Response to reviewers

General response: We have dropped the analysis of surface ozone throughout the paper to accommodate the suggestions of the editor and the second reviewer. We believe that our ozone analysis is important by itself, but we agree that such analysis is not essential to the present study and can be placed in a separate paper. As a result, we have dropped section 7 entirely.
We have identified and fixed small bugs in processing meteorological figures (Figs. S1-8). The fix only has an insignificant impact on our analyses and results.
We would like to thank the editor and reviewers for useful comments and suggestions.

Reviewer 1

General Comments
The authors used the GEOS-Chem model to investigate the uncertainties associated with both meteorology and chemistry in simulating tropospheric NO2 column and surface ozone. Authors identified several important meteorological fields and chemical processes that may affect the simulated NO2 column and conducted and analyzed multiple sensitivity simulations by perturbing the selected parameters within the reasonable ranges to their best knowledge. They quantified errors in simulated NO2 columns associated with the uncertainties of 10 selected parameters. The uncertainty in simulating NO2 column could have implication for uncertainty of top-down NOx emission estimate using inverse modeling methodology based on GEOS-Chem model and satellite retrievals.
I find this topic about uncertainty quantification is interesting and important. The uncertainty quantified could be used to improve the corresponding model processes and also used for top-down emission inversion. It highlights the importance to understand model biases before deriving top-down emissions with traditional inversion methodology. However, this paper is not well presented. I do have some specific comments and questions about the results as listed below. I would only suggest this paper for publication if the authors address these comments.

Response: We have accommodated comments of the two reviewers. Thank you very much for your suggestions.

Specific Comments
1. The title is not appropriate. This study focuses on the uncertainties in simulating NO2 columns instead of inverse modeling although it has implication for inverse modeling studies. The current title is misleading. In addition, the model uncertainty is always related to a specific model, although it may have implication for other models. Therefore, the title should be something like "modeling uncertainties of tropospheric NO2 columns in GEOS-Chem".

Response: The title has been changed to 'Modeling uncertainties for tropospheric nitrogen dioxide columns affecting satellite-based inverse modeling of nitrogen oxides emissions'. As discussed in the paper (i.e., first paragraphs of sections 4 and 5), many errors discussed here for model meteorology/chemistry are not just specific to GEOS-Chem but also relevant to other CTMs. Therefore our analyses are of interest not only for GEOS-Chem users but (to certain extent)

for users of other models. Therefore we feel it is not necessary to specify 'GEOS-Chem' in the title.

2. In the abstract, authors suggest "a possible systematic model bias such that the top-down emissions will be overestimated by the same magnitudes if the model is used for emission inversion without corrections". I don't agree with this. In the inverse modeling methodology (e.g., Martin et al., 2003; Zhao and Wang, 2009), the model error is accounted for, therefore, as long as the model error is estimated correctly, the inversed emission is still valid.

Response: Martin et al., 2003, etc. assume that model errors are random rather than systematic, therefore they will not affect the mean value of top-down emissions (i.e., the accuracy is not affected while the precision is affected). If the errors are systematic, they will affect the mean value (i.e., accuracy) of top-down emissions. Therefore our statement should be correct.

3. In section 2.1, the meteorological parameters investigated in this study are only compared with surface measurements to estimate uncertainties. It has limit, since some parameters such as RH may have significant vertical gradient. Please discuss it.

Response: A discussion has been added in the end of Sect. 2.1.

4. One of my major concerns about tuning meteorological parameters off-line in GEOS-Chem is that offline tuning may cause meteorological inconsistence, since all these meteorological parameters are not independent. For example, authors tuned surface air temperature in section 4.1, it actually very likely affects the PBLH, water vapor and RH, which are fixed by authors artificially. However, authors again tuned water vapor and PBLH in other sensitivity simulations. Therefore, the NO2 sensitivities due to these tunings should not be added linearly as that authors formulated (equation 1) in section 6. The interaction effect among parameters should be accounted to correctly estimate the overall errors of model.

Response: Nonlinearity exists in the impact of errors in both meteorological and chemical parameters, but is not fully accounted for in the present study, as stated in the abstract and Sect. 6. In addition, the true value of meteorological/chemical parameters is difficult to obtain. For example, the actual vertical gradient of met fields is not known. For these reasons, we clearly stated in the abstract and Sect. 6 that the modification attempt is on the first order (i.e., considering the linear impact only) and should not be interpreted literally for error apportionment. Further study (e.g., time-consuming adjoint modeling and coupled chemistry-climate modeling) is needed to fully account for the nonlinearity effect. Nevertheless, such first-order analysis indeed provides useful information regarding model errors and impacts on NO2 columns. We further clarify the nonlinearity issue in Sect. 6 of the revised manuscript.

5. In section 2.3, MODIS AOD is known having some problems over land and overestimated compared with MISR. It's better to evaluate model aerosol simulation using MISR instead of MODIS over land.

Response: MISR measurements are sparse not to allow for a more robust statistical analysis. Our comparison of MODIS data with AERONET and CARSNET (Che et al., 2009; MODIS pixels being within 0.3 degree of ground stations) AOD over China for 2006 shows a slope of 1.01, an intercept of 0.05 and a correlation of 0.91. This suggests the validity of MODIS AOD for our purposes.
Reference: Che et al., Instrument calibration and aerosol optical depth validation of the China Aerosol Remote Sensing Network, JGR, 114, D03206, doi: 10.1029/2008JD011030, 2009.

6. In section 4, line 3 of page 8, you may not want to say that precipitation is not important for simulating NOx, since precipitation is important for wet scavenging of HNO3 and aerosols.

Response: We agree and have removed the sentence.

7. Line 29 of page 10, please give the reference that GEOS-5 simulated PBLH is overestimated in the daytime. Is your conclusion of GEOS-5 biases based on studies over the US? Can you apply it over China? Any difference between the US and China?

Response: As specified in the paper, the estimate that GEOS-5 PBLH may be slightly overestimated in the daytime is based on our analysis on temperature errors. The nighttime bias is determined mainly based on analysis over the U.S. (Liu and Liang, 2010). The negative nighttime temperature bias found for China seems to support (at least the sign of) the nighttime underestimate in PBLH.
In our subsequent post-model modification (Sect. 6), we chose not to include modifications on PBLH as we feel it is too uncertain to decide if the model PBLH is biased in the daytime. (The effect of nighttime PBLH biases is small, as stated in Sect. 6.)
We believe the PBLH issue needs further studies using observations over China, when available.

8. In section 4.5, lightning in GEOS-Chem is parameterized based on the lightning observations over the U.S. and it's very sensitive to regions and years. Authors should be very careful when discussing GEOS-Chem simulated lightning NOx over East Asia unless they can constrain GEOS-Chem lightning production over East Asia using observations.

Response: The analysis of lightning NOx can be found in detail in Lin (2012) and references therein. Lightning NOx is indeed constrained by the OTD/LIS climatology (i.e., for each month, the multi-year mean is fixed at the OTD/LIS value while allowing for variations from one year to another and from one day to another). This paper presents the main results of previous papers in discussing the effect of lightning on nitrogen burdens.

9. When authors discuss the relative importance of uncertain parameters, it should be pointed out that the relative importance of parameters are related to their uncertain ranges assigned by authors. The uncertain ranges are also quite uncertain. So please discuss it.

Response: This point is discussed in the revised manuscript (Sect. 6).

10. One of the important points of this study is to estimate the overall uncertainty of GOES-Chem in simulating NO2 column, which can be used for future inverse modeling study using GEOS-Chem. However, authors didn't report an overall error based on their estimate. How to use the estimated model uncertainties for top-down inversion? Martin et al. [2003] estimated 30% relative error of GEOS-Chem simulated NO2 column. How is it compared with authors' estimate? Which uncertainty value should people use for future inverse modeling using GEOS-Chem?

Response: Considering errors in the 10 non-PBLH parameters, the model underestimates NO2 columns by 18% in July and 8% in January averaged over East China, as stated in the abstract, Sect. 6 and conclusion. These values are compared with various previous studies in Sect. 6. They represent systematic errors of the model, while the value in Martin et al. (30%) was assumed to be random error (so that all aspects of model errors considered in their paper were added in quadrature). Note that Martin et al. used a much older version of GEOS-Chem with different met fields and chemical parameters so that results should not be compared literally between the two studies. Our recent papers suggest the errors to be 30-40% and random (e.g., Lin and McElroy, 2011).
Based on the present study and previous works, it appears that for GEOS-Chem simulations over East China, a negative systematic error of 10-20% plus a random error of 30% is appropriate, conservatively speaking. This information has been added in the revised manuscript (Sect. 6 and conclusion).

Technical Comments

1. Line 30 of page 24, "about 0-20%" to "up to 20%".

Response: The discussion on surface ozone has been dropped.

2. Figure 1 is not necessary to be included in this paper. It's not well described and discussed.

Response: Figure 1 is important for understanding the nitrogen chemistry and clarifying sensitivity simulations. We feel it should be kept. We have added a summary of NOx sinks in the introduction referring to Fig. 1.

3. Remove figure 13-16, it's better to just summarize the statistical metrics from all these figures in a table. The figures are not very informative. This paper includes too many figures and some of them are not in good quality, such as Figure 5 and 7. The contour colors can be improved to have less blue.

Response: Figs. 15-16 are dropped together with the discussion on surface ozone. We think the individual panels in Figs. 13-14 are necessary as some readers may be interested not just in the overall (averaged) errors but also errors in specific locations.
The color and figure quality issue is due to file conversion from doc to pdf – they are fine in our WORD file. This will be improved at the production stage.

4. In Figure 2, why is R_square not improved in the modified model? Does it mean that

modification doesn't improve model in terms of capturing variance of satellite retrievals?

Response: That R2 is not improved is in part because post-model modifications are too dramatic at some locations (e.g., see the left side of Fig. 2d). Also, R2 is difficult to improve since it is already at a high level (0.64 in July and 0.71 in January) and the remaining (unexplained) variance may be due to satellite errors as well as un-included model errors.

Reviewer 2

General Comments

The authors attempt to improve the agreement of NO2 column simulated with GEOS-Chem with OMI observations over East Asia. The current study is motivated by a particularly large under prediction of NO2columns over polluted regions in China. NO2columns are simulated for a range of model parameters to find the localized linear response of NO2column to each parameter. Finally simulated columns are modified based on a selected set of parameters and values according to the linear response determined. The authors select a reasonable range of uncertainties. For chemical parameters these ranges are based on recent laboratory and field studies. For meteorological parameters they are largely based on satellite and monitoring networks for meteorological parameters.

The discussion of the uncertainty of simulated NO2 over a range of model parameters is interesting, but the paper attempts to do more than that. The paper in its current form reads in part as a review, in part an uncertainty analysis, and in part an attempt at a novel means of improving simulated NO2. I highly recommend removing discussion of ozone chemistry (very little attention is given to it any way), removing discussion of the localized linear response and focusing fully on the uncertainty analysis. I suggest this paper for publication only if the authors respond to the comments above and the more specific comments that follow.

Response: This paper attempts to conduct a comprehensive analysis of key aspects of model uncertainties affecting NO2 simulations relevant to top-down constraint of NOx emissions. For this purpose, various sensitivity simulations combined with a summary of previous works (for processes/parameters not explicitly tested here) are necessary. Furthermore, while model sensitivity to individual parameters are important to know, it is also important to understand the overall model uncertainty by considering these various parameters together. We feel the reader will be very interested in the overall systematic model error. For this purpose, we make efforts to use a linearized and simplified post-model scaling approach to evaluate the overall model error. Therefore it is better to include the results of such post-model modification.

We agree with the reviewer that discussion on surface ozone is not essential to this study and have thus dropped the discussion.

Specific Comments

I. Introduction

P. 14273 L 12; "Even for areas…" to the end of the paragraph seems unnecessary and contrary to the rest of the study. Isn't the purpose of this paper to see how much of a model-satellite discrepancy could be due to parameters other than NOx emissions?

Response: The statement here says it is difficult to use measured (in situ) NOx concentrations alone to evaluate model simulations without knowledge of the accuracy of emissions. It follows that it is important to evaluate the model sensitivity/response to errors in meteorological and chemical parameters with specified NOx emissions.

P 14274 ¶1 The bulk of this paragraph should be moved to section 3

Response: The paragraph follows the prior paragraph to introduce GEOS-Chem and justify the logics of choosing the model as representative of CTMs in such top-down studies. It also introduces the study domain and time. Thus it belongs to the introduction section. Model details are indeed discussed in Sect. 3.

II. Ground and Space measurements

P. 14275 ¶1+2 Are the data filtered for clear-sky conditions? What are biases in these datasets? Is there reason to believe them over GEOS-Chem? Do you compare these datasets or GEOS to surface measurements (e.g., dewpoint)

Response: All measurement data used here include both cloudy and clear-sky conditions. Errors are small in the ground measurements (surface air temperature, RH, surface pressure, wind speed and precipitation). Satellite measurements (tropospheric water vapor path, cloud fraction and cloud optical depth) are monthly data and contain larger errors, but are expected to be much better than GEOS-5 data and have been used to evaluate assimilated meteorological datasets (e.g., see Liu et al., 2009). By comparison, assimilation data like GEOS-5 are well known to have large difficulty in representing cloud microphysics, convection and the hydrological cycle.
We use ground and satellite meteorological measurements to evaluate GEOS-5 met fields.

P. 14275 ¶3 Please briefly state how NO2 column is processed (cloud fraction, viewing zenith angle, data quality flags, etc.). For the uninitiated please state basic instrument characteristics (time of day, pixel size, repeats, etc.)

Response: As indicated in the paper, the information is presented in Lin (2012). A brief summary has been added in the revised manuscript.

P. 14275 ¶4 What is meant by valid data?

Response: 'valid MODIS data' means those indicated in the MODIS product specification to be valid. We have removed the word 'valid' not to cause confusion.

III. GEOS-Chem simulation and comparison with OMI retrievals

Please include a figure that shows simulated and observed NO2, AOD, COD at the product resolution, not just at meteorological stations in a figure. In addition, I would have benefitted in the discussion from a figure of NOx sinks (see sect. 5 comments)

Response: As stated in the paper, comparison of modeled and observed NO2 at the product resolution is presented in Lin (2012), thus it will be redundant to repeat the work here. We present AOD and COD at 284 met stations in order to be consistent with comparisons for other meteorological parameters. Comparisons on the product resolution lead to similar conclusions, so that we feel it is not necessary to repeat the work.
A figure for NOx sink pathways was presented, i.e., Fig. 1. We have added a brief summary of NOx sinks in the introduction (2nd paragraph) in the revised manuscript.

Throughout the paper, the authors refer to northwest China etc. Please either qualify these descriptions with adjectives (e.g., remote northwest China or urbanized eastern china) or designate 3-4 regional categories in a figure, perhaps the one mentioned above.

Response: 'North China' refers to a more specific area so that we have added a rough longitude-latitude range the first time it is referred to. Its name was capitalized. Other regions are more general and are called based on their relative locations in East China so that it is not necessary and more difficult to assign lon-lat ranges.
Naming 'remote northwest China' etc. may be misleading when discussing meteorological parameters (such as PBLH in Sect. 4.4), since the parameters are discussed for both urbanized and cleaner areas of a given region (e.g., northwest).

IV. Sensitivity of GEOS-Chem simulations to meteorological parameters
In general, many of the effects presented here and in section 5 are relatively uniform across the domain, rather modest and well known (e.g., water vapor). For parameters of these types, I feel the flow of the paper would benefit from fewer figures and a more succinct summary, with a table entry for the range of percentage change. Also, Please clarify that what parameters independently vary and what do not. (e.g. Sect. 4.1 –Water vapor does not change with temperature in your simulation, but biogenic emissions do)

Response: The impacts of meteorological errors, although understood qualitatively, have not been well quantified, especially over China. Also, regional dependence is obvious and is of interest to some readers. Thus we feel it is better to present figures in the paper.
Meteorological parameters, if changed implicitly by perturbation of a given parameter, are specified in the paper. We have clarified this issue in the revised manuscript.

Heading - Please include "NO2column" somewhere in the title.

Response: Done.

Sect 4.3 I don't believe sensitivity tests of cloud optical depths are pertinent for NO2 column observed in clear skies unless you have filtered model observations for 1PM clear-sky conditions.

Response: The occurrence (time, frequency, and magnitude) of clouds in GEOS-5 differs from the occurrence in satellite measurements, as can be expected for an assimilated meteorological dataset with large difficulty in representing the hydrological cycle. Selecting days with low cloud

fractions in satellite NO2 product do not mean the respective GEOS-Chem results are avoid of the impact of clouds, as can be inferred from our results.

Sect 4.4 Please separately identify effects of NO2 vertical distribution versus changes in mass due to nonlinear chemistry with regards to PBL analysis.

Response: The increases in NO_2 columns are due mainly to reduced vertical gradient of NO_2 in the lower troposphere and the subsequent application of height-dependent AK. We have added this information in the end of Sect. 4.4.

Sect 4.5 Briefly state where lightning NOx is most important. Would it affect the remote regions strongly?

Response: As shown in Lin (2012) and cited in the present paper, lightning NOx is more important in the north (north of 30N) in July. It affects NO2 abundances significantly in remote areas of Mongolia (the northern boundary province of East China).

Sect 4.5 There is no discussion of current results for lightning, only review. Please include some discussion or remove from analysis.

Response: The lightning analysis is done in Lin (2012) using the same model. The current paper presents the main results of Lin (2012) as lightning is an important factor for NO2 abundances. In the revised manuscript, we further point out the implications of lightning errors for top-down constraint.

Sect 4.5 Is northeast China a good location to study lightning NOx emissions? Are there better, more remote regions with comparable lightning activity? It seems to me that any analysis would be flawed by the affect of boundary layer transport to altitude.

Response: As a normal practice of top-down emission estimate, lightning emissions are un-adjusted when constraining anthropogenic emissions. The present study discusses the effect of errors in lightning emissions on model NOx abundances and subsequent anthropogenic emission constraint, rather than attempting to constrain lightning emissions. Constraint of lightning emissions is done in Lin (2012). The effect of vertical transport of boundary-layer NO2 on the constraint of lightning emissions is discussed in Lin (2012).

V. Sensitivity of GEOS-Chem simulations to chemical parameters
Please include “NO2column” somewhere in the heading title.

Response: Done.

Sect 5.1 – I am surprised that the increase of NO2 across the domain is uniform for a 30% decrease in kNO2+OH. I have found that NO2 simulated in NOx suppressed regions does not depend on the rate constant. When NOx dominates OH reactivity, OH should increase for any

decrease in the rate constant, and the product of kNO2+OH[NO2][OH] should be nearly constant. Please check this result.

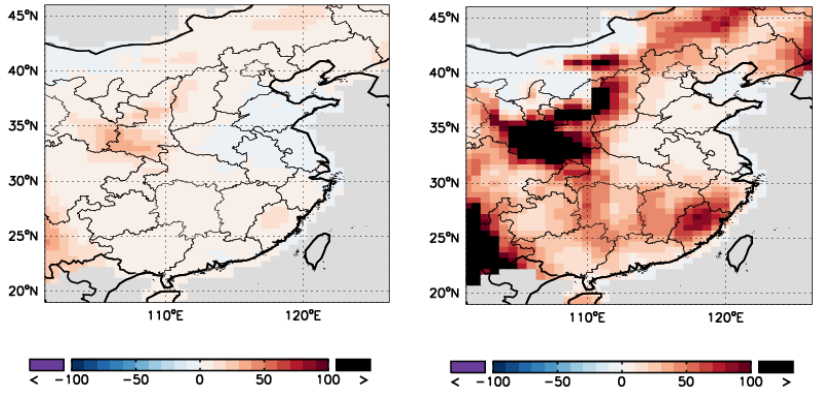
Response: The regional dependence is clearly shown in Figs. 7a,b. We have not found any bugs in our setup that may lead to the results presented.
The abundance of OH depends not only on OH+NO2 reaction but also on production from ozone photolysis, reactions of OH with CO/CH4/VOC and reaction of NO+HO2. In particular, reactions with CO and CH4 are primary sinks of OH in the troposphere. A unit of percentage change in K_OH+NO2 will not lead to the same amount of change in [OH] due to these other reactions. And changes in OH + NO2 reaction affect ozone and subsequent production of OH. (An extreme case to understand this would be to turn off OH+NO2 reaction, i.e., k decreases by 100%. In this case one can expect increases in NO2.)

Sect 5.3 – Heading should include Isoprene, OH-recycling, PAN and alkyl nitrates

Response: The heading is updated.

Sect 5.3 – The following combination of findings surprises me.
A 40% increase of OH only leads to a 0-4% decrease of NO2 column. No isoprene-OH chemistry results in a 40-50% increase of NO2 and a 100% increase of OH What fraction of NO2sinks is OH, AN, and PANs? What happens to isoprene+NO3 at night? -- I suggest adding a figure or pie chart in the beginning of the manuscript that shows simulated HNO3, alkly nitrates, and PANs to provide overview of the various NOx sinks.

Response: Below are the percentage OH changes for the two cases in July.
Areas with large OH changes are those with high VOC (isoprene) and low NOx, therefore OH+NO2 reaction is a comparatively less important sink of NOx while formation of organic nitrates (e.g., isoprene nitrates) is more important. From Fig. 9e, it is obvious that isoprene nitrate formation is a very effective NOx sink in these areas. From Fig. 7a, one can also see changes in k_OH+NO2 do not have large impacts on NO2 in these areas compared to other areas.
In these high isoprene and low NOx areas, when OH is regenerated from isop+OH reaction, it leads to enhanced OH+NO2 reaction with small impacts on isoprene nitrate formation, therefore NO2 columns are reduced. When isop+OH reaction is turned off, by comparison, the formation of isoprene nitrates is also turned off leading to increased NO2. We have clarified this point in the revised manuscript.
The nighttime chemistry is not important for afternoon NO2 columns due to short lifetime of NOx, especially in July discussed here.
A figure indicating the sink pathways of NOx has been presented (Fig.1). We further summarize the pathways in the revised manuscript (introduction, 2nd paragraph) to accommodate the reviewer's suggestions above.



100% OH regeneration NO isop+OH reaction

Sect 5.4 – Please state the regional representativeness of the observed Cu aerosol mass fractions reported.

Response: We have specified the cities reported. Initial results of our on-going project summarizing Cu concentrations across East China also support the high Cu loading and its large fraction in PM2.5.

Sect 5.6 Heading, Please change heading title to “Emissions of non-NOx species”

Response: We discussed in this section emissions of non-NOx species as well as emissions of NOx.

Sect 5.6 – Propene is not a good proxy for aromatics because of differences in alkyl nitrate formation potential and possible differences in PAN formation. Regardless, the effect seems very small.

Response: We have added this information in the revised manuscript.

Sect 5.7 – Please state a range for the magnitude and the sign of the resolution-dependent biases predicted over urban centers by the Valin et al (2011) modeling study.

Response: The importance of model resolution is time and region dependent, as shown in Valin et al. and as can be expected. Valin et al. (2011) only present limited case study (7 days in July) to allow for statistically significant evaluation. We thus feel it is not very important to provide their numbers here. Those who are interested in this specific topic are referred to Valin et al.

What happens for a 50%, 100% increase of NOx emissions? Many of these parameters tested here depend nonlinearly on NOx concentration. Please comment on their effects at NOx concentrations that are more pertinent to what is observed.

Response: The nonlinear dependence of nitrogen chemistry on NOx emissions are discussed in detail in previous studies (Martin et al., 2006; Lin and McElroy, 2011; Lin, 2012) and summarized in the end of Sect. 5.6. The nonlinearity effect is about 10% for both July and January averaged

over East China, with larger impacts in some individual locations.

VI. Modifying model NO2 columns accounting for errors in meteorology and chemistry

As mentioned previously, I don't think this section is needed. A simple table of uncertainties for each parameter and scatter plot of simulation and observation are sufficient. The authors should justify why HO2+ NO was not selected for this analysis.

Response: As discussed above, we believe this section is important as it combines most parameters analyzed in previous sections to provide an estimate of overall systematic model error. Without such work, it is difficult to have good understanding of overall model error.
Reaction of HO2+NO to produce HNO3 is discussed in Sect. 5.2 but not included in deriving the overall model error here. This is because such reaction is found only in limited lab experiments and needs to be confirmed by more lab experiments from more research groups. Considering the current uncertainty of this reaction, we are cautious about including it in the CTM. We have added a discussion in this section.

Sect 6 –P 14294 L 3. and L4 . Please remove "best estimate"

Response: Done.

VII. Implications for surface ozone.

I feel this section is beyond the scope of this paper and no observations are presented for comparison.

Response: This section has been dropped.

VIII. Conclusions

Please comment on the likelihood that the difference between model and observation is simply due to an underestimate of emissions in the context of the range of model errors presented.

Response: Such discussion is done in the last paragraph.

Technical comments:

Repeated Instances:

-Please say increase and decrease by 30% instead of scaled by 130% in many examples throughout the manuscript. Also, please check that all captions match the image shown (e.g. Fig. 7 NO2increases but the caption says the NO2removal rate constant is scaled 130%)

Response: 'Decreased by 30%' is different from 'scaled down by 130%'. For example, a value of one would become 0.7 if decreased by 30%, but would become 0.77 if scaled down by 130% (i.e., 1/1.3).
The captions/figures are fixed.

-Readability would be improved by reducing the number of works cited and using (e.g.,

⋯) formalism.

Response: Done.