Review of Cooper et al.

March 25, 2020

Cooper et al. "Effects of *a priori* shape assumptions on comparisons between satellite NO_2 columns and model simulations" presents a case study using synthetic OMI observations and the GEOS-Chem adjoint to show how different methods of accounting for the vertical sensitivity of satellite NO_2 measurements when comparing to model NO_2 fields affects emissions inferred from said comparison. This paper may have been a better fit for GMD rather than ACP, as it primarily touches on model comparison, but is also relevant to the remote sensing community as it informs what information must be contained in satellite products for effective model comparison, and to the broader atmospheric community for understanding possible sources of error in constrained emissions, so ACP is also appropriate.

I do have two concerns about the experimental design. First, I question whether using scattering weights computed for average OMI observing geometry in a $4^\circ \times 5^\circ$ grid cell is appropriate for creating synthetic observations. Second, some of the choices for prior test cases are not very relevant to current NO_2 retrievals. I will address these more below. If the authors can address these concerns, then this manuscript should be published in ACP.

Major concerns

Representativeness of average scattering weights

In sect. 3.2.1, line 226, the authors say: "To represent typical conditions, average scattering weight profiles for each grid box are found by averaging scattering weights for OMI observations during July 2010." I have two questions about this. First, it is ambiguous whether the mean scattering weights in question are found by averaging weights providing in the product (implied by "...average scattering weight profiles for each grid box are found by averaging scattering weights for OMI observations during July 2010...") or by using the average viewing conditions to compute the scattering weight vector for the average viewing angles, albedo, surface pressure, etc. (implied by "...OMI scattering weights are calculated using the LIDORT radiative transfer model (Spurr, 2002) by providing LIDORT with the observation geometry of the OMI observations and aerosol profiles from the GEOS-Chem base simulation..."). I'm assuming the latter, but this could be made clearer.

Assuming that the authors calculated their own scattering weights from the average viewing conditions, my second concern is that while this simplifies the problem of computing synthetic observations for each model step, it may not adequately represent the variation in OMI measurements during that time. Scattering weights depend nonlinearly on observation geometry, so:

$$w(z|\overline{\theta}_s, \overline{\theta}_v, \overline{\phi}, \overline{a}_{\text{surf}}, \overline{p}_{\text{surf}}) \neq \frac{1}{c_{\text{obs}}} \sum_{i=1}^{c_{\text{obs}}} w(z|\theta_{s,i}, \theta_{v,i}, \phi_i, a_{\text{surf},i}, p_{\text{surf},i})$$

where θ_s is the solar zenith angle, θ_v the viewing zenith angle, φ the relative azimuth angle, a the surface reflectivity, p_{surf} the surface pressure, and overlined quantities represent grid cell averages. In other words, the vector of scattering weights corresponding to the average observation conditions is not guaranteed to be the same as the average of scattering weight vectors for all individual observations. Concretely, consider two observations, one with a viewing zenith angle of 0° and one at ~ 60°. The average of these two observations' scattering weights is not guaranteed to be the same as for an observation of 30°.

That being said, it may well be close enough, especially averaged over a 4°×5° grid cell. If the authors can show that the difference between using mean scattering weights and the mean of synthetic observations computed using individual OMI observation scattering weights is within the measurements' uncertainty for at least a few days of synthetic observations, then I think that would be adequate.

Thank you for your comments. The method we use to calculate average scattering weight profiles is to provide LIDORT with observation geometry from individual OMI observations, and then average the resulting scattering weight profiles. We have edited the text to clarify this procedure (Line 224):

"Scattering weights are calculated using the LIDORT radiative transfer model (Spurr, 2002) by providing LIDORT with the observation conditions of OMI observations during July 2010, which are used to represent typical viewing conditions of low earth orbit satellite observations, and aerosol profiles from the GEOS-Chem base simulation. To represent typical conditions, these representative scattering weight profiles for each grid box are used to produce the synthetic slant columns."

The difference between using an average scattering weight profile and using individual observation scattering weights is indeed small when averaged over a 4°x5° grid. We now discuss this on Line 229:

"Tests performed for all 4°x5° grid boxes used here indicate that the mean relative difference between an air mass factor calculated using an average scattering weight profile and the average of air mass factors using observation-specific scattering weight profiles is less than 4%."

Relevance of prior test cases

Of the shape factor test cases described in sect. 3.2.2, the SF_{trop} and SF_{BL} cases are not particularly relevant for satellite measurements. Both of the two main global OMI NO_2 retrievals (NASA SP3, Krotkov et al. 2017; QA4ECV, see Williams et al. 2016 for the NO_2 profiles) use ~ 1° resolution for their NO_2 profiles. Therefore, the SF_{BL} and SF_{trop} cases, which assume one profile globally are not representative of any major satellite product. In fact, the more relevant question, assuming that 4°×5° is still a common resolution for adjoint modeling, is what happens if the satellite prior is *higher* resolution than the model profile.

In my opinion, a sixth test case similar to SF_{prior} but using a set of priors from a 2°×2.5° or 1°×1.25° GEOS-Chem simulation would add value to the paper by studying the effect of having the satellite product's prior at higher spatial resolution than the adjoint model. Also, if the SF_{BL} and SF_{trop} cases are retained, it should be clearly stated that they represent extreme cases that do not represent any modern NO_2 product.

Thank you for this suggestion. We now include a test SF_{finer} that uses a set of priors based on a 2°x2.5° GEOS-Chem simulation. We also note in Section 3.2.2 that the SF_{BL} and SF_{trop} cases are extreme cases that do not represent any modern NO_2 product (Line 264):

"The SF_{BL} and SF_{trop} tests do not represent any modern retrieval algorithms, but are used as extreme examples of using an *a priori* that assumes no spatial variability"

Other primary concerns

- **In sect. 3.2, line 210, the authors say that they use one observation per grid box per hour. But OMI will only observe a given location twice per day, maximum, and usually only once per day at about 13:30 local standard time. Are you then filtering these once-per-hour observations down to the ones OMI would actually observe?**

  OMI viewing geometries are used here only as an example of typical viewing geometries of low earth orbit satellite instruments for the scattering weight calculations. The synthetic observations used here are not meant to be synthetic OMI observations or represent the spatial or temporal sampling of OMI. We have clarified this in the text at line 224 as previously noted:

  "…by providing LIDORT with the observation geometry of OMI observations during July 2010, which are used to represent typical viewing geometries of low earth orbit satellite observations, and aerosol profiles from the GEOS-Chem base simulation."

- **I don't follow Eq. (20). Specifically why n_a shows up on the right hand side. Given that:**

$$M(\mathbf{n}) = \frac{\sum_i \mathbf{w}_i \mathbf{n}_i}{\sum_i \mathbf{n}_i}$$

and

$$\mathbf{A}_i(\mathbf{n}) = \frac{\mathbf{w}_i}{M(\mathbf{n})}$$

then to compute $M(\mathbf{n}_m)$ given $\mathbf{A}(\mathbf{n}_a)$ you only need to multiply $\mathbf{A}(\mathbf{n}_a)$ by $M(\mathbf{n}_a)$ to extract the necessary scattering weights:

$$M(\mathbf{n}_m) = \frac{\sum_i \mathbf{w}_i \mathbf{n}_{m,i}}{\sum_i \mathbf{n}_{m,i}}$$

$\mathbf{w}_i = \mathbf{A}_i(\mathbf{n}_a) \cdot M(\mathbf{n}_a)$

$$\therefore M(\mathbf{n}_m) = \frac{\sum_i \mathbf{A}_i(\mathbf{n}_a) M(\mathbf{n}_a) \mathbf{n}_{m,i}}{\sum_i \mathbf{n}_{m,i}} = M(\mathbf{n}_a) \frac{\sum_i \mathbf{A}_i(\mathbf{n}_a) \mathbf{n}_{m,i}}{\sum_i \mathbf{n}_{m,i}}$$

I don't think you need $\mathbf{n}_a$ to compute $M(\mathbf{n}_m)$ as long as $M(n_a)$ is included in the satellite data (both the NASA SP3 and QA4ECV OMI NO_2 products include the tropospheric AMFs), and given only the AKs and prior profile, it would be difficult if not impossible to compute $M(n_a)$. That means the statement on line 320 about needing the a priori profiles in the dataset is incorrect.

Thank you for noticing the error in Equation 20. The n_a term on the right hand side should indeed be n_m. The text at line 341 has been adjusted accordingly:

"This is most straightforward when scattering weights (rather than averaging kernels) are provided alongside retrieved column data, as scattering weights and shape factors are independently calculated, however simulation-based air mass factors can be calculated using the averaging kernel and *a priori*-based air mass factor via Eq. 19."

Minor corrections

- **For Eq. (8), it would be good to make clear that A(*z*) is the column averaging kernel and therefore a vector, since in Rodgers and Conner, the capital A is typically the full AK matrix. But I agree that following the convention in Eskes and Boersma (2003) is best.**
  We now clarify that A(z) is the column averaging kernel on Line 109.

- **Eq. (9) doesn't seem to be used anywhere else in the paper, and technically is inconsistent with the implicit definition of w(*z*) in Eq (8). Recommend removing Eq. (9).**

  This equation has been removed.

- **In Eq. (17) and Eq. (18) it's unclear what is being summed. Recommend using *I* subscripts to make clear what terms are iterated in the sum.**

  This has been changed.

- **On line 198 in sect. 3.1, the authors say that Δ is computed using either Eq. (12) or (17). Given that much of sect. 2 was spent establishing that these two equations differ, this should be clarified. If I understood correctly, which equation is used effectively depends on which shape factor was used for a given test. If so, I recommend saying that explicitly.**

  This is now explicitly stated on line 193: "(Δ, from either Eq. (11) if using a simulation-based air mass factor or Eq. (16) if using the retrieval *a priori*-based air mass factor)"

- **For the different shape factors, have you considered the impact of profiles simulated by a model with systematically, rather than randomly (as in SF_{n30}), different emissions?**

  Thank you for this suggestion. We now include a test SF_{diffem} that considers *a priori* profiles simulated by a model with systematically different emissions.

References

Eskes, H. J. and Boersma, K. F.: Averaging kernels for DOAS total-column satellite retrievals, Atmospheric Chemistry and Physics, 3, 1285–1291, doi:10.5194/acp-3-1285-2003, URL https://www.atmos-chem-phys.net/3/1285/2003/, 2003.

Krotkov, N. A., Lamsal, L. N., Celarier, E. A., Swartz, W. H., Marchenko, S. V., Bucsela, E. J., Chan, K. L., Wenig, M., and Zara, M.: The version 3 OMI NO2 standard product, Atmospheric Measurement Techniques, 10, 3133–3149, doi:10.5194/amt-10-3133-2017, URL https://doi.org/10.5194/amt-10-3133-2017, 2017.

Williams, J. E., Boersma, K. F., Sager, P. L., and Verstraeten, W. W.: The high-resolution version of TM5-MP for optimised satellite retrievals: Description and Validation, doi: 10.5194/gmd-2016-125, URL https://doi.org/10.5194/gmd-2016-125, 2016.