A simplified non-linear chemistry-transport model for analyzing NO_2 column observations: STILT-NO_x

Dien Wu[1], Joshua L. Laughner[2], Junjie Liu[2,1], Paul I. Palmer[3,4,2], John C. Lin[5], and Paul O. Wennberg[1,6]

[1]Division of Geological and Planetary Sciences, California Institute of Technology, Pasadena, USA
[2]Jet Propulsion Laboratory, California Institute of Technology, Pasadena, USA
[3]School of GeoSciences, University of Edinburgh, Edinburgh, UK
[4]National Centre for Earth Observation, University of Edinburgh, Edinburgh, UK
[5]Department of Atmospheric Sciences, University of Utah, Salt Lake City, USA
[6]Division of Engineering and Applied Science, California Institute of Technology, Pasadena, USA

Correspondence: Dien Wu (dienwu@caltech.edu)

Abstract. Satellites monitoring air pollutants (e.g., nitrogen oxides, $NO_x = NO + NO_2$) or greenhouse gases (GHGs) are widely utilized to understand the spatiotemporal variability and evolution of emission characteristics, chemical transformations, and atmospheric transport over anthropogenic "hotspots". Recently, the joint use of space-based long-lived GHGs (e.g., carbon dioxide, CO_2) and short-lived pollutants has made it possible to improve our understanding of emission characteristics. Some previous studies, however, lack consideration of the non-linear NO_x chemistry or complex atmospheric transport. Considering the increase in satellite data volume and the demand for emission monitoring at higher spatiotemporal scales, it is crucial to construct a local-scale emission optimization system that can handle both long-lived GHGs and short-lived pollutants in a coupled and effective manner. This need motivates us to develop a Lagrangian chemical transport model that accounts for NO_x chemistry and fine-scale atmospheric transport (STILT-NO_x); and investigate how physical and chemical processes, anthropogenic emissions, and background may affect the interpretation of tropospheric NO_2 columns (tNO_2).

Interpreting emission signals from tNO_2 commonly involves either an efficient statistical model or a sophisticated chemical transport model. To balance computational expenses and chemical complexity, we describe a simplified representation of the NO_x chemistry that bypasses an explicit solution of individual chemical reactions while preserving the essential non-linearity that links NO_x emissions to its concentrations. This NO_x chemical parameterization is then incorporated into an existing Lagrangian modeling framework that is widely applied in the GHG community. We further quantify uncertainties associated with the wind field and chemical parameterization and evaluate modeled columns against retrieved columns from the TROPOspheric Monitoring Instrument (TROPOMI v2.1). Specifically, simulations with alternative model configurations of emissions, meteorology, chemistry, and inter-parcel mixing are carried out over three US power plants and two urban areas across seasons. Using EPA-reported emissions for power plants with non-linear NO_x chemistry improves the model-data alignment in tNO_2 (a high bias of $\leq$10% on an annual basis), compared to simulations using either EDGAR or without chemistry (bias approaching 100%). The largest model-data mismatches are associated with substantial biases in wind directions or conditions of slower atmospheric mixing and photochemistry. More importantly, our model development illustrates (1) how NO_x chemistry affects the relationship between NO_x and CO_2 in terms of the spatial and seasonal variability and (2) how assimilating tNO_2 can

quantify systematic biases in modeled wind directions and emission distribution in prior inventories of NO_x and CO_2, which laid a foundation for a local-scale multi-tracer emission optimization system.

1 Introduction

Emissions of air pollutants (APs) and greenhouse gases (GHGs) adversely impact urban ecosystems and environments, human health, and the climate via the moderation of energy budgets (Myhre et al., 2014; Watts et al., 2021). APs and GHGs are directly inter-connected considering they are co-emitted from many combustion sources, suggesting that reductions in GHGs may bring co-benefits in mitigating APs (Cifuentes et al., 2001; West et al., 2013; Lin et al., 2018). Although quantifying emissions in GHGs and APs and understanding their underlying drivers at all scales are equally important, emission estimates beyond a county or city become more relevant in addressing policy-relevant topics such as emission mitigation.

Space-based remote sensors offer an objective perspective to monitoring global air quality and GHGs. These new data enable us to uncover the spatial variability along with the temporal trend and perturbation of anthropogenic emissions. Air quality-related observations have been among the first to demonstrate the capability of satellite remote sensing to globally diagnose air quality (Duncan et al., 2016; Laughner and Cohen, 2019; Jin et al., 2020), constrain emissions across time, space, and sectors (Jiang et al., 2018; Goldberg et al., 2019; Tang et al., 2019; Qu et al., 2022), and evaluate real-world decisions (Lamsal et al., 2011; Demetillo et al., 2020). Leveraging satellite observations in understanding the spatiotemporal distribution of emissions within cities is still limited compared to those city-total estimates. Data and analysis uncertainty further present the main challenge in extracting robust combustion signals from remotely sensed measurements and these uncertainties are amplified in attempts to resolve dynamic flows and heterogeneous combustion activities within cities (Valin et al., 2013; Goldberg et al., 2022; Souri et al., 2022).

Making full use of existing and upcoming satellites that retrieve concentrations of APs and GHGs offers an informative way to target urban emissions from different sources at a policy-relevant scale of a few km. Combining satellite observations of species with different atmospheric lifetimes has enabled studies to diagnose chemical conditions and meteorological processes (Jin et al., 2017; Lama et al., 2022), identify urban plumes, and constrain emissions for the tracer of interest (Wunch et al., 2009; Yang et al., 2023), and obtain observation-based ratios between tracers (Silva and Arellano, 2017; Wu et al., 2022; MacDonald et al., 2022) to infer structural changes in combustion activities (Reuter et al., 2014; Miyazaki and Bowman, 2023). In light of the rapid rise in satellite data volume, it is beneficial to have an analysis system that adequately accounts for the important local-scale processes in interpreting the abundance of GHG and APs in a coupled manner (**Fig. 1**). Analogies to such local-scale systems in a global context include the AP-focused Tropospheric Chemical Reanalysis (TCR-2, Miyazaki et al., 2020) and the GHG-focused Carbon Monitoring System-Flux (CMS, Hurtt et al., 2022). Only a few recent multi-tracer modeling systems aim to bridge CO_2 and NO_2 column measurements (Reuter et al., 2014; Kaminski et al., 2022; Hakkarainen et al., 2023), albeit limits in their modeling tools (elaborated in the next paragraph). In addition, as emphasized in Reuter et al. (2014), most multi-tracer studies rely on emission ratio/conversion ratio from inventories, which can be problematic.

In efforts to interpret CO_2 or NO_x emission signatures from satellite observations, most prior studies used either statistical or inversion approaches. The former approach involves the use of Gaussian plume or Exponentially-Modified Gaussian (EMG) models with input from simple wind information to derive emissions of CO_2 and NO_x (or lifetime if for NO_x) purely from observations in a computationally efficient manner without relying much on prior assumptions of emissions (Nassar et al., 2022; Beirle et al., 2011). These statistical approaches only provide a plume-integrated emission estimate that can be sensitive to the input wind speed and chemical lifetime. Multiple satellite overpasses need to be aggregated with wind direction aligned for a robust fit in the EMG model to obtain emission and lifetimes. It is challenging to infer and evaluate sub-grid cell variations in emissions. The more sophisticated inverse approach involves the use of a chemical transport model (CTM) that comprehensively accounts for atmospheric transport and chemical transformation and a coupled inversion or data assimilation system (e.g., Liu et al., 2022; Qu et al., 2022). CTMs are, however, computationally expensive and often involve hundreds of species and their coupling reactions. Most CTMs used in AP-related studies are Eulerian models, which may suffer from complications caused by rigid model grids (Wohltmann and Rex, 2009; Valin et al., 2011). Motivated by these approaches that rely on a constant lifetime or solve for individual chemical reactions, we have built a modeling framework to balance the advantages and imperfections —i.e., to simplify the chemical transformation process that preserves the non-linear relationship between NO_x emissions and the observed concentration field together with a high-resolution atmospheric transport using a Lagrangian Particle Dispersion Model (LPDM).

LPDMs have been increasingly utilized for emission estimates over the past decades. For instance, the Stochastic Time-Inverted Lagrangian Transport Model (STILT, Lin et al., 2003) building upon HYSPLIT (Stein et al., 2015) has been well adapted to analyze emission signals from all sorts of measurement platforms. STILT was designed to better describe the movement of air parcels only relevant to an observation site and explicitly provide the source-receptor relationship (i.e., the Jacobian matrix) to facilitate efficient atmospheric inversions for optimizing emissions. Besides, LPDMs themselves possess inherent numerical and computational advantages, such as avoiding artificial smoothing of concentration fields by spurious numerical diffusion in confined model boxes (Wohltmann and Rex, 2009; Lin et al., 2013). More importantly, the Lagrangian transport perspective is intuitively coupled with box models that handle chemical reactions. Noticeable examples include STOCHEM (Collins et al., 1997), ATLAS (Wohltmann and Rex, 2009), CLaMS v2.0 (Konopka et al., 2019), and HYSPLIT-based variations including HYSPLIT-CheM (Stein et al., 2000), ELMO-2 for ozone (Strong et al., 2010), and STILT-chem (Wen et al., 2012). These Lagrangian chemical models describe the chemical reactions of each species or lumped group with similar functional groups to calculate chemical transformation along trajectories but vary in the complexity of implemented chemistry and parameterization for turbulent mixing and numerical diffusion. Despite these prior modeling efforts, Lagrangian chemical models are more often adopted to inform the origins of APs but are less commonly used to constrain emissions. Such under-appreciation is in part a result of the heavy computational expenses in solving chemical changes at high frequency via ordinary differential equations (similar to most Eulerian CTMs) and the reliance on external meteorological fields.

To reduce computational costs in dealing with complex chemistry, studies have proposed machine learning techniques or defaulted to a constant-lifetime assumption as a shortcut. Machine learning techniques have been applied to approximate the chemical mechanisms (Keller and Evans, 2019; Huang and Seinfeld, 2022), predict OH field with observational constraint

(Zhu et al., 2022), and calculate emissions (He et al., 2022). Other studies have assumed a constant first-order lifetime to estimate NO_x emissions and emission ratios between NO_x and CO_2 (Lee et al., 2014; Hakkarainen et al., 2023). However, unlike chemically passive species such as CO_2, the chemical tendency of NO_x is not independent of atmospheric advection and turbulent mixing because of the chemically-driven non-linearity between the NO_x lifetime and the NO and NO_2 concentrations (Laughner and Cohen, 2019). More specifically, during the day NO_x is lost through two more permanent pathways of (1) NO_2 + OH to nitric acid and (2) NO + peroxy radicals (RO_2) with a minor branch in producing alkyl nitrates, ANs (**POINT 3 in Fig. 1**). The two pathways compete with one another and either may dominate depending on chemical conditions. Such non-linear dependence of NO_x lifetime or chemical tendency with NO_x concentration must be accounted for to estimate NO_x emissions from atmospheric NO_2 concentrations. Such non-linearity will affect the interpretation of tracer-to-tracer emission ratios from observed enhancement ratios.

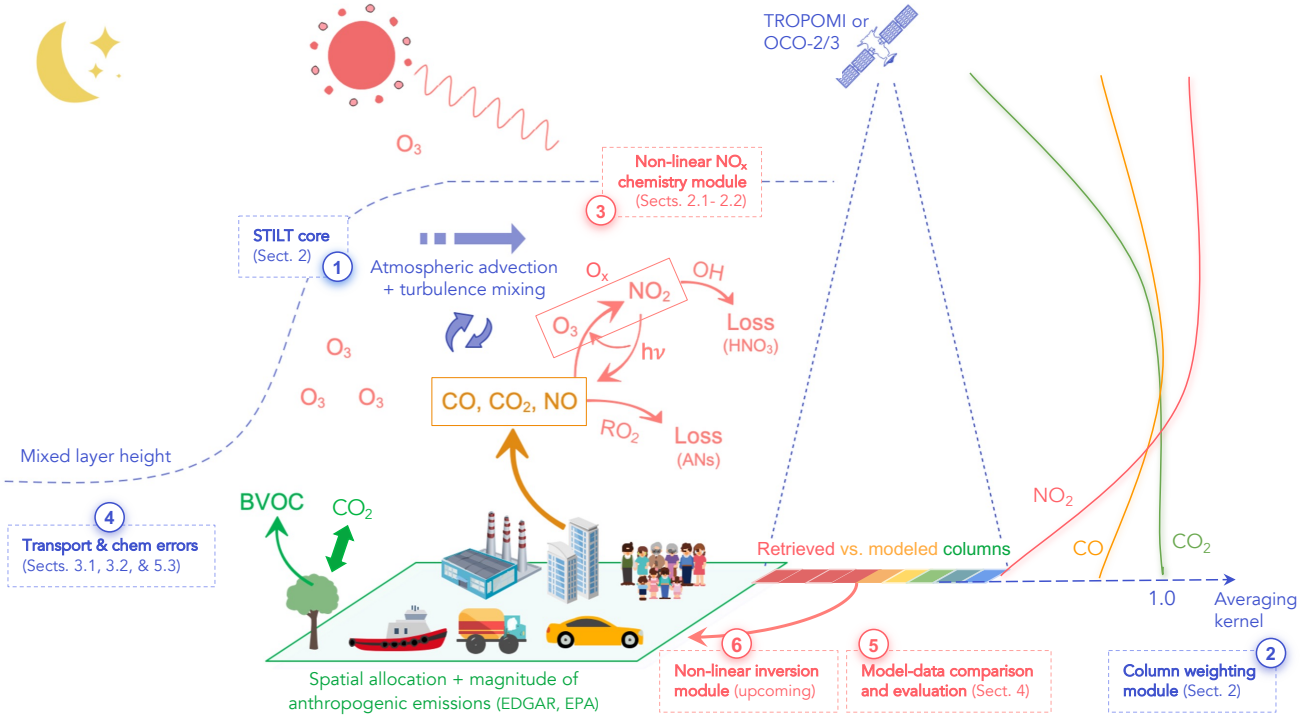

Figure 1. A conceptual diagram of our proposed local-scale multi-tracer modeling framework in interpreting column observations. It contains a road map for this study (POINTs 1 through 5). The diagram highlights key biogenic/physical/chemical processes for quantifying NO_x, CO, and CO_2 around cities based on space-based measurements (pixels from red to blue): atmospheric conditions (wind speed & PBLH for vertical mixing, horizontal mixing/diffusion lengths), chemical conditions (photolysis rate and NO_x regimes, regional versus local oxidant conditions), the spatial distribution of emissions (urban vs. power plant), and sensitivities of the column abundance to individual vertical levels (averaging kernel).

In this study, we present a non-linear modeling framework, STILT-NO_x, to simulate tropospheric column-average NO_2 mixing ratio (tNO_2) as retrieved from TROPOMI. Note that initial NO_2 vertical column density [VCD, molec cm^{-2}] is converted to tNO_2 [ppb] by dividing by a dry air VCD. The dry air VCD is calculated by integrating a profile of the ideal gas number

density of air minus a modeled water vapor profile. As illustrated in **Fig. 1**, the overarching goal of this framework is to facilitate emission optimizations over global anthropogenic hotspots by simulations of the concentrations of key trace gases of CO_2, CO, and NO_x at the local scale. To do so, the current work aims to equip the STILT model with simplified chemistry that avoids explicit calculations of chemical reactions while preserving the non-linearity that ties the NO_x concentrations to its emission (**POINT 3 in Fig. 1**). The proposed STILT-NO_x framework is comprised of four components, which correspond respectively to points 1 to 4 in **Fig. 1** and will be coupled to an upcoming non-linear flux inversion module (**POINT 6**).

1. the HYSPLIT-STILT core that resolves fine-scale atmospheric advection and turbulence; and calculates the sensitivity of concentration anomalies to upwind fluxes ("footprint") (Lin et al., 2003; Fasoli et al., 2018; Loughner et al., 2021); with an additional simplified inter-parcel mixing scheme (**Sect. 2.3**);
2. a column weighting module to simulate atmospheric columns (and uncertainties) that incorporates pressure weighting functions and retrieval-specific averaging kernel profiles (X-STILT, Wu et al., 2018);
3. a simplified chemistry module that describes NO_x chemical tendency (**Sect. 2.1**) and how much NO_x is presented as NO_2 (NO_2–to–NO_x ratio, **Sect. 2.2**);
4. an error analysis module that quantifies errors and biases in wind fields and chemical parameters (**Sect. 3**) following methods initially proposed in Lin and Gerbig (2005) and Wu et al. (2018), which can be used for future flux inversions.

We illustrate the skill of this framework using comparisons of modeled tNO_2 and those diagnosed from TROPOMI over 3 US power plants and 2 cities across seasons (**Sect. 4**). Lastly, we discuss possible future advances in **Sect. 5.3** and demonstrate the benefits of applying this framework, especially on the quantification of CO_2 emissions, emission ratios between NO_x and CO_2, and "near-field" wind biases in **Sects. 5.1 and 5.2**.

2 STILT-NO_x model descriptions

Building upon the HYSPLIT-STILT atmospheric transport core, the STILT-NO_x framework traces the origin of the atmospheric column observed by the satellite and calculates changes of NO_x concentrations due to emissions, inter-parcel mixing, and chemical transformations at the (sub-)minute scale. The STILT-NO_x simulations are conducted in three steps (**Fig. 2**).

First, the backward-trajectory mode records the lat/long/pressure coordinates of air parcels originating from the same atmospheric column sampled by satellites and being driven by the Eulerian meteorological fields (**STEP 1 in Fig. 2**). In this work, we tested two meteorological fields when they are available for each examined region, namely from the Global Forecast System (GFS0p25) and the High-Resolution Rapid Refresh (HRRR) with a respective horizontal grid spacing of 0.25° and 3km (Rolph et al., 2017). As most anthropogenic and all soil sources of NO_x are from the surface, air parcels are evenly distributed and released from the surface to 2 km which is slightly above the typical planetary boundary layer (PBL) height (Wu et al., 2018). To evaluate how representative enhancements between 0 and 2 km are compared to the total tropospheric column enhancements (which can include sources from lightning and aviation), we analyzed vertical distributions of NO_x mixing ratios from TCR-2 (Miyazaki et al., 2020). TCR-2 is a global chemical reanalysis that includes full physical and chemical processes for

various species and assimilates multiple satellite products of NO_2, ozone, CO, and SO_2. As a result, monthly mean NO_x concentrations over the $2^\circ \times 2^\circ$ area around the top 1000 cities is quite insignificant for pressure $\leq$ 700 hPa compared to huge signals within the PBL (**Supplement Fig. S1**). Although 0 to 2 km columns include most anthropogenic enhancements over urban areas, we subtracted a local NO_2 background from the total tropospheric columns to minimize the non-anthropogenic influences with a plume detection algorithm following Kuhlmann et al. (2019). The model-data comparisons with background subtracted are discussed in **Sect. 4.1**.

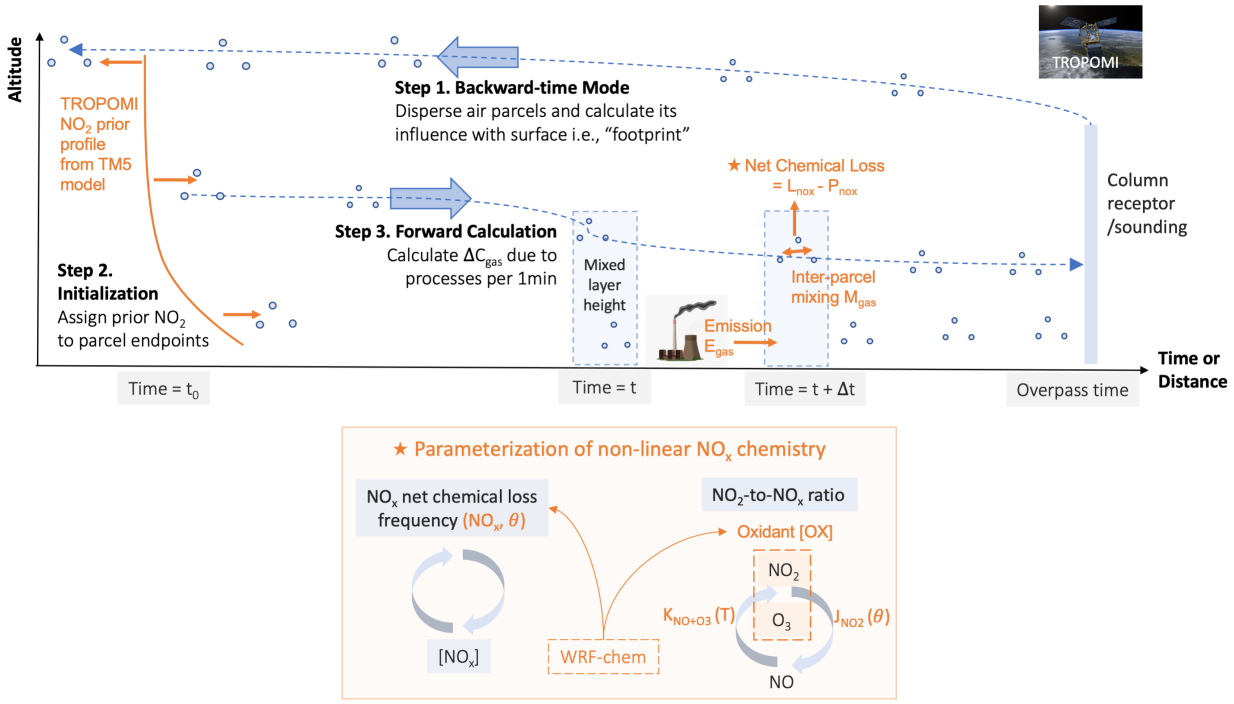

Figure 2. A schematic of STILT-NO_x for simulating concentrations in three steps. **STEP 1** — routine backward-time calculation: record locations of air parcels at each timestamp (Δt) of 1 min or less and their influence from potential fluxes ("footprint"). **STEP 2** — initial condition: the trajectory endpoint at time = t_0 is given a concentration from 4D fields (e.g., TM5 in the case of NO_x). **STEP 3** — forward-time concentration calculation: updates change in concentrations due to emissions, net chemical losses, and inter-parcel mixing along each trajectory at a timescale of $\leq$ 1 min. To clarify, **STEP 3** made use of trajectories originating from a column receptor stretching from the surface to 2 km generated from **STEP 1**.

After being released from a given TROPOMI sounding at the overpass time ($\sim$1 pm local time for nadir soundings), air parcels are dispersed backward in time for 12 hours (time at t_0 in **Fig. 2**). **STEP 1** also provides the STILT "footprint" $[\mathrm{ppm}/(\mu\mathrm{mol\ m^{-2}\ s^{-1}})]$ per air parcel per timestamp (Lin et al., 2003). STILT footprint of a given air parcel is proportional to the time this parcel spends in a small area (of $\sim$100 meters) and describes how the downwind concentration may be altered

if this air parcel is influenced by emissions. A much more complete description of STILT can be found in Lin et al. (2003); Fasoli et al. (2018). The footprint concept, by definition, relies on atmospheric transport and only accounts for concentration changes due to emissions, but not chemical transformations.

Next, $\mathrm{NO_x}$ concentrations at the endpoints of the model trajectory are extracted from the Tracer Model version 5, Massively Parallel version (TM5-MP) to serve as the initial conditions (**STEP 2 in Fig. 2**). TM5-MP is an auxiliary dataset whose $\mathrm{NO_2}$ vertical profiles serve as the prior knowledge facilitating the stratosphere-troposphere separation in L2 $\mathrm{NO_2}$ retrieval (Van Geffen et al., 2022). Here, we simply assume that most $\mathrm{NO_x}$ is presented as $\mathrm{NO_2}$ at nighttime, despite the apparent caveat in neglecting $\mathrm{NO_3}$ chemistry and heterogenous reactions involving $\mathrm{N_2O_5}$.

Once $\mathrm{NO_x}$ is initialized at the time t_0 for the endpoint of every trajectory, we proceed with (**STEP 3 in Fig. 2**) to estimate changes in concentrations due to emissions, chemical transformation, and inter-particle mixing. Mathematically, the concentration per air parcel per timestamp ($\mathrm{C_{p,t}}$) relies on that from the last timestamp following **Eq. 1**:

$$\mathrm{C_{p,t} = C_{p,t-\Delta t} + \Delta C_{emis,p,t}(E, F_{p,t}) + \Delta C_{chem,p,t}(C_{p,t}, \theta_{p,t}) + \Delta C_{mix,p,t}(C_{p,t}, \overline{C_{p_{ngb},t}})} \quad (1)$$

where the time interval for updating concentrations, $\Delta \mathrm{t}$, is defaulted to 1 min or reduced to sub-minute when $\mathrm{C_t}$ becomes nonphysically negative to ensure numerical stability. Concentration gains from emissions, $\Delta \mathrm{C_{emis}}$, result from multiplying STILT parcel-specific footprints ($\mathrm{F_{p,t}}$) with prior emissions (E) from EDGARv6.1 (Crippa et al., 2022) and EPA (United States Environmental Protection Agency, 2022) for power plant cases in this study. We neglect soil $\mathrm{NO_x}$ emissions given the relatively small contributions in cities. Unlike sophisticated CTMs which resolve chemical reactions of an individual or lumped groups of species, concentration anomalies due to chemical reactions, $\Delta \mathrm{C_{chem}}$, are solved in an explicit first-order fashion involving a "net chemical tendency" with a unit of ppb $\mathrm{hr^{-1}}$. Such a chemical tendency ($\mathrm{R_{NO_x}}$ in **Eq. 2b**) is parameterized offline as functions of $\mathrm{NO_x}$ concentrations and solar zenith angles, θ, which is explained in **Sect. 2.1**. The final term, $\Delta \mathrm{C_{mix,p,t}}$, denotes the concentration exchange between a given air parcel and its volumetric neighborhood ($\mathrm{p_{ngb}}$), which is explained in **Sect. 2.3**.

Following these steps, we obtain modeled $\mathrm{NO_x}$ mixing ratio for every trajectory released between the surface and 2 km based on $\mathrm{NO_x}$ curves described in **Sect. 2.1**. To compare against TROPOMI tropospheric $\mathrm{NO_2}$ columns, we account for the fraction of $\mathrm{NO_x}$ that is present as $\mathrm{NO_2}$ (**Sect. 2.2**) and properly weight modeled $\mathrm{NO_2}$ from different altitudes according to pressure weighting function and averaging kernel profiles following Wu et al. (2018). Such an approach in applying averaging kernel (**Fig. 1**) to modeled profiles is equivalent to a more commonly used approach, which re-calculated retrieved $\mathrm{tNO_2}$ as "seen" from the CTM by re-calculating air mass fraction based on modeled $\mathrm{NO_x}$ profiles as investigated in Goldberg et al. (2022). In addition, we evaluate the modeled meteorology and chemistry using a separate set of STILT-$\mathrm{NO_x}$ simulations with "true" $\mathrm{NO_x}$ emissions from EPA for three US power plants (**Sect. 4.1**).

2.1 $\mathrm{NO_x}$ net chemical tendency, $\mathrm{R_{NO_x}}$, and uncertainty

Inspired by the theoretical non-linear curves of $\mathrm{NO_x}$ lifetimes as functions of $\mathrm{NO_2}$ vertical column density and volatile organic compound reactivity ($\mathrm{VOC_R}$) based on a box model in Laughner and Cohen (2019), we extract similar non-linear parameterizations using the Weather Research and Forecasting model coupled with Chemistry (WRF-Chem v4.0.2, Grell et al., 2005).

Focusing primarily on polluted environments, we carried out WRF-Chem simulations for three mid-latitude cities and extracted model outputs from a $2^\circ \times 2^\circ$ region centered around each city. Three cities, namely Los Angeles in the US, Shanghai in China, and Madrid in Spain represent typical megacities in North America, Asia, and Europe. Their varied climatic conditions and sectoral emissions of NO_x, VOC, and GHGs provide a holistic view of the variability of NO_x chemical tendency. While our analyses extended to power plants and cities beyond these three training sites when compared to TROPOMI data (**Sect. 4**), it helps assess the broader applicability of our chemical parameterizations.

Appendix A describes our specific WRF-Chem settings used to generate look-up tables of NO_x chemical loss tendencies, which we will refer to as "NO_x curves" (**Fig. 3**). Of the WRF-Chem settings, the chosen chemical mechanism (RADM2, Stockwell et al., 1990) is the most relevant to the accuracy of these NO_x curves. Despite uncertainties in these WRF-Chem simulations, what matters the most for reproducing the NO_x tendency is how NO_x varies with for example solar zenith angle and ozone, rather than the exact accuracy of NO_x concentrations themselves (from WRF-Chem). Thus, non-chemical components (prior emissions, boundary conditions, and physical processes) in this specific WRF-Chem configuration do not necessarily need to be "perfect" or optimized against observations. We clarify that WRF-Chem simulations had been performed to facilitate the parameterization of NO_x tendency within STILT-NO_x but are not required when running STILT-NO_x.

By leveraging WRF-Chem's chemical diagnostic capability, we derive the net chemical tendency of NO_x within each hour [R_{NO_x}, ppb hr^{-1}] for every model grid within the lower 12 vertical levels (x, y, z). R_{NO_x} is calculated specifically from the cumulative changes in NO and NO_2 concentrations solely due to chemical reactions (i.e., "chem_no2" and "chem_no" in WRF-Chem registry) following **Eqs. 2**:

$$\sum_{h_0}^{h} \Delta C_{NO_x}(x,y,z) = \sum_{h_0}^{h} \Delta C_{NO}(x,y,z) + \sum_{h_0}^{h} \Delta C_{NO_2}(x,y,z) \tag{2a}$$

$$R_{NO_x}(x,y,z,h) = P_{NO_x}(x,y,z,h) - L_{NO_x}(x,y,z,h) = \frac{\sum_{h_0}^{h} \Delta C_{NO_x}(x,y,z) - \sum_{h_0}^{h-1} \Delta C_{NO_x}(x,y,z)}{1\ hr} \tag{2b}$$

where model hour h denotes the start time of each hour interval in the WRF-Chem outputs and z denotes the index of model vertical levels (i.e., from 1 to 12). $\sum_{h_0}^{h} \Delta C_{NO_x}$ describes the cumulative net changes to NO_x concentration given chemical reactions from the initial model hour h_0.

WRF-Chem pixel-specific hourly NO_x rate changes, R_{NO_x}, are then grouped by both SZA (θ) bins with a spacing of 2° and C_{NO_x} bins with equal spacing in log10 scale (**Fig. 3a**). θ is chosen given the close relation to solar radiation under clear-sky conditions and controls the photolysis frequency of ozone and OH production (Rohrer and Berresheim, 2006) when ozone and water vapor abundance remain unchanged. Because the intention in using STILT-NO_x is to inform the relationship between emission sources and satellite NO_2 columns, which are almost always filtered to remove cloudy scenes (i.e., quality assurance of ≥ 0.7), the choice of θ without considering cloud coverage is reasonable. Specifically, these net chemical changes explicitly contain all NO_x-relevant reactions within the WRF-Chem/RADM2 scheme, such as the recycling of NO_x from oxidized odd-nitrogen species like peroxyacetyl nitrate.

The above grouping procedure of R_{NO_x} based on a finite number of bins of C_{NO_x} and θ unavoidably reduces the variability of R_{NO_x} that were directly derived from WRF-Chem. To assess the extent to which the R_{NO_x} variability can be explained by

the selected binning feature variables, we performed a sensitivity test to quantify the deviation of bin-averaged R_{NO_x} from the initial R_{NO_x}. Generally, the R_{NO_x} variability is better preserved over polluted regimes with higher NO_x level > 1 ppb than over low-NO_x regimes (**Supplement Figs. S2ab**). Choosing C_{NO_x} alone better explains the R_{NO_x} variability than choosing SZA or air temperature alone. Including additional variables (e.g., air temperature, NO_2–to–NO_x ratio, and VOC_R) on top of our default choice of SZA and C_{NO_x} marginally improves the prediction of R_{NO_x} except for the inclusion of ozone. However, estimating ozone remains a challenging problem, thereby ozone is not included as a feature variable in this study.

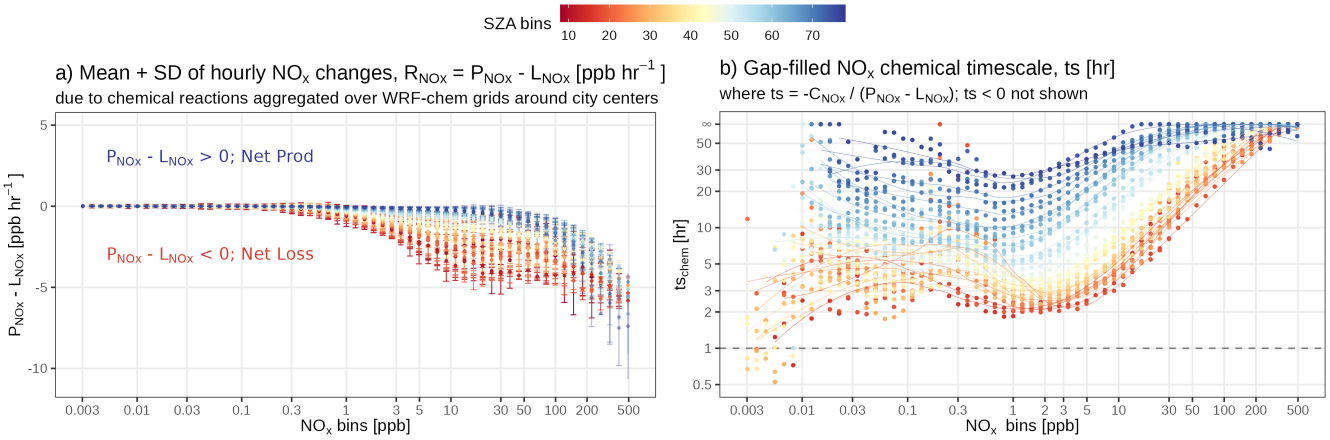

Figure 3. A diagram of NO_x net chemical loss tendency [R_{NO_x}, ppb hr^{-1}] as functions of NO_x concentration (C_{NO_x}) and solar zenith angle (θ). The net loss timescale was first calculated for each 12 km grid cell of all WRF-Chem simulations for 3 cities (with specific model setups summarized in **Appendix A**) and then aggregated into multiple bins of NO_x concentration [ppb]. The NO_x bins are equally divided in the logarithmic space. The solid dots and error bars denote the average and standard deviation of R_{NO_x} within each combined θ and C_{NO_x} bin. For the net loss timescale, only positive values are displayed given the logarithmic scale of the y-axis in panel b and data points with values > 72 hours are simply treated as infinite.

As a net result, R_{NO_x} is mostly negative during the day, meaning NO_x is removed from the system. R_{NO_x} is large with small spread at low θ of $\leq 20°$ and gradually decreases during the day. R_{NO_x} becomes positive as approaches nighttime hours (**Supplement Fig. S2c**) and its variability peaks during sunset when $\theta \in [80°, 100°]$ with a fractional uncertainty of over 100% (blue error bars in **Figs. 3a**) considering the transition to nighttime chemistry. When focusing on the daytime portion with $\theta <$ 70° and $C_{NO_x} \geq 1$ ppb, the spread in R_{NO_x} among WRF-Chem urban pixels ranges from 12.2% to 67.9% according to varied θ and C_{NO_x} (red to yellow error bars in **Fig. 3a**) with an average uncertainty of 41.2%. When focusing on the nighttime portion with $\theta \geq 70°$ and $C_{NO_x} \geq 1$ ppb, the spread in R_{NO_x} spans from 27.9% to over 100% with an average uncertainty of 96.3% largely skewed by the high uncertainty around the dusk hours. Lastly, the average daytime uncertainty in the NO_x tendency at medium to high NO_x concentrations (i.e., 41.2%) will be propagated into chemical uncertainties in tNO_2 for cases of power plants and urban areas, which is further described in **Sect. 3**.

Given the further fluctuation in R_{NO_x} with C_{NO_x}, we define a "net loss timescale" [hr] as $ts_{NO_x} = -C_{NO_x}/R_{NO_x}$ and distinguish it from the conventional chemical lifetime that only accounts for chemical losses. For reference, a positive (or negative) timescale corresponds to a net loss (or production) of NO_x (**Fig. 3b**). The contribution from NO_x production is minor

during noon hours. The non-linear dependence of ts_{NO_x} with C_{NO_x} is largely driven by several NO_x loss pathways: predominately by the loss processes of NO_2 + OH and the formation of alkyl nitrates during the daytime and by the NO_3 chemistry and heterogeneous chemistry at nighttime (**Supplement Fig. S2**). Here, we do not differentiate NO_x curves by VOC_R despite its critical role in determining the turning point when NO_x is mainly lost to either nitric acid or alkyl nitrates (Laughner and Cohen, 2019). We instead perform a sensitivity study of the impact on NO_x curves for three VOC_R intervals in **Sect. 5.3**. Note that these NO_x curves should be considered as a first-order approximation and can certainly be improved upon to evaluate more complex parameterization (**Sect. 5.3**). When it comes to calculating chemical changes within STILT-NO_x per air parcel per timestamp (i.e., $\Delta C_{chem,p,t}$ in **Eq. 1**), such a loss timescale is looked up according to parcel-specific θ and C_{NO_x} to enable the non-linearity core (**Fig. 2** bottom).

2.2 NO_2–to–NO_x ratio

As only the vertical column density of NO_2 is retrieved, the fraction of NO_x present as NO_2 as TROPOMI passed over is an important component of our analysis. Prior studies estimated such ratios using a constant value (e.g., of 0.75) at noon hours across seasons with a 10% uncertainties (Beirle et al., 2011, 2019; Goldberg et al., 2022), monthly mean climatology of ozone from reanalysis (Beirle et al., 2021), and CTMs. NO_x is primarily emitted as NO but converted to NO_2 via the reaction with ozone. During the daytime, NO_2 is photolyzed back to NO with a photolysis frequency, J_{NO_2}. Thus, NO_2–to–NO_x ratio scales with the ratio of ozone and J_{NO_2} (**Eq. 3a**).

Considering the close coupling between NO_2 and O_3, their sum O_x in **Eq. 3b**, is a key indicator for atmospheric oxidant capability in understanding urban air chemistry (Clapp and Jenkin, 2001; Fujita et al., 2016) and informing chemical dynamics (e.g., during COVID, Parker et al., 2020; Lee et al., 2020). O_x levels within PBL can be regarded as a NO_x–independent component related to regional ozone inflow plus a NO_x–dependent component that non-linearly varies with local NO_x and VOC_R conditions (Clapp and Jenkin, 2001; Jenkin, 2004). The complexity in the local O_x-NO_x non-linearity is caused by key reactions behind NO_x curves, which are discussed in **Sect. 5.3.1**.

For simplification, we prescribed a typical O_x level of 50 ppb in the first version of STILT-NO_x and calculate the NO_2–to–NO_x ratio via **Eqs. 3** assuming steady-state:

$$J_{NO_2}(\theta)[NO_2] = k_{NO+O_3}(T_A, P)[O_3][NO] \quad (3a)$$

$$[O_x] = [O_3] + [NO_2] \quad (3b)$$

where J_{NO_2} relies on θ for daytime and the reaction rate coefficient of NO with O_3 (k_{NO+O_3}) is a function of air temperature T_A and pressure P (**Fig. 2** bottom). The inclusion of O_x in calculating NO_2–to–NO_x ratio is to avoid non-physical infinite conversion of NO to NO_2 at high-emitting sources following the titration of ambient ozone. Sensitive tests were performed to reveal how biases in prescribed O_x level may modify the modeled tNO_2 (**Sect. 3**). Typical NO_2–to–NO_x ratios over examined mid-latitude targets across seasons are summarized in **Sect. 5.2**. In the future, satellite observations of tropospheric ozone could be used to add the additional complexity of variable O_x.

2.3 Inter-parcel mixing

Eulerian chemical models usually suffer from too strong mixing or numerical diffusion within their model grid; while Lagrangian models (equivalent to possess extremely high "spatial resolution") may lack any mixing between air parcels that are normally assumed to be independent of one another (Lin et al., 2013; Brunner, 2012). Such lack of mixing has negligible impact on passive tracers as mixing alters only the spatial distribution of concentration among air parcels but not the resultant concentration averaged across parcels at the receptor. However, non-linear processes alter both the spatial distribution of parcel-specific concentrations together with the average resultant concentration. As a result, the calculation of the total $\mathrm{NO_x}$ tendency will be sensitive to how inter-parcel mixing is parameterized. Common ways to realize turbulence mixing are through (1) stochastic processes followed by the exchanging/averaging properties of air parcels found within a certain mixing length (e.g., STILT-chem, Wen et al., 2012), (2) implemented deformation- and instability- driven schemes that rely on atmospheric stability and wind shear/stress characteristics (e.g., CLaMS, McKenna et al., 2002; Konopka et al., 2019), and (3) diffusion approaches that require the vertical gradient of concentrations (e.g., CiTTyCAT and ELMO-2, Pugh et al., 2012; Strong et al., 2010).

Here we follow the STILT-chem approach to enable a process of exchanging concentrations per timestamp among the air parcels in close proximity to each other ($\Delta \mathrm{C_{mix}}$ term back in **Eq. 1**), which smooth the horizontal gradient of concentrations among those air parcels. Specifically, at the timestamp of t, the concentration for a given air parcel p is updated based on the concentration gradient between p and its neighborhood according to a mixing timescale (τ_{mix}) within a grid volume with a mixing length scale of a horizontal area and the mixed layer height for the height as follows:

$$\mathrm{C'_{p,t}} = \mathrm{C_{p,t}} \exp\left(-\frac{\Delta \mathrm{t}}{\tau_{\mathrm{mix}}}\right) + \overline{\mathrm{C}}_{\mathrm{p_{ngb},t}} \left[1 - \exp\left(-\frac{\Delta \mathrm{t}}{\tau_{\mathrm{mix}}}\right)\right] \tag{4}$$

where $\exp\left(-\frac{\Delta t}{\tau_{mix}}\right)$ implies the degree of horizontal mixing and $\overline{C}(t)$ represents the average concentration among air parcels within the mixing volume. The update of $\mathrm{C'_{p,t}}$ from $\mathrm{C_{p,t}}$ responds to $\Delta \mathrm{C_{mix}}$ in **Eq. 1**. A relatively fast mixing time scale of 3 hours and a horizontal mixing length of 1 km is used for testing the mixing impact on modeling $\mathrm{tNO_2}$. Although we neglect the mixing in the free troposphere and the mixing between the mixed layer and the free troposphere in this first model version, we tested a spectrum of the horizontal mixing scales and include possible future improvements (**Sect. 5.3.2**).

3 Model uncertainty in $\mathrm{tNO_2}$ due to wind and chemistry

As atmospheric transport and chemical transformation are the two main components in any CTMs, we assess how uncertainties tied to the modeled wind field, $\mathrm{NO_x}$ loss timescale, and $\mathrm{NO_2}$–to–$\mathrm{NO_x}$ ratio may contribute to uncertainties in $\mathrm{tNO_2}$ in ppb.

$$\sigma^2_{sim} = \sigma^2_{trans} + \sigma^2_{ts} + \sigma^2_{nn}. \tag{5}$$

Here we briefly describe how various $\mathrm{tNO_2}$ uncertainties were approximated based on our understanding of errors in respective model parameters/inputs (i.e., wind error, $\mathrm{NO_x}$ chemical tendency, or $\mathrm{O_x}$ levels). To approximate $\mathrm{tNO_2}$ uncertainties due to transport errors, we followed previous approaches to first assess the GFS- and HRRR- modeled wind profiles against radiosonde, calculate respective error statistics including wind error, correlation time and length scales, and lastly propagate

wind error statistics into errors in column concentrations. Mathematically, σ^2_{trans} in **Eq. 5** is derived from the difference in the variance of STILT-NO_x air parcel-specific NO_2 concentrations between the original simulation and a second simulation with wind error (Lin and Gerbig, 2005; Wu et al., 2018). The derivations of modeled wind errors and contributions to tNO_2 errors are elaborated in **Appendix B**. To evaluate the impact due to errors associated with chemical parameters, we perturbed the NO_x curves or the O_x level according to 20 perturbing factors. Perturbed curves/parameters are used to generate 20 new sets of tNO_2 fields, of which their respective standard deviation among perturbations serves as the chemical uncertainty [ppb] due to NO_x net loss timescale and NO_2–to–NO_x ratio (σ_{ts}, σ_{nn} in **Eq. 5**). These 20 perturbing factors were randomly selected from a normal distribution $N(\mu = 1, \sigma_{param})$. Here we tested out σ_{param} of 40% for NO_x loss timescales according to uncertainties in the chemical tendency (**Fig. 3**) and a σ_{param} of 40% for O_x level (**Eqs. 3**).

Due to heavy computational expenses in conducting such wind and chemical perturbation analyses for all overpasses and locations, we only ran error analyses for a total of six overpasses over a power plant and a city. To cover seasonal changes in NO_2 signals and their uncertainties, overpasses in varied seasons are examined for the New Madrid power plant on Feb 8, June 15, and Dec 8, 2020, and Phoenix on Feb 7, May 27, and Dec 23, 2020. Two winter cases with relatively large signals are shown in **Fig. 4a**. Considering the non-linearity between chemical tendency and NO_x concentration, sounding-specific uncertainties for all six cases are presented against modeled tNO_2 in **Fig. 4b**. When conducting those perturbations, other model parameters like meteorological field and emissions remain unchanged.

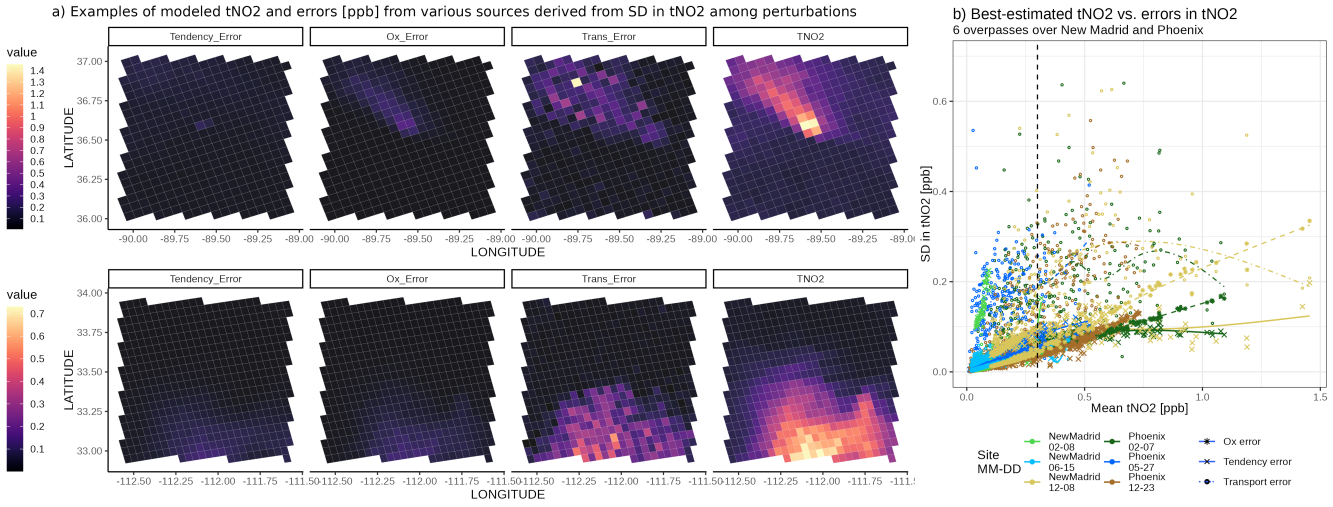

Figure 4. (a) Demonstrations of best-estimated tNO_2 ("TNO2") and their uncertainties [ppb] due to random u-/v- wind errors ("Trans_Error"), NO_x chemical tendency ("Tendency_Error"), and O_x levels ("Ox_Error") on Dec 8, 2020, over the New Madrid power plant and Dec 23, 2020, over Phoenix. (b) Scaling between uncertainties and mean tNO_2 signals over six overpasses for the two targets with smooth splines fitted (crosses with solid lines for tendency errors, stars with dashed lines for Ox errors; and circles with dotted-dashed lines for transport errors). Colors differentiate the sites and TROPOMI overpass times.

As a result, the average percent error in u-/v- wind speed in the PBL is roughly 22% for the New Madrid case (**Supplement Fig. S3a**), which contributes to 50% uncertainty in tNO_2 at the sounding level (3rd column in **Fig. 4a**). Higher transport errors

may more frequently occur if an intensive point source is in the area or over pixels on the border of the NO_2 plumes with moderate signals of about 0.2 to 0.5 ppb (e.g., dots in **Fig. 4b**). This is because small deviation in modeled wind vectors causes air parcels to either "hit" or "miss" the intensive source. The transport uncertainty appears to first correlate positively with the signals and then decreases when signals are sufficiently high, e.g., > 0.7 ppb. Such a decline may be associated with hyper-near-field soundings, where deviation in wind fields may not alter modeled signals as modeled air parcels will always experience large influence from the emission source (dotted-dashed lines in **Fig. 4b**). Compared to power plants, cities may be associated with a more homogeneous transport uncertainty if emissions are more homogeneous and better mixed in the PBL.

Given a roughly 40% uncertainty in O_x levels or NO_x chemical tendency, chemical uncertainties in ppb remain small when modeled signals are compared (**Fig. 4a**). Uncertainties from chemical tendency first increase with tNO_2 signals and gradually plateau for tNO_2 beyond 0.7 ppb, likely because NO_x is lost slowly when the NO_x concentration stays high and further perturbations in chemical tendency are less impactful. In contrast, uncertainty from O_x levels appears to consistently scale against the signals, i.e., more apparent for soundings adjacent to the power plant with reasons explained as follows: When the certain perturbed O_x level approaches zero, the amount of NO that can be oxidized as NO_2 becomes minimal (2nd column in **Fig. 4a**). This case mimics the scenario where O_3 can be titrated in proximity to an intense release of NO before the ozone-depleted plume air is mixed with the ambient ozone-rich air. Nevertheless, considering the entire sample, the percent errors due to chemical parameters remains relatively low (13% to 18% for six cases in **Fig. 4b**).

Whether chemical or meteorological errors dominate the total model errors fundamentally depends on tropospheric NO_2 signals which further rely on factors like atmospheric stability with wind errors, chemical tendency, and emission distribution. Such dependence leads to spatial gradient and seasonal variations in estimated errors as seen from the above examples. In brief, our limited perturbation experiments suggest that transport uncertainties dominate the total modeled uncertainties, except for a few hyper-near-field soundings where chemical uncertainties become more substantial.

For future emission optimizations, uncertainties in the emissions, retrieval, and background should also be included. Despite the significant advance in the TROPOMI NO_2 retrieval version of v2 compared to v1 (Van Geffen et al., 2022), v2 retrieval is associated with a fractional uncertainty (normalized over retrieved tNO_2) of $\sim$30 to 50% for most soundings within the plume. Uncertainties in NO_x emissions between inventories can serve as the prior uncertainty, which is substantial at the pixel level (**Supplement Figs. S4, S5**). Besides the regional wind assessment, a novel plume rotation algorithm based on model-data NO_2 plumes is proposed in **Sect. 5.2** to quantify near-field wind biases.

4 Model-data evaluations and comparisons

The tropospheric NO_2 mixing ratio at a given sounding location is influenced by the regional inflow, atmospheric advection and turbulence mixing, underlying emission characteristics, and chemical changes en route to the sounding (**Fig. 2**). Modeled tropospheric NO_2 mixing ratios using a variety of model configurations are compared against retrieved values from TROPOMI. Such model-data comparisons help evaluate the overall model performance and the roles of individual physical and chemical processes with a naming convention of <MET>_<EMISS>_<GAS>_<PROC> explained as follows:

345 * <MET> represent meteorological fields of either 0.25° GFS or 3km HRRR that are used to drive STILT air parcels.

* <EMISS> represent two prior NO_x emission inventories. EDGARv6.1 with monthly mean emissions and the latest year available of 2018 is the primary one for simulating all cases. Hourly mean emissions from EPA reports are only used to evaluate modeled chemistry and meteorology for several US power plants (**Sect. 4.1**).

* <GAS> represent the simulated species with a default string of "TNO2" without subtracting a localized tNO_2 background. A separate comparison with background subtracted is shown in multi-track comparisons (**Fig. 6**, **Sect. 4.1**).

* <PROC> denotes the physical and chemical processes considered per run. Two main configurations include (1) "DEF" runs with both inter-parcel mixing and chemical parameterization included, and (2) "NOCHEM" runs with mixing but without considering the NO_x chemical tendency. The "NOCHEM" runs do account for the NO_2–to–NO_x conversion but as a constant ratio of 0.74 according to EMG-based studies.

355 Only model-data comparisons using TROPOMI v2 are shown. As satellite averaging kernel and observed tNO_2 differ substantially between v1 and v2, modeled concentrations are weighted by the version-specific AKs to yield apple-to-apple comparisons. Changes in AKs, retrieved and modeled values between versions are summarized in **Supplement Fig. S6**.

4.1 Model validation: US power plants

The New Madrid power plant along the Mississippi River is a 1,300-megawatt coal-fired power station (GEM, 2021), which 360 ranks first in 2020 among US power plants regarding NO_x emissions provided by EPA. Thomas Hill and Martin Lake power plants ranked second and third in 2020, respectively. We also report results for an overpass over the Intermountain power plant in Utah where the surrounding complex terrain is difficult to be modeled properly. Let us start with two examples to illustrate plumes modeled by different model configurations (**Sect. 4.1.1**) and then present model-data comparisons over dozens of overpasses of the three power plants (**Sect. 4.1.2**).

365 4.1.1 Single-track demonstration

The correction in NO_x emissions greatly improves the model-data alignment. For example, EDGAR-based simulations substantially underestimate or overestimate the tropospheric columns (**Fig. 5a1, 5c1**), as EDGAR emissions are almost 1/3 or twice of the reported hourly EPA emissions for New Madrid or Intermountain power plant, respectively (**Supplement Fig. S7**). EPA-based simulations align better with retrieved values from TROPOMI v2 despite deviations over the far-field region (**Fig. 5a2 370 vs. 5a6**). Such improvements in model-data alignment are also inferred from the linear regression slopes reported in **Fig. 5bd**. Not accounting for NO_x chemistry or lifetime elevates NO_2 concentrations both within the plume and over the background even if EPA emissions are assumed to be "correct" (**Fig. 5a3, 5a5**). The inter-parcel mixing with a 3-hour mixing timescale redistributes NO_x concentrations among adjacent air parcels but leads to a minimal impact of $\leq 5\%$ of the modeled tNO_2 at individual column receptors (thereby not shown).

375 The choice of meteorological fields with different spatial resolutions insignificantly affects the modeled signals except for cases surrounded by complex topography and flows. For example, the HRRR-based plumes resemble the GFS-based plumes

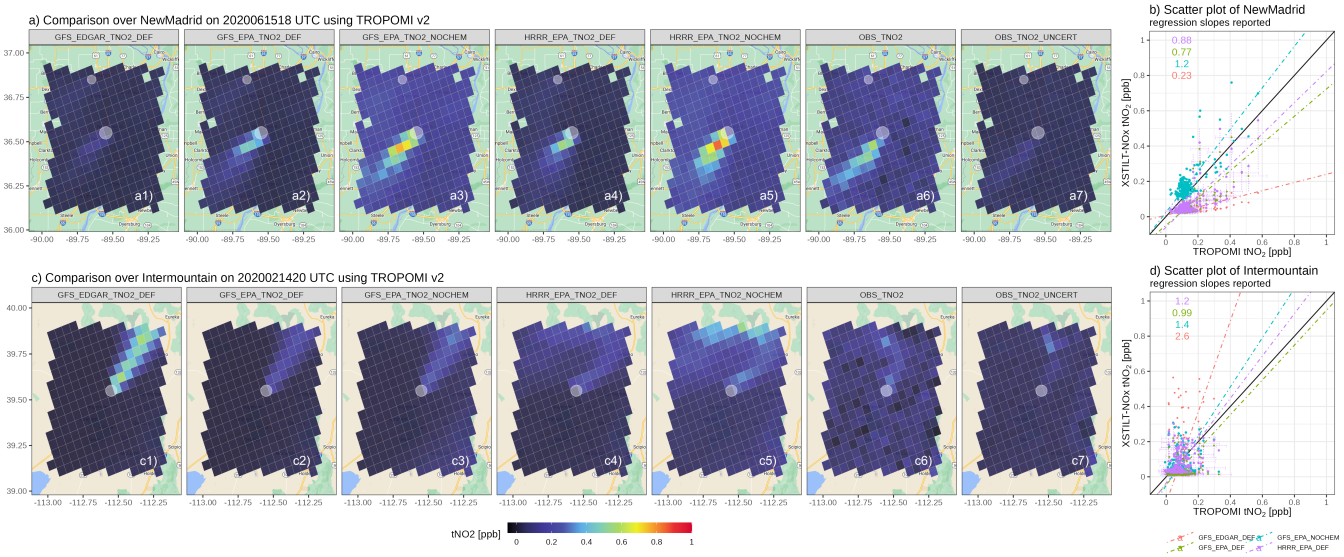

Figure 5. Maps (a) and scatter plot (b) of modeled plumes based on several model configurations (first five columns) versus retrieved plumes + uncertainties from TROPOMI v2.3 (last two columns) for the New Madrid power plant on June 15, 2020, and Intermountain, Utah power plant on Feb 14, 2020. Varied model configurations are labeled on the top of each panel, following the naming convention of "<MET>_<Emiss>_<GAS>_<PROC>" explained in the list of **Sect. 4**. In particular, "_DEF" and the "_NOCHEM" denote the modeled columns using the default (with mixing and chemistry) and non-chemistry configurations. Grid cells with intensive NO emissions from EDGARv6 are labeled as white circles with sizes denoting the relative emission magnitude. The type II linear regression slope is fitted for each configuration (dotted-dashed line) and modeled and retrieval uncertainties are added (dashed error bars).

for the New Madrid power plant (**Fig. 5a2 vs. 5a4**), which is also revealed by their similar wind error statistics (**Supplement Fig. S3a**). However, complex terrain and stable PBL during wintertime complicate and usually worsen the model performance as a result of increased meteorological errors. On Feb 14, 2020, EPA-based plumes over the Intermountain power plant in Utah using two meteorological fields differ substantially from each other and they both deviate from the observed plume regarding the plume shape. GFS delineates the mean wind direction within its coarser 0.25° grid box while 3km HRRR offers more spatial variability in wind directions (**Fig. 5c2 vs. 5c4**). Yet, precisely capturing the curvature in the wind vector is extremely challenging even using 3km meteorological fields (**Fig. 5c6**) and more difficult using Gaussian plume approaches that rely on only one effective wind vector. Such model-data mismatch in plume shapes can further help quantify the wind biases, which are discussed in **Sect. 5.2.1**.

Besides modeling challenges, retrieval uncertainty cannot be neglected, as it ranges from 22% to 31% of the retrieved signal for the New Madrid case (**Fig. 5a7**) and up to 100% for the Intermountain case (**Fig. 5c7**) at the sounding level. When using retrieved data and averaging kernel from TROPOMI v1, the regression slope becomes 1.18 and 1.25 (**Supplement Fig. S8**), indicating that modeled plumes using both meteorological fields are larger than observed plumes. While using TROPOMI v2, the respective slopes are 0.88 and 1.2 (**Fig. 5bd**). This again emphasizes the substantial uncertainty in retrieved signals, large

enough to even alter the conclusion of whether emissions are underestimated or overestimated for a single overpass; and the need for analyzing multiple overpasses for evaluations (**Sect. 4.1.2**).

4.1.2 Multi-track evaluation

To provide a broad impression of the model performance, we expand the model-data comparisons to a total of 50 TROPOMI overpasses across all seasons in 2020 including 34 overpasses for the New Madrid power plant and 9 and 7 summertime overpasses for the Thomas Hill and Martin Lake power plants, respectively. These overpasses are selected based on their relatively intense signals compared to the surrounding. Model-data comparisons for all overpasses are shown on maps in **Supplement Figs. S9 –S12** with linear regression slopes reported and summarized in **Fig. 6** and **Table S1**.

Cases with slopes deviating significantly from 1 are usually associated with substantial near-field wind directional biases. For instance, modeled wind vectors on March 11, April 28, and Sept 9, 2020, have directional biases of > 30 degrees (**Supplement Figs. S9b, S10b**), which explain the respective abnormal linear regression slopes of -1.75, 0.49, and 3.2 (**Fig. 6**). EDGAR-based simulations are biased too high or too low by a factor of two or more compared to observed values from TROPOMI v2.3 (green dots in **Fig. 6**), driven by biases in EDGAR emissions (**Supplement Fig. S7**). The "NOCHEM" simulations without the account of NO_x losses overestimate tNO_2 by a factor of two across all seasons and three power plants, regardless of the meteorological or emission fields adopted (empty circles in **Fig. 6**). "Upgrading" meteorological fields to a higher resolution seems to contribute less to the improvement of model-data agreements than "correcting" emissions or chemistry. In the end, modeled values with NO_x chemistry and correct EPA emissions using either GFS or HRRR yield the best agreement with retrieved values from TROPOMIv2 (orange dots and lines in **Fig. 6**). Aggregating results of all overpasses, simulations using the "best" knowledge of emissions, the simplified chemistry, two different meteorological fields, and inter-parcel mixing are slightly high biased (regression slope up to 1.2, **Table S1**). RMSE values between observed and modeled tNO_2 when enabling NO_x chemistry range from 0.11 to 0.15 ppb (**Table S1**), which is comparable to the random uncertainty in the NO_2 retrieval of 0.09 ppb.

Statistics discussed above compared the *total* tropospheric NO_2 columns from the model and TROPOMI for soundings around each power plant. It is noticeable that modeled tNO_2 uncontaminated by emissions (i.e., background tNO_2) are sometimes slightly lower than observed background tNO_2 (**Supplement Figs. S9a, S10a**), possibly because higher chemical uncertainties are related to low-NO_x regimes and non-anthropogenic NO_x sources from soil and lightning are excluded from current simulations but can play a bigger role of tNO_2 over rural regions (Goldberg et al., 2022; Shah et al., 2022). In particular, column contributions from lightning NO_x emissions aloft may be amplified since TROPOMI NO_2 retrieval has a higher sensitivity towards the free troposphere than PBL. Since our current model setup only accounts for anthropogenic NO_x sources below 2 km, we conducted an additional test by subtracting the background tNO_2 from total tNO_2 to arrive at observed anthropogenic enhancements (the second paragraph in **Sect. 2**), assuming soundings within or outside the plumes have equal contributions from the nearby non-anthropogenic NO_x sources. After subtracting the tNO_2 background, the model-data comparison based on observed tNO_2 enhancements does not change dramatically (e.g., orange crosses vs. orange solid dots in **Fig. 6**).

In summary, using more accurate NO_x emissions with chemistry considerably improves the model-data comparison. Increasing the spatial resolution of meteorological fields has less impact on cases with relatively flat terrain. Larger model-data mismatches generally are associated with larger wind directional biases. Modeled values in tNO_2 may be slightly biased low in summer months from April to June and high in winter months from Nov to Feb with minimal annual biases, assuming EPA emissions and observed tNO_2 are unbiased.

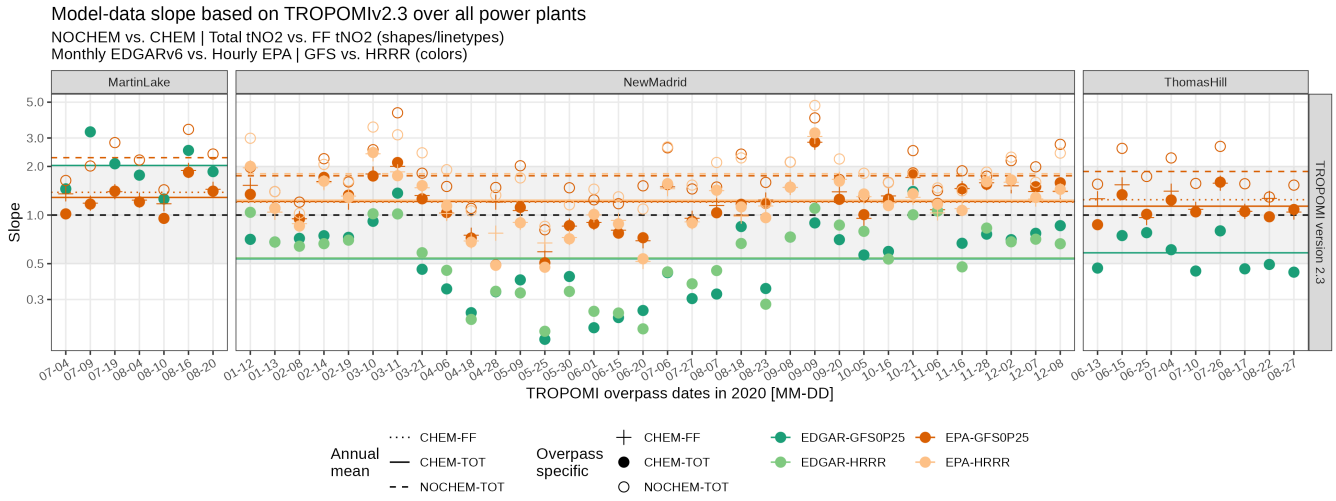

Figure 6. A summary figure of the linear slope between the observed tNO_2 and simulated tNO_2 using a variety of model configurations over all three US power plants. Model configurations include simulations (1) with or without NO_x chemistry parameterization (empty vs. solid dots), (2) using default EDGAR or scaled emissions from EPA (green vs. orange dots), (3) using 0.25° GFS or 3km HRRR (dark green/orange vs. light green/orange dots), (4) using total tNO_2 or background-removed local enhancements (e.g., CHEM-FF as crosses). Annual mean slopes are displayed as horizontal solid or dashed lines. The model evaluation here uses TROPOMI v2.3 and emissions from EDGARv6.1. Evaluations based on TROPOMI v1.3 and annual mean EDGARv5 emissions are shown in **Supplement Fig. S13a**.

4.2 Model application: two cities

We now move to city cases including an industrial city, Baotou in China, and one of the fastest growing megacities in the US, Phoenix. As CO_2 and NO_x are commonly co-emitted into the atmosphere, observed XCO_2 enhancements derived from OCO-3 Snapshot Area Mapping (SAM) mode are displayed with observed tNO_2 (**Fig. 7**). Background XCO_2 is defined as the mean values over the background region that is determined by NO_2 plumes (modified from the background approach in Wu et al., 2018). Both cities possess relatively richer OCO-3 SAM observations co-located with TROPOMI data. Since no "true" NO_x emissions are available for cities, EDGAR is utilized as the prior emission inventory for simulating tNO_2 and optimizing NO_x emissions.

We simulated 18 and 12 TROPOMI overpasses respectively for Baotou and Phoenix (**Supplement Figs. S14, S15**) and first presented one example per city in **Fig. 7**. Baotou is surrounded by four point sources suggested by EDGARv6 but one

large source in the city center informed by both the observed tNO_2 and XCO_2 enhancements on Oct 3, 2020 (**Fig. 7a**). Such a mismatch is confirmed by the comparison of normalized tNO_2 across all 18 TROPOMI overpasses with various wind speeds and directions (**Supplement Fig. S14b**), suggesting that EDGAR very likely misallocated anthropogenic NO_x sources. Similarly, the largest emission source to the east of the city center of Phoenix according to EDGAR seems suspicious and may again be misplaced once simulating more overpasses (**Supplement Fig. S15b**). The observed plume is more concentrated near the city center compared to the HRRR-derived plume that disperses farther away from the city center on Dec 23, 2020 (**Fig. 7b**). Such a spatial offset of the tNO_2 plumes is likely due to an overestimation in the modeled wind speed, pushing the plumes to the southern edge while diluting tNO_2 values over the urban core.

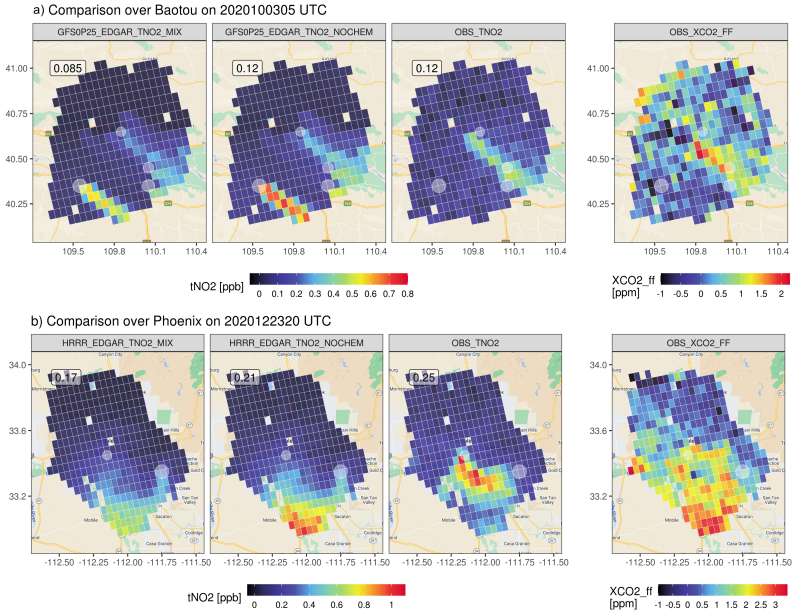

Figure 7. An example of GFS-based tNO_2 plumes over Baotou on Oct 3, 2020 (a) and the HRRR-based tNO_2 plumes over Phoenix on Dec 23, 2020 (b). Modeled plumes are generated using annual mean ENO_x from EDGAR with top emitters highlighted in light-grey circles. Both observed tNO_2 and the anthropogenic XCO_2 enhancements from OCO-3 are plotted. XCO_2 enhancements calculated from a local background have been averaged based on the TROPOMI sounding size. Overpass time differences between TROPOMI and OCO-3 for the two cases are < 1 hour. TROPOMI observations are cropped to match the boundary of the available OCO-3 soundings.

When more overpasses are examined, the model captures well the seasonal variation in tNO_2 —i.e., higher/lower values in winter/summer months (**Supplement Figs. S14a, S15a**). Other than emission biases that affect all cases, a few overpasses stand out for Baotou with poorer agreements with TROPOMI likely owing to (1) clear biases in wind direction on May 31, Aug 9, and Dec 15, 2020, and Feb 19, 2021, and (2) a likely overestimation in STILT footprint that may trigger several effects on Sept 29, March 29, 2020, and Oct 16, 2021 (**Supplement Fig. S14a**). Although STILT can characterize sub-grid cell turbulent mixing by its stochastic nature, the quarter-degree GFS may be insufficient to resolve the complex terrain and air flows, contributing to biases in wind directions and PBL heights over mountainous locations (Lin et al., 2017) such as over Baotou. Deviations

in PBLH may cause a cascade of effects: deciding up to what height the emissions are diluted, whether such height is above or below the emission/plume height, and chemical changes along the way. Such effects may be magnified under low-mixing low-wind conditions where the model particularly struggles with the accuracy of PBLH. Without much mixing between plume and background air, the prescribed available O_x level may be overestimated adjacent to intensive NO_x sources. Overestimation in NO_2 concentration may further be amplified considering the dependence of NO_x rate changes on its concentration. Hence, concentrations of chemically reactive species under low-mixing scenarios are extremely challenging to be modeled properly with an extreme during the nighttime.

5 Discussions

Our ultimate goal is to explore what can be learned about the emission characteristics from anthropogenic hotspots with the joint use of space-based NO_2, CO, and CO_2 plumes. As an intermediate step, this study is informed by previous efforts in extracting and constraining urban CO_2 emissions from satellites using Lagrangian framework (Wu et al., 2018; Roten et al., 2022) and extends it to the interpretation of tropospheric NO_2 satellite data. To diagnose NO_x emissions from NO_2 column signals, we need to effectively account for how NO_x evolves in air parcels from its initial source to locations sampled by TROPOMI, making the Lagrangian perspective an ideal candidate. Now we discuss when and how such a framework can be of most use and possible future improvements.

5.1 Model advantages and flexibility

At the urban extent stretches a few hundred kilometers, our framework accounts for atmospheric transport and chemical transformation in a more rigorous way than typical statistical approaches such as the EMG method, and in a more efficient way than full-chemistry models that explicitly resolve individual chemical reactions.

Another advantage of the STILT-NO_x design is that each of the three main components (trajectory calculations representing air transport, chemical production or loss of the target specie, and optimization of emission) are independent and each can reuse previously saved output from the others. For example, if one wanted to test how sensitive model concentrations were to the chosen chemical scheme, the simulations of atmospheric transport can be reused via the storage of trajectory-based modeling, thereby reducing the computational cost. Our prototype demonstrates a global solution of the NO_x chemical tendency parameterized by the one set of "NO_x curves" in **Fig. 3**. Although simplification may be thought of as a limitation, one can easily replace those default curves with alternatives that are tailored toward a specific region or regime of interest. Such flexibility can inform us of the influence on modeled columns from NO_x curves derived from different chemical mechanisms. Similarly, one can investigate the sole meteorological influence by diversifying the meteorological and mixing parameters. Moreover, because air parcels in LPDMs are not tied to a certain atmospheric tracer, we can estimate concentrations of various species along model trajectories. It allows us to constrain emissions for multiple atmospheric constituents in a consistent framework, which may shed light on tracer-tracer analyses (**Sect. 5.2**).

The Lagrangian modeling approach has its inherent benefits. Firstly, the generation and recording of trajectories can easily reveal the source regions only relevant to a specific satellite sounding and the sub-city scale variations in emission characteristics (Wu et al., 2022). In addition to storing lat/long coordinates and extrapolated meteorological quantities along every trajectory at each timestamp, STILT-NO_x outputs and records NO_x concentration changes due to every process including emission, net chemical changes, and inter-parcel mixing at minute scales. Those trajectory-level concentration changes are further driven by several model configurations listed in **Sect. 4**, which facilitates model debugging and comprehends modeled results. See **Sect. 5.2** for one of the applications. Secondly, the spatial resolution of concentration calculations is not bounded by the rigid boundary of model grid cells, which is particularly important for dealing with non-linear processes for chemically active species. As demonstrated in several studies (e.g., Valin et al., 2011), the grid-average concentration may undergo excessive mixing in Eulerian models, and the concentration-driven chemical tendency varies with the adopted spatial resolution. While the Lagrangian perspective solves for concentration changes at extremely high spatiotemporal resolutions, inter-parcel mixing schemes can be implemented to "smooth" the concentration gradients, whereas it may be challenging to "recover" the sub-grid cell concentration gradients in the Eulerian framework unless increasing the spatial resolution.

More broadly, the proposed simplified parameterization of the non-linear NO_x tendency or NO_x curves is not limited to the STILT framework and can potentially be incorporated into other Lagrangian modeling frameworks or even Eulerian frameworks with a fine spatial resolution to resolve the local variability in chemistry.

5.2 Implications for constraining urban CO_2 emission and emission ratios

The knowledge learned from analyzing NO_2 plumes can be transferrable to constraining bottom-up CO_2 emissions. Two main sources of biases influencing the urban CO_2 emission constraint include biases in wind direction and emission locations. Model-data mismatches in NO_2 columns have shown great value in easily identifying the biases with emission locations even without deploying atmospheric inverse analyses (**Sect. 4.2**), especially for point sources in urban areas when plumes from multiple sources are less overlapping with one another. Additionally, diagnosing the NO_x emissions can facilitate CO_2 emission estimates in two ways: it can reveal systematic biases in near-field wind directions (**Sect. 5.2.1**) and by quantifying the NO_x to CO_2 emission ratios for point and/or area sources assist in sector-based attribution (**Sect. 5.2.2**).

5.2.1 Quantifying wind bias

In addition to leveraging limited radiosonde measurements, the obtained modeled and retrieved NO_2 plumes can be used to *quantify* wind biases, to improve the accuracy of top-down CO_2 emission constraints, whether or not employing conventional atmospheric inversions. To do so, we conducted a second wind assessment involving a plume rotation algorithm. In brief, a NO_2 plume from either model or retrieval is rotated clockwise (α from -180 to -5° with a spacing of 5°) or counter-clockwise (from 5 to 180°) around the emission source and then resampled onto the original TROPOMI pixels (**Supplement Fig. S16**). tNO_2 from an original and a rotated plume are multiplied to arrive at a cross-product of tNO_2 in ppb^2, analogous to the concept of "cross-correlation". The original or the rotated plume can be chosen from either model or observations and their

normalized cross-product can be expressed as a function of rotating angles (colors in **Fig. 8a**). More details on intermediate steps to calculate such a function are described in **Appendix C**.

As a result, the width of the Gaussian-shaped curve of the cross-product (as measured by the σ parameter of a Gaussian fit) reflects the bias in the plume shape resulting from horizontal dispersion. A larger area under the Gaussian curve indicates a greater overlap between the initial plume and the rotated plume. More importantly, deviations in the central line of the Gaussian fit away from zero (as measured by the μ parameter) imply possible biases in the "near-field" wind direction for each TROPOMI overpass (**Fig. 8b**). Specifically, wind directional biases of both GFS and HRRR appear to be smaller from May to early Sept than the remaining months (**Fig. 8b**). A few outliers stand out due to large wind biases on March 11, April 28, Sept 20, Oct 5, Oct 16, and Nov 16, 2020.

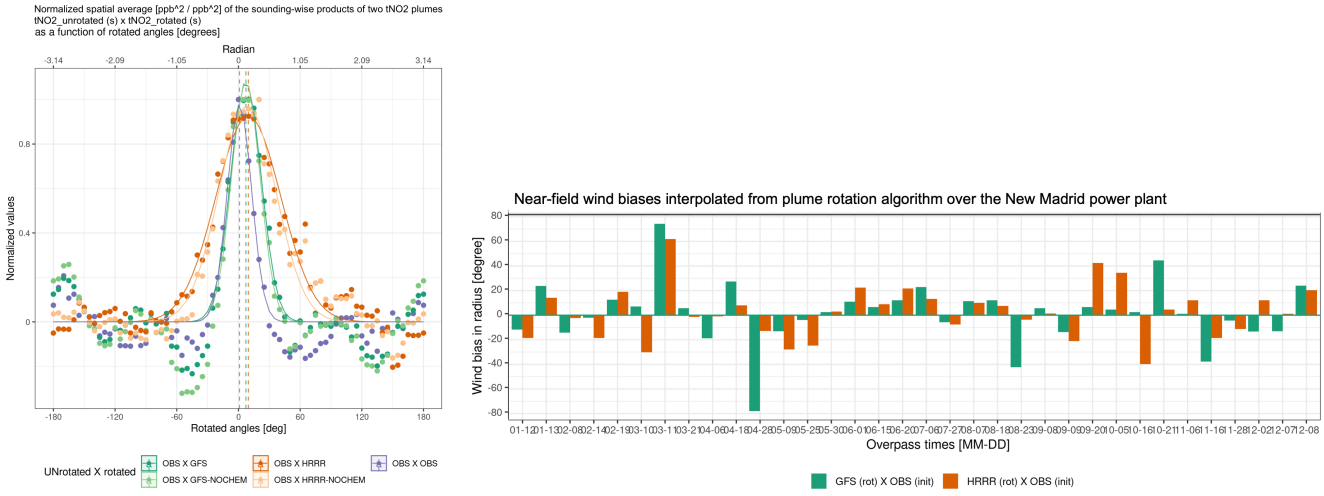

Figure 8. (a) An example of the normalized spatial mean of the sounding-wise product $[\mathrm{ppb}^2/\mathrm{ppb}^2]$ between an un-rotated observed $\mathrm{tNO_2}$ and a rotated plume for the New Madrid power plant on June 15, 2020. Gaussian-like curves are fitted to each set with mean and standard deviation indicating modeled wind biases. Five sets of rotated plumes include observed $\mathrm{tNO_2}$ (purple) and simulated $\mathrm{tNO_2}$ driven by GFS (dark green) or HRRR (dark orange) with or without the account of $\mathrm{NO_x}$ lifetime (light green or light orange). The horizontal dashed lines denote the μ parameter that can translate into wind bias in degrees or radians. (b) "near-field" wind directional bias quantified by the modeled $\mathrm{tNO_2}$ plumes using 3km HRRR (orange bars) or 0.25° GFS (green bars) and retrieved $\mathrm{tNO_2}$ plumes for every examined TROPOMI overpass (y-axis, in degrees) following the rotation algorithm in panel (a).

By identifying those outliers with strong wind directional biases, one can consider either removing those cases or assigning a larger observational uncertainty when attempting to constrain emissions of $\mathrm{NO_x}$, CO and $\mathrm{CO_2}$, assuming their emissions are mostly co-located. Alternatively, we can use this rotating algorithm to create a model plume with minimized wind directional bias before being fed into atmospheric inversions or data assimilation systems (which usually deal with random uncertainties). A more sophisticated approach would be to optimize the emission and wind field simultaneously (Liu et al., 2017). More investigations may be needed to examine the degree of freedom of such a wind-emission optimization framework.

5.2.2 Quantifying emission ratios between NO_x and GHGs

Our modeling development offers additional insights into the discrepancy between emission ratios at the sources and directly-observed enhancement ratios between two species with different chemical lifetimes. The joint use of NO_2 and CO_2 has enabled the calculation of emission ratios by adopting a spatial constant NO_x lifetime (MacDonald et al., 2022; Hakkarainen et al., 2023), and the constraint of CO_2 emissions using NO_2 plumes by adopting inventory-based emission ratios (Zheng et al., 2020; Zhang et al., 2023). However, inventory-based emission ratios might not be well constrained, and the impact of how NO_x decays over time and space on observed ENO_x –to–ECO_2 emission ratios have not been comprehensively assessed, which may impair the ability to accurately quantify such observed emission ratios (Kuhlmann et al., 2021).

Thanks to the ability of our model in tracking NO_x and NO_2 concentrations along trajectories with different model configurations, we can provide an assessment of the influence of atmospheric chemistry on estimating emission ratios. Specifically, the impact from NO_x net losses from each satellite sounding (s) is specified as the ratio of modeled tropospheric NO_x with chemistry over that without chemistry: $\gamma_{ts,s} = NO_{x,CHEM,s}/NO_{x,NOCHEM,s}$. Because NO_x is simply treated as a passive tracer like CO_2 in the NOCHEM simulations, γ_{ts} are naturally smaller than one. Lower γ_{ts} corresponds to faster NO_x chemical frequency and more chemical losses en route to the sounding location, suggesting that NO_2–to–CO_2 enhancement ratios derived directly from satellites need to be scaled up to render the ENO_x–to–ECO_2 emission ratios at source locations.

We calculated γ_{ts} for every sounding and present their distribution as histograms in **Fig. 9a** or as a function of the distance from the emission source (**Fig. 9b**). γ_{ts} ranges from 0.24 to 0.61 for three power plants and from 0.42 to 0.84 for three cities, where lower values correspond to summer months (green bars in **Fig. 9a**). That is to say, the directly observed NO_2 –to–CO_2 enhancement ratios may have to be scaled up by 1.2 to even 4 times across seasons to properly "recover" the NO_x being lost en route from emission sources to the sounding locations. Not properly accounting for such an effect leads to an underestimation of derived emission ratios from satellites. More importantly, discrepancies between enhancement ratios and emission ratios, reflected by γ_{ts}, are not spatially uniform. γ_{ts} gradually decline as soundings move away from the emission sources (**Fig. 9b**). Soundings located farther downwind from emission sources tend to undergo more chemical transformations, likely because NO_x losses become more rapid as NO_x concentrations become lower by atmospheric dispersion (triggering positive feedback). We clarify that only downwind soundings affected by major NO_x emissions are included in **Fig. 9b**; and simulations with or without chemistry have included the effect of atmospheric dispersion as distance increases. Furthermore, how quickly γ_{ts} decline with distance depends on the wind speed and heterogeneity in emissions. For example, the faster the wind may be or the more isolated emissions there are, the steeper γ_{ts} decline with distance. γ_{ts} at the distance of zero are much lower than one in summer, which suggests that chemical transformation can affect the NO_x inflow. We further observe slight differences in the distribution of γ_{ts} for cities versus power plants. Histograms of both tropospheric NO_x (**Supplement Fig. S17**) and γ_{ts} over cities are associated with a wider spread than power plants because cities contain a wider spectrum of emission types and intensities.

Lastly, enhancement ratios need to be adjusted considering the NO_2–to–NO_x ratio and differences in averaging kernels among two retrievals. The medians of our estimated NO_x–to–NO_2 ratio over power plants and cities range from 1.33 to 1.66,

which generally aligns with previous studies of around 1.32 (Beirle et al., 2011; Goldberg et al., 2022). Our estimates are lower in winter than in summer and can be as large as 2 or 3 for a few soundings experiencing intense NO_x sources (**Fig. 9c**).

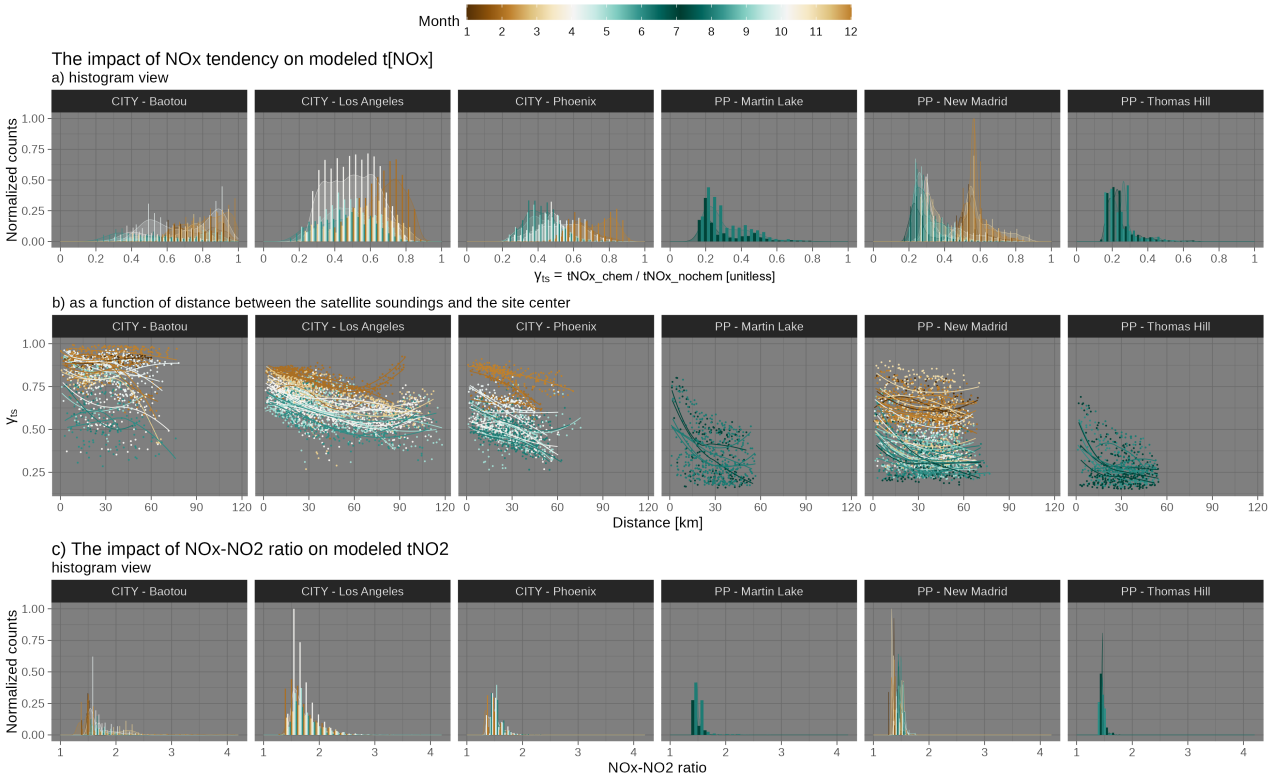

Figure 9. A quantitative metric of the impact from NO_x chemistry on tropospheric NO_x and NO_2 columns. Ratios in tNO_x between simulations with and without chemistry are calculated as $\gamma_{ts,s} = tNO_{x,CHEM,s}/tNO_{x,NOCHEM,s}$, which is displayed as a histogram (a) and as a function of the distance of the satellite sounding from the site center (b). Soundings in summertime overpasses are colored in dark green whereas brown for soundings in the dormant months. Soundings of all overpasses for all city and power plant cases are included in histograms. Only downwind soundings in the NO_2 plumes are included in the distance panel (b) with a smooth spline fitted per overpass to reveal the anti-correlation. The ratio between modeled tropospheric NO_x column versus the tropospheric NO_2 column is derived from each sounding to reveal the NO_2–to–NO_x ratio influence (c). As a reference, most previous studies adopted a constant NO_x–to–NO_2 ratio (reciprocal of NO_2–to–NO_x ratio) of 1.32 and can reach 2 in a hyper-near-field area of a major NO_x source.

5.3 Limitation and room for improvements

The diversity of VOC emissions, the vertical profiles of emissions, and the extent of inter-parcel mixing may impact the modeled results. Perhaps one of the biggest limitations of the current NO_x chemical representation lies in not directly accounting for VOCs, which may affect (a) the "sweet spot" on NO_x curves where two NO_x loss pathways reach their maximum and (b) the O_x -based NO_2–to–NO_x ratios (**Sect. 5.3.1**). Moreover, the influence of representations of emission profile on modeled tNO_2 can be magnified when further considering the TROPOMI NO_2 averaging kernel (**Sect. 5.3.2**). Simulations of point sources like power plants may be more sensitive to these factors compared to simulations of areal sources.

5.3.1 The impact from VOC_R

To investigate the impact of VOCs on NO_x curves, we calculated the VOC reactivity against OH from existing WRF-Chem results based on the following formula: $VOC_R = \sum_{i=1}^{n} k_{OH+VOC_i}[VOC_i]$ and generated separate sets of NO_x curves for 4 respective VOC_R intervals of [0.1, 1), [1, 3), [3, 10), and [10, 50) s^{-1} with a coarse SZA bin spacing of 10°. Curves become much noisier at night and in pristine environments with extremely low NO_x ($\leq$ 0.1 ppb) where WRF-Chem/RADM2 may be less suitable (thereby not shown in **Fig. 10**).

When considering lower SZAs (consistent with TROPOMI overpass time of 1 pm local time), the general non-linear characteristic of these NO_x curves holds as VOC_R increases (**Fig. 10**). Higher VOC_R relative to lower NO_x concentration favors the oxidation of VOCs by OH and the associated minor loss pathway of NO + RO_2 to form alkyl nitrates with a minor branching ratio, over the competing major NO_x loss pathway of NO_2 + OH (**Fig. 1**). With rising VOC_R, the NO_x chemical tendency becomes more positive (P - L, **Supplement Fig. S18a**) and the net loss timescale elongates (e.g., ts_min from 2 to 4 hours in **Fig. 10**). Moreover, NO_x is required to reach a higher level to compete with the reactions involving VOCs, evident by the shift in the trough of the NO_x curves (e.g., NOx_min from 1.4 to 3.7 ppb in **Fig. 10**). To put it in context, the NO_x curves shown in **Fig. 3** represent typical patterns as long as VOC_R remains below 10 s^{-1}.

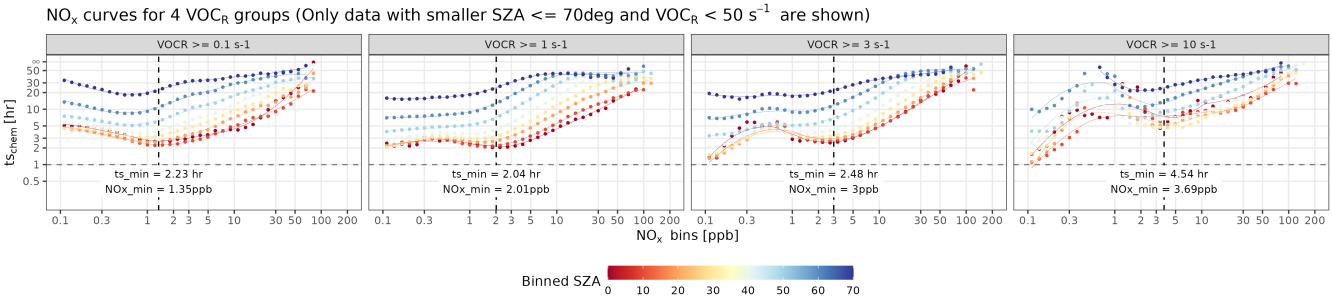

Figure 10. (a) Similar to **Fig. 3b**, but differentiated by 4 intervals of VOC_R and SZA bins smaller than 70 degrees with a spacing of 10 degrees. All panels here utilized model results from the same WRF-Chem simulations described in Sect. 2 and **Appendix A**.

VOC_R may also affect the O_x level and the NO_2–to–NO_x ratio. The prescribed O_x level of 50 ppb (**Sect. 2.2**) overlooks the nonlinear O_x variability related to VOC_R (Murphy et al., 2007; Li et al., 2022). In NO_x-limited scenarios, OH favors the oxidation of VOCs, and local-scale O_x is predominately produced by NO + RO_2 or HO_2, suggesting higher O_x levels with increased NO_x concentrations. The omission of NO + RO_2 or HO_2 in **Eqs. 3** could lead to an underestimation of the NO_2–to–NO_x ratio, which likely explains the modeled tNO_2 being consistently lower than observations over background regions. Conversely, under NO_x-saturated conditions, the consumption of OH by NO_x may limit the VOC oxidation and O_x production, leading to a decline in O_x level as NO_x concentration rises. Consequently, the NO_2–to–NO_x ratio might be overestimated when true O_x levels fall below 50 ppb (particularly under stagnant atmospheric mixing) or underestimated due to the absence of NO + RO_2 reactions. Nevertheless, our pre-determined O_x level of 50 ppb acts as a first-order limit to prevent unrealistic conversion from NO to NO_2 at extremely high NO_x levels when O_3 is being titrated.

To address these limitations, one potential approach is to leverage formaldehyde concentrations retrieved from TROPOMI. Recent studies revisited the use of the formaldehyde-to-NO_2 ratios (i.e., FNR) from satellites as a means of inferring O_3 production rates (Goldberg et al., 2022; Souri et al., 2022). Our WRF-Chem simulations, which were used to parameterize the NO_x chemical tendency, show that modeled formaldehyde generally increases with VOC_R with varying slopes influenced by SZA and NO_x concentrations (**Supplement Fig. S18c**); and O_3 concentrations scale non-linearly with FNRs with O_3 concentration approaching a background value at high FNR > 10 (**Supplement Fig. S18d**). Even though satellite-based FNRs may theoretically help probe O_3 or O_x concentration to better parameterize NO_2–to–NO_x ratios, Souri et al. (2022) stressed that retrieval errors especially from formaldehyde (40 to 90% with $\leq$ 50% over cities) and inherent chemical errors of the predictive power of FNRs may hinder the broad application of space-based FNRs at the current stage. Nonetheless, sensitivity analyses in **Sect. 3** indicate an overall chemical uncertainty in tNO_2 of about 10 to 20% with respect to NO_2 signals, even if perturbed O_x level is much lower than 50 ppb (**Fig. 4**).

5.3.2 Uncertainties in non-chemical processes

Besides simplification of chemical reactions, modeled tNO_2 values can be subject to a few physical processes and parameters, including emission profiles, inter-parcel mixing scales, and dry deposition.

The underlying STILTv2 (Fasoli et al., 2018) accounted for a gradual growth of the mixed layer height over the hyper-near-field area around emissions. Yet, by convolving the STILT footprint with NO_x emissions, we assumed that emissions originate from the surface and are uniformly mixed over the mixed layer without considering the possible uneven distribution of emissions from different vertical levels. In reality, under stable atmospheric conditions, the stack heights or plume heights of emission sources can sometimes extend above the shallow PBL. Our current assumption may thus lead to an overestimation in modeled concentrations, and such biases can in turn affect the estimate of NO_x tendency. More importantly, changes in the vertical profile of emissions can lead to changes in concentration per model level, which affect the tropospheric columns as the typical averaging kernel profile is far from uniform within the PBL. Recall that TROPOMI NO_2 AKs decreases rapidly towards the surface (**Fig. 1**). Hence, placing an emission plume at the surface or an elevated altitude (e.g., 400 m) can cause a discrepancy in modeled column concentrations. In addition, if the wind shear is strong over an intensive point source (likely the Intermountain example in **Fig. 5c**), assumptions in the injection height and vertical profile of emission plumes may affect the modeled plume shape and possibly deviate the estimated near-field wind bias following **Sect. 5.2.1**. Noticeably, Maier et al. (2022) investigated the influence of inaccurate representation of emission profiles on the flask-like modeled concentrations by implementing a time-varying sector-specific emission profile into STILT. Such an impact on column concentrations may be minimized but yet requires future in-depth investigations, particularly over point sources.

Accounting for inter-parcel mixing was an important aspect when developing Lagrangian chemical models. Omitting inter-parcel mixing makes solving for non-linear processes (such as chemical NO_x loss) problematic. On the contrary, Eulerian models suffer from excessive numerical diffusion. Mixing that is too strong smooths the spatial gradient of concentration and can lower the concentration within the fixed model grids, which may cause slight shifts in NO_x regimes. Valin et al. (2011) suggested that a spatial resolution of 4 to 12 km is sufficient to capture the non-linearity in NO_x loss rate. As for Lagrangian

models, efforts can be made to enable the flux exchange between air parcels via deformations (Konopka et al., 2019; McKenna et al., 2002). In addition to the inter-pacel mixing within the mixed layer (ML), several other turbulent mixing processes require future investigation, including (1) horizontal mixing in the free troposphere (FT), (2) vertical mixing between the ML and FT, and (3) mixing between tracked air parcels with the untracked surrounding background. For example, Real et al. (2008) utilized a linear relaxation with exponential decay of the plume concentrations towards the background based on a timescale of 2 days to address the third mixing process. The second mixing process requires future modifications involving the determination of entrainment zones and mixing hyperparameters for such ML-FT exchange.

The original STILT model realized vertical mixing by diluting surface emissions across the ML height (Lin et al., 2003) and we further enabled an exchange in pollutants' concentrations with prescribed mixing length- and time-scales representing typical horizontal mixing rates (**Sect. 2.3**). As final sensitivity tests, we simulated $\mathrm{tNO_2}$ based on a spectrum of mixing hyperparameters for the New Madrid power plant. Uncertainties in the mixing parameters result in minimal uncertainties on the sounding-level modeled $\mathrm{tNO_2}$ values (**Supplement Fig. S19**). For example, differences in modeled $\mathrm{tNO_2}$ between the mixing and non-mixing simulations become larger as mixing becomes faster and for receptors/soundings located on the edge of the plume (i.e., only a small fraction of the trajectories encountered power plant emission in **Supplement Fig. S20**). Uncertainties in the prescribed mixing hyperparameters contribute even less to the modeled values over urban areas (i.e., $<$ 10% for Phoenix cases), where emissions are generally better mixed than at power plants. In addition, such mixing influence can vary with the spatial resolution of the emission inventory used in the simulations.

The dry deposition of $\mathrm{NO_2}$ was not factored into this study, which could lead to an overestimation of modeled $\mathrm{NO_2}$. For future model implementations, it is possible to track loss of $\mathrm{NO_2}$ concentrations due to dry deposition by calculating "dry deposition velocities" (e.g., Wesely, 1988) when air parcels descend close to the surface, e.g., 50 meters above the surface (Wen et al., 2012).

6 Summary

In developing STILT-$\mathrm{NO_x}$, we aim to quantify anthropogenic $\mathrm{NO_x}$ emission signals for power plants and cities using remote sensors using a novel Lagrangian chemical system that preserves the non-linear relationship between $\mathrm{NO_x}$ concentrations and emissions. This development is motivated by the desire to reduce computational costs by replacing the conventional kinetics-based approach to solve for concentration changes with a simplified parameterization relying on as few variables as possible. Such a simplified parameterization can be improved and adopted by other Lagrangian models. This work expands the capability of (X-)STILT in tracing the origins of chemically reactive species to simulate their concentrations at satellite soundings (**Fig. 1**). Although uncertainties exist in modeling atmospheric transport, mixing, and chemical processes, this study covers the key $\mathrm{NO_x}$ chemical mechanisms, various error sources, model validation, and the benefit of using $\mathrm{NO_x}$ to constrain $\mathrm{CO_2}$ emissions and tracer-tracer emission ratios.

To evaluate our modeling system, which consists of the HYSPLIT-STILT core and modules of column weighting, simplified chemistry, and error analyses (**Fig. 1**), we compared modeled tropospheric $\mathrm{NO_2}$ columns using EPA-reported emissions

against observed columns from TROPOMI over three power plants in the US (**Fig. 6**). The largest model-data discrepancies are found for overpasses with substantial wind directional biases. Across three power plants and seasons, the systematic bias informed by the model-data regression slope appears to be small when using EPA emissions. Switching NO_x emissions from prior emissions to EPA (usually with a scaling factor of 2 to 3) greatly improves the model-data agreement, followed by the impact of whether to turn on the NO_x chemistry. Upgrading to a higher-resolution meteorological field minimally alleviated model-data mismatches but should be considered for regions with complex terrain. Subtracting the background NO_2 is necessary, especially over stagnant days and regions with strong non-anthropogenic emission influences. Based on our limited case studies, NO_2 simulations of power plants are usually more challenging compared to urban areas with more of the areal source for several reasons: from atmospheric mixing, spatial heterogeneity and vertical profiles of emissions, to the exposure of ambient ozone-rich air when estimating the NO_2–to–NO_x ratios.

Our comprehensive analyses on modeling tNO_2 further shed light on the estimation of CO_2 emissions at the local scale. Modeling two species in a consistent modeling framework makes the quantification of two key bias terms easier, namely from wind directions and emission locations. For example, we demonstrate the use of model-data NO_2 plumes to obtain a quantitative value of the directional biases associated with the modeled wind (**Sect. 5.2.1**) and biases with the emission distribution in prior inventories like EDGARv6 (**Sect. 4.2**). As growing interest arise from the joint use of GHG and air pollutants, we also investigated the differences between NO_2–to–CO_2 enhancement ratios and the ENO_x–to–ECO_2 emission ratios (**Fig. 9**). Such differences between the two tracer-to-tracer ratios vary across seasons and space, which is again driven by the non-linearity between the emissions and concentrations. For instance, to be consistent with emission ratios at the sources, observation-based enhancement ratios need to be scaled up by 2 to 3 times in the summer months due to faster photochemistry. Soundings with a separation of 60 km from the site center need to be scaled up further by roughly 1.3 times than near-field soundings concerning changes in chemical tendency.

STILT-NO_x in conjunction with the forthcoming local-scale multi-tracer non-linear modeling/inversion system (**Fig. 1**) can be employed to simultaneously constrain emissions from multiple species of both GHGs and key air pollutants along with their respective emission ratios, allowing for improvements in sectoral attributions. Such a framework has the potential to be scaled up to a large number of cities for estimating emissions of NO_x, CO_2, and possibly other tracers from space-based sensors.

Code and data availability. TROPOMI v2.3 of the Level 2 NO_2 data were accessed from https://data-portal.s5p-pal.com/products/no2.html. The official DOI of TROPOMI is http://doi.org/10.5270/S5P-9bnp8q8.The official DOI of the OCO-3 Level 2 B10p4r XCO_2 data is https://doi.org/10.5067/970BCC4DHH24. EDGARv6.1 emissions are accessed from https://data.jrc.ec.europa.eu/dataset/df521e05-6a3b-461c-965a-b703fb62313e and have been preprocessed. The STILT-NOx v1 model is built on previous efforts of the X-STILT model in modeling NO2. The exact version used in the discussion paper is archived on Zenodo with a doi of https://zenodo.org/record/8057850.

Appendix A: WRF-Chem setups

We used meteorological fields from the Global Forecast System ($0.25^\circ \times 0.25^\circ$ GFS-FNL, NCEP, 2015) to generate hourly outputs at a grid spacing of 12 km for five consecutive days in each month of 2020. The first day is regarded as the spin-up time to stabilize the model whose concentration fields are excluded from the following analyses. Anthropogenic emissions of air pollutants and VOCs are adopted respectively from EDGARv6.1 (Crippa et al., 2022) and EDGARv4.3.2 (Huang et al., 2017) with biogenic VOC emissions derived from the Model of Emissions of Gases and Aerosols from Nature (MEGANv2, Guenther et al., 2012). No lightning or soil $\mathrm{NO_x}$ source is included in WRF-Chem simulations. The boundary condition of chemicals relies on the CAM-CHEM model (Buchholz et al., 2019). The most important part is the gas phase photochemistry scheme, which is driven by 2nd generation of the Regional Acid Deposition Model (RADM2, Stockwell et al., 1990) with Dry and wet depositions included. RADM2 is well-suited under polluted environments but may miss several key aromatic components for pristine environments dominated by BVOCs (Stockwell et al., 1997).

Appendix B: Technical notes on regional wind assessment

We assess the wind uncertainty associated with two meteorological fields that drive (X-)STILT, namely the 3 km HRRR and 0.25° GFS. The first approach targets regional wind error statistics by comparing modeled wind fields (both HRRR and GFS) against true wind observations at radiosonde stations. The wind error statistic is further used to translate wind errors to uncertainties in $\mathrm{tNO_2}$. Radiosonde balloons are normally launched at 00 or 12 UTC. U- and V-component wind observations for only levels below 2 km over the 24 hours ahead of the TROPOMI overpass time are selected. We then estimate random uncertainties of u-/v-component wind speed (i.e., RMSE in $\mathrm{m\ s^{-1}}$) and normalize RMSEs over mean wind speed to yield fractional uncertainties (%) for every overpass (**Supplement Fig. S3**). For example, fractional wind uncertainties over Missouri (around 20 to 40%) are generally smaller than uncertainties over mountainous lands in Utah (> 40%), which relates to the model's capabilities in resolving topography and topographic flows. In addition, HRRR-based winds at radiosonde stations appear to be more erroneous compared to GFS-based winds High-resolution models provide better descriptions of the surface land cover type, terrain height, and surface roughness, which may improve the spatial variability of PBLH (Lin et al., 2017) and wind vectors. Without true wind measurements, it remains unclear whether higher-resolution models can capture more accurate fine-scale meteorology. Nevertheless, the radiosonde analysis provides an overall picture of the regional wind uncertainty.

To propagate wind error statistics to transport uncertainty in concentrations, a wind error component is added to the mean wind component and the turbulence component when generating backward trajectories. Transport uncertainties in $\mathrm{tNO_2}$ are defined as the differences in variations of parcel-specific $\mathrm{NO_2}$ mixing ratio with the proper vertical weighting of AK and PWF between the perturbed and the initial set of the trajectories.

Appendix C: Technical notes on near-field wind bias quantification

As introduced in **Sect. 5.2.1**, a $\mathrm{NO_2}$ plume from either model or retrieval is rotated clockwise (α from -180 to -5° with a spacing of 5°) or counter-clockwise (from 5 to 180°) around the emission source and then resampled onto the original TROPOMI pixels (**Supplement Fig. S16**). We then multiply gridded $\mathrm{tNO_2}$ from the initial plume with gridded $\mathrm{tNO_2}$ from each rotated plume under each rotating angle, α. The cross-product of two $\mathrm{tNO_2}$ plumes [XP, $\mathrm{ppb^2}$] measures how two $\mathrm{tNO_2}$ plumes are similar in terms of their spatial structures, as one is rotated around its source location for 360° (**Eq. C1**):

$$\mathrm{XP}(\mathrm{x_s}, \mathrm{y_s}, \alpha) = \mathrm{tNO_2^{initial}}(\mathrm{x_s}, \mathrm{y_s})\ \mathrm{tNO_2^{rotated}}(\mathrm{x_s}, \mathrm{y_s}, \alpha) \tag{C1}$$

We next calculate the square-root-mean of these sounding-specific cross-products per rotating angle (**Fig. 8a**). The peak of the Gaussian shape suggests when two plume signals reach the maximum correlation, while the wing suggests when one plume signal starts to decouple with another plume signal. The plume that undergoes rotation ($\mathrm{tNO_2^{rotated}}$) can either be a modeled plume with different model configurations (e.g., GFS or HRRR; with or without chemistry) or an observed plume. For example, root-mean-products (RMP) based on the simulations without chemistry displays a high bias compared to RMP using observed $\mathrm{tNO_2}$, which implies that the entire modeled scene including the background signal is biased high when $\mathrm{NO_x}$ lifetime is not included. When an observed plume was rotated to match its original self, their $\mathrm{tNO_2}$ product can serve as a baseline. Normalizing the cross-products offers a diagnostic (**Fig. 8**), analogous to the concept of the "cross-correlation coefficient".

Author contributions. DW, POW, and JLiu designed the modeling experiments and contributed to the interpretation of results. DW realized the STILT-$\mathrm{NO_x}$ model code and conducted WRF-Chem simulations. Specific insights from individuals: POW, JLLaughner, and PIP —$\mathrm{NO_x}$ chemical parameterization; JLLaughner —WRF-Chem model setup; JCLin —Lagrangian inter-parcel mixing. All authors contribute to the manuscript writing and editing.

Competing interests. The authors declare no conflict of interest.

Acknowledgements. The analysis is supported by the National Aeronautics and Space Administration with grant number of 80NSSC21K1064. A portion of the research was carried out at the Jet Propulsion Laboratory, California Institute of Technology, under a contract with the National Aeronautics and Space Administration (80NM0018D0004). The computations presented here were conducted in the Resnick High-Performance Computing Center, a facility supported by the Resnick Sustainability Institute at the California Institute of Technology. We acknowledge the use of the WRF-Chem preprocessor tool of mozbc provided by the Atmospheric Chemistry Observations and Modeling Lab (ACOM) of NCAR. The authors acknowledge the NOAA Air Resources Laboratory (ARL) for the provision of the GFS and HRRR meteorological files used in this publication, which were downloaded from the READY website (http://www.ready.noaa.gov, last access: 1 May 2018). The first author extends appreciation toward Kazuyuki Miyazaki (JPL) for discussions on $\mathrm{NO_x}$ modeling and Rob Nelson and Annmarie Eldering (JPL) for the OCO-3 data.

References

Beirle, S., Boersma, K. F., Platt, U., Lawrence, M. G., and Wagner, T.: Megacity emissions and lifetimes of nitrogen oxides probed from space, Science, 333, 1737–1739, 2011.

Beirle, S., Borger, C., Dörner, S., Li, A., Hu, Z., Liu, F., Wang, Y., and Wagner, T.: Pinpointing nitrogen oxide emissions from space, Science advances, 5, eaax9800, 2019.

Beirle, S., Borger, C., Dörner, S., Eskes, H., Kumar, V., de Laat, A., and Wagner, T.: Catalog of NO_x emissions from point sources as derived from the divergence of the NO_2 flux for TROPOMI, Earth System Science Data, 13, 2995–3012, 2021.

Brunner, D.: Atmospheric chemistry in lagrangian models–overview, Lagrangian Modeling of the Atmosphere, edited by: Lin, JC, Brunner, D., Gerbig, C., Stohl, A., Luchar, A., and Webley, P., Geophysical Monograph Series, 200, 2012.

Buchholz, R., Emmons, L., and Tilmes, S.: The CESM2 Development Team: CESM2.1/CAM-chem Instantaneous Output for Boundary Conditions, UCAR/NCAR–Atmospheric Chemistry Observations and Modeling Laboratory, Subset used Jan 2020–Dec 2020, 2019.

Cifuentes, L., Borja-Aburto, V. H., Gouveia, N., Thurston, G., and Davis, D. L.: Hidden health benefits of greenhouse gas mitigation, 2001.

Clapp, L. J. and Jenkin, M. E.: Analysis of the relationship between ambient levels of O_3, NO_2 and NO as a function of NO_x in the UK, Atmospheric Environment, 35, 6391–6405, 2001.

Collins, W., Stevenson, D. S., Johnson, C., and Derwent, R.: Tropospheric ozone in a global-scale three-dimensional Lagrangian model and its response to NO_x emission controls, Journal of atmospheric chemistry, 26, 223–274, 1997.

Crippa, M., Guizzardi, D., Muntean, M., Schaaf, E., MONFORTI-FERRARIO, F., BANJA, M., PAGANI, F., SOLAZZO, E., et al.: EDGAR v6.1 global air pollutant emissions, http://data.europa.eu/89h/df521e05-6a3b-461c-965a-b703fb62313e, type: data set, 2022.

Demetillo, M. A. G., Navarro, A., Knowles, K. K., Fields, K. P., Geddes, J. A., Nowlan, C. R., Janz, S. J., Judd, L. M., Al-Saadi, J., Sun, K., et al.: Observing nitrogen dioxide air pollution inequality using high-spatial-resolution remote sensing measurements in houston, Texas, Environmental Science & Technology, 54, 9882–9895, 2020.

Duncan, B. N., Lamsal, L. N., Thompson, A. M., Yoshida, Y., Lu, Z., Streets, D. G., Hurwitz, M. M., and Pickering, K. E.: A space-based, high-resolution view of notable changes in urban NO_x pollution around the world (2005–2014), Journal of Geophysical Research: Atmospheres, 121, 976–996, 2016.

Fasoli, B., Lin, J. C., Bowling, D. R., Mitchell, L., and Mendoza, D.: Simulating atmospheric tracer concentrations for spatially distributed receptors: updates to the Stochastic Time-Inverted Lagrangian Transport model's R interface (STILT-R version 2), Geoscientific Model Development, 11, 2813–2824, 2018.

Fujita, E. M., Campbell, D. E., Stockwell, W. R., Saunders, E., Fitzgerald, R., and Perea, R.: Projected ozone trends and changes in the ozone-precursor relationship in the South Coast Air Basin in response to varying reductions of precursor emissions, Journal of the Air & Waste Management Association, 66, 201–214, 2016.

Goldberg, D. L., Lu, Z., Streets, D. G., de Foy, B., Griffin, D., McLinden, C. A., Lamsal, L. N., Krotkov, N. A., and Eskes, H.: Enhanced capabilities of TROPOMI NO_2: estimating NO_x from North American cities and power plants, Environmental science & technology, 53, 12 594–12 601, 2019.

Goldberg, D. L., Harkey, M., de Foy, B., Judd, L., Johnson, J., Yarwood, G., and Holloway, T.: Evaluating NO_x emissions and their effect on O_3 production in Texas using TROPOMI NO_2 and HCHO, Atmospheric Chemistry and Physics, 22, 10 875–10 900, 2022.

Grell, G. A., Peckham, S. E., Schmitz, R., McKeen, S. A., Frost, G., Skamarock, W. C., and Eder, B.: Fully coupled "online" chemistry within the WRF model, Atmospheric Environment, 39, 6957–6975, 2005.

Guenther, A., Jiang, X., Heald, C. L., Sakulyanontvittaya, T., Duhl, T. a., Emmons, L., and Wang, X.: The Model of Emissions of Gases and Aerosols from Nature version 2.1 (MEGAN2.1): an extended and updated framework for modeling biogenic emissions, Geoscientific Model Development, 5, 1471–1492, 2012.

Hakkarainen, J., Ialongo, I., Oda, T., Szelag, M. E., O'Dell, C. W., Eldering, A., and Crisp, D.: Building a bridge: Characterizing major anthropogenic point sources in the South African Highveld region using OCO-3 carbon dioxide Snapshot Area Maps and Sentinel-5P/TROPOMI nitrogen dioxide columns, Environmental Research Letters, 2023.

He, T.-L., Jones, D., Miyazaki, K., Bowman, K. W., Jiang, Z., Chen, X., Li, R., Zhang, Y., and Li, K.: Inverse modelling of Chinese NOx emissions using deep learning: integrating in situ observations with a satellite-based chemical reanalysis, Atmospheric Chemistry and Physics, 22, 14 059–14 074, 2022.

Huang, G., Brook, R., Crippa, M., Janssens-Maenhout, G., Schieberle, C., Dore, C., Guizzardi, D., Muntean, M., Schaaf, E., and Friedrich, R.: Speciation of anthropogenic emissions of non-methane volatile organic compounds: a global gridded data set for 1970–2012, Atmospheric Chemistry and Physics, 17, 7683–7701, 2017.

Huang, Y. and Seinfeld, J. H.: A Neural Network-Assisted Euler Integrator for Stiff Kinetics in Atmospheric Chemistry, Environmental Science & Technology, 56, 4676–4685, 2022.

Hurtt, G. C., Andrews, A., Bowman, K., Brown, M. E., Chatterjee, A., Escobar, V., Fatoyinbo, L., Griffith, P., Guy, M., Healey, S. P., et al.: The NASA Carbon Monitoring System Phase 2 synthesis: scope, findings, gaps and recommended next steps, Environmental Research Letters, 17, 063 010, 2022.

Jenkin, M. E.: Analysis of sources and partitioning of oxidant in the UK—Part 2: contributions of nitrogen dioxide emissions and background ozone at a kerbside location in London, Atmospheric Environment, 38, 5131–5138, 2004.

Jiang, Z., McDonald, B. C., Worden, H., Worden, J. R., Miyazaki, K., Qu, Z., Henze, D. K., Jones, D. B., Arellano, A. F., Fischer, E. V., et al.: Unexpected slowdown of US pollutant emission reduction in the past decade, Proceedings of the National Academy of Sciences, 115, 5099–5104, 2018.

Jin, X., Fiore, A. M., Murray, L. T., Valin, L. C., Lamsal, L. N., Duncan, B., Folkert Boersma, K., De Smedt, I., Abad, G. G., Chance, K., et al.: Evaluating a space-based indicator of surface ozone-NO_x-VOC sensitivity over midlatitude source regions and application to decadal trends, Journal of Geophysical Research: Atmospheres, 122, 10–439, 2017.

Jin, X., Fiore, A., Boersma, K. F., Smedt, I. D., and Valin, L.: Inferring changes in summertime surface Ozone–NO_x–VOC chemistry over US urban areas from two decades of satellite and ground-based observations, Environmental science & technology, 54, 6518–6529, 2020.

Kaminski, T., Scholze, M., Rayner, P., Houweling, S., Voßbeck, M., Silver, J., Lama, S., Buchwitz, M., Reuter, M., Knorr, W., et al.: Assessing the Impact of Atmospheric CO2 and NO2 Measurements From Space on Estimating City-Scale Fossil Fuel CO2 Emissions in a Data Assimilation System, Frontiers in Remote Sensing, 3, 2022.

Keller, C. A. and Evans, M. J.: Application of random forest regression to the calculation of gas-phase chemistry within the GEOS-Chem chemistry model v10, Geoscientific Model Development, 12, 1209–1225, 2019.

Konopka, P., Tao, M., Ploeger, F., Diallo, M., and Riese, M.: Tropospheric mixing and parametrization of unresolved convective updrafts as implemented in the Chemical Lagrangian Model of the Stratosphere (CLaMS v2.0), Geoscientific model development, 12, 2441–2462, 2019.

Kuhlmann, G., Broquet, G., Marshall, J., Clément, V., Löscher, A., Meijer, Y., and Brunner, D.: Detectability of CO2 emission plumes of cities and power plants with the Copernicus Anthropogenic CO2 Monitoring (CO2M) mission, Atmospheric Measurement Techniques, 12, 6695–6719, 2019.

Kuhlmann, G., Henne, S., Meijer, Y., and Brunner, D.: Quantifying CO_2 emissions of power plants with CO_2 and NO_2 imaging satellites, Frontiers in Remote Sensing, 2, 14, 2021.

Lama, S., Houweling, S., Boersma, K. F., Aben, I., Van Der Gon, H. A. D., and Krol, M. C.: Estimation of OH in urban plume using TROPOMI inferred NO_2/CO, Authorea Preprints, 2022.

Lamsal, L., Martin, R., Padmanabhan, A., Van Donkelaar, A., Zhang, Q., Sioris, C., Chance, K., Kurosu, T., and Newchurch, M.: Application of satellite observations for timely updates to global anthropogenic NO_x emission inventories, Geophysical Research Letters, 38, 2011.

Laughner, J. L. and Cohen, R. C.: Direct observation of changing NO_x lifetime in North American cities, Science, 366, 723–727, 2019.

Lee, H.-J., Kim, S.-W., Brioude, J., Cooper, O., Frost, G., Kim, C.-H., Park, R., Trainer, M., and Woo, J.-H.: Transport of NO_x in East Asia identified by satellite and in situ measurements and Lagrangian particle dispersion model simulations, Journal of Geophysical Research: Atmospheres, 119, 2574–2596, 2014.

Lee, J. D., Drysdale, W. S., Finch, D. P., Wilde, S. E., and Palmer, P. I.: UK surface NO< sub> 2</sub> levels dropped by 42% during the COVID-19 lockdown: impact on surface O< sub> 3</sub>, Atmospheric Chemistry and Physics, 20, 15 743–15 759, 2020.

Li, C., Zhu, Q., Jin, X., and Cohen, R. C.: Elucidating Contributions of Anthropogenic Volatile Organic Compounds and Particulate Matter to Ozone Trends over China, Environmental Science & Technology, 56, 12 906–12 916, 2022.

Lin, J. C. and Gerbig, C.: Accounting for the effect of transport errors on tracer inversions, Geophysical Research Letters, 32, 2005.

Lin, J. C., Gerbig, C., Wofsy, S., Andrews, A., Daube, B., Davis, K., and Grainger, C.: A near-field tool for simulating the upstream influence of atmospheric observations: The Stochastic Time-Inverted Lagrangian Transport (STILT) model, Journal of Geophysical Research: Atmospheres, 108, 2003.

Lin, J. C., Brunner, D., Gerbig, C., Stohl, A., Luhar, A., and Webley, P.: Lagrangian modeling of the atmosphere, John Wiley & Sons, 2013.

Lin, J. C., Mallia, D. V., Wu, D., and Stephens, B. B.: How can mountaintop CO_2 observations be used to constrain regional carbon fluxes?, Atmospheric Chemistry and Physics, 17, 5561–5581, 2017.

Lin, J. C., Mitchell, L., Crosman, E., Mendoza, D. L., Buchert, M., Bares, R., Fasoli, B., Bowling, D. R., Pataki, D., Catharine, D., et al.: CO_2 and carbon emissions from cities: Linkages to air quality, socioeconomic activity, and stakeholders in the Salt Lake City urban area, Bulletin of the American Meteorological Society, 99, 2325–2339, 2018.

Liu, F., Tao, Z., Beirle, S., Joiner, J., Yoshida, Y., Smith, S. J., Knowland, K. E., and Wagner, T.: A new method for inferring city emissions and lifetimes of nitrogen oxides from high-resolution nitrogen dioxide observations: a model study, Atmospheric Chemistry and Physics, 22, 1333–1349, 2022.

Liu, X., Mizzi, A. P., Anderson, J. L., Fung, I. Y., and Cohen, R. C.: Assimilation of satellite NO_2 observations at high spatial resolution using OSSEs, Atmospheric Chemistry and Physics, 17, 7067–7081, 2017.

Loughner, C. P., Fasoli, B., Stein, A. F., and Lin, J. C.: Incorporating features from the stochastic time-inverted lagrangian transport (STILT) model into the Hybrid Single-Particle Lagrangian Integrated Trajectory (HYSPLIT) model: a unified dispersion model for time-forward and time-reversed applications, Journal of Applied Meteorology and Climatology, 60, 799–810, 2021.

MacDonald, C. G., Laughner, J. L., Hedelius, J. K., Nassar, R., Mastrogiacomo, J.-P., and Wunch, D.: Estimating Enhancement Ratios of Nitrogen Dioxide, Carbon Monoxide, and Carbon Dioxide using Satellite Observations, Atmospheric Chemistry and Physics Discussions, pp. 1–30, 2022.

Maier, F., Gerbig, C., Levin, I., Super, I., Marshall, J., and Hammer, S.: Effects of point source emission heights in WRF–STILT: a step towards exploiting nocturnal observations in models, Geoscientific Model Development, 15, 5391–5406, 2022.

McKenna, D. S., Konopka, P., Grooß, J.-U., Günther, G., Müller, R., Spang, R., Offermann, D., and Orsolini, Y.: A new Chemical Lagrangian Model of the Stratosphere (CLaMS) 1. Formulation of advection and mixing, Journal of Geophysical Research: Atmospheres, 107, ACH–15, 2002.

Miyazaki, K. and Bowman, K.: Predictability of fossil fuel CO2 from air quality emissions, Nature communications, 2023.

Miyazaki, K., Bowman, K., Sekiya, T., Eskes, H., Boersma, F., Worden, H., Livesey, N., Payne, V. H., Sudo, K., Kanaya, Y., et al.: Updated tropospheric chemistry reanalysis and emission estimates, TCR-2, for 2005–2018, Earth System Science Data, 12, 2223–2259, 2020.

Murphy, J. G., Day, D. A., Cleary, P. A., Wooldridge, P. J., Millet, D. B., Goldstein, A. H., and Cohen, R. C.: The weekend effect within and downwind of Sacramento–Part 1: Observations of ozone, nitrogen oxides, and VOC reactivity, Atmospheric Chemistry and Physics, 7, 5327–5339, 2007.

Myhre, G., Shindell, D., and Pongratz, J.: Anthropogenic and natural radiative forcing, 2014.

Nassar, R., Moeini, O., Mastrogiacomo, J.-P., O'Dell, C. W., Nelson, R. R., Kiel, M., Chatterjee, A., Eldering, A., and Crisp, D.: Tracking CO_2 emission reductions from space: A case study at Europe's largest fossil fuel power plant, Frontiers in Remote Sensing, 3, 98, 2022.

NCEP: NCEP GDAS/FNL 0.25 Degree Global Tropospheric Analyses and Forecast Grids, https://doi.org/10.5065/D65Q4T4Z, place: Boulder CO, 2015.

Parker, H., Hasheminassab, S., Crounse, J., Roehl, C., and Wennberg, P.: Impacts of traffic reductions associated with COVID-19 on southern California air quality, Geophysical research letters, 47, e2020GL090 164, 2020.

Pugh, T., Cain, M., Methven, J., Wild, O., Arnold, S., Real, E., Law, K. S., Emmerson, K., Owen, S., Pyle, J., et al.: A Lagrangian model of air-mass photochemistry and mixing using a trajectory ensemble: the Cambridge Tropospheric Trajectory model of Chemistry And Transport (CiTTyCAT) version 4.2, Geoscientific Model Development, 5, 193–221, 2012.

Qu, Z., Henze, D. K., Worden, H. M., Jiang, Z., Gaubert, B., Theys, N., and Wang, W.: Sector-based top-down estimates of NO_x, SO_2, and CO emissions in East Asia, Geophysical Research Letters, 49, e2021GL096 009, 2022.

Real, E., Law, K. S., Schlager, H., Roiger, A., Huntrieser, H., Methven, J., Cain, M., Holloway, J., Neuman, J., Ryerson, T., et al.: Lagrangian analysis of low altitude anthropogenic plume processing across the North Atlantic, Atmospheric Chemistry and Physics, 8, 7737–7754, 2008.

Reuter, M., Buchwitz, M., Hilboll, A., Richter, A., Schneising, O., Hilker, M., Heymann, J., Bovensmann, H., and Burrows, J.: Decreasing emissions of NOx relative to CO2 in East Asia inferred from satellite observations, Nature Geoscience, 7, 792–795, 2014.

Rohrer, F. and Berresheim, H.: Strong correlation between levels of tropospheric hydroxyl radicals and solar ultraviolet radiation, Nature, 442, 184–187, 2006.

Rolph, G., Stein, A., and Stunder, B.: Real-time environmental applications and display system: READY, Environmental Modelling & Software, 95, 210–228, 2017.

Roten, D., Lin, J. C., Kunik, L., Mallia, D., Wu, D., Oda, T., and Kort, E. A.: The information content of dense carbon dioxide measurements from space: a high-resolution inversion approach with synthetic data from the OCO-3 instrument, Atmospheric Chemistry and Physics Discussions, pp. 1–43, 2022.

Shah, V., Jacob, D. J., Dang, R., Lamsal, L. N., Steenrod, S., et al.: Nitrogen oxides in the free troposphere: Implications for tropospheric oxidants and the interpretation of satellite NO_2 measurements, UMBC Faculty Collection, 2022.

Silva, S. J. and Arellano, A.: Characterizing regional-scale combustion using satellite retrievals of CO, NO_2 and CO_2, Remote Sensing, 9, 744, 2017.

Souri, A. H., Johnson, M. S., Wolfe, G. M., Crawford, J. H., Fried, A., Wisthaler, A., Brune, W. H., Blake, D. R., Weinheimer, A. J., Verhoelst, T., et al.: Characterization of Errors in Satellite-Based HCHO/NO_2 Tropospheric Column Ratios with Respect to Chemistry, Column to PBL Translation, Spatial Representation, and Retrieval Uncertainties, Atmospheric Chemistry and Physics Discussions, pp. 1–43, 2022.

Stein, A. F., Lamb, D., and Draxler, R. R.: Incorporation of detailed chemistry into a three-dimensional Lagrangian–Eulerian hybrid model: Application to regional tropospheric ozone, Atmospheric Environment, 34, 4361–4372, 2000.

Stein, A. F., Draxler, R. R., Rolph, G. D., Stunder, B. J., Cohen, M., and Ngan, F.: NOAA's HYSPLIT atmospheric transport and dispersion modeling system, Bulletin of the American Meteorological Society, 96, 2059–2077, 2015.

Stockwell, W. R., Middleton, P., Chang, J. S., and Tang, X.: The second generation regional acid deposition model chemical mechanism for regional air quality modeling, Journal of Geophysical Research: Atmospheres, 95, 16 343–16 367, 1990.

Stockwell, W. R., Kirchner, F., Kuhn, M., and Seefeld, S.: A new mechanism for regional atmospheric chemistry modeling, Journal of Geophysical Research: Atmospheres, 102, 25 847–25 879, 1997.

Strong, J., Whyatt, J., Hewitt, C., and Derwent, R.: Development and application of a Lagrangian model to determine the origins of ozone episodes in the UK, Atmospheric Environment, 44, 631–641, 2010.

Tang, W., Arellano, A. F., Gaubert, B., Miyazaki, K., and Worden, H. M.: Satellite data reveal a common combustion emission pathway for major cities in China, Atmospheric Chemistry and Physics, 19, 4269–4288, 2019.

United States Environmental Protection Agency, E.: Power Sector Emissions Data., Washington,DC:OfficeofAtmosphericProtection, CleanAirMarketsDivision.AvailablefromEPA'sAirMarketsProgramDatawebsite:https://campd.epa.gov, 2022.

Valin, L., Russell, A., Hudman, R., and Cohen, R.: Effects of model resolution on the interpretation of satellite NO_2 observations, Atmospheric Chemistry and Physics, 11, 11 647–11 655, 2011.

Valin, L., Russell, A., and Cohen, R.: Variations of OH radical in an urban plume inferred from NO_2 column measurements, Geophysical Research Letters, 40, 1856–1860, 2013.

Van Geffen, J., Eskes, H., Compernolle, S., Pinardi, G., Verhoelst, T., Lambert, J.-C., Sneep, M., Ter Linden, M., Ludewig, A., Boersma, K. F., et al.: Sentinel-5P TROPOMI NO_2 retrieval: impact of version v2. 2 improvements and comparisons with OMI and ground-based data, Atmospheric Measurement Techniques, 15, 2037–2060, 2022.

Watts, N., Amann, M., Arnell, N., Ayeb-Karlsson, S., Beagley, J., Belesova, K., Boykoff, M., Byass, P., Cai, W., Campbell-Lendrum, D., et al.: The 2020 report of the Lancet Countdown on health and climate change: responding to converging crises, The Lancet, 397, 129–170, 2021.

Wen, D., Lin, J. C., Millet, D. B., Stein, A. F., and Draxler, R. R.: A backward-time stochastic Lagrangian air quality model, Atmospheric Environment, 54, 373–386, 2012.

Wesely, M. L.: Improved parametarizations for surface resistance to gaseous dry deposition in regional-scale numerical models, EPA, pp. 600–3, 1988.

West, J. J., Smith, S. J., Silva, R. A., Naik, V., Zhang, Y., Adelman, Z., Fry, M. M., Anenberg, S., Horowitz, L. W., and Lamarque, J.-F.: Co-benefits of mitigating global greenhouse gas emissions for future air quality and human health, Nature climate change, 3, 885–889, 2013.

Wohltmann, I. and Rex, M.: The Lagrangian chemistry and transport model ATLAS: Validation of advective transport and mixing, Geoscientific Model Development, 2, 153–173, 2009.

Wu, D., Lin, J. C., Fasoli, B., Oda, T., Ye, X., Lauvaux, T., Yang, E. G., and Kort, E. A.: A Lagrangian approach towards extracting signals of urban CO_2 emissions from satellite observations of atmospheric column CO_2 (XCO_2): X-Stochastic Time-Inverted Lagrangian Transport model ("X-STILT v1"), Geoscientific Model Development, 11, 4843–4871, 2018.

Wu, D., Liu, J., Wennberg, P. O., Palmer, P. I., Nelson, R. R., Kiel, M., and Eldering, A.: Towards sector-based attribution using intra-city variations in satellite-based emission ratios between CO_2 and CO, Atmospheric Chemistry and Physics, 22, 14 547–14 570, 2022.

Wunch, D., Wennberg, P., Toon, G., Keppel-Aleks, G., and Yavin, Y.: Emissions of greenhouse gases from a North American megacity, Geophysical research letters, 36, 2009.

Yang, E. G., Kort, E. A., Ott, L. E., Oda, T., and Lin, J. C.: Using Space-Based CO2 and NO2 Observations to Estimate Urban CO2 Emissions, Journal of Geophysical Research: Atmospheres, p. e2022JD037736, 2023.

Zhang, Q., Boersma, K. F., Zhao, B., Eskes, H., Chen, C., Zheng, H., and Zhang, X.: Quantifying daily NO_x and CO_2 emissions from Wuhan using satellite observations from TROPOMI and OCO-2, Atmospheric Chemistry and Physics, 23, 551–563, 2023.

Zheng, B., Geng, G., Ciais, P., Davis, S. J., Martin, R. V., Meng, J., Wu, N., Chevallier, F., Broquet, G., Boersma, F., et al.: Satellite-based estimates of decline and rebound in China's CO_2 emissions during COVID-19 pandemic, Science Advances, 6, eabd4998, 2020.

Zhu, Q., Laughner, J. L., and Cohen, R. C.: Combining Machine Learning and Satellite Observations to Predict Spatial and Temporal Variation of near Surface OH in North American Cities, Environmental Science & Technology, 56, 7362–7371, https://doi.org/10.1021/acs.est.1c05636, pMID: 35302754, 2022.