A Comparison of Regression Methods for Inferring Near-Surface NO2 with Satellite Data

Eliot J Kim[1], Tracey Holloway[1], Ajinkya Kokandakar[1], Monica Harkey[1], Stephanie Elkins[2], Daniel Goldberg[3], and Colleen Heck[1]

[1]University of Wisconsin-Madison
[2]Massachusetts Institute of Technology
[3]George Washington University

March 14, 2024

Abstract

Nitrogen dioxide (NO2) is emitted during high temperature combustion from anthropogenic and natural sources. Human exposure to high NO2 concentrations causes cardiovascular and respiratory illnesses. The EPA operates ground monitors across the U.S. which take hourly measurements of NO2 concentrations, providing precise measurements for assessing human pollution exposure but with sparse spatial distribution. Satellite-based instruments capture NO2 amounts through the atmospheric column with global coverage at regular spatial resolution, but do not directly measure surface NO2. This study compares regression methods using satellite NO2 data from the TROPospheric Ozone Monitoring Instrument (TROPOMI) to estimate annual surface NO2 concentrations in varying geographic and land use settings across the continental U.S. We then apply the best-performing regression models to estimate surface NO2 at 0.01o by 0.01o resolution, and we term this estimate as quasi-NO2 (qNO2). qNO2 agrees best with measurements at suburban sites (cross-validation (CV) R2 = 0.72) and away from major roads (CV R2 = 0.75). Among U.S. regions, qNO2 agrees best with measurements in the Midwest (CV R2 = 0.89) and agrees least in the Southwest (CV R2 = 0.65). To account for the non-Gaussian distribution of TROPOMI NO2, we apply data transforms, with the Anscombe transform yielding highest agreement across the continental U.S. (CV R2 = 0.78). The interpretability, minimal computational cost, and health relevance of qNO2 facilitates use of satellite data in a wide range of air quality applications.

Hosted file

`qNO2_final_draft_02-05-24_submit.docx` available at https://authorea.com/users/745260/articles/717412-a-comparison-of-regression-methods-for-inferring-near-surface-no2-with-satellite-data

Hosted file

`qno2_supp_02-05-24_submit.docx` available at https://authorea.com/users/745260/articles/717412-a-comparison-of-regression-methods-for-inferring-near-surface-no2-with-satellite-data

A Comparison of Regression Methods for Inferring Near-Surface NO_2 with Satellite Data

Eliot J. Kim[1*], Tracey Holloway[1,2], Ajinkya Kokandakar[3], Monica Harkey[1], Stephanie Elkins[4], Daniel L. Goldberg[5], Colleen Heck[1]

[1]Nelson Institute Center for Sustainability and the Global Environment (SAGE), University of Wisconsin-Madison, Madison, WI, USA.

[2]Department of Atmospheric and Oceanic Science, University of Wisconsin-Madison, Madison, WI, USA.

[3]Department of Statistics, University of Wisconsin-Madison, Madison, WI, USA.

[4]Department of Earth, Atmospheric, and Planetary Sciences, Massachusetts Institute of Technology, Cambridge, MA, USA.

[5]Department of Environmental and Occupational Health, George Washington University, Washington, D.C., USA.

Corresponding author: Eliot J. Kim (ejkim23@wisc.edu)

*now with Global Modeling and Assimilation Office, NASA Goddard Space Flight Center, Greenbelt, MD, USA.

Key Points:

- We compare regression methods to estimate surface nitrogen dioxide concentrations at 0.01^{o} resolution using satellite and land use data.
- Multivariate linear regression with Anscombe-transformed inputs has strongest agreement with surface nitrogen dioxide measurements.
- Regression methods provide accurate, low-bias concentration estimates with minimal computational and data requirements.

Abstract

Nitrogen dioxide (NO_2) is emitted during high temperature combustion from anthropogenic and natural sources. Human exposure to high NO_2 concentrations causes cardiovascular and respiratory illnesses. The EPA operates ground monitors across the U.S. which take hourly measurements of NO_2 concentrations, providing precise measurements for assessing human pollution exposure but with sparse spatial distribution. Satellite-based instruments capture NO_2 amounts through the atmospheric column with global coverage at regular spatial resolution, but do not directly measure surface NO_2. This study compares regression methods using satellite NO_2 data from the TROPospheric Ozone Monitoring Instrument (TROPOMI) to estimate annual surface NO_2 concentrations in varying geographic and land use settings across the continental U.S. We then apply the best-performing regression models to estimate surface NO_2 at 0.01^o by 0.01^o resolution, and we term this estimate as quasi-NO_2 (qNO2). qNO2 agrees best with measurements at suburban sites (cross-validation (CV) $R^2 = 0.72$) and away from major roads (CV $R^2 = 0.75$). Among U.S. regions, qNO2 agrees best with measurements in the Midwest (CV $R^2 = 0.89$) and agrees least in the Southwest (CV $R^2 = 0.65$). To account for the non-Gaussian distribution of TROPOMI NO_2, we apply data transforms, with the Anscombe transform yielding highest agreement across the continental U.S. (CV $R^2 = 0.77$). The interpretability, minimal computational cost, and health relevance of qNO2 facilitates use of satellite data in a wide range of air quality applications.

Plain Language Summary

Nitrogen dioxide (NO_2) is an air pollutant which causes cardiovascular and respiratory illnesses and reacts in the atmosphere to form other harmful pollutants. This necessitates accurate and reliable quantification of NO_2 concentrations in the air. Ground monitors directly observe NO_2 concentrations near the Earth's surface. However, monitors do not have sufficient spatial coverage to quantify NO_2 at large scales. Satellite-based instruments capture NO_2 amounts across the Earth at increasingly high spatial resolution. However, satellite instruments cannot directly observe surface NO_2 concentrations. In this study, we compare regression methods for estimating surface NO_2 over the continental U.S. using satellite data and auxiliary land-use variables. We find that NO_2 estimated using multivariate regression models with transforms applied to inputs result in the highest agreement with surface NO_2 among the regression methods we investigated. We then use the regression models to quantify surface NO_2 concentration across the U.S. at 0.01° by 0.01° spatial resolution. Our work leverages the precision of ground observations and the high resolution of satellite data to accurately quantify surface NO_2. The interpretable, generalizable, and easily applicable methods used in our study will facilitate the use of satellite data for air quality and human health assessments.

1 Introduction

1.1 Background

Nitrogen dioxide (NO_2) is an air pollutant with harmful impacts on human health. Exposure to high concentrations of NO_2 is closely associated with hospital admissions and mortality for a range of respiratory and cardiovascular diseases (Mills et al., 2015). NO_2 pollution accounts for a significant portion of asthma cases among children worldwide (Anenberg et al., 2022; Chowdhury et al., 2021). Given these health effects, NO_2 is regulated by the United States Environmental Protection Agency (EPA) under the National Ambient Air Quality Standards (NAAQS), which requires the annual mean concentration of NO_2 to remain below 53 parts per billion (ppb) in inhabited areas.

In addition to directly harming human health, NO_2 acts as a reactant in the troposphere to form other harmful air pollutants. In the presence of volatile organic compounds (VOCs) and sunlight, NO_2 reacts to form tropospheric ozone (O_3) which in turn damages human health, increases mortality, and harms ecosystems (Ashmore, 2005; Jerrett et al., 2009; Sillman, 1999). NO_2 also contributes to the formation of particulate nitrate (NO_3^-), a component of fine particulate matter ($PM_{2.5}$) which causes cardiovascular, respiratory, and birth-related illnesses and impairments (Behera & Sharma, 2012; Feng et al., 2016).

NO_2 is emitted from both anthropogenic and natural sources, mainly through high temperature combustion from biomass burning and fossil fuels (M. Lee et al., 1997). Thus, NO_2 serves as a tracer for air pollution from traffic, industrial sites, and other point sources. NO_2 is therefore important for estimating emissions of greenhouse gases that are co-emitted during combustion, such as CO_2 (Goldberg et al., 2019; Konovalov et al., 2016; Levy et al., 2014). Anthropogenic activity is the dominant source of NO_2 in industrialized North America, Europe, and Asia (van der A et al., 2008). Natural sources of NO_2 include soils and lightning (Olivier et al., 1998). Because NO_2 has a relatively short lifetime of several hours, it remains concentrated near its source, resulting in distinct spatial gradients in concentration that are strongly correlated to emissions (L. N. Lamsal et al., 2011; Pommier, 2023). Thus, reliable quantification of NO_2 concentration is critical for characterizing emissions from human activity and for measuring human air pollutant exposure in urban, roadside, and industrial areas with high NO_2 concentrations.

The EPA maintains a national network of ground-based monitors that provide ambient air pollution data known as the Air Quality System (AQS).[1] Although AQS monitors provide hourly measurements of NO_2 concentrations, their sparse and irregular spatial distribution renders them insufficient for capturing the spatiotemporal variability of NO_2 at regional and national scales. Ground monitors have limited usefulness for comprehensive assessments of human exposure to air pollution (Guay et al., 2011).

Satellite data products provide global coverage of column NO_2 on a high-resolution spatial grid, but have daily frequency as opposed to hourly ground measurements. Satellite data offer the potential to bridge the spatial gaps in ground-based monitor data for capturing surface NO_2 concentrations (Holloway et al., 2021). However, satellites do not directly measure NO_2 at the surface and instead detect NO_2 amounts through the atmospheric column with greater sensitivity to mid-tropospheric background NO_2 (Dang et al., 2023). Since NO_2 sources are concentrated at the surface, NO_2 vertical column densities (VCD) measured by satellites have varying strengths of correlation with surface NO_2 depending on spatiotemporal scale, season, region, and the characteristics of the surface and satellite data (van der A et al., 2008; Bechle et al., 2013; Goldberg et al., 2021; Griffin et al., 2019, 2021; Ialongo et al., 2020; Judd et al., 2020; L. N. Lamsal et al., 2008; Lamsal et al., 2015; H. J. Lee et al., 2023; Penn & Holloway, 2020; Pommier, 2023; Yu & Li, 2022).

The highest resolution global satellite NO_2 data currently available comes from the TROPOspheric Monitoring Instrument (TROPOMI) onboard the Sentinel-5 Precursor satellite, launched by the European Space Agency (ESA) in October 2017 (van Geffen et al., 2020;

[1] https://www.epa.gov/system/files/documents/2022-08/aqs_user_guide.pdf

Veefkind et al., 2012). TROPOMI follows a lineage of remote sensing spectrometers including the Global Ozone Monitoring Experiment (GOME), the Scanning Image Spectrometer for Atmospheric Chartography (SCIAMACHY), and the Ozone Monitoring Instrument (OMI). TROPOMI provides column NO_2 data at a peak resolution of 3.5 km by 5.5 km at nadir, a significant improvement over the 13.0 km by 24.0 km peak resolution of the OMI NO_2 data product (Veefkind et al., 2012). The smaller pixel size of TROPOMI enables an unprecedented scale of observation, such as distinguishing signals from individual sources at the scale of individual cities (Ialongo et al., 2020). By capturing the spatial heterogeneities in NO_2 at a finer scale, TROPOMI provides opportunities for significant improvements in satellite-based quantification of surface NO_2.

1.2 Literature Review

To best leverage the global coverage and high spatial resolution of satellite NO_2 data, it is critical to investigate the agreement between column NO_2 amounts and surface NO_2 concentrations across varying spatiotemporal scales. As detailed below, prior studies have utilized chemical transport models, statistical methods, and machine learning to investigate satellite column NO_2 to estimate surface NO_2 at daily to annual time scales and site-specific to global spatial scales.

Vertical profiles of mixing ratios from chemical transport models (CTM) have been used to derive surface NO_2 concentrations from satellite data. Commonly used CTMs include the global three-dimensional Goddard Earth Observing System-Chemistry (GEOS-Chem) model and the regional-scale Community Multi-Scale Air Quality (CMAQ) model and Comprehensive Air Quality Model with extensions (CAMx) (Bechle et al., 2013; Gu et al., 2017; Lamsal et al., 2015). Cooper et al. (2020) applied GEOS-Chem vertical profiles to both OMI and TROPOMI column NO_2 to correct for inaccuracies in vertical mixing assumptions in satellite products. Their work showed that TROPOMI-derived surface NO_2 had lower variance and greater ability to capture emissions sources at high resolution than OMI-derived surface NO_2. Gu et al. (2017) compared ground monitor NO_2 in China with both unadjusted OMI NO_2 and OMI surface NO_2 derived using CMAQ NO_2 profiles. Using the CMAQ-adjusted OMI NO_2, they found 0.03 greater correlation coefficients (R) for January 2014 and 0.05 greater R-values for July 2014. However, the use of chemical transport models in near-real-time requires meteorological reanalysis data and emissions inventories, significant computational resources, and additional re-gridding steps to accommodate for the lower spatial resolution of models. In this study, we use TROPOMI column NO_2 without CTM-based adjustments, to provide surface NO_2 estimates with minimal computational burden.

Machine learning methods are well-suited for estimation and prediction problems with complex input datasets, as is the case for air quality estimation and forecasting. Several recent studies implement machine learning methods using TROPOMI column NO_2 as well as meteorological and land use data inputs to estimate surface NO_2 concentrations (Chi et al., 2021; Grzybowski et al., 2023; Li et al., 2022; Qin et al., 2020). Ghahremanloo et al. (2021) trained convolutional neural networks to predict surface NO_2 concentrations over Texas using TROPOMI column NO_2, vegetation, land-use, and meteorological data as inputs. Their machine learning method improved (R = 0.91) had stronger agreement with surface NO_2 than multiple linear regression (R = 0.77). Kim et al. (2021) used tree-based ensemble machine learning methods with TROPOMI NO_2, land-use, meteorological, and topographic variables to predict hourly surface NO_2 over Switzerland and northern Italy. Their model achieved R^2 of 0.54 for monitors held-out from model training and R^2 of 0.84 for all monitors. Chan et al. (2021) estimated surface NO_2 concentrations over Germany for 2018 through 2020 at weekly to seasonal time-scales using artificial neural networks and TROPOMI NO_2 reprojected to 0.5 by 0.5 km resolution, resulting in R^2 of 0.64. These studies demonstrate that machine learning models can accurately estimate surface NO_2 from large, multi-dimensional input data sets. However, the usability of machine learning models is limited by their significant computational demands and their inherent lack of interpretability. Here, we investigate the ability of regression models with column NO_2 input to estimate surface NO_2. Regression models have minimal computational demands and are straightforward to interpret, enabling a broad range of applications.

Previous regression-based studies have shown strong agreement between surface measurements and TROPOMI column NO_2. Griffin et al. (2019) compared TROPOMI NO_2 and surface measurements in the Canadian Oil Sands and found an R^2 of 0.67, demonstrating the improved capability of TROPOMI in capturing fine-scale surface NO_2 variations compared to OMI. Yu and Li (Yu & Li, 2022) explored the agreement of TROPOMI NO_2 with surface monitors in China's Xinjiang Province, finding a province-wide R^2 of 0.78. Their work also explored meteorological and economic factors, using annual GDP of industry as a proxy for industrial activity. Goldberg et al. (2021) investigated the weekly and diurnal variability of TROPOMI NO_2 as well as the impact of temperature. Their study found an R^2 of 0.66 between annual-average EPA surface NO_2 and TROPOMI column NO_2 across the continental U.S.

Land use regression (LUR) studies incorporate land use and road data to estimate surface NO_2. Early literature conducted seasonal to annual-scale measurement campaigns of surface NO_2 to generate data for LUR. These studies improved on spatial interpolation methods while achieving particularly strong performance in urban areas with fine-scale gradients in NO_2 concentrations (Beelen et al., 2013; Henderson et al., 2007; Hoek et al., 2008). Novotny et al. included OMI-derived surface NO_2 as input to LUR models, resulting in a median R^2 of 0.76 on the hold-out set of monitors over the continental U.S. (Beelen et al., 2013; Henderson et al., 2007; Hoek et al., 2008; Novotny et al., 2011). Lee et al. (2023)used multivariate regression, land use data, and TROPOMI column NO_2 to estimate NO_2 across 89 monitor sites in California, attaining an R^2 of 0.76. Further, Lee et al. estimated surface NO_2 on a 500 meter-resolution grid across California.

Spatial statistical methods including kriging, geographically and temporally weighted regression (GTWR), and fuzzy models have been used to estimate ground-level NO_2 concentrations (Yeganeh et al., 2018, p. 201). Kriging applied to NO_2 ground monitors provides adequate performance in areas with clustered monitors, and incorporating satellite NO_2 data improves prediction at locations far from monitors (Young et al., 2016, p. 201). GTWR improved on ordinary least squares (OLS) regression for predicting ground-level NO_2, with a cross-validation R^2 of 0.60 for GTWR compared to 0.44 for OLS at a daily scale over central and eastern China (Qin et al., 2017). These statistical models provide accurate predictions of surface NO_2 but require the inclusion of chemical transport model profiles, meteorological data, and other information which are not readily available in a real-time prediction context.

Here, we investigate TROPOMI NO_2 to capture spatial heterogeneities in the distribution of ambient NO_2 at the surface across the continental U.S. We compare regression methods for estimating surface NO_2 concentrations in varied land use settings (urban/suburban/rural and highway proximity) and geographies (seven distinct U.S. regions). We then apply the regression models to provide a reliable metric of surface NO_2 across CONUS. This metric provides an easily interpretable, high-resolution estimate of surface NO_2 with minimal data and computational requirements. Recognizing the limitations of an annual average metric, we term this quantity “quasi-NO2” (qNO2 for short). We assess the performance of this metric on regional and national scales, investigate spatial patterns and potential causes of biases, and evaluate the applicability of qNO2 across different use cases. We anticipate qNO2, with its high spatial resolution and ease-of-use, will facilitate air quality and health impact assessments.

2 Methods

2.1 Surface Monitor Data

Hourly NO_2 measurements over the U.S. were obtained from the EPA Air Quality Service (AQS) for 2019 (US EPA, 2013). AQS monitors use a chemiluminescence method which measures the amount of NO that is converted from NO_2 by a molybdenum oxide converter (Fontijn et al., 1970). Other oxidized nitrogen compounds such as nitric acid (HNO_3) and peroxyacetyl nitrate (PAN) are also converted to NO by these converters, causing an overestimation of NO_2 when there are high concentrations of HNO_3 or PAN (Steinbacher et al., 2007). Interference is observed to be highest during afternoon hours for urban areas and in the summer season for rural areas (Dunlea et al., 2007; Steinbacher et al., 2007). This positive monitor bias is often corrected when used in comparison with satellite data (Cooper et al., 2020; Lamsal et al., 2015). Following the reasoning previously described in Penn and Holloway, and given the annual scale of our analysis, we do not apply a bias correction factor to the monitor data (Penn & Holloway, 2020). EPA NO_2 is used without bias corrections for many health impacts studies and regulatory purposes, such as determining attainment of the NAAQS across the nation. To remain consistent with the US air quality management community, we use the monitor data without bias correction.

EPA NO_2 data were filtered to only include monitors for which at least 75% of 2019 hourly measurements were considered "valid" by EPA quality control checks. Then, for each of the remaining 402 monitors, all valid 2019 hourly measurements were averaged to obtain the final "ground-truth" dataset. Our filtering method aligns with the criterion implemented in prior annual average NO_2 studies (Novotny et al., 2011; Penn & Holloway, 2020).

We use two monitor classifications provided by the EPA as input variables for regression modeling: “location setting”, which consists of urban (n=152), suburban (n=146), and rural (n=104) classes, and "road proximity", which has non-near-road (n=333) and near-road (n=69) classes. We use the term "location setting" rather than "land use" because our classification scheme is more general than traditional land use datasets. Near-road monitors are located near highways in metropolitan areas. 57% of these monitors are within 20 meters of a highway and 89% are within 30 meters (Watkins, 2016). Our dataset includes 49 near-road monitors in urban areas, 20 near-road monitors in suburban areas, and no near-road monitors in rural areas.

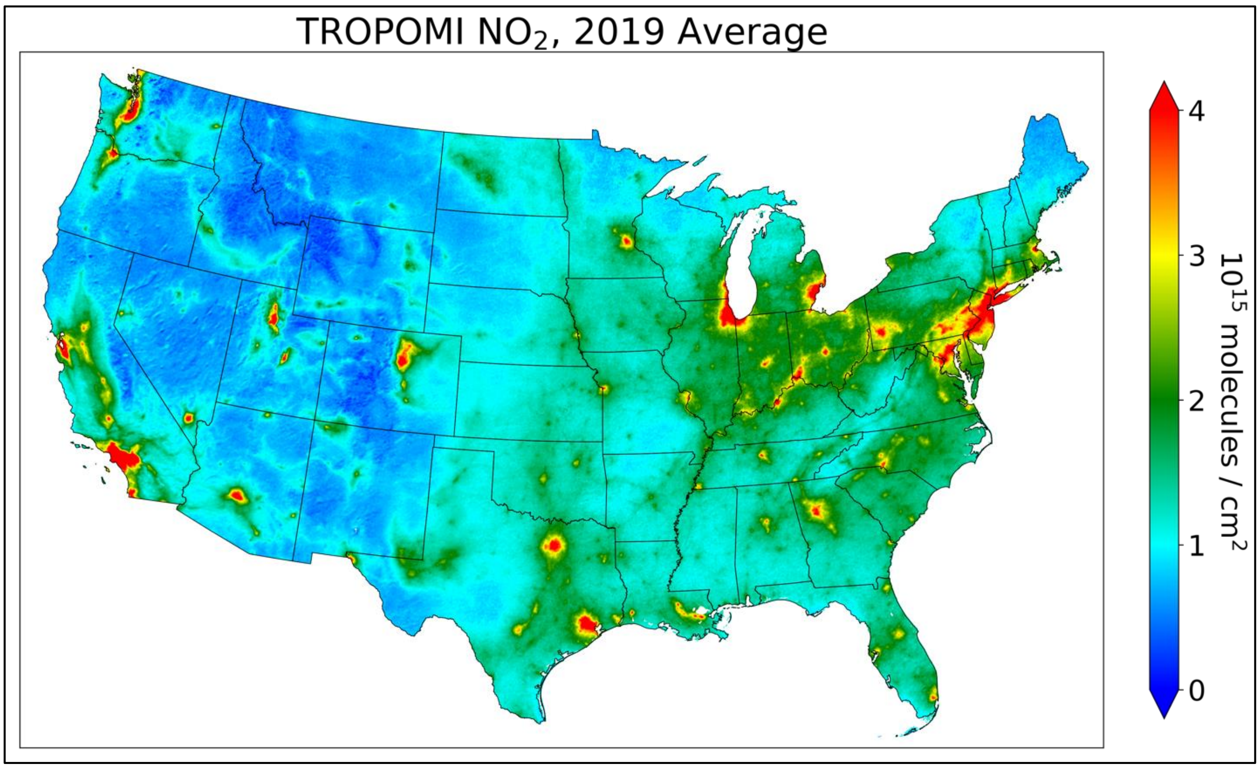


Figure 1: 2019 annual average TROPOMI NO_2 gridded at 0.01^o by 0.01^o resolution across the continental United States.

2.2 Satellite Data

We use 2019 annual average TROPOMI column NO_2 as an input variable for regression models. TROPOMI measures the slant column density (SCD) using a differential optical absorption spectroscopy (DOAS) technique, separating the column into stratospheric and tropospheric components. Air mass factors (AMFs) are then used to convert the SCDs into vertical column densities (VCDs) (van Geffen et al., 2020). Current AMFs are subject to uncertainty and may be a partial cause of low bias in column NO_2 observations in urban areas (Judd et al., 2020). The highest resolution of TROPOMI is 3.5 km by 5.5 km at nadir (resolution increased from 3.5 km by 7.0 km on August 6th, 2019). TROPOMI has an approximate overpass time of 1:30PM local time (Veefkind et al., 2012). We averaged surface measurements for 1-2PM, to match the TROPOMI overpass time, and for the full 24-hours, and found similar correlation between TROPOMI and surface NO_2 for both time ranges. We use 24-hour mean NO_2 monitor measurements in our work, consistent with the method of Lee et al. (2023). We use the method used in Goldberg et al. (2021) to re-grid TROPOMI NO_2 to a 0.01^o by 0.01^o grid (approximately 1 km by 1 km. Figure 1 shows the re-gridded TROPOMI NO_2 data used in our work.

2.3 Road and Location Setting Data

To characterize surface NO_2 concentration across the full domain, we applied our regression models for each TROPOMI NO_2 CONUS grid cell. To apply the regression models, we classified each grid cell by road proximity and location setting, as defined in Section 2.1.

We use road data from the U.S. Census Bureau TIGER/Line Primary Roads dataset.[2] All EPA NO_2 monitors classified as "near-road" are within the same TROPOMI grid cell as a Census Bureau "primary road." Thus, to create the near-road dataset for the full 0.01^o by 0.01^o CONUS grid, TROPOMI NO_2 grid cells overlapping with any segment of a TIGER/Line "primary road" were classified as "near-road." All remaining grid cells were classified as "non-near-road." We used ArcGIS Pro 3.0 to re-grid the TIGER/Line data onto the TROPOMI NO_2 grid. Figure S1 shows primary roads on the 0.01^o by 0.01^o CONUS grid along with near-road EPA monitors.

We determined the location setting classification of each TROPOMI grid cell based on the National Center for Education Statistics (NCES) Education Demographic and Geographic Estimates (EDGE) locale classification[3]. The NCES dataset provides boundaries for four categories of locales across the U.S.: City, Suburban, Town, and Rural. We used ArcGIS Pro 3.0 to determine the locale class with the most area covered in each TROPOMI grid cell. To align

[2]Data available at https://www2.census.gov/geo/tiger/TIGER2021/PRIMARYROADS/.

[3]Accessible at https://nces.ed.gov/programs/edge/Geographic/LocaleBoundaries.

NCES and EPA classifications when applying our regression models across the full CONUS grid, we classified NCES "City" grid cells as EPA "Urban," NCES "Suburban" as EPA "Suburban," and NCES "Rural" and "Town" as EPA "Rural." Figure S2 shows the location setting classifications for each grid cell. Supplemental Text S1 details the agreement between NCES and EPA location setting classifications.

2.4 Regression Methods

We fit simple (SLR) and multivariate linear regression (MLR) models to evaluate the relationship between surface monitor NO_2 and TROPOMI column NO_2, with the output of SLR termed qNO2 SLR and the output of MLR termed qNO2 MLR. Through this analysis, we aim to 1) understand under which conditions NO_2 satellite data best represents surface NO_2 concentrations and 2) compare the performance of different satellite NO_2-based regression methods for estimating surface NO_2. The TROPOMI NO_2 measurements have a Poisson-like distribution (Figure S4). Thus, to satisfy the normality and constant variance assumptions of linear regression, we fit additional regression models with the log transform and Anscombe transform applied to TROPOMI NO_2 inputs. Equation 1 gives the Anscombe transform for positive real number $\boldsymbol{x}$.

$$\boldsymbol{a}(\boldsymbol{x}) = \mathbf{2}\sqrt{\boldsymbol{x} + \frac{\mathbf{3}}{\mathbf{8}}} \#(\mathbf{1})$$

The distribution for the log and Anscombe transformed-TROPOMI NO_2 has greater symmetry than the non-transformed distribution (Figure S4). The Anscombe transform ensures transformed values remain positive, whereas log-transformed values may be negative and thus result in negative regression outputs (Anscombe, 1948). The outputs of MLR with log transform of TROPOMI NO_2 are termed qNO2 logMLR, and the outputs of MLR with Anscombe transform of TROPOMI NO_2 are termed qNO2 anscMLR.

While the resolution of TROPOMI NO_2 is much finer than previous NO_2 satellite data products, we still expect the kilometer-scale data to be insufficient to capture emissions near individual major roads, which can have sharp decay gradients over hundreds of meters (Kimbrough et al., 2017). Thus, we separately conduct simple linear regression on near-road and not-near road monitors. To further compare TROPOMI NO_2 performance over different location settings, we also conduct simple linear regressions on each location setting class: urban, suburban, and rural.

We fit multivariate regression models with three input variables: TROPOMI column NO_2 concentration (no transform, log transform, Anscombe transform), road proximity, and location setting. Road proximity is a binary variable representing EPA monitor "near-road" and "non-near-road" classification. Location setting is a categorical variable with three levels corresponding to EPA monitor classification: urban, suburban, and rural. Multivariate regression provides a single interpretable model for calculating qNO2 across the U.S., facilitating interpretation and application by stakeholders.

2.5 Evaluation Methods

We evaluate regression model performance using four metrics: coefficient of determination (R^2), root mean squared error (RMSE), mean fractional error (MFE), and mean fractional bias (MFB).

$$\boldsymbol{RMSE = \sqrt{\frac{\sum_{i=1}^{N}(qNO2[i] - EPA\ NO_2[i])^2}{N}}}\ \#(\mathbf{2})$$

$$\boldsymbol{MFB = \frac{1}{N}\left(\frac{\sum_{i=1}^{N}(qNO2[i] - EPA\ NO_2[i])}{\sum_{i=1}^{N}\frac{(qNO2[i] + EPA\ NO_2[i])}{2}}\right)}\#(\mathbf{3})$$

$$\boldsymbol{MFE = \frac{1}{N}\left(\frac{\sum_{i=1}^{N}|qNO2[i] - EPA\ NO_2[i]|}{\sum_{i=1}^{N}\frac{(qNO2[i] + EPA\ NO_2[i])}{2}}\right)}\#(\mathbf{4})$$

R^2 is the proportion of variance in the output that is captured by the model. RMSE is a metric of absolute error with the same units as the model output (ppb), calculated using Equation 1. MFB is a unitless metric of relative bias. For instance, MFB = 0.67 indicates that the model output is an overestimate of observed values by a factor of 2. MFB = 0.4 indicates that the model output is an overestimate by a factor of 1.5. MFE is a unitless metric of relative error. MFB and MFE are used to measure the relative performance of qNO2, enabling comparison between models fit on different classes and regions.

The multivariate regression models were evaluated for seven distinct regions of the continental U.S.: Northeast, Southeast, Midwest, Rockies, Southwest, Northwest, and Southern California. These regions were selected due to their distinct topographical and meteorological conditions, and to ensure a similar number of EPA monitors (n=51 ~ 63) in each region. Figure S5 shows region divisions and monitor locations.

We implement random and spatial cross-validation methods to assess the generalization ability of the multivariate regression models. Generalizability is an important factor for the utilization of qNO2 as a near-real-time metric for surface NO_2 in spatial and temporal domains beyond those evaluated in this work. We conduct random cross-validation using k-fold and Monte Carlo. We also use each of the seven regions as cross-validation "folds" to conduct spatial cross-validation. Cross-validation experiments are described in greater detail in Supplemental Text S2.

After model training and evaluation, we compute qNO2 MLR, logMLR, and anscMLR for the full CONUS TROPOMI NO_2 grid and discuss the spatial variation of qNO2 values and qNO2 performance metrics across road proximity and location setting classes as well as U.S. regions. Additional analyses are presented for three metropolitan areas with some of the highest TROPOMI NO_2 levels in the United States: Los Angeles, Dallas-Fort Worth, and New York City.

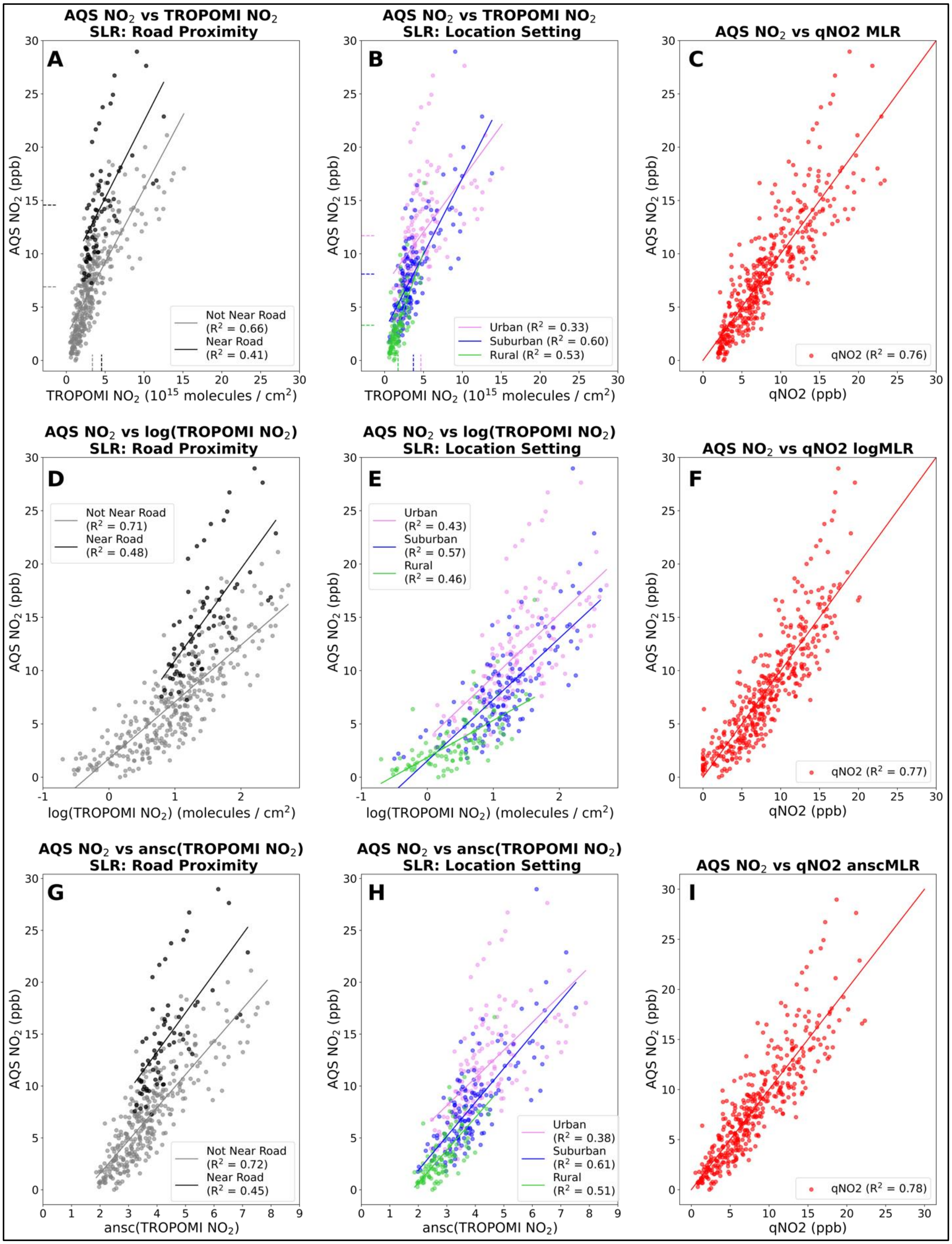

AQS NO2 vs TROPOMI NO2
SLR: Road Proximity
A
Not Near Road (R2 = 0.66)
Near Road (R2 = 0.41)
AQS NO2 (ppb)
TROPOMI NO2 (1015 molecules / cm2)
AQS NO2 vs TROPOMI NO2
SLR: Location Setting
B
Urban (R2 = 0.33)
Suburban (R2 = 0.60)
Rural (R2 = 0.53)
AQS NO2 vs qNO2 MLR
C
qNO2 (R2 = 0.76)
qNO2 (ppb)
AQS NO2 vs log(TROPOMI NO2)
SLR: Road Proximity
D
Not Near Road (R2 = 0.71)
Near Road (R2 = 0.48)
log(TROPOMI NO2) (molecules / cm2)
AQS NO2 vs log(TROPOMI NO2)
SLR: Location Setting
E
Urban (R2 = 0.43)
Suburban (R2 = 0.57)
Rural (R2 = 0.46)
AQS NO2 vs qNO2 logMLR
F
qNO2 (R2 = 0.77)
AQS NO2 vs ansc(TROPOMI NO2)
SLR: Road Proximity
G
Not Near Road (R2 = 0.72)
Near Road (R2 = 0.45)
ansc(TROPOMI NO2)
AQS NO2 vs ansc(TROPOMI NO2)
SLR: Location Setting
H
Urban (R2 = 0.38)
Suburban (R2 = 0.61)
Rural (R2 = 0.51)
AQS NO2 vs qNO2 anscMLR
I
qNO2 (R2 = 0.78)

Figure 2: Simple and multiple linear regression results for TROPOMI NO_2 with no transform, log transform, and Anscombe transform. **a)** Simple regression models trained separately on near-road (black line-of-best-fit) and non-near-road (gray line-of-best-fit) EPA NO_2 monitors. Dashed lines indicate mean near-road and non-near-road TROPOMI and EPA NO_2 values. **b)** Simple regression models trained separately on urban (pink line-of-best-fit), suburban (blue line-of-best-fit), and rural (green line-of-best-fit) EPA NO_2 monitors. Dashed lines indicate mean urban, rural, and suburban TROPOMI and EPA NO_2 values. **c)** qNO2 MLR trained on all monitors. y = x line in red. **d)** Simple regression models with log transform of TROPOMI NO_2 input trained separately on near-road and non-near-road EPA NO_2 monitors. **e)** Simple regression models with log transform of TROPOMI NO_2 input trained separately on urban, suburban, and rural EPA NO_2 monitors. **f)** qNO2 logMLR trained on all monitors. y = x line in red. **g)** Simple regression models with Anscombe transform of TROPOMI NO_2 input trained separately on near-road and non-near-road EPA NO_2 monitors. **h)** Simple regression models with Anscombe transform of TROPOMI NO_2 input trained separately on urban, suburban, and rural EPA NO_2 monitors. **i)** qNO2 anscMLR trained on all monitors. y = x line in red.

3 Results and Discussion

3.1 Regression Results

We present SLR and MLR results for surface NO_2 estimation, their relative performance across U.S. regions, and the impact of transforms applied to TROPOMI NO_2. Figure 2 shows the relationship between TROPOMI NO_2 and EPA NO_2, separated by road proximity and location setting classes. SLR with TROPOMI NO_2 as the sole input resulted in an R^2 of 0.55 when evaluated over all monitors. We fit separate SLR models for near-road and non-near-road monitors (Figure 2a). SLR with TROPOMI NO_2 captures the majority of variance in surface NO_2 concentrations at non-near-road monitors (R^2 = 0.66) but does not fully capture near-road variation (R^2 = 0.41). Surface monitors better detect NO_2 near major roads compared to TROPOMI NO_2 because the kilometer-scale resolution of TROPOMI cannot fully capture fine-scale NO_2 concentration gradients. However, while SLR at near-road sites has higher absolute error than non-near-road sites, fractional error and bias is lower at near-road sites than non-near-road sites. Thus, SLR with TROPOMI NO_2 can be useful as a nearly unbiased estimate in data-sparse settings near major roads. To account for the difference in performance between near-road and non-near-road sites, we include road proximity as a binary variable in the MLR models, aligning with several prior studies which include road proximity information in satellite NO_2-based statistical models to estimate surface NO_2 (Grzybowski et al., 2023; Henderson et al., 2007; Kim et al., 2021; H. J. Lee et al., 2023; Novotny et al., 2011; Yeganeh et al., 2018; Young et al., 2016). We display performance metrics for SLR and MLR models in Table S1.

In addition to classification by proximity to major roads, we separated monitor sites by their location setting (urban, suburban, rural) and fit SLR models to each class. We found TROPOMI NO_2 best captures surface concentrations at suburban sites (R^2 = 0.60, RMSE = 2.77 ppb), captures around half of concentration variance at rural sites (R^2 = 0.53, RMSE = 1.80 ppb), and has the poorest performance at urban sites (R^2 = 0.33, RMSE = 3.95 ppb) (Figure 2b). Column NO_2 does not fully capture surface NO_2 concentration peaks in urban areas but has stronger performance in suburban and rural areas, which have lower and more uniform NO_2 concentrations. However, urban and suburban sites have lower relative error and bias than rural sites. SLR with TROPOMI NO_2 is useful as a low bias estimate of urban and suburban surface NO_2. In rural areas, SLR-based estimates have moderate positive bias. To account for the differing performance of column NO_2 in capturing surface concentrations across location settings, our MLR models include location setting as a categorical variable.

We then fit a multiple linear regression (MLR) model with TROPOMI NO_2, road proximity, and location setting variables as inputs and surface NO_2 concentration estimates as the output. MLR on all monitor sites results in an R^2 of 0.76, greater than the full-domain SLR R^2 of 0.55. Thus, incorporating road proximity and location setting information aids in surface NO_2 estimation. MLR also has lower absolute (RMSE = 2.58 ppb) and fractional error (MFE = 0.29) than SLR (RMSE = 3.59 ppb, MFE = 0.40) and results in a lower positive bias (MFB = 0.09) than SLR (0.15). Thus, in locations with readily available data on major roads and basic land use classifications, we recommend the use of the MLR model for surface NO_2 estimation. Table S2 shows coefficients for the SLR models of each site classification and for the MLR model. Supplemental Text S3 includes additional analysis of regression coefficients. As noted in Section 2.4, we term the surface NO_2 estimates of the MLR model as qNO2 MLR.

2019 annual average TROPOMI NO_2 amounts over CONUS have a log-normal distribution (Figure S4). To better satisfy the assumption of normality in regression and to improve regression performance, we applied a log-transform and Anscombe transform to TROPOMI NO_2 and compared performance with the corresponding no-transform models. For SLR, both log

(Figure 2d) and Anscombe (Figure 2g) transformed-TROPOMI NO_2 have greater R^2 than no-transform SLR when fit on all sites. For MLR, both transforms resulted in marginal improvements in performance (Figures 2f,i). Performance metrics and regression coefficients for log-transform models are presented in Tables S3 and S4, respectively. Performance metrics and regression coefficients for Anscombe-transform models are presented in Tables S5 and S6, respectively. Following the naming convention defined in Section 2.4, we term the output of the MLR with log transform as qNO2 logMLR and the output of MLR with Anscombe transform qNO2 as anscMLR.

We specify the model configuration with the best surface NO_2 estimation performance for each site classification. For rural sites, SLR with no transform has the highest R^2 and lowest RMSE, but SLR with log-transform has the lowest rural fractional bias. SLR with log transform has the highest R^2, lowest RMSE, and lowest fractional bias for near-road sites. For urban sites, MLR with log transform has the highest R^2. The difference in performance between MLR and SLR is greatest at urban sites, which indicates the value of road proximity information for estimating urban surface NO_2. For non-near-road and suburban sites, MLR with Anscombe transform has the best performance. MLR with Anscombe transform has the best performance over all monitors, with an overall R^2 of 0.78.

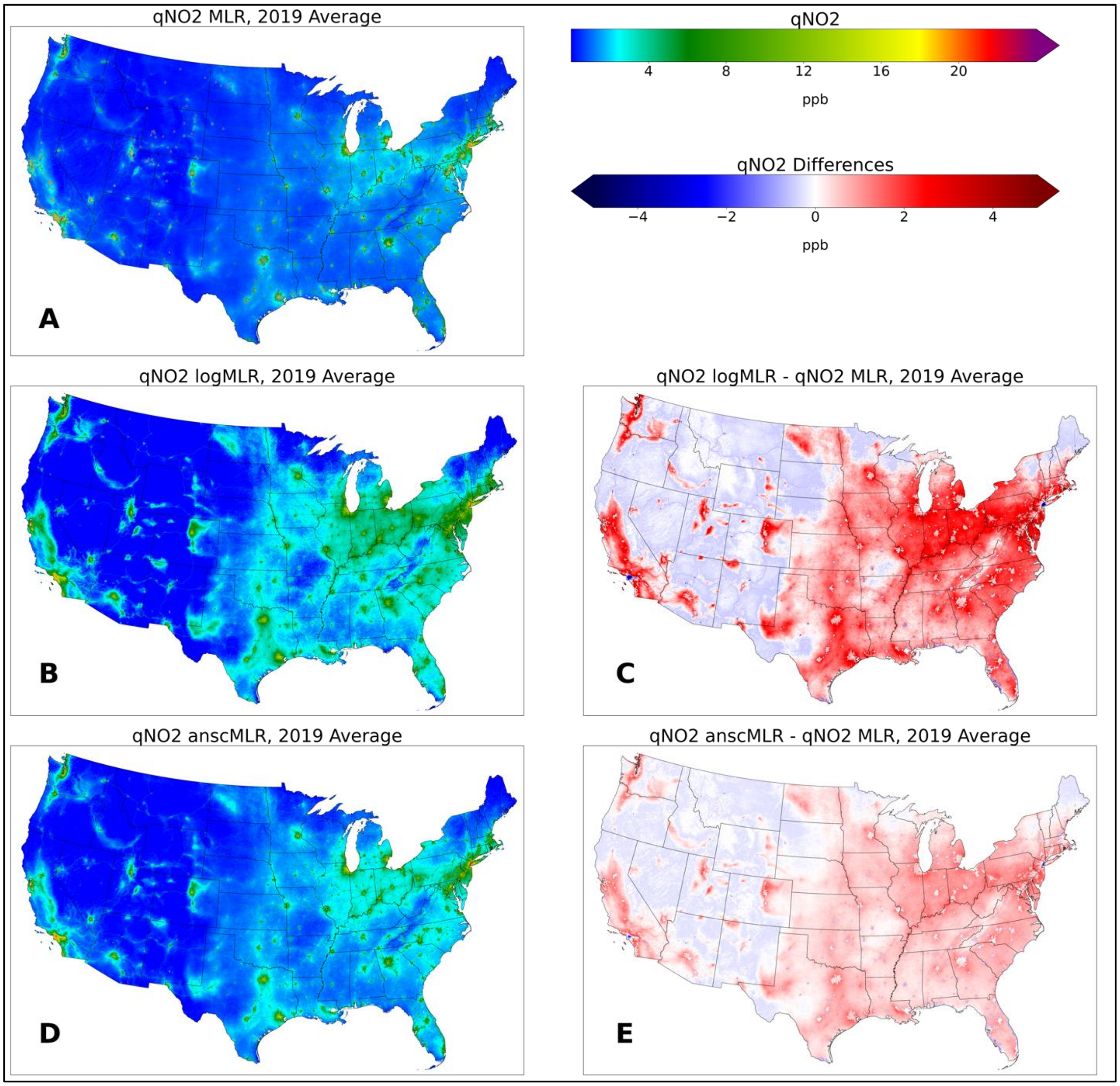
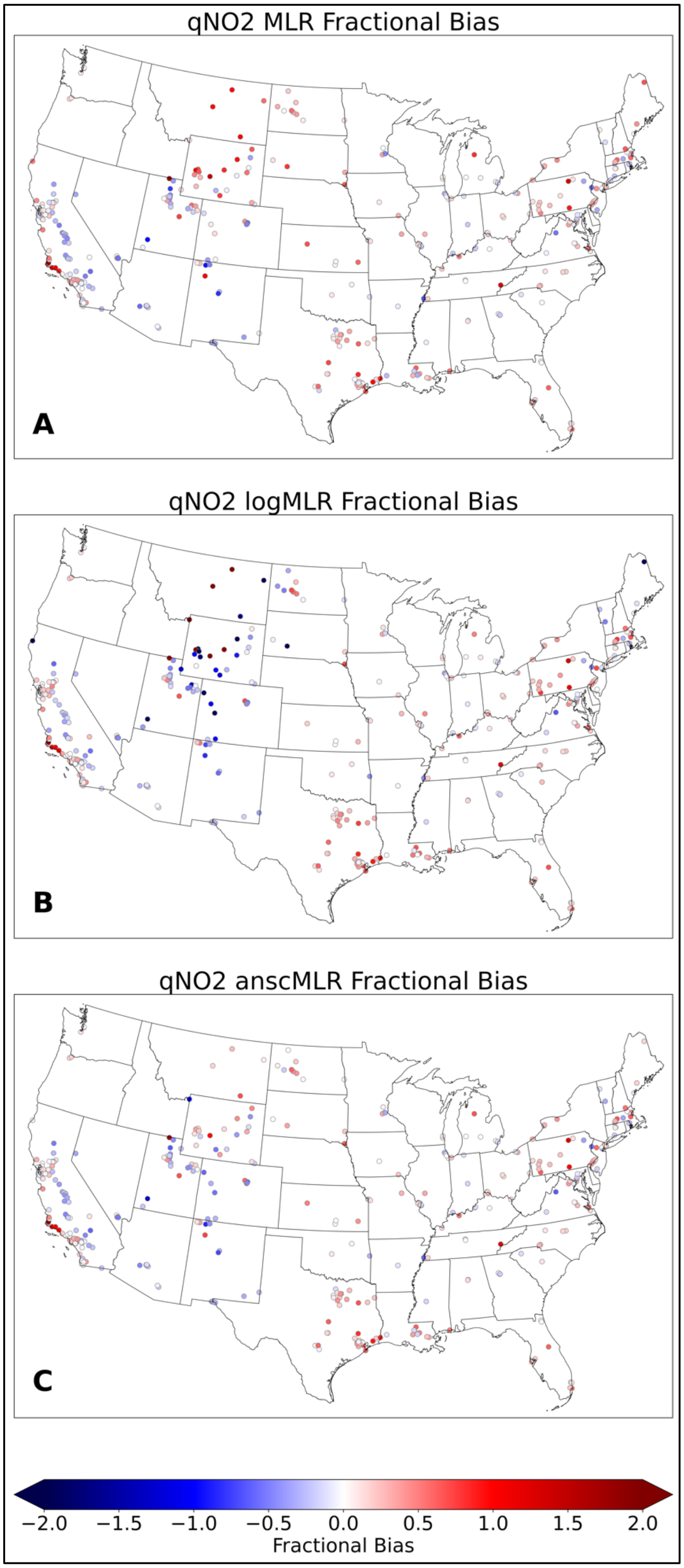


Figure 3: **a)** 2019 qNO2 gridded across the continental United States, computed using multiple linear regression (qNO2 MLR). **b)** 2019 qNO2 gridded across the continental United States, computed using multiple linear regression with log transform of the TROPOMI NO_2 input (qNO2 logMLR). **c)** The difference between qNO2 logMLR and qNO2 MLR. Red indicates areas where qNO2 logMLR is greater than qNO2 MLR, and blue indicates areas where qNO2 logMLR is less than qNO2 MLR. **d)** 2019 qNO2 gridded across the continental United States, computed using multiple linear regression with Anscombe transform of TROPOMI NO_2 input (qNO2 anscMLR). **e)** The difference between qNO2 anscMLR and qNO2 MLR. Red indicates areas where qNO2 logMLR is greater than qNO2 MLR, and blue indicates areas where qNO2 logMLR is less than qNO2 MLR.

3.2 qNO2 Computation

To analyze spatial patterns of surface NO_2 estimates, we computed qNO2 MLR for all TROPOMI 0.01^o by 0.01^o grid cells over CONUS, displayed in Figure 3a. qNO2 MLR is highest in major cities and along major highways across the U.S. The Great Lakes and much of the eastern half of the U.S. have high overall qNO2 MLR concentrations, while the Mountain West and Northern New England have lower overall concentrations. Western North Dakota and the Permian Basin in western Texas have elevated qNO2 MLR levels compared to the surrounding rural areas, coinciding with the high oil industry activity in both regions.

We also computed qNO2 logMLR (Figure 3b) and anscMLR (Figure 3d) at 0.01^o by 0.01^o resolution across CONUS. Figure 3c displays the difference between qNO2 logMLR and qNO2 MLR for each grid cell. qNO2 logMLR is greater than qNO2 MLR across the eastern half of the United States, particularly around the Great Lakes, Texas, and the Mid-Atlantic. qNO2 logMLR is also greater than qNO2 MLR in the California Central Valley and in areas around Seattle, Portland, Salt Lake City, Phoenix, and Denver, as well as the Bakken oil fields in North Dakota and Permian Basin in Texas. qNO2 logMLR and qNO2 MLR are close in value in most urban areas and throughout most of the rural western U.S. qNO2 logMLR and MLR have the greatest difference in the Los Angeles and New York City areas, where qNO2 logMLR concentrations are more than 4 ppb lower than qNO2 MLR. Figure 3e shows the difference between qNO2 anscMLR and qNO2 MLR for each grid cell. qNO2 anscMLR follows a similar spatial pattern of differences to qNO2 MLR as qNO2 logMLR, but with a lower magnitude of difference. Overall, qNO2 logMLR and anscMLR have greater spatial spread of NO_2 from urban areas and greater background concentrations in the eastern U.S. as well as lower maximum concentrations compared to qNO2 MLR.


qNO2 MLR Fractional Bias
A
qNO2 logMLR Fractional Bias
B
qNO2 anscMLR Fractional Bias
C
−2.0
−1.5
−1.0
−0.5
0.0
0.5
1.0
1.5
2.0
Fractional Bias

453

Figure 4: **a)** Fractional bias between EPA NO_2 and qNO2 MLR at EPA monitor locations (n=402) across the continental United States. Red indicates monitor locations where qNO2 is relatively high compared to the measured NO_2 concentration. Blue indicates monitor locations where qNO2 is relatively low compared to the measured NO_2 concentration. **b)** Fractional bias between EPA NO_2 and qNO2 logMLR at EPA monitor locations. **c)** Fractional bias between EPA NO_2 and qNO2 anscMLR at EPA monitor locations.

3.3 Regional Evaluation

We evaluated qNO2 in seven U.S. regions to investigate the variability of satellite-surface agreement between large spatial domains with similar topographic and meteorological conditions. qNO2 MLR best aligns with surface NO_2 in the Midwest states (R^2 = 0.88). Northeast, Southeast, Rockies, and Southern California regions have comparable qNO2 MLR performance with R^2 values ranging from 0.72 to 0.76. The Southwest (R^2 = 0.65) and Northwest (R^2 = 0.66) regions have the lowest qNO2 MLR performance (Table S7). The strong performance in the Midwest and relatively weak performance in the Western U.S. suggests that elevation gradient may be an additional variable that could be included to further improve MLR performance.

All regions have positive mean fractional bias except the Northwest, which has an MFB of -0.06 indicating that qNO2 is a slight underestimate of surface NO_2. Rockies region has the greatest MFE (0.46) and MFB (0.15). This may be due to the larger proportion of rural sites in the Rockies region with very low NO_2 concentrations, which inflates relative error metrics. For rural and remote areas with low background NO_2 concentrations, absolute error metrics are more relevant for assessing model performance.

qNO2 logMLR exhibits similar regional variability as qNO2 MLR. R^2 in the Northeast, Midwest, Northwest, and Southern California is slightly higher compared to qNO2 MLR, while R^2 in the Southeast is slightly lower (Table S8). qNO2 anscMLR has slightly higher R^2 than qNO2 MLR and logMLR in all regions (Table S9).

qNO2 performance varies within regions as well as between regions. Figure 4 displays the fractional bias of qNO2 at each EPA monitor. qNO2 MLR (Figure 4a) overestimates surface NO_2 relative to the measured value along the California coast, Wyoming, Montana, the Dakotas, and Texas. qNO2 MLR underestimates surface NO_2 in the California Central Valley and the

Southwest. Sites in the Midwest and Southeast have low overall bias. qNO2 anscMLR (Figure 4c) and qNO2 MLR have similar spatial variation in fractional bias across EPA monitor sites, but qNO2 anscMLR has lower fractional bias in Wyoming and Montana. qNO2 logMLR (Figure 4b) also has similar spatial fractional bias variation as MLR and anscMLR but has a much greater degree of bias in Wyoming and Montana.

3.4 Urban Case Studies

Figure 5 shows qNO2 MLR, logMLR, and anscMLR over three large U.S. metropolitan areas: Los Angeles, CA; Dallas-Fort Worth, TX; and New York City, NY-NJ-CT-PA. qNO2 MLR in Los Angeles (Figure 5a,d,g) is greater than 20 ppb in the city center. qNO2 MLR decreases sharply between the metropolitan area and the surrounding rural areas. qNO2 logMLR has a lower maximum level in the city center and a more gradual decrease towards the surrounding rural areas than qNO2 MLR. The urban-rural concentration gradient for qNO2 anscMLR is steeper than qNO2 logMLR but less steep than qNO2 MLR. All qNO2 models indicate concentrations greater than 18 ppb along the major highways extending south and east from central LA. Among the qNO2 models, qNO2 MLR (RMSE = 3.65 ppb) and anscMLR (3.52 ppb) have the lowest error in Los Angeles, while logMLR (RMSE = 7.78 ppb) has the highest error.

Dallas-Fort Worth (Figure 5b,e,h) has lower overall qNO2 than Los Angeles, with maximum qNO2 of 16 to 18 ppb along major highways. The qNO2 models estimate similar concentration levels in the metropolitan area, but logMLR and anscMLR have a broader radius of high concentrations than qNO2 MLR. In Dallas, qNO2 has high accuracy, with logMLR having the lowest error (RMSE = 1.56 ppb).

New York City (Figure 5c,f,i) has comparable peak qNO2 levels to Los Angeles, as the urban core and adjacent highways have qNO2 concentrations greater than 20 ppb. As in Los Angeles and Dallas-Fort Worth, qNO2 logMLR and anscMLR over New York City have smoother gradients toward the edges of the metropolitan area than qNO2 MLR. The spatial patterns of qNO2 anscMLR are a combination of the sharp gradients and high peak concentrations of qNO2 MLR and the smoother gradients of qNO2 logMLR. Among the qNO2 models, anscMLR results in the lowest error (RMSE = 3.49 ppb) while logMLR has the highest error (RMSE = 7.04 ppb).

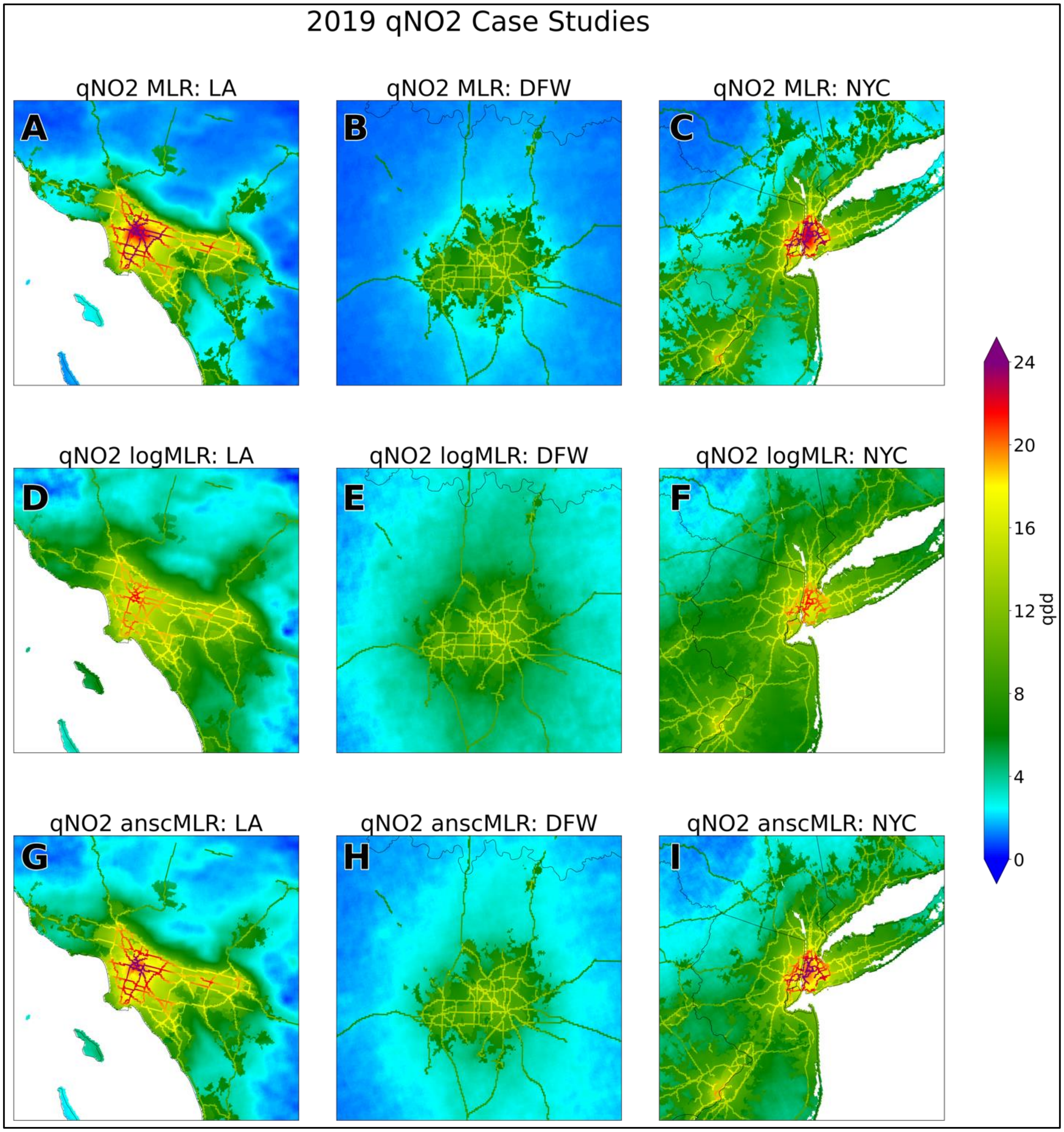


Figure 5: qNO2 MLR, qNO2 logMLR, and qNO2 anscMLR over three selected large U.S. metropolitan areas. **a)** qNO2 MLR over Los Angeles, CA. **b)** qNO2 MLR over Dallas-Fort Worth, TX. **c)** qNO2 MLR over New York-Newark-Jersey City, NY-NJ-CT-PA. **d)** qNO2 logMLR over LA. **e)** qNO2 logMLR over Dallas. **f)** qNO2 logMLR over NYC. **g)** qNO2 anscMLR over LA. **h)** qNO2 anscMLR over Dallas. **i)** qNO2 anscMLR over NYC.

3.5 Cross-Validation

We implemented k-fold and Monte Carlo cross-validation (CV) to investigate the generalizability of qNO2 on data sets held out from model fitting. Table S10 displays k-fold CV results and Table S11 displays Monte Carlo CV results.

Both CV methods indicate that qNO2 anscMLR performs well on unseen data. k-fold CV resulted in similar mean holdout set performance for k = 5 and k =10 with R^2 of 0.74. However, using k = 20 resulted in a mean holdout set performance of R^2 = 0.71. Smaller holdout sets are more likely to be unrepresentative of the population distribution, thus resulting in poor evaluation performance. Monte Carlo CV using holdout set sizes of 25% and 50% indicated strong evaluation performance on the holdout data, with anscMLR R^2 of 0.77. When evaluated over a sufficiently large set of unseen data points, qNO2 anscMLR exhibits strong generalization ability. Further, the difference between holdout set and training set performance is small, indicating that the anscMLR model is not overfit to the training data. This finding supports the use of qNO2 anscMLR as a reliable metric for future surface NO2 estimation beyond the domain of our analysis.

We also conduct cross-validation using the seven CONUS regions by leaving one region out for evaluation and fitting anscMLR models on the remaining six regions. As with non-cross-validated regional evaluation detailed in Section 3.2, qNO2 anscMLR generalizes well to Midwest monitors with an R^2 of 0.89 and has the lowest generalization performance for Southwest (R^2 = 0.65), Pacific Northwest (R^2 = 0.66), and Northeast sites (R^2 = 0.69) (Table S12). The similar results between cross-validated and non-cross-validated region-wise evaluation indicate that qNO2 is generalizable to new geographic contexts.

4 Conclusions

We fit regression models with TROPOMI NO_2, location setting, and road proximity inputs to estimate 2019 annual average surface NO_2 concentrations at 0.01^o by 0.01^o resolution across the continental U.S. Among the regression models studied, qNO2 anscMLR has the strongest overall performance. qNO2 anscMLR is the best estimate for surface NO_2 at non-near-road sites (anscMLR R^2 = 0.76) and suburban sites (anscMLR R^2 = 0.74). We also investigate qNO2 spatial patterns over large U.S. urban areas, compare qNO2 performance across U.S. regions, and assess the generalizability of qNO2. We find that qNO2 performs best in the Midwest, with cross-validated anscMLR R^2 of 0.89.

Using easily accessible data and interpretable methods, we demonstrate comparable or improved performance over prior regression-based studies which use satellite NO_2 to estimate surface NO_2. Novotny et al. (2011) used GEOS-Chem to derive surface NO_2 concentrations from OMI, which was then used as regression input along with land use to estimate surface NO_2 at 30-meter resolution. Their work resulted in an R^2 of 0.77 and an MAE of 2.40 ppb evaluated at EPA NO_2 monitors across CONUS. The slightly stronger performance of qNO2 anscMLR using a three-variable regression model without GEOS-Chem-based column NO_2 adjustments highlights the improved ability of the higher resolution TROPOMI to capture surface NO_2 compared to prior satellite products. Goldberg et al. (2021) found an R^2 of 0.66 between TROPOMI NO_2 and EPA NO_2 at non-near-road sites. Using the same 0.01^o by 0.01^o TROPOMI dataset, we apply the Anscombe transform to TROPOMI NO_2 which results in 0.06 greater R^2 at non-near-road sites. Lee et al. (2023) used multivariate regression to analyze TROPOMI NO_2 agreement with 2018-2019 annual-average surface NO_2 over California at 0.5 by 0.5 km resolution. Their final regression models included land use and road proximity inputs. Meteorological inputs were initially considered but were removed because they did not contribute to model performance. Their work achieved an R^2 of 0.76 and RMSE of 2.51 ppb. These metrics are comparable to qNO2 anscMLR metrics computed using the same cross-validation method as Lee et al., for all CONUS monitor sites (CV R^2 = 0.75 and RMSE = 2.64 ppb). We further find that qNO2 anscMLR has cross-validation R^2 of 0.76 and RMSE of 2.63 ppb in California, in close agreement with the Lee et al. results while using simpler input variables.

The regression models in this work can be applied to estimate surface NO_2 in any region with adequate road and location setting data, thus enabling NO_2 exposure assessments in areas with sparse or no monitor coverage. Additionally, since both road density and location setting are relatively static over time, surface NO_2 concentrations of additional years in the TROPOMI record can be estimated quickly. In addition to characterizing surface NO_2, our analysis of satellite-surface agreement across spatial scales, contexts, and model configurations informs the application of satellite NO_2 products in different domains.

The annual-average scale of our analysis is suitable for characterizing long-term spatial trends of surface NO_2 but is less applicable for studies of short-range pollution events and trends. For example, this methodology is less applicable for inferring NO_2 from biomass burning events because of their short-term time scale (less than one week) and tendency to be obscured from satellite measurements since they produce high-density smoke which is often indistinguishable from clouds (Griffin et al., 2021).

The spatial distribution of EPA monitors presents a potential source of bias for qNO2. In the eastern half of the U.S., clusters of monitors are evenly distributed, mainly near urban areas. In the western U.S., monitors are concentrated in rural Wyoming, western North Dakota, and throughout California but are sparse in Washington and Oregon. Thus, monitor measurements may not be fully representative of NO_2 spatial patterns over the U.S., impacting the generalizability of qNO2 to less-represented location settings and regions. Spatial models which account for error correlations between monitors in proximity may help to account for the inconsistent distribution of surface monitors.

We anticipate that the results presented here will inform analysis of data from TEMPO (Tropospheric Emissions: Monitoring of Pollution), a new NASA geostationary satellite instrument launched in April 2023 which captures hourly column NO_2 during all daylight hours at 2.1 km by 4.4 km resolution over the entire continental United States (Zoogman et al., 2017). The greater spatial and temporal resolution from TEMPO will expand the scope of air quality analyses. For example, the methods in this work can be extended to compare hourly TEMPO observations with hourly ground monitor measurements of NO_2. Further, greater spatial resolution will enable investigation of satellite-surface agreement over finer-scale emissions sources such as industrial sites in addition to major roads.

Acknowledgements

EK, TH, MH, CH, and SE were supported by NASA Grant 80NSSC21K0427 for the NASA Health and Air Quality Applied Sciences Team (HAQAST). DG was supported by NASA Grant 80NSSC21K0511 for the NASA Health and Air Quality Applied Sciences Team (HAQAST). EK was also supported by the Hilldale Undergraduate Research Fellowship at the University of Wisconsin-Madison.

Open Research

TROPOMI NO_2 data can be obtained here: 10.5067/MEASURES/MINDS/DATA203. The road data used in our work was derived from shape files accessible at https://www2.census.gov/geo/tiger/TIGER2021/PRIMARYROADS/. Location setting data was obtained from https://nces.ed.gov/programs/edge/Geographic/LocaleBoundaries. The above data re-gridded to the custom 0.01^o by 0.01^o grid used in this work is available at https://doi.org/10.5281/zenodo.10601063. EPA AQS data is accessible at https://aqs.epa.gov/aqsweb/airdata/download_files.html. All code for the analysis and visualizations presented in this study are available at https://doi.org/10.5281/zenodo.10582277.

References

van der A, R. J., Eskes, H. J., Boersma, K. F., van Noije, T. P. C., Van Roozendael, M., De Smedt, I., et al. (2008). Trends, seasonal variability and dominant NOx source derived from a ten year record of NO2 measured from space. *Journal of Geophysical Research: Atmospheres*, *113*(D4). https://doi.org/10.1029/2007JD009021

Anenberg, S. C., Mohegh, A., Goldberg, D. L., Kerr, G. H., Brauer, M., Burkart, K., et al. (2022). Long-term trends in urban NO2 concentrations and associated paediatric asthma incidence: estimates from global datasets. *The Lancet Planetary Health*, *6*(1), e49–e58. https://doi.org/10.1016/S2542-5196(21)00255-2

Anscombe, F. J. (1948). The Transformation of Poisson, Binomial, and Negative-Binomial Data. *Biometrika*, *35*(3–4), 246–254. https://doi.org/10.1093/biomet/35.3-4.246

Ashmore, M. R. (2005). Assessing the future global impacts of ozone on vegetation. *Plant, Cell & Environment*, *28*(8), 949–964. https://doi.org/10.1111/j.1365-3040.2005.01341.x

Bechle, M. J., Millet, D. B., & Marshall, J. D. (2013). Remote sensing of exposure to NO2: Satellite versus ground-based measurement in a large urban area. *Atmospheric Environment*, *69*, 345–353. https://doi.org/10.1016/j.atmosenv.2012.11.046

Beelen, R., Hoek, G., Vienneau, D., Eeftens, M., Dimakopoulou, K., Pedeli, X., et al. (2013). Development of NO2 and NOx land use regression models for estimating air pollution exposure in 36 study areas in Europe – The ESCAPE project. *Atmospheric Environment*, *72*, 10–23. https://doi.org/10.1016/j.atmosenv.2013.02.037

Behera, S. N., & Sharma, M. (2012). Transformation of atmospheric ammonia and acid gases into components of $PM_{2.5}$: an environmental chamber study. *Environmental Science and Pollution Research International*, *19*(4), 1187–1197. https://doi.org/10.1007/s11356-011-0635-9

Chan, K. L., Khorsandi, E., Liu, S., Baier, F., & Valks, P. (2021). Estimation of Surface NO2 Concentrations over Germany from TROPOMI Satellite Observations Using a Machine Learning Method. *Remote Sensing*, *13*(5), 969. https://doi.org/10.3390/rs13050969

Chi, Y., Fan, M., Zhao, C., Sun, L., Yang, Y., Yang, X., & Tao, J. (2021). Ground-level NO2 concentration estimation based on OMI tropospheric NO2 and its spatiotemporal characteristics in typical regions of China. *Atmospheric Research*, *264*, 105821. https://doi.org/10.1016/j.atmosres.2021.105821

Chowdhury, S., Haines, A., Klingmüller, K., Kumar, V., Pozzer, A., Venkataraman, C., et al. (2021). Global and national assessment of the incidence of asthma in children and adolescents from major sources of ambient NO2. *Environmental Research Letters*, *16*(3), 035020. https://doi.org/10.1088/1748-9326/abe909

Cooper, M. J., Martin, R. V., McLinden, C. A., & Brook, J. R. (2020). Inferring ground-level nitrogen dioxide concentrations at fine spatial resolution applied to the TROPOMI satellite instrument. *Environmental Research Letters*, *15*(10), 104013. https://doi.org/10.1088/1748-9326/aba3a5

Dang, R., Jacob, D. J., Shah, V., Eastham, S. D., Fritz, T. M., Mickley, L. J., et al. (2023). Background nitrogen dioxide (NO_2) over the United States and its implications for satellite observations and trends: effects of nitrate photolysis, aircraft, and open fires. *Atmospheric Chemistry and Physics*, *23*(11), 6271–6284. https://doi.org/10.5194/acp-23-6271-2023

Dunlea, E. J., Herndon, S. C., Nelson, D. D., Volkamer, R. M., San Martini, F., Sheehy, P. M., et al. (2007). Evaluation of nitrogen dioxide chemiluminescence monitors in a polluted urban environment. *Atmospheric Chemistry and Physics*, *7*(10), 2691–2704. https://doi.org/10.5194/acp-7-2691-2007

Feng, S., Gao, D., Liao, F., Zhou, F., & Wang, X. (2016). The health effects of ambient PM2.5 and potential mechanisms. *Ecotoxicology and Environmental Safety*, *128*, 67–74. https://doi.org/10.1016/j.ecoenv.2016.01.030

Fontijn, Arthur., Sabadell, A. J., & Ronco, R. J. (1970). Homogeneous chemiluminescent measurement of nitric oxide with ozone. Implications for continuous selective monitoring of gaseous air pollutants. *Analytical Chemistry*, *42*(6), 575–579. https://doi.org/10.1021/ac60288a034

van Geffen, J., Boersma, K. F., Eskes, H., Sneep, M., ter Linden, M., Zara, M., & Veefkind, J. P. (2020). S5P TROPOMI NO_2 slant column retrieval: method, stability, uncertainties and comparisons with OMI. *Atmospheric Measurement Techniques*, *13*(3), 1315–1335. https://doi.org/10.5194/amt-13-1315-2020

Ghahremanloo, M., Lops, Y., Choi, Y., & Yeganeh, B. (2021). Deep Learning Estimation of Daily Ground-Level NO2 Concentrations From Remote Sensing Data. *Journal of Geophysical Research: Atmospheres*, *126*(21), e2021JD034925. https://doi.org/10.1029/2021JD034925

Goldberg, D. L., Lu, Z., Oda, T., Lamsal, L. N., Liu, F., Griffin, D., et al. (2019). Exploiting OMI NO2 satellite observations to infer fossil-fuel CO2 emissions from U.S. megacities. *Science of The Total Environment*, *695*, 133805. https://doi.org/10.1016/j.scitotenv.2019.133805

Goldberg, D. L., Anenberg, S. C., Kerr, G. H., Mohegh, A., Lu, Z., & Streets, D. G. (2021). TROPOMI NO2 in the United States: A Detailed Look at the Annual Averages, Weekly

Cycles, Effects of Temperature, and Correlation With Surface NO2 Concentrations. *Earth's Future*, *9*(4), e2020EF001665. https://doi.org/10.1029/2020EF001665

Griffin, D., Zhao, X., McLinden, C. A., Boersma, F., Bourassa, A., Dammers, E., et al. (2019). High-Resolution Mapping of Nitrogen Dioxide With TROPOMI: First Results and Validation Over the Canadian Oil Sands. *Geophysical Research Letters*, *46*(2), 1049–1060. https://doi.org/10.1029/2018GL081095

Griffin, D., McLinden, C. A., Dammers, E., Adams, C., Stockwell, C. E., Warneke, C., et al. (2021). Biomass burning nitrogen dioxide emissions derived from space with TROPOMI: methodology and validation. *Atmospheric Measurement Techniques*, *14*(12), 7929–7957. https://doi.org/10.5194/amt-14-7929-2021

Grzybowski, P. T., Markowicz, K. M., & Musiał, J. P. (2023). Estimations of the Ground-Level NO2 Concentrations Based on the Sentinel-5P NO2 Tropospheric Column Number Density Product. *Remote Sensing*, *15*(2), 378. https://doi.org/10.3390/rs15020378

Gu, J., Chen, L., Yu, C., Li, S., Tao, J., Fan, M., et al. (2017). Ground-Level NO2 Concentrations over China Inferred from the Satellite OMI and CMAQ Model Simulations. *Remote Sensing*, *9*(6), 519. https://doi.org/10.3390/rs9060519

Guay, M., Stieb, D. M., & Smith-Doiron, M. (2011). Assessment of long-term exposure to air pollution in a longitudinal national health survey. *Journal of Exposure Science & Environmental Epidemiology*, *21*(4), 337–342. https://doi.org/10.1038/jes.2010.37

Henderson, S. B., Beckerman, B., Jerrett, M., & Brauer, M. (2007). Application of Land Use Regression to Estimate Long-Term Concentrations of Traffic-Related Nitrogen Oxides and Fine Particulate Matter. *Environmental Science & Technology*, *41*(7), 2422–2428. https://doi.org/10.1021/es0606780

Hoek, G., Beelen, R., de Hoogh, K., Vienneau, D., Gulliver, J., Fischer, P., & Briggs, D. (2008). A review of land-use regression models to assess spatial variation of outdoor air pollution. *Atmospheric Environment*, *42*(33), 7561–7578. https://doi.org/10.1016/j.atmosenv.2008.05.057

Holloway, T., Miller, D., Anenberg, S., Diao, M., Duncan, B., Fiore, A. M., et al. (2021). Satellite Monitoring for Air Quality and Health. *Annual Review of Biomedical Data Science*, *4*, 417–447. https://doi.org/10.1146/annurev-biodatasci-110920-093120

Ialongo, I., Virta, H., Eskes, H., Hovila, J., & Douros, J. (2020). Comparison of TROPOMI/Sentinel-5 Precursor NO_2 observations with ground-based measurements in Helsinki. *Atmospheric Measurement Techniques*, *13*(1), 205–218. https://doi.org/10.5194/amt-13-205-2020

Jerrett, M., Finkelstein, M. M., Brook, J. R., Arain, M. A., Kanaroglou, P., Stieb, D. M., et al. (2009). A Cohort Study of Traffic-Related Air Pollution and Mortality in Toronto, Ontario, Canada. *Environmental Health Perspectives*, *117*(5), 772–777. https://doi.org/10.1289/ehp.11533

Judd, L. M., Al-Saadi, J. A., Szykman, J. J., Valin, L. C., Janz, S. J., Kowalewski, M. G., et al. (2020). Evaluating Sentinel-5P TROPOMI tropospheric NO_2 column densities with airborne and Pandora spectrometers near New York City and Long Island Sound. *Atmospheric Measurement Techniques*, *13*(11), 6113–6140. https://doi.org/10.5194/amt-13-6113-2020

Kim, M., Brunner, D., & Kuhlmann, G. (2021). Importance of satellite observations for high-resolution mapping of near-surface NO2 by machine learning. *Remote Sensing of Environment*, *264*, 112573. https://doi.org/10.1016/j.rse.2021.112573

Kimbrough, S., Chris Owen, R., Snyder, M., & Richmond-Bryant, J. (2017). NO to NO2 conversion rate analysis and implications for dispersion model chemistry methods using Las Vegas, Nevada near-road field measurements. *Atmospheric Environment*, *165*, 23–34. https://doi.org/10.1016/j.atmosenv.2017.06.027

Konovalov, I. B., Berezin, E. V., Ciais, P., Broquet, G., Zhuravlev, R. V., & Janssens-Maenhout, G. (2016). Estimation of fossil-fuel CO_2 emissions using satellite measurements of "proxy" species. *Atmospheric Chemistry and Physics*, *16*(21), 13509–13540. https://doi.org/10.5194/acp-16-13509-2016

Lamsal, L. N., Martin, R. V., van Donkelaar, A., Steinbacher, M., Celarier, E. A., Bucsela, E., et al. (2008). Ground-level nitrogen dioxide concentrations inferred from the satellite-borne Ozone Monitoring Instrument. *Journal of Geophysical Research: Atmospheres*, *113*(D16). https://doi.org/10.1029/2007JD009235

Lamsal, L. N., Martin, R. V., Padmanabhan, A., van Donkelaar, A., Zhang, Q., Sioris, C. E., et al. (2011). Application of satellite observations for timely updates to global anthropogenic NOx emission inventories. *Geophysical Research Letters*, *38*(5). https://doi.org/10.1029/2010GL046476

Lamsal, Lok N., Duncan, B. N., Yoshida, Y., Krotkov, N. A., Pickering, K. E., Streets, D. G., & Lu, Z. (2015). U.S. NO2 trends (2005–2013): EPA Air Quality System (AQS) data versus improved observations from the Ozone Monitoring Instrument (OMI). *Atmospheric Environment*, *110*, 130–143. https://doi.org/10.1016/j.atmosenv.2015.03.055

Lee, H. J., Liu, Y., & Chatfield, R. B. (2023). Neighborhood-scale ambient NO2 concentrations using TROPOMI NO2 data: Applications for spatially comprehensive exposure assessment. *Science of The Total Environment*, *857*, 159342. https://doi.org/10.1016/j.scitotenv.2022.159342

Lee, M., Heikes, B. G., Jacob, D. J., Sachse, G., & Anderson, B. (1997). Hydrogen peroxide, organic hydroperoxide, and formaldehyde as primary pollutants from biomass burning. *Journal of Geophysical Research: Atmospheres*, *102*(D1), 1301–1309. https://doi.org/10.1029/96JD01709

Levy, I., Mihele, C., Lu, G., Narayan, J., & Brook, J. R. (2014). Evaluating Multipollutant Exposure and Urban Air Quality: Pollutant Interrelationships, Neighborhood Variability, and Nitrogen Dioxide as a Proxy Pollutant. *Environmental Health Perspectives*, *122*(1), 65–72. https://doi.org/10.1289/ehp.1306518

Li, M., Wu, Y., Bao, Y., Liu, B., & Petropoulos, G. P. (2022). Near-Surface NO2 Concentration Estimation by Random Forest Modeling and Sentinel-5P and Ancillary Data. *Remote Sensing*, *14*(15), 3612. https://doi.org/10.3390/rs14153612

Mills, I. C., Atkinson, R. W., Kang, S., Walton, H., & Anderson, H. R. (2015). Quantitative systematic review of the associations between short-term exposure to nitrogen dioxide and mortality and hospital admissions. *BMJ Open*, *5*(5), e006946. https://doi.org/10.1136/bmjopen-2014-006946

Novotny, E. V., Bechle, M. J., Millet, D. B., & Marshall, J. D. (2011). National Satellite-Based Land-Use Regression: NO2 in the United States. *Environmental Science & Technology*, *45*(10), 4407–4414. https://doi.org/10.1021/es103578x

Olivier, J. G. J., Bouwman, A. F., Van der Hoek, K. W., & Berdowski, J. J. M. (1998). Global air emission inventories for anthropogenic sources of NOx, NH3 and N2O in 1990.

Environmental Pollution, *102*(1, Supplement 1), 135–148. https://doi.org/10.1016/S0269-7491(98)80026-2

Penn, E., & Holloway, T. (2020). Evaluating current satellite capability to observe diurnal change in nitrogen oxides in preparation for geostationary satellite missions. *Environmental Research Letters*, *15*(3), 034038. https://doi.org/10.1088/1748-9326/ab6b36

Pommier, M. (2023). Estimations of NOx emissions, NO2 lifetime and their temporal variation over three British urbanised regions in 2019 using TROPOMI NO2 observations. *Environmental Science: Atmospheres*, *3*(2), 408–421. https://doi.org/10.1039/D2EA00086E

Qin, K., Rao, L., Xu, J., Bai, Y., Zou, J., Hao, N., et al. (2017). Estimating Ground Level NO2 Concentrations over Central-Eastern China Using a Satellite-Based Geographically and Temporally Weighted Regression Model. *Remote Sensing*, *9*(9), 950. https://doi.org/10.3390/rs9090950

Qin, K., Han, X., Li, D., Xu, J., Loyola, D., Xue, Y., et al. (2020). Satellite-based estimation of surface NO2 concentrations over east-central China: A comparison of POMINO and OMNO2d data. *Atmospheric Environment*, *224*, 117322. https://doi.org/10.1016/j.atmosenv.2020.117322

Sillman, S. (1999). The relation between ozone, NOx and hydrocarbons in urban and polluted rural environments. *Atmospheric Environment*, *33*(12), 1821–1845. https://doi.org/10.1016/S1352-2310(98)00345-8

Steinbacher, M., Zellweger, C., Schwarzenbach, B., Bugmann, S., Buchmann, B., Ordóñez, C., et al. (2007). Nitrogen oxide measurements at rural sites in Switzerland: Bias of conventional measurement techniques. *Journal of Geophysical Research: Atmospheres*, *112*(D11). https://doi.org/10.1029/2006JD007971

US EPA, O. (2013, August 1). Air Quality System (AQS) [Data and Tools]. Retrieved February 15, 2022, from https://www.epa.gov/aqs

Veefkind, J. P., Aben, I., McMullan, K., Förster, H., de Vries, J., Otter, G., et al. (2012). TROPOMI on the ESA Sentinel-5 Precursor: A GMES mission for global observations of the atmospheric composition for climate, air quality and ozone layer applications. *Remote Sensing of Environment*, *120*, 70–83. https://doi.org/10.1016/j.rse.2011.09.027

Watkins, N. (2016). *Near-road Air Quality Monitoring Network: Status and Data*. Presented at the 2016 National Monitoring Conference.

Yeganeh, B., Hewson, M. G., Clifford, S., Tavassoli, A., Knibbs, L. D., & Morawska, L. (2018). Estimating the spatiotemporal variation of NO2 concentration using an adaptive neuro-fuzzy inference system. *Environmental Modelling & Software*, *100*, 222–235. https://doi.org/10.1016/j.envsoft.2017.11.031

Young, M. T., Bechle, M. J., Sampson, P. D., Szpiro, A. A., Marshall, J. D., Sheppard, L., & Kaufman, J. D. (2016). Satellite-Based NO_2 and Model Validation in a National Prediction Model Based on Universal Kriging and Land-Use Regression. *Environmental Science & Technology*, *50*(7), 3686–3694. https://doi.org/10.1021/acs.est.5b05099

Yu, Z., & Li, X. (2022). The Temporal–Spatial Characteristics of Column NO2 Concentration and Influence Factors in Xinjiang of Northwestern Arid Region in China. *Atmosphere*, *13*(10), 1533. https://doi.org/10.3390/atmos13101533

Zoogman, P., Liu, X., Suleiman, R. M., Pennington, W. F., Flittner, D. E., Al-Saadi, J. A., et al. (2017). Tropospheric Emissions: Monitoring of Pollution (TEMPO). *Journal of*

Quantitative Spectroscopy & Radiative Transfer, *186*, 17–39. https://doi.org/10.1016/j.jqsrt.2016.05.008