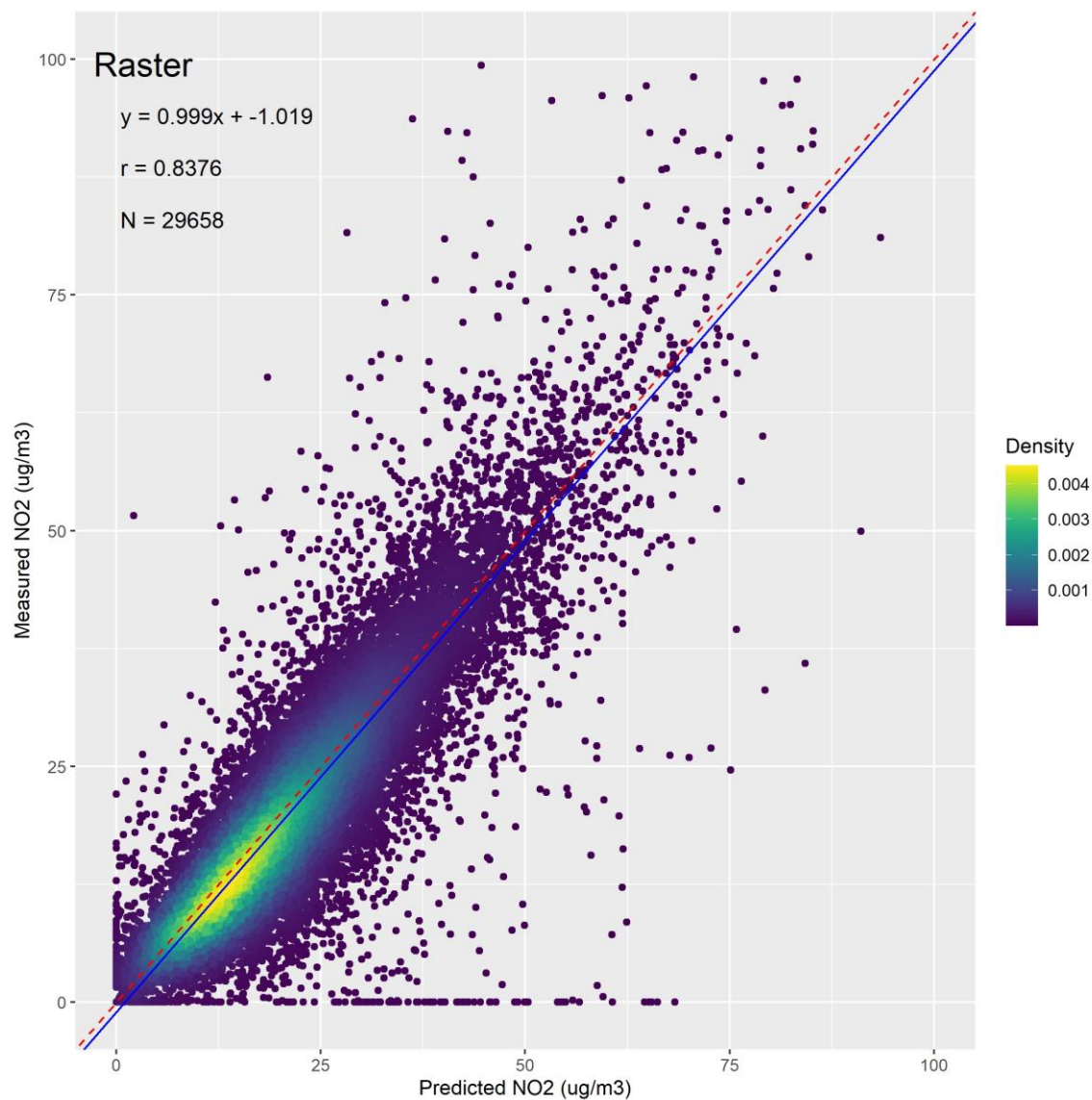
**Figure 9: The Density Plots between the Measured Ground-Level NO<sub>2</sub> Concentration and the Predicted NO<sub>2</sub> Concentration from the GWPR Results.**

(The red dashed line is the 1:1 line. The blue line is the regression line.)



**Figure 10: Time Series Comparisons of Monthly Means of Tropospheric, Measured Ground-Level, Predicted Ground-Level NO<sub>2</sub> Concentration and Their 95% Confidence Interval**  
 (Unit: tropospheric NO<sub>2</sub> concentration [ $\mu\text{g}/\text{m}^2$ ], measured and predicted Ground-level NO<sub>2</sub> concentration [ $\mu\text{g}/\text{m}^3$ ])

781



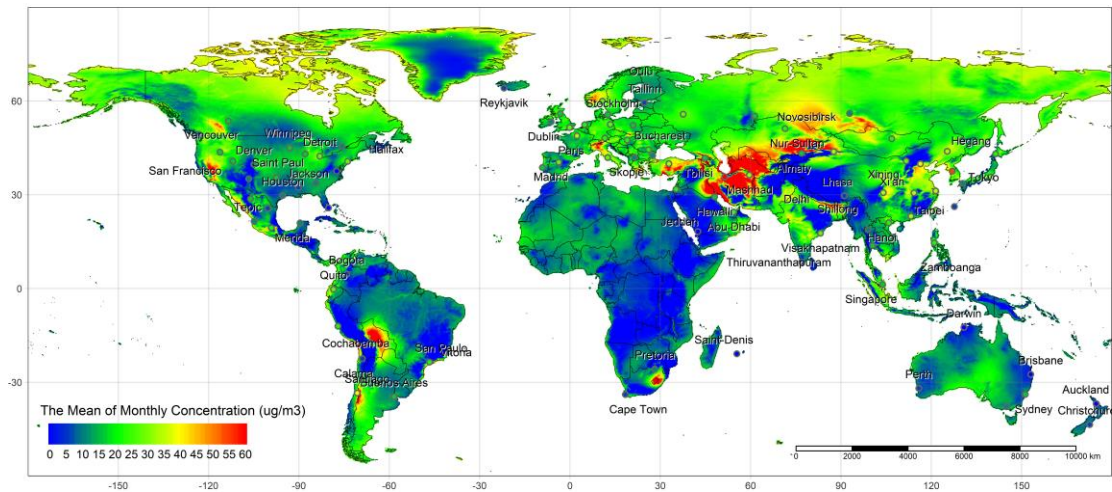
782

783 **Figure 11: The Density Plots between the Measured Ground-Level NO<sub>2</sub>**  
784 **Concentration and the Predicted NO<sub>2</sub> Concentration from the Predicted Grid**  
785 **Datasets**

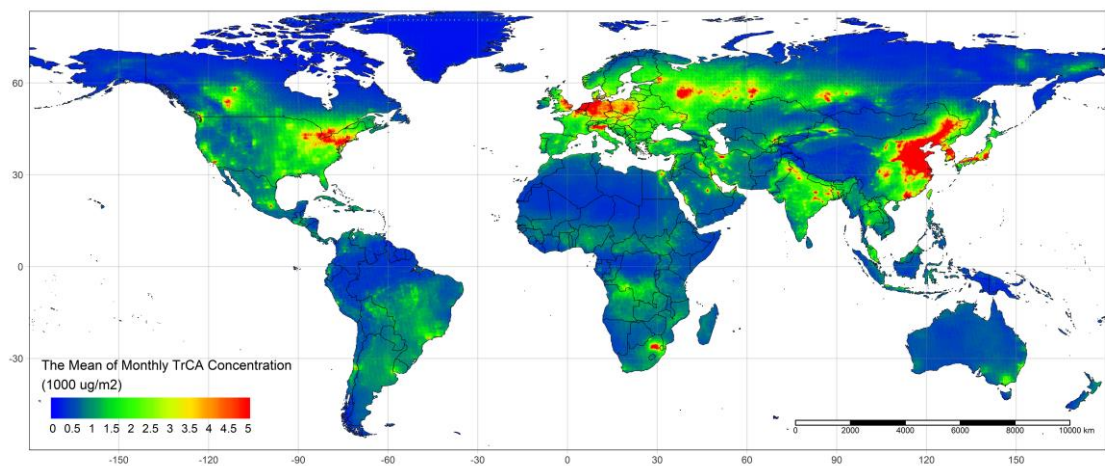
786 (The red dashed line is the 1:1 line. The blue line is the regression line.)

787

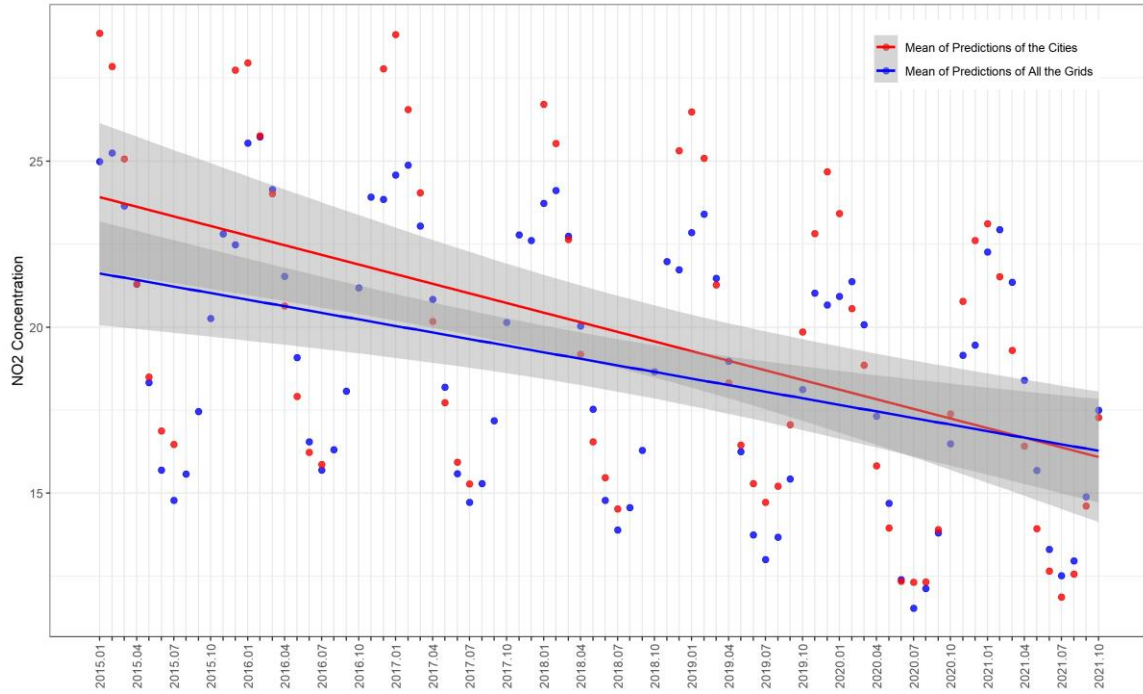




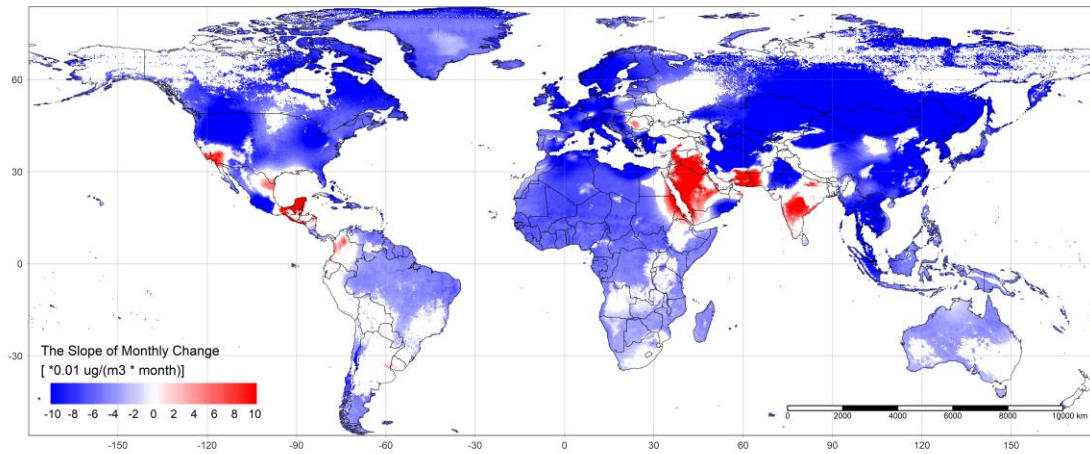
**Figure 12: Map of Average Monthly Ground-Level NO<sub>2</sub> Concentration ( $\mu\text{g}/\text{m}^3$ ),  
Jan 2015 – Oct 2021**



**Figure 13: Map of Average Monthly TrCA NO<sub>2</sub> Concentration ( $1000 \mu\text{g}/\text{m}^2$ ), Jan  
2015 – Oct 2021**



**Figure 14: Monthly Change Trends of Average Predicted Ground-Level NO<sub>2</sub> Concentration ( $\mu\text{g}/\text{m}^3$ ), Jan 2015 – Oct 2021**



**Figure 15: Map of Monthly Change Trends of Ground-Level NO<sub>2</sub> Concentration ( $0.01 \mu\text{g}/(\text{m}^3 \times \text{month})$ ), Jan 2015 – Oct 2021**

**Table 1: Yearly Statistical Indicators Summary of the GWPR Results**

Year	N	R <sup>2</sup>	RMSE ( $\mu\text{g}/\text{m}^3$ )	MAE ( $\mu\text{g}/\text{m}^3$ )	r	Slope	Intercept
2015	3115	86.46%	5.151	3.476	0.930	0.843	3.623
2016	3315	81.64%	6.126	3.617	0.904	0.794	4.508
2017	2936	85.20%	5.442	3.355	0.925	0.803	4.242
2018	3560	74.99%	8.881	4.071	0.869	0.689	6.461
2019	5790	71.34%	7.475	3.566	0.846	0.676	6.418
2020	6023	77.32%	5.317	3.335	0.881	0.827	2.912
2021	4947	51.55%	9.619	3.115	0.718	0.520	7.831
Total	29686	74.45%	7.171	3.480	0.863	0.724	5.373

Note: The Ideal value of slope is 1, and the ideal value of intercept is 0.

805

**Table 2: Statistical Indicators Summary of 10-Fold Cross Validation**

Statistical Indicator of Training								Statistical Indicator of Testing						
N	R2	RMSE $\mu\text{g}/\text{m}^3$	MAE $\mu\text{g}/\text{m}^3$	r	Slope	Intercep t		N	R2	RMSE $\mu\text{g}/\text{m}^3$	MEA $\mu\text{g}/\text{m}^3$	r	Slope	Intercep t
1	26718	89.68%	7.746	3.800	0.839	0.681	6.195	2968	89.73%	7.669	4.096	0.831	0.691	6.139
2	26718	89.47%	7.843	3.861	0.837	0.684	6.151	2968	89.90%	7.422	4.367	0.825	0.733	5.296
3	26718	89.54%	7.828	3.846	0.839	0.683	6.148	2968	91.42%	6.754	3.994	0.844	0.799	4.160
4	26718	92.53%	6.464	3.719	0.876	0.749	4.926	2968	68.55%	15.57	4.399	0.627	0.358	12.275
5	26718	89.96%	7.640	3.763	0.844	0.692	6.009	2968	91.02%	7.155	4.055	0.854	0.711	5.592
6	26718	90.04%	7.637	3.828	0.847	0.694	5.960	2968	90.13%	7.267	4.025	0.826	0.793	4.044
7	26718	89.61%	7.777	3.790	0.839	0.681	6.226	2968	92.39%	6.546	3.997	0.875	0.762	4.511
8	26718	90.55%	7.404	3.724	0.853	0.705	5.756	2968	89.39%	7.875	4.171	0.834	0.683	6.137
9	26718	90.95%	7.268	3.742	0.862	0.712	5.602	2968	89.98%	7.418	4.160	0.828	0.781	4.391
10	26712	90.27%	7.515	3.759	0.849	0.693	5.976	2974	90.76%	7.332	4.086	0.856	0.710	5.718

Note: The Ideal value of slope is 1, and the ideal value of intercept is 0. When the values of statistical indicators are close between training results and testing result, the predictions of model are reliable.

806

807



**Table 3: Statistical Indicators Summary of the Leave-One-Out Cross Validation of OK Interpolation**

	Variable	N	R <sup>2</sup>	MRSE $\mu g/m^3$	MAE $\mu g/m^3$	r	Intercept	Slope	On Point R <sup>2</sup>
Coefficient Interpolation	TrCA of NO <sub>2</sub> Terrain	530	67.54%	8.49E-04	3.81E-04	0.826	-9.31E-05	1.109	99.87%
	Atmospheric Pressure	530	77.99%	0.148	0.079	0.883	0.003	1.021	99.86%
	Temperature	530	60.11%	0.320	0.102	0.778	0.015	1.080	99.91%
	NDVI	530	79.28%	16.815	6.491	0.892	1.065	1.053	99.93%
	Precipitation	530	75.46%	14.210	5.288	0.869	0.682	1.037	99.91%
	PBLH	530	72.38%	0.004	0.002	0.852	0.001	1.054	99.83%
	2016 Year	530	72.38%	0.004	0.002	0.852	0.001	1.054	99.85%
	2017 Year	530	71.89%	3.161	1.378	0.848	0.081	1.021	99.87%
	2018 Year	530	55.09%	6.727	1.986	0.743	0.218	1.009	99.83%
	2019 Year	530	72.50%	3.642	1.555	0.852	0.150	1.028	99.87%
	2020 Year	530	75.28%	3.273	1.443	0.868	0.170	1.013	99.87%
	2021 Year	530	76.82%	3.791	1.643	0.877	0.223	1.029	99.87%
Mean Value Interpolation	Measured Ground- Level NO <sub>2</sub>	530	34.50%	7.773	5.633	0.626	4.505	0.748	98.40%
	TrCA of NO <sub>2</sub> Terrain	530	79.85%	1143.454	738.228	0.896	75.632	0.946	99.59%
	Atmospheric Pressure	530	67.70%	36.269	21.535	0.823	-17.663	1.018	99.79%
	Temperature	530	84.21%	2.597	1.564	0.919	-0.821	1.046	99.92%
	NDVI	530	54.86%	0.095	0.068	0.748	0.059	0.876	99.44%
	Precipitation	530	86.86%	0.022	0.012	0.932	-0.002	1.009	99.94%
	PBLH	530	77.98%	101.276	53.343	0.883	-18.025	1.024	99.91%

808

809

810

## 811 **Reference:**

- 812 Bechle, M.J., Millet, D.B., & Marshall, J.D. (2015). National Spatiotemporal Exposure  
813 Surface for NO<sub>2</sub>: Monthly Scaling of a Satellite-Derived Land-Use Regression, 2000–  
814 2010. *Environmental Science & Technology*, 49, 12297-12305
- 815 Beckerman, B., Jerrett, M., Brook, J.R., Verma, D.K., Arain, M.A., & Finkelstein, M.M.  
816 (2008). Correlation of nitrogen dioxide with other traffic pollutants near a major  
817 expressway. *Atmospheric Environment*, 42, 275-290
- 818 Beenstock, M., & Felsenstein, D. (2019). *The econometric analysis of non-stationary*  
819 *spatial panel data*. Springer
- 820 Bigdeli, M., Taheri, M., & Mohammadian, A. (2021). Spatial sensitivity analysis of  
821 COVID-19 infections concerning the satellite-based four air pollutants levels.  
822 *International Journal of Environmental Science and Technology*, 18, 751-760
- 823 Bivand, R.S., Pebesma, E.J., Gómez-Rubio, V., & Pebesma, E.J. (2008). *Applied spatial*  
824 *data analysis with R*. Springer
- 825 Breusch, T.S., & Pagan, A.R. (1980). The Lagrange Multiplier Test and its Applications to  
826 Model Specification in Econometrics. *The Review of Economic Studies*, 47, 239
- 827 Brunekreef, B., & Holgate, S.T. (2002). Air pollution and health. *The Lancet*, 360, 1233-  
828 1242
- 829 Brunsdon, C., Fotheringham, A.S., & Charlton, M.E. (2010). Geographically Weighted  
830 Regression: A Method for Exploring Spatial Nonstationarity. *Geographical Analysis*, 28,  
831 281-298
- 832 Brunsdon, C., Fotheringham, S., & Charlton, M. (1998). Geographically Weighted  
833 Regression. *Journal of the Royal Statistical Society: Series D (The Statistician)*, 47, 431-  
834 443
- 835 Chang, Y., Wang, S., Zhou, Y., Wang, L., & Wang, F. (2019). A Novel Method of  
836 Evaluating Highway Traffic Prosperity Based on Nighttime Light Remote Sensing. *Remote*  
837 *Sensing*, 12, 102
- 838 Chi, Y., Fan, M., Zhao, C., Sun, L., Yang, Y., Yang, X., & Tao, J. (2021). Ground-level  
839 NO<sub>2</sub> concentration estimation based on OMI tropospheric NO<sub>2</sub> and its spatiotemporal  
840 characteristics in typical regions of China. *Atmospheric Research*, 264, 105821
- 841 Chiusolo, M., Cadum, E., Stafoggia, M., Galassi, C., Berti, G., Faustini, A., Bisanti, L.,  
842 Vigotti, M.A., Dessì M.P., Cernigliaro, A., Mallone, S., Pacelli, B., Minerba, S., Simonato,  
843 L., & Forastiere, F. (2011). Short-Term Effects of Nitrogen Dioxide on Mortality and  
844 Susceptibility Factors in 10 Italian Cities: The EpiAir Study. *Environmental Health*  
845 *Perspectives*, 119, 1233-1238
- 846 Croissant, Y., & Millo, G. (2008). Panel Data Econometrics in R: The plm Package.  
847 *Journal of Statistical Software*, 27
- 848 Curier, R.L., Kranenburg, R., Segers, A.J.S., Timmermans, R.M.A., & Schaap, M. (2014).  
849 Synergistic use of OMI NO<sub>2</sub> tropospheric columns and LOTOS–EUROS to evaluate the  
850 NO<sub>x</sub> emission trends across Europe. *Remote Sensing of Environment*, 149, 58-69
- 851 Di, Q., Amini, H., Shi, L., Kloog, I., Silvern, R., Kelly, J., Sabath, M.B., Choirat, C.,  
852 Koutrakis, P., Lyapustin, A., Wang, Y., Mickley, L.J., & Schwartz, J. (2020). Assessing  
853 NO<sub>2</sub> Concentration and Model Uncertainty with High Spatiotemporal Resolution across  
854 the Contiguous United States Using Ensemble Model Averaging. *Environmental Science*  
855 *& Technology*, 54, 1372-1384

- Fan, H., Zhao, C., & Yang, Y. (2020). A comprehensive analysis of the spatio-temporal variation of urban air pollution in China during 2014–2018. *Atmospheric Environment*, 220, 117066
- Fotheringham, A., Brunson, C., & Charlton, M. (2002). *Geographically Weighted Regression: The Analysis of Spatially Varying Relationships*. John Wiley & Sons
- Fotheringham, A.S., Crespo, R., & Yao, J. (2015). Geographical and Temporal Weighted Regression (GTWR). *Geographical Analysis*, 47, 431-452
- Fotheringham, A.S., & Oshan, T.M. (2016). Geographically weighted regression and multicollinearity: dispelling the myth. *Journal of Geographical Systems*, 18, 303-329
- Geddes, J.A., Martin, R.V., Boys, B.L., & Van Donkelaar, A. (2016). Long-Term Trends Worldwide in Ambient NO<sub>2</sub> Concentrations Inferred from Satellite Observations. *Environmental Health Perspectives*, 124, 281-289
- Gollini, I., Lu, B., Charlton, M., Brunson, C., & Harris, P. (2015). GWmodel: An R package for exploring spatial heterogeneity using geographically weighted models. *Journal of Statistical Software*, 63
- Hamra, G.B., Laden, F., Cohen, A.J., Raaschou-Nielsen, O., Brauer, M., & Loomis, D. (2015). Lung Cancer and Exposure to Nitrogen Dioxide and Traffic: A Systematic Review and Meta-Analysis. *Environmental Health Perspectives*, 123, 1107-1112
- Hu, X., Waller, L.A., Al-Hamdan, M.Z., Crosson, W.L., Estes, M.G., Estes, S.M., Quattrochi, D.A., Sarnat, J.A., & Liu, Y. (2013). Estimating ground-level PM<sub>2.5</sub> concentrations in the southeastern U.S. using geographically weighted regression. *Environmental Research*, 121, 1-10
- Irie, H., Boersma, K.F., Kanaya, Y., Takashima, H., Pan, X., & Wang, Z.F. (2012). Quantitative bias estimates for tropospheric NO<sub>2</sub> columns retrieved from SCIAMACHY, OMI, and GOME-2 using a common standard for East Asia. *Atmospheric Measurement Techniques*, 5, 2403-2411
- Jiang, M., Sun, W., Yang, G., & Zhang, D. (2017). Modelling Seasonal GWR of Daily PM<sub>2.5</sub> with Proper Auxiliary Variables for the Yangtze River Delta. *Remote Sensing*, 9, 346
- Kang, S. (1985). A note on the equivalence of specification tests in the two-factor multivariate variance components model. *Journal of Econometrics*, 28, 193-203
- Kasparoglu, S., Incecik, S., & Topcu, S. (2018). Spatial and temporal variation of O<sub>3</sub>, NO and NO<sub>2</sub> concentrations at rural and urban sites in Marmara Region of Turkey. *Atmospheric Pollution Research*, 9, 1009-1020
- Kim, M., Brunner, D., & Kuhlmann, G. (2021). Importance of satellite observations for high-resolution mapping of near-surface NO<sub>2</sub> by machine learning. *Remote Sensing of Environment*, 264, 112573
- Larkin, A., Geddes, J.A., Martin, R.V., Xiao, Q., Liu, Y., Marshall, J.D., Brauer, M., & Hystad, P. (2017). Global Land Use Regression Model for Nitrogen Dioxide Air Pollution. *Environmental Science & Technology*, 51, 6957-6964
- Lelieveld, J., Evans, J.S., Fnais, M., Giannadaki, D., & Pozzer, A. (2015). The contribution of outdoor air pollution sources to premature mortality on a global scale. *Nature*, 525, 367-371
- Li, C., & Managi, S. (2021a). Contribution of on-road transportation to PM<sub>2.5</sub>. *Scientific Reports*, 11, 21320

901 Li, C., & Managi, S. (2021b). Spatial Variability of the Relationship between Air Pollution  
 902 and Well-being. *Sustainable Cities and Society*, 103447  
 903 Li, C., & Managi, S. (2022). Impacts of air pollution on COVID-19 case fatality rate: a  
 904 global analysis. *Environmental Science and Pollution Research*, Jan 4:1–14.  
 905 Li, J., & Heap, A.D. (2011). A review of comparative studies of spatial interpolation  
 906 methods in environmental sciences: Performance and impact factors. *Ecological  
 907 Informatics*, 6, 228-241  
 908 Li, L., & Wu, J. (2021). Spatiotemporal estimation of satellite-borne and ground-level NO<sub>2</sub>  
 909 using full residual deep networks. *Remote Sensing of Environment*, 254, 112257  
 910 Li, T., Wang, Y., & Yuan, Q. (2020). Remote Sensing Estimation of Regional NO<sub>2</sub> via  
 911 Space-Time Neural Networks. *Remote Sensing*, 12, 2514  
 912 Lin, C.-A., Chen, Y.-C., Liu, C.-Y., Chen, W.-T., Seinfeld, J.H., & Chou, C.C.K. (2019).  
 913 Satellite-Derived Correlation of SO<sub>2</sub>, NO<sub>2</sub>, and Aerosol Optical Depth with  
 914 Meteorological Conditions over East Asia from 2005 to 2015. *Remote Sensing*, 11, 1738  
 915 Liu, L., Zhang, X., Xu, W., Liu, X., Lu, X., Chen, D., Zhang, X., Wang, S., & Zhang, W.  
 916 (2017). Estimation of monthly bulk nitrate deposition in China based on satellite NO<sub>2</sub>  
 917 measurement by the Ozone Monitoring Instrument. *Remote Sensing of Environment*, 199,  
 918 93-106  
 919 Liu, Y., Ming, T., Peng, C., Wu, Y., Li, W., De Richter, R., & Zhou, N. (2021). Mitigating  
 920 air pollution strategies based on solar chimneys. *Solar Energy*, 218, 11-27  
 921 Mackerron, G., & Mourato, S. (2009). Life satisfaction and air quality in London.  
 922 *Ecological Economics*, 68, 1441-1453  
 923 Meng, L., Liu, J., Tarasick David, W., Randel William, J., Steiner Andrea, K., Wilhelmsen,  
 924 H., Wang, L., & Haimberger, L. (2021). Continuous rise of the tropopause in the Northern  
 925 Hemisphere over 1980–2020. *Science Advances*, 7, eabi8065  
 926 Newell, K., Kartsonaki, C., Lam, K.B.H., & Kurmi, O.P. (2017). Cardiorespiratory health  
 927 effects of particulate ambient air pollution exposure in low-income and middle-income  
 928 countries: a systematic review and meta-analysis. *The Lancet Planetary Health*, 1, e368-  
 929 e380  
 930 Nickolay, K., Lok, L., Sergey, M., Edward, C., Eric, B., William, S., Joanna, J., & the OMI  
 931 core team (2019). OMI/Aura NO<sub>2</sub> Total and Tropospheric Column Daily L2 Global  
 932 Gridded 0.25 degree x 0.25 degree V3. In N.G.S.F. Center (Ed.): Goddard Earth Sciences  
 933 Data and Information Services Center (GES DISC)  
 934 Ogen, Y. (2020). Assessing nitrogen dioxide (NO<sub>2</sub>) levels as a contributing factor to  
 935 coronavirus (COVID-19) fatality. *Science of The Total Environment*, 726, 138605  
 936 OMI Team (2012). Ozone Monitoring Instrument (OMI) data user's guide. In  
 937 Orellano, P., Reynoso, J., Quaranta, N., Bardach, A., & Ciapponi, A. (2020). Short-term  
 938 exposure to particulate matter (PM<sub>10</sub> and PM<sub>2.5</sub>), nitrogen dioxide (NO<sub>2</sub>), and ozone (O<sub>3</sub>)  
 939 and all-cause and cause-specific mortality: Systematic review and meta-analysis.  
 940 *Environment International*, 142, 105876  
 941 Pebesma, E.J. (2004). Multivariable geostatistics in S: the gstat package. *Computers &  
 942 Geosciences*, 30, 683-691  
 943 Qin, K., Rao, L., Xu, J., Bai, Y., Zou, J., Hao, N., Li, S., & Yu, C. (2017). Estimating  
 944 Ground Level NO<sub>2</sub> Concentrations over Central-Eastern China Using a Satellite-Based  
 945 Geographically and Temporally Weighted Regression Model. *Remote Sensing*, 9, 950

Rice, M.B., Ljungman, P.L., Wilker, E.H., Gold, D.R., Schwartz, J.D., Koutrakis, P., Washko, G.R., O'Connor, G.T., & Mittleman, M.A. (2013). Short-Term Exposure to Air Pollution and Lung Function in the Framingham Heart Study. *American Journal of Respiratory and Critical Care Medicine*, 188, 1351-1357

Schoeberl, M.R., Douglass, A.R., Hilsenrath, E., Bhartia, P.K., Beer, R., Waters, J.W., Gunson, M.R., Froidevaux, L., Gille, J.C., Barnett, J.J., Levelt, P.F., & Decola, P. (2006). Overview of the EOS aura mission. *IEEE Transactions on Geoscience and Remote Sensing*, 44, 1066-1074

Shen, Y., Jiang, F., Feng, S., Zheng, Y., Cai, Z., & Lyu, X. (2021). Impact of weather and emission changes on NO<sub>2</sub> concentrations in China during 2014–2019. *Environmental Pollution*, 269, 116163

Taylor, R. (1990). Interpretation of the Correlation Coefficient: A Basic Review. *Journal of Diagnostic Medical Sonography*, 6, 35-39

Wang, B., & Chen, Z. (2013). An intercomparison of satellite-derived ground-level NO<sub>2</sub> concentrations with GMSMB modeling results and in-situ measurements – A North American study. *Environmental Pollution*, 181, 172-181

Wu, X., Nethery, R.C., Sabath, M.B., Braun, D., & Dominici, F. (2020). Air pollution and COVID-19 mortality in the United States: Strengths and limitations of an ecological regression analysis. *Science Advances*, 6, eabd4049

Xiang, Y., Zhang, T., Liu, J., Lv, L., Dong, Y., & Chen, Z. (2019). Atmosphere boundary layer height and its effect on air pollutants in Beijing during winter heavy pollution. *Atmospheric Research*, 215, 305-316

Yao, Y., Pan, J., Liu, Z., Meng, X., Wang, W., Kan, H., & Wang, W. (2021). Ambient nitrogen dioxide pollution and spreadability of COVID-19 in Chinese cities. *Ecotoxicology and Environmental Safety*, 208, 111421

Yousefian, F., Faridi, S., Azimi, F., Aghaei, M., Shamsipour, M., Yaghmaeian, K., & Hassanvand, M.S. (2020). Temporal variations of ambient air pollutants and meteorological influences on their concentrations in Tehran during 2012–2017. *Scientific Reports*, 10, 292

Yu, D., Zhang, Y., Wu, X., Li, D., & Li, G. (2021). The varying effects of accessing high-speed rail system on China's county development: A geographically weighted panel regression analysis. *Land Use Policy*, 100, 104935

Zhao, Zhou, Li, Cao, He, Yu, Li, Elvidge, Cheng, & Zhou (2019). Applications of Satellite Remote Sensing of Nighttime Light Observations: Advances, Challenges, and Perspectives. *Remote Sensing*, 11, 1971

Zheng, Yang, Wu, & Marinello (2019). Spatial Variation of NO<sub>2</sub> and Its Impact Factors in China: An Application of Sentinel-5P Products. *Remote Sensing*, 11, 1939