

Interactive comment on “Spatial and temporal representativeness of point measurements for nitrogen dioxide pollution levels in cities” by Ying Zhu et al.

Anonymous Referee #1

Received and published: 15 April 2020

This manuscript presents an approach to merge a combination of mobile and stationary in-situ observations of NO₂ with ground-based remote sensing of NO₂ to produce a temporally consistent NO₂ concentration surface within an urban area at 20m x 20m spatial resolution. Using this surface, they infer the frequency of exceedances during the sampling campaign relative to current WHO NO₂ guideline concentrations. The authors discuss mismatches between the sampling location at typical monitoring stations and concentrations observed on roads and at street level, and conclude that a variety of monitoring techniques are important to derive representative air pollutant concentrations.

C1

The authors can be commended for their attempt to derive an approach that synthesizes NO₂ observations across multiple relevant scales, from the micro environment to integrated urban scales. They discuss several important insights that may be of interest to other investigators who measure NO₂ with in-situ and remote sensing techniques. There are a number of novel measurements described here, and the authors have put in a great deal of effort to interpret them clearly. However, several elements of their methods and analysis could benefit from further clarification and evaluation, as I outline below. I have a number of general concerns that should be addressed before publication, followed by a list of specific comments.

General comments:

(1) Mobile measurements: More details would be very helpful. It's not clear at all how many, when, and what routes were taken for the mobile measurements. The authors simply state that mobile measurements “were performed in June and July 2016” and that “varying routes” were taken. Were mobile deployments performed every day over those two months? What were the routes like? I suggest the authors at least provide a table of the dates and times that mobile measurements were made, and a map of each of the routes (unless they were all identical, or the number of mobile deployments was truly large) as Supplementary Information, and include a short description in the main text (at least N=?). Since I assume the route was not identical in each case, it might be useful to include a panel in Figure 3 showing how many spatially overlapping measurements are included in each pixel average (i.e. how many days are included in each 20 x 20m average?). Also, it's not clear until much later in the manuscript that the authors explain these observations are gridded to 20 x 20 m. This should be included in the description here.

(2) Long path DOAS: I'm also a little confused about how all the different LP-DOAS measurements were used. The methods describe four different absorption paths, shown in Figure 1. But when/how is each absorbing path used? It seems to me that the LP-DOAS measurements are only used in the normalization step (Section 3.1.1).

C2

Were all the LP-DOAS measurements averaged together to get each day's diurnal cycle? If so, how? Or were different absorption paths used for different parts of the city? Or were the results from only one absorption path used? (If the latter, why are all paths described and shown?)

(3) Diurnal normalization: In my reading of this manuscript, the algorithm for diurnal "normalization" based on LP-DOAS measurements above the city street levels strikes me as one of the central developments presented in this manuscript, underpinning their main results. But the approach is not evaluated in any way. In my opinion, some evaluation is warranted, since it is a key aspect of their final results. I think a satisfactory evaluation could be quite easy to accomplish using the available ground-based local air quality monitoring data for which you have full diurnal observations. For example, I would propose: Why not take a random sampling of NO₂ measurements at various dates and times from the stationary monitoring sites, and see if the correct diurnal profile/24-hr average can be retrieved based on the LP-DOAS normalization for that date. Repeating this experiment for, say, 50 different days from different times of day, at all three of the monitoring stations at coincident times, could quite nicely result in statistics describing how accurate this diurnal normalization approach is, depending on the time and location sampled (at least for these stationary cases). Also, just to be clear, the output of this algorithm is essentially a "retrieved" 24-hour average concentration from a measurement at any instant in time, correct? I think this is worth spelling out very clearly. In other words, instead of calling it a "normalized" concentration, why not call it an "inferred daily average" or something like this? This would be much more intuitive. Perhaps these are not mathematically the same? (But if not, why can these numbers be compared to the WHO 24-hour average guideline concentration?)

(4) Comparison of NO₂ concentrations at different heights and distances: I may not have understood correctly, but is this section based on measurements from a single monitoring station over two individual days? At the beginning of Section 3.2, the authors state "measurements were conducted over two days" in March, but then later in

C3

the same section, the authors state "we repeated the 1.5 m to 4 m measurement height comparison on several days at different seasons and derived the same factor". This is confusing and requires clarification. Please be specific regarding the deployments. As written, it seems like the authors have derived the rules for vertical and horizontal NO₂ diffusion based on one specific monitoring location from two days in March, and applied that generally for their mobile results across the whole city taken in June and July. Wouldn't the dilution and chemistry of NO₂ be specific to each micro environment, where the monitoring is done? Buildings, wind patterns, light availability, I would expect to all play a role in the spatial scale of NO₂ concentrations at such small scales, which would vary depending on location. Likewise, I'm surprised that these spatial scales are independent of local meteorology, and seasonality (driving O₃ concentrations and other factors for example). I think this requires some evaluation, clarification, and some reasonable caveats. Finally, isn't there prior work to cite, related to how NO₂ is observed to decay by distance from, say, a highway? How do your results compare? Are the spatial scales that different?

(5) Comparison with OMI: It's not clear to me what is gained from the comparison with OMI in Section 3.3, and how it is relevant to this study. Essentially, the authors have demonstrated that the LP-DOAS and OMI observations share seasonality. They also conclude that the OMI-inferred values are lower than the LP-DOAS values. I don't see much new here, but perhaps the authors could clarify. How exactly does this comparison support the main conclusions of this manuscript, which involve extremely small spatial scales within a small period of time that would not be well captured by OMI?

(6) Main result: In my opinion, one of the main results of this manuscript is how many exceedances of the WHO guideline would actually be inferred with a more dense monitoring strategy, compared to the sparse observations currently available. But this result gets glossed over in the rest of the discussion and concluding remarks. I'd be curious if the authors agree that this one of the most important results, and if they think it

C4

deserves more prominent position in the abstract/conclusion.

Specific comments:

Introduction:

Page 2, Line 11: "NO₂ levels are often strongly correlated with many other toxic air pollutants". Perhaps this is indeed well-known, but I think this statement merits some citations.

Page 2, Line 25: "measures" should be "measure"

Page 2, Line 25: The authors refer to "defined rules" in terms of measurement height and distance in several places throughout the manuscript, but have mentioned what these rules are anywhere. I assume these "rules" may vary by jurisdiction. Are there consistent, e.g. European-wide, rules? Does the WHO provide location guidelines? I think at least providing the "rule" for the local agency would be insightful to share somewhere. Even better, share the rules for other places if they vary.

Page 3, Lines 15-20: It seems to me that certain portions of the introduction, including these lines, would be better placed in the Methods section of this manuscript

Methods:

Page 5, Line 13: The names of each urban station means very little to someone unfamiliar with the region. I later discovered these locations are indeed on the map in Figure 3c. I suggest the authors mention here that these locations can be found on this map.

Page 6, Lines 4-11: The authors explain that the OMI VCDs are gridded onto a 0.02 x 0.02 grid following approaches from previous work, and later in the manuscript it looks like the authors use monthly averages. Can the authors be slightly more clear here? Are the observations from each day gridded separately, then I assume simple monthly averages are calculated from these daily surfaces?

C5

Page 6, Lines 13-20: Can the authors be clear about what model profiles are being used? I.e., what time of day, and I assume from the same month/year as the satellite observations? Is the surface estimate calculated each day, prior to calculating the monthly average? Or is the surface estimate performed on the monthly average columns? I believe the authors answered these questions in the response to the quick review, but why not include these details in the text for the benefit of readers? Regarding the conversion to surface concentrations, can the authors describe the logic behind using spatially interpolated profiles across four grid cells? Doesn't this represent interpolating over an area of 4x5 degrees? I am curious why they would not simply use the spatially coincident model grid cell.

Page 6, Lines 21-22: I'm not sure I agree entirely here. What if the diurnal pattern of sources or chemistry is different over the area of the satellite pixel, compared to the area covered by the LP-DOAS absorption path. Wouldn't that cause a mismatch between the diurnal cycle integrated over each?

Results and Discussion:

Page 7, Line 11: "as a general indicator". An indicator of what?

Page 7, Line 13: Why have the concentrations been averaged over the three locations in this case, which presumably represent quite different micro environments? Why not just report the average concentration of each separately, and within each of their 10-km radius? Section 3.1.1: As I mention above, this really strikes me as belonging in the Methods section. This means you could also keep Figure 2 before Figure 3. As it is now, you reference Figure 3 in the text before discussing Figure 2.

Figure 3c: You could be very clear here that this is a zoomed-in plot of the diurnally normalized data.

Table 1: You could be very clear here that the WHO AQG is according to the 24-hr average guideline.

C6

Page 10, Line 17: This is assuming that the WHO guidelines are based on the same monitoring “rules” that are applied by the environmental reporting agency. Is this true? What are the WHO guidelines for monitor placement? Are these the same rules employed by the local air quality agency? Please share these rules explicitly for the benefit of the readers.

Interactive comment on Atmos. Chem. Phys. Discuss., <https://doi.org/10.5194/acp-2019-1198>, 2020.