







## An Observing System Simulation Experiment Analysis of How Well Geostationary Satellite Trace-Gas Observations Constrain NO<sub>x</sub> Emissions in the US

Chia-Hua Hsu<sup>1,2,3</sup> , Daven K. Henze<sup>1</sup> , Arthur P. Mizzi<sup>1,3,4,5</sup>, Gonzalo González Abad<sup>6</sup> , Jian He<sup>2,3</sup>, Colin Harkins<sup>2,3</sup> , Aaron R. Naeger<sup>7</sup>, Congmeng Lyu<sup>2,3</sup>, Xiong Liu<sup>6</sup> , Christopher Chan Miller<sup>6,8</sup>, R. Bradley Pierce<sup>9</sup>, Matthew S. Johnson<sup>10</sup> , and Brian C. McDonald<sup>3</sup>

<sup>1</sup>Department of Mechanical Engineering, University of Colorado Boulder, Boulder, CO, USA, <sup>2</sup>Cooperative Institute for Research in Environmental Sciences, University of Colorado Boulder, Boulder, CO, USA, <sup>3</sup>NOAA Chemical Sciences Laboratory, Boulder, CO, USA, <sup>4</sup>NASA Earth Exchange, NASA Ames Research Center, Moffett Field, CA, USA, <sup>5</sup>Bay Area Environmental Research Institute, Moffett Field, CA, USA, <sup>6</sup>Center for Astrophysics, Harvard & Smithsonian, Cambridge, MA, USA, <sup>7</sup>Earth System Science Center, University of Alabama in Huntsville, Huntsville, AL, USA, <sup>8</sup>Harvard John A. Paulson School of Engineering and Applied Sciences, Harvard University, Cambridge, MA, USA, <sup>9</sup>Space Science and Engineering Center, University of Wisconsin-Madison, Madison, WI, USA, <sup>10</sup>Earth Science Division, NASA Ames Research Center, Moffett Field, CA, USA

**Key Points:**

- True NO<sub>x</sub> emissions can be recovered using half as many simulation days when assimilating synthetic Tropospheric Emissions: Monitoring of Pollution (TEMPO) observations rather than TROPospheric Monitoring Instrument
- Assimilating synthetic TEMPO observations improve emissions inversion accuracy by 13%–42% across different regions of US
- The best estimates of NO<sub>x</sub> emissions are achieved by using short data assimilation window (e.g., 30 min) and updating concentrations/emissions jointly

**Supporting Information:**

Supporting Information may be found in the online version of this article.

**Correspondence to:**

B. C. McDonald,  
[Brian.Mcdonald@noaa.gov](mailto:Brian.Mcdonald@noaa.gov)

**Citation:**

Hsu, C.-H., Henze, D. K., Mizzi, A. P., González Abad, G., He, J., Harkins, C., et al. (2024). An observing system simulation experiment analysis of how well geostationary satellite trace-gas observations constrain NO<sub>x</sub> emissions in the US. *Journal of Geophysical Research: Atmospheres*, 129, e2023JD039323. <https://doi.org/10.1029/2023JD039323>

Received 31 MAY 2023

Accepted 26 NOV 2023

**Author Contributions:**

**Conceptualization:** Chia-Hua Hsu, Daven K. Henze, Arthur P. Mizzi, Brian C. McDonald

**Formal analysis:** Chia-Hua Hsu

**Investigation:** Chia-Hua Hsu, Daven K. Henze, Arthur P. Mizzi, Congmeng Lyu, Brian C. McDonald

**Methodology:** Chia-Hua Hsu, Daven K. Henze, Arthur P. Mizzi

**Resources:** Gonzalo González Abad, Jian He, Colin Harkins, Aaron R. Naeger, Congmeng Lyu, Xiong Liu, Christopher Chan Miller, R. Bradley Pierce, Matthew S. Johnson, Brian C. McDonald

**Abstract** We investigate the benefit of assimilating high spatial-temporal resolution nitrogen dioxide (NO<sub>2</sub>) measurements from a geostationary (GEO) instrument such as Tropospheric Emissions: Monitoring of Pollution (TEMPO) versus a low-earth orbit (LEO) platform like TROPospheric Monitoring Instrument (TROPOMI) on the inverse modeling of nitrogen oxides (NO<sub>x</sub>) emissions. We generated synthetic TEMPO and TROPOMI NO<sub>2</sub> measurements based on emissions from the COVID-19 lockdown period. Starting with emissions levels prior to the lockdown, we use the Weather Research and Forecasting Model coupled with Chemistry/Data Assimilation Research Testbed (WRF-Chem/DART) to assimilate these pseudo-observations in Observing System Simulation Experiments to adjust NO<sub>x</sub> emissions and quantify how well the assimilation of TEMPO versus TROPOMI measurements recovers the lockdown-induced emissions changes. We find that NO<sub>x</sub> emission biases can be ameliorated using half as many simulation days when assimilating GEO observations, and the estimated NO<sub>x</sub> emissions in 23 out of 29 major urban regions in the US are more accurate. The root mean square error and coefficient of determination of posterior NO<sub>x</sub> emissions are reduced by 12.5%–41.5% and 1.5%–17.1%, respectively, across different regions. We conduct sensitivity experiments that use different data assimilation (DA) configurations to assimilate synthetic GEO observations. Results demonstrate that the temporal width of the DA window introduces –10% to –20% biases in the emissions inversion and constraining both NO<sub>x</sub> concentrations and emissions simultaneously yields the most accurate NO<sub>x</sub> emissions estimates. Our work serves as a valuable reference on how to appropriately assimilate GEO observations for constraining NO<sub>x</sub> emissions in future studies.

**Plain Language Summary** Nitrogen oxides (NO<sub>x</sub>) are major air pollutants and precursors to tropospheric ozone and secondary inorganic aerosols. The diverse natural and anthropogenic sources of NO<sub>x</sub> pose a challenge for NO<sub>x</sub> emissions estimates. Inverse modeling techniques which use observations to infer emissions can be applied to improve our understanding of anthropogenic NO<sub>x</sub> emissions. This study aims to compare the ability of the new geostationary (GEO) instrument Tropospheric Emissions: Monitoring of Pollution (TEMPO) and the existing low-earth orbit instrument TROPospheric Monitoring Instrument (TROPOMI) to constrain NO<sub>x</sub> emissions. Synthetic TEMPO and TROPOMI NO<sub>2</sub> measurements are generated and assimilated to constrain NO<sub>x</sub> emissions in an idealized experiment in which the “true” emissions are known. The results show the true NO<sub>x</sub> emissions can be retrieved using half as many simulation days when assimilating GEO NO<sub>2</sub> observations. Moreover, the experiment that assimilates GEO NO<sub>2</sub> observations improves the accuracy of estimated NO<sub>x</sub> emissions by 12.5%–41.5% and 1.5%–17.1% in terms of root mean square error and coefficient of determination, respectively, across different air quality regions. The NO<sub>x</sub> emissions in most urban regions are better constrained when assimilating GEO NO<sub>2</sub> data. We also propose best practices for assimilating GEO NO<sub>2</sub> observations, which can serve as reference for future research.

**Software:** Chia-Hua Hsu, Arthur P. Mizzi, Jian He, Colin Harkins  
**Supervision:** Daven K. Henze, Arthur P. Mizzi, Brian C. McDonald  
**Visualization:** Chia-Hua Hsu  
**Writing – original draft:** Chia-Hua Hsu  
**Writing – review & editing:** Daven K. Henze, Arthur P. Mizzi, Gonzalo González Abad, Matthew S. Johnson, Brian C. McDonald

## 1. Introduction

Nitrogen oxides ( $\text{NO}_x = \text{NO} + \text{NO}_2$ ) are important air pollutants, and they contribute to the formation of tropospheric ozone ( $\text{O}_3$ ) and aerosols (Seinfeld & Pandis, 2012) which degrade air quality and lead to premature mortality (Jerrett et al., 2009; Pope et al., 2009). Current estimates of  $\text{O}_3$  and fine particulate matter ( $\text{PM}_{2.5}$ ) health impacts in the US are 20,000 and 45,000 premature deaths per year, respectively (Dedoussi et al., 2020). To formulate effective policies for reducing this problem, it is critical to identify the main sources of  $\text{NO}_x$  contributing to air pollution and to evaluate trends in  $\text{NO}_x$  emissions stemming from existing emission mitigation policies. Sources of  $\text{NO}_x$  in the contiguous United States (CONUS) consist mainly of fossil-fuel combustion such as mobile source engines, energy generation, and industrial processes (EPA, 2020). Other activities such as wildfires (Wiedinmyer et al., 2011), lightning (Nault et al., 2017), and microbes in soil (Vinken et al., 2014) are also important emitters of  $\text{NO}_x$  in specific seasons and regions. The diverse and dynamically evolving natural and anthropogenic sources of  $\text{NO}_x$  pose a challenge for emissions estimation. Given the large uncertainty and variability of  $\text{NO}_x$  emissions, keeping up with current  $\text{NO}_x$  emissions is critical for reliable air quality forecasting (Campbell et al., 2021).

Traditionally, a “bottom-up” approach is used to develop  $\text{NO}_x$  emission inventories across different sectors, such as the US Environmental Protection Agency’s National Emissions Inventory (NEI) (EPA, 2020). This approach builds emissions inventories by multiplying detailed sectoral-based activity data and corresponding emissions factors, which has the advantage of offering better insight into sector-specific emissions. Bottom-up emissions inventories can be used to identify the major sources of emissions, aid in the development of mitigation strategies, and track progress toward emissions reduction goals. However, the development of bottom-up inventories is laborious and time-consuming, which usually means that bottom-up emission inventories lag several years from present day. Uncertainties in emissions factors and activity statistics could also result in poor representation of emissions in terms of their spatial-temporal variation and magnitude (Fujita et al., 2012; Gately et al., 2015; McDonald, McKeen, et al., 2018). Several modeling studies have reported large uncertainty associated with  $\text{NO}_x$  emissions in the NEI when evaluating the model simulated  $\text{NO}_x$  against in situ, aircraft, and remote-sensing measurements in the U.S. Anderson et al. (2014) found the mobile  $\text{NO}_x$  emissions in NEI-2011 over the Baltimore/Washington region were overestimated by 51%–70%. Kim et al. (2011) found that industrial and ship  $\text{NO}_x$  emissions in the NEI-2005 over the Houston Ship Channel could potentially be biased high by 70%. McDonald, McKeen, et al. (2018) compared the  $\text{NO}_x$  emissions from the Fuel-based Inventory of motor-Vehicle emissions (FIVE) against NEI and found that mobile source  $\text{NO}_x$  emissions in the NEI-2014 were likely biased high by ~28% in the southeast U.S., but Li et al. (2021) in a more recent analysis of FIVE versus NEI-2017 found them to be similar.

Alternatively, the “top-down” approach, which estimates  $\text{NO}_x$  emissions by combining chemical transport model (CTM) simulations with observations through inverse modeling techniques or directly infers  $\text{NO}_x$  emissions from the remote sensing observation (e.g., Beirle et al., 2011), can improve our understanding of  $\text{NO}_x$  sources and chemistry in a more timely manner. Nevertheless, discrepancies in the CTM and biases in the observations could introduce uncertainties in the top-down emissions, which require further assessment to ensure their reliability (Elguindi et al., 2020). Satellite nitrogen dioxide ( $\text{NO}_2$ ) retrievals are well suited for  $\text{NO}_x$  emissions inversion given their wide geographic coverage and have been extensively used in constraining  $\text{NO}_x$  emissions from local to global scales (e.g., Dix et al., 2022; Duncan et al., 2013; Martin et al., 2003; Müller & Stavrou, 2005; Qu et al., 2017). Two advanced data assimilation (DA) methods are typically utilized to conduct top-down emissions estimates:

1. The four-dimensional variational assimilation (4D-VAR) optimizes the emissions by minimizing the cost function using the adjoint model (e.g., Cao et al., 2022; Choi et al., 2022; Elbern et al., 2007; Qu et al., 2019; Stavrou et al., 2013).
2. The ensemble Kalman filter (EnKF) approach utilizes the flow-dependent error covariance generated by the ensemble of model simulations to relate the observation information to emissions (e.g., Barbu et al., 2009; Huang et al., 2022; Ma et al., 2019; Miyazaki et al., 2012, 2017; Peng et al., 2018; Zhang, Li, Wang, et al., 2021; Zhang, Li, Wei, et al., 2021).

Recent studies of  $\text{NO}_x$  emissions top-down estimates mainly use data from low-earth orbit (LEO) instruments such as the Ozone Monitoring Instrument (OMI) (Lamsal et al., 2021) and the TROPOspheric Monitoring Instrument (TROPOMI) (van Geffen et al., 2020). However, the low temporal resolution of LEO measurements

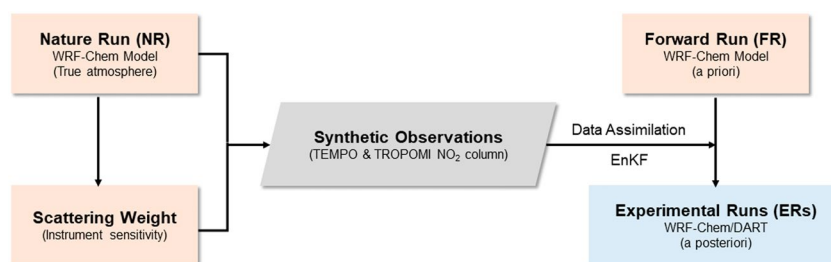
may restrict the ability of emissions inversion to identify key features of emissions sources, such as diurnal variability. The recent and upcoming launches of several geostationary (GEO) satellite instruments, consisting of the Geostationary Environment Monitoring Spectrometer (GEMS) (Kim et al., 2020) over Asia (launched February 2020), the Tropospheric Emissions: Monitoring of Pollution (TEMPO) (Zoogman et al., 2017) over North America (launched April 2023), and Sentinel-4 over Europe (anticipated launch in late 2024), are expected to be utilized more in top-down emissions inversion as they provide unprecedented high spatiotemporal resolution  $\text{NO}_2$  measurements during the daytime. These spatially and temporally dense observations can resolve fine-scale emission patterns, such as major point sources, the difference between the urban core and surrounding suburban regions, and the tracking of emissions from miscellaneous activities (e.g., lightning  $\text{NO}_x$ , biomass burning).

An Observing System Simulation Experiment (OSSE) (Lahoz et al., 2010) is a modeling experiment designed to assess the potential impact of new observing systems (e.g., GEO instruments) on operational forecast when actual observational data is unavailable. The synthetic observations are created and assimilated in an OSSE to quantify the improvements due to the new observing network. For example, Zoogman et al. (2014) conducted an OSSE and demonstrated that assimilation of future TEMPO  $\text{O}_3$  observations can greatly improve  $\text{O}_3$  simulations in the Intermountain West of US by capturing 82% of the high-ozone days. Similarly, Shu et al. (2022) reported that assimilating  $\text{O}_3$  measurement from GEMS reduced the root-mean-square-errors (RMSE) of simulated surface  $\text{O}_3$  concentration by 7.2%–19.2% over urban regions in East Asia. The modeled  $\text{O}_3$  vertical profiles were also improved in the middle to upper troposphere.

Nevertheless, most published OSSE studies have only focused on evaluating the benefits of updating trace-gas concentrations using GEO data (e.g., Barré et al., 2016; Quesada-Ruiz et al., 2020; Shu et al., 2022; Timmermans et al., 2019; Zoogman et al., 2014), while the influence of the assimilation GEO measurement on top-down emissions estimates has not been fully investigated. Liu et al. (2017) examined the ability of TEMPO  $\text{NO}_2$  measurements to constrain  $\text{NO}_x$  emissions. The results demonstrated that the emissions inversion perform better when the error in the forecast wind field is less than  $1.5 \text{ m s}^{-1}$ . However, Liu et al. (2017) only studied the Denver metropolitan region, and they did not explore the improvement of assimilating GEO measurements on emissions inversion compared to the currently existing LEO platforms. As a result, here we propose an OSSE to investigate the potential benefit of assimilating TEMPO  $\text{NO}_2$  observations in constraining anthropogenic  $\text{NO}_x$  emissions versus TROPOMI data over the CONUS. This also helps inform the design of atmospheric composition instruments for follow-on geostationary satellite missions, such as the Geostationary Extended Observations (GeoXO) planned over North America in the 2030s–2050s. In this work, we use WRF-Chem/Data Assimilation Research Testbed (DART) (Mizzi et al., 2016, 2018), a state-of-the-art regional ensemble chemical weather forecast/DA system, to assimilate synthetic TEMPO/TROPOMI  $\text{NO}_2$  observations and optimize  $\text{NO}_x$  emissions in our OSSE. Our first objective is to assess how GEO  $\text{NO}_2$  measurements improve  $\text{NO}_x$  emissions inversion compared to LEO  $\text{NO}_2$  measurements by conducting an OSSE during the COVID-19 lockdown period when the traffic/economic activities and their associated anthropogenic emissions are substantially decreased (Forster et al., 2020; Gkatzelis et al., 2021). We take advantage of this unprecedented reduction in emissions as an ideal and realistic scenario to investigate if a top-down emissions estimate approach can properly constrain  $\text{NO}_x$  emissions perturbations caused by the COVID-19 pandemic lockdown. Second, we investigate the optimal approaches for assimilating GEO  $\text{NO}_2$  observations by assessing the performance of  $\text{NO}_x$  emissions inversion from various DA configurations in terms of DA window, data filtering, and which parameters to optimize.

## 2. Data and Methods

Figure 1 depicts our OSSE framework, which consists of four stages. We choose the first week of April 2020 as the study period because the COVID-19 lockdown rapidly changed anthropogenic emissions. We first conduct a *nature run* (NR) using the Weather Research and Forecasting model coupled with Chemistry (WRF-Chem) (Grell et al., 2005) to represent the “true” atmosphere. The NR is driven by “true” anthropogenic emissions that consider the societal disruption of the COVID-19 pandemic on anthropogenic emissions. Next, the synthetic GEO (TEMPO) and LEO (TROPOMI)  $\text{NO}_2$  observations are sampled from the NR according to their observing strategies (described in Section 2.2). We chose TROPOMI as the representative of the LEO instrument because it is a newer instrument that provides high-quality observations and is well-known in the research community. The spatial resolution of TEMPO and TROPOMI is similar, which also makes them suitable for comparison. Finally, the experimental runs (ERs) are initialized from a *forward run* (FR), which is a parallel WRF-Chem simulation



**Figure 1.** Flowchart of the Observing System Simulation Experiment (OSSE) for assimilation of synthetic TEMPO and TROPOMI NO<sub>2</sub> slant column density observations.

driven by business-as-usual (BAU) emissions (a priori emissions) that is not affected by the COVID-19 perturbations. The synthetic TEMPO and TROPOMI measurements are then assimilated by WRF-Chem/DART to update the BAU NO<sub>x</sub> emissions. By comparing the posterior NO<sub>x</sub> emissions from the ERs (e.g., assimilating TEMPO or TROPOMI data) with the true emissions, we evaluate the benefit of assimilating TEMPO data and propose best practices for NO<sub>x</sub> emissions top-down estimates using GEO observations.

### 2.1. WRF-Chem Configurations

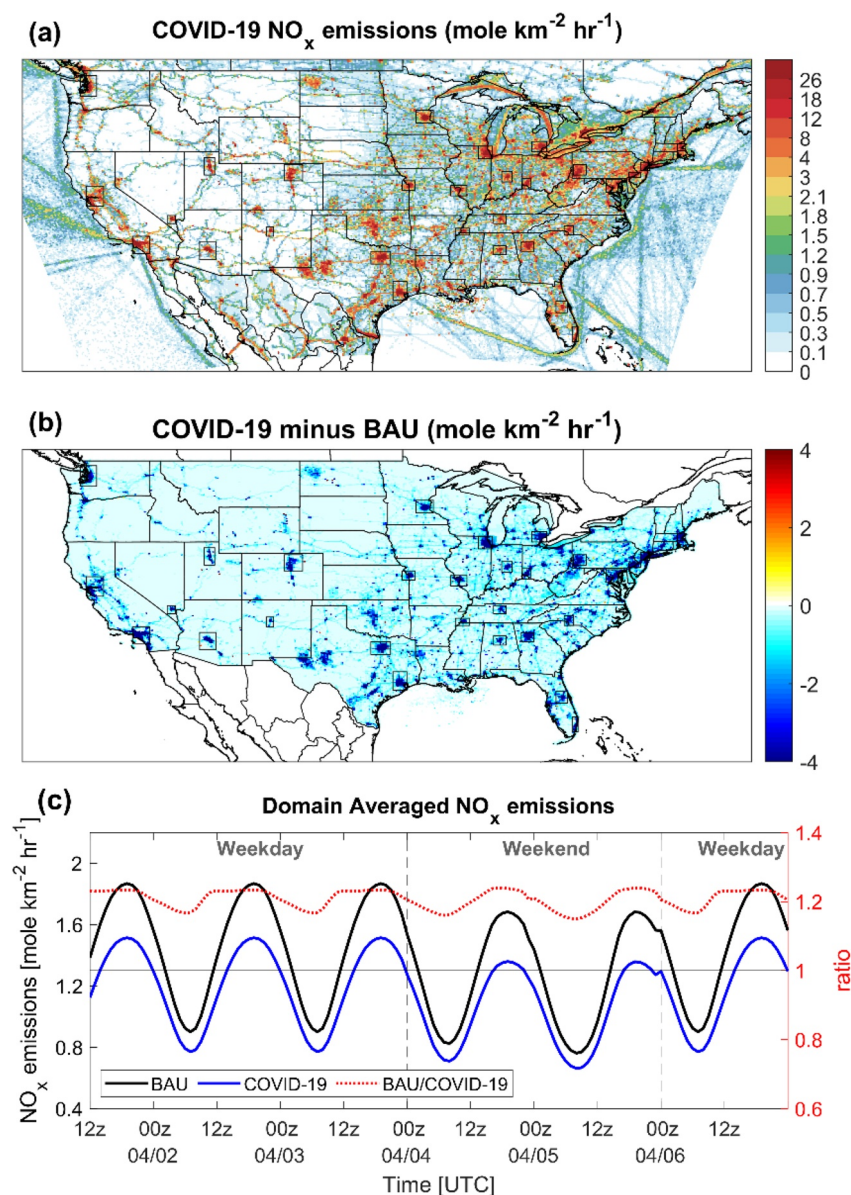
For both the NR and FR we use WRF-Chem version 4.2.2. The horizontal spatial resolution is 12 × 12 km<sup>2</sup> over the CONUS (Figure S1 in Supporting Information S1), with 50 vertical layers spanning from the surface to 50 hectopascal (hPa). The gas-phase and aerosol chemistry are based on the RACM\_ESRL\_VCP (Coggon et al., 2021) and MADE/VBS (Ahmadov et al., 2012) schemes. The meteorological and chemical initial and boundary condition (BCs) are derived from the NCEP North American Mesoscale analysis (<https://rda.ucar.edu/datasets/ds609.0/>) and Realtime Air Quality Modeling System (RAQMS, <http://raqms-ops.ssec.wisc.edu/>) (Pierce et al., 2003), respectively. More details regarding the model configuration are provided in Table S1 in Supporting Information S1. The simulation period spans 30 March to 06 April 2020, with the first 2 days reserved for model spin-up.

The anthropogenic emissions used in this study are a hybrid of several bottom-up inventories, which are summarized in Table 1. The NR is driven by emissions from the COVID-19 lockdown period (COVID-19 emissions), while the FR utilizes the same model configuration as the NR, but it is driven by BAU emissions. The FR represents a priori trace-gas concentrations without the disturbance of the pandemic, which also provides the initial conditions (ICs) for the WRF-Chem/DART ERs (Section 2.4). The BAU and COVID-19 anthropogenic emissions were grouped by emission sector and, where applicable, adjusted from base year inventories using monthly scaling factors developed from applicable energy and economic data sets for March and April in 2019 and 2020, respectively. The emissions outside of CONUS were assumed to be the same between BAU and COVID-19

**Table 1**  
Emissions Inventories for Each Sector

Sources	Emissions inventory
Mobile	Fuel-based Inventory of motor-Vehicle Emissions (FIVE) (McDonald et al., 2014; McDonald, McKeen, et al., 2018)
Power plant	Continuous Emissions Monitoring System (CEMS)
Other point and area sources	NEI 2017 (EPA, 2020)
Oil and Gas	Fuel-based Oil and Gas (FOG) inventory (Francoeur et al., 2021)
Canada/Mexico	Copernicus Atmospheric Monitoring Service (CAMS) version 4.2 (Dombia et al., 2021)
Other	Volatile Chemical Products (VCPs) (McDonald, de Gouw, et al., 2018)
Biogenic <sup>a</sup>	Biogenic Emissions Inventory System (BEISv3.14)
Fire <sup>b</sup>	N/A

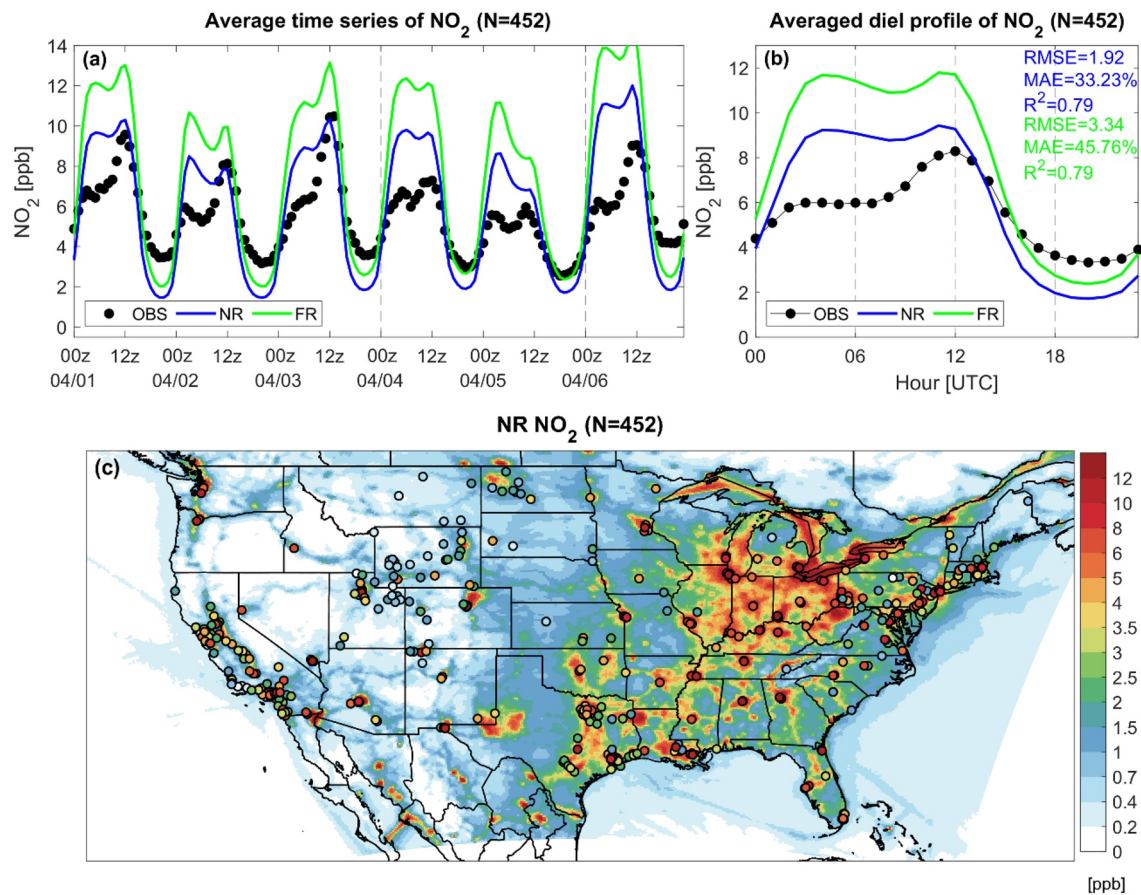
<sup>a</sup>Isoprene emissions are reduced by a factor of two following Li et al. (2021) to reduce the biases in the simulated formaldehyde concentration. <sup>b</sup>We did not include wildfire emissions in this study because wildfires were relatively inactive between March and April 2020 over the CONUS.



**Figure 2.** (a) Spatial distribution of the 6-day averaged (01–06 April) COVID-19 NO<sub>x</sub> emissions. The black boxes highlight 29 major urban areas in CONUS, (b) difference between BAU and COVID-19 NO<sub>x</sub> emissions, and (c) time series [GMT] of domain averaged BAU and COVID-19 NO<sub>x</sub> emissions and their relative ratio (i.e., BAU/COVID-19).

scenarios. The COVID-19 perturbations reduce the overall US NO<sub>x</sub>, carbon monoxide (CO), volatile organic compounds (VOCs), and sulfur dioxide (SO<sub>2</sub>) emissions by 23%, 30%, 20%, and 25%, respectively, compared to BAU emissions. The decreases in NO<sub>x</sub> and CO emissions are mainly from mobile sources (Harkins et al., 2021). The drop of VOCs emissions is driven by changes in the oil and gas sector and the volatile chemical products in the urban areas, while industrial and power plant sources dominate the changes in the SO<sub>2</sub> emissions. The reduction for other emissions such as ammonia (NH<sub>3</sub>) and fine particles (PM<sub>2.5</sub>) are relatively small (within 2%). A list of data sources used to generate the COVID-19 and BAU emissions input for WRF-Chem can be found at <https://csl.noaa.gov/groups/csl7/measurements/2020covid-aqs/emissions/>.

The spatial and temporal variations of COVID-19 and BAU NO<sub>x</sub> emissions are shown in Figure 2. COVID-19 NO<sub>x</sub> emissions are lower than the BAU over the CONUS, with the greatest drop occurring in metropolitan areas, where traffic and economic activities were greatly impacted by pandemic lockdowns (Figure 2b), and the two inventories are statistically distinct at the 95% confidence level (see Text S1 in Supporting Information S1). The



**Figure 3.** (a) Averaged time series of NR (FR) simulated surface NO<sub>2</sub> and Air Quality System (AQS) NO<sub>2</sub>, (b) Diurnal variation of NR (FR) and AQS NO<sub>2</sub>, the root-mean-square-errors, mean absolute error, and  $R^2$  are also listed, (c) Six-day averaged spatial map of NR simulated NO<sub>2</sub> overlay with AQS NO<sub>2</sub> data.  $N$  represents the total AQS monitoring sites.

diurnal and daily variation of domain-averaged BAU and COVID-19 NO<sub>x</sub> emissions are followed similar diurnal patterns (Figure 2c), with COVID-19 emissions being 20% lower than BAU. These two inventories also show the weekend-weekday variation in NO<sub>x</sub> emissions, with emissions lower on weekends (04–05 April) than on weekdays.

A robust OSSE demands the NR to properly represent the real atmospheric trace gas concentrations and to differ sufficiently from the FR. Figure 3 shows the spatial-temporal variation of NR and FR simulated surface NO<sub>2</sub> concentrations against surface observations from US EPA Air Quality System (AQS) data. The NR can reasonably capture the daily variation of surface NO<sub>2</sub> (Figure 3a), whereas the FR has persistent positive biases when compared to the NR and AQS observations. In terms of the diurnal cycle of NO<sub>2</sub>, the NR shows overall consistency between simulated surface NO<sub>2</sub> and hourly observations (Figure 3b), with a RMSE of 1.92 ppb and a mean absolute error (MAE) of 33.23%. On the contrary, FR displays larger error statistics when driven by BAU emissions, and the substantial difference between the NR and FR (RMSE = 1.84 and MAE = 31.24%) can be seen. The NR also accurately reflects the spatial distribution of surface NO<sub>2</sub> over the CONUS, with a spatial correlation coefficient of 0.71 (Figure 3c). The NO<sub>2</sub> hotspots in urban areas and the NO<sub>2</sub> gradient between the eastern US and the remote mountainous regions in the western US are well captured by NR. As a result, we believe our NR is a reliable proxy for the real-world NO<sub>2</sub> concentration for generating the synthetic satellite NO<sub>2</sub> measurements. The positive bias in the FR is driven by the errors in the input emissions, which can be constrained by DA.

## 2.2. Synthetic TEMPO and TROPOMI NO<sub>2</sub> Observations

The TEMPO and TROPOMI instruments are both solar-backscatter spectrometers that measure total column amounts of trace gases such as NO<sub>2</sub>, formaldehyde (HCHO), sulfur dioxide (SO<sub>2</sub>), and O<sub>3</sub> in the ultraviolet-visible spectral ranges (Veefkind et al., 2012; Zoogman et al., 2017). TROPOMI was launched onboard the Copernicus

Sentinel 5 Precursor (S5p) satellite on 13 October 2017. The satellite is in a sun-synchronous, LEO (817 km) with a daily equator overpass time of approximately 13:30 local solar time, allowing for one (up to two) NO<sub>2</sub> column measurements at a specific location on the same day. The pixel sizes at the nadir are 3.5 × 7 km<sup>2</sup> (3.5 × 5.6 km<sup>2</sup> since 6 August 2019), with near-global coverage in 1 day. TEMPO launched in April 2023. It is hosted by an Intelsat 40e communications satellite and is in geostationary orbit (35,786 km) at 91.0° W longitude with a Field of Regard (FoR) covering North America. It will collect hourly NO<sub>2</sub> column measurements during the daytime (e.g., 8–17 local solar time) with a native resolution of 2 × 4.5 km<sup>2</sup> at the center of its FoR.

We generate and assimilate synthetic TEMPO and TROPOMI tropospheric NO<sub>2</sub> slant column density (SCD) retrievals to evaluate which data set can better constrain NO<sub>x</sub> emissions. A synthetic tropospheric NO<sub>2</sub> SCD ( $\Omega$ ) is calculated as:

$$\Omega = \sum_k SW_k X t_k + \varepsilon \quad (1)$$

where  $\Omega$  (molec cm<sup>-2</sup>) is the vertical integration of the product of the NO<sub>2</sub> number concentrations ( $X t_k$ ) from the NR and the scattering weights ( $SW_k$ ) at layer  $k$  which runs from the surface to the tropopause, and  $\varepsilon$  (molec cm<sup>-2</sup>) is a random Gaussian error. The calculation of  $SW_k$  is described in the next section.

### 2.2.1. Scattering Weight Calculation

The SW at a given altitude describes the sensitivity of the retrieved NO<sub>2</sub> to the abundance of the absorber (e.g., NO<sub>2</sub> concentration) at that altitude (Lamsal et al., 2021). For a given observation location, the SW vector is derived based on a Look-Up-Table (LUT). The SW LUT is constructed using version 2.8 of the Vector Linearized Discrete Ordinate Radiative Transfer (VLIDORT) model (Spurr, 2006) assuming a Rayleigh atmosphere. The LUT provides information on SWs as a function of altitude (expressed as atmospheric pressure), solar zenith angle (SZA), viewing zenith angle (VZA), relative azimuth angle (RAA), surface albedo, surface (cloud) pressure, and total ozone column. Aerosols are not considered despite being a significant source of uncertainty (Jung et al., 2019; Lorente et al., 2017). There are two reasons for that: (a) it is not easy to obtain aerosol information at the pixel level and (b) it will increase the size of the LUT quite significantly. The 22 ozone profiles employed in the VLIDORT simulations were derived using OMI ozone profile retrievals (PROFOZ) (Liu et al., 2010). They provide climatological values for three latitudinal bands for different total columns at tropical, midlatitude, and polar regions. The selection of the LUT nodes is based on minimizing the interpolation errors. Here, we use the same LUT to compute SWs for TEMPO and TROPOMI observations given that we could not get the operational LUTs for the real TROPOMI data, and we assume clear-sky conditions for the SW calculation. The average synthetic TROPOMI NO<sub>2</sub> SW vertical profiles resembled the actual profiles (not shown). Therefore, we expect the impact of using the same LUTs to be relatively minor.

We collect input data for the SW computation from several sources. Surface pressure is provided by the NR. The total ozone column is calculated by combining the outputs of the NR and RAQMS, with RAQMS supplying the ozone concentration above the top of the WRF-Chem model (50 hPa). Hourly surface albedo is derived from the geometry-dependent surface Lambertian-equivalent reflectivity product (Qin et al., 2019) calculated at the 440 nm wavelength. The pixel location and SW pressure levels of TEMPO are obtained from the TEMPO NO<sub>2</sub> proxy data set (Naeger et al., 2021). The derivation of TEMPO proxy data can be found in Text S2 in Supporting Information S1. The satellite viewing geometry parameters (e.g., SZA, VZA, and RAA) for TEMPO are computed hourly following Liu et al. (2017) using MATLAB with inputs of the location, terrain height, and local time of each TEMPO pixel, as well as the TEMPO sensor's location (0°N, 91°W, and 35,786 km). The TROPOMI pixel location, SW pressure levels, SZA, VZA, and RAA are derived from TROPOMI NO<sub>2</sub> level 2 swath data in April 2020. Figure S2 in Supporting Information S1 depicts the hourly averaged SW vertical profile from the surface to the top of the atmosphere. The TEMPO and TROPOMI SWs have 47 and 34 total vertical layers, respectively. The vertical profiles of TEMPO SW show significant diurnal variation, with greater sensitivity to the near-surface atmosphere during midday hours (between 16 and 21 UTC) and increased sensitivity to the middle and upper atmosphere during sunrise and sunset hours. In contrast, TROPOMI SW profiles show very little diurnal variability due to the instrument's similar local overpass time (at around 13:30 local solar time).

### 2.2.2. Observation Error Estimation and Perturbed Observations

We begin by calculating the “perfect tropospheric NO<sub>2</sub> SCD ( $\Omega_r$ )” by solving for the first term on the right-hand side of Equation 1 for TEMPO and TROPOMI. A random Gaussian error ( $\varepsilon$ ) is then added to the “perfect observation” to create a “perturbed observation”—the synthetic tropospheric NO<sub>2</sub> SCD ( $\Omega$ ) which will be assimilated

in the ERs. The error ( $\epsilon$ ) is computed by taking a random sample from a Gaussian distribution that describes the instrument's error characteristics which is defined as:

$$\epsilon = N(0, \sigma_o). \quad (2)$$

$\sigma_o$  is the retrieval/observation error, and its characteristics are associated with  $\Omega$  via the following relationship:

$$\sigma_o = \gamma\Omega \quad (3)$$

where  $\gamma$  represents relative mean uncertainty which is the ratio of  $\sigma_o$  to its corresponding observation value ( $\Omega$ ). The magnitude of  $\gamma$  for  $\text{NO}_2$  column measurements can differ between polluted and clean regions in the actual observations. For example, the polluted pixels generally have lower  $\gamma$  values (Boersma et al., 2004). Therefore, Liu et al. (2017) estimated the synthetic TEMPO  $\text{NO}_2$  observation error by assuming two distributions of  $\gamma$  values for polluted and clean pixels, respectively. However, the relationship between the  $\gamma$  value and the  $\text{NO}_2$  actual SCD varies in a complex manner and may depend on other parameters such as SZA and VZA. Figure S3 in Supporting Information S1 shows  $\gamma$  as a function of  $\text{NO}_2$  SCD, SZA, and VZA for TEMPO proxy data and actual TROPOMI data. In general,  $\gamma$  displays a non-linear and inversely proportional response to the  $\text{NO}_2$  SCD, with low (high)  $\gamma$  values associated with polluted (clean) pixels for both sensors (Figures S3a and S3d in Supporting Information S1). The  $\gamma$  value also exhibits a linear relationship with SZA and VZA for TEMPO, while there is no clear relationship in the TROPOMI data.

To better reproduce the instrument's error characteristics and estimate the observation error ( $\sigma_o$ ) for the synthetic  $\text{NO}_2$  SCD products, we developed a multiple linear regression approach that predicts the magnitude of  $\gamma$  as a function of SZA, VZA, and  $\Omega$  for each pixel of synthetic TEMPO and TROPOMI data. The training data set for the regression analysis is compiled from TEMPO proxy data and level 2 data from actual TROPOMI  $\text{NO}_2$  observations, and only observations with low cloud cover and high retrieval quality flag (qf) values are used (qf > 0.75 and cloud fraction < 0.2 for TROPOMI data; qf = 0, and cloud fraction < 0.2 for TEMPO proxy data). It should be noted that the  $\gamma$  value for TEMPO is calculated based on the total column data in the TEMPO proxy data set (i.e., the ratio of total SCD error to total SCD) due to the lack of tropospheric column error in TEMPO proxy data set, and we simply assume the error characteristics are the same in the total and tropospheric column when estimating TEMPO tropospheric SCD error. The impact of TEMPO observation error assumption on emissions inversion will be discussed in Section 4.

We divided the training data into groups based on  $\text{NO}_2$  SCD, SZA, and VZA, with detailed data binning setups provided in Table S2 in Supporting Information S1. Multiple linear regression was then performed on each group to fit a prediction equation as follows:

$$\gamma_m = a + b \ln(\Omega/10^{15}) + c(\Omega/10^{15})^{-1.5} + dVZA + eSZA. \quad (4)$$

$a$ ,  $b$ ,  $c$ ,  $d$ , and  $e$  are regression coefficients for corresponding predictor, and  $\gamma_m$  is the predicted relative mean uncertainty. For TEMPO, we conduct the regression analysis hourly from 11 to 01 UTC while all overpasses from TROPOMI are used to fit a single regression equation. In the data fitting process, we use both SZA and VZA as predictors for TROPOMI, but only VZA for TEMPO because incorporating SZA in the regression analysis degraded the goodness of fit. Figure S4 in Supporting Information S1 shows examples of  $\gamma_m$  for TEMPO and TROPOMI, and Equation 4 can successfully predict the overall trend of relative mean uncertainty in the training data. Nevertheless, Equation 4 fails to capture the spread of  $\gamma$  from the training data set. To properly reflect the uncertainty associated with  $\gamma$ , we compute the standard deviation of the residual ( $\eta$ ) of Equation 4 and estimate  $\eta$  as a function of  $\text{NO}_2$  SCD in each data group using another linear regression, as follows:

$$\eta = f + g(\Omega/10^{15}) \quad (5)$$

The  $f$  and  $g$  represent the regression coefficients for predicting the  $\eta$ . Figure S5 in Supporting Information S1 shows  $\eta$  value and the corresponding regression line for the same data group as in Figure S4 in Supporting Information S1. In general,  $\eta$  is inversely proportional to the  $\text{NO}_2$  SCD, and the estimated  $\eta$  is then used to perturb  $\gamma_m$  by assuming  $\gamma_m$  followed a Gaussian distribution, which is defined as follows:

$$\gamma_f = N(\gamma_m, \eta). \quad (6)$$



$\gamma_f$  can better capture the trend and dispersion of the input training data (Figure S6 in Supporting Information S1) after imposing a random perturbation.

The regression model from Equations 4–6 is used to estimate the  $\gamma_f$  for synthetic NO<sub>2</sub> column. However, the synthetic  $\Omega$  is initially unknown; we thus use the  $\Omega_{tr}$ , SZA, and VZA from our synthetic TEMPO and TROPOMI NO<sub>2</sub> data as inputs to predict  $\gamma_f$ . Once  $\gamma_f$  is known, we use Equation 7 to calculate the observation error for individual pixels which is defined as:

$$\sigma_o = \gamma_f \Omega_{tr}. \quad (7)$$

The observation error is then applied to perturb the  $\Omega_{tr}$  and create the final perturbed synthetic TEMPO and TROPOMI  $\Omega$  products by Equations 2 and 1.

It should be emphasized that NO<sub>2</sub> column observations are always non-negative, which means that any perturbation imposed on them must be done with caution. When calculating a random Gaussian perturbation by Equation 2, a negative NO<sub>2</sub> column might occur if the observation error exceeds 35% of the observed value. Positive biases in the perturbed observations can emerge if the negative NO<sub>2</sub> column is reset to zero or a small positive value. This happens more often in the synthetic TROPOMI NO<sub>2</sub> data set because of its larger  $\gamma_f$  value. To prevent positive biases in the synthetic TROPOMI data set, we have set the upper limit of the  $\gamma_f$  value to 0.35 (35%). Additionally, we define the time of synthetic NO<sub>2</sub> observations to be precisely (and only) on the hour. For example, all TEMPO and TROPOMI pixels within 18–19 UTC are defined at 18 UTC for simplicity.

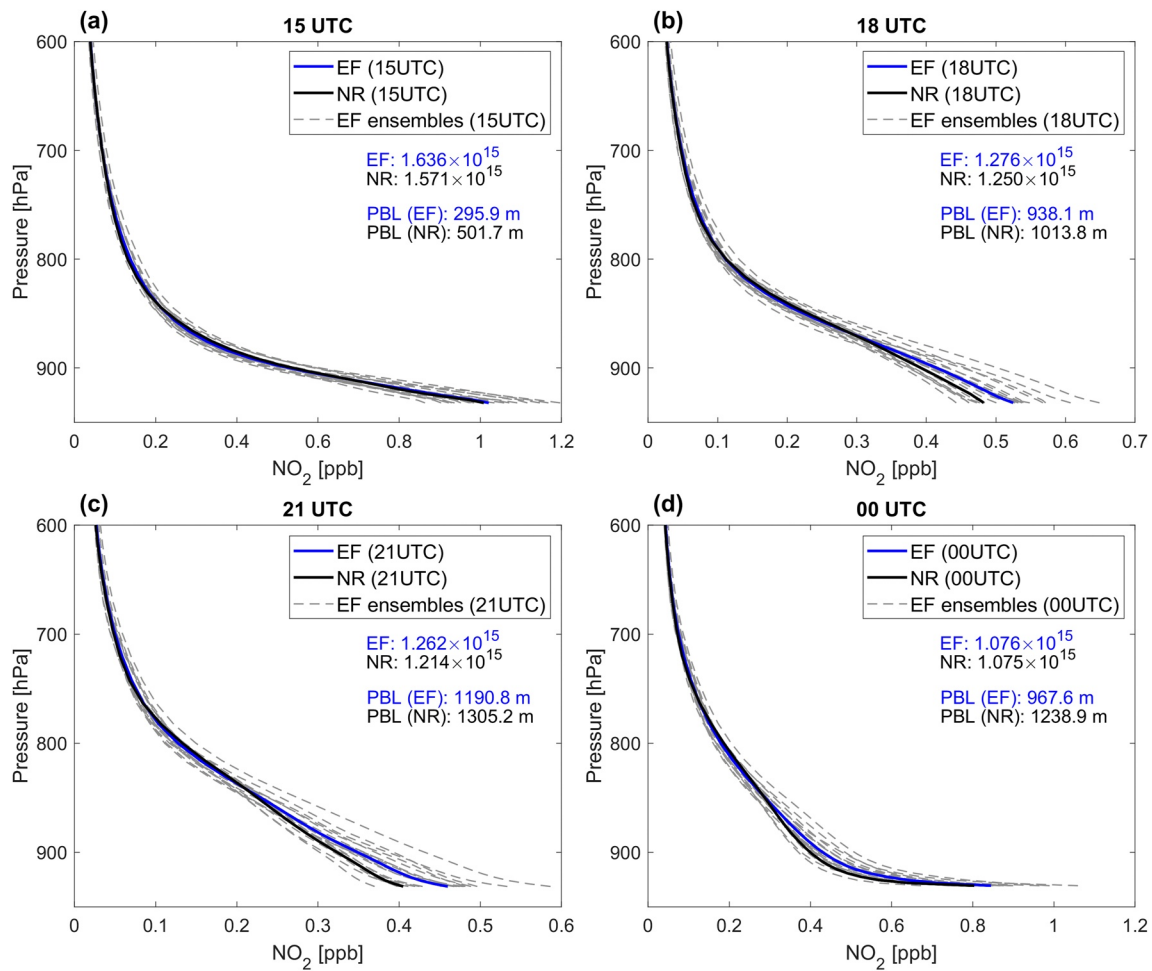
### 2.3. Experimental Runs With WRF-Chem/DART

WRF-Chem/DART combines WRF-Chem with the Data Assimilation Research Testbed (DART) to conduct chemical-weather ensemble forecasts and assimilate synthetic TEMPO and TROPOMI NO<sub>2</sub> observations to update NO<sub>x</sub> emissions. DART is an open-source EnKF based DA system (Anderson et al., 2009). It offers users a variety of EnKF analysis algorithms and aids in managing covariance inflation and localization during DA. WRF-Chem/DART can assimilate a wide range of data types, including in situ and remote sensing meteorological/chemical observations to constrain their associated ICs/BCs. The state augmentation approach (J. L. Anderson, 2003; Liu et al., 2017; Ma et al., 2019) is used to constrain the emissions, for example, by adding the NO<sub>x</sub> emissions to the WRF-Chem state vector during the assimilation step which allows the EnKF to use ensemble correlations between observations and states to constrain the elements of the state vectors (J. L. Anderson, 2003). This enables us to use satellite-based NO<sub>2</sub> observations to constrain NO<sub>x</sub> emissions.

The ERs are initialized from the FR output at 00 UTC on 1 April. The meteorological ICs ensemble is created by randomly perturbing the FR output. The perturbation for each member is sampling from NCEP background error covariance using the WRF DA system (WRFDA), as described by Liu et al. (2017). The *pert\_wrf\_bc* module in the WRF-Chem/DART generates random perturbations for each member of the meteorological lateral BCs ensemble using the random variation approach. The ensemble of chemical ICs and BCs are created following Mizzi et al. (2016) and Ma et al. (2019) with imposition of spatial-temporal ensemble error decorrelations. We apply the spatial-temporal correlated Gaussian perturbation with a standard deviation of 30% around a defined mean (the FR and RAQMS output). The horizontal and vertical correlation of the perturbations is modeled by an isotropic exponential decay function with correlation lengths of 300 and 1 km, respectively (Evensen, 1994, 2003). The temporal correlation length is assumed to be 24 hr and to decline exponentially with time (Boynard et al., 2011; Gaubert et al., 2014).

We produce emission ensembles in the same way as we generate chemical ICs perturbations (Mizzi et al., 2016, 2018). The hourly emissions ensemble is created by imposing a Gaussian perturbation around the BAU emissions with a standard deviation of 30%. This value is consistent with other inverse modeling studies which typically assume an uncertainty range of 30%–50% for NO<sub>x</sub> emissions (Ma et al., 2019; Qu et al., 2017; Sekiya et al., 2022; Souri et al., 2020). The emissions perturbations have the same spatial correlation length as chemical ICs, and the temporal correlation length is set to 72 hr.

One concern in our OSSE is we used the same numerical model for both NR and ERs (i.e., identical twin OSSEs) which may present overly optimistic forecast skill and underestimate the background error in the ERs (Lahoz et al., 2010). We address those shortcomings by altering the WRF-Chem model configurations in the ERs to create some independence between the NR and the ERs, and this approach is also adopted in other OSSE studies



**Figure 4.** Three-day average (04/01–03)  $\text{NO}_2$  vertical profiles over CONUS from the nature run (black curve), ensemble mean of ensemble forecast (EF) (blue curve), and individual members of EF (gray curves) at (a) 15, (b) 18, (c) 21, and (d) 00 UTC. The average TEMPO  $\text{NO}_2$  SCD and planetary boundary layer height are also listed.

(e.g., Chang et al., 2022; Descheemaeker et al., 2019; Ye et al., 2022). We use 10 different WRF-Chem physics configurations (see Table S3 in Supporting Information S1) to conduct the ensemble forecasts, but the same chemical mechanism is still employed in NR and ERs.

To quantify the benefit of altering model configurations, we conducted an ensemble forecast (EF) initialized from the NR and driven by the perturbed true  $\text{NO}_x$  emissions. Figure 4 displays the average  $\text{NO}_2$  profile over CONUS from NR and EF, and some members (gray curves) in ER display greater divergence from the NR profile (black curve). The presence of these outliers could increase the ensemble spread, thereby mitigating the problem of underestimating the background error when using the same model. Additionally, the EF also tend to predict a higher concentration of  $\text{NO}_2$  (blue curves) in the lower troposphere which results from the lower planetary boundary layer (PBL) height. The PBL height in the EF is biased low by 8%–41% compared to the NR, which enhances the  $\text{NO}_2$  concentration within PBL. The bias in the simulated  $\text{NO}_2$  vertical profile also causes differences (+3%–4%) in the synthetic TEMPO  $\text{NO}_2$  column. This result provides evidence that altering the physics option leads to a distinct  $\text{NO}_2$  profile in the ER, which can be treated as a surrogate of model error. In summary, we anticipate that using different physics options in the ERs will help to address the problem of identical twin OSSEs and enhances the robustness of our conclusion.

Finally, we did not generate synthetic meteorological observations to constrain the forecast meteorology. Instead, we used the hourly grid analysis nudging technique in WRF-Chem to nudge the ER meteorology forecasted meteorology (e.g., wind fields, temperature, and moisture) toward the NR simulated meteorology.

**Table 2**  
WRF-Chem/Data Assimilation Research Testbed Configurations

WRF-Chem/DART configurations	
Filter type	EAKF
Ensemble members	20
DA cycle	3-hr
Prior adaptive inflation	1 and 0.9
Prior inflation damping <sup>a</sup>	0.9 and 0.65
Posterior inflation	Section 2.4
Localization function	Gaspari-Cohn
Horizontal localization half-width	150 km
Vertical localization	Off
Chemical ICs/BCs and emissions uncertainty	30%
Horizontal correlation length	300 km
Vertical correlation length	1 km
Temporal correlation length	24 and 72 hr

<sup>a</sup>The inflation damping of 0.65 is applied to TEMPO-exp1 and TEMPO-exp2 (Table 3), and 0.9 is used for other ERs.

## 2.4. Data Assimilation Configuration

We use a 20-member ensemble to run continuous 3-hr forecast-data assimilation cycles at 0, 3, 6, ..., and 21 UTC. The small ensemble used in this work is mainly a result of computational cost, and emissions inversion studies that use a regional model with a higher horizontal resolution often employ 10–30 ensemble members as well (e.g., Liu et al., 2017; Ma et al., 2019; Timmermans et al., 2019; Zhang, Li, Wang, et al., 2021; Zhang, Li, Wei, et al., 2021). The modeled concentrations and emissions are updated using the ensemble adjustment Kalman filter (EAKF) (J. L. Anderson, 2001, 2003) in WRF-Chem/DART. The Gaspari-Cohn (GC) localization function (Gaspari & Cohn, 1999) is applied to reduce the effect of spurious ensemble correlation across observations and distant state variables due to small ensemble size (20). In this study, the vertical localization is disabled and the horizontal localization half-width for the GC function is set to 150 km. To lessen the erroneous correlations between observations and various kinds of state variables, the state variable localization technique (Ma et al., 2019; Miyazaki et al., 2012) is also used. We eliminate the ensemble correlations to other variables and solely permit TEMPO and TROPOMI NO<sub>2</sub> observations to update the NO<sub>x</sub> concentration and anthropogenic emissions.

The prior adaptive inflation (J. L. Anderson, 2007) in WRF-Chem/DART is utilized to maintain the ensemble spread of chemical states and emissions for covariance inflation. However, we find that this approach is insufficient to

support the emissions ensemble spread. The NO<sub>x</sub> emissions ensemble spread collapses after several DA cycles due to a lack of an emissions model to forecast/grow the error with time. Therefore, a posterior inflation approach from Miyazaki et al. (2012) is used to avoid the contracting of the emissions ensemble standard deviation. We artificially inflate the posterior emissions ensemble spread back to a minimum predefined value (i.e., 50% of the BAU emissions ensemble spread) when the posterior spread falls below this. This inflation uses centering about the ensemble mean, so there is no impact to the ensemble mean emissions. The detailed WRF-Chem/DART configurations are summarized in Table 2.

## 2.5. Emissions Adjustment Scheme

We update the instantaneous NO<sub>x</sub> emissions at the DA cycle hour using the synthetic TEMPO and TROPOMI NO<sub>2</sub> measurements, and WRF-Chem/DART computes the 3-D emissions scaling factor ( $S$ ) for each emission ensemble member as follows:

$$S^i(t) = \frac{e_a^i(t)}{e^i(t)}. \quad (8)$$

$S^i(t)$  is a 3-D emissions scaling factor for individual members  $i$  at DA cycle time  $t$  (hour), while  $e_a^i(t)$  and  $e^i(t)$  (mole km<sup>-2</sup> hr<sup>-1</sup>) are 3-D posterior and prior NO<sub>x</sub> emissions member  $i$  at DA cycle time  $t$  (hour). The emissions scaling factor is advanced in time without damping to adjust the prior emissions and obtain the adjusted prior NO<sub>x</sub> emissions for the next 3 hours as follows:

$$e_p^i(t : t + 3) = e^i(t : t + 3) S^i(t) \quad (9)$$

$e_p^i$  is the adjusted prior NO<sub>x</sub> emissions for the hour over the course of  $t$  to  $t + 3$  for individual members  $i$ , which are used to drive the WRF-Chem ensemble forecast and serve as the prior emissions for the next DA cycle to calculate the new posterior emissions and update emissions scaling factor. It is noted that  $e^i(t)$  is equivalent to  $e_p^i(t)$  for the first DA cycle. With this emissions adjustment scheme, the entire historical emissions adjustment, as well as the structure of posterior emissions ensemble covariance, is preserved and propagated forward in time. Because no satellite observations can be assimilated to update the emissions during the night, we apply the emissions scaling factor from the last DA cycle before sunset to adjust NO<sub>x</sub> emissions overnight. Our emissions adjustment scheme corrects the error in the emissions diurnal pattern at the DA cycle hour. We propagate the

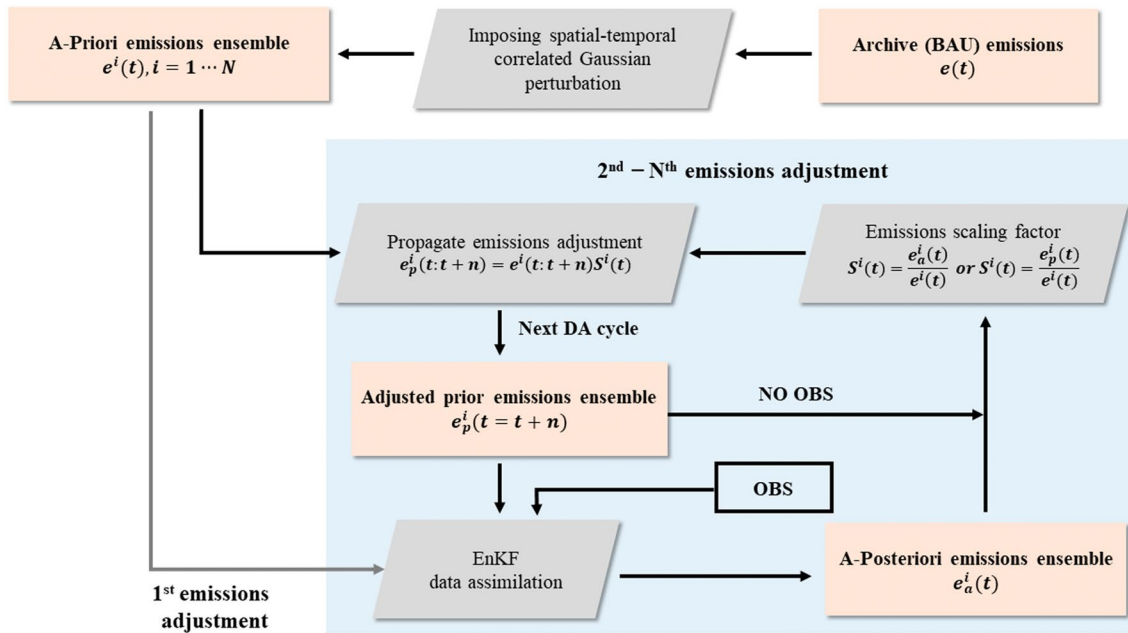


Figure 5. Flowchart of dynamical emissions adjustment scheme.  $N$  is the ensemble size;  $n$  is the data assimilation cycle hours ( $n = 3$  in this study).

emissions scaling factor during the intra-cycle period; therefore, the emissions adjustment scheme may not be able to resolve the intra-cycle diurnal pattern. The process of our emissions adjustment scheme is also summarized in Figure 5.

## 2.6. Experimental Design

### 2.6.1. Two True $\text{NO}_x$ Emissions Scenarios

We use a 6-day study period beginning at 00 UTC on 1 April and ending at 00 UTC on 7 April 2020. The series of ERs were summarized in Table 3. We propose two NRs with different scenarios for true  $\text{NO}_x$  emissions. In the first scenario, the true  $\text{NO}_x$  emissions are consistent with the COVID-19 lockdown period emissions (Figure 2), and we run five ERs, including four TEMPO and one TROPOMI run. The length of the DA window,

**Table 3**  
Configuration for the Different WRF-Chem/DART Experimental Runs

Experiment	DA window (hours)	Hour (UTC) of OBS being assimilated	Update $\text{NO}_x$ concentration	Update $\text{NO}_x$ emissions	#OBS being assimilated
<i>COVID-19 emissions case</i>					
TEMPO-exp	0 <sup>a</sup>	15, 18, 21, 00	Yes	Yes	752,091
TEMPO-exp1	$\pm 1.5^b$	14–01	Yes	Yes	2,144,243
TEMPO-exp2	0 and $\pm 1.5$	15, 17–22, 00	Yes	Yes	1,697,575
TEMPO-EMS	0	15, 18, 21, 00	No	Yes	752,091
TROPOMI-exp	$\pm 1.5$	Varies with each date	Yes	Yes	330,261
<i>Time-varying true <math>\text{NO}_x</math> emissions case</i>					
TEMPO-TVE	0	15, 18, 21, 00	Yes	Yes	752,091
TROPOMI-TVE	$\pm 1.5$	Varies with each date	Yes	Yes	330,261

<sup>a</sup>The synthetic observations time are defined at the hour. As a result, observations at cycle time are assimilated when using 0-hr DA window. The 0-hr DA window is equivalent to a  $\pm 0.5$ -hr DA window when assimilating operational TEMPO data. <sup>b</sup>For  $\pm 1.5$ -hr DA windows, the windows width spans from 13.5–16.5, 16.5–19.5, 19.5–22.5, and 22.5–1.5 UTC for the cycle time at 15, 18, 21, and 00 UTC, respectively.

a parameter in EnKF that controls the hour/amount of observation assimilated in each DA cycle, is the primary distinction between the TEMPO experiments. We start with an experiment (TEMPO-exp) that utilizes a 0-hr DA window (i.e., equivalent to a  $\pm 0.5$ -hr DA window when assimilating operational TEMPO data) and assimilates the TEMPO observations at the DA cycle hour (e.g., 15, 18, 21, and 00 UTC). This is the most restrictive DA window which only assimilates one full scan TEMPO data over CONUS per cycle-time. The main disadvantage of this configuration is that it uses only a small portion ( $\sim 30\%$ ) of the available daily TEMPO observations. In the next experiment (TEMPO-exp1), we extend the DA window to  $\pm 1.5$ -hr for all DA cycles to assimilate 3 hours of TEMPO observations per DA cycle. By comparing TEMPO-exp and TEMPO-exp1, we investigate the effect of DA window on the constrained emissions. We find that using a  $\pm 1.5$ -hr DA window during the early morning and evening DA cycles (e.g., the 15 and 00 UTC cycles, day-to-night transition cycles, referred to as the “transition cycles”) introduces large biases in emissions estimates. As a result, we propose another TEMPO experiment (TEMPO-exp2) that employs a  $\pm 1.5$ -hr DA window at 18 and 21 UTC and a 0-hr DA window for DA cycles at 15 and 00 UTC. The results from those experiments are examined in Section 3.2. We also conduct a TEMPO experiment that solely constrains the  $\text{NO}_x$  emissions and does not constrain the  $\text{NO}_x$  concentrations (TEMPO-EMS). This experiment enables us to document the effect of constraining the  $\text{NO}_x$  concentrations on the  $\text{NO}_x$  emissions. For the TROPOMI experiment (TROPOMI-exp), we use a  $\pm 1.5$ -hr DA window to assimilate all the available observations. The 15, 18, and 21 UTC DA cycles are typically when TROPOMI provides observations within the WRF-Chem domain. We will compare TEMPO-exp and TROPOMI-exp in Section 3.1 to assess the impacts of assimilating geostationary observations as opposed to polar orbiting observations. By comparing the four TEMPO experiments, we can recommend best practices for the use of GEO  $\text{NO}_2$  observations.

In the second scenario, we make the inverse modeling more challenging by modifying the spatial and temporal variation of the true  $\text{NO}_x$  emissions. The goal is to further examine the ability of the differing observation types (GEO—TEMPO vs. LEO—TROPOMI) to constrain the  $\text{NO}_x$  emissions. Specifically, we add daily fluctuations to the true weekday emissions to create temporal variations in the daily  $\text{NO}_x$  emissions. We generate the perturbations by first binning the weekday  $\text{NO}_x$  emissions rate by percentile (e.g., from clean to polluted areas), and then applying a magnitude-dependent scaling factor. We assume areas with high emissions rates, such as cities or highways, have greater daily variability, whereas emissions in rural areas are less disturbed. Figure S7 in Supporting Information S1 depicts the emissions scaling factors for various emission rates as well as the time series of time-varying true  $\text{NO}_x$  emissions. The time-varying true  $\text{NO}_x$  emissions are used to generate a new NR for the second scenario. Then, we generate the corresponding synthetic TEMPO and TROPOMI  $\text{NO}_2$  observations as discussed in Section 2.2, which are assimilated in the corresponding ERs (TEMPO-TVE and TROPOMI-TVE). Comparison of TEMPO-TVE and TROPOMI-TVE provides guidance as to the effectiveness of inverse modeling when there are daily variations in the emissions.

### 2.6.2. Observation Data Preprocessing

The computational cost of directly assimilating raw TEMPO  $\text{NO}_2$  observations is expensive due to high spatiotemporal resolution of the measurements (i.e.,  $10^6$  to  $10^7$  pixels per hour). To reduce the cost, we first used an inverse error weighted superobservation technique (Hartung et al., 2008) based on the observation error (not considering the distance to the grid center or the pixel sizes) to reduce the data volume. However, we found that assimilating superobservations introduced negative biases in the emissions inversion. This is due to the observation error in our synthetic data is proportional to the observed column magnitude, thus pixels with lower  $\text{NO}_2$  column concentrations have larger impacts in the inverse error weighted superobing calculation. As a result, the superobservation  $\text{NO}_2$  column tended to be biased low compared to the raw data. Therefore, we simply use spatial data thinning of the synthetic TEMPO and TROPOMI  $\text{NO}_2$  observations by retaining the observations that are closest to the centroid of the model grid point within which the horizontal location of the observations falls. This means that we have only one observation for each model grid cell. In terms of data screening for quality control, we omit synthetic TEMPO and TROPOMI  $\text{NO}_2$  data that have  $\text{SZA} > 70^\circ$  or maximum cloud fraction  $> 0.5$  (as given by NR). Figure S8 in Supporting Information S1 shows the spatial distribution of averaged synthetic TEMPO and TROPOMI  $\text{NO}_2$  SCD, and the number of measurements during 01 to 06 April. Over CONUS, TEMPO data show better spatial coverage and provides 4–5 times more observations than TROPOMI. However, regions in the northeast, northwest, and Texas have comparatively few satellite observations because of cloudy skies (Figures S8c and S8d in Supporting Information S1). The total number of observations assimilated in each ER are also provided in Table 3.

### 3. Results

Before examining the performance of the  $\text{NO}_x$  emissions inversion from individual ERs, we first perform observational space diagnostics. The left column of Figure S9 in Supporting Information S1 shows the time series of prior and posterior RMSEs, and they are defined as follows:

$$\text{prior RMSE} = \sqrt{\frac{1}{N} \sum_{i=1}^N (y_i^o - \bar{y}_i^f)^2} \quad (10)$$

$$\text{posterior RMSE} = \sqrt{\frac{1}{N} \sum_{i=1}^N (y_i^o - \bar{y}_i^a)^2} \quad (11)$$

where  $N$  is the total number of the assimilated  $\text{NO}_2$  column measurements,  $y_i^o$  is an assimilated observation, and  $\bar{y}_i^f$  and  $\bar{y}_i^a$  are the ensemble mean of expected and analysis observations. The posterior RMSE is lower than the prior RMSE for all ERs at the cycle time, indicating that the DA system performed properly in terms of assimilating TEMPO and TROPOMI data.

We next evaluate the ratio of the prior RMSE to the prior total spread, which is defined as follows:

$$\text{prior total spread} = \sqrt{\frac{1}{N} \sum_{i=1}^N (\sigma_i^o + \sigma_i^f)^2} \quad (12)$$

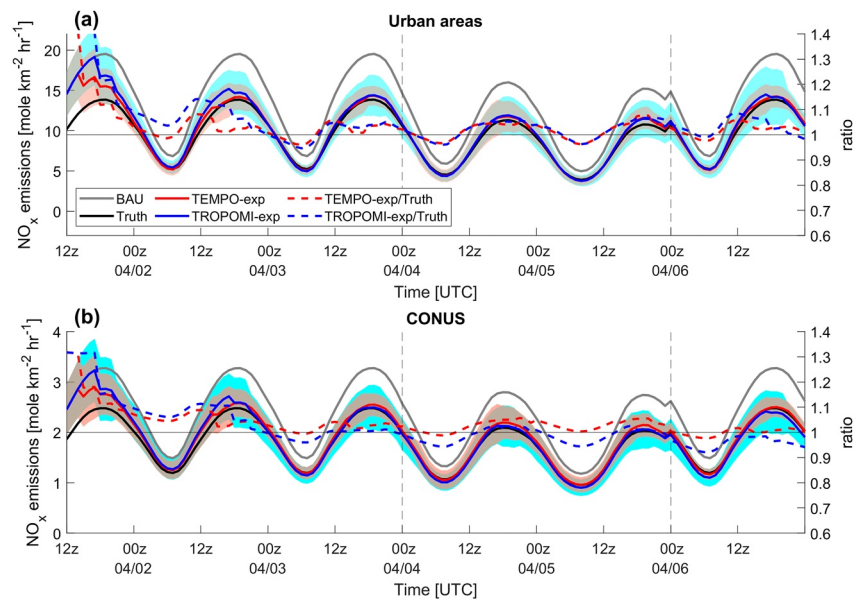
where  $\sigma_i^o$  is the observation error variance and  $\sigma_i^f$  is the prior ensemble variance in the observation space. The ratio is a chi-square variate (Ménard & Chang, 2000) and can be used to verify whether the DA system is appropriately tuned (Raeder et al., 2012). If the observation error, model ensemble spread, and covariance inflation are adequately specified, the prior RMSE and total spread should be balanced. We multiply the observation error variance by 0.70 to balance the prior RMSE and total spread for all our ERs. The right column of Figure S9 in Supporting Information S1 shows the time series of the ratio, and the averaged ratio of each experiment ranging from 0.88 to 1.07, suggesting that the prior RMSE and total spread are comparable, and the DA system is well balanced.

#### 3.1. Advantage of TEMPO Over TROPOMI Data for $\text{NO}_x$ Emissions Inversion

We quantify the impact of assimilating TEMPO observations on the posterior emissions by comparing the posterior  $\text{NO}_x$  emissions estimated from TEMPO-exp (GEO) and TROPOMI-exp (LEO) to the true COVID-19  $\text{NO}_x$  emissions.

Figure 6 shows the time series of the average posterior  $\text{NO}_x$  emissions over 29 cities and the CONUS. The names and geographic information of 29 urbans are provided in Table S4 in Supporting Information S1. TEMPO-exp recovers the true  $\text{NO}_x$  emissions within a single day, particularly in metropolitan areas. At 0 UTC on 2 April, the ratio of posterior  $\text{NO}_x$  emissions to true  $\text{NO}_x$  emissions (shown by the red dashed curve in Figure 6) for urban areas approaches one, and the error in  $\text{NO}_x$  emissions over the CONUS is significantly reduced. In contrast, TROPOMI-exp takes two to 3 days to retrieve the true  $\text{NO}_x$  emissions over urban regions. Similar to the result in urban areas, the posterior  $\text{NO}_x$  emissions in sub-urban/rural regions (e.g., bottom 80%  $\text{NO}_x$  emissions) from TEMPO-exp are substantially closer to the true  $\text{NO}_x$  emissions, whereas TROPOMI-exp emissions exhibit a larger persistent negative bias (5%–10%) in comparison to the true  $\text{NO}_x$  emissions (not shown). Additionally, TEMPO-exp is better at tracking temporal fluctuations of true  $\text{NO}_x$  emissions, including the weekend drop and subsequent weekday rebound across the CONUS. On the other hand, TROPOMI-exp estimated  $\text{NO}_x$  emissions that are biased low by ~5% during the weekend (4 and 5 April) and continued cycling during the following weekdays does not eliminate that bias.

Figure 7 illustrates the spatial distribution of the difference between the ER emissions and the true  $\text{NO}_x$  emissions, along with the corresponding domain-averaged coefficient of determination ( $R^2$ ), normalized mean bias (NMB), and RMSE. At the beginning of the study period, positive biases are present since TEMPO-exp and TROPOMI-exp were initially driven by BAU emissions (Figure 7a), with metropolitan regions having the highest degrees



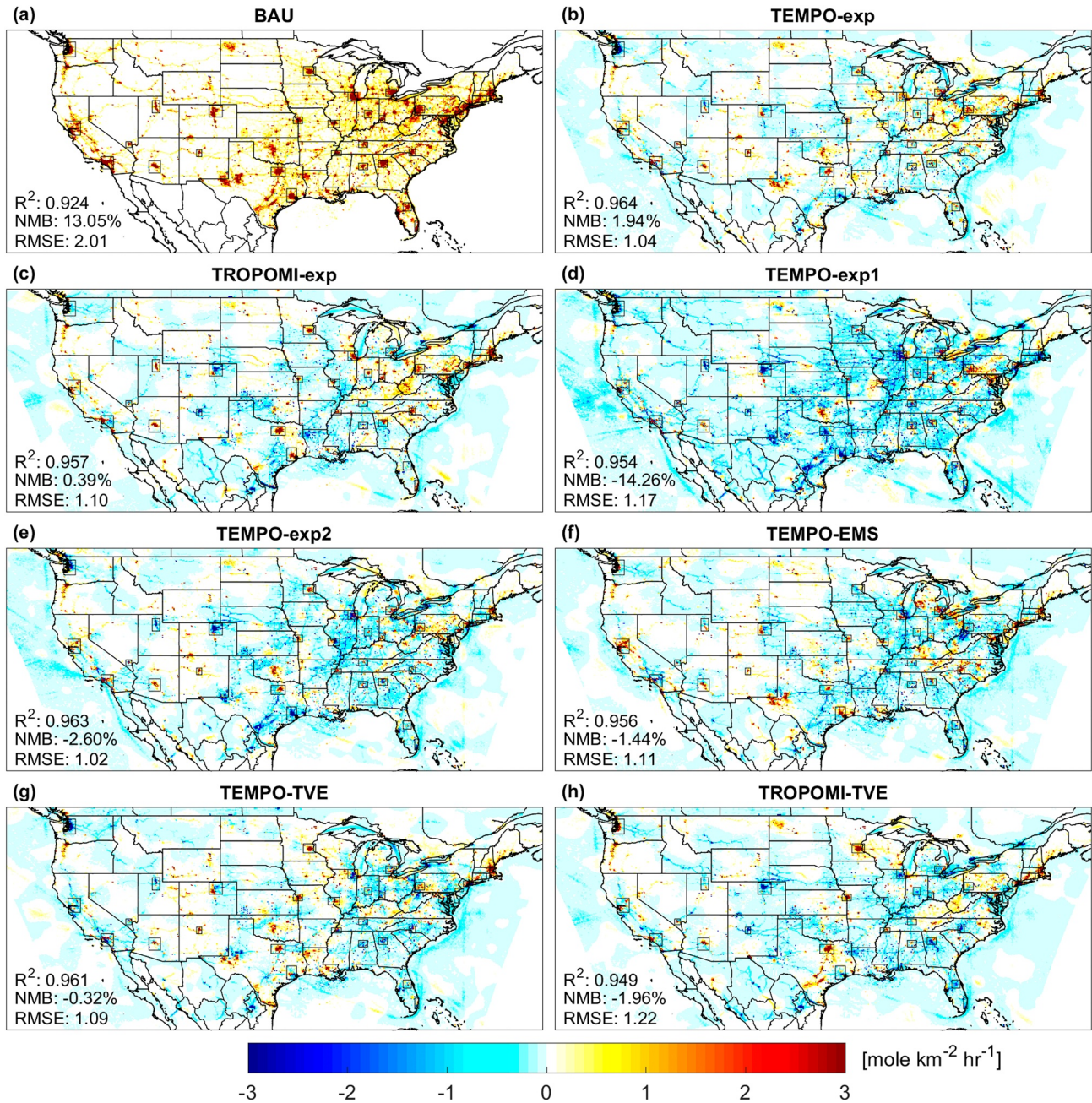
**Figure 6.** Time series of the TEMPO-exp and TROPOMI-exp top-down NO<sub>x</sub> emissions (left axis) and their relative proportion to the true NO<sub>x</sub> emissions (right axis). (a) Averaged NO<sub>x</sub> emissions over 29 urban areas and (b) averaged NO<sub>x</sub> emissions over the CONUS. The red and blue shaded areas represent 95% error bars from TEMPO-exp and TROPOMI-exp respectively.

of inaccuracy. With cycling, TEMPO-exp (Figure 7b) significantly reduces the emission biases in most regions, with the exception of a few metropolitan areas (e.g., Seattle, Denver, and Dallas) and western Texas, which still have somewhat higher errors. Conversely, the spatial distribution of TROPOMI-exp posterior emissions are not well constrained (Figure 7c). Several cities (e.g., Pittsburgh, Chicago, and Minneapolis) and areas from the central to western US have higher error and lower domain-averaged skill metrics compared to TEMPO-exp.

We further examine the variation of performance of the NO<sub>x</sub> emissions inversion from regional to urban scales. Figure 8 shows the  $R^2$ , RMSE, and NMB of posterior NO<sub>x</sub> emissions in each of the US EPA's regulatory regions (RRs) across the CONUS. A map defining the RRs is provided in Figure S10 in Supporting Information S1. TEMPO-exp outperforms TROPOMI-exp by 1.5%–17.1% and 12.5%–41.5% in terms of  $R^2$  and RMSE for all RRs, respectively. Some RRs, such as New England (RR1), show significant improvement. For NMB, TEMPO-exp also surpasses TROPOMI-exp in most RRs, except for RR2 (New York), RR5 (Great Lakes), and RR8 (Mountains and Plains) which have slightly larger biases. The performance of the emissions inversion across RR1-3 shows greater error in terms of RMSE and NMB compared to the other RRs in the TEMPO-exp, which may be linked to the relative sparsity of observations over the northeastern CONUS due to cloud cover, as illustrated in Figure S8c in Supporting Information S1. The small DA window in TEMPO-exp further limits the available observations across these regions.

At the urban scale, Figure 9a shows a Taylor diagram of the posterior NO<sub>x</sub> emissions from the TEMPO-exp and TROPOMI-exp from 29 cities across the CONUS. TEMPO-exp recovers the true NO<sub>x</sub> emissions more accurately than TROPOMI-exp. The TEMPO-exp (red dots) show better agreement with the true emissions than the TROPOMI-exp emissions (blue dots) in 23 cities. However, Seattle is an exception, where the TEMPO-exp emissions deviate significantly from the true emissions because they were constrained by relatively distant and less representative observations during the first 4 days of cycling due to broader spatiotemporal coverage of TEMPO data. The posterior NO<sub>x</sub> emissions over Seattle eventually converge to the true emissions on 5–6 April, when more direct observations were available to constrain emissions (not shown).

In summary, we have demonstrated that TEMPO-exp recovers the true NO<sub>x</sub> emissions at least twice as fast as TROPOMI-exp and better resolves temporal fluctuations of the true NO<sub>x</sub> emissions. TEMPO-exp also provides more accurate NO<sub>x</sub> emissions compared to TROPOMI-exp from continental to urban scales based on the skill metrics discussed previously. We assume that the high temporal coverage of TEMPO data, which allows constraining emissions more frequently, is the major contributor to the improvement of emissions inversion and



**Figure 7.** Six-day average spatial distributions of the differences between the various experimental run emission estimates and the true  $\text{NO}_x$  emissions. (a) BAU, (b) TEMPO-exp, (c) TROPOMI-exp, (d) TEMPO-exp1, (e) TEMPO-exp2, (f) TEMPO-EMS, (g) TEMPO-TVE, and (h) TROPOMI-TVE. The black boxes highlight 29 urban regions over the CONUS. The domain averaged  $R^2$ , NMB, and RMSE are also listed.

can be further explored in the future study. The benefit of frequent emissions updating is also highlighted in Ma et al. (2019) when assimilating surface measurements.

### 3.2. Impact of the DA Window on Emissions Inversion

#### 3.2.1. 1.5-Hour DA Window for All Cycles

Geostationary satellites provide hourly daytime  $\text{NO}_2$  column observations which can assist in constraining the strength and spatial-temporal variability of  $\text{NO}_x$  emissions. To utilize the most observations per day, we extend the DA window from 0- to  $\pm 1.5$ -hr Figure 10 compares the temporal evolution of the ratio of the posterior emissions



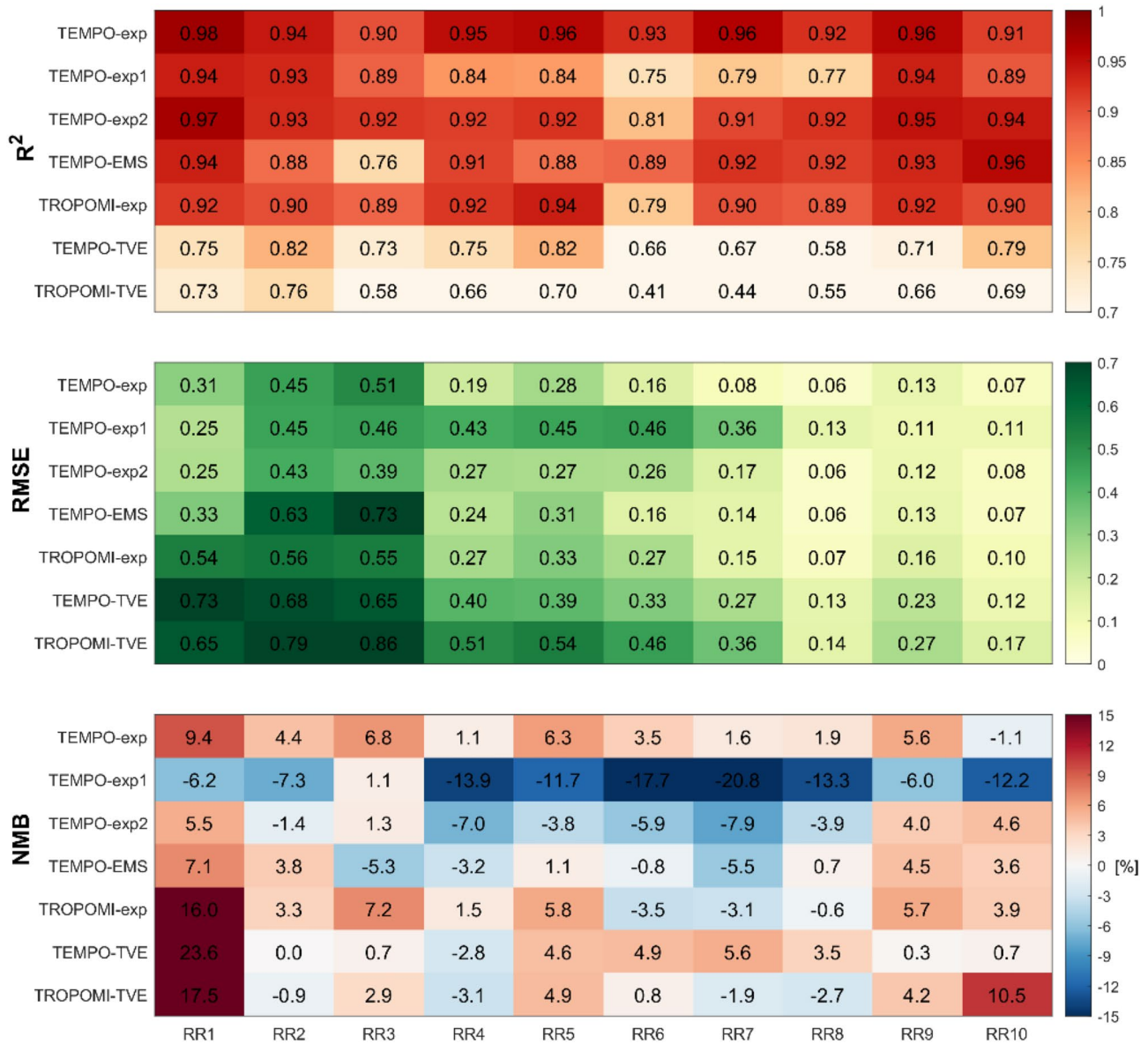
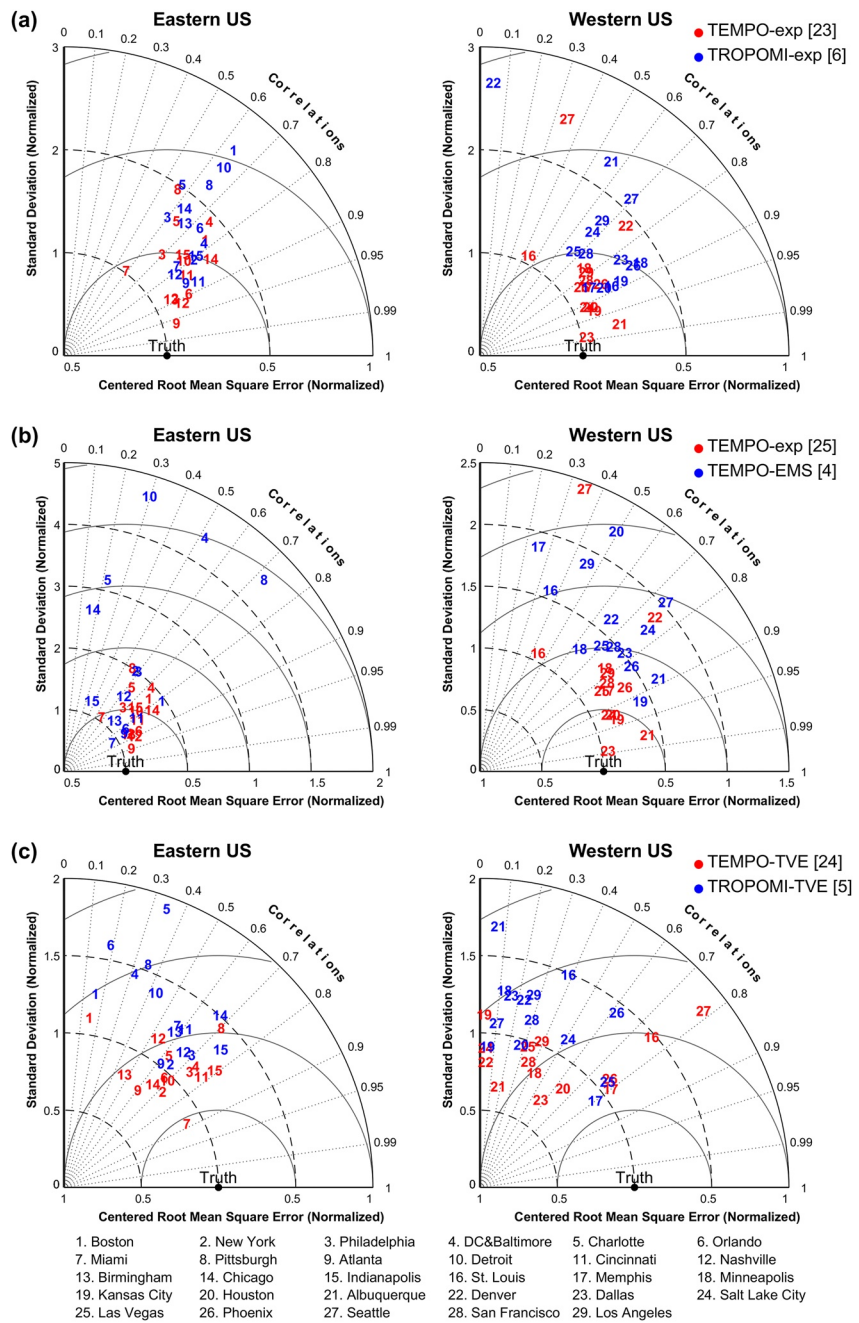


Figure 8. The 6-day average skill scores in EPA's 10 regulatory regions for the experimental runs. The  $R^2$ , RMSE, and NMB are displayed from top to bottom. The deeper the color, the greater the value for each metric.

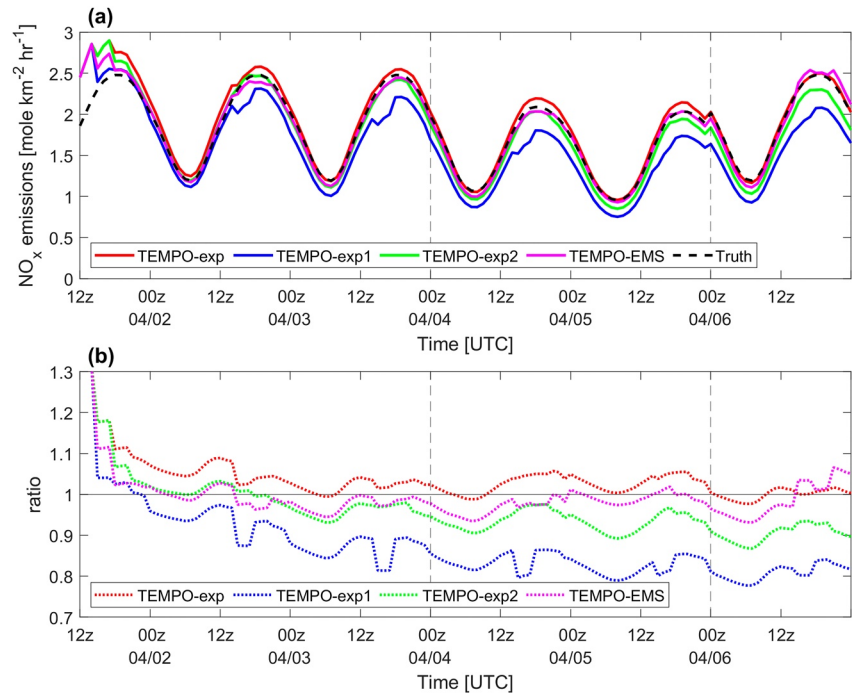
to the true emissions over the CONUS from all TEMPO ERs. The  $\text{NO}_x$  emissions estimated by TEMPO-exp1, which uses a  $\pm 1.5$ -hr DA window for all cycles, are biased low by 15%–20% compared to true  $\text{NO}_x$  emissions. The  $\text{NO}_x$  emissions ratios consistently show a sharp decrease at 15 UTC with a rebound at 18 UTC, and the emissions ratios also show a small drop at 0 UTC. Spatially, TEMPO-exp1 substantially underestimates true  $\text{NO}_x$  emissions at most locations throughout the domain (Figure 7d), particularly over the central US. As compared to TEMPO-exp, TEMPO-exp1 degrades  $R^2$  by 4%–12% and RMSE by 40%–300% for most RRs, and six RRs are biased low with NMB ranging from  $-11.7\%$  to  $-20.8\%$  (Figure 8).

To better understand the ER posterior emissions skill drops at 15 and 00 UTC, we examine the emissions adjustment results on 3 and 4 April to determine how the biases emerge. At first this result seems counter-intuitive because we expect to increase the inversion skill by assimilating a greater number of TEMPO observations. We start by computing the relative innovation in the observational space, defined as  $(y_i^o - \bar{y}_i^f) / \bar{y}_i^f$ , to quantify the impact of observations. For the 15 and 0 UTC cycles, the hours of observations being assimilated are from 14–16 to 23–01 UTC, respectively. Figure 11 depicts the relative innovation and the number of observations as



**Figure 9.** Taylor diagrams for posterior  $\text{NO}_x$  emissions from the different ERs. The cities in eastern and western US are roughly separated by the Mississippi river. (a) TEMPO-exp versus TROPOMI-exp, (b) TEMPO-exp versus TEMPO-EMS, and (c) TEMPO-TVE versus TROPOMI-TVE. The number in parentheses to the right of the experiment name shows the number of urban areas that performed better for the referenced experiment compared to the companion experiment.

function of the observation hour. First, the number of available observations varies during the transition hours. This happens because the TEMPO instrument only provides observations when sunlight is available, and the hours away from sunrise/sunset (e.g., 16 and 23 UTC) provide more satellite observations (Figures 11a and 11b), which may have a greater impact on emissions adjustments. Second, in the EnKF analysis step, the model background forecast at the cycle time hours is compared to measurement from the different observation hours, which can introduce representative error into the emissions inversion. For example, the model background forecast at 15 UTC is compared to observations from 14 to 16 UTC. Figure 11a shows the relative innovation distribution from observations hour 15 UTC is normally distributed and unbiased, with a mean close to zero, indicating no



**Figure 10.** (a) Time series of ensemble mean NO<sub>x</sub> emissions from the TEMPO ERs and the true NO<sub>x</sub> emissions averaged across the CONUS and (b) time series of the ratio between posterior NO<sub>x</sub> emissions from the TEMPO ERs to the true NO<sub>x</sub> emissions.

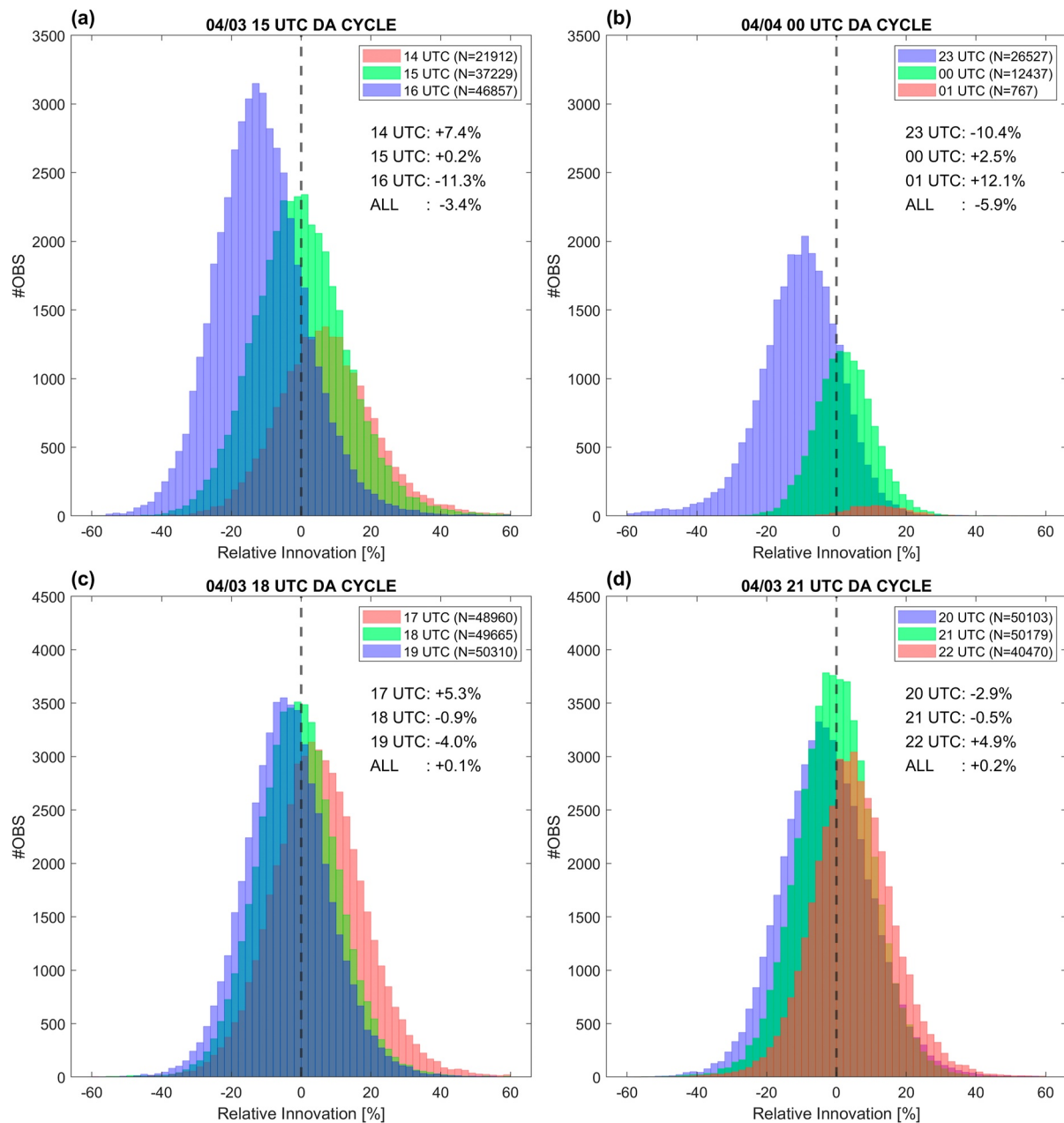
representative error. However, there are substantial positive and negative biases for the innovation from non-cycle time observation hour, suggesting increased representativeness error. The level of bias for observation hour 16 UTC (−11.3%) is substantially higher than that for observations from 14 UTC (+7.4%). The same pattern can also be seen for the 00 UTC cycle (Figure 11b), where greater negative biases are associated with the observation from 23 UTC.

The presence of the representative error is originated from the diurnal variation of the synthetic TEMPO NO<sub>2</sub> measurement which is driven by the NO<sub>2</sub> vertical profile changes as shown in Figures S11a and S11c in Supporting Information S1. The diurnal variation of the SW vertical profile may have an additional impact on the hourly variation of NO<sub>2</sub> column since there are greater SWs appearing during the transition hours (Figure S2 in Supporting Information S1).

Given the influence of representative error and the uneven number of observations provided from the different observation hours, the emissions inversion at the 15 and 00 UTC cycles are dominated by observations from hours 16 and 23 UTC (i.e., more observation and negative innovation), respectively, resulting in a decrease in prior NO<sub>x</sub> emissions and introducing a systematically negative bias. The same pattern occurs in the other days of the study period.

### 3.2.2. 1.5-Hour DA Window for Midday Cycles

To mitigate the effect of representative error during the transition cycles, we propose a new experiment (TEMPO-exp2) that uses the ±1.5-hr DA window only at 18 and 21 UTC while using the 0-hr DA window for the remaining cycles. By decreasing the size of the DA window during the transition cycles, the domain averaged NMB in posterior NO<sub>x</sub> emissions is reduced by 11.6% in TEMPO-exp2 (Figure 7d) compared to TEMPO-exp1, and the R<sup>2</sup> and RSME are also improved. Nevertheless, as shown in Figure 10b, the posterior NO<sub>x</sub> emissions from TEMPO-exp2 are biased low by 4%–10% across the CONUS during the study period. On a regional scale, the posterior NO<sub>x</sub> emissions in RR4 through RR8 (southeast to central CONUS) remain biased low by 3.9%–7.9%, despite having a higher R<sup>2</sup> and lower RMSE than TEMPO-exp1 (Figure 8). Consequently, we examined the relative innovations from the 18 and 21 UTC cycles on April 3 (Figures 10c and 10d) and find smaller representative



**Figure 11.** Distribution of relative innovation (%) for DA cycle at (a) 15 UTC on 04/03 from TEMPO-exp1, (b) 00 UTC on 04/04 from TEMPO-exp1, (c) 18 UTC on 04/03 from TEMPO-exp2, and (d) 21 UTC on 04/03 from TEMPO-exp2. The total number of assimilated observations ( $N$ ) is also provided.

error associated with the assimilated observations. The relative innovations only change by  $\pm 5\%$  compared to a variation of  $\pm 10\%$  for observation hours in the 15 and 00 UTC cycles.

Next, we compare the WRF-Chem prior and posterior TEMPO  $\text{NO}_2$  SCDs to the observation during the 17 to 22 UTC observation hours, and the results are summarized in Table 4. We find that the posterior  $\text{NO}_2$  SCD for the observation hours that correspond to the cycle time (i.e., 18 and 21 UTC) display the best agreement with measurements, with the lowest RMSE, NMB, and highest  $R^2$ . This result is expected because it effectively reproduces the results in TEMPO-exp. However, the posterior  $\text{NO}_2$  SCD from the daytime side of the DA window (i.e., observation hour 19 and 20 for the 18 and 21 UTC cycle, respectively) tend to exhibit greater goodness of fit with the observations, especially in terms of RMSE and regression slope as shown in Table 4. In contrast, assimilation of the nighttime side observations (observations hours 17 and 22 UTC) enhances the bias. For example, the

**Table 4**

The RMSE (molec cm<sup>-2</sup>), NMB (%), R<sup>2</sup>, and the Regression Slopes for the Prior and Posterior Expected NO<sub>2</sub> SCDs, and the Ensemble Correlations (R) Between the Expected Prior NO<sub>2</sub> SCD Ensemble and NO<sub>x</sub> Emission Ensemble

Hours		18 UTC DA cycle			21 UTC DA cycle		
		17	18	19	20	21	22
A Priori	#OBS	287,201	286,120	284,426	285,938	286,290	236,628
	RMSE	2.34 × 10 <sup>14</sup>	2.05 × 10 <sup>14</sup>	2.36 × 10 <sup>14</sup>	2.31 × 10 <sup>14</sup>	1.86 × 10 <sup>14</sup>	2.04 × 10 <sup>14</sup>
	NMB	-2.21%	3.45%	6.87%	6.25%	2.59%	-3.28%
	R <sup>2</sup>	0.897	0.928	0.899	0.905	0.939	0.928
	Slope	1.016	0.971	0.943	0.922	0.968	1.042
A Posteriori	RMSE	2.07 × 10 <sup>14</sup>	1.46 × 10 <sup>14</sup>	1.70 × 10 <sup>14</sup>	1.55 × 10 <sup>14</sup>	1.38 × 10 <sup>14</sup>	1.64 × 10 <sup>14</sup>
	NMB	-5.32%	-0.34%	2.49%	4.23%	1.62%	-2.58%
	R <sup>2</sup>	0.929	0.962	0.941	0.95	0.964	0.957
	Slope	1.058	1.013	0.988	0.962	1	1.048
	Ensemble R <sup>a</sup>	0.173	0.185	0.179	0.104	0.090	0.084

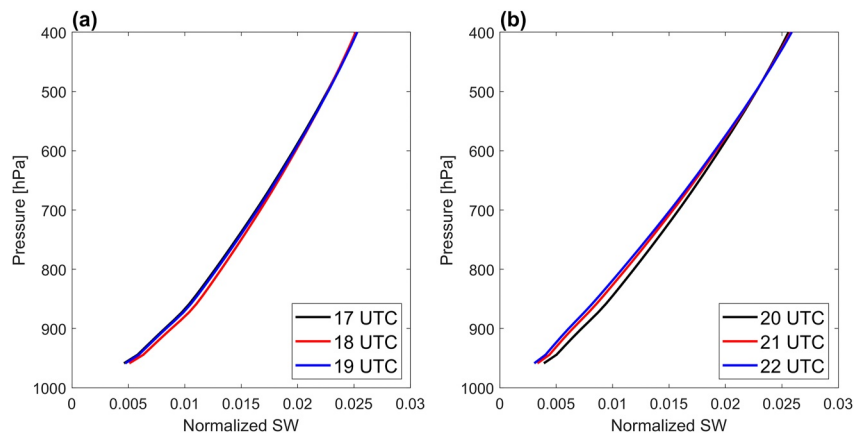
<sup>a</sup>The average ensemble correlation (R) between any two observations hours for hours ranging from 17 to 22 UTC are significantly different at the 95% level based on the student's *t*-test.

NMB deteriorates from -2.21% to -5.32%, and the regression slope degrades from 1.016 to 1.058 for 17 UTC observation hour after DA.

We further investigate the ensemble correlation between expected NO<sub>2</sub> SCD and NO<sub>x</sub> emissions since the observations update the emissions via their correlations in the EnKF (J. L. Anderson, 2003). The ensemble correlations for observations hours between 17 and 22 UTC are also listed in Table 4, and the average ensemble correlation is defined as  $\frac{1}{N} \sum_{k=1}^N \text{corr}(\text{SCD}_k, E_{i,j}^t)$  where *N* is the total number of NO<sub>2</sub> column observations for each observation hour, SCD<sub>*k*</sub> is the *k*th forecasted NO<sub>2</sub> SCD ensemble, and E<sub>*i,j*</sub><sup>*t*</sup> is the prior emissions rate ensemble at grid (*i,j*) which is spatially closest to SCD<sub>*k*</sub> at cycle time *t*. Table 4 shows that the expected NO<sub>2</sub> SCDs from observation hours 19 and 20 UTC are better correlated with the prior NO<sub>x</sub> emissions rate than those from observation hours 17 and 22 UTC. This suggests that observations from hours 19 and 20 UTC contribute more toward constraining the NO<sub>x</sub> emissions. The increased correlations for daytime side observations can be related to the diurnal variation of the SW profiles. Figure 12 shows average normalized SW profiles (i.e., normalized by the vertical summation of SW). For the 18 UTC cycle, the SW profile from observation hour 18 UTC has the greatest sensitivity to the lower troposphere, followed by 19 and 17 UTC. For the 21 UTC cycle, the observation hour 20 UTC SW profile (i.e., daytime side) has the greatest sensitivity to the lower troposphere compared to other observation hour. We hypothesize that the greater SW sensitivity to the lower troposphere leads to greater correlations with the simulated near-surface NO<sub>2</sub> concentrations and NO<sub>x</sub> emissions. As a result, observations from 19 to 20 UTC can have larger impacts on the emissions inversion than those from observation hours 17 and 22 UTC. This process also enhances the representative error (i.e., negative innovation) from observation hours 19 and 20 UTC, resulting in a negative bias in the emissions inversion.

In this section, we showed that using a ±1.5-hr DA window to assimilate synthetic TEMPO observations introduces biases into the posterior NO<sub>x</sub> emissions due to the unrepresentativeness of observations from the non-cycle time observations hours. The occurrence of representative error is primarily due to the hourly/diurnal variations of synthetic TEMPO NO<sub>2</sub> SCD, and this problem is more pronounced in the transition time cycles (e.g., 15 and 0 UTC) owing to larger variation in the synthetic NO<sub>2</sub> column and volume of observation. For the midday cycles, observations from hours that are closest to midday (e.g., hours 19 and 20 UTC for the 18 and 21 UTC cycles, respectively) show greater impact because their associated SWs have greater sensitivity to the lower troposphere, which leads to systematic biases in the posterior NO<sub>x</sub> emissions.

We would like to note that despite the NR simulated NO<sub>2</sub> properly representing the surface NO<sub>2</sub> diurnal fluctuation, and the hourly variation of TEMPO NO<sub>2</sub> SCD proxy data being similar to our synthetic product (Figures S11a and S11b in Supporting Information S1), the diurnal pattern in the actual TEMPO NO<sub>2</sub> SCD data are still



**Figure 12.** Average normalized SW profiles for observation hours (a) 17–19 and (b) 20–22 UTC. The lower part of the SW profile within the troposphere (bottom 35 layers) are plotted.

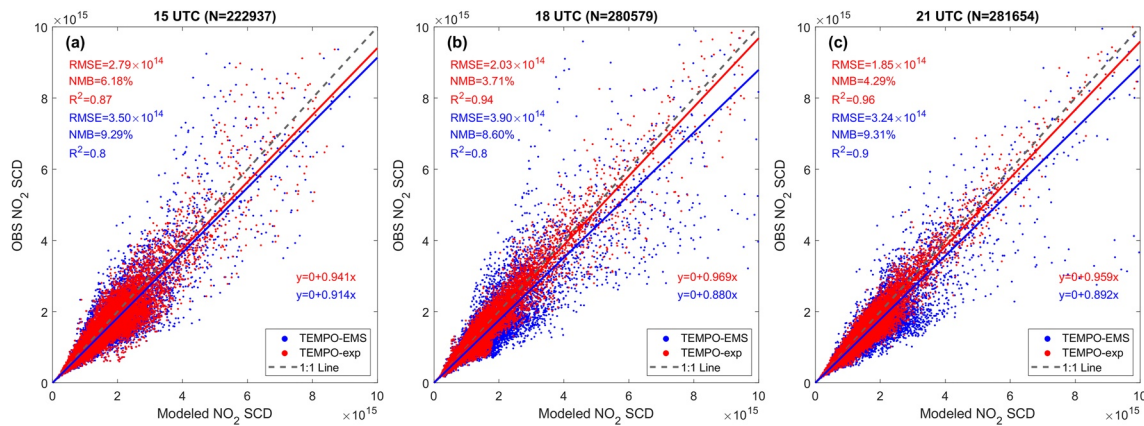
unknown. Therefore, the representative error issue might exist if future operational TEMPO  $\text{NO}_2$  SCD data display a similar diurnal pattern as our synthetic data. Additionally, the conclusion drawn from Section 3.2 may only be applicable in the context of using EnKF algorithm because the definition of the DA window differs in other approaches such as 4D-VAR, which may not see the representative issue.

### 3.3. Impact of Adjusting $\text{NO}_x$ Concentration on $\text{NO}_x$ Emissions Inversion

The assimilation of remotely sensed  $\text{NO}_2$  observations can optimize both  $\text{NO}_x$  concentrations and emissions simultaneously, whereas many previous inverse modeling studies have focused only on adjusting emissions without updating related species concentrations (e.g., Qu et al., 2017; Sourì et al., 2020). However, Liu et al. (2017) pointed out that solely updating the emissions rate can lead to large inaccuracies in the estimated emissions. To investigate the influence of updating concentrations on emissions inversion, we conduct a TEMPO experiment (TEMPO-EMS) that only allows observations to update  $\text{NO}_x$  emissions, and the estimated  $\text{NO}_x$  emissions are then compared with those of TEMPO-exp.

Figure 10 depicts the temporal evolution of the TEMPO-EMS posterior  $\text{NO}_x$  emissions and their proportion to true  $\text{NO}_x$  emissions. During the first 2 days (01–02 April) of the study period, the TEMPO-EMS  $\text{NO}_x$  emissions outperform the TEMPO-exp results as it recovers the true  $\text{NO}_x$  emissions more quickly. However, during the rest of the study period (03–06 April), the TEMPO-EMS emissions have a low bias of 2%–5%, as shown in Figure 10b. The spatial error pattern of TEMPO-EMS emissions (Figure 7f) is similar to that of TEMPO-exp (Figure 7b), although the domain averaged  $R^2$ , RMSE, and NMB are slightly worse. On a regional scale, the  $R^2$  and RMSE between TEMPO-EMS and TEMPO-exp are comparable throughout most RRs (Figure 8). However, on the urban scale, the posterior  $\text{NO}_x$  emissions from TEMPO-EMS are substantially less accurate, as demonstrated in the Taylor diagram in Figure 9b. Most cities in the western part of the CONUS, as well as several eastern cities (e.g., Chicago, Pittsburgh, Washington, D.C., Detroit, and Charlotte) show much larger deviations from true  $\text{NO}_x$  emissions compared to TEMPO-exp. This suggests that constraining only the emissions may introduce substantial error into the top-down emissions, especially over polluted area.

Solely adjusting emissions implicitly assumes that the discrepancy between the forecasted and observed  $\text{NO}_2$  columns is only related to emissions. However, errors in the model transport, physics, and chemistry (i.e., model error/biases) can contribute to errors in the predicted  $\text{NO}_2$  columns. When model error exists such as our ER configuration (Section 2.3), adjusting emissions alone is insufficient to address the errors in the simulated  $\text{NO}_2$  concentration. Model error can accumulate and propagate forward in time, resulting in over-adjustments in emissions to compensate for this source of error. Figure 13 displays the errors of the expected TEMPO  $\text{NO}_2$  SCD at 15, 18, and 21 UTC. At these cycle hours, TEMPO-EMS shows larger errors (RMSE and NMB) and a worse  $R^2$  than TEMPO-exp in the predicted  $\text{NO}_2$  column concentrations. The RMSE in TEMPO-EMS remains at a high level over each DA cycle and is 25%–90% greater than in TEMPO-exp, indicating that adjusting  $\text{NO}_x$  emissions alone cannot effectively improve the  $\text{NO}_2$  concentration forecast.



**Figure 13.** Expected and observed TEMPO NO<sub>2</sub> SCD at (a) 15, (b) 18, and (c) 21 UTC. The red and blue dots are the values from TEMPO-exp and TEMPO-EMS, respectively. The linear regression lines, RMSE, NMB, and R<sup>2</sup> are also presented.

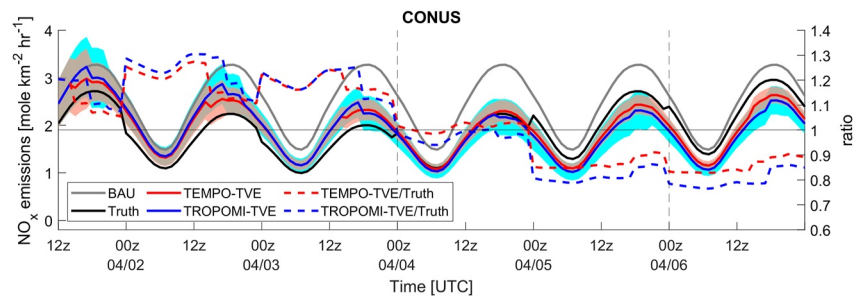
These results highlight the importance of adjusting both NO<sub>x</sub> concentration and emissions simultaneously for optimal NO<sub>x</sub> emissions inversion in the presence of model errors, which is likely to always be the case.

### 3.4. Reducing Time-Varying Uncertainties in NO<sub>x</sub> Emissions

For the cases addressed in previous sections, only three emission profiles (weekday, Saturday, and Sunday) need to be recovered, and they are relatively similar to each other, making the emissions inversion somewhat simple. In reality, the NO<sub>x</sub> emissions may have daily variations in the diurnal pattern due to changes in energy consumption or traffic flow caused by holidays, weather, or other short-term events. Therefore, we conduct an experiment (TEMPO-TVE and TROPOMI-TVE) where we introduce daily temporal variability to the true NO<sub>x</sub> emissions and making emissions inversion more challenging. Specifically, the perturbed true NO<sub>x</sub> emissions follow a V-shape pattern of a decreasing daily mean NO<sub>x</sub> emission for the first 3 days and an increasing mean emission for the subsequent days, as shown in Figure 14. With this perturbation scheme, polluted regions, such as cities, exhibit higher daily variation than rural areas, as illustrated in Figure S7a in Supporting Information S1.

Figure 14 shows the posterior NO<sub>x</sub> emissions time series over CONUS for the TEMPO-TVE and TROPOMI-TVE experiments. In the scenario with time-varying true NO<sub>x</sub> emissions, both TEMPO-TVE and TROPOMI-TVE struggle to recover the true emissions, but TEMPO-TVE performs better in tracking the daily variability of NO<sub>x</sub> emissions throughout the simulation. The spatial distribution of the discrepancy between the posterior and true NO<sub>x</sub> emissions are relatively small in TEMPO-TVE (see Figure 7g). Only cities in the central US (e.g., Dallas, Kansas City) and Seattle show larger errors. In contrast, the estimated NO<sub>x</sub> emissions from TROPOMI-TVE display greater errors in the central US, with emissions outside urban regions and on highways biased low (Figure 7f), leading to larger RMSE and NMB compared to TEMPO-TVE. On the regional scale, the TEMPO-TVE results display higher R<sup>2</sup> (2%–60%) for all RRs compared to the TROPOMI-TVE results as shown in Figure 8, especially for the regions over the central US (i.e., RR6 and RR7). The RMSE for most RRs, except RR1, are also outperformed in TEMPO-TVE with an improvement of 10%–40% over TROPOMI-TVE. Similarly, the NMB shows lower errors in most of RRs compared to TEMPO-TVE (Figure 8). In urban regions, the TEMPO-TVE estimated NO<sub>x</sub> emissions shows better agreement with the true NO<sub>x</sub> emissions over 24 city areas, while posterior NO<sub>x</sub> emissions from TROPOMI-TVE display larger deviations from the truth (Figure 9c). Therefore, we conclude that assimilating geostationary observations can better constrain the daily variability of NO<sub>x</sub> emissions rate.

Additionally, TEMPO-TVE only uses 30% of the observations each day. We anticipate that the performance of emissions inversion could be further improved if more synthetic TEMPO data are assimilated by increasing the frequency of emissions updating (e.g., 1 hr DA cycle) when the true emissions consist of a more complex spatial-temporal variation pattern.



**Figure 14.** Time series of the averaged posterior  $\text{NO}_x$  emissions over the CONUS for TEMPO-TVE and TROPOMI-TVE (left axis) and their relative proportion to the perturbed true  $\text{NO}_x$  emissions (right axis). The red and blue shaded areas are 95% error bars from TEMPO-TVE and TROPOMI-TVE, respectively.

#### 4. Discussion

As mentioned in Section 2.2, the synthetic TEMPO  $\text{NO}_2$  observation error is estimated based on the total column from the TEMPO proxy data. However, this approach may underestimate the observation error, as demonstrated by the  $\gamma$  value calculated from the real TROPOMI  $\text{NO}_2$  total SCD data being substantially smaller than the one based on the tropospheric SCD (Figure S12 in Supporting Information S1). Underestimating the observation error could lead to a more accurate synthetic observation and overestimate the influence of synthetic observation on the state variables during DA. To investigate the impact of the synthetic observation errors on the emissions inversion, we assumed that the error characteristics of synthetic TEMPO data follow real TROPOMI data and regenerated the synthetic TEMPO  $\text{NO}_2$  data. These data are used to constrain  $\text{NO}_x$  emissions in an additional ER with a 0-hr DA window (TEMPO-exp3). It is noted that the  $\gamma$  value upper limit (35%) is still applied to prevent bias in the synthetic TEMPO  $\text{NO}_2$  data.

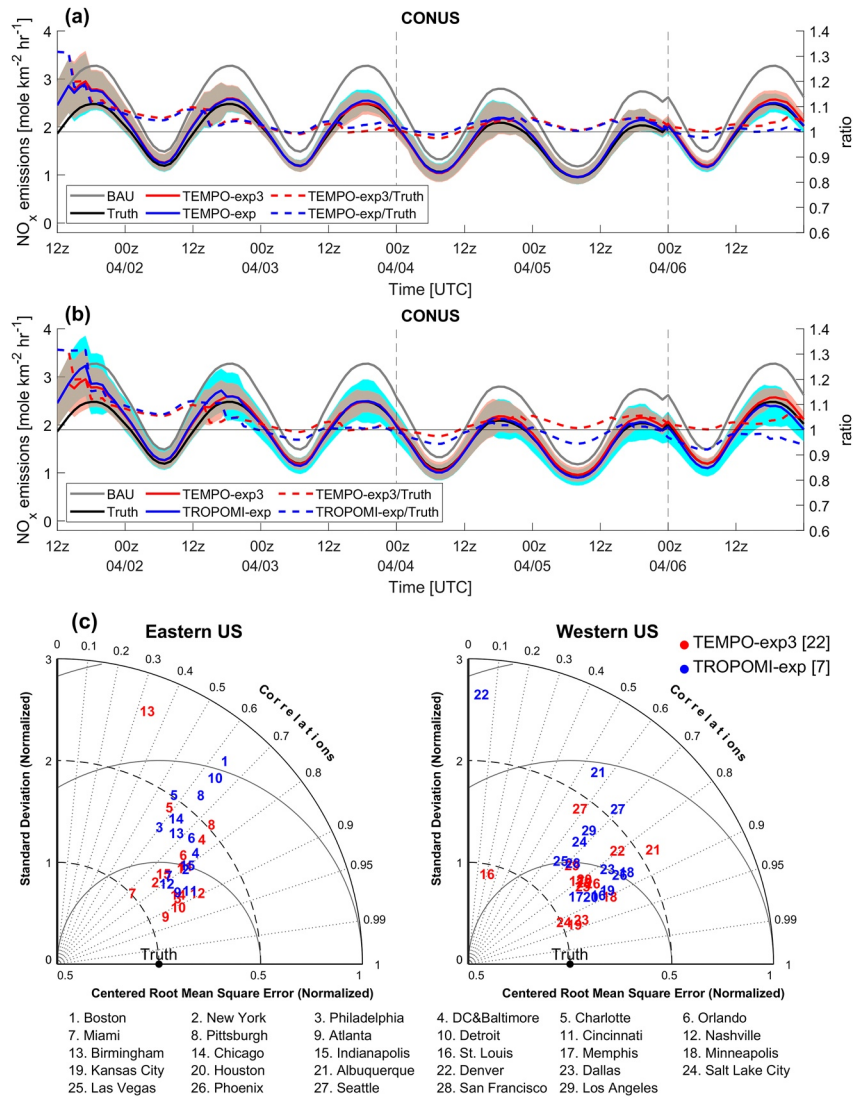
We first compare the average CONUS posterior  $\text{NO}_x$  emission time series from TEMPO-exp3 to TEMPO-exp, as shown in Figure 15a. In general, the posterior emissions from TEMPO-exp3 and TEMPO-exp are similar, but the skill of the emissions inversion in TEMPO-exp3 degrades slightly. The RMSE and NMB increased from 1.04% to 1.09% and 1.94% to 3.35%, respectively, when compared to TEMPO-exp, especially on the last day of simulation (6 April), which displayed a large discrepancy between these two ERs. This is owing to the larger observation error and the less accurate synthetic observations assimilated in TEMPO-exp3. Next, we compare TEMPO-exp3 to TROPOMI-exp (Figure 15b), and TEMPO-exp3 still outperforms TROPOMI-exp in terms of emissions inversion performance, as TEMPO-exp3 can resolve true  $\text{NO}_x$  emissions faster and better. The TEMPO-exp3 also shows better performance in constraining  $\text{NO}_x$  emissions in most urban areas, as shown in Figure 15c.

As a result, the conclusion drawn from Section 3.1 that assimilating synthetic TEMPO observations improves the skill of  $\text{NO}_x$  emissions inversion is still valid, even though the TEMPO observation error is assumed to be the same as TROPOMI data. We also anticipate the conclusion obtained from other TEMPO sensitivity runs will remain the same because the emissions inversion results from those ERs are either subjected to the diurnal variation of synthetic TEMPO  $\text{NO}_2$  data (e.g., TEMPO-exp1 and TEMPO-exp2) or the DA configuration (e.g., TEMPO-EMS), neither of which are directly related to the magnitude observation error.

Additionally, the consequence of setting an upper bound for the value of  $\gamma$  (i.e., 0.35) leads to an underestimation of synthetic TROPOMI  $\text{NO}_2$  observation error by 30%–50% in polluted areas when compared to real data, and the differences can be even larger in non-polluted regions. This indicates that the impact of our synthetic TROPOMI  $\text{NO}_2$  data on emissions inversion may be overestimated, particularly in non-polluted locations.

One method for relaxing the upper bound of  $\gamma$  is to decouple the observation error estimation algorithm from the process of generating perturbed synthetic column observations. Assuming the observation error follows a log-normal distribution could be another approach that allows for imposing large errors on synthetic observations while preserving the positiveness of the synthetic  $\text{NO}_2$  column observation.





**Figure 15.** (a) Time series of the averaged posterior  $\text{NO}_x$  emissions over the CONUS for TEMPO-exp3 and TEMPO-exp, (b) Same as (a) but for TEMPO-exp3 and TROPOMI-exp, (c) Taylor diagrams for posterior  $\text{NO}_x$  emissions from the TEMPO-exp3 and TROPOMI-exp.

## 5. Conclusions

We conducted an OSSE to assess how assimilating geostationary  $\text{NO}_2$  observations from TEMPO (GEO) improves the inverse modeling of  $\text{NO}_x$  emissions compared to TROPOMI (LEO). The emissions inversion is accomplished with WRF-Chem/DART using an EAKF with the state augmentation method. Several TEMPO ERs were conducted with different DA configurations to study the optimal strategy for constraining  $\text{NO}_x$  emissions using GEO  $\text{NO}_2$  observations.

Our results show that assimilating synthetic TEMPO observations allows for the recovery of true  $\text{NO}_x$  emissions at least twice as fast as TROPOMI. The posterior  $\text{NO}_x$  emissions from the TEMPO ER (TEMPO-exp) are also more accurate than the TROPOMI ER in most RRs and urban areas over CONUS. Specifically, the RMSE and  $R^2$  of the posterior  $\text{NO}_x$  emissions are improved by 12.5%–41.5% and 1.5%–17.1%, respectively, across the different RRs. Additionally, the TEMPO posterior  $\text{NO}_x$  emissions better capture the daily fluctuations of true emissions, while constraining the emissions with synthetic TROPOMI observations shows larger biases when trying to recover the more challenging time-varying true emissions.

We find that assimilating synthetic TEMPO observations with a  $\pm 1.5$ -hr DA window degrades the performance of the top-down  $\text{NO}_x$  emissions estimates. That degradation is primarily subject to the diurnal variability of our synthetic TEMPO  $\text{NO}_2$  data. In the transition cycles (15 and 00 UTC), large representativeness errors associated with the non-cycle time observation hours as well as more observations from the daytime side observation hours lead to substantial negative biases ( $-20\%$ ) in the posterior  $\text{NO}_x$  emissions. Using a  $\pm 1.5$ -hr DA window only during the midday cycles (18 and 21 UTC) can still cause a  $-10\%$  bias in the posterior  $\text{NO}_x$  emissions (TEMPO-exp2); this is owing to the scattering weights associated with the daytime side observation hours having greater sensitivity in the lower troposphere, which enhances representativeness error impact on the posterior. The  $\text{NO}_x$  emissions inversion also performs better when optimizing both  $\text{NO}_x$  concentration and emissions simultaneously under the influence of model error, particularly in polluted regions.

As a result, we propose the following best practices for constraining  $\text{NO}_x$  emissions with TEMPO observations:

1. Utilize a short DA window (e.g., 30 min) to avoid the potential representativeness errors associated with observations that are different from the DA cycle time hour. This is especially true for the day-night transition DA cycles;
2. Increase the frequency of assimilation cycling, use shorter cycling periods (e.g., 1 or 2 hr) to better account for the diurnal variations in the model and observations. This exploits the high temporal resolution of the geostationary observations and assimilates more observations without increasing the DA window size;
3. Jointly constrain concentrations and emissions.

It is worth noting that some of the best practices obtained in this work (e.g., using a short DA window at transition DA cycles) may only be applicable when the future operational TEMPO  $\text{NO}_2$  data display a similar diurnal pattern as our synthetic data and when the EnKF method is used.

We would also like to highlight certain challenges associated with assimilating remote sensing observations and emissions inverse modeling using the EnKF technique that require further investigation. For example, the approach for how to properly conduct vertical localization for non-local observations (e.g., TEMPO and TROPOMI  $\text{NO}_2$  columns) and the impact on emissions inversion have not yet been examined. Moreover, performing chemical DA and emissions inversion in log space can be further developed and implemented in the OSSE or real case study. Such an approach has been used to constrain methane emissions (Chen et al., 2022; Cui et al., 2019) and may improve the performance of  $\text{NO}_x$  emissions inversion because the distribution of trace-gas concentrations and emissions tends to be log-normal (Deeter et al., 2007; de Souza & Ozonur, 2019; Yuan et al., 2015), which violates the underlying Gaussian error assumption in the EnKF analysis process and may result in a sub-optimal posterior estimation.

This study can guide future top-down  $\text{NO}_x$  emissions estimation using geostationary observations, although the same model is utilized for NR and ERs. Future OSSEs should use two distinct models for more representative results. Despite these obstacles, we expect that operational geostationary satellite observations will improve the skill of top-down emission estimates and our ability to track the impact of specific emissions regulations or sudden societal adjustments on changes in air quality.

## Data Availability Statement

The WRF-Chem/DART code used in this study as well as the BAU and COVID-19 adjusted bottom-up emission inventory files developed by NOAA CSL can be found at: <https://csl.noaa.gov/groups/csl7/measurements/2020-covid-aqs/emissions/>.

## References

- Ahmadov, R., McKeen, S. A., Robinson, A. L., Bahreini, R., Middlebrook, A. M., de Gouw, J. A., et al. (2012). A volatility basis set model for summertime secondary organic aerosols over the eastern United States in 2006. *Journal of Geophysical Research*, 117(D6), D06301. <https://doi.org/10.1029/2011jd016831>
- Anderson, D. C., Loughner, C. P., Diskin, G., Weinheimer, A., Canty, T. P., Salawitch, R. J., et al. (2014). Measured and modeled CO and NO<sub>y</sub> in DISCOVER-AQ: An evaluation of emissions and chemistry over the eastern US. *Atmospheric Environment*, 96, 78–87. <https://doi.org/10.1016/j.atmosenv.2014.07.004>
- Anderson, J., Hoar, T., Raeder, K., Liu, H., Collins, N., Torn, R., & Avellano, A. (2009). The data assimilation research testbed: A community facility. *Bulletin of the American Meteorological Society*, 90(9), 1283–1296. <https://doi.org/10.1175/2009bams2618.1>

## Acknowledgments

We acknowledge support from the NOAA GeoXO program, NOAA NRDD Project (19533), NOAA Cooperative Agreement (NA17OAR4320101 and NA22OAR4320151), NASA 80NSSC21K1343, and the NASA Earth eXchange (NEX). We also thank NOAA's High Performance Computing Program and NASA's High End Computing (HEC) Program. The continued development of WRF-Chem/DART has supported in part by the NASA Ames Research Center (ARC), the NASA Postdoctoral Program, and the Colorado Department of Public Health and Environment. It has utilized resources from the NASA HEC Program of the NASA Advanced Supercomputing Division at ARC, and the RMACC Summit supercomputer, which is supported by the National Science Foundation (Awards ACI-1532235 and ACI-1532236), the University of Colorado Boulder, and Colorado State University. The Summit supercomputer is a joint effort of the University of Colorado Boulder and Colorado State University. Matthew S. Johnson acknowledges the internal funding from NASA's Earth Science Research and Analysis Program. Dr. Arthur P. Mizzi conceived of and designed the vertical and temporal decorrelation scheme for generating emissions ensemble.

- Anderson, J. L. (2001). An ensemble adjustment Kalman filter for data assimilation. *Monthly Weather Review*, *129*(12), 2884–2903. [https://doi.org/10.1175/1520-0493\(2001\)129<2884:AEAKFF>2.0.CO;2](https://doi.org/10.1175/1520-0493(2001)129<2884:AEAKFF>2.0.CO;2)
- Anderson, J. L. (2003). A local least squares framework for ensemble filtering. *Monthly Weather Review*, *131*(4), 634–642. [https://doi.org/10.1175/1520-0493\(2003\)131<0634:ALLSFF>2.0.CO;2](https://doi.org/10.1175/1520-0493(2003)131<0634:ALLSFF>2.0.CO;2)
- Anderson, J. L. (2007). An adaptive covariance inflation error correction algorithm for ensemble filters. *Tellus Series A-Dynamic Meteorology and Oceanography*, *59*(2), 210–224. <https://doi.org/10.1111/j.1600-0870.2006.00216.x>
- Barbu, A. L., Segers, A. J., Schaap, M., Heemink, A. W., & Buitjes, P. J. H. (2009). A multi-component data assimilation experiment directed to sulphur dioxide and sulphate over Europe. *Atmospheric Environment*, *43*(9), 1622–1631. <https://doi.org/10.1016/j.atmosenv.2008.12.005>
- Barré, J., Edwards, D., Worden, H., Arellano, A., Gaubert, B., Da Silva, A., et al. (2016). Barré the feasibility of monitoring carbon monoxide in the lower troposphere from a constellation of northern hemisphere geostationary satellites: Global scale assimilation experiments (Part II). *Atmospheric Environment*, *140*, 188–201. <https://doi.org/10.1016/j.atmosenv.2016.06.001>
- Beirle, S., Boersma, K. F., Platt, U., Lawrence, M. G., & Wagner, T. (2011). Megacity emissions and lifetimes of nitrogen oxides probed from space. *Science*, *333*(6050), 1737–1739. <https://doi.org/10.1126/science.1207824>
- Boersma, K. F., Eskes, H. J., & Brinksma, E. J. (2004). Error analysis for tropospheric NO<sub>2</sub> retrieval from space. *Journal of Geophysical Research*, *109*(D4), D04311. <https://doi.org/10.1029/2003jd003962>
- Boynard, A., Beekmann, M., Foret, G., Ung, A., Szopa, S., Schmechtig, C., & Coman, A. (2011). An ensemble assessment of regional ozone model uncertainty with an explicit error representation. *Atmospheric Environment*, *45*(3), 784–793. <https://doi.org/10.1016/j.atmosenv.2010.08.006>
- Campbell, P. C., Tong, D., Tang, Y., Baker, B., Lee, P., Saylor, R., et al. (2021). Impacts of the COVID-19 economic slowdown on ozone pollution in the U.S. *Atmospheric Environment*, *264*, 118713. <https://doi.org/10.1016/j.atmosenv.2021.118713>
- Cao, H., Henze, D. K., Zhu, L., Shephard, M. W., Cady-Pereira, K., Dammers, E., et al. (2022). 4D-Var inversion of European NH<sub>3</sub> emissions using CrIS NH<sub>3</sub> measurements and GEOS-Chem adjoint with Bi-directional and uni-directional flux schemes. *Journal of Geophysical Research: Atmospheres*, *127*(9), e2021JD035687. <https://doi.org/10.1029/2021jd035687>
- Chang, C. C., Yang, S. C., & Penny, S. G. (2022). A regional hybrid gain data assimilation system and preliminary evaluation based on radio occultation reflectivity assimilation. *SOLA*, *18*, 33–40. <https://doi.org/10.2151/sola.2022-006>
- Chen, Z., Jacob, D. J., Nessler, H., Sulprizio, M. P., Lorente, A., Varon, D. J., et al. (2022). Methane emissions from China: A high-resolution inversion of TROPOMI satellite observations. *Atmospheric Chemistry and Physics*, *22*(16), 10809–10826. <https://doi.org/10.5194/acp-22-10809-2022>
- Choi, J., Henze, D. K., Cao, H., Nowlan, C. R., González Abad, G., Kwon, H., et al. (2022). An inversion framework for optimizing non-methane VOC emissions using remote sensing and airborne observations in northeast Asia during the KORUS-AQ field campaign. *Journal of Geophysical Research: Atmospheres*, *127*(7), e2021JD035844. <https://doi.org/10.1029/2021jd035844>
- Coggon, M. M., Gkatzelis, G. I., McDonald, B. C., Gilman, J. B., Schwantes, R. H., Abuhassan, N., et al. (2021). Volatile chemical product emissions enhance ozone and modulate urban chemistry. *Proceedings of the National Academy of Sciences of the United States of America*, *118*(32), e2026653118. <https://doi.org/10.1073/pnas.2026653118>
- Cui, Y. Y., Henze, D. K., Brioude, J., Angevine, W. M., Liu, Z., Bousseret, N., et al. (2019). Inversion estimates of lognormally distributed methane emission rates from the Haynesville-Bossier oil and gas production region using airborne measurements. *Journal of Geophysical Research: Atmospheres*, *124*(6), 3520–3531. <https://doi.org/10.1029/2018jd029489>
- Dedoussi, I. C., Eastham, S. D., Monier, E., & Barrett, S. R. H. (2020). Premature mortality related to United States cross-state air pollution. *Nature*, *578*(7794), 261–265. <https://doi.org/10.1038/s41586-020-1983-8>
- Deeter, M. N., Edwards, D. P., & Gille, J. C. (2007). Retrievals of carbon monoxide profiles from MOPITT observations using lognormal a priori statistics. *Journal of Geophysical Research*, *112*(D11), D11311. <https://doi.org/10.1029/2006jd007999>
- Deschneemaeker, M., Plu, M., Maréchal, V., Claeys, M., Olivier, F., Aoun, Y., et al. (2019). Monitoring aerosols over Europe: An assessment of the potential benefit of assimilating the VIS04 measurements from the future MTG/FCI geostationary imager. *Atmospheric Measurement Techniques*, *12*(2), 1251–1275. <https://doi.org/10.5194/amt-12-1251-2019>
- de Souza, A., & Ozonur, D. (2019). Statistical behavior of O<sub>3</sub>, O<sub>x</sub>, NO, NO<sub>2</sub>, and NO<sub>x</sub> in urban environment. *Ozone: Science & Engineering*, *42*(1), 66–78. <https://doi.org/10.1080/01919512.2019.1602468>
- Dix, B., Francoeur, C., Li, M., Serrano-Calvo, R., Levelt, P. F., Veeffkind, J. P., et al. (2022). Quantifying NO<sub>x</sub> emissions from U.S. Oil and gas production regions using TROPOMI NO<sub>2</sub>. *ACS Earth and Space Chemistry*, *6*(2), 403–414. <https://doi.org/10.1021/acsearthspacechem.1c00387>
- Doombia, T., Granier, C., Elguindi, N., Bouarar, I., Darras, S., Brasseur, G., et al. (2021). Changes in global air pollutant emissions during the COVID-19 pandemic: A dataset for atmospheric modeling. *Earth System Science Data*, *13*(8), 4191–4206. <https://doi.org/10.5194/essd-13-4191-2021>
- Duncan, B. N., Yoshida, Y., de Foy, B., Lamsal, L. N., Streets, D. G., Lu, Z., et al. (2013). The observed response of ozone monitoring instrument (OMI) NO<sub>2</sub> columns to NO<sub>x</sub> emission controls on power plants in the United States: 2005–2011. *Atmospheric Environment*, *81*, 102–111. <https://doi.org/10.1016/j.atmosenv.2013.08.068>
- Elbern, H., Strunk, A., Schmidt, H., & Talagrand, O. (2007). Emission rate and chemical state estimation by 4-dimensional variational inversion. *Atmospheric Chemistry and Physics*, *7*(14), 3749–3769. <https://doi.org/10.5194/acp-7-3749-2007>
- Elguindi, N., Granier, C., Stavrou, T., Darras, S., Bauwens, M., Cao, H., et al. (2020). Intercomparison of magnitudes and trends in anthropogenic surface emissions from bottom-up inventories, top-down estimates, and emission scenarios. *Earth's Future*, *8*(8), e2020EF001520. <https://doi.org/10.1029/2020ef001520>
- Environmental Protection Agency. (2020). *National emissions inventory (NEI) 2017, version 1*. Office of Air Quality Planning and Standards, U.S. Environmental Protection Agency.
- Evensen, G. (1994). Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics. *Journal of Geophysical Research*, *99*(C5), 10143–10162. <https://doi.org/10.1029/94jc00572>
- Evensen, G. (2003). The ensemble Kalman filter: Theoretical formulation and practical implementation. *Ocean Dynamics*, *53*(4), 343–367. <https://doi.org/10.1007/s10236-003-0036-9>
- Forster, P. M., Forster, H. I., Evans, M. J., Gidden, M. J., Jones, C. D., Keller, C. A., et al. (2020). Current and future global climate impacts resulting from COVID-19. *Nature Climate Change*, *10*(10), 913–919. <https://doi.org/10.1038/s41558-020-0883-0>
- Francoeur, C. B., McDonald, B. C., Gilman, J. B., Zarzana, K. J., Dix, B., Brown, S. S., et al. (2021). Quantifying methane and ozone precursor emissions from oil and gas production regions across the contiguous US. *Environmental Science & Technology*, *55*(13), 9129–9139. <https://doi.org/10.1021/acs.est.0c07352>
- Fujita, E. M., Campbell, D. E., Zielinska, B., Chow, J. C., Lindhjem, C. E., DenBleyker, A., et al. (2012). Comparison of the MOVES2010a, MOBILE6.2, and EMFAC2007 mobile source emission models with on-road traffic tunnel and remote sensing measurements. *Journal of the Air & Waste Management Association*, *62*(10), 1134–1149. <https://doi.org/10.1080/10962247.2012.699016>

- Gaspari, G., & Cohn, S. E. (1999). Construction of correlation functions in two and three dimensions. *Quarterly Journal of the Royal Meteorological Society*, 125(554), 723–757. <https://doi.org/10.1002/qj.49712555417>
- Gately, C. K., Hutyra, L. R., & Sue Wing, I. (2015). Cities, traffic, and CO<sub>2</sub>: A multidecadal assessment of trends, drivers, and scaling relationships. *Proceedings of the National Academy of Sciences of the United States of America*, 112(16), 4999–5004. <https://doi.org/10.1073/pnas.1421723112>
- Gaubert, B., Coman, A., Foret, G., Meleux, F., Ung, A., Rouil, L., et al. (2014). Regional scale ozone data assimilation using an ensemble Kalman filter and the CHIMERE chemical transport model. *Geoscientific Model Development*, 7(1), 283–302. <https://doi.org/10.5194/gmd-7-283-2014>
- Gkatzelis, G. I., Gilman, J. B., Brown, S. S., Eskes, H., Gomes, A. R., Lange, A. C., et al. (2021). The global impacts of COVID-19 lockdowns on urban air pollution: A critical review and recommendations. *Elementa: Science of the Anthropocene*, 9(1), 00176. <https://doi.org/10.1525/elementa.2021.00176>
- Grell, G. A., Peckham, S. E., Schmitz, R., McKeen, S. A., Frost, G., Skamarock, W. C., & Eder, B. (2005). Fully coupled “online” chemistry within the WRF model. *Atmospheric Environment*, 39(37), 6957–6975. <https://doi.org/10.1016/j.atmosenv.2005.04.027>
- Harkins, C., McDonald, B. C., Henze, D. K., & Wiedinmyer, C. (2021). A fuel-based method for updating mobile source emissions during the COVID-19 pandemic. *Environmental Research Letters*, 16(6), 065018. <https://doi.org/10.1088/1748-9326/ac0660>
- Hartung, J., Knapp, G., & Sinha, B. K. (2008). *Statistical meta-analysis with applications*. John Wiley & Sons, Inc.
- Huang, Y., Wei, J., Jin, J., Zhou, Z., & Gu, Q. (2022). CO fluxes in western Europe during 2017–2020 winter seasons inverted by WRF-Chem/data assimilation research testbed with MOPITT observations. *Remote Sensing*, 14(5), 1133. <https://doi.org/10.3390/rs14051133>
- Jerrett, M., Burnett, R. T., Pope, C. A., Ito, K., Thurston, G., Krewski, D., et al. (2009). Long-term ozone exposure and mortality. *New England Journal of Medicine*, 360(11), 1085–1095. <https://doi.org/10.1056/nejmoa0803894>
- Jung, Y., González Abad, G., Nowlan, C. R., Chance, K., Liu, X., Torres, O., & Ahn, C. (2019). Explicit aerosol correction of OMI formaldehyde retrievals. *Earth and Space Science*, 6(11), 2087–2105. <https://doi.org/10.1029/2019ea000702>
- Kim, J., Jeong, U., Ahn, M., Kim, J. H., Park, R. J., Lee, H., et al. (2020). New Era of air quality monitoring from space: Geostationary environment monitoring spectrometer (GEMS). *Bulletin of the American Meteorological Society*, 101(1), E1–E22. <https://doi.org/10.1175/BAMS-D-18-0013.1>
- Kim, S.-W., McKeen, S. A., Frost, G. J., Lee, S.-H., Trainer, M., Richter, A., et al. (2011). Evaluations of NO<sub>x</sub> and highly reactive VOC emission inventories in Texas and their implications for ozone plume simulations during the Texas Air Quality Study 2006. *Atmospheric Chemistry and Physics*, 11(22), 11361–11386. <https://doi.org/10.5194/acp-11-11361-2011>
- W. Lahoz, B. Khattatov, & R. Ménard (Eds.) (2010). *Data assimilation making sense of observations* (p. 718). Springer.
- Lamsal, L. N., Krotkov, N. A., Vasilkov, A., Marchenko, S., Qin, W., Yang, E. S., et al. (2021). Ozone Monitoring Instrument (OMI) Aura nitrogen dioxide standard product version 4.0 with improved surface and cloud treatments. *Atmospheric Measurement Techniques*, 14(1), 455–479. <https://doi.org/10.5194/amt-14-455-2021>
- Li, M., McDonald, B. C., McKeen, S. A., Eskes, H., Levelt, P., Francoeur, C., et al. (2021). Assessment of updated fuel-based emissions inventories over the contiguous United States using TROPOMI NO<sub>2</sub> retrievals. *Journal of Geophysical Research: Atmospheres*, 126(24), e2021JD035484. <https://doi.org/10.1029/2021JD035484>
- Liu, X., Bhartia, P. K., Chance, K., Spurr, R. J. D., & Kurosu, T. P. (2010). Ozone profile retrievals from the ozone monitoring instrument. *Atmospheric Chemistry and Physics*, 10(5), 2521–2537. <https://doi.org/10.5194/acp-10-2521-2010>
- Liu, X., Mizzi, A. P., Anderson, J. L., Fung, I. Y., & Cohen, R. C. (2017). Assimilation of satellite NO<sub>2</sub> observations at high spatial resolution using OSSEs. *Atmospheric Chemistry and Physics*, 17(11), 7067–7081. <https://doi.org/10.5194/acp-17-7067-2017>
- Lorente, A., Folkert Boersma, K., Yu, H., Dörner, S., Hilboll, A., Richter, A., et al. (2017). Structural uncertainty in air mass factor calculation for NO<sub>2</sub> and HCHO satellite retrievals. *Atmospheric Measurement Techniques*, 10(3), 759–782. <https://doi.org/10.5194/amt-10-759-2017>
- Ma, C., Wang, T., Mizzi, A. P., Anderson, J. L., Zhuang, B., Xie, M., & Wu, R. (2019). Multiconstituent data assimilation with WRF-Chem/DART: Potential for adjusting anthropogenic emissions and improving air quality forecasts over eastern China. *Journal of Geophysical Research: Atmospheres*, 124(13), 7393–7412. <https://doi.org/10.1029/2019jd030421>
- Martin, R. V., Jacob, D. J., Chance, K., Kurosu, T. P., Palmer, P. I., & Evans, M. J. (2003). Global inventory of nitrogen oxide emissions constrained by space-based observations of NO<sub>2</sub> columns. *Journal of Geophysical Research*, 108(D17), 4537. <https://doi.org/10.1029/2003JD003453>
- McDonald, B. C., de Gouw, J. A., Gilman, J. B., Jathar, S. H., Akherati, A., Cappa, C. D., et al. (2018). Volatile chemical products emerging as largest petrochemical source of urban organic emissions. *Science*, 359(6377), 760–764. <https://doi.org/10.1126/science.aag0524>
- McDonald, B. C., McBride, Z. C., Martin, E. W., & Harley, R. A. (2014). High-resolution mapping of motor vehicle carbon dioxide emissions. *Journal of Geophysical Research: Atmospheres*, 119(9), 5283–5298. <https://doi.org/10.1002/2013jd021219>
- McDonald, B. C., McKeen, S. A., Cui, Y. Y., Ahmadov, R., Kim, S.-W., Frost, G. J., et al. (2018). Modeling ozone in the Eastern U.S. using a fuel-based mobile source emissions inventory. *Environmental Science & Technology*, 52(13), 7360–7370. <https://doi.org/10.1021/acs.est.8b00778>
- Ménard, R., & Chang, L. (2000). Assimilation of stratospheric chemical tracer observations using a Kalman filter. Part II:  $\chi^2$ -Validated results and analysis of variance and correlation dynamics. *Monthly Weather Review*, 128(8), 2672–2686. [https://doi.org/10.1175/1520-0493\(2000\)128<2672:AOSCTO>2.0.CO;2](https://doi.org/10.1175/1520-0493(2000)128<2672:AOSCTO>2.0.CO;2)
- Miyazaki, K., Eskes, H., Sudo, K., Boersma, K. F., Bowman, K., & Kanaya, Y. (2017). Decadal changes in global surface NO<sub>x</sub> emissions from multi-constituent satellite data assimilation. *Atmospheric Chemistry and Physics*, 17(2), 807–837. <https://doi.org/10.5194/acp-17-807-2017>
- Miyazaki, K., Eskes, H. J., & Sudo, K. (2012). Global NO<sub>x</sub> emission estimates derived from an assimilation of OMI tropospheric NO<sub>2</sub> columns. *Atmospheric Chemistry and Physics*, 12(5), 2263–2288. <https://doi.org/10.5194/acp-12-2263-2012>
- Mizzi, A. P., Arellano, A. F., Jr., Edwards, D. P., Anderson, J. L., & Pfister, G. G. (2016). Assimilating compact phase space retrievals of atmospheric composition with WRF-Chem/DART: A regional chemical transport/ensemble Kalman filter data assimilation system. *Geoscientific Model Development*, 9(3), 965–978. <https://doi.org/10.5194/gmd-9-965-2016>
- Mizzi, A. P., Edwards, D. P., & Anderson, J. L. (2018). Assimilating compact phase space retrievals (CPSRs): Comparison with independent observations (MOZAIC in situ and IASI retrievals) and extension to assimilation of truncated retrieval profiles. *Geoscientific Model Development*, 11(9), 3727–3745. <https://doi.org/10.5194/gmd-11-3727-2018>
- Müller, J. F., & Stavrou, T. (2005). Inversion of CO and NO<sub>x</sub> emissions using the adjoint of the IMAGES model. *Atmospheric Chemistry and Physics*, 5(5), 1157–1186. <https://doi.org/10.5194/acp-5-1157-2005>
- Naeger, A. R., Newchurch, M. J., Moore, T., Chance, K., Liu, X., Alexander, S., et al. (2021). Revolutionary air-pollution applications from future tropospheric emissions: Monitoring of pollution (TEMPO) observations. *Bulletin of the American Meteorological Society*, 102(9), E1735–E1741. <https://doi.org/10.1175/bams-d-21-0050.1>
- Nault, B. A., Laughner, J. L., Wooldridge, P. J., Crouse, J. D., Dibb, J., Diskin, G., et al. (2017). Lightning NO<sub>x</sub> emissions: Reconciling measured and modeled estimates with updated NO<sub>x</sub> chemistry. *Geophysical Research Letters*, 44(18), 9479–9488. <https://doi.org/10.1002/2017gl074436>

- Peng, Z., Lei, L., Liu, Z., Su, J., Ding, A., Ban, J., et al. (2018). The impact of multi-species surface chemical observation assimilation on air quality forecasts in China. *Atmospheric Chemistry and Physics*, 18(23), 17387–17404. <https://doi.org/10.5194/acp-18-17387-2018>
- Pierce, R. B., Al-Saadi, J. A., Schaack, T., Lenzen, A., Zapotocny, T., Johnson, D., et al. (2003). Regional air quality modeling system (RAQMS) predictions of the tropospheric ozone budget over east Asia. *Journal of Geophysical Research*, 108(D21), 8825. <https://doi.org/10.1029/2002jd003176>
- Pope, C. A., Ezzati, M., & Dockery, D. W. (2009). Fine-particulate air pollution and life expectancy in the United States. *New England Journal of Medicine*, 360(4), 376–386. <https://doi.org/10.1056/nejmsa0805646>
- Qin, W., Fasnacht, Z., Haffner, D., Vasilkov, A., Joiner, J., Krotkov, N., et al. (2019). A geometry-dependent surface Lambertian-equivalent reflectivity product for UV–Vis retrievals – Part 1: Evaluation over land surfaces using measurements from OMI at 466 nm. *Atmospheric Measurement Techniques*, 12(7), 3997–4017. <https://doi.org/10.5194/amt-12-3997-2019>
- Qu, Z., Henze, D. K., Capps, S. L., Wang, Y., Xu, X., Wang, J., & Keller, M. (2017). Monthly top-down NO<sub>x</sub> emissions for China (2005–2012): A hybrid inversion method and trend analysis. *Journal of Geophysical Research: Atmospheres*, 122(8), 4600–4625. <https://doi.org/10.1002/2016jd025852>
- Qu, Z., Henze, D. K., Theys, N., Wang, J., & Wang, W. (2019). Hybrid mass balance/4D-var joint inversion of NO<sub>x</sub> and SO<sub>2</sub> emissions in East Asia. *Journal of Geophysical Research: Atmospheres*, 124(14), 8203–8224. <https://doi.org/10.1029/2018jd030240>
- Quesada-Ruiz, S., Attié, J. L., Lahoz, W. A., Abida, R., Ricaud, P., El Amraoui, L., et al. (2020). Benefit of ozone observations from Sentinel-5P and future Sentinel-4 missions on tropospheric composition. *Atmospheric Measurement Techniques*, 13(1), 131–152. <https://doi.org/10.5194/amt-13-131-2020>
- Raeder, K., Anderson, J. L., Collins, N., Hoar, T. J., Kay, J. E., Lauritzen, P. H., & Pincus, R. (2012). DART/CAM: An ensemble data assimilation system for CESM atmospheric models. *Journal of Climate*, 25(18), 6304–6317. <https://doi.org/10.1175/JCLI-D-11-00395.1>
- Seinfeld, J. H., & Pandis, S. N. (2012). *Atmospheric chemistry and physics: From air pollution to climate change*. John Wiley & Sons.
- Sekiya, T., Miyazaki, K., Eskes, H., Sudo, K., Takigawa, M., & Kanaya, Y. (2022). A comparison of the impact of TROPOMI and OMI tropospheric NO<sub>2</sub> on global chemical data assimilation. *Atmospheric Measurement Techniques*, 15(6), 1703–1728. <https://doi.org/10.5194/amt-15-1703-2022>
- Shu, L., Zhu, L., Bak, J., Zoogman, P., Han, H., Long, X., et al. (2022). Improved ozone simulation in East Asia via assimilating observations from the first geostationary air-quality monitoring satellite: Insights from an Observing System Simulation Experiment. *Atmospheric Environment*, 274, 119003. <https://doi.org/10.1016/j.atmosenv.2022.119003>
- Souri, A. H., Nowlan, C. R., González Abad, G., Zhu, L., Blake, D. R., Fried, A., et al. (2020). An inversion of NO<sub>x</sub> and non-methane volatile organic compound (NMVOC) emissions using satellite observations during the KORUS-AQ campaign and implications for surface ozone over East Asia. *Atmospheric Chemistry and Physics*, 20(16), 9837–9854. <https://doi.org/10.5194/acp-20-9837-2020>
- Spurr, R. J. (2006). VLIDORT: A linearized pseudo-spherical vector discrete ordinate radiative transfer code for forward model and retrieval studies in multilayer multiple scattering media. *Journal of Quantitative Spectroscopy and Radiative Transfer*, 102(2), 316–342. <https://doi.org/10.1016/j.jqsrt.2006.05.005>
- Stavrakou, T., Müller, J. F., Boersma, K. F., Van Der, A. R. J., Kurokawa, J., Ohara, T., & Zhang, Q. (2013). Key chemical NO<sub>x</sub> sink uncertainties and how they influence top-down emissions of nitrogen oxides. *Atmospheric Chemistry and Physics*, 13(17), 9057–9082. <https://doi.org/10.5194/acp-13-9057-2013>
- Timmermans, R., Segers, A., Curier, L., Abida, R., Attié, J. L., El Amraoui, L., et al. (2019). Impact of synthetic space-borne NO<sub>2</sub> observations from the Sentinel-4 and Sentinel-5P missions on tropospheric NO<sub>2</sub> analyses. *Atmospheric Chemistry and Physics*, 19(19), 12811–12833. <https://doi.org/10.5194/acp-19-12811-2019>
- van Geffen, J., Boersma, K. F., Eskes, H., Sneep, M., ter Linden, M., Zara, M., & Veeffkind, J. P. (2020). S5P TROPOMI NO<sub>2</sub> slant column retrieval: Method, stability, uncertainties and comparisons with OMI. *Atmospheric Measurement Techniques*, 13(3), 1315–1335. <https://doi.org/10.5194/amt-13-1315-2020>
- Veeffkind, J. P., Aben, I., McMullan, K., Förster, H., de Vries, J., Otter, G., et al. (2012). TROPOMI on the ESA Sentinel-5 precursor: A GMES mission for global observations of the atmospheric composition for climate, air quality and ozone layer applications. *Remote Sensing of Environment*, 120, 70–83. <https://doi.org/10.1016/j.rse.2011.09.027>
- Vinken, G. C. M., Boersma, K. F., Maasakkers, J. D., Adon, M., & Martin, R. V. (2014). Worldwide biogenic soil NO<sub>x</sub> emissions inferred from OMI NO<sub>2</sub> observations. *Atmospheric Chemistry and Physics*, 14(18), 10363–10381. <https://doi.org/10.5194/acp-14-10363-2014>
- Wiedinmyer, C., Akagi, S. K., Yokelson, R. J., Emmons, L. K., Al-Saadi, J. A., Orlando, J. J., & Soja, A. J. (2011). The fire INventory from NCAR (FINN): A high resolution global model to estimate the emissions from open burning. *Geoscientific Model Development*, 4(3), 625–641. <https://doi.org/10.5194/gmd-4-625-2011>
- Ye, H., You, W., Zang, Z., Pan, X., Wang, D., Zhou, N., et al. (2022). Observing system simulation experiment (OSSE)-quantitative evaluation of lidar observation networks to improve 3D aerosol forecasting in China. *Atmospheric Research*, 270, 106069. <https://doi.org/10.1016/j.atmosres.2022.106069>
- Yuan, B., Kaser, L., Karl, T., Graus, M., Peischl, J., Campos, T. L., et al. (2015). Airborne flux measurements of methane and volatile organic compounds over the Haynesville and Marcellus shale gas production regions. *Journal of Geophysical Research: Atmospheres*, 120(12), 6271–6289. <https://doi.org/10.1002/2015jd023242>
- Zhang, Q., Li, M., Wang, M., Mizzi, A., Huang, Y., Wei, C., et al. (2021). CO<sub>2</sub> flux over the contiguous United States in 2016 inverted by WRF-Chem/DART from OCO-2 XCO<sub>2</sub> retrievals. *Remote Sensing*, 13(15), 2996. <https://doi.org/10.3390/rs13152996>
- Zhang, Q., Li, M., Wei, C., Mizzi, A. P., Huang, Y., & Gu, Q. (2021). Assimilation of OCO-2 retrievals with WRF-Chem/DART: A case study for the midwestern United States. *Atmospheric Environment*, 246, 118106. <https://doi.org/10.1016/j.atmosenv.2020.118106>
- Zoogman, P., Jacob, D. J., Chance, K., Liu, X., Lin, M., Fiore, A., & Travis, K. (2014). Monitoring high-ozone events in the US Intermountain West using TEMPO geostationary satellite observations. *Atmospheric Chemistry and Physics*, 14(12), 6261–6271. <https://doi.org/10.5194/acp-14-6261-2014>
- Zoogman, P., Liu, X., Suleiman, R., Pennington, W., Flittner, D., Al-Saadi, J., et al. (2017). Tropospheric emissions: Monitoring of pollution (TEMPO). *Journal of Quantitative Spectroscopy and Radiative Transfer*, 186, 17–39. <https://doi.org/10.1016/j.jqsrt.2016.05.008>

## References From the Supporting Information

- Boccippio, D. J., Cummins, K. L., Christian, H. J., & Goodman, S. J. (2001). Combined satellite- and surface-based estimation of the intracloud–cloud-to-ground lightning ratio over the continental United States. *Monthly Weather Review*, *129*(1), 108–122. [https://doi.org/10.1175/1520-0493\(2004\)132<0103:aratim>2.0.co;2](https://doi.org/10.1175/1520-0493(2004)132<0103:aratim>2.0.co;2)
- Grell, G. A., & Dévényi, D. (2002). A generalized approach to parameterizing convection combining ensemble and data assimilation techniques. *Geophysical Research Letters*, *29*(14), 38. <https://doi.org/10.1029/2002gl015311>
- Hong, S. Y., Dudhia, J., & Chen, S. H. (2004). A revised approach to ice microphysical processes for the bulk parameterization of clouds and precipitation. *Monthly Weather Review*, *132*(1), 103–120. [https://doi.org/10.1175/1520-0493\(2004\)132<0103:aratim>2.0.co;2](https://doi.org/10.1175/1520-0493(2004)132<0103:aratim>2.0.co;2)
- Hu, L., Keller, C. A., Long, M. S., Sherwen, T., Auer, B., Da Silva, A., et al. (2018). Global simulation of tropospheric chemistry at 12.5 km resolution: Performance and evaluation of the GEOS-Chem chemical module (v10-1) within the NASA GEOS earth system model (GEOS-5 ESM). *Geoscientific Model Development*, *11*(11), 4603–4620. <https://doi.org/10.5194/gmd-11-4603-2018>
- Johnson, M. S., Strawbridge, K., Knowland, K. E., Keller, C., & Travis, M. (2021). Long-range transport of Siberian biomass burning emissions to North America during FIREX-AQ. *Atmospheric Environment*, *252*, 118241. <https://doi.org/10.1016/j.atmosenv.2021.118241>
- Joiner, J., Vasilkov, A. P., Gupta, P., Bhartia, P. K., Veefkind, P., Sneep, M., et al. (2012). Fast simulators for satellite cloud optical centroid pressure retrievals; evaluation of OMI cloud retrievals. *Atmospheric Measurement Techniques*, *5*(3), 529–545. <https://doi.org/10.5194/amt-5-529-2012>
- Keller, C. A., Knowland, K. E., Duncan, B. N., Liu, J., Anderson, D. C., Das, S., et al. (2021). Description of the NASA GEOS composition forecast modeling system GEOS-CF v1.0. *Journal of Advances in Modeling Earth Systems*, *13*(4), e2020MS002413. <https://doi.org/10.1029/2020ms002413>
- Knowland, K. E., Keller, C. A., & Lucchesi, R. A. (2020). File specification for GEOS-CF products. GMAO Office Note No. 17 (Version 1.1) (p. 37). Retrieved from [http://gmao.gsfc.nasa.gov/pubs/office\\_notes](http://gmao.gsfc.nasa.gov/pubs/office_notes)
- Mlawer, E. J., Taubman, S. J., Brown, P. D., Iacono, M. J., & Clough, S. A. (1997). Radiative transfer for inhomogeneous atmospheres: RRTM, a validated correlated-k model for the longwave. *Journal of Geophysical Research*, *102*(D14), 16663–16682. <https://doi.org/10.1029/97jd00237>
- Mukul Tewari, N., Tewari, M., Chen, F., Wang, W., Dudhia, J., LeMone, M., et al. (2004). Implementation and verification of the unified NOAA land surface model in the WRF model. In *Presented at the 20th conference on weather analysis and forecasting/16th conference on numerical weather prediction (Formerly Paper Number 17.5)* (pp. 11–15).
- Nakanishi, M., & Niino, H. (2009). Development of an improved turbulence closure model for the atmospheric boundary layer. *Journal of the Meteorological Society of Japan. Ser. II*, *87*(5), 895–912. <https://doi.org/10.2151/jmsj.87.895>