

Response to Reviewer 1:

We thank the reviewer for the time and review. Your feedback has improved our manuscript.

In their paper, X. Liu and co-authors discuss the impact of future geostationary NO₂ observations on air quality models and the potential of these observations to improve estimates of emissions.

General assessment: The highlight of this paper is a documentation of the sensitivity of NO_x emission estimates on the quality of the meteorological fields and the winds in particular. As shown, these issues are especially critical when the resolution of the observations and the models advances to the km scale. This sensitivity is demonstrated in a clear way, and to my opinion is an important aspect to keep in mind for the development of future regional analysis systems. As such, I am in favor of publishing these results in ACP.

The paper is well written, has a good introduction, but the overview of the method C1 and tools is very short and condensed. The implementation of the OSSE setup is not provided with enough detail for other groups to repeat the experiment, see my detailed comments below. I therefore would urge the authors to expand these parts of the paper.

We have made major revisions in section 2~4. In section 2, we provide detailed description on the assimilation system including the forecast model, DART facility (with localization settings), initial and boundary condition ensemble, emission update scheme and synthetic meteorological and chemical observations. In section 3, we redesign the assimilation experiment.

In order to judge the importance of the meteorological uncertainties for emission estimates, it is important to provide estimates of real-life uncertainties in state-of-the-art regional weather analysis systems, and to compare these uncertainties with the OSSE setup. This would be a valuable addition to the conclusions section.

We agree that it's important to compare the identified wind uncertainties in this work with state-of-the-art regional weather analysis. The available atmospheric reanalysis products (<https://climatedataguide.ucar.edu/climate-data/atmospheric-reanalysis-overview-comparison-tables>) provide a dynamically consistent estimate of the climate state at each time step. These assimilation scheme and models ingest all available observations every 6-12 hours and produce global and regional reanalysis at mesoscales (above 32 km). The products with such assimilation frequency and spatial resolution do not provide analysis uncertainties at the fine scale as in this work, which is hourly assimilation at the resolution of 3 km. In this work, we downscale NARR (32 km) to initialize the WRF forecast and use WRFDA to generate the ensemble. WRFDA system uses a state of the art method for generating the ensembles/quantifying the meteorological uncertainty. It uses the NMC method to generate the static error covariance and then uses the static error covariance to perturb the meteorological fields. Since the WRF meteorological forecast errors are comparable to other regional system, we would expect the uncertainties of the associated ensembles to be representative and reasonable. So we revised the text as follows:

“For meteorology ensemble, random perturbations were added to each member by sampling the NCEP background error covariance using WRF Data Assimilation System (<http://www2.mmm.ucar.edu/wrf/users/wrfda>).”

Since we redesign our experiment to assimilate meteorological observations hourly, the meteorological uncertainties becomes small enough to allow chemical observations to constrain emissions successfully in our results. Then we analyze whether the wind uncertainties are important for emission estimation in REA. Based on our result in REA run, we conclude as following:

“Last but not least our results show that the information of wind *uncertainties* is not important for NO_x emission estimation. Assimilations of NO₂ only with the meteorology constrained from an hourly weather assimilation product would perform as well as the joint assimilation of meteorology and chemistry.”

It seems the authors make very optimistic assumptions on the uncertainty of the TEMPO retrievals. Is this justified and are the conclusions sensitive to the choice made for the observation error?

In ensemble assimilation method, the adjustment on the emissions depends on the relative magnitude of TEMPO observation uncertainties and the prior emission uncertainties. We agree that the uncertainty of TEMPO retrieval in this study is optimistic compared to current OMI retrieval. However, previously we have conducted experiments with TEMPO observation of 30%. We find the emission estimation results not promising because the observation uncertainties are too large to let the observations influence the emissions. In addition, the prior emission uncertainty is set to be 30% in our OSSE. This can also be very optimistic given the dispute on the accuracy of bottom-up inventory. In real application, with a reasonable estimation of prior emission uncertainty which will be higher than 30%, it is highly possible that more weights will be given to TEMPO observations to adjust the emissions. The choice of TEMPO retrieval uncertainty in this study is limited by the estimation of the prior emission uncertainties. In real application, even if the reported TEMPO uncertainty is too large to adjust the emissions, the conclusions of our OSSEs on the optimal approach of NO_x emission estimation will not change.

Last line p2: 3.3 to 5 molecules/cm² ?

We change the text as, “biases of 34% (3.3 to 5.0×10^{15} molecules/cm²) are found in the modeled average NO₂ column over Los Angeles at resolutions of 96 km compared with 12 km.”

p3, l27: “idealized profile setting provided by WRF-Chem.” This is unclear and needs to be explained.

In our updated experiments we initialize WRF-Chem using the global chemical transport model outputs that will provide a realistic initial condition for chemical species. We modified the text as, “We use the global chemical model output from MOZART to initialize the chemical simulation on the outer domain and to provide the chemical boundary condition.”

p4, top: Please discuss the state augmentation approach for NO_x in more detail: how are the emissions perturbed as compared to the inventory, and how is the ensemble constructed.

Compared with the NEI 2011 which is referred as the truth, the CR emissions are perturbed by adding a 30% low bias uniformly in temporal and spatial scale. This is used as the ensemble mean of emissions in the assimilation process to construct the emission ensemble. Each ensemble member is then generated by adding a perturbation to the ensemble mean. The resulting emission ensemble in this work shows ~35% uncertainty presented by the ensemble standard deviation. The correlation between emission perturbations of two grid point is modeled by a simple isotropic exponential decay function with a characteristic correlation length of 50 km.

We discuss the state augmentation approach for emissions in section 2.4 in detail. We described the ensemble construction in section 2.3. To better explain the emission perturbation, we added the text as below:

“We consider the NR as the true atmosphere and sample meteorological and NO₂ observations from the NR. The control run (CR) is a parallel model calculation to the NR and suffers from imperfect model input and parametrization. The differences between the NR and the CR in this study are the emission inputs and the initial conditions for the meteorology. ... In all experimental runs, we bias the CR initial emissions to be 30% below the reference model and examine the ability of the assimilation to recover the reference emissions.”

p5, l1: “We chose the 10 km distance based on sensitivity experiments” How are these experiments done?

We changed the text to the following: We perform univariate experiment by tuning the localization distance from 5km, 10km, 20km to 50km and keeping other settings in the system unchanged. By comparing the posterior emission RMSE of REF (described in section 3) of all experiments, we find the smallest RMSE with the localization distance of 10km.

p5, l26: “We calculate a layer dependent Box-Air Mass Factor (BAMF) : : : follow the latest version of the NASA : : :” It is unclear from the text how this is done. If the authors compute this themselves, then which RTM is used? How is the geometry computed? What about the terrain parameters. Please provide these details.

We didn’t calculate BAMF by running a RTM directly as it is not required when we follow the method in (Bucsela et al., 2013). “The scattering weights are computed and stored a priori in six-dimensional look-up tables (LUT) generated from a RTM. The six LUT parameters are solar zenith angle (SZA), viewing zenith angle (VZA), relative azimuth angle (RAA), terrain reflectivity (Rt), terrain pressure (Pt), and atmospheric pressure level, (p).”. We modified the text as below:

“The geometry related parameters (SZA, VZA and RAA) are computed hourly for each TEMPO observation using Matlab functions sun_position.m and geodetic2aer.m with inputs of the location and time of each TEMPO observation, and the location (36.5°N, 100°W) and altitude (35,786 km) of the TEMPO sensor. The terrain reflectivity and terrain pressure are sampled from the WRF-Chem nature run (NR, see section 3) for each TEMPO pixel.”

p6, l3: “The other parameters are sampled from the model run”. Which parameters?

Other parameters indicate terrain reflectivity and terrain pressure. We clarified this in the updated text.

p6, l14: The mean uncertainty of 7.5 % seems very optimistic. Is this justified somehow? Would the results be very sensitive to this choice?

We agree that it is very optimistic to set the TEMPO observation uncertainty as 7.5%, which is the lower bound of current observation error for OMI NO₂ tropospheric column product. This paper focuses on the optimal approach for NO_x emission estimations from TEMPO. It is difficult to separate systematic and random contributions to the uncertainty in TEMPO measurements. Future experiments should explore variations in systematic and random uncertainties in the measurements.

p6, l29: “scaled to be 70%” A uniform scaling is very idealized, and may be more easily recovered in a DA system than spatially-varying emission perturbations. It seems logical to add also a more random

perturbation in emissions to reflect the uncertainty in the emission spatial distribution. Would this impact the conclusions?

The conclusions in this work will not change by using a spatial-varying emission perturbation. We agree that a uniform emission perturbation is idealized. Success in recovering emissions with a uniform scaling can't not guarantee the success to recover emissions with spatially-varying perturbations. However, the approach that didn't work in the ideal case will fail in the realistic scenario with spatially-inhomogeneous emission perturbations. In our new experiments, we show the shortcomings of some emission estimation strategy in the scenario of ideal emission perturbation. We added the following text in section 5.

"In our OSSE, we start with an ideal case of a priori emissions, which is spatially-uniform 70% of true emissions. Success in recovering emissions with a uniform scaling can't not guarantee the success to recover emissions with spatially-varying perturbations. However, the approach that didn't work in the ideal case will fail in the realistic scenario with spatially-inhomogeneous emission perturbations. The degradation of emission estimation in the ENS.2 and ENS.3 will still hold for correcting emission errors under more complex circumstances, such as spatially- or temporally- varying emission errors. For future work, further analysis is required to examine whether spatially-varying emission errors can be reduced under the condition of low wind RMSE."

p7, l29: Please discuss explicitly the formula to compute the uncertainty. In line 31 the RMS is defined, but "uncertainty" is unclear.

We added the text as follows:

"We also analyze the uncertainty of the prior and posterior estimates. The uncertainty is expressed by the 1- σ standard deviation of the ensemble."

p8, l11: "initial uncertainty of 41.70 mol/(km²hr)". What is this absolute number? Is it an average over the red domain in Fig. 1?

The sentence is removed in our updated discussion section 4.

Fig.3: It would be helpful to indicate with symbols at what times an analysis is produced (for the green and blue curves).

Since we updated the experiments in this paper, the original Figure 3 is removed.

p8, l18: Please explain how this correlation is computed.

We removed the discussion of correlation in our updated paper.

P10, l3: "future TEMPO NO2 observations will enable us to constrain surface emissions on a city scale". I would claim that this is not fully demonstrated by the authors, because the OSSE setup is still idealized. In particular, the same model is used to construct reality (the nature run), and emissions are idealized, which will lead to too optimistic results. Please formulate the conclusions in a more careful way. Please also report on typical wind uncertainties which are within reach in present-day high-resolution regional meteorological analysis systems. This may be compared with RMS and uncertainty values as reported in table 2. Is a performance like reported for e.g. ENS-H feasible in reality?

We clarified our point as follows:

“We test the emission constraints from TEMPO NO₂ observations in an ideal case assuming no errors associated with the modeled meteorology. In the experiment of joint assimilation of meteorology and chemical NO₂, we find that the emission estimation is successful in the morning but degrades in the afternoon when the prior wind RMSE grows above 1 m/s. Considering the dependence of errors in estimated emissions on the wind forecast errors, we recommend to guarantee the accuracy in modeled wind with RMSE below 1 m/s for the success of chemical assimilation to infer emissions at the scale of our model grid.”

Reference:

Dee, Dick, Fasullo, John, Shea, Dennis, Walsh, John & National Center for Atmospheric Research Staff (Eds). Last modified 21 Jan 2016. "The Climate Data Guide: Atmospheric Reanalysis: Overview & Comparison Tables." Retrieved from <https://climatedataguide.ucar.edu/climate-data/atmospheric-reanalysis-overview-comparison-tables>.

Response to Reviewer 2:

We appreciate the referee's comments, which have improved our manuscript, as detailed below:

General comments: The focus of this paper is to test the ability of estimating emissions of NO₂ with data assimilation and future high resolution satellite instruments by doing an Observing System Simulation Experiment (OSSE), since the temporal and spatial variability of NO₂ are high and hourly observations at a resolution of at least 5km would be required. I think the title should include “using OSSEs”, otherwise it suggests that real data experiments have been carried out.

We updated the title as “Assimilation of satellite NO₂ observations at high spatial resolution using OSSEs”

The abstract and the introductory section are excellent. One of the main experimental conclusions is that the assimilation of meteorology and of NO₂ should be done simultaneously, supporting Kang et al. (2012) who pointed out the need to do simultaneous assimilation of the meteorology and atmospheric CO₂ and thus succeeded (in an OSSE framework) in estimating accurately the surface carbon fluxes even in the absence of any a priori information.

Sections 2-4 are not written clearly enough, in my opinion, at least not for a person who is not very familiar with the details of the subject, like this reviewer. In addition, much of the results revolve around an unrealistic experiment: having the TEMPO NO₂ observations at very high spatial and temporal resolution, and at the same time assimilating the winds every 12hr or even 24hr, which, not surprisingly, ruins the results. I would suggest to replace this with a more realistic experiment: Assume that you have access to a meteorology data assimilation performed every hour, so that you have access to the mean winds but not to the uncertainties. Are the wind uncertainties essential? Experiments that we have performed suggest that the wind uncertainties are not so important for CO₂, but they may be for NO₂ because of the smaller scales involved.

We have made major revisions in section 2~4. In section 2, we provide detailed description on the assimilation system including the forecast model, DART facility (with localization settings), initial and

boundary condition ensemble, emission update scheme and synthetic meteorological and chemical observations. In section 3, we redesign the assimilation experiment. First, we incorporated MADIS weather observations to increase the volume of wind observations. This enables us to perform hourly joint assimilation of meteorology and chemistry in the run named ENS.1. Following the reviewer's suggestion, we have REA assimilation run in which the meteorology is extracted from an existing reanalysis. REA reinitializes the meteorological state every hour with the best estimate of meteorological states generated by ENS.1. The result discussion and conclusion are also updated accordingly in section 4.

My recommendation is that this paper is important and could be accepted after major revision. Detailed comments:

P3, line 28: "After a spin-up time of 40 hours on the outer domain, the inner domain simulation is initialized and constrained through one-way nesting in both meteorology and chemistry". It is not clear whether this is in addition of the data assimilation, and for how long it is done.

This part only introduces the domain setup and how simulation on the outer domain serves for the inner domain. There is no assimilation on the outer domain. We describe the experiment details in section 3. The inner domain simulation is initialized on 06:00 am on 2014/07/02 and assimilation starts on 10:00 am 2014/07/02 (P8 line5). We extended the model spin-up time and modified the text as below:

"We use the global chemical model output from MOZART to initialize the chemical simulation on the outer domain and to provide the chemical boundary condition. After a spin-up time of four days on the outer domain, the inner domain simulation is initialized and constrained through one-way nesting in both meteorology and chemistry."

P4 line 1: "We use the RADM2 gas phase chemical mechanism for its simplicity (Stockwell et al., 1990)" I don't know what this means.

We provide descriptions of RADM2 chemical mechanism as below.

"We use the widely-used regional acid deposition model version 2 (RADM2) as the gas phase chemical mechanism. There are 59 species and 157 reactions to represent both inorganic and organic chemical reactions under tropospheric conditions. It includes the chemical losses of NO_x through reaction with OH radical to form nitric acid, and other NO_x sinks as peroxyacyl nitrates and alkyl nitrate. RADM2 predicts peak NO₂ concentrations and the timing that the peak occurs very well when the mechanism is tested against the representative environmental chamber experiments."

Line 12: Since Kang et al. (2011, 2012) were the first to estimate surface fluxes of carbon as evolving parameters, perhaps they should be referenced.

We modified the text as follows in section 2.4:

"By including emissions in the ensemble state vector, emissions are estimated as hourly evolving parameters. Estimation of time-evolving emissions using data assimilation is first presented in carbon fluxes estimation (Kang et al., 2011, 2012). Such an approach provides emission information more than a monthly mean, or an average for a specific time period. For NO_x emission estimation problem, emissions in cities show significant variation from urban to suburban. The observed columns show strong spatial

variation dominated by emission hotspot as a result of the short lifetime. Combined with high-resolution observations, the goal in this work is to constrain a time-static emission error as we perturb emissions of each hour in the same way.”

Line 19-24. This is an interesting persistence approach to estimate the emission forecast model. I am not sure whether it will always work in the presence of error spin down or if the surface fluxes vary substantially. It would be good to show how well it works in this case, perhaps under varying surface fluxes.

In all model runs, the surface fluxes are varying from hour to hour. In the OSSE, the emission bias in the control run is -30% uniformly though 24 hours. We agree that in reality, the errors in emissions could vary from hour to hour. So there may not be such a constant relative bias. However, the simplicity of this assumption makes it useful for this initial trial. We hope that other experiments with alternate assumptions will emerge in future research.

P5 line 6: “showed” not “suggest”

We changed to “showed”.

Lines 13-14: Are these simulated observations obtained from NARR?

The simulated meteorological observations are sampled from the nature run (NR, see in section 3). We described the simulated MADIS observations in detail in section 2.5.

Line 24: Does you model changes in the “truth” in the simulated TEMPO observations or are the fluxes fixed?

The NO_x emissions are specified by NEI 2011 dataset. This emission provides hourly-varying emission for a typical weekday in summertime. The fluxes does not vary from day to day. NEI 2011 is used as the emission input for nature run without any perturbation. We added this information in section 2.1:

“Anthropogenic emissions for WRF-Chem are from the National Emission Inventory (NEI) 2011 for a typical July weekday 30 at native 4×4 km resolution. The NEI 2011 provides hourly-varying emission for a typical weekday in summertime. The emissions do not vary from day to day.”

In section 3, we added as:

“The original NEI 2011 is used as the emission input for the NR without any emission perturbation. We begin by creating a NR and a CR simulation on the outer domain of 12 km resolution (d01) without assimilating observations using a simulation setup as described above in section 2.1. We impose a difference to the CR by using emissions in the CR that are scaled to be 70% of the NR emissions.”

P6 line 10: It would be good to show a companion figure showing the diurnal cycle.

The diurnal cycle of TEMPO NO₂ column is now shown in Figure 4.

Line 14 “a mean uncertainty of 7%, which is lower than the 35%...” unclear.

This sentence is not in the revised paper.

Line 19: AMF: I couldn't find its definition. In general there is a profusion of use of abbreviations whose original definition is difficult to find in the paper. I can deal with Control run (CR) and Nature Run (NR) but then new capitals appear (AR) defined somewhere else, and I get confused. If all the acronyms were defined in the same place it would be OK, but my pdf reader, when I search for the definition of AR, gives me hundreds of words that contain "ar".

We defined box- air mass factor as BAMF previously, but not AMF. We thank the reviewer for pointing out this issue. We provide the definition of abbreviations in the text as follows.

“Finally, as TEMPO is expected to be operational no sooner than 2018, it is reasonable to expect the retrieval error dominated by air mass factor (AMF) in polluted region will be reduced as a result of future improvements in AMF simulation”

We removed AR which confuses the reader. We define nature run (NR) and control (CR) as below:

“The original NEI 2011 is used as the emission input for the NR without any emission perturbation. We consider the NR as the true atmosphere and sample meteorological and NO₂ observations from the NR. The control run (CR) is a parallel model calculation to the NR and suffers from imperfect model input and parametrization. The differences between the NR and the CR in this study are the emission inputs and the initial conditions for the meteorology.”

P7 line 24: “the 17:00LT assimilation cycle each day” suggests that the assimilation may have been done for many days, and I really don't know yet whether this is true, or just one day was simulated (with constant true emission?).

In the updated experiments, we extended the assimilation to three days longer. We modified the text as below:

“We run assimilation experiments from 10:00 LT 2014/07/02 to 18:00 LT 2014/07/05 with an assimilation window of one hour.”

Along these lines, it would be really strange for a real life researcher to use detailed TEMPO observations every hour and then do atmospheric assimilation only every 12 or even 24 hours. The experiments ENS-O and ENS-T are not realistic (and have horrible acronyms). I would drop these experiments. The results shown in Figure 3 are very obvious: when you assimilate wind observations the uncertainty in the winds becomes much smaller, and the transport errors likewise.

As indicated above, we redesigned the assimilation experiment as described in section 3. First, we incorporated MADIS weather observations to greatly increase the volume of wind observations. This enables us to perform hourly joint assimilation of meteorology and chemistry in the run named ENS.1. Also we add ENS.2 to determine whether or not concentrations should be updated when observations are assimilated to constrain unobserved emissions. In ENS.3, we use the meteorology of the next day to initialize the CR ensemble so that there is some difference between the CR ensemble mean and the NR in the meteorology. REA reinitializes the meteorological state every hour with the best estimate of meteorological states generated by ENS.1. We hope the reviewer finds the strategy and acronyms an improvement.

P8 line 7 The acronym of “No meteorological observations assimilated” should NOT be called “BIAS-MET”, the errors are much more than a bias!

We removed this run with no meteorological observations assimilated in the new experiment design.

Fig 3: Label “Local time (hours)”

We added “local time (hours)” in Figure 3~6.

Fig 4: Labels: “Longitude” “Latitude” in the figure, not in the figure caption.

The original Figure 4 is no longer in our updated paper.

Assimilation of satellite NO₂ observations at high spatial resolution using OSSEs

X. Liu¹, A. Mizzi², J. Anderson³, I. Fung¹ and R.C. Cohen^{1,4}

¹Department of Earth and Planetary Science, University of California at Berkeley, Berkeley, CA, USA

5 ²Atmospheric Chemistry Observation and Modeling Laboratory, National Center for Atmospheric Research, Boulder, CO, USA

³Institute for Mathematics Applied to Geosciences, National Center for Atmospheric Research, Boulder, CO, USA

⁴Department of Chemistry, University of California at Berkeley, Berkeley, CA, USA

10 *Correspondence to:* Ronald C. Cohen (rccohen@berkeley.edu)

Abstract. Observations of trace gases from space-based instruments offer the opportunity to constrain chemical and weather forecast and reanalysis models using the tools of data assimilation. To date attempts at assimilation of nitrogen dioxide (NO₂) satellite remote sensing data have focused on updating emissions and concentrations. These initial efforts evaluated updates at length scales of ~100 km using once a day measurements from satellites with ground pixels of 13 km × 24 km or
15 larger. In the boundary layer, NO₂ has a lifetime of order five hours and corresponding e-folding concentration variations near urban and point sources occur on spatial scales of order 50-75 km. Accurate observation and modeling of these variations require spatial resolution of order 4 km. In addition because of the short lifetime, NO₂ variations are more strongly coupled to short time scale meteorological parameters than longer lived chemicals such as CO or CO₂. In the next few years, we anticipate the launch of several instruments with ~3 km spatial resolution. In addition, some of these instruments will be
20 in geostationary orbits and thus have hourly revisit times. In anticipation of these instruments, we investigate the potential of high space and time resolution column measurements to serve as constraints on urban NO_x emissions using a geostationary observation simulator coupled to a data assimilation system. Our results show the dependence of errors in estimated emissions on the wind forecast errors. Modeled wind vectors must have RMSE below 1 m/s for the success of chemical assimilation to infer NO_x emissions at the scale of the model grid with RMSE less than 30 mol/(km²·hr). We demonstrate
25 that the simultaneous update of NO_x emission and concentrations outperforms the approach of updating emissions only. We recommend carrying out meteorology assimilations to stabilize from the initial transport errors before starting the emission inversion. Last but not least our results show that the information of wind *uncertainties* is not important for NO_x emission estimation. Chemical assimilations of NO₂ only with the meteorology extracted from an hourly weather assimilation analysis would perform as well as the joint assimilation of meteorology and chemistry.

30

1 Introduction

Weather and climate act in concert with emissions to establish the concentrations of chemicals and aerosols in the boundary layer. To improve forecasts of exposure and understand the factors that affect public health and the productivity of agriculture and animal husbandry, we require accurate models of both emissions and the boundary layer meteorology to define the surface layer concentrations that determine the exposure of humans, animals and plants. There remain substantial uncertainties in even the best models of emissions and even more so in the best models of boundary layer dynamics (for example, Hu et al., 2010). Current uncertainties in the surface NO₂ emission inventories in the U.S. are thought to be of order 50% (Krotkov et al., 2016; Travis et al., 2016). Comparable uncertainties affect estimates of the planetary boundary layer (PBL) height and mixing rates that redistribute emissions from the surface (Kretschmer et al., 2012, 2014; Lauvaux and Davis, 2014).

Over the last decade, data assimilation techniques have been used to constrain model forecasts and reanalyses of atmospheric constituents (Arellano Jr. et al., 2007; Edwards et al., 2009; Claeysman et al., 2011; Lahoz et al., 2012; Pagowski and Grell, 2012; Bowman, 2013; Gaubert et al., 2014; Hache et al., 2014; Saide et al., 2014; Zoogman et al., 2014; Barré et al., 2015; Bousserez et al., 2016; Mizzi et al., 2016). Assimilation of chemicals can be extended to optimize model inputs, such as emissions thereby providing insight into how to improve the processes that govern the model performance (Elbern et al., 2007; Barbu et al., 2009; Chatterjee et al., 2012; Miyazaki et al., 2012b; Koohkan et al., 2013; Yumimoto, 2013; Cui et al., 2015; Guerrette and Henze, 2015; Turner et al., 2015).

To date most efforts to incorporate satellite remote sensing in data assimilation have focused on long-lived chemicals such as CO, CH₄ or CO₂ and regional and continental scale aspects of emissions. Processes that govern variability of emissions within an urban center require new approaches that use high spatial and temporal resolution models and observations. NO₂ has a lifetime of only a few hours and thus exhibits concentration changes that are substantial on spatial scales of 50-75 km. Observations of variations in NO₂ are thus uniquely suited to study emissions and meteorology at the scales of cities. Averaged measurements of NO₂ have been shown to be promising for evaluation of absolute emissions and trends (Russell et al., 2012; Miyazaki et al., 2016) as well as providing information on the coupling of boundary layer winds to chemical lifetime (Beirle et al., 2011; Valin et al., 2013). Current space based instruments have resolution that is too low to provide direct information on lifetimes and emissions from a single overpass. Instead, analyses have focused on averages of the data that wash out some of the key details about emission location and chemical lifetime.

New instruments with spatial resolution of a few kilometers will soon change that situation. The TROPOspheric Monitoring Instrument (TROPOMI, launch date of 2017) will be the first to provide spatial resolution sufficient to observe these NO₂ changes on a single overpass. TROPOMI will view the atmosphere from low earth orbit and provide one image per day. We also anticipate the launch of three geostationary satellites, the Geostationary Environmental Monitoring Spectrometer (GEMS), the Tropospheric Emissions: Monitoring of POLLution (TEMPO) and Sentinel-4, which will provide observations at higher temporal resolution with hourly repeat at locations in Asia, North America and Europe, respectively (Zoogman et al.,

2017). The spatial resolution of these new low earth orbit (LEO) and Geostationary (GEO) instruments will be sufficient to provide ~10 samples within the advection distance that is determined by the chemical lifetime of NO₂. This dense sampling will permit characterization of multi-exponential or non-exponential behavior where current analyses are typically forced to assume single exponential decay. To take full advantage of these measurements within a data assimilation system, we will need to model the NO₂ column at similar spatial resolution. This is both because the spatial scales of important variation in atmospheric plumes are on the order of 4 km and because of the steep non-linearity in the lifetime of NO₂ as a function of the NO₂ concentration. For example, biases of 34% (3.3 to 5.0×10^{15} molecules/cm²) are found in the modeled averaged NO₂ column over Los Angeles at resolutions of 96 km compared to 12 km. For a point source, such as a power plant, model convergence is observed only at a grid resolution of 4 km or smaller (Valin et al., 2011).

In this study, we describe a high spatial and temporal resolution chemical transport ensemble data assimilation system with joint assimilation of meteorology and chemistry to adjust NO_x emissions on scales consistent with the temporal scale of NO_x evolution. We use that forecast/assimilation system to investigate the factors that influence the capability of TEMPO NO₂ observations to accurately constrain NO_x emissions. Our long-term goal is to estimate hour-to-hour variations in NO_x emissions at the scale of model grid point resolution (3 km) and to use these variations to understand the processes controlling the emissions. The remainder of this paper is organized as follows: In section 2, we describe the forecast/data assimilation system, the system setup, observations, and the TEMPO NO₂ simulator – the simulation of column NO₂ that would be observed by TEMPO. In section 3, we describe the experimental design including a series of assimilation experiments that guide optimization of the emission estimation performance. In section 4, we assess the performance of meteorology and chemistry assimilation. We then discuss the results and provide insight into the potential accuracy of NO₂ emission fields derived from geostationary NO₂ observations. We conclude in section 5.

2 The data assimilation system

The forecast/data assimilation system used here is WRF-Chem/DART as described by Mizzi et al. (2016). It consists of the following elements: the forecast model, the assimilation engine, and observations of meteorological and chemical states to be assimilated.

2.1 WRF-Chem model description

The core meteorological and chemical forecast model is the regional online chemical transport model WRF-Chem v3.4.1 (www2.acd.ucar.edu/wrf-chem). The model domain is a one-way nest with an outer domain of 12 km resolution covering western North America and an inner domain of 3 km resolution focused on the city of Denver, CO (Figure 1). The 3 km domain is 660 km by 840 km. The model has 30 vertical levels between the surface and an upper boundary of 100 mb and 10 levels within the boundary layer (~1.5 km). Simulations of meteorology on the outer domain are initialized and constrained at the lateral boundary by North American Regional Reanalysis (NARR) data from National Centers for Environmental

Prediction (NCEP). The NARR data have a native horizontal resolution of 32 km with 45 pressure levels and 3 h temporal resolution. We use the global chemical model output from MOZART to initialize the chemical simulation on the outer domain and to provide the chemical boundary condition. After a spin-up time of four days on the outer domain, the inner domain simulation is initialized and constrained through one-way nesting in both meteorology and chemistry.

- 5 Anthropogenic emissions for WRF-Chem are from the National Emission Inventory (NEI) 2011 Version 1 at native $4\times 4\text{ km}^2$ resolution. The NEI 2011 provides hourly-varying emission for a typical weekday in summertime. The emissions do not vary from day to day. Biogenic emissions are calculated online with the simulation results by Model of Emissions of Gases and Aerosols from Nature (MEGAN). Fire emissions are not included. We use the widely-used regional acid deposition model version 2 (RADM2) as the gas phase chemical mechanism (Stockwell et al., 1990). There are 59 species and 157
- 10 reactions to represent both inorganic and organic chemical reactions under tropospheric conditions. It includes the chemical losses of NO_x through reaction with OH radical to form nitric acid, and other NO_x sinks as peroxyacyl nitrates and alkyl nitrate. RADM2 predicts peak NO_2 concentrations and the timing that the peak occurs very well when the mechanism is tested against the representative environmental chamber experiments.

2.2 DART ensemble assimilation system

- 15 WRF-Chem/DART (data assimilation research testbed) is a regional multivariate data assimilation system developed by the National Center for Atmospheric Research (NCAR) to analyze meteorological variables and chemical variables simultaneously (Mizzi et al., 2016). We use the Ensemble Adjustment Kalman Filter (EAKF) in DART to analyze the states with an ensemble size of 30. Details of the EAKF algorithm and its implementation in DART are documented in (Anderson, 2001; Anderson and Collins, 2007; Anderson et al., 2009). In this study the system is extended to assimilate synthetic
- 20 TEMPO NO_2 column observations. As emissions are not prognostic variables of the forecast model, we implement a state augmentation approach to include emissions in the state variables (Aksoy, 2006). The chemical state variables include the NO_2 concentration, NO_x emissions. The meteorological state variables are U, V, W, T, QVAPOR, QCLOUD, QRAIN, QICE, QSNOW (MU and PH are used in vertical coordinate transforms, T2, Q2, U10, V10, PSFC used for surface data assimilation forward operators.) based on the settings used in meteorology data assimilation (Romine et al., 2013). Adaptive
- 25 spatially and temporally varying inflation is applied to the prior state to assist in maintaining ensemble spread. We summarize the DART configuration details in Table 2.

2.2.1 Spatial localization

- In ensemble methods the correlations among spatially remote variables in the prior ensemble are regarded as spurious correlations due to the small ensemble size (30). To compensate for this under-sampling issue, spatial localization is
- 30 introduced to reduce the prior correlations based on the distance between the observed/modeled state variables (Houtekamer and Mitchell 2001). In this study, we apply the fifth-order distance-dependent Gaspari and Cohn (GC) function (Gaspari et al., 1999) to reduce the spurious impact of observations on spatially remote state variables. The scaling distance in the GC

function is defined by a half-width parameter, two times which is the distance where the GC function goes to zero. With a data assimilation window of one hour and a maximum wind speed of 3~5 m/s, an observation of column NO₂ contains information about emissions that occurred during the last hour within 10 km. We use the half-width distance in spatial localization as 10 km and demonstrate this as the optimal value based on sensitivity experiments with localization distance of 5 km, 10 km, 20 km and 50 km. Because of the high density of TEMPO NO₂ observations (2×4.5 km²), the update of chemical state variables is mostly determined by the local observations.

2.2.2 Variable localization

Similar to the concept of spatial localization, variable localization techniques have been introduced (Arellano Jr. et al., 2007) to reduce spurious correlations among observations and different types of state variables. For example, for CO₂ flux estimation, Kang et al. (2011) showed that the performance of data assimilation using a variable localization that zeros out the prior error covariance between meteorological variables and CO₂ flux is better than using a standard full covariance approach. Here we isolate the influence of meteorological observations on chemical variables and vice versa. In our next paper we will discuss the potential benefit from assimilations of chemical observations to improve the meteorological fields.

2.3 Initial and boundary condition ensemble

We generate the initial chemical ensemble by adding the perturbations to the mean state of the fine domain forecast. In the ensemble method the generated ensemble should represent the error statistics of the initial guess of the model state (Evensen 2003). The correlation between perturbations of chemical state variables is modeled by a simple isotropic exponential decay function with a characteristic correlation length of 50 km. For meteorology ensemble, random perturbations were added to each member by sampling the NCEP background error covariance using WRF Data Assimilation System (<http://www2.mmm.ucar.edu/wrf/users/wrfda>) (Barker et al., 2012). The options in WRFDA setting is summarized in Table 3. The ensemble member lateral boundary condition perturbations are generated in a similar manner as the initial ensemble using the fixed-covariance perturbation technique. The boundary condition for the analysis time is adjusted to match the analysis from DART. The tendencies for the later times in the forecast are adjusted to match the change in the boundary condition for the analysis time.

2.4 Emission update scheme

By including emissions in the ensemble state vector, emissions are estimated as hourly evolving parameters. Estimation of time-evolving emissions using data assimilation is first presented in carbon fluxes estimation (Kang et al., 2011, 2012). Such an approach provides emission information more than a monthly mean, or an average for a specific time period. For NO_x emission estimation problem, emissions in cities show significant variation from urban to suburban. The observed columns show strong spatial variation dominated by emission hotspot as a result of the short lifetime. Combined with high-resolution

observations, the goal in this work is to constrain a time-static emission error as we perturb emissions of each hour in the same way.

A challenge for updating the emissions in the augmented state vector is the absence of an emission forecast model to evolve the emission variables forward in time. The bottom-up inventory to be optimized provides hourly-resolved emissions for each model grid point. Instead of treating the emission variables of each hour at a specific location as independent parameters, we update the emission scaling factors at each assimilation cycle. For a given model grid point, we define the emissions of truth (e_i^t), prior (e_i^{prior}) and posterior (e_i^{post}) at the i th cycle. Since we start the assimilation with 70% of true emissions, we have $e_i^{prior} = 0.7e_i^t$, for each cycle. After one assimilation cycle, we compute the scaling factor (S_i) as follows: $S_i = e_i^{post} / e_i^{prior}$. Then we update the prior emissions at the next cycle as $e_{i+1}^{prior} = S_i \cdot e_i^{prior}$. This prescription enables us to derive spatial 2-D emission scaling factors which play the role of an emission forecast model.

2.5 Synthetic meteorological and chemical observations

Observations assimilated in the system include meteorological observations and NO₂ trace gas retrievals from the TEMPO OSSE. For meteorological observations, we assimilated synthetic observations of temperature, wind and humidity from the NCEP Meteorological Assimilation Data Ingest System (MADIS) (<https://madis.noaa.gov/>). MADIS is a meteorological observational database and data delivery system that provides observations that cover the globe. MADIS ingests data from NOAA data sources and non-NOAA providers, decodes the data then encodes all of the observational data into a common format with uniform observational units and time stamps. For wind observations, the assimilated observation types include standard aviation routine weather report (METAR), wind profilers, aircraft-based observations (ACARS), national mesonet data and satellite data. Among them the mesonet wind data is the most abundant with ~1000 observations located in the mapping domain in Figure 2. The observation errors are the default values from the DART facility that are defined based on NCEP statistics (Romine et al., 2013)

The GEOstationary Coastal and Air Pollution Events (GEO-CAPE) mission (Fishman et al., 2012) aims at improving our understanding of both coastal ecosystems and air-quality from regional to continental scales. As the first phase of the GEO-CAPE implementation, TEMPO (Zoogman et al., 2017), launch date circa 2019, will provide hourly measurements of NO₂, HCHO, tropospheric ozone, aerosols, and cloud parameters during the daytime. TEMPO will measure solar back scattered light in the UV-Vis spectral range. Implemented on a geostationary platform, TEMPO retrievals will achieve hourly observations of NO₂ vertical column density (VCD) at a native spatial resolution of 2×4.5 km during the day-lit period. TEMPO's high spatiotemporal resolution will allow a more detailed assessment of emission inventories, e.g. urban scale and large power plant NO₂ emissions and mobile emissions that show significant spatial and temporal variations due to urban transit patterns, than is possible with existing LEO observations.

As the TEMPO has not been launched yet, we generate synthetic TEMPO NO₂ observations by simulating the instrument's observing characteristics. We carried out a model run, i.e. a forward integration of WRF-Chem for the period from July 2nd

to July 7th 2014 with NO₂ emissions specified by NEI 2011 ('truth'). We calculate a layer dependent Box-Air Mass Factor (BAMF) representing the sensitivity of the retrieved NO₂ in a specific layer to the true value in the atmosphere. The BAMF of NO₂, as an optically thin absorber, is a vector and determines the measurement sensitivity to NO₂ molecules at 35 pressure levels. In the calculation of BAMFs, we follow the latest version of the NASA standard product retrieval (Level 2, Version 2.1, Collection 3) algorithm (Bucsela et al., 2013) assuming the TEMPO measurement has similar characteristics to OMI. We assume clear-sky conditions for all observing scenes. Cloudy-sky scenes affect only the number of observations available as the cloudy scenes are usually discarded in the data filtering process. The elements of the BAMF vector are computed as a function of solar zenith angle (SZA), viewing zenith angle (VZA), relative azimuth angle (RAA), terrain reflectivity (Rt), terrain pressure (Pt), atmospheric pressure level, (p) and the NO₂ profile (Bucsela et al., 2013). The viewing parameters are computed by simulating viewing geometry based on the location of ground pixels in relation to the observing instrument. The geometry related parameters (SZA, VZA and RAA) are computed hourly for each TEMPO observation using Matlab functions `sun_position.m` and `geodetic2aer.m` with inputs of the location and time of each TEMPO observation, and the location (36.5°N, 100°W) and altitude (35,786 km) of the TEMPO sensor. The terrain reflectivity and terrain pressure are sampled from the WRF-Chem nature run (NR, see section 3) for each TEMPO pixel. All the parameters have an hourly frequency consistent with the TEMPO temporal observing pattern. Consequently, the NO₂ profile with high-spatial-temporal resolution captures the diurnal variation of NO₂ and its urban-rural contrast. This contrast is essential to accurate interpretation of the measured spectrum (Russell et al., 2011; Laughner et al., 2016).

To generate synthetic TEMPO data, the modeled 3-D concentration fields from the NR are sampled in as similar a manner to the planned TEMPO measurements as the transport model permits: using the computed BAMF vertically; hourly frequency; 2×4.5 km nadir resolution and variations following the Earth's curvature horizontally. Figure 2 shows an example of the spatial distribution of TEMPO data over Denver, CO.

We describe the observation error as a relative value (σ_{rel}) and a random draw from a Gaussian distribution to avoid using a fixed value. The magnitude of the mean uncertainty of the NO₂ column is different between clean and polluted areas (Boersma et al., 2004). We follow their categorization of clean versus polluted regions and summarize the mean and standard deviation of a Gaussian distribution for each scenario in Table 4. For polluted regions, we give mean uncertainty of 7.5%, which is lower than the 35% minimum in the OMI NO₂ retrievals. First, most of these errors are systematic affecting comparison of different cities but have smaller variation across a single, small area scene of observations. Second, a relatively lower observation error improves the efficiency of data assimilation and helps to examine the sensitivity to other parameters. Finally, as TEMPO is expected to be operational no sooner than 2018, it is reasonable to expect the retrieval error dominated by air mass factor (AMF) in polluted region will be reduced as a result of future improvements in AMF simulation (Laughner et al., 2016). The synthetic observations assimilated are obtained by sampling the NR using the TEMPO observation simulator and adding observation error as $y^{obs} = N(y^{tr}, \sigma^2)$, where y^{tr} is the TEMPO NO₂ observations sampled from the truth, and σ is the observation error standard deviation computed as $\sigma = y^{tr} \cdot \sigma_{rel}$.

3 Assimilation experiments

We begin by performing observing system simulation experiments (OSSEs) in the context of a perfect model. The original NEI 2011 is used as the emission input for the NR without any emission perturbation. We consider the NR as the true atmosphere and sample meteorological and NO₂ observations from the NR. The control run (CR) is a parallel model calculation to the NR and suffers from imperfect model input and parametrization. The differences between the NR and the CR in this study are the emission inputs and the initial conditions for the meteorology. We begin by creating a NR and a CR simulation on the outer domain of 12 km resolution (d01) without assimilating observations using a simulation setup as described above in section 2.1. We impose a difference to the CR by using emissions in the CR that are scaled to be 70% of the NR emissions. We apply the identical forecast model (WRF-Chem) for both the NR and the CR to isolate the behavior of the ensemble filter algorithm from the influence of the model errors. Then the NR and the CR on the inner domain of 3km (d02) are initialized from the corresponding d01 simulations respectively on 06:00 local time (LT) on July 2nd 2014. At the time of initialization, the NR and CR on d02 share the same meteorological fields and differ in NO_x concentrations due to different emission inputs. Our next step is to generate a 30-member ensemble from the CR. We use WRFDA to generate an ensemble in meteorological variables (Barker et al., 2012). For chemical states, we give an ensemble in NO_x emissions and concentrations using the method described above in section 2.3. The forecast of the CR ensemble is the prior estimate of the states and will be combined with the observations in the assimilation cycle to yield the posterior states. By comparing the posterior emissions with the “true emission”, we evaluate the data assimilation performance. We run assimilation experiments from 10:00 LT 2014/07/02 to 18:00 LT 2014/07/05 with an assimilation window of one hour. We assimilate ~20,000 weather observations in each assimilation window and ~9,000 TEMPO NO₂ column observations in each daytime assimilation window.

We design a series of experiments to explore the optimal approach to estimate NO_x emissions as shown in Table 1. In all experimental runs, we bias the CR initial emissions to be 30% below the reference model and examine the ability of the assimilation to recover the reference emissions. First, a reference assimilation run (REF) is conducted without including the meteorological ensemble so that the NR and CR ensembles have identical meteorological simulations. This shows the best case scenario to constrain emissions assuming no errors associated with meteorology. In practice, the modeled meteorology is different from the true atmosphere due to errors in the model initial conditions, parameterizations and resolutions. In a real case labelled as ENS, we initialize an ensemble in both meteorology initial conditions and NO_x emissions in which both weather observations and TEMPO NO₂ columns are assimilated. In ENS.1 the CR ensemble is generated by adding perturbations to the CR mean state, so the CR ensemble mean meteorology is very close to the NR because CR and NR differ in NO_x emissions only. For the chemistry part, the assimilated TEMPO NO₂ observations will update both NO₂ concentration and NO_x emissions every hour. In ENS.2 we only allow NO₂ observations to update NO_x emissions and keep the meteorology assimilation the same as ENS.1. By comparing ENS.1 and ENS.2 we can determine whether or not concentrations should be updated when observations are assimilated to constrain unobserved emissions. In ENS.3, we use

the meteorology of the next day to initialize the CR ensemble so that there is some difference between the CR ensemble mean and the NR in the meteorology. To be specific, the CR meteorology ensemble on 2014/07/03 9:00 LT is used as the CR ensemble on 2014/07/02 9:00 LT. This is to mimic our imperfect knowledge about the atmosphere state and its uncertainty in reality. ENS.1 and ENS.3 differ only in the meteorology initial ensemble and by comparing these two runs, we can see the sensitivity of the NO₂ assimilation to the meteorology initialization. Our final experiment REA mimics a general approach in chemistry-only data assimilation where the meteorology is extracted from an existing reanalysis. REA reinitializes the meteorological state every hour with the best estimate of meteorological states generated by ENS.1. As a result, REA has a single run of meteorology but still an ensemble of NO₂ emissions and concentrations which allow it to assimilate TEMPO NO₂ observations. Same as ENS.1, REA updates emissions and concentrations simultaneously.

4 Results

We evaluate the assimilation result by comparing with the NR states. We calculate the root mean square errors (RMSE) of observed quantities by $\sqrt{\sum_i^n (y_i^m - y_i^t)^2 / n}$, where y_i^m and y_i^t are the model and true values for the i th observation respectively, and n is the total number of observations of interest. We also calculate the RMSE of model states by $\sqrt{\sum_i^n (x_i^m - x_i^t)^2 / n}$, where x_i^m and x_i^t are the model and true states at the i th model grid point respectively, and n is the total number of grid points of interest. For the wind variable, the grid points of interest are all the points located within a sub-model space as shown in Figure 2, containing the lowest 10 model levels vertically. Because NO_x is located mostly in the boundary layer, the NO₂ transport error is determined by the meteorological errors in the lowest 10 model levels. For NO_x emission variables, the grid points of interest are categorized as emission points with emissions greater than 50 mol/(km²·hr). Our analysis does not include emissions below 50 mol/(km²·hr) because the observations over such low emission regions have large uncertainty and are not constrained. We also analyze the uncertainty of the prior and posterior estimates. The uncertainty is expressed by the 1- σ standard deviation of the ensemble.

4.1 Wind assimilation

The success of ensemble-based assimilation relies on how well the ensemble system represents the uncertainty. One way to test the success of an OSSE is to compare the RMSE computed with respect to the “true” observations with the ensemble spread directly. Figure 3 shows the evolution of the RMSE and spread for mesonet observations of zonal wind for ENS.1 and ENS.3. Overall, for each experiment the variation and magnitude of prior ensemble spread are similar to those of the prior RMSE, indicating that the ensemble develops a good amount of spread for the success of OSSE.

We find the errors in the observation space of mesonet winds are reduced by 50% on average from the prior to the posterior. The prior wind RMSE shows peaks in the afternoon and this results in the most significant error reduction over the time period. The posterior wind RMSE shows a temporal average of 0.39 m/s and 0.47 m/s in ENS.1 and ENS.3 respectively. Because ENS.1 is initialized with a meteorology ensemble with its mean close to the truth, the wind RMSE on the first day is

low and gradually grows to about 1 m/s. In contrast, the prior wind RMSE in ENS.3 is as high as 2 m/s on the first day as a result of using an initial meteorology ensemble that is very different from the truth. The wind RMSE evolution in the two experiments becomes very similar after the afternoon of 2014/07/04. We conclude that the ensemble wind assimilation system performance is independent of the initialization approach after the first day.

5 4.2 TEMPO NO₂ assimilation

We assimilate hourly TEMPO NO₂ column observations and take their difference with the modeled column to correct the predicted NO₂. Figure 4 shows the TEMPO NO₂ column RMSE evolution for all experiments. With perfect knowledge of meteorology, REF shows significant reduction in TEMPO NO₂ RMSE in the first three update cycles, and succeeds in recovering the true emissions (Figure 5). The prior TEMPO NO₂ RMSE in the last three days varies below 3×10^{14} molecules/cm² as a result of perfect NO₂ transport and improved emissions. This ideal case with the assumption of perfect meteorology sets the upper limit of error reduction in NO₂ concentrations by assimilating the TEMPO NO₂ observations. Compared with REF, ENS.1 shows prior TEMPO NO₂ RMSE of $5 \sim 10 \times 10^{14}$ molecules/cm² due to the errors in NO₂ transport and emissions. By assimilating NO₂ observations, the TEMPO NO₂ RMSE is reduced by more than 50% from the prior to the posterior indicating the potential of TEMPO NO₂ observations to improve the modeled atmospheric NO₂ composition for the chemical reanalysis product. Without updating the NO₂ concentrations in ENS.2, there is no reduction in the TEMPO NO₂ RMSE as expected. We find the TEMPO NO₂ RMSE varies above 1×10^{15} molecules/cm², being the largest among all experiments because the emission estimations show very poor results (shown in section 4.3). The TEMPO NO₂ RMSE development in ENS.3 is very similar to ENS.1 except for the first day when ENS.3 shows higher errors in the wind field, which contribute to the NO₂ transport errors. We find the NO₂ forecast using a single meteorology field in REA is very similar to the ensemble NO₂ forecast in ENS.1. This is because there is very little difference between the one-hour meteorology forecast and the ensemble forecast. In addition, the emission estimation results are also very similar. This is different from the previous study on CO₂ forecast which shows that for a 6-hour forecast, the CO₂ transport driven by a single meteorological field has weaker vertical mixing and stronger CO₂ vertical gradient when compared to the mean of the ensemble CO₂ forecasts initialized by the ensemble meteorological field (Liu et al., 2011).

We compare the TEMPO NO₂ column spread in REF and ENS.1 in Figure 5. For both experiments, the prior NO₂ column spread varies at the same magnitude with the prior RMSE (Figure 4), indicating good quality of the NO₂ ensemble spread. The NO₂ forecast uncertainty represented by the NO₂ ensemble spread results from the uncertainties in NO₂ transport and emissions when the uncertainties in chemical production and removal processes are not included in this study. The uncertainties in NO₂ transport are determined by the prior wind ensemble spread, which shows peaks in the afternoon and stays as low as ~ 0.5 m/s at other times for zonal wind (Figure 3). The prior NO_x emission uncertainties are 60% after inflation (Figure 6). Under these circumstances, the mean prior TEMPO NO₂ column spread is 4.55×10^{14} molecules/cm² in REF which does not include NO₂ transport uncertainties, and is 7.03×10^{14} molecules/cm² in ENS.1 which takes uncertainties in transport and emissions into account. So the NO₂ transport contributes to 35% of the total NO₂ forecast uncertainties in

our assimilation setup. The TEMPO NO₂ column spread in REF is very stable because it is determined by the constant emission spread of 60%. ENS.1 shows fluctuations in the evolution of TEMPO NO₂ column spread which corresponds to the wind spread variation with increasing spread in the afternoon.

4.3 NO_x emission estimation

5 In our emission update scheme, the TEMPO NO₂ observations at time t are assimilated to generate a scaling factor for emissions at $t-1$ hour. This scaling factor is used to correct the prior emission at time t . In this way, the model-observation difference in the NO₂ column will correct the emission of an hour ago instead of the current emission. This approach is reasonable because errors in NO₂ concentration result from errors in previous emissions. Considering the short NO₂ lifetime of three hours in summer daytime, emissions from the previous hour have a large contribution to the NO₂ total mass at the
10 current time.

We show the time evolution of the averaged city emissions for all experiments in Figure 6. For all experiments, the posterior emission ensemble spread is reduced compared to the prior spread, suggesting the effectiveness of assimilated NO₂ columns in constraining the emission uncertainties. We should mention that we ignore the emission correction of the first assimilation cycle. We notice that the first update produces a significant over-correction on emission because of the accumulated
15 underestimation of NO₂ from the previous time. By neglecting the first update, the prior emission ensemble mean of the second cycle is still 70% of the truth. During the nighttime when TEMPO observations are not available, we calculate the ratio of posterior ensemble mean to the truth in the last cycle of daytime and use this together with the nighttime true emissions to derive the ensemble mean for the nighttime emissions. The prior and posterior emission ensemble of each nighttime hour is the same.

20 Not surprisingly, under the condition of perfect knowledge in meteorological fields, assimilating TEMPO NO₂ observations successfully improves the emissions within the first few updates. The estimated emissions agree well with the true emissions throughout the assimilation period. This demonstrates the capability of a geostationary NO₂ column observing system to constrain city-scale emissions and the reliability of the ensemble-based assimilation method to project the observed information to emissions.

25 We find the errors in estimated emissions correlate with the wind errors. In ENS.1, the posterior emission is corrected to the truth at the second cycle and stays close to the truth throughout the first day. The good performance on the first day benefits from an initial meteorology ensemble with its mean close to the truth. For the following three days, the emission estimates succeed in recovering the true emissions during the morning and show deviations from the truth in the afternoon as a result of the increased error in boundary layer winds. Figure 7 shows the dependence of errors in the inverted emissions to the
30 prior wind RMSE. The emission errors show high sensitivity to the wind errors with the slope of the regression line of $32.5 \text{ mol}\cdot\text{km}^{-2}\text{hr}^{-1}/(\text{m}\cdot\text{s}^{-1})$. With the RMSE of model predicted wind vectors of 1 m/s, the errors in the estimated emissions are $30 \text{ mol}\cdot\text{km}^{-2}\text{hr}^{-1}$ on average. For the daytime cycles, the prior emission ensemble spread after inflation is approximately 60%

and is reduced by more than half after assimilation (Figure 6). Even though the posterior ensemble mean does not match with the truth in the afternoon, the truth falls within the range of the posterior ensemble spread with a few exceptions. We find the simultaneous update of emission and concentration performs better than the emission update only scheme with an hourly assimilation window. ENS.2 is a parallel assimilation run with ENS.1 but updates emissions only. As shown in Figure 6, the estimated emissions have very large differences from the truth and the posterior ensemble spread does not cover the truth. For example, at 10:00 July 3rd, the posterior ensemble mean (red) is very close to the truth. As a result of this, we have a very good prior ensemble estimate (black) at 11:00. However, the posterior emission at 11:00 is largely underestimated compared with the truth. This is because the posterior emissions from 7:00 to 9:00 are overestimated which results in overestimated NO₂ concentrations at 10:00 and 11:00. As a result, even though the prior emissions from 10:00 to 11:00 are good, the model still overestimates NO₂ at 11:00 due to the NO₂ overestimation at 10:00. Without updating the concentrations, the observed differences in NO₂ columns are dominated by the NO₂ concentration errors of an hour ago and should not be attributed to the emissions.

We also find that the emission estimation should start after the meteorology assimilation becomes stable. As a comparison to ENS.1, ENS.3 is initialized with a meteorology ensemble that is very different from the truth. On the first day, the prior wind RMSE varies from 1 m/s to 2 m/s (Figure 3) and leads to enhanced NO₂ transport errors. As a result, the emission estimations are not successful for the first day. After the afternoon of the second day (07/03), the wind RMSE evolution is similar between ENS.1 and ENS.3 and as a result, the emission estimations perform in a similar way. We recommend carrying out meteorology assimilations to stabilize from the initial transport errors before assimilating chemical observations to constrain the emissions.

With an hourly re-initialization of meteorology, the NO₂ transport error statistics are not important in emission estimation in the current practice of a single meteorological field to transport NO₂. The emission estimation performance in REA is very similar with that in ENS.1 (Figure 6). This is because the difference in the one-hour NO₂ forecast driven by an ensemble meteorological field and a single ensemble mean field is very small. Though the wind uncertainties represented by the meteorological ensemble reach 1.5 m/s in the afternoon, our results show that the information of wind uncertainties is not important for NO_x emission estimation.

Finally we examine the emission estimation performance in ENS.1 at the scale of the model grid (3 km). As shown in Figure 8, the true emission shows high spatial variation from city center to suburban as well as distinct point emission sources. In the example of the emission estimate at 9:00 am, the posterior emission recovers the truth very well with the posterior RMSE of 21.6 mol/(km²·hr). In contrast, the emission estimate at 4:00 pm shows RMSE of 46.5 mol/(km²·hr) due to relatively high wind errors. The posterior underestimates the emissions significantly all over the city except for the regional overestimation in the east. The emission hot spot of ~250 mol/(km²·hr) in the city center is not fully represented in the posterior estimate.

5 Summary and conclusions

In this study, we demonstrate the ability to estimate NO_x emissions by assimilating column NO_2 and meteorological observations into a data assimilation system comprised of regional CTM WRF-Chem and DART-EAKF. Such an ensemble-based data assimilation system is appealing for regional chemical transport studies as it allows the flexibility and efficiency to assimilate observations with various scales and of various chemical species. This is especially true, at present, in light of the availability of future long-term multi-species observations from satellite, such as GEO-CAPE and TROPOMI.

Previous work has shown that NO_x concentrations and columns vary at fine scales necessitating high spatiotemporal resolution to make use of them in the assimilation. In the coupled chemical and meteorological data assimilation system, we apply an OSSE framework to estimate NO_x emissions in Denver by jointly assimilating MADIS observation of meteorological variables as well as future TEMPO NO_2 columns. In meteorological assimilation we successfully reduced the posterior wind RMSE below 0.5 m/s in Denver to better represent the NO_2 transport. The prior wind RMSE and spread show peaks in the afternoon thus increasing the errors in NO_2 transport. We find that the meteorological uncertainties contribute 35% to the total NO_2 forecast uncertainties considering the emission uncertainties of 60%. Assimilation of TEMPO NO_2 columns reduces errors in the predicted NO_2 concentration by more than 50%, which demonstrates the potential of future geostationary observations to constrain NO_2 chemical weather.

One of the goals of this work is to investigate the optimal strategy to estimate NO_x emissions. We test the upper limit of emission constraint from TEMPO NO_2 observations in an ideal case assuming no errors associated with the modeled meteorology. In the experiment of joint assimilation of meteorology and chemical NO_2 , we find that the emission estimation is successful in the morning but degrades in the afternoon when the prior wind RMSE grows above 1 m/s. Considering the dependence of errors in estimated emissions on the wind forecast errors, we recommend to guarantee the accuracy in modeled wind with RMSE below 1 m/s for the success of chemical assimilation to infer emissions at the scale of our model grid. We show that the simultaneous update of NO_x emissions and concentrations outperforms the approach of updating emissions only. We recommend carrying out meteorology assimilations to stabilize from the initial transport errors before starting the emission inversion. Last but not least our results show that the information of wind uncertainties is not important for NO_x emission estimation. Assimilations of NO_2 only with the meteorology constrained from an hourly weather assimilation product would perform as well as the joint assimilation of meteorology and chemistry.

The emissions optimization scheme used in this study assumes no errors in the diurnal variation of emissions, which enables us to optimize the emission scaling factors with hourly frequency. For future work, if there is a large uncertainty in the temporal variability, the emissions at different time steps can be treated as independent state variables to allow the possibility of updating the prior temporal pattern of emissions. In our OSSE, we start with an ideal case of a priori emissions, which is spatially-uniform 70% of true emissions. Success in recovering emissions with a uniform scaling can't guarantee success in recovering emissions with spatially-varying perturbations. However, an approach that didn't work in the ideal case will fail in the realistic scenario with spatially-inhomogeneous emission perturbations. The degradation of emission estimation in the

ENS.2 and ENS.3 will still hold for correcting emission errors under more complex circumstances, such as spatially- or temporally- varying emission errors. For future work, further analysis is required to examine whether spatially-varying emission errors can be reduced under the condition of low wind RMSE.

Finally, we note that the methodology developed in this study can also be applied to emissions estimates of other chemical species or time-evolving parameters. For chemical species with longer chemical lifetimes, care should be taken in constraining the errors in transport and boundary conditions and preserving chemical covariance, as emission-induced forcing will be less dominant.

Acknowledgements. The authors gratefully acknowledge support from the NASA Grant NNX14AH046 and NNX15AE376. We thank N. Collins (NCAR/IMAGe) and T. Hoar (NCAR/IMAGe) for the assistance with DART. We would like to acknowledge high-performance computing support from Yellowstone (ark:/85065/d7wd3xhc) provided by NCAR's Computational and Information Systems Laboratory, sponsored by the National Science Foundation. We also thank the reviewers of this manuscript for their constructive suggestions.

References

- Anderson, J., Hoar, T., Raeder, K., Liu, H., Collins, N., Torn, R. and Avellano, A.: The Data Assimilation Research Testbed: A Community Facility, *Bull. Am. Meteorol. Soc.*, 90(9), 1283–1296, doi:10.1175/2009BAMS2618.1, 2009.
- Anderson, J. L.: An Ensemble Adjustment Kalman Filter for Data Assimilation., *Mon. Weather Rev.* Dec2001, 129(12), 2001.
- Anderson, J. L. and Collins, N.: Scalable Implementations of Ensemble Filter Algorithms for Data Assimilation, *J. Atmos. Ocean. Technol.*, 24(8), 1452–1463, doi:10.1175/JTECH2049.1, 2007.
- Arellano Jr., A. F., Raeder, K., Anderson, J. L., Hess, P. G., Emmons, L. K., Edwards, D. P., Pfister, G. G., Campos, T. L. and Sachse, G. W.: Evaluating model performance of an ensemble-based chemical data assimilation system during INTEX-B field mission, *Atmos. Chem. Phys.*, 7(21), 5695–5710, doi:10.5194/acp-7-5695-2007, 2007.
- Barbu, A. L., Segers, A. J., Schaap, M., Heemink, A. W. and Builtjes, P. J. H.: A multi-component data assimilation experiment directed to sulphur dioxide and sulphate over Europe, *Atmos. Environ.*, 43(9), 1622–1631, doi:10.1016/j.atmosenv.2008.12.005, 2009.
- Barker, D., Huang, X.-Y., Liu, Z., Auligné, T., Zhang, X., Rugg, S., Ajjaji, R., Bourgeois, A., Bray, J., Chen, Y., Demirtas, M., Guo, Y.-R., Henderson, T., Huang, W., Lin, H.-C., Michalakes, J., Rizvi, S. and Zhang, X.: The Weather Research and Forecasting Model's Community Variational/Ensemble Data Assimilation System: WRFDA, *Bull. Am. Meteorol. Soc.*, 93(6), 831–843, doi:10.1175/BAMS-D-11-00167.1, 2012.
- Barré, J., Gaubert, B., Arellano, A. F. J., Worden, H. M., Edwards, D. P., Deeter, M. N., Anderson, J. L., Raeder, K., Collins, N., Tilmes, S., Francis, G., Clerbaux, C., Emmons, L. K., Pfister, G. G., Coheur, P.-F. and Hurtmans, D.: Assessing the

- impacts of assimilating IASI and MOPITT CO retrievals using CESM-CAM-chem and DART, *J. Geophys. Res. Atmos.*, 120(19), 10,501–10,529, doi:10.1002/2015JD023467, 2015.
- Beirle, S., Boersma, K. F., Platt, U., Lawrence, M. G. and Wagner, T.: Megacity emissions and lifetimes of nitrogen oxides probed from space., *Science*, 333(6050), 1737–9, doi:10.1126/science.1207824, 2011.
- 5 Boersma, K. F., Eskes, H. J. and Brinksma, E. J.: Error analysis for tropospheric NO₂ retrieval from space, *J. Geophys. Res.*, 109(D4), D04311, doi:10.1029/2003JD003962, 2004.
- Bousserez, N., Henze, D. K., Rooney, B., Perkins, A., Wecht, K. J., Turner, A. J., Natraj, V. and Worden, J. R.: Constraints on methane emissions in North America from future geostationary remote sensing measurements, *Atmos. Chem. Phys.*, 16, 6175–6190, doi:10.5194/acp-16-6175-2016, 2016.
- 10 Bowman, K. W.: Toward the next generation of air quality monitoring: Ozone, *Atmos. Environ.*, 80, 571–583, doi:10.1016/j.atmosenv.2013.07.007, 2013.
- Bucsela, E. J., Krotkov, N. A., Celarier, E. A., Lamsal, L. N., Swartz, W. H., Bhartia, P. K., Boersma, K. F., Veefkind, J. P., Gleason, J. F. and Pickering, K. E.: A new stratospheric and tropospheric NO₂ retrieval algorithm for nadir-viewing satellite instruments: applications to OMI, *Atmos. Meas. Tech.*, 6(10), 2607–2626, doi:10.5194/amt-6-2607-2013, 2013.
- 15 Chatterjee, A., Michalak, A. M., Anderson, J. L., Mueller, K. L. and Yadav, V.: Toward reliable ensemble Kalman filter estimates of CO₂ fluxes, *J. Geophys. Res. Atmos.*, 117(22), 1–17, doi:10.1029/2012JD018176, 2012.
- Claeyman, M., Attié, J.-L., Peuch, V.-H., El Amraoui, L., Lahoz, W. A., Josse, B., Ricaud, P., von Clarmann, T., Höpfner, M., Orphal, J., Flaud, J.-M., Edwards, D. P., Chance, K., Liu, X., Pasternak, F. and Cantié, R.: A geostationary thermal infrared sensor to monitor the lowermost troposphere: O₃ and CO retrieval studies, *Atmos. Meas. Tech.*, 4(2), 297–317, doi:10.5194/amt-4-297-2011, 2011.
- 20 Cui, Y. Y., Brioude, J., McKeen, S. a., Angevine, W. M., Kim, S.-W., Frost, G. J., Ahmadov, R., Peischl, J., Bousserez, N., Liu, Z. and Ryerson, T. B.: Top-down estimate of methane emissions in California using a mesoscale inverse modeling technique: The South Coast Air Basin, *J. Geophys. Res. Atmos.*, 120, 6698–6711, doi:10.1002/2014JD023002, 2015.
- Edwards, D. P., Arellano, A. F. and Deeter, M. N.: A satellite observation system simulation experiment for carbon monoxide in the lowermost troposphere, *J. Geophys. Res.*, 114(D14), D14304, doi:10.1029/2008JD011375, 2009.
- 25 Elbern, H., Strunk, A., Schmidt, H. and Talagrand, O.: Emission rate and chemical state estimation by 4-dimensional variational inversion, *Atmos. Chem. Phys.*, 7(14), 3749–3769, doi:10.5194/acp-7-3749-2007, 2007.
- Fishman, J., Iraci, L. T., Al-Saadi, J., Chance, K., Chavez, F., Chin, M., Coble, P., Davis, C., DiGiacomo, P. M., Edwards, D., Eldering, A., Goes, J., Herman, J., Hu, C., Jacob, D. J., Jordan, C., Kawa, S. R., Key, R., Liu, X., Lohrenz, S., Mannino, A., Natraj, V., Neil, D., Neu, J., Newchurch, M., Pickering, K., Salisbury, J., Sosik, H., Subramaniam, A., Tzortziou, M., Wang, J. and Wang, M.: The United States' Next Generation of Atmospheric Composition and Coastal Ecosystem Measurements: NASA's Geostationary Coastal and Air Pollution Events (GEO-CAPE) Mission, *Bull. Am. Meteorol. Soc.*, 93(10), 1547–1566, doi:10.1175/BAMS-D-11-00201.1, 2012.
- 30 Gaspari, G., Gaspari, G., Cohn, S. E. and Cohn, S. E.: Construction of correlation functions in two and three dimensions, *Q. J. R. Meteorol. Soc.*, 125(April 1998), 723–757, doi:10.1002/qj.49712555417, 1999.

- Gaubert, B., Coman, A., Foret, G., Meleux, F., Ung, A., Rouil, L., Ionescu, A., Candau, Y. and Beekmann, M.: Regional scale ozone data assimilation using an ensemble Kalman filter and the CHIMERE chemical transport model, *Geosci. Model Dev.*, 7(1), 283–302, doi:10.5194/gmd-7-283-2014, 2014.
- Guerrette, J. J. and Henze, D. K.: Development and application of the WRFPLUS-Chem online chemistry adjoint and WRFDA-Chem assimilation system, *Geosci. Model Dev.*, 8(6), 1857–1876, doi:10.5194/gmd-8-1857-2015, 2015.
- Hache, E., Attié, J.-L., Tourneur, C., Ricaud, P., Coret, L., Lahoz, W. A., El Amraoui, L., Josse, B., Hamer, P., Warner, J., Liu, X., Chance, K., Höpfner, M., Spurr, R., Natraj, V., Kulawik, S., Eldering, A. and Orphal, J.: The added value of a visible channel to a geostationary thermal infrared instrument to monitor ozone for air quality, *Atmos. Meas. Tech.*, 7(7), 2185–2201, doi:10.5194/amt-7-2185-2014, 2014.
- 10 Hu, X.-M., Nielsen-Gammon, J. W. and Zhang, F.: Evaluation of Three Planetary Boundary Layer Schemes in the WRF Model, *J. Appl. Meteorol. Climatol.*, 49(9), 1831–1844, doi:10.1175/2010JAMC2432.1, 2010.
- Kang, J.-S., Kalnay, E., Liu, J., Fung, I., Miyoshi, T. and Ide, K.: “Variable localization” in an ensemble Kalman filter: Application to the carbon cycle data assimilation, *J. Geophys. Res.*, 116(D9), D09110, doi:10.1029/2010JD014673, 2011.
- 15 Kang, J.-S., Kalnay, E., Miyoshi, T., Liu, J. and Fung, I.: Estimation of surface carbon fluxes with an advanced data assimilation methodology, *J. Geophys. Res. Atmos.*, 117(D24), n/a–n/a, doi:10.1029/2012JD018259, 2012.
- Koohkan, M. R., Bocquet, M., Roustan, Y., Kim, Y. and Seigneur, C.: Estimation of volatile organic compound emissions for Europe using data assimilation, *Atmos. Chem. Phys.*, 13(12), 5887–5905, doi:10.5194/acp-13-5887-2013, 2013.
- Kretschmer, R., Gerbig, C., Karstens, U. and Koch, F.-T.: Error characterization of CO₂ vertical mixing in the atmospheric transport model WRF-VPRM, *Atmos. Chem. Phys.*, 12(5), 2441–2458, doi:10.5194/acp-12-2441-2012, 2012.
- 20 Kretschmer, R., Gerbig, C., Karstens, U., Biavati, G., Vermeulen, a., Vogel, F., Hammer, S. and Totsche, K. U.: Impact of optimized mixing heights on simulated regional atmospheric transport of CO₂, *Atmos. Chem. Phys.*, 14(14), 7149–7172, doi:10.5194/acp-14-7149-2014, 2014.
- 25 Krotkov, N. a., McLinden, C. a., Li, C., Lamsal, L. N., Celarier, E. a., Marchenko, S. V., Swartz, W. H., Bucsela, E. J., Joiner, J., Duncan, B. N., Boersma, K. F., Veefkind, J. P., Levelt, P. F., Fioletov, V. E., Dickerson, R. R., He, H., Lu, Z. and Streets, D. G.: Aura OMI observations of regional SO₂ and NO₂ pollution changes from 2005 to 2015, *Atmos. Chem. Phys.*, 16, 4605–4629, doi:10.5194/acp-16-4605-2016, 2016.
- Lahoz, W. A., Peuch, V.-H., Orphal, J., Attié, J.-L., Chance, K., Liu, X., Edwards, D., Elbern, H., Flaud, J.-M., Claeysman, M. and Amraoui, L. El: Monitoring Air Quality from Space: The Case for the Geostationary Platform, *Bull. Am. Meteorol. Soc.*, 93(2), 221–233, doi:10.1175/BAMS-D-11-00045.1, 2012.
- 30 Laughner, J. L., Zare, A. and Cohen, R. C.: Effects of daily meteorology on the interpretation of space-based remote sensing of NO₂, *Atmos. Chem. Phys.*, 16, 15247–15264, doi:10.5194/acp-16-15247-2016, 2016.
- Lauvaux, T. and Davis, K. J.: Planetary boundary layer errors in mesoscale inversions of column-integrated CO₂ measurements, *J. Geophys. Res. Atmos.*, 119(2), 490–508, doi:10.1002/2013JD020175, 2014.
- 35 Liu, J., Fung, I., Kalnay, E. and Kang, J.-S.: CO₂ transport uncertainties from the uncertainties in meteorological fields, *J. Geophys. Res. Lett.*, 38(12), n/a–n/a, doi:10.1029/2011GL047213, 2011.

- Miyazaki, K., Eskes, H. J., Sudo, K., Takigawa, M., van Weele, M. and Boersma, K. F.: Simultaneous assimilation of satellite NO₂, O₃, CO, and HNO₃ data for the analysis of tropospheric chemical composition and emissions, *Atmos. Chem. Phys.*, 12, 9545–9579, doi:10.5194/acp-12-9545-2012, 2012.
- Miyazaki, K., Eskes, H., Sudo, K., Boersma, K. F., Bowman, K. and Kanaya, Y.: Decadal changes in global surface NO_x emissions from multi-constituent satellite data assimilation, *Atmos. Chem. Phys.*, 17, 807–837, doi:10.5194/acp-17-807-2017, 2017.
- Mizzi, A. P., Arellano Jr., A. F., Edwards, D. P., Anderson, J. L. and Pfister, G. G.: Assimilating compact phase space retrievals of atmospheric composition with WRF-Chem/DART: a regional chemical transport/ensemble Kalman filter data assimilation system, *Geosci. Model Dev.*, 9(3), 965–978, doi:10.5194/gmd-9-965-2016, 2016.
- 10 Pagowski, M. and Grell, G. A.: Experiments with the assimilation of fine aerosols using an ensemble Kalman filter, *J. Geophys. Res. Atmos.*, 117(D21), n/a–n/a, doi:10.1029/2012JD018333, 2012.
- Romine, G. S., Schwartz, C. S., Snyder, C., Anderson, J. L. and Weisman, M. L.: Model Bias in a Continuously Cycled Assimilation System and Its Influence on Convection-Permitting Forecasts, *Mon. Weather Rev.*, 141(4), 1263–1284, doi:10.1175/MWR-D-12-00112.1, 2013.
- 15 Russell, A. R., Perring, A. E., Valin, L. C., Bucsela, E. J., Browne, E. C., Wooldridge, P. J. and Cohen, R. C.: A high spatial resolution retrieval of NO₂ column densities from OMI: method and evaluation, *Atmos. Chem. Phys.*, 11(16), 8543–8554, doi:10.5194/acp-11-8543-2011, 2011.
- Russell, A. R., Valin, L. C. and Cohen, R. C.: Trends in OMI NO₂ observations over the United States: effects of emission control technology and the economic recession, *Atmos. Chem. Phys.*, 12(24), 12197–12209, doi:10.5194/acp-12-12197-2012, 2012.
- 20 Saide, P. E., Kim, J., Song, C. H., Choi, M., Cheng, Y. and Carmichael, G. R.: Assimilation of next generation geostationary aerosol optical depth retrievals to improve air quality simulations, *Geophys. Res. Lett.*, 9188–9196, doi:10.1002/2014GL062089, Received, 2014.
- Stockwell, W. R., Middleton, P., Chang, J. S. and Tang, X.: The second generation regional acid deposition model chemical mechanism for regional air quality modeling, *J. Geophys. Res.*, 95(D10), 16343, doi:10.1029/JD095iD10p16343, 1990.
- Travis, K. R., Jacob, D. J., Fisher, J. a., Kim, P. S., Marais, E. a., Zhu, L., Yu, K., Miller, C. C., Yantosca, R. M., Sulprizio, M. P., Thompson, A. M., Wennberg, P. O., Crounse, J. D., St. Clair, J. M., Cohen, R. C., Laughner, J. L., Dibb, J. E., Hall, S. R., Ullmann, K., Wolfe, G. M., Pollack, I. B., Peischl, J., Neuman, J. a. and Zhou, X.: NO_x emissions, isoprene oxidation pathways, vertical mixing, and implications for surface ozone in the Southeast United States, *Atmos. Chem. Phys.*, 16(21), 13561–13577, doi:10.5194/acp-16-13561-2016, 2016..
- 30 Turner, A. J., Jacob, D. J., Wecht, K. J., Maasakkers, J. D., Lundgren, E., Andrews, A. E., Biraud, S. C., Boesch, H., Bowman, K. W., Deutscher, N. M., Dubey, M. K., Griffith, D. W. T., Hase, F., Kuze, A., Notholt, J., Ohyama, H., Parker, R., Payne, V. H., Sussmann, R., Sweeney, C., Velasco, V. A., Warneke, T., Wennberg, P. O. and Wunch, D.: Estimating global and North American methane emissions with high spatial resolution using GOSAT satellite data, *Atmos. Chem. Phys.*, 15(12), 7049–7069, doi:10.5194/acp-15-7049-2015, 2015.
- Valin, L. C., Russell, A. R., Hudman, R. C. and Cohen, R. C.: Effects of model resolution on the interpretation of satellite NO₂ observations, *Atmos. Chem. Phys.*, 11(22), 11647–11655, doi:10.5194/acp-11-11647-2011, 2011.

Valin, L. C., Russell, A. R. and Cohen, R. C.: Variations of OH radical in an urban plume inferred from NO₂ column measurements, *Geophys. Res. Lett.*, 40(9), 1856–1860, doi:10.1002/grl.50267, 2013.

Yumimoto, K.: Impacts of geostationary satellite measurements on CO forecasting: An observing system simulation experiment with GEOS-Chem/LETKF data assimilation system, *Atmos. Environ.*, 74, 123–133, doi:10.1016/j.atmosenv.2013.03.032, 2013.

Zoogman, P., Jacob, D. J., Chance, K., Worden, H. M., Edwards, D. P. and Zhang, L.: Improved monitoring of surface ozone by joint assimilation of geostationary satellite observations of ozone and CO, *Atmos. Environ.*, 84, 254–261, doi:10.1016/j.atmosenv.2013.11.048, 2014.

Zoogman, P., Liu, X., Suleiman, R. M., Pennington, W. F., Flittner, D. E., Al-Saadi, J. a., Hilton, B. B., Nicks, D. K., Newchurch, M. J., Carr, J. L., Janz, S. J., Andraschko, M. R., Arola, a., Baker, B. D., Canova, B. P., Chan Miller, C., Cohen, R. C., Davis, J. E., Dussault, M. E., Edwards, D. P., Fishman, J., Ghulam, a., González Abad, G., Grutter, M., Herman, J. R., Houck, J., Jacob, D. J., Joiner, J., Kerridge, B. J., Kim, J., Krotkov, N. a., Lamsal, L., Li, C., Lindfors, a., Martin, R. V., McElroy, C. T., McLinden, C., Natraj, V., Neil, D. O., Nowlan, C. R., O'Sullivan, E. J., Palmer, P. I., Pierce, R. B., Pippin, M. R., Saiz-Lopez, a., Spurr, R. J. D., Szykman, J. J., Torres, O., Veeffkind, J. P., Veihelmann, B., Wang, H., Wang, J. and Chance, K.: Tropospheric emissions: Monitoring of pollution (TEMPO), *J. Quant. Spectrosc. Radiat. Transf.*, 186(2017), 17–39, doi:10.1016/j.jqsrt.2016.05.008, 2017.

Table 1. The experimental set up of each assimilation run. The three ensemble runs assimilate NO₂ observations every hour, and differ in treatment of meteorology forecast.

Experiment	Met Assim	Chem Assim	note
REF	No	Yes	Truth meteorology
ENS.1	Yes	Yes	Ensemble of meteorology and chemistry
ENS.2	Yes	Yes	Only update emissions
ENS.3	Yes	Yes	Initial meteorology ensemble is from the next day
REA	No	Yes	Using ensemble mean from ENS.1

Table 2. DART configurations.

Parameter	Value
Filter type	EAKF
Adaptive inflation	1.0, 0.6 (initial mean, spread)
Inflation damping	0.9
Adaptive localization threshold	2000
Localization type	Gaspari-Cohn
Horizontal localization half-width for meteorology (chemical) observation	50 km (10 km)
Outlier threshold	3.0
Ensemble members	30

5

Table 3. WRFDA configurations.

Parameter	Value
cv_options	3 (NCEP background error model)
je_factor (ensemble covariance weighting factor)	1.0

10

Table 4. Relative observation uncertainty in synthetic TEMPO NO₂ column for each scenario.

Type	NO ₂ column	Gaussian distribution
Clean	<0.3×10 ¹⁵ molec. com ⁻²	N (200%, 100%)
Polluted	>=0.3×10 ¹⁵ molec. com ⁻²	N (7.5%, 2.5%)

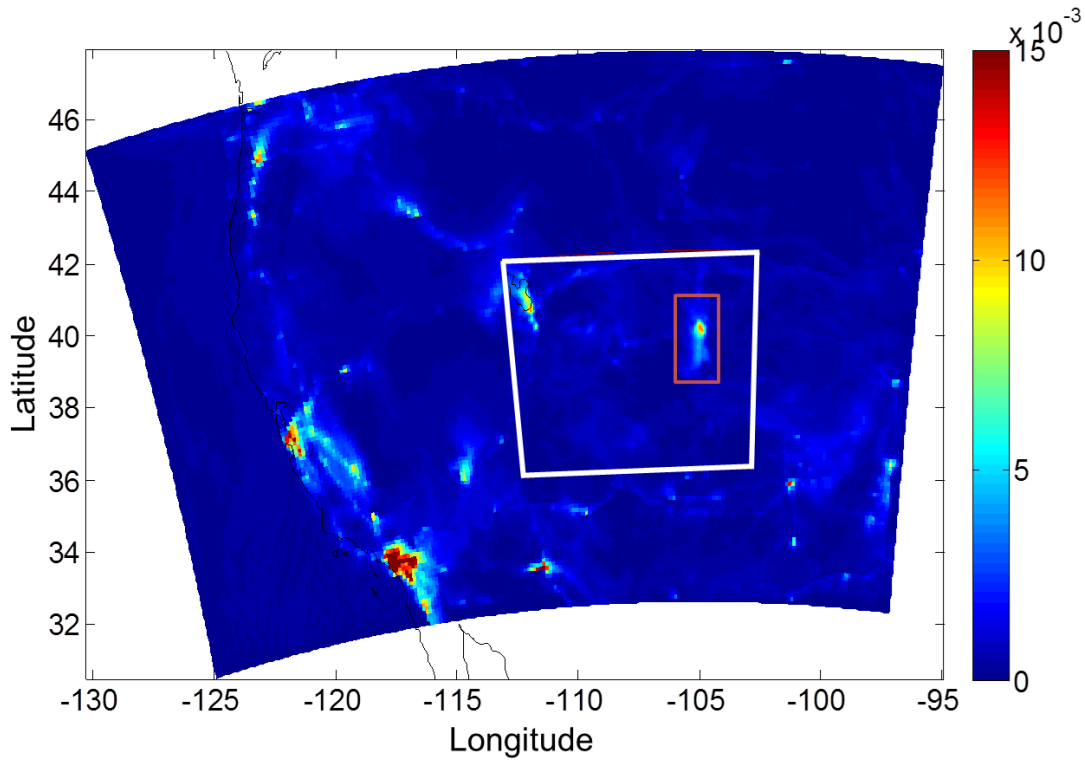


Figure 1. Model domain setup with 12 km outer domain and 3 km inner domain (white square). Data assimilation is performed on the inner domain. Meteorological observations on the inner domain are assimilated. TEMPO NO₂ observations inside the red rectangle are assimilated.

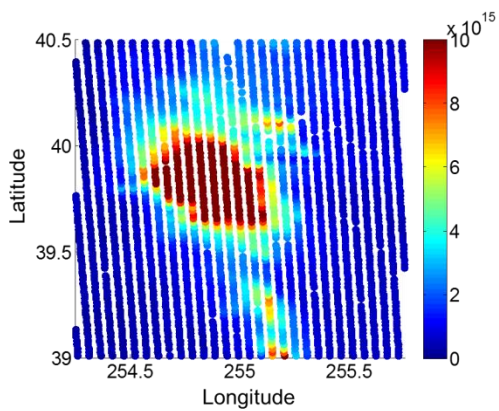
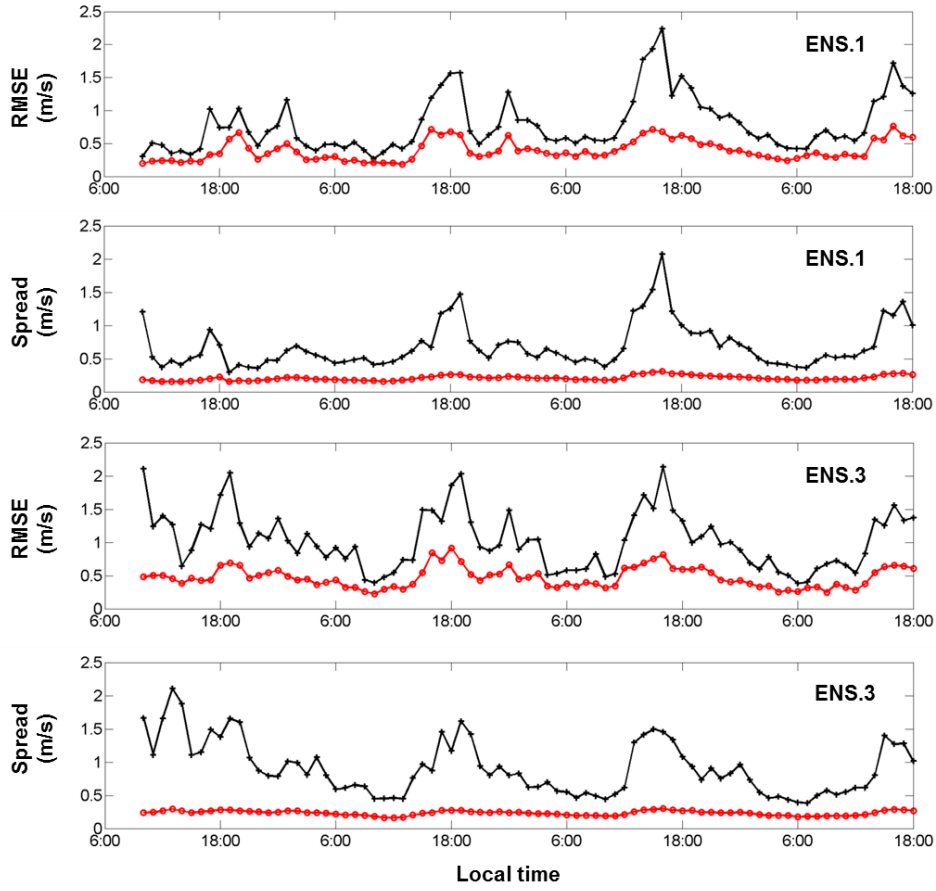


Figure 2. Example of synthetic TEMPO NO₂ column observations over Denver, CO at 17:00 LT July 2nd in 2014.



5 **Figure 3.** Time evolution of prior (black) and posterior (red) RMSE and spread of surface mesonet zonal wind observation in Denver from July 2nd 10:00 to 5th 18:00 for ENS.1 (top) and ENS.3 (bottom).

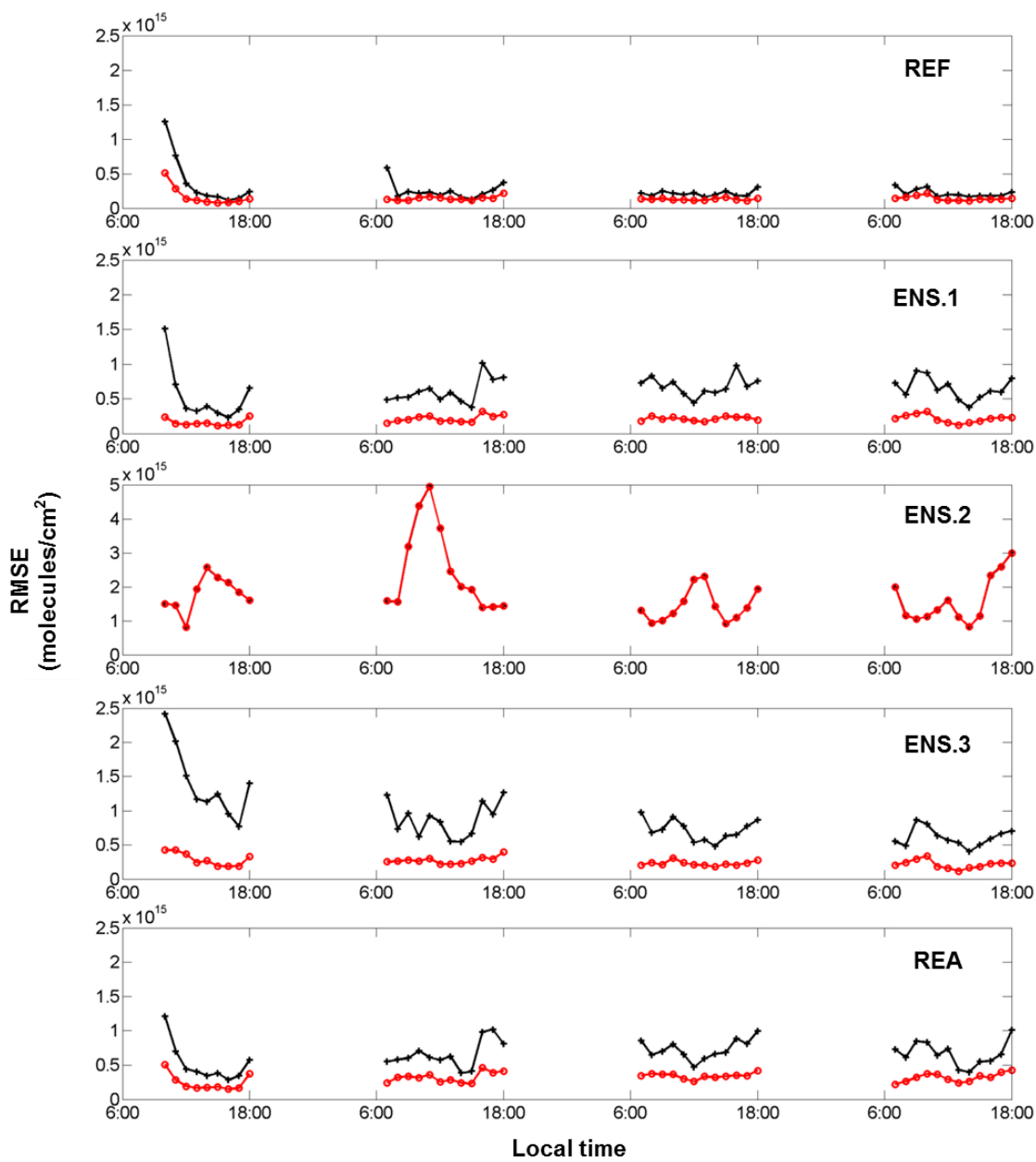


Figure 4. Time evolution of prior (black) and posterior (red) RMSE of Denver TEMPO NO₂ column observation from July 2nd 10:00 to 5th 18:00 for REF, ENS.1, ENS.2, ENS.3 and REA (from top to bottom).

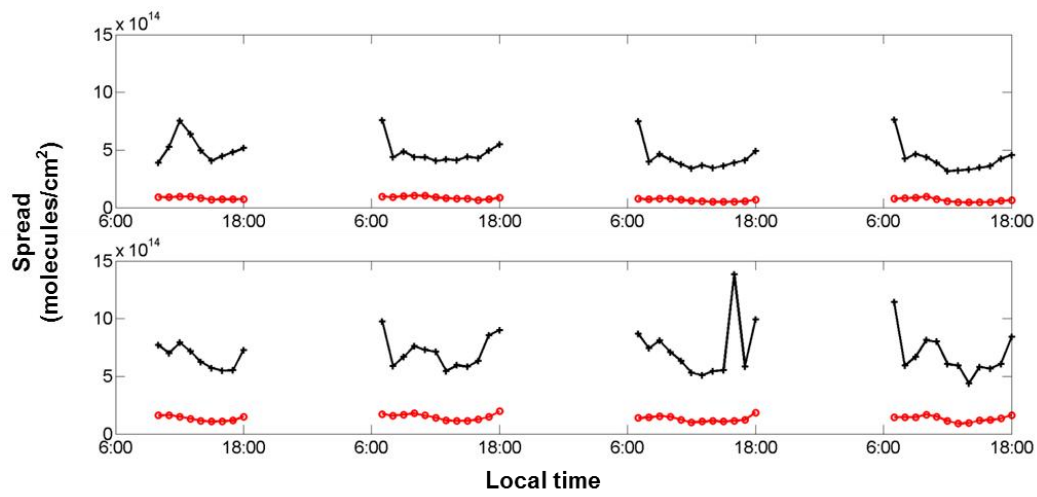


Figure 5. Time evolution of prior (black) and posterior (red) spread of Denver TEMPO NO₂ column observation from July 2nd 10:00 to 5th 18:00 for REF (top) and ENS.1 (bottom).

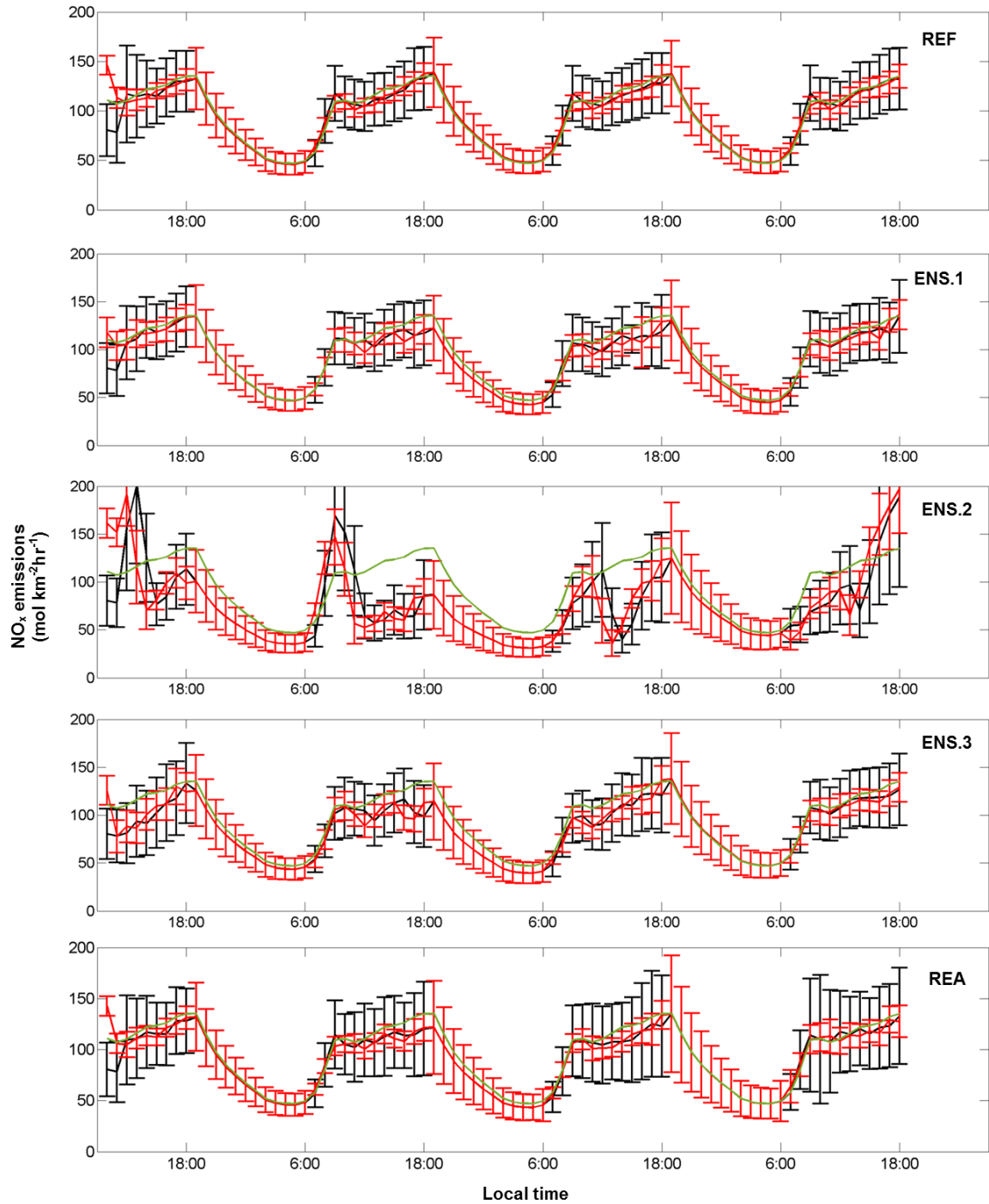


Figure 6. Time evolution of averaged Denver city emission of prior (black), posterior (red) and truth (green) for REF, ENS.1, ENS.2, ENS.3 and REA (from top to bottom). The error bar is defined by the ensemble spread and represents the uncertainty of the prior and posterior estimates.

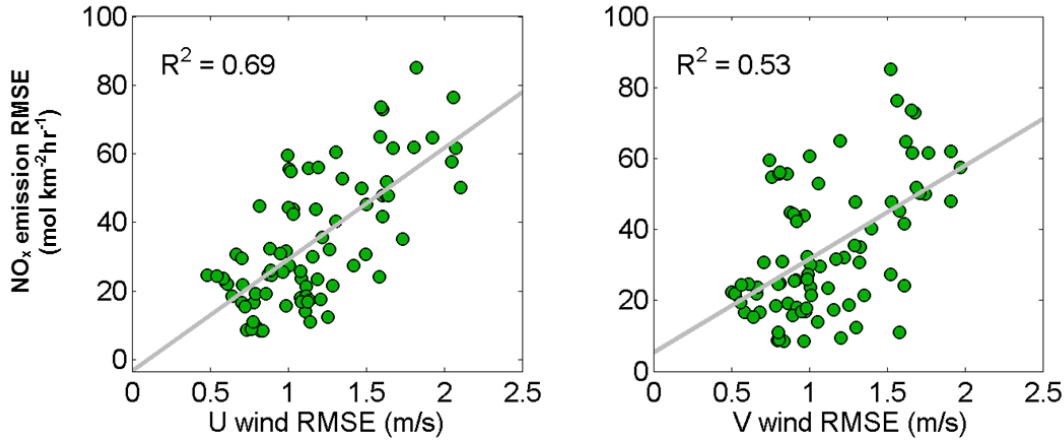


Figure 7. The scatter plot between the prior RMSE of boundary layer wind vectors and urban NO_x emission posterior RMSE over four-day daytime assimilation time period in ENS.1.

5

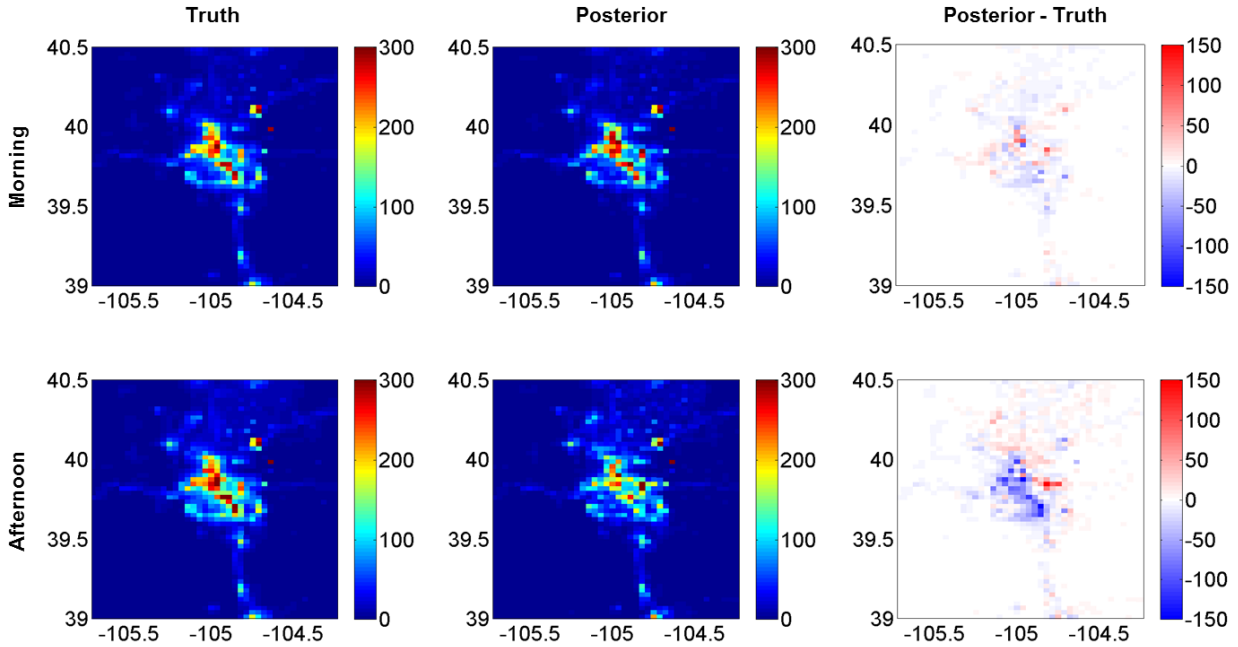


Figure 8. The emission estimation results in ENS.1 at 9:00 am (top) 4:00 pm (bottom) on July 3rd of truth, posterior and the difference between truth and posterior (from left to right). The unit is mol km⁻²hr⁻¹.