

Estimating Monthly Global Ground-Level NO₂ Concentrations Using Geographically Weighted Panel Regression

Abstract

Predicting long-term ground-level nitrogen dioxide (NO₂) 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), using Aura Satellite, detects monthly global tropospheric column amounts (TrCA) of NO₂ molecules. However, the relationship between the ground-level NO₂ concentration and TrCA of NO₂ molecules remains elusive because NO₂ molecules in the air are not vertically evenly distributed. Here, we examine the relationship among satellite-derived data, measured ground-level NO₂ concentration, and several controlling meteorological variables from January 2015 to October 2021. The geographically weighted panel regression (GWPR) is built and applied. The accuracy of raster prediction is 69.61%. The coefficient of correlation between predicted and measure value is 0.8376. The root mean square error and mean absolute error are 7.84 and 4.07 $\mu\text{g}/\text{m}^3$, respectively. Moreover, the GWPR is reliable, indicated by the 10-fold cross-validation. The GWPR can analyze unbalanced panel data and capture the spatial variability of the relationship. Based on the GWPR estimation, the 82 monthly global ground-level NO₂ concentrations are predicted from January 2015 to October 2021. 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 NO₂ Concentration; GWPR model; spatial non-stationarity; OMI

Introduction

Air pollutant, nitrogen dioxide (NO₂), 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 NO₂ concentration is positive (Brunekreef and Holgate 2002). Specifically, a 10 $\mu\text{g}/\text{m}^3$ increase in NO₂ concentration in 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 NO₂ should be lower than 40 $\mu\text{g}/\text{m}^3$. Additionally, other air pollutants, such as particulate matter and ozone, are related to approximately 3.3 million premature deaths per year globally, and NO₂ is one of the critical factors attributable to these pollutants (Beckerman et al. 2008; Lelieveld et al. 2015). Evidence shows a link between NO₂ 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 indicate that in areas with high NO₂ concentration, people are

more easily infected by the coronavirus and more difficult to recover, controlling for several other crucial factors, including age, gender, restriction policy, among others (Li and Managi 2022; Ogen 2020; Wu et al. 2020; Yao et al. 2021). Besides, air pollution affects human well-being directly (Li and Managi 2021b). For example, an increase of $10 \mu g/m^3$ in annual average NO_2 is related to a roughly 50% decrease of life satisfaction in London (Mackerron and Mourato 2009). The ground-level NO_2 concentration affects human well-being and public health the most (Lelieveld et al. 2015; Li and Managi 2021b, 2022). However, the satellite instrument, the Ozone Monitoring Instrument (OMI), focuses mainly on the tropospheric column amounts (TrCA) and the total column amounts (ToCA) of NO_2 (**Figure 1**) (OMI Team 2012). The association ground-level NO_2 concentration with TrCA or ToCA remains unclear because the NO_2 molecules are not vertically evenly distributed in the air. Besides, there is a lack of ground-level NO_2 concentration measurements in low-income countries, which induces broad uncertainty of ground-level NO_2 distribution (Larkin et al. 2017). In order to help to detect the adverse impacts of NO_2 on public health and human well-being, the global time-series ground-level NO_2 data is strongly desired.

Figure 1 is located here.

Geographically weighted regression (GWR) is an advanced method that estimates air pollution 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 are developed, such as land-use regression (Bechle et al. 2015; Larkin et al. 2017), chemical transport model (Geddes et al. 2016), linear regression model (Lin et al. 2019; Liu et al. 2017), GIS-based multi-source and multi-box model

(Wang and Chen 2013), geographically and temporally weighted regression (Li et al. 2020; Qin et al. 2017), among others (Kim et al. 2021). All these models show relatively high accuracy on the estimations. In order to further improve the accuracy, we fully take 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 data set, while the statistically local models are based on partial data set taking each observing subject as the center (Beenstock and Felsenstein 2019).

Currently, the GWR can only analyze the cross-sectional data (data collected by observing many subjects at one period of time) or the 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, GTWR (Fotheringham et al. 2015; Qin et al. 2017) and geographically weighted panel regression (GWPR) (Yu et al. 2021), to regress the balanced panel data (data collected by observing many subjects at all periods during the study period). GTWR is a spatially and temporally non-stationary model. Integrating spatial and temporal distance into spatiotemporal distance poses fresh unknowns, though Euclidean distance is still a powerful tool (Fotheringham et al. 2015; Li and Managi 2021a; Qin et al. 2017). Furthermore, the recent GTWR requires that the input data set should be a balanced panel data set (Gollini et al. 2015). Yu et al.'s GWPR improves GWR by using the least squares

dummy variable estimator, but models with other effects are under development (Yu et al. 2021). This method works when the panel data have only a handful of periods. Moreover, neither of these two methods can solve the problem with unbalanced panel data (except balanced panel data, all panel data are unbalanced) due to the issues in the spatially weighted matrix. However, this limitation is induced by the current toolboxes in most statistical software, rather than the GWPR algorithm. In fact, most observed data sets are unbalanced, including our data set. To analyze the panel data by GWPR, we build a new R package, “GWPR.light”, and release it on CRAN (The Comprehensive R Archive Network).

In this paper, firstly, according to a series of statistical tests and model comparisons, we choose the GWPR with fixed time-fixed effects. This GWPR model excels at accuracy. Secondly, 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 ground-level and TrCA of NO₂. Thirdly, we calibrate the optimal bandwidth, critical to building the spatially weighted matrix. Fourthly, we estimate the spatially non-stationary coefficients by the GWPR model and convert the coefficients into raster data through the interpolation method. Finally, this study provides the monthly global ground-level NO₂ concentration from January 2015 to October 2021.

Materials and Methodology

Materials

Satellite-Derived NO₂ Estimation

The satellite instrument, OMI, provides vertical column amounts of NO₂. The OMI/Aura NO₂ Total and Tropospheric Column Daily L2 Global Gridded 0.25 degree × 0.25 degree V3 (OMNO2G) is the global daily grids data set with a 0.25-degree resolution to gauge TrCA and ToCA of NO₂ (**Figure 1**) 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). This satellite was launched on July 15th 2004, and began providing daily global TrCA, stratospheric column amounts and ToCA of NO₂ observation (unit: *molecules/cm²*) from October 1st 2004 (Nickolay et al. 2019; OMI Team 2012). The OMI-NO₂ data sets, especially the 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 the most harmfully (Hamra et al. 2015; Rice et al. 2013). However, the ground-level NO₂ concentration is unavailable in many areas, particularly in rural areas and most developing countries. Therefore, the conversion or the estimation from TrCA or ToCA to ground-level data becomes a critical point. The EOS Aura orbit is a sun-synchronous polar orbit at an altitude of 705 km, and the satellite goes over each location between the local time 12:00 and 15:00 (OMI Team 2012; Schoeberl et al. 2006). Due to the orbit altitude, the ToCA of NO₂ covers the NO₂ molecules in the troposphere, stratosphere, mesosphere, and thermosphere, while TrCA of NO₂ takes those in only the troposphere into account. Because the troposphere is most affected by human activities and

air pollution, there has the most substantial impact on human health (Lelieveld et al. 2015; Li and Managi 2022) and well-being (Li and Managi 2021b), the TrCA of NO₂ is an ideal independent variable to predict the ground-level NO₂ concentration. However, owing to anthropogenic climate change, tropopause height is slowly rising (Meng et al. 2021). Though this continuous rise is relatively tiny in comparison to the tropopause height, whether it influences the estimation of the TrCA of NO₂ remains elusive. In order to choose the satellite NO₂ variable, we put both the ToCA and TrCA of NO₂ in the analyses as the primary variable, respectively. Both the goodness of fit of the model taking the TrCA of NO₂ as the dependent variable (R^2 : 0.7445) and cross validation results are slightly better than the other model using the ToCA of NO₂ (R^2 : 0.7439). Therefore, the TrCA of NO₂ is employed in the final analysis.

We convert the unit of the TrCA of NO₂ from *molecules/cm²* to microgram 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 NO₂ 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/cm²*, N_A is the Avogadro constant ($6.022,140,857 \cdot 10^{23} \text{ mol}^{-1}$), and MM_{NO_2} is the molar mass of NO₂ (46.006,6 *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 (Schoeberl et al. 2006). In other words, each location could at most be observed twice in a month rather than every day. Furthermore, other auxiliary satellite data, such as land surface temperature, NDVI, among others, are

obtained from EOS Terra and Aqua satellites, whose satellite revisit periods are also 16 days. For these reasons, we convert the daily TrCA of NO₂ from the OMNO2G data set into the monthly average TrCA of NO₂.

Ground-Level NO₂ Measurements

This data set 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 data set included 539 cities, but nine were dropped, owing to its no more than five monthly records from January 2015 to October 2021. This study exploits the monthly average NO₂ 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 data set records the medians of ground-level NO₂ in view of the data from different measurement points in a specific city. The ground-level measured NO₂ concentration data are daily. To connect them with the satellite data, we convert them into monthly average values.

Figure 2 is located here.

The unit of monthly average ground-level measured NO₂ concentration data is PPB, while the units the monthly average TrCA of NO₂ is microgram per square meter ($\mu g/m^2$). To make the units of these two data sets consistent, we convert the unit of monthly average ground-level measured NO₂ concentration data from PPB to $\mu g/m^3$ by 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 NO₂ concentration, whose unit is $\mu g/m^3$, $RGN02$ is the NO₂ concentration before conversion, whose unit is PPB, P is the terrain atmospheric pressure (mmHg), which could 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 0.750,06 $mmHg/hPa$), MW_{NO2} is 46.0066, 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.

Auxiliary Data

Auxiliary data, including terrain atmospheric pressure, temperature, NDVI, precipitation, and PBLH, are taken as the control variables to improve the model's performance. A considerable amount of previous studies involves terrain atmospheric pressure in the analyses of NO₂ concentration (Chi et al. 2021; Li et al. 2020). Terrain atmospheric pressure data are from OMNO2G, with a 0.25-arc-degree resolution. Although the temporal resolution of this data set is daily, it has the same problem as the

TrCA of NO₂ data set, a long satellite revisit period of Aura satellite. So, we convert it into a monthly average value. Additionally, to make the calculation understandable, the unit of land surface temperature has been converted into Celsius.

The temperature data are provided by the Moderate Resolution Imaging Spectroradiometer (MODIS, <https://modis.gsfc.nasa.gov/>). The products, MOD11C3 and MYD11C3, are based on the MODIS Terra and Aqua satellites' observations and measurements, including the day-time and nighttime average temperature, with a 0.05-arc-degree spatial resolution (<https://modis.gsfc.nasa.gov/data/dataproduct/mod11.php>). Due to the 16-day revisit period, there are one or two global raster data based on each satellite's observation every month. In a month, we take the mean of all the raster data, involving the day-time and nighttime average temperature in a month, to obtain the monthly average temperature raster. Of note, the spatial resolution of the monthly average temperature raster is still 0.05 arc degree. To make it consistent with the monthly average TrCA of NO₂, we reduce the resolution to 0.25 arc degree, by applying the average value.

The monthly NDVI data are acquired from the MODIS products, MOD13C2 and 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 become 0.25 arc degree. The NDVI value ranges from -1 to 1. Because the ground-level NO₂ data is from cities, the situation in rural areas is not carefully gauged. To enable the model to predict the environment in rural areas, we incorporated 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 could depict landscape, but land cover data is discrete and yearly, not applicable for this analysis. Therefore, we employ the NDVI to

slash the influence of the difference between urban and rural areas. Additionally, to represent the human activity density, we also put the nighttime light data into the 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). The temporal and spatial resolutions are one month and 0.25 arc degree, respectively. This data set 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 data set twice, that is, the original GLDAS_NOAH025_M data set boundary extends roughly 0.5 arc degree into the seas and oceans. The original unit of this data set is $kg/(m^2 \cdot s)$, but the most values are less than $0.0001 kg/(m^2 \cdot s)$. To improve the readability and robustness of the calculation, we convert its unit to $kg/(m^2 \cdot h)$.

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). The spatial resolution of this data set is $0.5^\circ \times 0.625^\circ$. We employ the bilinear interpolation method to impute a new data set with a 0.25-arc-degree resolution. The unit of PBLH is meter (m).

In the model design stage, we consider many other variables, including humidity, nighttime light, wind speed, cloud pressure, aerosol optical depth, ultraviolet aerosol index,

among others. However, none of these variables are effective due to either their poor improvement in the prediction accuracy or 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 so strong, only 0.15. Therefore, we keep the precipitation and temperature in the analysis. Moreover, the results of GWPR are spatial point data frames. To secure the coefficient raster data for the prediction, we have to employ interpolation methods, which would apparently cause some slight errors. If the benefit from 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 NO₂, terrain atmospheric pressure, temperature, NDVI, precipitation, and PBLH as independent variables.

Descriptive Statistics

We establish an unbalanced dataset based on the above data sources for modeling and prediction with time stamps and geographical locations from January 2015 to October 2021. The number of complete observations in this data set is 29,686. The unbalance of data expresses in both temporal and spatial extent. Every city has at least five-month, and at most 65-month 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 degree, while the largest distance is 24.710 arc degrees. As shown in **Figure 2**, most cities with the data are located in Northern Hemisphere, while only a few

of them distribute in South America and Africa, which drives increased difficulty of the interpolation.

Figure 3 illustrates and summarizes the statistical distributions and descriptive statistics of the variables. The concentration of ground-level NO₂ has a mean value of 19.47 $\mu\text{g}/\text{m}^3$ and a standard deviation (SD) of 14.19 $\mu\text{g}/\text{m}^3$. Its 1st quantile, median, 3rd quantile are 10.56 $\mu\text{g}/\text{m}^3$, 16.59 $\mu\text{g}/\text{m}^3$, and 24.81 $\mu\text{g}/\text{m}^3$, respectively. Its statistical distribution is one peak and right-skewed because the skewness is 7.17. The 1st quantile, median, 3rd quantile of the monthly TrCA of NO₂ are 1347.52 $\mu\text{g}/\text{m}^2$, 2139.54 $\mu\text{g}/\text{m}^2$, and 3955.21 $\mu\text{g}/\text{m}^2$, with a mean value of 3411.90 $\mu\text{g}/\text{m}^2$ and an SD of 3806.50 $\mu\text{g}/\text{m}^2$. Due to the 4.17 skewness, the statistical distribution of the monthly TrCA of NO₂ is one right-skewed peak. The monthly terrain atmospheric pressure has a mean of 974.32 hPa, an SD of 60.43 hPa, a 1st quantile of 967.6 hPa, a median of 998.6 hPa, and a 3rd quantile of 1009.5 hPa. Its statistical distribution is one left-skewed peak, whose skewness is -2.37. The monthly average temperature has a mean of 14.98 °C and an SD of 10.33 °C, whose 1st quantile, median, 3rd quantile are 7.36 °C, 16.17 °C, and 22.99 °C. The shape of the distribution is one slightly left-skewed peak because of its skewness of -0.39. The NDVI has a mean value of 0.47, an SD of 0.17, a 1st quantile of 0.36, a median of 0.48, and a 3rd quantile of 0.60. The statistical distribution is one peak and slightly left-skewed (skewness: -0.34). The mean, SD, 1st quantile, median, and 3rd quantile of the precipitation are 0.12 $\text{kg}/(\text{m}^2 \cdot \text{h})$, 0.11 $\text{kg}/(\text{m}^2 \cdot \text{h})$, 0.05 $\text{kg}/(\text{m}^2 \cdot \text{h})$, 0.09 $\text{kg}/(\text{m}^2 \cdot \text{h})$, and 0.15 $\text{kg}/(\text{m}^2 \cdot \text{h})$, respectively. The shape of its statistical distribution is one right-skewed peak because its skewness is 2.25. The mean, SD, 1st quantile, median, and 3rd quantile of the PBLH are

755.92m, 313.13m, 550.6m, 712.8m, and 896.4m, respectively. The statistical distribution of the PBLH is one right-skewed peak (skewness: 1.79).

Figure 3 is located here.

Figure 4 shows the linear trends and correlation coefficients between the concentration of ground-level NO₂ and other variables. In **Figure 4**, r and p represent Pearson's correlation coefficient and p-value, respectively (Taylor 1990). The r is calculated as follows:

$$r = \frac{\sum_{i=1}^n (X1_i - \bar{X1})(X2_i - \bar{X2})}{\sqrt{\sum_{i=1}^n (X1_i - \bar{X1})^2} \sqrt{\sum_{i=1}^n (X2_i - \bar{X2})^2}} \quad (3)$$

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:

$$t = \frac{r\sqrt{n-2}}{\sqrt{1-r^2}} \quad (4)$$

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

$$p = 2 \times \Pr(T > t) \quad (5)$$

where T follows a t distribution with $n - 2$ degrees of freedom, and $\Pr(T > t)$ represents the probability of $T > t$. According to **Figure 4**, all correlation coefficients are significant at the 0.1 level. Albeit the coefficient of the correlation between ground-level NO₂ concentration and terrain atmospheric pressure is only -0.02, we still consider that it affects the ground-level NO₂ concentration as demonstrated by **Equation 2**.

Figure 4 is located here.

303 ***Methodology***

304 *Spatially Stationary Global Model*

305 We first assume that the relationships between the ground-level concentration of
306 NO₂ and other independent variables are spatially stationary. Those relationships do not
307 vary by spatial contexts (Brunsdon et al. 1998; Fotheringham et al. 2002). We apply the
308 three typical global models, random effects model (REM), fixed effects model (FEM), and
309 POLS, to estimate the relationships among the variables. The only difference among these
310 three global models is the time-fixed effects in the models. Therefore, substantial tests
311 about the time-fixed effects are required.

312 We select global models with different time-fixed effects, namely, REM, FEM, and
313 POLS. In order to compare FEM with POLS, we apply the F test for individual effects.
314 FEM is better than POLS since the test result is significant, and the null hypothesis that no
315 time-fixed effects are needed is rejected (Breusch and Pagan 1980; Croissant and Millo
316 2008). Based on the Breusch-Pagan Lagrange Multiplier test for random effects, the
317 significant result indicates that REM is appropriate because the null hypothesis that no
318 time-fixed effects are needed is not accepted, either (Breusch and Pagan 1980). We also
319 run the Hausman test to detect the reasonable model between FEM and REM (Kang 1985).
320 The significant result indicates FEM is preferred, as the null hypothesis that the preferred
321 model is REM is rejected. Here, FEM in our analysis is illustrated as follows:

$$GNO2_{it} = \beta X'_{it} + \alpha_i + \mu_{it} \quad (6)$$

322 where $GNO2_{it}$ denotes the monthly average ground-level concentration of NO₂ ($\mu g/m^3$)
323 in measurement i during t month, X_{it} denotes a matrix of independent variables, including
324 TrCA of NO₂ ($\mu g/m^2$), terrain atmospheric pressure (hPa), temperature ($^{\circ}C$), NDVI,

precipitation ($kg/(m^2 \cdot h)$), PBLH (m), and year dummy variables in measurement i during t month, α_i denotes the time-fixed effects, μ_{it} denotes an idiosyncratic error, and β is a vector of parameters to be estimated. Some ignored variables, such as air pollution policy in countries, economic status, among others, are changed year by year. We, therefore, add the year dummy variables to the model.

Spatially Non-stationary 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 ground-level NO₂ concentration, such as the goodness of fit of models, dramatically improves. Basically, the GWPR divides the total sample into numerous sub-samples, 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 as the neighbors of each other. Currently, two types of bandwidth are widely used, fixed distance bandwidth and adaptive distance bandwidth (Gollini et al. 2015). Literally, 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 an integer of 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). Taking a specific bandwidth (b), the mean square prediction error is calculated as follows:

$$MSPE(b) = \frac{m \sum_j [y_j - \widehat{y_j(b)}]^2}{(m - p + 1)^2} \quad (7)$$

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 the bandwidth b , and p is the number of the parameters in the analysis. It must be noted that the bandwidth b can be either the fixed distance bandwidth or the adaptive distance bandwidth here. The GWPR is an improvement of GWR, considering time-fixed effects. In the GWR, previous studies generally assume that **Equation 7** is a U-shape 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 degree to 20 arc degrees, and the examined bandwidth increases 0.25 arc degree every time because the spatial resolution of all the grid data is 0.25 arc degree. The optimal fixed distance bandwidth is 2.25 arc degrees (**Figure 5**). 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 data set. Both 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 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 5 is located here.

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

$$\mathbf{w}_i = \begin{cases} [1 - (\frac{\mathbf{d}}{b_i})^2]^2, & d_k \in \mathbf{d} \text{ if } d_k \leq b_i \\ 0, & d_k \notin \mathbf{d} \text{ if } d_k > b_i \end{cases} \quad (8)$$

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 weights vectors, the total sample is divided into 529 sub-samples, taking every city as the center, respectively.

The GWPR estimates the coefficients in every location in light of the spatially weighted matrix. In global models statistical tests, FEM is the preferred model, and in the locally statistically test, most local regressions show significant time-fixed effects.

Therefore, the GWPR model is also following FEM. The GWPR model is expressed as follows:

$$GNO2_{it} = \beta_i X'_{it} + \alpha_i + \epsilon_{it} \quad (9)$$

where β_i denotes a vector of parameters in the regression with the sub-sample taking the city i as the center. In terms of the equation expression, the only difference between FEM (**Equation 6**) and the GWPR based on FEM (**Equation 9**) is that β_i varies spatially. It must be noted that the algorithm would change to the GWPR based on POLS if we directly use the GWR method. According to cross validation, the accuracy and stability of the GWPR based on POLS is lower than the GWPR based on FEM. In this way, the new method, GWPR, is necessary for this analysis.

To estimate the coefficient in the GWPR, because the time-fixed effects (α_i) is unknown (Croissant and Millo 2008), we need to expand **Equation 9**, as follows:

$$GNO2_{it} - \hat{\theta}_i \overline{GNO2}_i = (1 - \hat{\theta}_i) \beta_{0i} + \beta_i (X_{it} - \hat{\theta}_i \overline{X}_i)' + (\epsilon_{it} - \hat{\theta}_i \bar{\epsilon}_i) \quad (10)$$

where $\hat{\theta}_i$ denotes the share of the individual effects in the total effects, $\overline{GNO2}_i$ denotes the mean of the ground-level NO2 concentration in the data set in the city i , β_{0i} is the estimated intercept in the city i , \overline{X}_i denotes a vector of means of independent variables in the city i , and $\bar{\epsilon}_i$ denotes the mean of the error in the city i . If the GWPR follows the data transformation of FEM, the $\hat{\theta}_i$ is always equal to 1. So β_{0i} is removed, in the GWPR based on FEM. To simplify Equation (10), we define the matrix of independent variables x_{it} and the dependent variable $gno2_{it}$ as follows:

$$x_{it} = X_{it} - \overline{X}_i \quad (11)$$

$$gno2_{it} = GNO2_{it} - \overline{GNO2}_i \quad (12)$$

$$\sigma_{it} = \epsilon_{it} - \bar{\epsilon}_i \quad (13)$$

408 In the light of **Equation 10-13**, the GWPR model is transformed as follows:

$$gno2_{it} = \beta_i \mathbf{x}'_{it} + \sigma_{it} \quad (14)$$

409 This algorithm becomes the typical GWR without the intercept. The coefficient estimation
410 could be displayed as follows:

$$\beta_i = [\mathbf{x}_{it}^T \mathbf{W}_i \mathbf{x}_{it}]^{-1} \mathbf{x}_{it}^T \mathbf{W}_i gno2_{it} \quad (15)$$

411

412 *Statistical Indicators*

413 Previous studies widely use several statistical indicators, including R^2 , root mean
414 square error (RMSE), mean absolute error (MAE), r between observed and predicted
415 values, and regression coefficients between observed and predicted values (Kim et al. 2021;
416 Li and Wu 2021; Liu et al. 2017), so we also apply to evaluate the estimation accuracy of
417 the GWPR model. R^2 is a critical statistical indicator describing the goodness of fit. The
418 high R^2 in the models means that the discrepancy between predicted NO_2 concentration
419 and measured ground-level NO_2 concentration is tiny. The R^2 calculation algorithm is
420 shown as follows:

$$R^2 = 1 - \frac{\sum_{k=1}^n (MNO2_k - PNO2_k)^2}{\sum_{k=1}^n (MNO2_k - \overline{MNO2})^2} \quad (16)$$

421 where n represents the observation number in the total sample, $MNO2_k$ represents the k th
422 record of the measured ground-level NO_2 concentration, $PNO2_k$ represents the k th record
423 of the predicted NO_2 concentration, and $\overline{MNO2}$ represents the mean of the measured
424 ground-level NO_2 concentration. It must be noted that this R^2 of the GWR or the GPWR is
425 the global value because here the total sample is used. The RMSE is also calculated,

indicating the differences between predicted NO₂ concentration and measured ground-level NO₂ concentration. The RMSE is sensitive to both systematic error and random error, shown as follows:

$$RMSE = \sqrt{\frac{1}{n} \sum_{k=1}^n (MNO2_k - PNO2_k)^2} \quad (17)$$

The MAE is calculated as follows:

$$MAE = \text{mean}(|MNO2_k - PNO2_k|) \quad (18)$$

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

$$MNO2_k = \alpha + \beta PNO2_k + \delta_k \quad (19)$$

where α is the intercept in the regression and the ideal value of α is 0, β is the slope and its ideal value is 1, and δ_k is a random error term.

Spatial Interpolation of the Coefficients

We use the ordinary Kriging (OK) method 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 NO₂ concentration raster data because the result is a spatial point data frame and the coefficients spatially vary. In the beginning, we consider the inverse distance weighting (IDW) and OK method. To choose the reasonable interpolation method, these two methods are examined by leave-one-out cross validation. The result shows that the OK method is more stable than IDW, consistent with the previous study (Li and Heap 2011). The general form of the OK method is expressed (Pebesma 2004) as follows:

$$\hat{\beta}(s_0) = \sum_{l=1}^m \lambda_l \beta(s_l) \text{ with } \sum_{l=1}^m \lambda_l = 1 \quad (20)$$

where $\hat{\beta}(s_0)$ represents the coefficient at unobserved location s_0 , $\beta(s_l)$ represent the coefficient value at known location s_l , λ_l is the spatial weight for the coefficient value at the l th location, and m is the number of known values. To obtain optimal λ_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 semi-variance 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 11** and **12** are necessary. To check the reliability of the interpolation process, we apply the leave-one-out cross validation. Simply, the leave-one-out cross-validation requires the number of folds equals the numbers of cities in our data set. In the cross validation process, one certain city is selected as the test set, while all other 529 cities are used as the training set. This process will repeat 530 times until all the cities have been selected as the test set.

Predictions of Ground-level NO₂ Concentration and Change Trends

The monthly global ground-level NO₂ concentration raster data are predicted based on the original raster data, coefficient rasters, and mean value rasters, following **Equation 14**. Furthermore, because the ground-level NO₂ concentration cannot be minus, 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 NO₂ concentrations of each grid are regressed with their month order. The first month of our

data set is January 2015, so the month order of data in January 2015 is the 1st. Lastly, the month order of October 2021 is the 82nd. The monthly change trends are estimated as follows:

$$PNO2_{LT} = \alpha_L + \beta_L MO_T + \varphi_{LT} \quad (20)$$

where $PNO2_{LT}$ is the predicted NO₂ concentration in the grid L in the T th month, MO_T is the month order of $PNO2_{LT}$, α_L is the intercept in the grid L , β_L is the monthly NO₂ concentration change trend in the grid L , and φ_{LT} is a random error term. It must be noted that the data from a certain grid should be performed the regression once if there are no less than 30 records. Additionally, every individual regression just uses the data from one grid.

Results

Validation of the GWPR Result

The overall accuracy of the GWPR based on FEM is 74.45%, the RMSE is 7.171 $\mu g/m^3$, and the is MAE 3.48 $\mu g/m^3$. **Figure 6** shows the relationship between predicted and measured ground-level NO₂ concentrations in the GWPR, and all coefficients in **Figure 6** are significant. The coefficient of the correlation (r) between predicted and measured values are 0.8632, following **Equation 3**, 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 roughly 1.021 significantly, and the ideal value is 1. To confirm whether the goodness of fit of a specific year is apparently lower than 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 is 9.619 and $4.071 \mu\text{g}/\text{m}^3$. Comparisons of monthly time series illustrate that the GWPR captures the monthly temporal variability accurately (**Figure 7**).

Figure 6 is located here.

Table 1 is located here.

Figure 7 is located here.

Table 2 reports the results of 10-fold cross validation. The total data set 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 accuracy of prediction for the test subset is 68.55%, and its RMSE, MEA, and coefficient (r) of the correlation between 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 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 values of its RMSE and MAE, this interpolated grid data set is still acceptable. There is a new indicator named on point R^2 in **Table 3**. This indicator is to depict 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 are overlapped. Hence, after OK interpolation, the predicted values from the imputed grid data set might differ slightly from the observed values. The on-point R^2 indicates that none of the interpolations lose the accuracy more than 2%, and most of them only slash 0.1% accuracy.

Table 3 is located here.

Prediction and Monthly Change Trends

The accuracy of the predicted grid data set 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 in the final prediction is lower than the result of GWPR. **Figure 8** demonstrates the relationship between predicted and measured ground-level NO_2 concentrations in the predicted grid data sets, and all coefficients in **Figure 8** are significant. The r is 0.8376, lower than the r in the GWPR result (0.8585). 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 9 presents the average monthly ground-level NO_2 concentration. The result shows that ground-level NO_2 concentration in Asia is relatively higher than in other

continents. Moreover, in western Asia, the ground-level NO₂ pollution is most severe. In most regions of Iran, Iraq, and Turkmenistan, the average values even exceed 100 $\mu\text{g}/\text{m}^3$, much higher than WHO's ambient NO₂ 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, partial regions in northern China, India, Chile, Bolivia, South Africa, northern Italy, and the west coast of the United States are also severely polluted, tallying with previous research (Di et al. 2020; Kim et al. 2021; Zheng et al. 2019). **Figure 10** shows the average monthly TrCA of NO₂ concentration from January 2015 to October 2021. The spatial distribution of severely polluted regions lines up with our predictions, which partially backs up our research. The monthly change trends of each grid are estimated following **Equation 20**. **Figure 11** illustrates the monthly change trends of each grid. All the values shown in **Figure 11** should be significant at the 0.1 level. **Figure 11** demonstrates that the ground-level NO₂ concentration is gradually decreasing in most regions. However, in the Red Sea region, Arabian Peninsula, Persian Gulf, the ground-level NO₂ concentration is increasing, even though there are already the most severely polluted areas. In addition, the changes in most places of India and Mexico also show an upward trend. In some areas of California and Nevada, the United States, the change trends are on the rise.

Figure 9 is located here.

Figure 10 is located here.

Figure 11 is located here.

Discussion

We employ the GWPR model to examine the association of satellite-derived data with measured ground-level NO₂ 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 about regional ground-level NO₂ concentration (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 data set. Our study predicts 82 monthly global ground-level NO₂ concentrations with 0.25-arc-degree spatial resolution from January 2015 to October 2021 (Shown in the **Supplementary Materials**). Also, our analysis demonstrates the monthly change trends of each grid.

In terms of accuracy, our study has made some improvements in the ground-level NO₂ concentration estimation based on the satellite data set. The accuracy of a study on ground-level NO₂ concentration over Central-Eastern China employing geographically and temporally weighted regression is 0.60, 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 research over the Wuhan urban agglomeration applying space-time neural networks obtains 0.69 accuracy and an 8.29- $\mu\text{g}/\text{m}^3$ RMSE, while our RMSE is 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, the NO₂ pollution is harmful, higher than 25 PPB (Di et al. 2020). Our prediction is in line with this study since the ground-level NO₂ in those places is predicted around 40 $\mu\text{g}/\text{m}^3$ (Under an ambient pressure of 1 atmosphere and a temperature of 25 °C, 1 PPB NO₂ roughly equals 1.88 $\mu\text{g}/\text{m}^3$). Kasparoglu et al. put forward that in urban sites of Marmara regression of Turkey, the NO₂ concentration are higher than 75 $\mu\text{g}/\text{m}^3$ (Kasparoglu et al. 2018), consistent with our study. Yousefian et al. propose that during 2012 - 2017, the NO₂ concentration in Tehran, the Capital of Iran, is always 1.5 - 2.5 times higher than WHO limits (40 $\mu\text{g}/\text{m}^3$) (Yousefian et al. 2020), where is also the most polluted area in our research. Kim et al.'s research hints that in an Alpine domain, the NO₂ in the 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. estimate the annual average ground-level NO₂ concentration in the main regions of China over 2014 – 2020 (Chi et al. 2021). Its results are similar to our predictions. Moreover, because the research on the ground-level NO₂ concentration in polar regions is rare, it is difficult to compare our study with previous. However, the low temperature and PBLH may account for the high NO₂ concentration in polar regions. Evidence shows that in winter, the NO₂ concentration is apparently higher than it 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 NO₂ concentration.

Our monthly grid data demonstrate that the temporal variation of the spatial distribution of the ground-level NO₂ concentration is evident (Shown in **Supplementary Materials** and **Figure 7**). In December, January, and February, the NO₂ 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: firstly, NO₂ can last longer or it is difficult to be removed in the low-temperature environment (Yousefian et al. 2020); secondly, residential energy for heating may produce more NO₂ (Fan et al. 2020); and thirdly, the low temperature changes other meteorological factors eventually leading to NO₂ unable to spread to other zones far from the surface (Fan et al. 2020; Shen et al. 2021). However, no matter which is the main reason, the policymakers should take this temporal variation into account in the air pollution policies.

However, some limitations remain in our study. Firstly, the number of measurement points is limited, and their spatial distribution is unbalanced. Our data set takes only 530 cities into account globally without rural areas. Especially in Africa and South America, there are few measurement points. Secondly, the spatial resolutions of one control variable, namely PBLH, are not good enough. Thirdly, some hidden variables are ignored. For instance, the research does not involve the impacts of air pollution policies and COVID-19, though we have used year dummy variables to reduce their effects. Fourthly, the spatial and temporal resolutions are relatively low. Future studies are better to focus on the following topics. Firstly, the GWPR model should be further optimized. For example, if the data set contains data islands based on a certain bandwidth, whether the GWRP is still statistically reliable remains unknown. Secondly, more confounders, such as land cover, human activity, among others, are needed in future analyses. Finally, the vertical distribution of NO₂ in air and its physical and chemical reasons are better to be carefully explored.

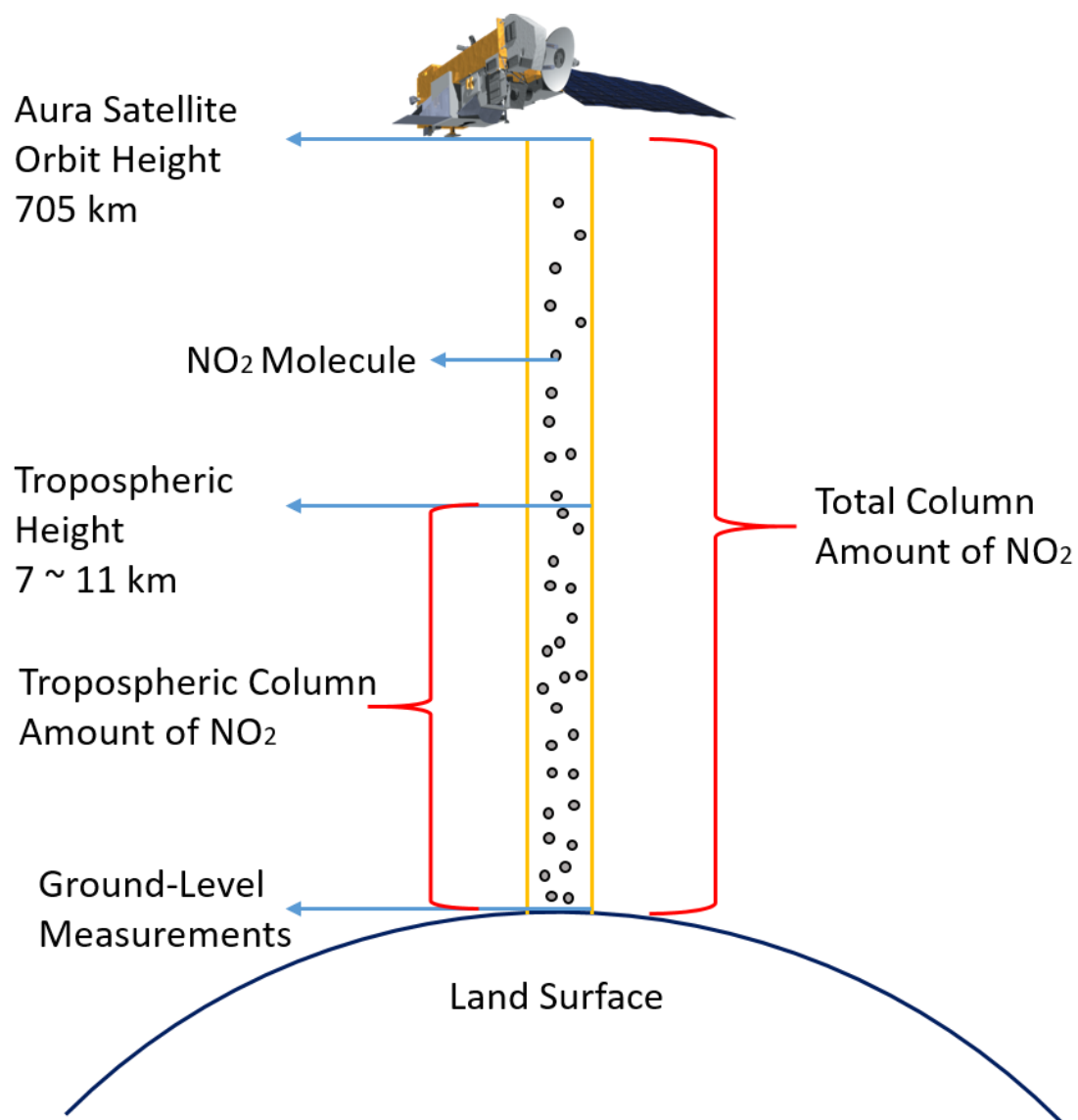
Conclusions

This study creates a GWPR model to detect the spatial variation of the relationship between the measured ground-level NO₂ concentration and satellite-derived data and predict the monthly ground-level NO₂ concentration based on the data from the OMI and other satellite platforms. The GWPR model illustrates the outstanding accuracy (0.7445), *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 NO₂ concentration 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

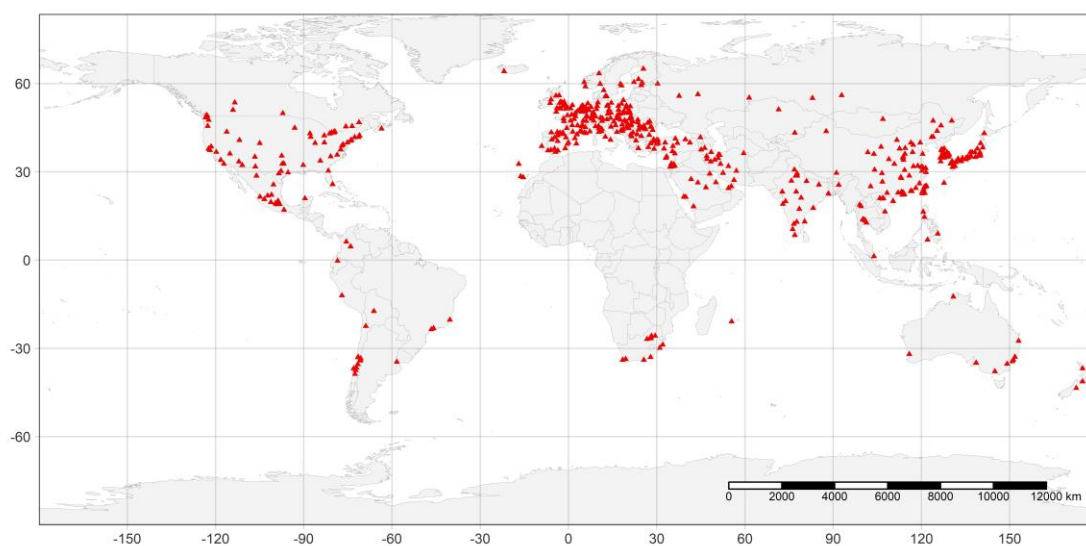
All data sources used in the analyses, along with fully reproducible code, are publicly available at https://github.com/MichaelChaoLi-cpu/Monthly_Global_Ground_Level_NO2.

638 **Figure**



639

640 **Figure 1: Illustration regarding TrCA, ToCA, and Ground-Level Measurements**



641

642

Figure 2: Locations of Cities with the Measurement Points

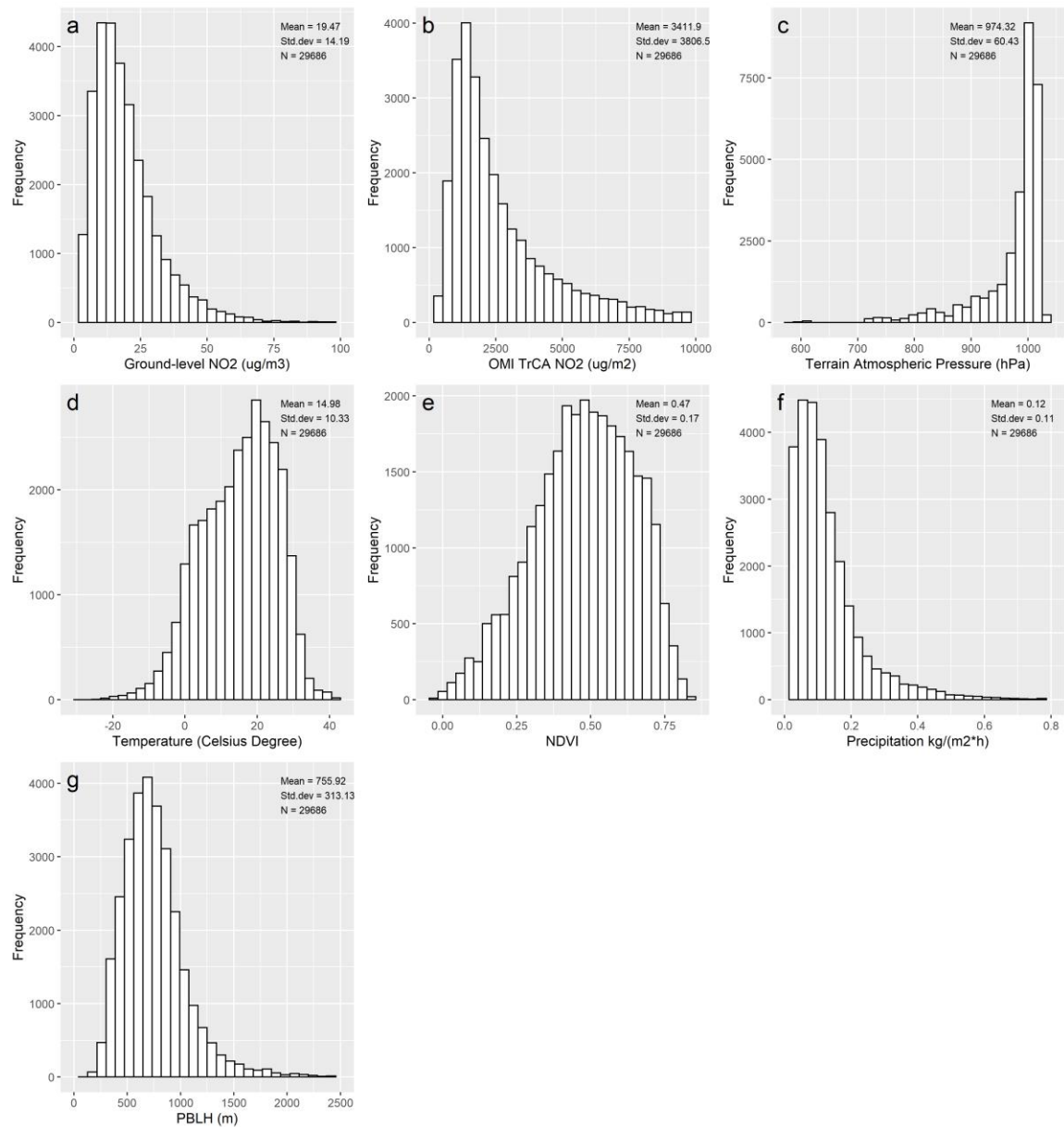


Figure 3: Statistical Distributions and Descriptive Statistics of the Variables

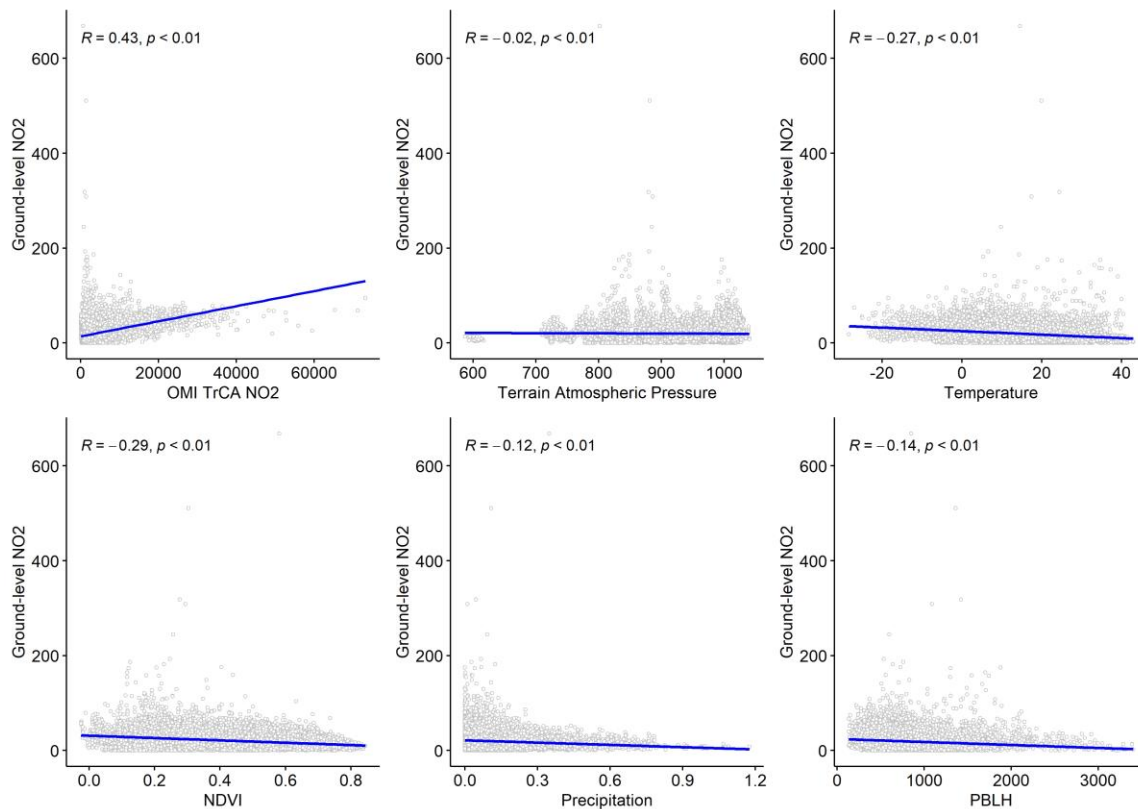


Figure 4: Linear Trends and Correlation Coefficients

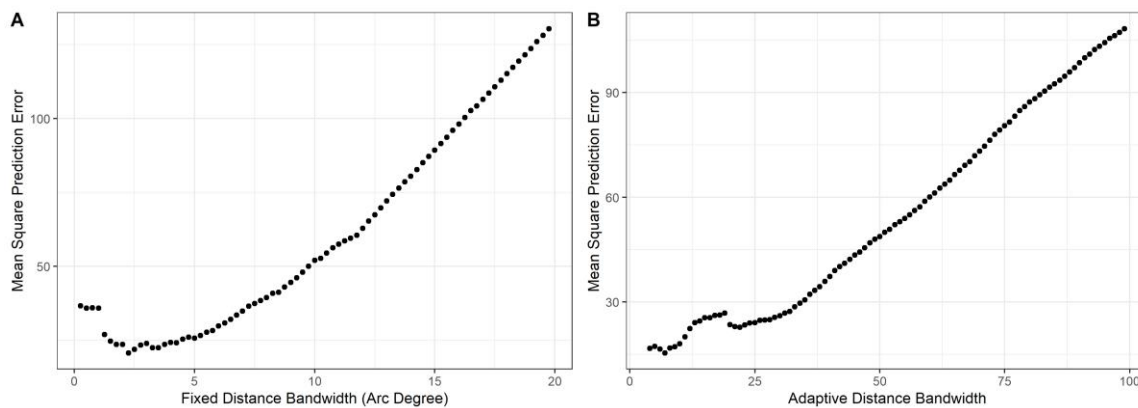


Figure 5: Step Increment Bandwidth Selection

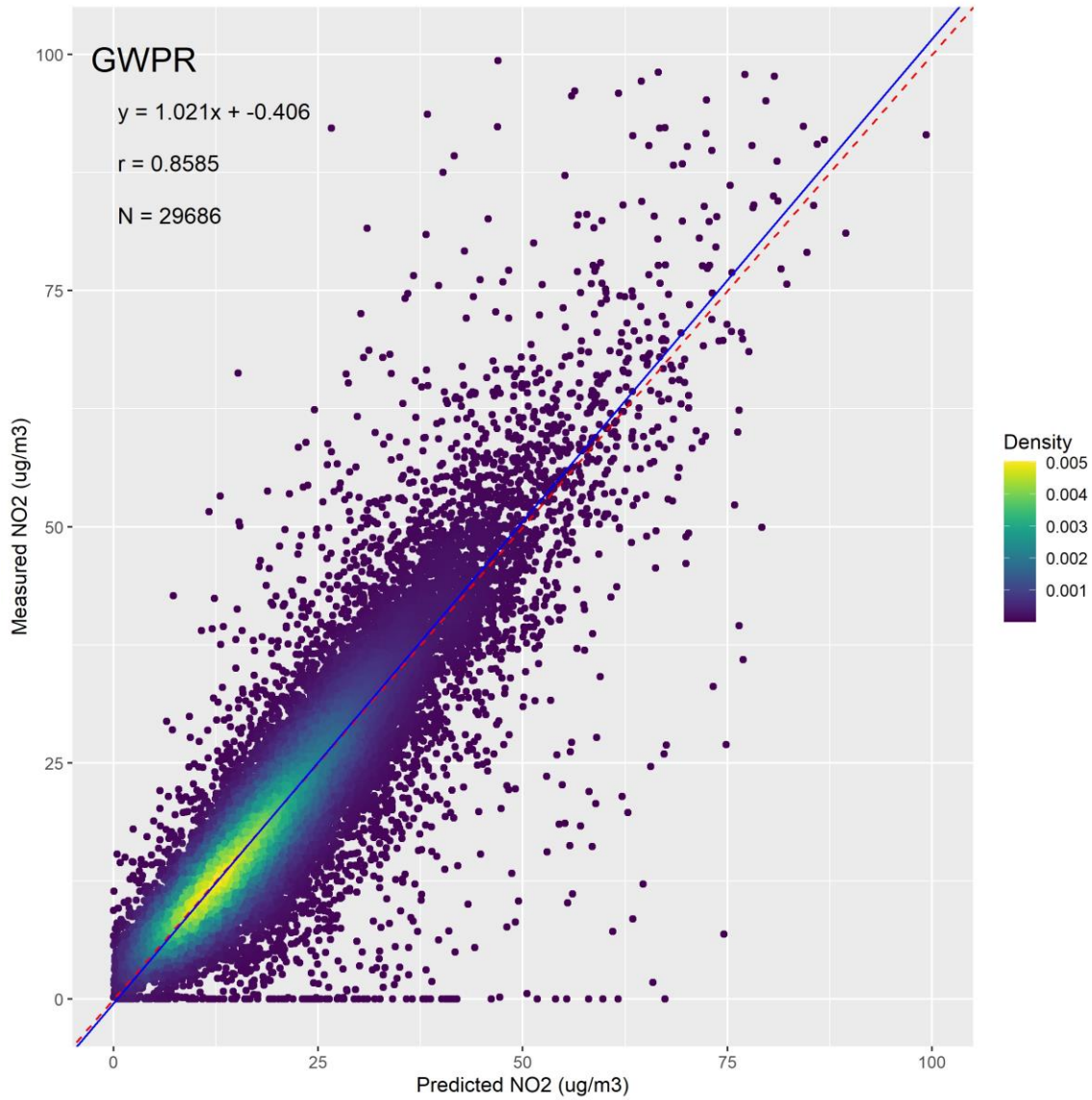
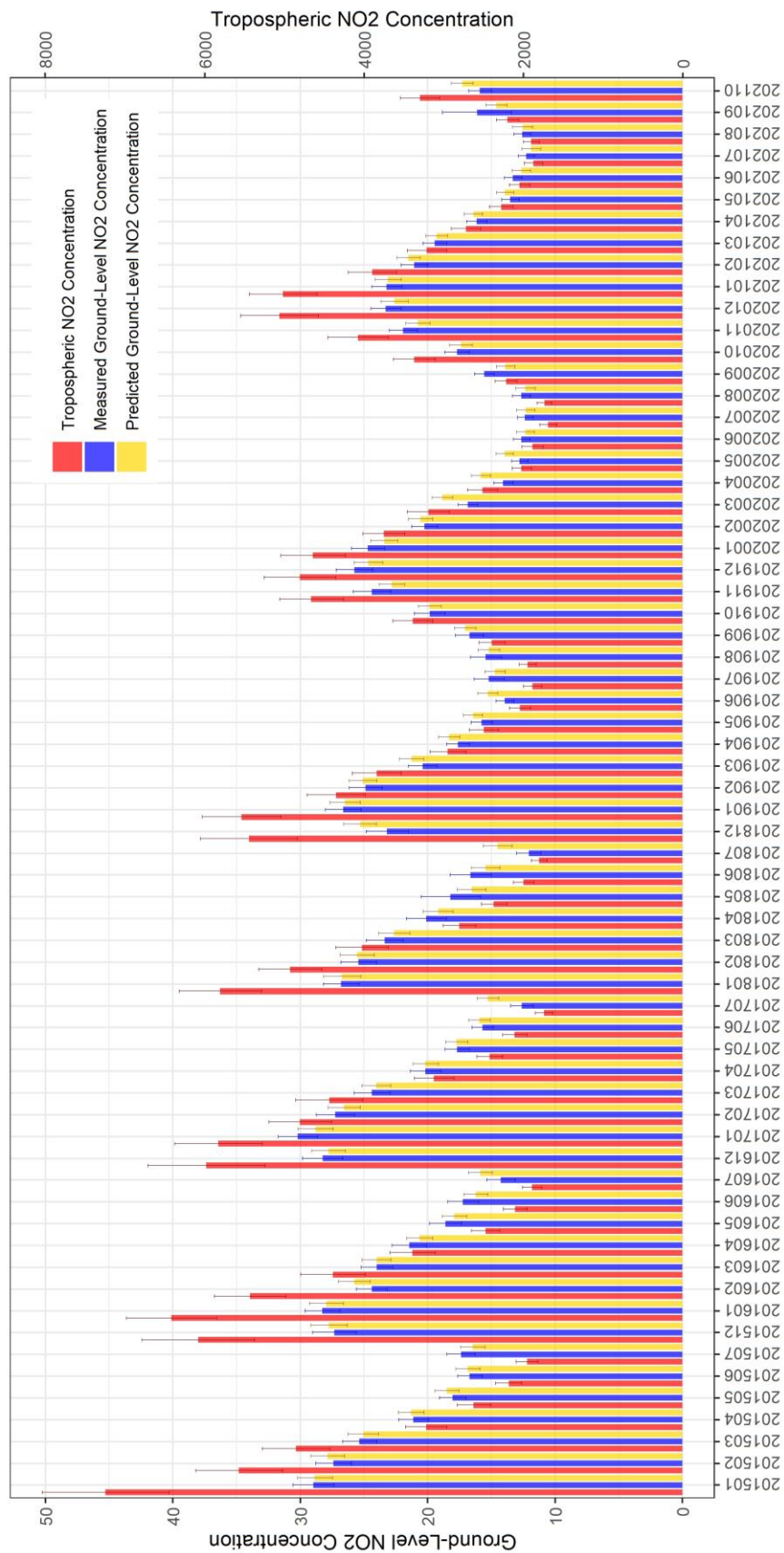


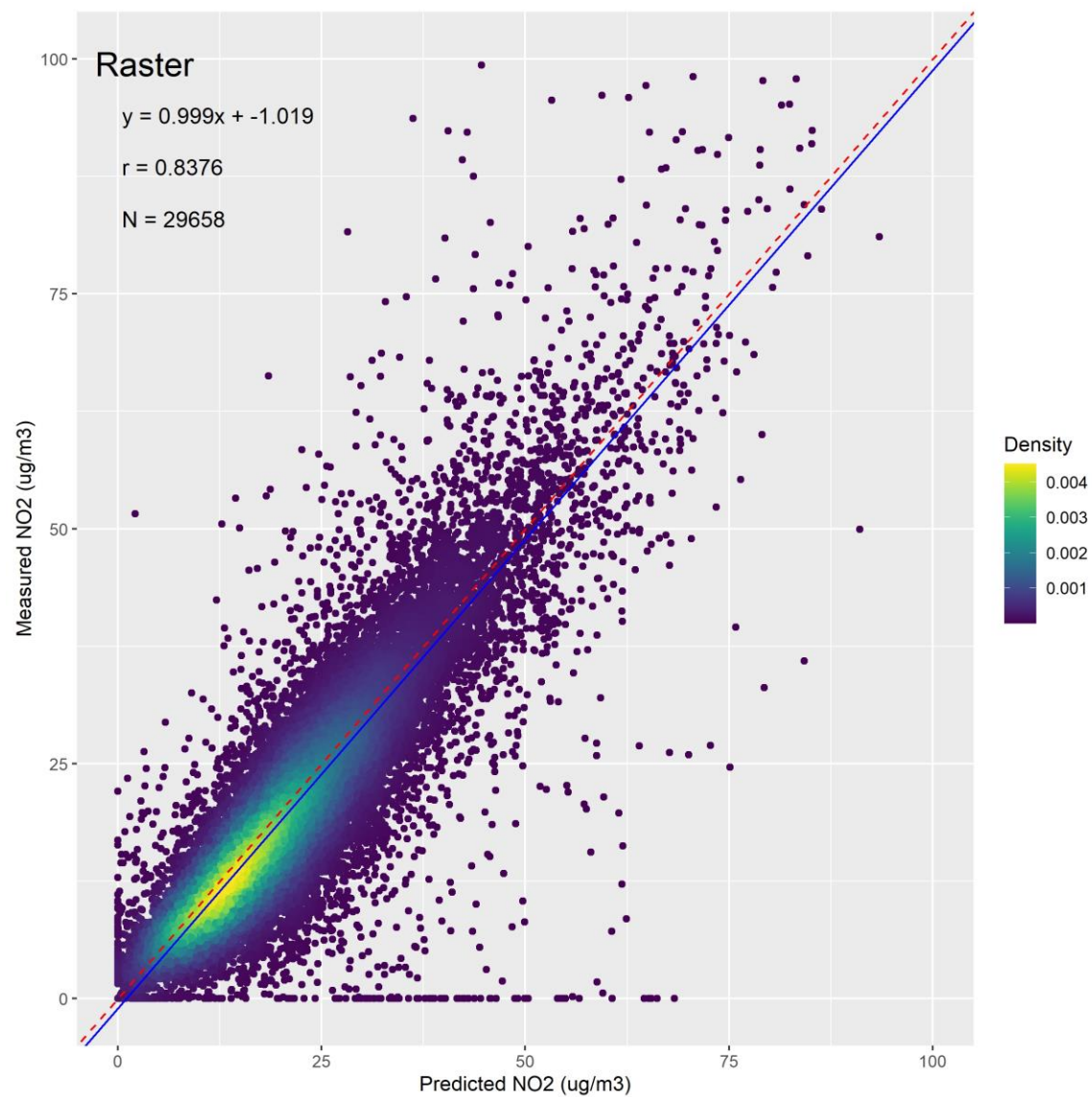
Figure 6: The Density Plots between the Measured Ground-Level NO₂ Concentration and the Predicted NO₂ Concentration from the GWPR Result.
 (The red dashed line is the 1:1 line. The blue line is the regression line.)



**Figure 7: Time Series Comparisons of Monthly Means of Tropospheric, Measured
Ground-Level, Predicted Ground-Level NO₂ Concentration and Their 95%
Confidence Interval**

(Unit: Tropospheric NO₂ Concentration [$\mu\text{g}/\text{m}^2$], Measured and Predicted Ground-
Level NO₂ Concentration [$\mu\text{g}/\text{m}^3$])

662



663

664 **Figure 8: The Density Plots between the Measured Ground-Level NO₂**
665 **Concentration and the Predicted NO₂ Concentration from the Predicted Grid Data**
666 **Sets.**

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

668

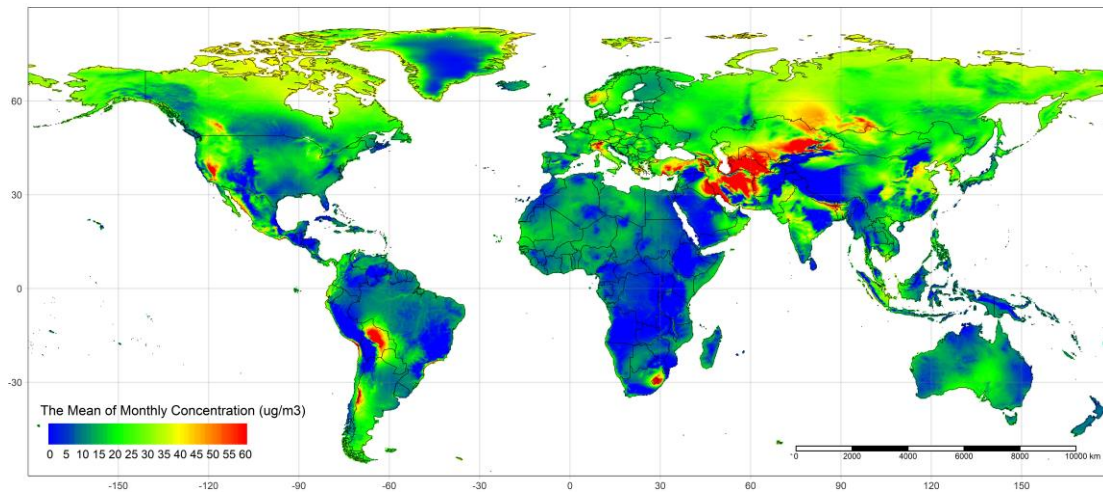


Figure 9: Average Monthly Ground-Level NO₂ Concentration, Jan 2015 – Oct 2021

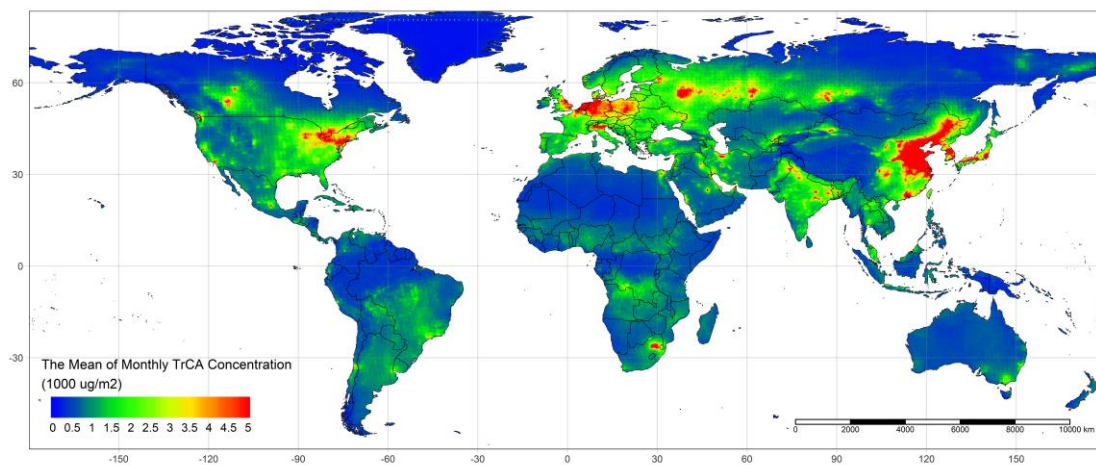
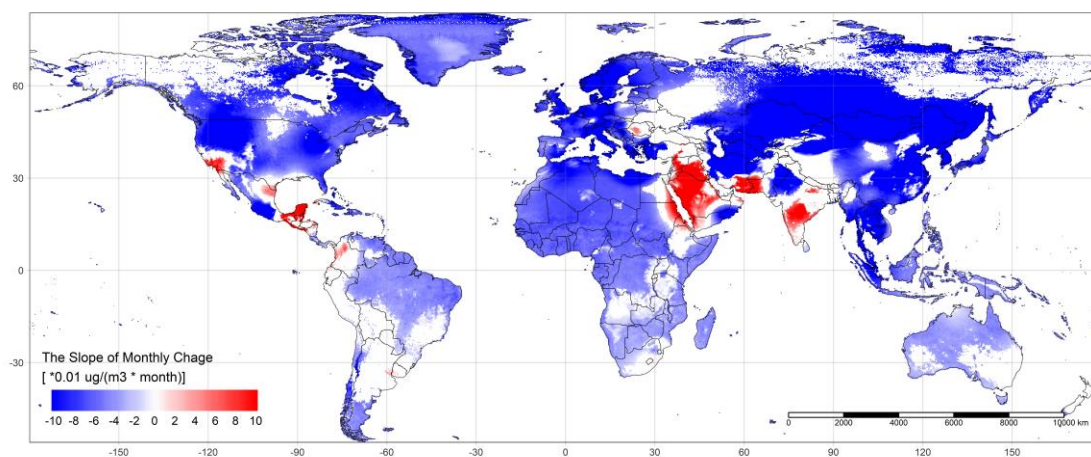


Figure 10: Average Monthly TrCA NO₂ Concentration, Jan 2015 – Oct 2021



**Figure 11: Monthly Change Trends of Ground-Level NO₂ Concentration, Jan 2015
– Oct 2021**

Table 1: Yearly Statistical Indicators Summary of the GWPR Results

| Year | N | R ² | 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.

680

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.

681

682

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

| | Variable | N | R ² | MRSE $\mu g/m^3$ | MAE $\mu g/m^3$ | r | Intercept | Slope | On Point R ² |
|------------------------------|--|-----|----------------|---------------------|--------------------|-------|-----------|-------|----------------------------|
| Coefficient Interpolation | TrCA of NO ₂ 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 ₂ | 530 | 34.50% | 7.773 | 5.633 | 0.626 | 4.505 | 0.748 | 98.40% |
| | TrCA of NO ₂ 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% |

683

684

685

686 **Reference:**

- 687 Bechle, M.J., Millet, D.B., & Marshall, J.D. (2015). National Spatiotemporal Exposure
688 Surface for NO₂: Monthly Scaling of a Satellite-Derived Land-Use Regression, 2000–
689 2010. *Environmental Science & Technology*, 49, 12297-12305
- 690 Beckerman, B., Jerrett, M., Brook, J.R., Verma, D.K., Arain, M.A., & Finkelstein, M.M.
691 (2008). Correlation of nitrogen dioxide with other traffic pollutants near a major
692 expressway. *Atmospheric Environment*, 42, 275-290
- 693 Beenstock, M., & Felsenstein, D. (2019). *The econometric analysis of non-stationary*
694 *spatial panel data*. Springer
- 695 Bigdeli, M., Taheri, M., & Mohammadian, A. (2021). Spatial sensitivity analysis of
696 COVID-19 infections concerning the satellite-based four air pollutants levels.
697 *International Journal of Environmental Science and Technology*, 18, 751-760
- 698 Breusch, T.S., & Pagan, A.R. (1980). The Lagrange Multiplier Test and its Applications to
699 Model Specification in Econometrics. *The Review of Economic Studies*, 47, 239
- 700 Brunekreef, B., & Holgate, S.T. (2002). Air pollution and health. *The Lancet*, 360, 1233-
701 1242
- 702 Brunsdon, C., Fotheringham, A.S., & Charlton, M.E. (2010). Geographically Weighted
703 Regression: A Method for Exploring Spatial Nonstationarity. *Geographical Analysis*, 28,
704 281-298
- 705 Brunsdon, C., Fotheringham, S., & Charlton, M. (1998). Geographically Weighted
706 Regression. *Journal of the Royal Statistical Society: Series D (The Statistician)*, 47, 431-
707 443
- 708 Chi, Y., Fan, M., Zhao, C., Sun, L., Yang, Y., Yang, X., & Tao, J. (2021). Ground-level
709 NO₂ concentration estimation based on OMI tropospheric NO₂ and its spatiotemporal
710 characteristics in typical regions of China. *Atmospheric Research*, 264, 105821
- 711 Chiusolo, M., Cadum, E., Stafoggia, M., Galassi, C., Berti, G., Faustini, A., Bisanti, L.,
712 Vigotti, M.A., Dessì M.P., Cernigliaro, A., Mallone, S., Pacelli, B., Minerba, S., Simonato,
713 L., & Forastiere, F. (2011). Short-Term Effects of Nitrogen Dioxide on Mortality and
714 Susceptibility Factors in 10 Italian Cities: The EpiAir Study. *Environmental Health*
715 *Perspectives*, 119, 1233-1238
- 716 Croissant, Y., & Millo, G. (2008). Panel Data Econometrics in R: The plm Package.
717 *Journal of Statistical Software*, 27
- 718 Curier, R.L., Kranenburg, R., Segers, A.J.S., Timmermans, R.M.A., & Schaap, M. (2014).
719 Synergistic use of OMI NO₂ tropospheric columns and LOTOS–EUROS to evaluate the
720 NO_x emission trends across Europe. *Remote Sensing of Environment*, 149, 58-69
- 721 Di, Q., Amini, H., Shi, L., Kloog, I., Silvern, R., Kelly, J., Sabath, M.B., Choirat, C.,
722 Koutrakis, P., Lyapustin, A., Wang, Y., Mickley, L.J., & Schwartz, J. (2020). Assessing
723 NO₂ Concentration and Model Uncertainty with High Spatiotemporal Resolution across
724 the Contiguous United States Using Ensemble Model Averaging. *Environmental Science*
725 *& Technology*, 54, 1372-1384
- 726 Fan, H., Zhao, C., & Yang, Y. (2020). A comprehensive analysis of the spatio-temporal
727 variation of urban air pollution in China during 2014–2018. *Atmospheric Environment*, 220,
728 117066
- 729 Fotheringham, A., Brunsdon, C., & Charlton, M. (2002). *Geographically Weighted*
730 *Regression: The Analysis of Spatially Varying Relationships*. John Wiley & Sons

731 Fotheringham, A.S., Crespo, R., & Yao, J. (2015). Geographical and Temporal Weighted
732 Regression (GTWR). *Geographical Analysis*, 47, 431-452

733 Fotheringham, A.S., & Oshan, T.M. (2016). Geographically weighted regression and
734 multicollinearity: dispelling the myth. *Journal of Geographical Systems*, 18, 303-329

735 Geddes, J.A., Martin, R.V., Boys, B.L., & Van Donkelaar, A. (2016). Long-Term Trends
736 Worldwide in Ambient NO₂ Concentrations Inferred from Satellite Observations.
737 *Environmental Health Perspectives*, 124, 281-289

738 Gollini, I., Lu, B., Charlton, M., Brunsdon, C., & Harris, P. (2015). GWmodel: An R
739 package for exploring spatial heterogeneity using geographically weighted models.
740 *Journal of Statistical Software*, 63

741 Hamra, G.B., Laden, F., Cohen, A.J., Raaschou-Nielsen, O., Brauer, M., & Loomis, D.
742 (2015). Lung Cancer and Exposure to Nitrogen Dioxide and Traffic: A Systematic Review
743 and Meta-Analysis. *Environmental Health Perspectives*, 123, 1107-1112

744 Hu, X., Waller, L.A., Al-Hamdan, M.Z., Crosson, W.L., Estes, M.G., Estes, S.M.,
745 Quattrochi, D.A., Sarnat, J.A., & Liu, Y. (2013). Estimating ground-level PM_{2.5}
746 concentrations in the southeastern U.S. using geographically weighted regression.
747 *Environmental Research*, 121, 1-10

748 Irie, H., Boersma, K.F., Kanaya, Y., Takashima, H., Pan, X., & Wang, Z.F. (2012).
749 Quantitative bias estimates for tropospheric NO₂ columns retrieved from SCIAMACHY,
750 OMI, and GOME-2 using a common standard for East Asia. *Atmospheric Measurement*
751 *Techniques*, 5, 2403-2411

752 Jiang, M., Sun, W., Yang, G., & Zhang, D. (2017). Modelling Seasonal GWR of Daily
753 PM_{2.5} with Proper Auxiliary Variables for the Yangtze River Delta. *Remote Sensing*, 9,
754 346

755 Kang, S. (1985). A note on the equivalence of specification tests in the two-factor
756 multivariate variance components model. *Journal of Econometrics*, 28, 193-203

757 Kasparoglu, S., Incecik, S., & Topcu, S. (2018). Spatial and temporal variation of O₃, NO
758 and NO₂ concentrations at rural and urban sites in Marmara Region of Turkey.
759 *Atmospheric Pollution Research*, 9, 1009-1020

760 Kim, M., Brunner, D., & Kuhlmann, G. (2021). Importance of satellite observations for
761 high-resolution mapping of near-surface NO₂ by machine learning. *Remote Sensing of*
762 *Environment*, 264, 112573

763 Larkin, A., Geddes, J.A., Martin, R.V., Xiao, Q., Liu, Y., Marshall, J.D., Brauer, M., &
764 Hystad, P. (2017). Global Land Use Regression Model for Nitrogen Dioxide Air Pollution.
765 *Environmental Science & Technology*, 51, 6957-6964

766 Lelieveld, J., Evans, J.S., Fnais, M., Giannadaki, D., & Pozzer, A. (2015). The contribution
767 of outdoor air pollution sources to premature mortality on a global scale. *Nature*, 525, 367-
768 371

769 Li, C., & Managi, S. (2021a). Contribution of on-road transportation to PM_{2.5}. *Scientific*
770 *Reports*, 11, 21320

771 Li, C., & Managi, S. (2021b). Spatial Variability of the Relationship between Air Pollution
772 and Well-being. *Sustainable Cities and Society*, 103447

773 Li, C., & Managi, S. (2022). Impacts of air pollution on COVID-19 case fatality rate: a
774 global analysis. *Environmental Science and Pollution Research*, Jan 4:1-14.

775 Li, J., & Heap, A.D. (2011). A review of comparative studies of spatial interpolation
 776 methods in environmental sciences: Performance and impact factors. *Ecological*
 777 *Informatics*, 6, 228-241
 778 Li, L., & Wu, J. (2021). Spatiotemporal estimation of satellite-borne and ground-level NO₂
 779 using full residual deep networks. *Remote Sensing of Environment*, 254, 112257
 780 Li, T., Wang, Y., & Yuan, Q. (2020). Remote Sensing Estimation of Regional NO₂ via
 781 Space-Time Neural Networks. *Remote Sensing*, 12, 2514
 782 Lin, C.-A., Chen, Y.-C., Liu, C.-Y., Chen, W.-T., Seinfeld, J.H., & Chou, C.C.K. (2019).
 783 Satellite-Derived Correlation of SO₂, NO₂, and Aerosol Optical Depth with
 784 Meteorological Conditions over East Asia from 2005 to 2015. *Remote Sensing*, 11, 1738
 785 Liu, L., Zhang, X., Xu, W., Liu, X., Lu, X., Chen, D., Zhang, X., Wang, S., & Zhang, W.
 786 (2017). Estimation of monthly bulk nitrate deposition in China based on satellite NO₂
 787 measurement by the Ozone Monitoring Instrument. *Remote Sensing of Environment*, 199,
 788 93-106
 789 Mackerron, G., & Mourato, S. (2009). Life satisfaction and air quality in London.
 790 *Ecological Economics*, 68, 1441-1453
 791 Meng, L., Liu, J., Tarasick David, W., Randel William, J., Steiner Andrea, K., Wilhelmsen,
 792 H., Wang, L., & Haimberger, L. (2021). Continuous rise of the tropopause in the Northern
 793 Hemisphere over 1980–2020. *Science Advances*, 7, eabi8065
 794 Newell, K., Kartsonaki, C., Lam, K.B.H., & Kurmi, O.P. (2017). Cardiorespiratory health
 795 effects of particulate ambient air pollution exposure in low-income and middle-income
 796 countries: a systematic review and meta-analysis. *The Lancet Planetary Health*, 1, e368-
 797 e380
 798 Nickolay, K., Lok, L., Sergey, M., Edward, C., Eric, B., William, S., Joanna, J., & the OMI
 799 core team (2019). OMI/Aura NO₂ Total and Tropospheric Column Daily L2 Global
 800 Gridded 0.25 degree x 0.25 degree V3. In N.G.S.F. Center (Ed.): Goddard Earth Sciences
 801 Data and Information Services Center (GES DISC)
 802 Ogen, Y. (2020). Assessing nitrogen dioxide (NO₂) levels as a contributing factor to
 803 coronavirus (COVID-19) fatality. *Science of The Total Environment*, 726, 138605
 804 OMI Team (2012). Ozone Monitoring Instrument (OMI) data user's guide. In
 805 Orellano, P., Reynoso, J., Quaranta, N., Bardach, A., & Ciapponi, A. (2020). Short-term
 806 exposure to particulate matter (PM₁₀ and PM_{2.5}), nitrogen dioxide (NO₂), and ozone (O₃)
 807 and all-cause and cause-specific mortality: Systematic review and meta-analysis.
 808 *Environment International*, 142, 105876
 809 Pebesma, E.J. (2004). Multivariable geostatistics in S: the gstat package. *Computers &*
 810 *Geosciences*, 30, 683-691
 811 Qin, K., Rao, L., Xu, J., Bai, Y., Zou, J., Hao, N., Li, S., & Yu, C. (2017). Estimating
 812 Ground Level NO₂ Concentrations over Central-Eastern China Using a Satellite-Based
 813 Geographically and Temporally Weighted Regression Model. *Remote Sensing*, 9, 950
 814 Rice, M.B., Ljungman, P.L., Wilker, E.H., Gold, D.R., Schwartz, J.D., Koutrakis, P.,
 815 Washko, G.R., O'Connor, G.T., & Mittleman, M.A. (2013). Short-Term Exposure to Air
 816 Pollution and Lung Function in the Framingham Heart Study. *American Journal of*
 817 *Respiratory and Critical Care Medicine*, 188, 1351-1357
 818 Schoeberl, M.R., Douglass, A.R., Hilsenrath, E., Bhartia, P.K., Beer, R., Waters, J.W.,
 819 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 NO₂ 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 NO₂ 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

Zheng, Yang, Wu, & Marinello (2019). Spatial Variation of NO₂ and Its Impact Factors in China: An Application of Sentinel-5P Products. *Remote Sensing*, 11, 1939