

Review of Judd et al. -- Evaluating the impact of spatial resolution on tropospheric NO₂ column comparisons within urban areas using high-resolution airborne data

The authors investigate the impact of spatial variability on correlative studies for the validation of satellite trace gas products with ground-based instruments. High resolution airborne imaging DOAS measurements from GeoTASO, ground-based Pandora, and two OMI satellite products are used.

The paper is generally well written and of significance for the validation of satellite trace gas retrievals. I therefore recommend publication in AMT after some minor revisions.

General comments

- The introduction should contain an overview of existing airborne imaging DOAS systems
- Information about the campaign is scattered in the manuscript. A solution could be a campaign sections, with a description of the target sites (urban/rural, # of inhabitants, industrial emitters...), as well as a description of the measurement conditions (Date, time of day, SZA, AOD, meteorology...) maybe as a table...
Here you should also give an overview of the flights presented in this study to help the reader
- I could not find any information about the DOAS fit settings used (except the fit window). Please provide that information (cross-sections, I0, Ring, ...), e.g. in a table.
- You often refer to differences in spatial resolution of the a priori inputs. It would be nice if you could provide the spatial (and temporal?) resolution of the SP and BEHR products.
- You often state that the Pandora Pandora measurements are representative up to a certain pollution scale. In my opinion this statement is not correct. The representativeness depends on the spatio-temporal variability of NO₂ at the Pandora location. You use an NO₂ threshold to filter out data with large variability, but the magnitude of NO₂ itself is not an indicator for the representativeness. I think you should amend the manuscript to reflect the differences between the physical reasons (variability) and the methodology (filtering by threshold).

Detailed comments & technical corrections:

Page	Line	Comment
2	26ff	You mention: "development of [...] instruments" but you then only write about GeoTASO. I think you also had GCAS in mind. I suggest to explicitly mention it. Here you could also refer to other instrument previously used.
2	35	"... such as NO ₂ ." NO ₂ is not a product, but a chemical species. Suggestion: NO ₂ tropospheric vertical column densities.
3	21	What is the field of view in degrees?
4	3	Is the spectral resolution constant over the spatial dimension? If not, how does it vary?
4	7	How many spectra are co-added for the 250m (or what is the speed of the aircraft)

4	13-18	I think this paragraph could be moved to the 'campaign' section suggested in the general comments
4	42	Are the inputs for the RT simulations generalized assumptions or do you perform specific calculations for each flight?
5	31	Do you account for diurnal changes in the stratospheric column, or do you assume a fixed value per campaign site?
7	5	"subtle influence of a varying NO ₂ shape factor is visible in the AMF", I assume you are referring to the rectangular pattern above the lake. Do you consider a change of ~50% to be subtle? Is this pattern coming from the CMAQ model grid boxes? Could it be that these patterns are caused by averaging of flights performed under varying geometries? It would be nice to also see the flight tracks. Maybe you can add them as lines in Figure 1.
7	40	Why do you use DU now? Please also write the molec / cm ² , e.g. 0.05DU (1.34e15 molec / cm ²)
8	13	What is the resolution of the SP and BEHR a priori data inputs?
12	17-20	For a larger AMF the a priori profile must be shifted towards higher altitude. Or in other words, the NO ₂ is in higher altitudes than assumed by the model. Is an uplift of NO ₂ likely in a sea breeze front? Please briefly explain the mechanism for an uplift of NO ₂ .
12	30	There is a small hill ~200-300 m, in the area around "CalTech" and "LA main street". Could it be that there are issues related to the surface air pressure in the RT simulations? No need for a detailed discussion, but you should have a look at the pressure profiles for this area.
13	F7	Please also add the date as a title in the figure as you did in figures before
14	F8	Please also add the date as a title in the figure as you did in figures before
15	6	"during these rasters": I think "rasters" should be replaced by flights / flight patterns or similar.
15	30	Do you have an explanation why the magnitude of NO ₂ levels is so different between L.A. / Chicago; or also between the different days. Is that related to wind speed? You should provide some information, see general comments.
15	34	How exactly do you do the coincidence analysis? In the co-added rasters you cannot take the median in a 750m radius. What is the GeoTASO time – the average time?
16	1	"solely" is a strong word. I am sure there are other minor reasons for mismatches. Maybe better use "driven by", "mainly caused by" or similar.
17	14	I do not agree, that the representativeness of Pandora measurements depends on the pollution scale, such as a threshold value. As you write in line 9ff. the Pandora measurements are representative in areas with small sub-satellite-pixel variability. It is true that small variability is usually found at stations in background (non-urban) regions with low pollution levels. Filtering by pollution levels is a very basic and simple approach to the problem of spatial representativeness. Though applying refined approaches it may be out of scope for this study, I think you should provide ideas how to improve this idea.
19	22	Do you have an explanation why the correlation decreases for large pixel sizes and low resolution a priori data?
20	32	"pollution scale appears to be..." for the investigated times and areas
20	34	Correct line break in unit
