

We would like to thank the reviewers for their thoughtful comments and efforts towards improving our manuscript. In the following, we highlight general concerns of reviewers that were common and our effort to address these concerns. We then address comments specific to each reviewer below.

Referees were generally concerned that the conclusions made in our original manuscript were too general for the range of conditions tested. Reviewers suggested that we generalize the study in terms of the following controlling parameters, which include

- VOC reactivity and its effects on the NO₂-OH relationship
- Variation of OH production rates with time of day, season, location, etc.
- Varying source distributions (e.g., randomly spaced sources, homogenous sources)
- Variations in atmospheric mixing (e.g. PBL height, wind speeds)
- The spatial scale over which biases are considered
- The effects of background NO₂ concentrations
- Memory effects of nighttime NO_x chemistry

Indeed, all of these parameters, and several others not mentioned, affect the interplay of NO₂ transport and removal computed on a discretized model grid. Instead of testing each one of these parameters that vary in space, time and representation in models separately, we include several example calculations that highlight the basic ideas at work and then add a general discussion of a plume's spatial characteristics, a term that combines source distribution, chemical and atmospheric mixing effects, in the context of model resolution. To accomplish this task, we have added contextual sentences throughout the manuscript, a full section titled "Implications for interpretation of satellite observations," references to observational datasets that capture the spatial variability of NO₂ in the atmosphere, and 2-D simulations that test for effects of increased VOC and effects of shifting an area source with respect to a coarse grid.

In addition to the suggestions above, the reviewers felt that the simulations presented in the original manuscript were disconnected from one another and were not adequately represented in the context of tropospheric chemistry and satellite-based observation of NO₂ column.

Many of the additions aimed at addressing the first set of general comments apply here as well; we have added a new section, better transition sentences, and a discussion of each model simulation in the context of the other simulations to the revised manuscript.

In the following sections, we respond to the comments that were more specific to each referee.

Responses to Referee #1

We thank the referee for the careful and insightful review of our manuscript. We address all of the concerns of the referee here. We note that we responded to the referee's general comment #1 and #4 under the set of responses common to all three referees.

The other comments include

Comment 1b) For the appendix, there is no description of the results at all, only the settings for the simulations are described.

Response 1b) The selected chemical parameters are standard for WRF-CHEM. Simulations with the selected parameters have been described in extensive detail elsewhere (e.g., Grell et al., 2005). In addition, we feel that Figures 6-9 show that the simulated chemistry is reasonable. If the reviewer believes there is some specific prediction that will enhance the discussion, we will gladly include it, but do not want to add extraneous information.

Comment 2) For the 2D studies, area sources should be used (instead of point sources), since this will provide the realistic results important for the comparisons to measurements. The results for the point sources will be misinterpreted, since the authors write that for area sources the biases are significantly smaller.

Response 2) We feel that biases over point sources and area sources are both relevant, and thus mention both. There are many examples of model satellite comparisons in the literature over point sources as well as area sources (e.g., Kim et al 2009, Martin et al., 2003). Furthermore, large coal-fired power plants, account for a large portion of anthropogenic NO_x emissions around the world.

Comment 3) Still there is also the effect of the averaging of the NO2 for the satellite pixels (with different cloud coverage, albedo, etc). How do these problems for comparisons go together?

Response 3) These are good questions, but we feel that this comparison is beyond the scope of our research in assessing model uncertainty due to grid resolution. There are many examples of research assessing satellite uncertainty and model-satellite comparisons (e.g., Russell et al, 2011; Boersma et al, 2008, Bucsela et al., 2008, Kim et al 2009). While this work is ongoing, the consensus is that the retrievals are improving and are reaching the 25% limit, hence we believe the values chosen for our example cases of 10 and 25 % are reasonable choices.

Comment 5) Nothing is said about the resolution in time. I heard that for CTMs the relation of the time resolution to the spatial resolution is important for a correct simulation. If this should not be the case for this study then explain why, otherwise a study on the effect of the time resolution on the results would be important.

Response 5) It is true the time step selected for the dynamical portion of a CTM must consider spatial resolution. However, we have selected a dynamical and chemical time step of 15 seconds for all model resolutions, adequate for accurately solving both chemistry and meteorology.

Referee #1 minor comments:

Referee #1 had several minor comments. We respond to those below.

The title should also make clear that the study is on “atmospheric chemistry” models

We prefer the shorter title and feel that it is clear from the title and the first sentence of abstract that this is a study of tropospheric NO₂.

One or two introducing sentences on a general level would be good. Especially it should be said that the study is on tropospheric chemistry and in particular NO₂.

Added one sentence to the abstract.

“Inference of NO_x emissions (NO+NO₂) from satellite observations of tropospheric NO₂ column requires knowledge of NO_x lifetime, usually provided by chemical transport models (CTMs).”

“(100s ppb)” and (100s ppt)”: please check if this is the conventional scientific writing.

Changed to (10-1000 ppb) and (10-100 ppt)

“High-quality satellite based observations...” The term high-quality is not defined and represents an impression or opinion of the authors, which should be skipped. Also it gives the bad impression of some kind of “commercial” since 2 (or 3) of 4 articles cited here are from Co-authors of this article!

We remove reference to the observations being “High-quality” and add more references to support the usefulness of the observations in answering science questions.

Page 4, Line 2: “... may have resolution-dependent biases” This is a very vague formulation considering the obvious effect and also the result of this study. I suggest rephrasing this sentence.

We change to “As a result, NO_x emissions derived from an inversion of satellite NO₂ observations will have resolution-dependent biases, the magnitude of which is not well known.”

Page 4, line 15 and 18: check if “1s-1” and “0 molec cm-3 s-1” are the correct values

We added a table that lists the parameters in the analytical NO_x-HO_x steady state relationship out of the text. $k_{\text{OH}+\text{VOC}}[\text{OH}]$ or VOC reactivity of 1 s^{-1} is typical for background CO and CH₄. $0 \text{ molec cm}^{-3} \text{ s}^{-1}$ is a correct value. We include the following text as a footnote.

“^aAlkyl-nitrate formation rate has been set to 0 in these simulations. Increases in this rate will slightly affect NO_x lifetime (Farmer et. al., 2011).”

Page 4, line19: For what resolution and what distance of the source is the figure 1? Or what data is averaged here ? And would results from 2d and 3d be similar?

There is a misunderstanding here. Figure 1 represents the relationship of OH to NO_x under different chemical conditions and has connection to transport or emissions. To better clarify, we add a table that lists the controlling parameters that had previously been listed in text (P 20248 L1-6), and add separate discussion of the steady-state NO_x-HO_x chemistry from discussion of plume models.

Page 20248 line 18: “NO₂ is numerically diluted ...” : Is this correct writing ? And if yes, I think it should be rephrased to avoid confusion. This is only distributing the large NO₂ over a wider area.

We feel that numerically diluted a more succinct way to describe that the NO₂ as being distributed over a wider area.

In multiple comments, the reviewer feels that it is better to discuss biases at any given location instead of comparing domain-average values. (e.g., Page 20248 line 20: Depending on the resolution, the bias averaged over this large domain seems for me not to be the important value, considering the motivation of the study. What is the maximum difference that can occur ?)

All known satellite-model comparisons make comparisons at a spatial scale of at least 50 x 50 km² even if the simulations are run at much finer scale (e.g., Kim et al 2009). Comparison at finer spatial scales would suffer from errors in model transport, which we show in Figure 5 (5e, black line) and are not the focus of this paper.

Page 20248 L29: “ ... corresponding to a chemical lifetime of about three hours and maximum OH.” Do you mean that of the three investigated source strengths, OH number density is largest for the intermediate ? Otherwise, for the maximum in OH, the lifetime of NO₂ is less than 2 h, see Fig. 1

We clarify by changing the sentence to

“...a gradient corresponding to a chemical lifetime of about three hours and near-maximum OH (Fig. 2b,e – solid line).”

P20249 L1 : Obviously an e-fold decay within 60 km can not be resolved by a resolution of 128 km.

While this is obvious, we make mention to provide continuity with discussion of other source strengths which mention both resolutions.

Page 20249 line 2: “... near the NO₂-OH crossover regime,...”: Is this term correct and the best choice ?

The term is often applied to the maximum in the relationship of NO₂ to ozone production, an analogous relationship to the NO₂-OH relationship. However, we change the sentence to read

“Because NO₂ concentrations predicted for an intermediate source are near the maximum in OH ...”

P20249 L16: “... and diffuses at 10 m² s⁻¹.” : what is meant here ?

This diffusion rate represents a small rate of turbulent mixing. We change the sentence to the following.

“Diffusion rates are set to 10 m² s⁻¹.”

Page 7 line 21: “... as would be observed by a satellite based instrument ...”

Mention that the observations of tropospheric NO₂ have their own uncertainties (albedo, cloud coverage, aerosols, averaging over the pixel size, ...), also causing problems for such a comparison. The pixel sizes for the three instruments could be given here, not every reader will know this.

These are excellent questions. We feel that this comparison is beyond the scope of this study, which aims to assess model uncertainty due to model resolution. There are many examples of research assessing satellite uncertainty (Russell et al, 2011, Hains et al., 2010, Boersma et al., 2008, Bucsela et al., 2008 , Kim et al., 2009). As for the spatial scale of satellite observations, we refrain from listing the footprints of the current instruments to avoid confusion. We want to focus on whether a model is accurately capturing the spatial scale at which NO₂ varies in the atmosphere, and not the spatial scale at which it is observed.

Response to Referee #2

We first note that reference to Referee #2 and Referee #3 was switched in the html and pdf formats. We are responding here to Referee #2 as determined in the pdf Interactive Comment. We thank the referee for the careful and insightful review of our manuscript. We address all of the concerns of the referee here. Many of the referee’s comments fell under the general umbrella of comments common to all three referees. Only comments not covered under the general comments section are covered here. Please see our response to the common comments above.

Note that the chemical lifetime of NO₂ defined in the paper is a function of both NO₂ and OH. In this sense, it is difficult to understand the reason why % deviation VCD in Figure 3 is so large for large source with 256 and 512 km resolutions.

There is a misunderstanding. We more clearly define the lifetime of NO₂ in the text and figure captions. For example, we add

“...NO₂ lifetime indicated ($k_{\text{NO}_2+\text{OH}}[\text{OH}]$)⁻¹.”

Why are the responses to various resolutions so different for the 1-D and the 2-D models although the emission rates are similar for the two models?

Dilution in 1d only occurs along the axis of propagation. In 2-D, the plume is allowed to diffuse horizontally, which leads to slightly different behavior for

small sources in 1-D and 2-D. Furthermore, the averaging effects responsible for biases at coarse resolution are second order in 2-D and first order in 1-D making comparison difficult.

We add the following text to the introduction of the 2-D source

“While 1-D models are illustrative, 2-D models are a better approximation of NO₂ column and provide some additional insights. For instance, we can consider the effects of horizontal diffusion as well as different source distributions.”

For WRF-Chem run for Four Corners and San Juan power plants, the finest resolution in the EPA National Emission Inventory is 4 km x 4 km. Please be specific if the dilutions in the emissions are the same for the two simulations with 4 km x 4km and 1 km x 1km resolutions.

Only point sources are included in the Four Corners simulations. WRF emissions processing has infinite resolution for point sources.

Regarding the sampling box [fig 7-9], are the results sensitive to the definition of the box? Looking at the columns in 12 km x 12km resolution run (Figure 7), the plumes extend out of the limits defined. I have the same question for the Los Angeles box.

The WRF-CHEM predicted biases certainly depend on the definition of the box, but not to the extent of the gross biases we are assessing. In choosing the sampling boxes, we balanced the representation of the example air basin (Four Corners - point source, Large urban source - LA, Large area source – SJ valley) with buffering contributions from long-lived, upper-tropospheric NO₂ column ($\sim 5 \times 10^{14}$ molecules cm⁻²). We add the following text to discussion of the WRF-CHEM simulations.

“We find that the exact numbers depend on the choice of the boundaries that surround each plume, but that the conclusions are independent of that choice.”

For San Joaquin Valley, many grids look like being affected by intermediate source of NO₂ (green color), but the simulations with coarse resolutions do not show positive bias in the NO₂ columns. More analyses of connecting the WRF-Chem model results to the 1-D or the 2D model are necessary.

Indeed, the referee is correct. However, as noted previously, direct comparison of WRF, which integrates full chemistry, with the steady-state midday plume model is difficult. We explain by adding the text below.

“And while the San Joaquin Valley appears to be an intermediate source of NO_x (Fig. 9), which according to the 2-D plume model would indicate that coarse resolution prediction of NO₂ should be biased high (Fig. 5h) it is important to consider the differences between the 2-D plume model, which simulates midday summertime chemistry at steady-state, with WRF-CHEM, which integrates the full diurnal cycle. In WRF-CHEM, all of these sources (Fig. 7-9), including the San Joaquin Valley, suppress OH throughout the morning when

NO_x concentrations are higher and HO_x production rate is lower. This leads to biases that start small in the morning hours and that grow with time of day.”

Comparison of the model NO₂ columns with satellite NO₂ columns over Four Corners and San Juan power plants can provide useful insights since the model simulations in this region are based on the emission inventory that included the measured NO_x emissions from these power plants.

We feel that this comparison is beyond the scope of this study, which aims to assess model uncertainty due to model resolution. There are many examples of research assessing satellite uncertainty and model-satellite comparisons (e.g., Russell et al, 2011, Hains et al., 2010, Kim et al., 2009, Boersma et al., 2008, Bucsela et al., 2008) with one specifically targeting the strategy mentioned (Kim et al., 2009).

One missing part in the manuscript is the impact of volatile organic compounds (VOC) including isoprene on the level of OH

We fully agree. VOC certainly affects the relationship of OH to NO₂. We do note that the role of HO_x recycling from oxidation of VOCs is still under development (e.g., Lelieveld et al., 2008) thus is difficult to generalize. We performed a set of 2-D point and area source simulations under 10x VOC concentration and include the following text.

“VOC reactivity affects predicted biases by altering the NO₂ concentration at which maximum OH occurs (Fig. 1, Table 1). For simulations with VOC reactivity increased by a factor of 10 and the same NO₂ emissions, we find that the pattern of predicted biases changes in the direction one expects based on the shifts in the NO₂-OH relationship for an increase in VOC. However, if the NO_x emission rates are increased along with VOC reactivity, the patterns of biases returns to those predicted at the original VOC conditions.”

Response to Referee #3

We first note that reference to Referee #2 and Referee #3 was switched in the html and pdf formats. We are responding here to Referee #3 as determined in the pdf Interactive Comment. We thank the referee for the careful and insightful review of our manuscript. We address all of the concerns of the referee here. Many of the referee’s comments fell under the general umbrella of comments common to all three referees. Only comments not covered under the general comments section are covered here. Please see our response to the common comments above.

Several major issues need to be addressed prior to consideration for publication. The title implies a general statement on the relation of resolution and NO₂. But the studies are rather specific and idealized. The reader is left wondering whether the conclusions are general. Some of this confusion could be avoided by adding to the title phrases such as “in the southwest US” or “in idealized environments”.

We feel that the manuscript addresses the issues posed in the title. We prefer the shorter title, but we modified the abstract and text such that it is obvious that the current study does not fully generalize these effects but is offering guidelines. Please see the common comments above.

How does diurnal variation affect the 1-d and 2-d simulations? For example, suppose the domain concentration was initialized at sunrise after nighttime chemistry and advection. How would the bias be affected? Is the resolution requirement specific to OMI (afternoon), or would it differ for a morning observation by GOME-2?

The referee is correct. Certainly the diurnal variation in photochemistry will affect these biases. Also, the lifetime of NO_x at nighttime will affect the initial conditions. Nonetheless, biases will persist at coarse resolutions. We do feel that the time of day effects are consistent enough and directly applicable to satellite-model comparisons at 10 AM and 1 PM. Thus, we add the following text discussing 10 AM biases predicted in WRF-CHEM.

“Biases in NO_2 column at 10 AM (not shown) are consistently smaller than those predicted at 1 PM because the NO_2 column predicted at 10 AM is exposed to much less OH through the nighttime and early morning hours. As a result, prediction of 10 AM NO_2 column to 25% accuracy as tested here requires slightly coarser model resolution than was necessary at 1 PM: 12 km over Four Corners, 24 km over San Joaquin Valley, and 48 km over Los Angeles. For 10% accuracy, model resolution of 4 km is necessary over Four Corners, 24 km over San Joaquin Valley and 48 km over Los Angeles.”

Why is 12km accuracy sufficient for Los Angeles, but 4km accuracy needed for the San Joaquin Valley for prediction to 10% accuracy? Does this tell something about the spatial extent of the source? Or is it simply a function of the choice of sub-domain?

We believe that differences in the biases predicted using a 2-D plume model (Fig. 5) indicate that source distribution plays a large role on the behavior and magnitude of predicted biases. The referee is correct that the sub-domain over which biases are computed does affect the exact value of the bias, but it does not affect the overall conclusion. We add the following text to discussion of the WRF-CHEM results.

“We find that the exact numbers depend on the choice of the boundaries that surround each plume, but that the conclusions are independent of that choice.”

Abstract. The first sentence about ozone production should be removed or better supported in the manuscript.

We feel that the references that we provide in the introduction (e.g., Sillman et al., 1990; Kumar et al., 1994; Gillani et al., 1996; Cohan et al., 2006; Wild and Prather, 2006) sufficiently support this statement made in the abstract. If the referee knows of studies that dispute that finding, we will definitely reconsider.