Atmos. Chem. Phys. Discuss.,
https://doi.org/10.5194/acp-2020-307-RC1, 2020

Atmospheric Chemistry and Physics Discussions

Interactive comment on "Improving NO_2 and ozone simulations through global NO_x emission inversions" *by* Zhen Qu et al.

Anonymous Referee #1

Received and published: 29 April 2020

Review of 'Improving NO2 and ozone simulations through global NOx emission inversions' by Zhen Qu et al.

General comments

Qu et al. have studied the impact of top-down NOx emission estimates derived from two OMI NO2 satellite data sets (NASA SP v3 and DOMINO v2) on NO2 and O3 simulations with the GEOS-Chem model. Previous work already showed (e.g. Verstraeten et al. [2015], studies by Miyazaki et al.) that O3 in the troposphere is generally better understood when NOx emissions are derived from satellite NO2 data than when taken from emission inventories.

Here, Qu et al. find substantial differences in the agreement of NO2 and O3 simulations

against independent measurements depending on whether data set NASA or data set DOMINO is used. This was to be expected given that it is well-known that the NASA and DOMINO datasets have considerable differences. A useful aspect of the study is that the authors now quantify the consequences of these differences, which is relevant because satellite data is increasingly used to improve model understanding of atmospheric composition.

What is disappointing however is that we do not learn much new. Simulations with the NASA emissions compare better to some metrics, and worse to others, but the authors do not explain why. This makes the manuscript a technical document, where it is left to the reader to figure out what emissions could work best for his/her particular purpose, without actual guidance on why that would be. The authors should do more to investigate why using one dataset leads to better agreement e.g. for surface O3 at remote sites, and the other for polluted sites. Aspects of spatial resolution, temporal representativeness, and vertical sensitivity should be taken into account when providing this guidance to potential users.

Another criticism is that the chain of technicalities is very long and that the experiments are set-up in a sub-optimal manner (for example comparing 2.5 degree simulations of surface NO2 to surface stations that are representative for much smaller domains). A major concern I have is with the lack of detail and clarity on how the adjoint incorporates the information from the satellite retrievals. From the manuscript I first suspected that monthly mean column NO2 data was simply used to estimate the emissions, suggesting that the highly variable and non-linear vertical sensitivities of the retrievals have not been used to interface the model with the satellite data. There are various studies pointing out how crucial it is to account for the vertical sensitivity of the NO2 retrievals, e.g. Miyazaki et al. [2017], Boersma et al. [2016] to name a few. Then I read the supplementary material and there the impression was given that at least the a priori profile shapes are made consistent between the NASA and DOMINO retrievals, but it remains unclear to what extent this has harmonized the data, and to what extent

vertical sensitivities between the two datasets are still fundamentally different.

Specific comments

P2, L42-43: the formation depends not only on the local NOx and VOC concentrations, but also on the radiative regime in which these occur.

P2, L65: different → differ

P3, L72: import →importer

P3, L78-81: Zhang et al. [2008] and Verstraeten et al. [2015] already showed that through optimizing NOx emissions in China, the simulated O3 over the Pacific and over the western US indeed improved.

Section 2.1 It is unclear in this manuscript how the adjoint accounts for (a) vertical sensitivity of the satellite retrievals, and (b) the diurnal cycle of NOx emissions. These aspects are important enough to describe in the manuscript, for (a) useful information is provided in the supplement but it is not clear whether the replacing of the a priori profiles by GEOS-Chem prior profiles was also applied in the research to estimate the emissions. The authors should clarify this in section 2, and also briefly quantify to what extent the differences in prior simulations have been minimized by this approach. For (b), some info is given but only late in the game (P7: The diurnal variations of NOx emission are constrained to be those of the prior emissions), and we do not learn what the diurnal cycle is in the first place. Please revise section 2 thoroughly with this in mind.

Then I have other questions:

- how does the adjoint approach account for other relevant aspects of data assimilation?

- how is the OMI data averaged spatially to the grid of GEOS-Chem, and how are superobservation errors incorporated?

- did the authors only use the mostly cloud-free OMI retrievals?

Section 2.2: OMI is suffering from the so-called row anomaly, which was absent until mid-2007, and then became gradually more important. How did the authors ensure that the growing impact of the row anomaly did not unduly affect their trends in NOx emissions?

Section 2.3: it remains unclear what type of surface station was used for the GEOS-Chem surface evaluation. Using urban background and regional stations seems appropriate to evaluate the large GC grid cells, but urban street stations should be excluded.

P5, L152-154: what explains the OMI-driven differences between the posterior NOx emissions, differences in tropospheric slant columns or in the AMFs? Presumably the latter, but since the a priori profile differences have been “minimized”, the differences must be in the assumptions on surface albedo and clouds. It would be best if the authors could shed more light on how the scattering weights or averaging kernels are different between the OMI NO2 retrievals. Please clarify.

P6, L173-174: the statement that “NO2 column simulations at $2^\circ \times 2.5^\circ$ in this study are likely to be underestimated and lead to high biases of posterior NOx emissions to match satellite NO2column concentrations” needs more evidence. The hypothesis that instant dilution leads to too much OH (by Valin et al. [2011]) may be valid for isolated NOx sources in otherwise pristine areas, but instant dilution of NOx emissions situated in high-background NO2 regions such as the eastern US or western Europe is probably of less concern.

P6, L193: what is the magnitude of the correction factors over China and the US? How do they vary by season?

P7, L195-199: this part is rather inconclusive. The GEOS-Chem simulations have been corrected for resolution (an increase) and surface measurements have been corrected down for molybdenum interference, and still GEOS-Chem with posterior emissions is

biased low by 20%-50%. What explains the persistent low bias?

P7, L224-225: OMI measurements frequently miss the high values of NO2 column densities that occur before or after its overpassing time. OMI was never designed to measure NO2 before or after its overpass time, so to say that OMI misses these high values is misleading. Please rephrase.

P7, L226: twice-per-day constraints on NOx emissions have been achieved in earlier studies based on SCIAMACHY + OMI (Boersma et al. [2008], GOME-2 + OMI [Lin et al., 2011], including via sophisticated assimilation schemes [Miyazaki et al., 2017].

P8, L237: the June peak in NO2 over China can be easily traced back to crop residu burning in that month – e.g. Stavrakou et al. [2016].

P8, L238-240: can you explain more why the DOMINO product would be more sensitive to soil NOx emissions? It's not because of the different a priori profiles assumed in the NASA and DOMINO retrievals?

P8, L243-244: please see my previous comment. The authors seem to know something very interesting here, but they don't show it. Is there any evidence that one retrieval would be more sensitive to NOx sources than the other? That would be extremely relevant to know more about. Since the satellite measurements are identical for the NASA and OMI retrievals, it must have to do with AMF differences , see e.g. Lorente et al. [2017]. But what drives the apparent difference in sensitivity – albedo, cloud fraction, cloud pressure?

P8, L246-256: Figure 5 – the daytime O3 simulations in China all seem strongly low biased relative to the observations. The other ozone metrics in China and all in the US match much better. Why is this?

P9, L271-272: "also not reflected"?

P9, L276: no reduction of NOx emissions in Europe? This is strange – NO2 tropopsheric columns are decreasing over Europe, and Miyazaki et al. [2017] showed reduc-

tions in for NOx emissions. Overall, Figure 6 looks very odd to me. DOMINO NO2 columns are 40% higher than NASA, but the NOx emissions inferred from DOMINO are more than 40% higher than the emissions inferred with NASA (L278-281). Also, Miyazaki et al. [2017] (Figure 9) still find reductions in NOx emissions over Europe between 2005 and 2014 based on the same DOMINO data, so how can you find increases? Please clarify.

P10, L295-297: I'm missing an explanation or hypothesis why NOx emissions from one dataset would do better than the other for different ozone metrics.

P10, L304 and L315: please clarify how the impact of meteorology and non-NOx sources on O3 changes was evaluated.

L306-307: "The trends of simulated MDA8 ozone are similar when using the NASA and the DOMINO posterior NOx emissions as inputs" – yes, but please also explain why the magnitude of the NASA-derived MDA8 O3 levels are biased high then.

P11, L332-333: the prior simulated O3 profiles in Figure 10 agree much better with the O3 sondes between 800-400 hPa than the assimilated profiles. I don't understand why that is, since the effect of the updated NOx emissions should be mostly felt in the lower 2 kms of the atmosphere. Or is this the impact of changes in background O3 in response to changing Asian emissions?

P13, L394-395: one important difference between this research and the work done by Miyazaki in a number of papers, is that the latter assimilates also other species relevant for NOx inversions and O3 simulations (e.g. CO, HNO3, SO2). It would be interesting to also discuss to what extent these additional constraints can help explain the "remaining differences between simulated and measured ozone".

P13, L398-400: the statement "Both OMI NO2 retrievals employed in this study use NO2 vertical shape factors from coarse resolution simulations, and therefore are biased low compared to in-situ measurements [Goldberg et al., 2017]." Brought up the

question (again) whether both OMI NO2 retrievals are at least consistent now in their use of the same coarse-resolution vertical shape factors (i.e. those from GEOS-Chem).

P13, L401: “retrievals also have not explicitly accounted for the aerosol optical effects, which are demonstrated to degrade the accuracy of NO2 column concentrations”. This is an overstatement. Only when AOD is very high (>0.5-1.0) there are indications that implicit corrections break down. Even in Liu et al. [2019] accounting for AOD did not solve the low bias in tropospheric NO2 which was not apparent in the DOMINO scheme without an explicit aerosol correction.

References

Boersma, K. F., Vinken, G. C. M., and Eskes, H. J.: Representativeness errors in comparing chemistry transport and chemistry climate models with satellite UV–Vis tropospheric column retrievals, Geosci. Model Dev., 9, 875-898, doi:10.5194/gmd-9-875-2016, 2016.

Miyazaki, K., Eskes, H., Sudo, K., Boersma, K. F., Bowman, K., and Kanaya, Y.: Decadal changes in global surface NOx emissions from multi-constituent satellite data assimilation, Atmos. Chem. Phys., 17, 807–837, https://doi.org/10.5194/acp-17-807-2017, 2017.

Stavrakou, T., Müller, J.-F., Bauwens, M., De Smedt, I., Lerot, C., Van Roozendael, M., Coheur, P.-F., Clerbaux, C., Boersma, K. F., van der A., R. J., and Song, Y.: Substantial underestimation of post-harvest burning emissions in the North China Plain revealed by multi-species space observations, Sci. Rep., 6(32307), doi:10.1038/srep32307, 2016.

Interactive comment on Atmos. Chem. Phys. Discuss., https://doi.org/10.5194/acp-2020-307, 2020.