Review of An et al.

September 17, 2020

Thank you to the authors for responding to my second round of comments. At this stage, all but one of my scientific concerns have been addressed. If this remaining concern is addressed, this paper should be published with minor revisions—the paper would benefit substantially from editing for clarity before final publication, but publication need not be contingent on this.

Specific comments

- One of my previous comments was:

  > *A separate concern to how dependent the spline interpolation is on the coordinates is how well the spline reproduces known points. So a better experiment might be to remove a set of points from the data the spline is fit to (e.g. remove all points of a specific SZA value), fit the spline to the limited dataset, and test how close it comes to those known points (similar in concept to withholding test data during machine learning training).*

  In their response, the authors say:

  > *We are not sure that which one of the specific data can be deleted without destroying the continuity of the original dataset. In the other words the specific data basically do not exist in this research because the DSHS is not officially working.*

  If the original dataset is so fragile that removing a single point makes it unusable, that raises a great deal of concern over how robust this method will be in practice. Additionally, unless I misunderstood rather badly, this test should not be contingent upon having the physical spectrometers operational. To produce Fig. 17, you *must* have an operational version of the database to carry out retrievals on synthetic spectra. It should be straightforward to carry out the test I proposed previously, and as the spline method is an important part of your algorithm, it should be rigorously tested. The reasons given in the response are not sufficient justification for why this test is not possible. I know this may seem like a strange point to reiterate, but in my own work I have seen cases where splines can lead to very unphysical interpolations, so it is critical to verify that any spline interpolation gives valid results.

- Regarding my previous comments about why each measurement only uses data from SHS1 at T1 and SHS2 at T2, I finally understand that the measurement of SHS2 at T1 and SHS1 and T2 can be used for different measurements (that measure different locations in the atmosphere) thus increasing the number of OH soundings. That was not and still is not clear in the paper.

- Regarding by previous question about whether the radiance intensity at each wavelength are coordinates for the lookup table, based on the authors' response, my understanding of the lookup table procedure is:

  1. The date, latitude, longitude, SZA, and SAA are fixed to their values for the observation being retrieved. This gives a 1D slice out of the 6D table where the only remaining degree of freedom is the OH concentration.
  2. The OH concentration is varied and the corresponding radiances compared against those observed. If they match, then the retrieval reports that OH concentration as the result. If not, the spline interpolation is used to find a better match.

  In other words, the tomographic database is essentially replacing the radiative transfer model in an inversion framework. Is this correct? I do not find Fig. 16 particularly helpful in understanding this process.

- Minor suggestion: add an arrow to Fig. 3 indicating the direction of flight of the spacecraft.