Evaluating NO\textsubscript{x} emissions and their effect on O\textsubscript{3} production in Texas using TROPOMI NO\textsubscript{2} and HCHO

Daniel L. Goldberg\textsuperscript{1}, Monica Harkey\textsuperscript{2}, Benjamin de Foy\textsuperscript{3}, Laura Judd\textsuperscript{4}, Jeremiah Johnson\textsuperscript{5}, Greg Yarwood\textsuperscript{2}, and Tracey Holloway\textsuperscript{2,6}

\textsuperscript{1}Department of Environmental and Occupational Health, Milken Institute of Public Health, George Washington University, Washington, DC, USA
\textsuperscript{2}Nelson Institute Center for Sustainability and the Global Environment (SAGE), University of Wisconsin–Madison, Madison, WI, USA
\textsuperscript{3}Department of Earth and Atmospheric Sciences, Saint Louis University, St. Louis, MO, USA
\textsuperscript{4}NASA Langley Research Center, Hampton, VA, USA
\textsuperscript{5}Ramboll, Novato, CA, USA
\textsuperscript{6}Department of Atmospheric and Oceanic Sciences, University of Wisconsin–Madison, Madison, WI, USA

Correspondence: Daniel L. Goldberg (dgoldberg@gwu.edu)

Received: 20 April 2022 – Discussion started: 27 April 2022
Revised: 25 July 2022 – Accepted: 5 August 2022 – Published: 26 August 2022

Abstract. The Tropospheric Monitoring Instrument (TROPOMI) on the Sentinel-5 Precursor (S5P) satellite is a valuable source of information to monitor the NO\textsubscript{x} emissions that adversely affect air quality. We conduct a series of experiments using a 4 × 4 km\textsuperscript{2} Comprehensive Air Quality Model with Extensions (CAMx) simulation during April–September 2019 in eastern Texas to evaluate the multiple challenges that arise from reconciling the NO\textsubscript{x} emissions in model simulations with TROPOMI. We find an increase in NO\textsubscript{2} (+17% in urban areas) when transitioning from the TROPOMI NO\textsubscript{2} version 1.3 algorithm to the version 2.3.1 algorithm in eastern Texas, with the greatest difference (+25%) in the city centers and smaller differences (+5%) in less polluted areas. We find that lightning NO\textsubscript{x} emissions in the model simulation contribute up to 24% of the column NO\textsubscript{2} in the areas over the Gulf of Mexico and 8% in Texas urban areas. NO\textsubscript{x} emissions inventories, when using locally resolved inputs, agree with NO\textsubscript{x} emissions derived from TROPOMI NO\textsubscript{2} version 2.3.1 to within 20% in most circumstances, with a small NO\textsubscript{x} underestimate in Dallas–Fort Worth (−13%) and Houston (−20%). In the vicinity of large power plant plumes (e.g., Martin Lake and Limestone) we find larger disagreements, i.e., the satellite NO\textsubscript{2} is consistently smaller by 40%–60% than the modeled NO\textsubscript{2}, which incorporates measured stack emissions. We find that TROPOMI is having difficulty distinguishing NO\textsubscript{2} attributed to power plants from the background NO\textsubscript{2} concentrations in Texas – an area with atmospheric conditions that cause short NO\textsubscript{2} lifetimes. Second, the NO\textsubscript{3}/NO\textsubscript{2} ratio in the model may be underestimated due to the 4 km grid cell size. To understand ozone formation regimes in the area, we combine NO\textsubscript{2} column information with formaldehyde (HCHO) column information. We find modest low biases in the model relative to TROPOMI HCHO, with −9% underestimate in eastern Texas and −21% in areas of central Texas with lower biogenic volatile organic compound (VOC) emissions. Ozone formation regimes at the time of the early afternoon overpass are NO\textsubscript{x} limited almost everywhere in the domain, except along the Houston Ship Channel, near the Dallas/Fort Worth International airport, and in the presence of undiluted power plant plumes. There are likely NO\textsubscript{x}-saturated ozone formation conditions in the early morning hours that TROPOMI cannot observe and would be well-suited for analysis with NO\textsubscript{2} and HCHO from the upcoming TEMPO (Tropospheric Emissions: Monitoring Pollution) mission. This study highlights that TROPOMI measurements offer a valuable means to validate emissions inventories and ozone formation regimes, with important limitations.
1 Introduction

Nitrogen oxides (NO\textsubscript{x} ≡ NO + NO\textsubscript{2}) are a group of reactive trace gases toxic to human health (Burnett et al., 2004; He et al., 2020; Khreis et al., 2017) that can be converted into other chemical species, including ozone and fine particulate matter (Jacob, 1999). There are some natural emissions of NO\textsubscript{3} (e.g., lightning and soil), but the majority of the NO\textsubscript{x} emissions are from anthropogenic sources (Van Vuuren et al., 2011). Anthropogenic NO\textsubscript{3} emissions in polluted areas can be estimated using NO\textsubscript{2} column measurements from satellites (Lamsal et al., 2011; Leue et al., 2001; Martin, 2003; Stavrakou et al., 2008) if the meteorology, NO\textsubscript{2} chemical lifetime, tropospheric/stratospheric components, and NO\textsubscript{3}/NO\textsubscript{2} ratio are all properly accounted for (Beirle et al., 2011; de Foy et al., 2014; Goldberg et al., 2020).

Satellite instruments can observe NO\textsubscript{2} from space because it has strong absorption features within the 400–465 nm wavelength region (Vandaele et al., 1998). By comparing observed spectra with a reference spectrum, the amount of NO\textsubscript{2} in the atmosphere between the instrument and the surface can be derived; this technique is called differential optical absorption spectroscopy (DOAS; Platt, 1994). The first satellite instrument to utilize the DOAS technique to observe NO\textsubscript{2} air pollution was the Global Ozone Monitoring Experiment (GOME; Burrows et al., 1999), launched in 1995 (320 × 40 km\textsuperscript{2} spatial resolution), and was followed by the Ozone Monitoring Instrument (OMI; Levelt et al., 2006), launched in 2004 with vastly improved pixel resolution (24 × 13 km\textsuperscript{2} at nadir) and instrument stability (Schenkeveld et al., 2017). Initial studies used OMI NO\textsubscript{2} satellite data to pinpoint NO\textsubscript{2} emissions in the vicinity of large power plants (Duncan et al., 2013; Kim et al., 2009; Russell et al., 2012) and in areas with high population densities (Boersma et al., 2008; Lamsal et al., 2008, 2010).

Tropospheric Monitoring Instrument (TROPOMI; Veefkind et al., 2012) builds upon the overwhelming success of OMI (Levelt et al., 2018) and has a pixel resolution and instrument stability that are even more advantageous for observing urban-scale NO\textsubscript{2} pollution. Most recently, TROPOMI has been used to estimate NO\textsubscript{3} emissions (Beirle et al., 2019; Dix et al., 2022; de Foy and Schauer, 2022; Goldberg et al., 2019b; Griffin et al., 2019; Lorente et al., 2019) and its changes during the COVID-19 lockdown period (Bauwens et al., 2020; Cooper et al., 2022; Goldberg et al., 2020; Liu et al., 2020; Souri et al., 2021; Sun et al., 2021; Wang et al., 2020). The high spatial resolution of TROPOMI makes it an excellent instrument to observe some of the fine-scale structure of NO\textsubscript{2} pollution, such as within cities (Demetillo et al., 2020; Geddes et al., 2021; Goldberg et al., 2021; Ialongo et al., 2020; Zhao et al., 2020), near power plants (Saw et al., 2021; Shikwambana et al., 2020), near ships (Geougoulas et al., 2020), in the presence of wildfires (Griffin et al., 2021; Jin et al., 2021), and in the presence of oil and gas operations (van der A et al., 2020; Dix et al., 2022; Ialongo et al., 2021).

Studies in the mid-2010s (Canty et al., 2015; Curier et al., 2014; Harkey et al., 2015; Kemball-Cook et al., 2015; Souri et al., 2016; Travis et al., 2016) described the synergistic use of satellite NO\textsubscript{2} and regional chemical transport model simulations to better quantify NO\textsubscript{x} emissions. These studies compared satellite data to model simulations directly, while also accounting for the vertical sensitivity differences between the satellite and model simulation. Results from these studies were mixed but generally found that satellite NO\textsubscript{2} was larger than the model data in rural areas and smaller than the model in urban areas. These studies suggested a potential overestimate of NO\textsubscript{3} emissions in U.S. urban areas and demonstrated the importance of stratospheric transport, lightning NO\textsubscript{2} emissions, soil NO\textsubscript{3} emissions, and NO\textsubscript{2} chemical recycling.

For simulations of 2018 and more recent years, TROPOMI data have been used for model evaluations, e.g., the Community Multiscale Air Quality (CMAQ) modeling system, Long Term Ozone Simulation European Operational Smog (LOTOS-EUROS) model, and Weather Research Forecast with Chemistry (WRF-Chem) model. Most studies show high correlations but larger NO\textsubscript{2} columns in the model in major urban areas and near large point sources. This result is persistent across regions including South Korea (Kim et al., 2020), Europe (Skoulidou et al., 2021), and North America (Lawal et al., 2021; Li et al., 2021). Judd et al. (2020) examined NO\textsubscript{2} in New York city, using TROPOMI version 1.3 (v1.3) NO\textsubscript{2} data and aircraft-/ground-based spectrometer measurements, and found that the satellite underestimated NO\textsubscript{2} by 19%–33%. Verhoelst et al. (2021) also found a satellite low bias (23%–51%) in v1.3 when compared to ground-based measurements, which suggests that an algorithm change is a necessary.

There appear to be the following three primary causes for the low bias in the v1.3 algorithm: (1) a persistent high bias of the cloud pressure retrieved with the Fast Retrieval Scheme for Clouds from the Oxygen A band (FRESCO) cloud algorithm (van Geffen et al., 2021), (2) the relatively coarse model a priori vertical NO\textsubscript{2} profiles (1° × 1°), which underestimate the near-surface NO\textsubscript{2} in polluted regions and are needed for the conversion of the satellite slant column into a vertical column (Goldberg et al., 2017), and (3) the spatial heterogeneity in pointwise data to gridded data comparisons (Souri et al., 2022). The TROPOMI version 2.3.1 (v2.3.1) NO\textsubscript{2} algorithm includes an improved way to estimate cloud pressure and addresses reason 1. Reason 2 can be remediated by incorporating high-resolution spatial information. Judd et al. (2020) reported that, when information from higher-resolution chemical transport models were included in the
calculation of the air mass factor, TROPOMI NO\textsubscript{2} values increased by approximately 12\%–14\% in an urban area. Reason 3 can be accounted for by comparing the satellite measurements to model simulations at similar spatial resolutions to the satellite.

We conduct a series of experiments using a high-resolution photochemical grid model simulation over eastern Texas and evaluate the multiple challenges that arise in evaluation with TROPOMI. We examine the impact of the revised TROPOMI algorithm (Sect. 3.1) and the impact of lightning emissions and other sources of NO\textsubscript{x} in the free troposphere (Sect. 3.2), accounting for TROPOMI’s vertical sensitivity (Sect. 3.3) and evaluating the ability of TROPOMI to resolve urban areas and power plants (Sect. 3.4). While each of these issues involves disparate aspects of model methodology and chemistry, in fact they are intertwined in the correct interpretation of satellite and model results. Based on these results, we consider the ability of TROPOMI to inform emission quantification (Sect. 4.1) and evaluate ozone sensitivity along with formaldehyde (HCHO) retrievals (Sect. 4.2). Based on these results, we offer best practice recommendations for TROPOMI model evaluation and future work.

2 Methods

2.1 CAMx model simulation

For our analysis, we use a 4 × 4 km\textsuperscript{2} Comprehensive Air quality Model with Extensions (CAMx) simulation, version 7.00, centered over eastern Texas and driven offline by the Weather Research Forecast (WRF) model (version 4.0.3). The 4 × 4 km\textsuperscript{2} domain is nested inside 12 × 12 and 36 × 36 km\textsuperscript{2} two-way domains, as shown in Fig. 1. We ran the WRF and CAMx models for the 2019 Texas ozone season from 15 March to 15 October. Only model outputs between 1 April through 30 September are used for this study. We used the 0.25° × 0.25° Global Forecast System data assimilation system as the initial conditions for the WRF meteorological model, which is also used for boundary conditions and analysis nudging on the 36 and 12 km domains. The WRF simulation had 43 vertical levels between the surface and 50 hPa, with approximately 21 layers below 700 hPa. The 43 WRF vertical layers were mapped to 28 vertical layers for the CAMx model simulations; all 21 layers below 700 hPa were mapped without merging. The CAMx simulation was utilized with the Carbon Bond version 6, revision 4 (CB6r4), chemical mechanism (Emery et al., 2016).

For this study, we use a projected 2020 Texas Commission on Environmental Quality (TCEQ) modeling inventory from a 2017 TCEQ inventory, which is different from the National Emissions Inventory (NEI). The 2020 modeling emissions inventory did not include the impacts of the socioeconomic response to COVID-19, which was advantageous for this application since we modeled the 2019 ozone season. TCEQ developed the 2020 modeling emissions inventory for the Dallas–Fort Worth and Houston–Galveston–Brazoria attainment demonstration State Implementation Plan (SIP) revision (Johnson et al., 2018). Within Texas, emissions were calculated using locally resolved inputs, such as mobile emissions from the MOtor Vehicle Emission Simulator (MOVES2014a) that are adjusted based on traffic statistics from the Highway Performance Monitoring System. Outside of Texas, NEI estimates were used, such as the default outputs from MOVES2014 and the 2014 EPA NEI.

We included hourly specific power plant emissions using measurements from the EPA’s Clean Air Markets Division (CAMD; https://www.epa.gov/airmarkets, last access: 24 August 2022) as inputs into the model simulation. Large power plants use continuous emissions monitoring systems (CEMSs) to report emissions of sulfur dioxide (SO\textsubscript{2}), NO\textsubscript{x}, and CO\textsubscript{2}, along with other parameters such as heat input, as required by the federal Clean Air Act. We downloaded hourly data from EPA’s Air Markets Program Data (AMPD) website for the continental U.S. for March through October 2019. Stack parameters were based on EPA’s 2017 NEI data. The 2017 NEI data, with matching facilities in Texas, were then adjusted to their 2019 annual totals. Table 1 provides the annual inventory NO\textsubscript{x} emission rates for four cities within a 50 km radius of the city center and three power plants examined in detail in this study.

Biogenic emissions were estimated for 2019 from the Model of Emissions of Gases and Aerosols from Nature (MEGAN) version 3.1 and fire emissions from Fire Inventory of NCAR (FINN) version 1. We included lightning NO\textsubscript{x} (LNO\textsubscript{x}) emissions, with the CAMx LNO\textsubscript{x} processor using the 2019 WRF meteorological data. The LNO\textsubscript{x} processor estimates hourly, grid-column-specific lightning flash rates (Luo et al., 2017; Price and Rind, 1992) using cloud top heights and convective available potential energy (CAPE) diagnosed from WRF temperature and moisture profiles. The processor then determines the ratio of intracloud light-
ning (IC) to cloud-to-ground lightning (CG) according to the approach of Price and Rind (1993), NO yield per flash estimated by Pickering et al. (2017), and vertical distribution of resulting NO emission rates following DeCaria et al. (2005). In-line inorganic iodine emissions (I_x) from saltwater areas and iodine chemistry are also included.

2.2 TROPOMI

TROPOMI was launched by the European Space Agency (ESA) for the European Union’s Copernicus S5P satellite mission on 13 October 2017. The satellite follows a sun-synchronous, low-Earth (825 km) orbit, with an Equator overpass time of approximately 13:30 LST (local solar time). TROPOMI measures total column amounts of several trace gases in the ultraviolet-visible-near-infrared (UV-VIS-NIR; e.g., NO_2 and HCHO) and shortwave infrared (SWIR; e.g., CO) spectral regions. At nadir, pixel sizes are 3.5 × 7 km^2 (modified to 3.5 × 5.5 km^2 on 6 August 2019), with the edges having slightly larger pixels sizes (~14 km wide) across a 2600 km swath, equating to 450 rows (van Geffen et al., 2020). The instrument observes the swath approximately once every second and orbits the Earth in about 100 min, resulting in daily global coverage. For domain-wide comparisons, we screened TROPOMI pixels for quality assurance flag values greater than 0.75. As a polar-orbiting satellite with an afternoon overpass, care must be taken in the interpretation of TROPOMI column retrievals as an indicator of near-surface emissions (Penn and Holloway, 2020; Streets et al., 2013). TROPOMI provides snapshots at the same time each day, except when limited by cloud cover, surface albedo, or instrument errors.

2.2.1 NO_2

NO_2 slant column densities are derived from radiance measurements in the 405–465 nm spectral window of the UV-VIS-NIR spectrometer. Tropospheric vertical column density data, which represent the vertically integrated number of NO_2 molecules per unit area between the surface and
the tropopause, are then calculated by subtracting the stratospheric portion and then converting the tropospheric slant column to a vertical column using an air mass factor (AMF). The AMF is a unitless quantity used to convert the slant column into a vertical column and is a function of the satel-

2.2.2 HCHO

HCHO slant column densities are derived from radiance measurements in the 328–359 nm spectral window of the UV-VIS-NIR spectrometer. In a similar manner to NO₂, HCHO is measured as a slant column and is converted from a slant column to a vertical column using an AMF and a pri-

2.2.3 Regridding and accounting for the vertical sensitivity of TROPOMI

For comparison with CAMx, we gridded TROPOMI data to the model to create a custom level 3 data product for comparison either with each other or with model data on a common grid. Though our level 3 data product is on an equivalent horizontal grid as the model, the satellite a priori (used in the retrieval) and CAMx have different vertical resolutions and distributions of NO₂. To limit artificial differences when doing the comparisons in this work, additional processing is done in two ways.

1. Applying the averaging kernel. The most user-friendly approach involves creating a model-simulated satellite NO₂ column using the CAMx profile and a TROPOMI data-product-specific averaging kernel, which may be described as the weights used to calculate a weighted vertical integral (we refer to this as AK). To apply the averaging kernel to the model simulation, we first inter-

2. Recalculating the AMF. In a second approach, we instead use daily partial vertical NO₂ columns from CAMx and the tropospheric averaging kernel to re-

2.3 Deriving NO_x emissions from TROPOMI NO₂

2.3.1 Exponentially modified Gaussian fitting method

To derive NO_x emissions from the polluted areas of eastern Texas, an exponentially modified Gaussian (EMG) function is fit to a collection of NO₂ plumes observed from TROPOMI. The original methodology, proposed by Beirle et al. (2011), involves the fitting of satellite line densities to an EMG function. Line densities are the integral of the column NO₂ retrieval perpendicular to the path of the plume, and the units are mass per distance. We rotate each day’s plume based on the wind direction, so that all daily plumes are art-

https://doi.org/10.5194/acp-22-10875-2022 Atmos. Chem. Phys., 22, 10875–10900, 2022
where $\alpha$ is the total number of NO$_2$ molecules observed near the pollution source, excluding the effect of background NO$_2$ and $\beta$. $x_0$ is the e-folding distance downwind, representing the length scale of the NO$_2$ decay, $\mu$ is the location of the apparent source relative to the assumed pollution source center, $\sigma$ is the standard deviation of the Gaussian function, representing the Gaussian smoothing length scale, and $\Phi$ is the Gaussian cumulative distribution function. Using the CURVEFIT function in IDL (Interactive Data Language), we determine the following five unknown parameters: $\alpha$, $x_0$, $\sigma$, $\mu$, and $\beta$, based on the independent (distance; $x$) and dependent (NO$_2$ column line density) variables.

Using the mean ERA5 100 m wind speed, $w$, the mean effective NO$_2$ lifetime $\tau_{\text{effective}}$, and the mean NO$_x$ emissions can be calculated from the fitted parameters $x_0$ and $\alpha$, as expressed in Eq. (2):

$$\text{NO}_x \text{Emissions} = 1.32 \left(\frac{\alpha}{\tau_{\text{effective}}}\right), \text{ where } \tau_{\text{effective}} = \frac{x_0}{w}.$$  \hfill (2)

Equation (2) yields emission estimates in units of moles per second. A factor of 1.32 is the mean column-averaged NO$_x$/NO$_2$ ratio and is the widely used value to convert the NO$_2$ to NO$_x$ in polluted regions (Beirle et al., 2021). Appendix C shows the variation in the NO$_x$/NO$_2$ across our domain.

### 2.3.2 Flux divergence method

Emissions were also estimated using the flux divergence method, as follows (Beirle et al., 2019):

$$\text{NO}_x \text{Emissions} = 1.32 \left( \nabla \cdot (\text{VCD} \cdot u) + \frac{\text{VCD}}{\tau} \right).$$  \hfill (3)

Fluxes of NO$_2$ were obtained by multiplying NO$_2$ vertical column densities (VCDs) with wind speeds ($u$) in orthogonal directions (along and across the swath tracks). The divergence of the fluxes yields an emission estimate in units of moles per meter squared per second. The fluxes can then be integrated across the 2-D urban area to obtain emission rates in analogous units as Eq. (2). Sinks of NO$_2$ are included in the equation by adding VCDs divided by the atmospheric lifetime of NO$_2$, $\tau$, which was taken from the EMG fit. Estimates of NO$_x$ emissions are obtained by multiplying the estimates by the ratio of NO$_2$ to NO$_x$, which is the same 1.32 value as the EMG method (Beirle et al., 2021). The fluxes were calculated using the same 100 m ERA5 wind product used for the EMG estimates. The winds were linearly interpolated to the daily swath grid. This method follows de Foy and Schauer (2022), with minor modifications. The quality assurance flag threshold was set to 0.75 to be consistent with EMG. The central 250 pixels (out of 450) were used for swaths, as these have a higher resolution than the outer bands and are critical for this method. We retrieved swaths from October 2019 through September 2021. Although this period does include the COVID-19 lockdown period, the October 2019 through September 2021 time frame does not show time-averaged NO$_2$ values more than 10% different than the year prior and is well within the uncertainty of this analysis. Using the method described in de Foy et al. (2014), two-dimensional Gaussian fits were obtained.

### 3 Results and discussion

#### 3.1 Comparison between TROPOMI version 1.3 and version 2.3.1 algorithms

To elucidate the effects of the recent TROPOMI NO$_2$ algorithm change from v1.3 to v2.3.1, we compare both within our model domain. As expected, the v2.3.1 algorithm yields consistently larger values than the v1.3 algorithm in most areas of our eastern Texas domain (Fig. 2). The largest increases by both magnitude and percentage occur in the most polluted areas. We find an average increase of +16.6% in urban counties, with a maximum increase of +45% in the most polluted section of eastern Houston. Increases exceeding +20% also occur in the vicinity of large point source emissions. In the rural areas of eastern Texas, we generally observe small increases of less than 5%. We fit a linear regression to a scatterplot of the tropospheric vertical columns from both algorithms in the urban counties and find a slope of 1.30 and a negative intercept, which further confirms that the algorithm change affects the most polluted areas more strongly than the moderate and less polluted areas.

#### 3.2 Effects of free tropospheric NO$_2$ and lightning NO$_x$

For this study, we conducted two CAMx simulations, i.e., with and without lightning NO$_x$ emissions. The tropospheric NO$_2$ vertical profiles for eastern Texas, Dallas, and Houston are shown in the left-hand side panels of Fig. 3. In a CAMx simulation without lightning NO$_x$, average NO$_2$ concentrations between 2.5–10 km averaged 20 ppt (parts per trillion) for the eastern Texas domain. This can be compared to free tropospheric (>2.5 km) NO$_2$ concentrations from the NASA Studies of Emissions, Atmospheric Composition, Clouds and Climate Coupling by Regional Surveys (SEAC4RS) campaign within our eastern Texas model domain but in 2013 instead of 2019. Measured NO$_2$ concentrations between 2.5 and 10 km averaged 50 ppt during the SEAC4RS campaign. This also compares the ∼40 ppt estimate from OMI using a cloud-slicing methodology in the central U.S. during June–August 2005–2007 (Marais et al., 2018). When lightning NO$_x$ emissions are included in CAMx, the free tropospheric NO$_2$ between 2.5–10 km increases from 20 to 33 ppt, but there is still a slight underestimate compared to SEAC4RS data between 2.5 and 6 km. The small underestimate shown in the CAMx simulation with lightning NO$_x$ emissions compared to the SEAC4RS data in the 2.5–6 km altitude range could be due to the decrease in anthropogenic NO$_x$ emis-
D. L. Goldberg et al.: Evaluating NO\textsubscript{x} emissions and their effect on O\textsubscript{3} production in Texas

Figure 2. The left panel shows the NO\textsubscript{2} tropospheric vertical column amounts from the TROPOMI NO\textsubscript{2} v2.3.1 algorithm screened with a quality assurance flag greater than 0.75. The center panel shows the ratio between the NO\textsubscript{2} tropospheric vertical column amounts from the v2.3.1 algorithm compared to the v1.3 algorithm. The right panel shows a scatterplot and linear fit between the two TROPOMI NO\textsubscript{2} products used in the center panel. The urban area is defined as the five counties surrounding the largest five cities (Houston, Dallas, Fort Worth, San Antonio, and Austin). The rural area is everywhere outside those counties.

sions between 2013 and 2019. Colocated vertical NO\textsubscript{2} measurements in time and space would be needed to evaluate this further.

In order to compare model simulation output to satellite data, it is important to understand free tropospheric NO\textsubscript{2} (Marais et al., 2018, 2021) and understand its effects on the satellite retrieval (Silvern et al., 2019). TROPOMI has greater sensitivity to the upper portion of the troposphere, and this must be accounted for in any comparison with model output. In the right panels of Fig. 3, we show the modeled shape profiles – the NO\textsubscript{2} vertical distribution normalized to a unitless quantity that integrates to unity over the depth of the troposphere – and the sensitivity of TROPOMI to NO\textsubscript{2} at different levels of the atmosphere (green line). In Texas during summer 2019, TROPOMI was 3 times as sensitive to NO\textsubscript{2} at an altitude of 10 km (tropospheric averaging kernel = 1.5) compared to the surface (tropospheric averaging kernel = 0.5). This demonstrates that NO\textsubscript{2} at the tropospheric–stratospheric interface (~12 km altitude), such as lightning NO\textsubscript{x} (Zhu et al., 2019) and cruising aircraft emissions, can have an outsized effect on the satellite measurement. To facilitate a comparison, model-simulated column amounts can be adjusted by applying the averaging kernel, which will be discussed in Sect. 3.3.

The inclusion of lightning NO\textsubscript{x} emissions increases seasonal column tropospheric NO\textsubscript{2} by an average of 0.16 × 10\textsuperscript{15} molec. cm\textsuperscript{-2} in the model simulation during April through September 2019 (Fig. 4). This increase varies spatiotemporally due to the prevalence of thunderstorms; however, when averaged over 6 months, the increase is relatively homogeneous. The inclusion of lightning NO\textsubscript{x} emissions most affects the satellite–model comparison in rural areas but is also relevant in urban areas. The 0.16 × 10\textsuperscript{15} molec. cm\textsuperscript{-2} increase yields an increase in the tropospheric column NO\textsubscript{2} of +7.8% in urban areas, +15% in the rural areas of eastern Texas, and up to +24% over the Gulf of Mexico. For the rest of this paper, only the CAMx simulation with the inclusion of lightning NO\textsubscript{x} emissions will be analyzed.

3.3 Applying the averaging kernel and recalculating the air mass factor

To compare a chemical transport model simulation to satellite data, one must account for the differing assumptions about the vertical NO\textsubscript{2} distributions between model and satellite. One can either apply the averaging kernel from the satellite instrument to the NO\textsubscript{2} column from the model simulation or use the NO\textsubscript{2} vertical profile from the model simulation and the averaging kernel to recalculate AMF and tropospheric NO\textsubscript{2} vertical column of the satellite measurement. Typically, studies use either one of the two methods; similar to Douros et al. (2022), we use both.

The comparison between the model and model with the tropospheric averaging kernel (AK) applied is shown in the left column of Fig. 5. Upon application of the AK, the tropospheric column NO\textsubscript{2} in the model simulation artificially increases in rural areas by +15.4%, while the urban NO\textsubscript{2} will artificially decrease. The latter is due to most NO\textsubscript{2} being below 2 km due to the large NO\textsubscript{x} emissions near the surface in urban areas where AK < 1.

Once the tropospheric averaging kernel is applied, it can be compared to the satellite directly (top row of Fig. 5). In Dallas–Fort Worth and Houston, there are lower amounts of NO\textsubscript{2} in the model simulation in the most polluted areas of
the city but generally good agreement (+0.4 %) when the five urban areas (Dallas, Fort Worth, Houston, San Antonio, Austin) are averaged together. In the rural areas of eastern Texas, there are slightly larger amounts (+10.7 %) in the model simulation than that observed by TROPOMI, but these absolute differences are small. The largest disagreements between CAMx and TROPOMI occur in the vicinity of large point sources, which we discuss further in Sect. 3.4

While applying the averaging kernel to a regional model simulation is an appropriate way to compare model simulations with satellite data, it does so by artificially adjusting the high-resolution model simulation to follow the coarse resolution (1.0° × 1.0°) of the TM5-MP model simulation used to originally process the AMF. Instead, incorporating the high-resolution model vertical profiles in the calculation of the AMF, while more computationally intensive, results in satellite measurements incorporating higher spatial resolution information; in urban areas, this yields satellite measurements that have greater spatial heterogeneity.

In the middle row of Fig. 5, we show a comparison between the model and the satellite with the CAMx-derived AMF. In this comparison, we obtain similar conclusions, as mentioned earlier, in which the model has systematically smaller NO₂ amounts than TROPOMI in Dallas–Fort Worth and Houston and larger amounts in rural areas. The agreement between the satellite measurement with a new AMF applied and model simulation is marginally different than when the averaging kernel is applied to the model simulation and compared to the satellite measurement directly. The percentage difference calculations differ primarily because the denominator (i.e., TROPOMI value) is a different magnitude in each case. We attribute this small difference to the rounding errors in the interpolation of the averaging kernel to the CAMx model pressure levels.

### 3.4 Localized TROPOMI vs. CAMx NO₂ comparison

We evaluate two versions of the TROPOMI seasonal average against the CAMx model simulation, i.e., TROPOMI v2.3.1 and TROPOMI v2.3.1 with CAMx AMFs. In Fig. 6, a comparison of these satellite products versus CAMx are shown for the following four metropolitan areas: Dallas–Fort Worth (DFW), San Antonio (SAT), Austin (AUS), and Houston (HOU). Comparing TROPOMI v2.3.1 to CAMx directly without the application of the averaging kernel (which is not recommended) suggests a model high bias of +8.4 % but moderately good association with each other ($r^2 = 0.70$). We then use the a priori profiles from the CAMx simulation to recalculate the AMF and find that the original model high bias in urban areas becomes a low bias of −0.1 % and becomes a larger low bias in the most polluted sections of the cities (consistent with our discussion in Sect. 3.3). The low model bias is most pronounced in eastern Houston and the downtown area of Dallas. For Dallas–Fort Worth, there also appears to some spatial misallocation because NO₂ near the DFW airport is larger in the model than the satellite, while NO₂ in the downtown areas of Dallas and Fort Worth is smaller in the model than the satellite. In San Antonio and Austin, there is a small model overestimate, which becomes worse near the large point sources on the periphery of the city. Overall, however, there is a generally good performance between CAMx NO₂ and TROPOMI NO₂, which is within 20 % in most cases. The value of 20 % is well within the expectation of TROPOMI accuracy and precision. The non-point NOₓ emissions input into the model simulation (e.g., mobile, nonroad, and area sources) are generally within the uncertainty of the satellite measurement, and we would not
Figure 4. NO$_2$ tropospheric vertical column amounts from the CAMx model (left panel) without and (center panel) with lightning NO$_x$ emissions averaged during April through September 2019 at the coincident TROPOMI overpass time (~19:00 UTC). Areas with invalid TROPOMI data are similarly screened out from the model output on a daily basis. The urban area is defined as the five counties surrounding the largest five cities (Houston, Dallas, Fort Worth, San Antonio, and Austin). The rural area is everywhere outside those counties. The right panel shows the fraction of the NO$_2$ column attributed to different layers of the atmosphere (below 2 km, above 2 km (attributed to Other), and above 2 km attributed to lightning NO$_x$ (LNO$_x$)) at six locations (Gulf of Mexico, rural central Texas, Austin, San Antonio, Dallas, and Houston). The fraction attributed to lightning NO$_x$ (LNO$_x$) is calculated as the NO$_2$ addition between the two simulations without and with lightning NO$_x$ emissions.

recommend a substantial alteration to the inventory for these sector emissions, except in the eastern Houston neighborhood. This exercise demonstrates the importance of utilizing the AMF when comparing satellite data to model simulations.

To evaluate the performance of TROPOMI in observing point source emissions, we compare TROPOMI NO$_2$ measurements at the locations of three power plants with stack measurements, i.e., Martin Lake, Limestone, and Sam Seymour (Fig. 7). In each case, TROPOMI substantially underestimates NO$_2$ at the locations of these power plants, even when the new algorithm and recalculated AMF are both applied. We have previously found better agreement between TROPOMI NO$_2$ and the stack measurements for the Colstrip power plant in Montana and San Juan/Four Corners complex in New Mexico (Goldberg et al., 2019b). The reason for the substantial disagreement in Texas is still unknown, but we do not believe this repudiates our prior evaluation for urban areas. NO$_x$ emissions from the power sector in the U.S. have declined by 76% between 2005 (3.63×10$^6$ t) and 2019 (0.86×10$^6$ t; https://campd.epa.gov/, last access: 24 August 2022). At these lower emission rates, it appears that TROPOMI has difficulty distinguishing NO$_2$ attributed to power plants from the background NO$_2$ concentrations, especially in areas such as Texas with atmospheric conditions that cause short NO$_2$ lifetimes, i.e., rapid plume dilution, high oxidation capacity due to large numbers of volatile organic compounds (VOCs) and large amounts of water vapor, and high solar elevation angles. Second, the NO$_x$/NO$_2$ ratio in the model may be underestimated due to

the 4 km grid cell size (Appendix C). The two power plants in New Mexico and Montana are located in areas with smaller background NO$_2$, lighter wind speeds, fewer VOCs and less water vapor, and higher elevations; all of these factors cause the satellite sensor to be more sensitive to the NO$_x$ emissions. TROPOMI does not have the same difficulty over urban areas because the larger aggregated NO$_x$ emissions are more easily distinguishable from background concentrations. Please see Appendix D for a discussion on this topic. Future work should focus on evaluating the NO$_2$ from power plants and the NO$_x$/NO$_2$ ratio as the plume evolves, such as in situ measurements from aircraft and ground-based vertical column instruments (e.g., Pandora; Herman et al., 2009).

4 Policy-relevant findings based on TROPOMI model evaluation

4.1 TROPOMI NO$_x$ emissions

In order to calculate NO$_x$ emissions directly, we need to account for the NO$_2$ lifetime and NO$_2$ background concentrations. The first technique we use is the exponentially modified Gaussian (EMG) method. We first apply the EMG method to the CAMx simulations of the Limestone power plant (latitude 31.42° N, longitude 96.25° W) NO$_x$ plume. By comparing the known emissions with the inferred top-down emissions, we can evaluate assumptions in the EMG model. The number of NO$_x$ emissions input into the model within a 12 km radius of the facility is 9.8 Gg yr$^{-1}$. The top-down EMG method applied to the CAMx simulation
Figure 5. The top row panels show the NO\textsubscript{2} tropospheric vertical column amounts from CAMx and TROPOMI v2.3.1 reprocessed with a priori profiles from the CAMx model with lightning NO\textsubscript{x} emissions and the difference averaged across April through September 2019. The middle row panels show the NO\textsubscript{2} tropospheric vertical column amounts from CAMx with the averaging kernel applied and the TROPOMI v2.3.1 product and difference averaged across April through September 2019. The bottom row shows the difference between the top and middle rows. Areas with invalid TROPOMI data are similarly screened out from the model out on a daily basis. The urban area is defined as the five counties surrounding the largest five cities (Houston, Dallas, Fort Worth, San Antonio, and Austin). The rural area is everywhere outside those counties.

yields a NO\textsubscript{x} emissions rate of 13.1 Gg yr\textsuperscript{-1}. The disagreement between the NO\textsubscript{x} emissions inventory (9.8 Gg yr\textsuperscript{-1}) and the inferred CAMx NO\textsubscript{x} emissions driven by the inventory (13.1 Gg yr\textsuperscript{-1}) must be due to incorrectly assumed effective wind speed likely driven by the meandering of the winds. Winds rarely have a consistent direction and instead meander due to boundary layer turbulence and frictional effects, yielding a slower effective speed in the wind direction over long distances (> 10 km). If we assume that the effective speed of the NO\textsubscript{2} plume is 25% slower than the unidirectional wind speed, then the inferred top-down emissions can be made to match the known emissions (9.8 Gg yr\textsuperscript{-1}).

Applying the CAMx-based effective plume speed to analysis of TROPOMI (25% slower than the unidirectional...
wind speed), we find that the TROPMI NO\textsubscript{x} product yields an estimated NO\textsubscript{x} emissions rate of 5.2 Gg yr\textsuperscript{-1}, and this increased to 6.0 Gg yr\textsuperscript{-1} when using the TROPMI v2.3.1 algorithm with a recalculated AMF (Table 2 and Fig. 8). Even with all known corrections applied, it appears that TROPMI is not capturing the full magnitude of NO\textsubscript{x} emissions from the power plant and vicinity (9.8 Gg yr\textsuperscript{-1}), which is consistent with the discussion in Sect. 3.4.

For the Dallas–Fort Worth area, if we apply the same method to the CAMx simulation, we obtain an effective NO\textsubscript{x} emissions rate of 55 Gg yr\textsuperscript{-1} from the metropolitan area. This is equivalent to the NO\textsubscript{x} emissions aggregated within a 47 km radius of the Dallas–Fort Worth metropolitan area (latitude 32.85° N, longitude 96.95° W) and is roughly equivalent to two-sigma of the Gaussian plume (σ = 23.7 km).

Using the TROPMI v2.3.1 algorithm, we calculate a top-down NO\textsubscript{x} emissions rate of 56 Gg yr\textsuperscript{-1}, and it increases to 62 Gg yr\textsuperscript{-1} when a CAMx AMF is used (Table 2 and Fig. 8). The difference between the 62 Gg yr\textsuperscript{-1} calculated directly from the TROPMI v2.3.1 with a recalculated AMF, and the 55 Gg yr\textsuperscript{-1} effective emissions rate from CAMx represents a small 13% low bias that is within the uncertainty of the satellite and the assumptions made to facilitate the comparison.
Table 2. NO\textsubscript{x} emission rates for Dallas–Fort Worth and the Limestone power plant from the TCEQ Emissions Inventory and various iterations of the TROPOMI NO\textsubscript{2} algorithm.

<table>
<thead>
<tr>
<th>Data source</th>
<th>Data source type</th>
<th>Dallas–Fort Worth NO\textsubscript{x} emissions (Gg yr\textsuperscript{-1})</th>
<th>Limestone PP NO\textsubscript{x} emissions (Gg yr\textsuperscript{-1})</th>
</tr>
</thead>
<tbody>
<tr>
<td>TCEQ Projected 2020 Inventory</td>
<td>Bottom up</td>
<td>55</td>
<td>9.8</td>
</tr>
<tr>
<td>TROPOMI NO\textsubscript{2} v2.3.1</td>
<td>Top down</td>
<td>56 ± 20</td>
<td>5.2 ± 1.9</td>
</tr>
<tr>
<td>TROPOMI NO\textsubscript{2} v2.3.1 CAMx AMFs</td>
<td>Top down</td>
<td>62 ± 22</td>
<td>6.0 ± 2.2</td>
</tr>
</tbody>
</table>

Figure 8. EMG method to derive NO\textsubscript{x} emissions from the TROPOMI NO\textsubscript{2} v2.3.1 with CAMx AMFs applied to (left panel) Dallas–Fort Worth and (right panel) Limestone power plant. The color bar for the right panel is halved to better show the NO\textsubscript{2} plume near Limestone. ERA5 100 m winds are used to rotate daily TROPOMI NO\textsubscript{2} plumes.

The technique was applied to other urban areas, but those cities have large point sources at the periphery of the urban areas which adversely affected the calculation of the effective NO\textsubscript{2} lifetime needed to calculate the NO\textsubscript{x} emissions.

The top-down approach can also calculate effective NO\textsubscript{2} lifetimes. Most top-down methods fit both the effective NO\textsubscript{2} lifetime and NO\textsubscript{x} emissions simultaneously and therefore have a seesaw relationship – as the lifetime increases, the NO\textsubscript{x} emissions decrease, given a constant NO\textsubscript{2} burden. Here, we visually inspect the plume to ensure that the NO\textsubscript{2} effective lifetime is reasonable (generally between 0.5–5 h), given the plume decay before proceeding. For Dallas–Fort Worth, the method calculates an effective NO\textsubscript{2} lifetime of 1.7 h. The same approach applied to CAMx yields an effective NO\textsubscript{2} lifetime of 1.1 h. This suggests that the effective NO\textsubscript{2} lifetime in CAMx is too short. The effective lifetime is a function of the chemical lifetime and dispersion lifetime as follows (de Foy et al., 2014):

$$\frac{1}{\tau_{\text{effective}}} = \frac{1}{\tau_{\text{chemical}}} + \frac{1}{\tau_{\text{dispersion}}}.$$  \hspace{1cm} (4)

The effective lifetime could be increased in a model simulation either by increasing the NO\textsubscript{2} chemical lifetime (e.g., slower photolysis, slowing the NO\textsubscript{2} + OH reaction rate, faster recycling of NO\textsubscript{2} (NO\textsubscript{2} = alkyl nitrates, peroxyacetyl nitrate (PAN), and HNO\textsubscript{3}) back to NO\textsubscript{2}) or by increasing the vertical mixing (less plume meandering at higher altitudes due to fewer surface frictional effects). Chemical NO\textsubscript{2} lifetimes are well-constrained by laboratory studies, so we hypothesize that too slow vertical transport may be the primary culprit for this disagreement, and this is also suggested by the analysis presented in Fig. 3, which suggests a low model bias in the free troposphere using measurements from the SEAC4RS campaign. Future vertical NO\textsubscript{2} measurements separated by altitude will be critical to answering this question.

The total error associated with the magnitude of the top-down versus bottom-up comparison is calculated to be 36 % and is the sum of the quadrature of five potential sources of error, namely the tropospheric vertical column measurement in urban areas (20 %), the wind speed and direction (25 %; Appendix B), the clear-sky bias (10 %), which for these purposes is a result of emissions being different on clear-sky days compared to cloudy days, the NO\textsubscript{x}/NO\textsubscript{2} ratio (10 %; Appendix C), and the random error of the statistical EMG fit (10 %; de Foy et al., 2014). This total uncertainty is approximately 20 % smaller than similar methods using OMI. For further information on this method or the uncertainties.
Table 3. NOx emission rates for Dallas–Fort Worth and the Limestone power plant from the TCEQ Emissions Inventory and various iterations of the flux divergence method using the TROPOMI NO2 v2.3.1 algorithm.

<table>
<thead>
<tr>
<th>Data source</th>
<th>Dallas–Fort Worth NOx emissions (Gg yr$^{-1}$)</th>
<th>Limestone PP NOx emissions (Gg yr$^{-1}$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>TCEQ projected 2020 inventory</td>
<td>55</td>
<td>9.8</td>
</tr>
<tr>
<td>TROPOMI NO2 v2.3.1, infinite NO2 lifetime</td>
<td>24 ± 9</td>
<td>1.6 ± 0.4</td>
</tr>
<tr>
<td>TROPOMI NO2 v2.3.1, short NO2 lifetime</td>
<td>62 ± 16</td>
<td>3.4 ± 1.1</td>
</tr>
</tbody>
</table>

associated with this method, please see the other literature (de Foy et al., 2014; Goldberg et al., 2019a; Lu et al., 2015; Verstraeten et al., 2018).

We then test the flux divergence method (Beirle et al., 2019, 2021; de Foy and Schauer, 2022) on the same two sources, i.e., Dallas and the Limestone power plant. We apply the flux divergence method to the native TROPOMI pixels rather than a regridded version of the data. Figure 9 shows that TROPOMI columns distinguish between a large hotspot over Dallas and a smaller one over Fort Worth. For the Dallas urban area, the algorithm identified 11 separate source regions which were each represented by a separate two-dimensional Gaussian method. The flux divergence method was able to resolve source regions with better detail, with estimates for some of the individual point sources and sub-areas within Dallas. In particular, the area including the Dallas/Fort Worth International Airport appears as a distinct source area. In Table 3, we show the NOx emissions aggregated for these two sources, using both an infinite NO2 lifetime and the effective short NO2 lifetime provided by the EMG method (τ = 1.7 h for Dallas–Fort Worth; τ = 0.5 h for Limestone PP). The results from the flux divergence method are consistent with the results from the EMG method in the Dallas area, provided that a short NO2 lifetime is assumed.

4.2 Evaluating ozone sensitivity using the HCHO–NO2 ratio

Satellite observations of formaldehyde (HCHO) can be combined with NO2 to determine the ozone sensitivity to NOx emissions using the formaldehyde to nitrogen dioxide column density ratio (FNR; Duncan et al., 2010; Jin et al., 2017; Jin and Holloway, 2015; Martin et al., 2004). HCHO may be used to estimate short-lived VOC emissions, both anthropogenic and biogenic combined, which often quickly oxidize to HCHO in the presence of sunlight and the hydroxyl (OH) radical (Wolfe et al., 2016; Zhu et al., 2017). In a similar manner to NO2, column HCHO can be compared to chemical transport models in order to better understand the spatial variability in VOC emissions. Harkey et al. (2020) found that a regional model captured the general spatial and temporal behavior of satellite estimates but tended to underestime column HCHO. In the western U.S., TROPOMI HCHO measurements have been rigorously evaluated using ground-based spectrometers, and the v1.1 algorithm was found to be biased low by approximately 25% (De Smedt et al., 2021).

We first compare the column HCHO comparison between CAMx and TROPOMI. Tropospheric column TROPOMI HCHO measurements using the v1.1 algorithm are biased low by approximately 25% (De Smedt et al., 2021). We then create a bias-corrected (b-c) product (multiply by a factor of 1.25) to account for this low bias. In Fig. 10, we compare the operational TROPOMI HCHO v1.1 product and TROPOMI HCHO v1.1 b-c product to CAMx tropospheric column amounts with and without the averaging kernel sampled at coincident timeframes. Since HCHO spatial patterns have less heterogeneity than NO2, due to a large fraction of HCHO originating from biogenic precursors during warm season months, column HCHO amounts are less sensitive to the application of the AK than with NO2. The difference between CAMx and CAMx with the averaging kernel applied is ±2.5% for area-wide averages. CAMx underestimates HCHO in central and western Texas, but in eastern Texas the magnitude and spatial patterns match better. The model bias is −7.9% in eastern Texas and −25.0% in central Texas, compared to the TROPOMI HCHO v1.1 b-c product. This model bias, in both cases, is within the uncertainty of the satellite retrieval.

We apply the FNR to TROPOMI and CAMx to determine how well CAMx is representing ozone formation regimes. Initial studies showed that, when the FNR in a region exceeds 2, the ozone formed is considered to be limited by the amount of NOx present in the air. When the FNR is below 0.5, the ozone formed is considered to be limited by the amount of VOCs. Ratio values between 0.5 and 2 indicate sensitivity to both NOx and VOCs (Duncan et al., 2010). More recent studies have demonstrated that the upper bound of the transitional regime could be as high as 4 (or even higher), depending on regional characteristics (Jin et al., 2017, 2020; Schroeder et al., 2017).

For this analysis, using the v1.1 HCHO and v1.3 NO2 algorithms is sufficient, since both products have similar biases related to the cloud schemes that may cancel out when a ratio is calculated. We use a value of 4 to indicate the transition between NOx and VOC sensitivity, while simultaneously noting that this value should not be static in all scenarios. In Fig. 11, the ratios from the satellite and model for each area are shown directly in the plot.

On a regional scale, there is excellent spatial agreement between the satellite and model. Ozone formation conditions are NOx limited (FNR > 4) throughout the vast majority of Texas; other studies have found similar conclusions within the last 5 years (Jin et al., 2020; Koplitz et al., 2021). Only along the Houston Ship Channel, near the DFW airport, and in the presence of undiluted power plant plumes are conditions potentially in the transitional regime. When aggregated on an urban scale, the model ratio values are marginally
lower than the satellite-derived ratios, especially in San Antonio and Austin. This low model bias is improved when the AMF of the NO$_2$ product is recalculated. Consistent with the analyses presented in Sect. 3.3 and 3.4, the model appears to be capturing both the HCHO and NO$_2$ spatial patterns with satisfactory performance, and therefore, the ozone production regimes are also captured well. The only areas of strong disagreement are in the presence of power plant plumes and large point sources, which TROPOMI appears to be not fully characterizing.

The downside of low-Earth orbiting instruments is the consistent measurement during the early afternoon. This early afternoon measurement time coincides with (1) a temporary dip in NO$_x$ emission rates, which are largest in the early morning and late afternoon, (2) the peak of the biogenic VOC emissions, which often peak at the time of the maximum daily 2 m temperature, and (3) stronger photolysis rates, which affect both NO$_2$ and HCHO.

We use the CAMx model to investigate the temporal variation in the FNR. In Fig. 12, we show diurnal cycles of column NO$_2$, column HCHO, and the FNR. The NO$_2$ diurnal cycle has a minimum in the early afternoon, driven mostly by the higher photolysis rates and, second, by the relatively lower NO$_x$ emission rates compared to the early morning and late afternoon. HCHO has broad peak in the afternoon, which is likely related to biogenic emissions and secondary formation. However, the HCHO diurnal cycle is flatter than we expected; this may be due to model difficulties in representing complex VOC chemistry for secondary HCHO production (Schroeder et al., 2016; Schwantes et al., 2022).

According to CAMx, the FNR has a temporary maximum in urban areas around 14:00 LT and a minimum around
Figure 11. The formaldehyde/NO$_2$ ratio (FNR) in Texas averaged across April through September 2019, using the (left panel) operational TROPOMI products, (center panel) operational TROPOMI HCHO product and TROPOMI NO$_2$ product with new AMFs, and (right panel) CAMx column amounts. Only CAMx data coincident with the overpass time and valid TROPOMI pixels are included. The ratios from the satellite and model for each area are shown directly in the plot.

Figure 12. Diurnal cycles of column NO$_2$, column HCHO, and the HCHO/NO$_2$ ratio from CAMx for these regions in our model domain, i.e., Houston, Dallas–Fort Worth, and rural eastern Texas (Cass County). The approximate TROPOMI overpass time of 13:30 LT (local time) is denoted by the dotted line.

08:00 LT, with a secondary minimum around 20:00 LT. In the rural areas of eastern Texas, the variation in the FNR is even more substantial than in the urban areas, and even in these rural areas, ozone production might be VOC limited during early morning hours. Therefore, an early afternoon satellite measurement suggesting NO$_x$-limited conditions does not eliminate the possibility of VOC-limited ozone formation conditions in the early morning. This suggests that targeted VOC controls in urban areas of Texas between 06:00–10:00 LT could be an effective way to further reduce ozone concentrations, in addition to expanded NO$_x$ controls at all hours. Upcoming observations from the Tropospheric Emissions Monitoring of Pollution (TEMPO) instrument, which will be located in geostationary orbit, will further help answer this question.

5 Conclusions

In this study, we find that TROPOMI NO$_2$ columns offer a valuable means to validate NO$_x$ emissions inventories, with important limitations. When using locally resolved inputs, simulated urban NO$_2$ columns in Texas agree with TROPOMI to within 20% in most areas. Using data from the newest TROPOMI NO$_2$ algorithm (v2.3.1) generally showed better agreement with the model. We find some evidence that NO$_x$ emissions in certain sections of Dallas–Fort Worth, TX, and Houston, TX, may be underestimated, but the underestimates are within the uncertainty of the methods presented herein.

In the rural areas of eastern Texas, we find generally good agreement to within 20% in most circumstances between the model and TROPOMI NO$_2$ when lightning NO$_x$ emissions are included. In rural regions of eastern Texas, > 50% of the column NO$_2$ appears to be above 2 km in altitude, demon-
strating the influence of the free tropospheric NO$_2$, including lightning. Lightning NO$_x$ emissions can represent up to 24 % of the column NO$_2$ in our eastern Texas domain and presumably would be larger in more isolated tropical regions. Since free tropospheric NO$_2$ has an outsized effect in rural areas, it is critical to have an accurate estimate of free tropospheric NO$_2$ before conducting a model to satellite comparison in these regions (Shah et al., 2022). More aircraft measurements between the top of the boundary layer and the stratosphere–troposphere interface would be helpful to better understand these regions (Shah et al., 2022). More aircraft measurements between the top of the boundary layer and the stratosphere–troposphere interface would be helpful to better understand and quantify free tropospheric NO$_2$.

Over large power plant plumes, however, we find statistically significant differences between the model and satellite measurements. Because the NO$_x$ emissions from these power plants are directly measured, we conclude that TROPOMI cannot distinguish NO$_2$ attributed to power plants from the background NO$_2$ concentrations in Texas. This limitation may be due to short NO$_2$ lifetimes characteristic of that region and, second, the NO$_x$/NO$_2$ ratio in the 4 km model simulation. More work should be dedicated to investigating NO$_2$ and NO$_x$ partitioning near power plant plumes, including aircraft and vertical profilers (e.g., Pandora).

In our comparison between TROPOMI and modeled HCHO, we find excellent agreement in far eastern Texas and the Ozarks but an underestimate in central Texas. This is consistent with Harkey et al. (2020), who showed a model underestimate in the western U.S. More work should be done to evaluate HCHO and VOCs in areas with assumed small amounts of biogenic emissions.

In a last step, we evaluate the ozone formation regimes at the time of the early afternoon TROPOMI overpass. We find that ozone production is NO$_x$ limited almost everywhere in the domain, except near the Houston Ship Channel, near the DFW airport, and in the presence of power plant plumes. There are likely NO$_x$-saturated ozone formation conditions in the early morning hours that TROPOMI cannot observe.

We are encouraged by the future observational strategies that could help tackle some of the remaining questions presented herein. In early 2023, TEMPO will be acquiring column NO$_2$ and HCHO measurements during all daylight hours in the presence of low amounts of clouds. When coupled with the current ground monitoring network, this will elucidate some of the unknown NO$_2$ and HCHO diurnal cycles, giving us more confidence in our understanding of NO$_x$ emissions, NO$_2$ chemistry, and satellite retrievals.

**Appendix A: CAMx model simulation performance**

We evaluated CAMx NO$_x$ and ozone surface concentrations using data collected at TCEQ Continuous Ambient Monitoring Stations (CAMSs). We evaluated the performance by five geographical subregions, namely Austin, San Antonio, Waco, Tyler, and Dallas–Fort Worth. NO$_x$ monitors deployed for routine monitoring have limitations for NO$_2$. These monitors measure NO, and consequently, NO$_2$ is chemically converted to NO for measurement. The converter also captures other compounds, including PAN and a portion of HNO$_3$ (Dickerson et al., 2019). These NO$_x$ monitors have a detection limit of around 1 ppb (parts per billion), but differentiation between NO and NO$_2$ is less accurate near the detection limit. Therefore, we compare both CAMx NO$_x$ (i.e., NO + NO$_2$) and NO$_x$ (i.e., NO + NO$_2$ + PAN compounds + HNO$_3$) to monitored NO$_x$ in Fig. A1. Hourly ozone measurements were aggregated to 8 h maximum daily averages (MDA8), and hourly NO$_2$ measurements were aggregated to early afternoon averages (12:00–15:00 CST) to correspond with the TROPOMI overpass time.

Figure A1 displays the O$_3$ and NO$_2$ performance in the CAMx simulation compared to ground monitors. High observed NO$_x$ detected by ground monitors in urban areas (e.g., > 10 ppb) are not resolved at the 4 km CAMx horizontal grid resolution. As discussed in Souri et al. (2022), care is needed when comparing pointwise measurements to concentrations spatially averaged over large (> 1 km) grid cells. For example, Dallas, Hinton St. (CAMS 0401), is located 0.9 km from a major freeway interchange and 200 m from a busy road (Mockingbird Lane). In contrast, Tyler Airport Relocated (CAMS 0082) is in a rural location, removed from busy roads, and the nearby airport is regional and not highly trafficked. When compared with monitored NO$_x$ in less polluted areas (i.e., < 10 ppb), CAMx NO$_x$ tends to be lower than measured NO$_x$, whereas CAMx NO$_x$ tends to be higher than measured NO$_x$. We therefore conclude that CAMx is consistent with the ambient NO$_x$ measurements within the limitations of the monitoring equipment capabilities and siting.

We present similar scatterplots for maximum daily 8 h average (MDA8) ozone in Fig. A1. CAMx shows skill in identifying low and high ozone days, with $R^2$ values from 0.56 (Austin) to 0.61 (Tyler). CAMx displays a positive ozone bias across all five regions, with the mean bias (MB) ranging from 4.8 ppb (Waco) to 10.1 ppb (San Antonio). Emery et al. (2017) define the criteria standards for MDA8 ozone as ±15 % for normalized mean bias (NMB) and < 25 % for normalized mean error (NME). Only Waco and Dallas–Fort Worth meet the criteria standard for NMB, while all regions, except San Antonio, meet the criteria standard for NME.
D. L. Goldberg et al.: Evaluating NO$_x$ emissions and their effect on O$_3$ production in Texas

Figure A1. CAMx model performance for (left) maximum daily averaged 8 h ozone (MDA8 O$_3$) and (right) midday 12:00–15:00 LT (local time) NO$_x$ and NO$_y$. The model output is compared to the EPA Air Quality System (AQS) ground observations for the five regions of interest in our eastern Texas domain (Austin, San Antonio, Tyler, Waco, and Dallas–Fort Worth).

Appendix B: ERA5 winds

To justify the use of the ERA5 100 m winds (as opposed to another vertical level or interval), we use the NO$_2$ column information from CAMx to determine the weighted-column midpoint. Using the shape profiles described in Fig. 3, we find that 50% of the tropospheric NO$_2$ column in the Dallas–Fort Worth area is below 227 m in altitude (and therefore 50% is above this); this is the weighted-column midpoint. Using the WRF simulation, we find that the 100 m wind speed is 6% slower than the 227 m wind speed in Dallas–Fort Worth. However, as we discuss in Sect. 4.1, errors due to wind meandering (∼25%) are far more critical.

We can then apply uncertainty bounds to this. In the most polluted sections of the city, the column midpoint would be lower (tens of meters), and in the least polluted sections of the city, the column midpoint can be as high as 500 m. While neither of these are appropriate for an area-wide average, they can constrain the uncertainties of the column midpoint. Using the WRF simulation, we find that winds at 500 m are 15% faster and surface winds are 24% slower than the 100 m wind speed (Table B1).
Table B1. Wind speeds at Dallas–Fort Worth for the April–September 2019 average at various vertical levels in comparison to the 100 m wind speed.

<table>
<thead>
<tr>
<th>Wind fields</th>
<th>Ratio to 100 m wind speed</th>
</tr>
</thead>
<tbody>
<tr>
<td>10 m winds</td>
<td>0.76</td>
</tr>
<tr>
<td>100 m winds</td>
<td>1</td>
</tr>
<tr>
<td>227 m winds</td>
<td>1.06</td>
</tr>
<tr>
<td>500 m winds</td>
<td>1.15</td>
</tr>
</tbody>
</table>

Appendix C: NO\textsubscript{x}/NO\textsubscript{2} ratio

To further investigate whether the NO\textsubscript{x}/NO\textsubscript{2} ratio used in our study is appropriate, we probe the CAMx simulation to calculate the NO\textsubscript{x}/NO\textsubscript{2} ratio for the partial columns below 2 km (Fig. C1). The NO\textsubscript{x}/NO\textsubscript{2} ratio above 2 km is inappropriate for use in the EMG method since the column above 2 km represents background conditions and is subtracted out when using the EMG method.

The NO\textsubscript{x}/NO\textsubscript{2} ratio for the partial column below 2 km in urban areas is 1.31 ± 0.02 (Dallas is 1.33, Austin is 1.30, San Antonio is 1.32, and Houston is 1.295). For urban areas, this represents an uncertainty in the NO\textsubscript{x}/NO\textsubscript{2} ratio of less than 10%. Our original assumption of using a NO\textsubscript{x}/NO\textsubscript{2} ratio of 1.32 is warranted.

However, the NO\textsubscript{x}/NO\textsubscript{2} ratio can vary more substantially near large point sources. In the grid cells of the large point source itself, the NO\textsubscript{x}/NO\textsubscript{2} ratio can be as large as 1.52. It is possible that the NO\textsubscript{x}/NO\textsubscript{2} ratio in the model may be underestimated due to the emissions being equally spread out across the 4 km grid cell. NO\textsubscript{x}/NO\textsubscript{2} ratios can be as large as 2 within 100 m downwind of major NO\textsubscript{x} sources, especially under low ozone (< 30 ppb) conditions (Kimbrough et al., 2017). However, further downwind (> 4 km) of these large point sources, the NO\textsubscript{x}/NO\textsubscript{2} ratio quickly converges back to a value of ~ 1.31.

Figure C1. The NO\textsubscript{x}/NO\textsubscript{2} ratio at 13:00 LT for April–September 2019 for partial NO\textsubscript{2} columns below 2 km in altitude, as simulated by CAMx.

Appendix D: Daily TROPOMI NO\textsubscript{2}

Daily images of the TROPOMI NO\textsubscript{2} vertical column densities are shown in Fig. D1. The top set of panels shows the daily images over Dallas–Fort Worth during July 2019. These daily images document that a NO\textsubscript{2} plume can be observed on every day on which there are no clouds. We also plot the daily ERA5 wind speed and direction on each daily panel. ERA5 winds appear to correctly identify the urban plume direction on each day.

The middle set and bottom set of panels (Martin Lake, TX, and Colstrip, MT, respectively), demonstrate the capability of TROPOMI of observing daily plumes from power plants during July 2019. For Colstrip (13 600 t NO\textsubscript{x} yr\textsuperscript{-1}), a plume signature can be visually located on every cloud-free day. However, in Martin Lake, TX (9500 t NO\textsubscript{x} yr\textsuperscript{-1}), a plume signature cannot be visually located on every cloud-free day, even though NO\textsubscript{x} emissions are of a similar order of magnitude as in Colstrip. This suggests that the location and atmospheric conditions in Texas are causing TROPOMI to not fully observe Martin Lake’s NO\textsubscript{x} emissions.
Figure D1. Daily TROPOMI NO$_2$ vertical column densities over three locations (Dallas–Fort Worth and Martin Lake and Colstrip power plants) during each day of July 2019; the July 2019 monthly average is denoted in the top left panel of each location aggregate.
Code and data availability. TROPOMI NO$_2$ v1.3 data (https://doi.org/10.5270/SSP-s4jjg54; European Space Agency, 2019) and TROPOMI HCHO v1.1 data (https://doi.org/10.5270/SSP-tjlxfd2; European Space Agency, 2022a) can be freely downloaded from the Copernicus Open Access Hub (https://5phub.copernicus.eu/dhus/; European Space Agency, 2022b) or NASA Earthdata Hub (https://disc.gsfc.nasa.gov/datacollection/S5P_L2__NO2____HiR_1.html; National Aeronautics and Space Administration, 2022a; https://disc.gsfc.nasa.gov/datacollection/S5P_L2__HCHO____1.html; National Aeronautics and Space Administration, 2022b; and https://disc.gsfc.nasa.gov/datacollection/S5P_L2__HCHO____HiR_1.html; National Aeronautics and Space Administration, 2022c). TROPOMI NO$_2$ v2.3.1 data can be freely downloaded from the S5P-PAL Data Portal (https://data-portal.s5p-pal.com/products/no2.html; European Space Agency, 2022c). NASA SEAC4RS data can be downloaded from NASA Data Archive (https://doi.org/10.5067/Aircraft/SEAC4RS/Aerosol-TraceGas-Cloud; National Aeronautics and Space Administration, 2022c) and was acquired by the UC Berkeley Cohen research team. ERA5 reanalysis hourly data on single levels (https://doi.org/10.24381/cds.adbb2da47; Copernicus Climate Data Store, 2022a) can be downloaded from Copernicus Climate Data Store (https://cds.climate.copernicus.eu/#/home; Copernicus Climate Data Store, 2022b). The IDL code to regrid and process the data is available upon request.

Author contributions. DLG processed the satellite data, produced most of the figures, and wrote the paper. MH, BdF, and LJ provided guidance with the processing of the satellite data and helped to develop several of the figures. JJ performed the model simulation, processed the model output, and compared the model simulation to the ground monitors. GY and TH conceptualized the study, obtained funding, and provided guidance and feedback throughout. All authors helped to edit the text of the paper.

Competing interests. The contact author has declared that none of the authors has any competing interests.

Disclaimer. Publisher’s note: Copernicus Publications remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

The findings, opinions, and conclusions are the work of the authors and do not necessarily represent findings, opinions, or conclusions of the AQRP or the TCEQ.

Acknowledgements. The authors would like to thank the reviewers at TCEQ for their input on the paper. We appreciate feedback from Ron Cohen, on the usage of the SEAC4RS data, and Ted Russell, on the intercomparison near power plants.

Financial support. This preparation of this work has been funded by the Texas Air Quality Research Program (AQRP; grant no. 20TCEQ-070) at the University of Texas at Austin through the Texas Emission Reduction Program (TERP) and the Texas Commission on Environmental Quality (TCEQ). This project has also been funded by grants from the NASA Health and Air Quality Applied Sciences Team (HAQAST, grant no. 80NSSC21K0511), NASA Health and Air Quality (HAQ, grant no. 80NSSC19K0193), and the NASA Atmospheric Composition Modeling and Analysis Program (ACMAP, grant no. 80NSSC19K0946). For the AQRP funding of this project, Daniel L. Goldberg was paid as a consultant through Ramboll.

Review statement. This paper was edited by Jeffrey Geddes and reviewed by two anonymous referees.

References


Luo, C., Wang, Y., and Koshak, W. J.: Development of a self-consistent lightning NOx simulation in large-scale...


National Aeronautics and Space Administration: NASA Earth Data Hub, NASA [data set], https://disc.gsfc.nasa.gov/dataset/SSP_L2_NO2____HiR_1.html (last access: 24 August 2022), 2022b.


National Aeronautics and Space Administration: NASA Earth Data Hub, NASA [data set], https://disc.gsfc.nasa.gov/dataset/SSP_L2_HCHO___HiR_1.html (last access: 24 August 2022), 2022d.


