Atmospheric Measurement Techniques Discussions

Atmos. Meas. Tech. Discuss., referee comment RC1
https://doi.org/10.5194/amt-2022-135-RC1, 2022

Comment on amt-2022-135

Anonymous Referee #2

Referee comment on "A new machine-learning-based analysis for improving satellite-retrieved atmospheric composition data: OMI SO_2 as an example" by Can Li et al., Atmos. Meas. Tech. Discuss., https://doi.org/10.5194/amt-2022-135-RC1, 2022

Li et al. presented an new machine learning scheme and applied it to OMI SO2. It is an interesting preliminary study and the work should be published after minor revision. The paper is well written and structured. Overall, the figures are not of sufficient resolution.

Main comment: although it is an interesting study, I am concerned by the fact that the NN function is heavily weighted by SRR. The SRR, as defined by the author, actually contains the desired result. Therefore, it is not a surprise that this works. The only real task of the NN is to resolve the RMS dependence. Also, the fact that the noise gets reduced is in a way artificial, as it is the direct result of the constrain on the clean pixels that should be SCD = zero. The reduction of the bias poses also the problem of a possible overcorrection. On long-term averages, are the weak emissions sources detected by OMI PCA still present in the NN data set?

In my opinion, what would be much stronger is to train the NN directly with the corresponding radiances. This is done to some extend (through a PCA transformation) but it is coming at the end of the paper. It is pity it is not put more in front.

Minor comments

L55: I agree with the necessity to improve the retrievals but is a 10% noise increase a real problem for addressing long-term trend monitoring? I do not think so. The appearance of instrument issue like row anomaly is more of a problem.

Figure 2b is confusing. Is the SRR unit less? The SCD is DU and RMS has no unit, thus SRR should be expressed in DU (?).

Section 2.2. The classification of pixels is quite complex. As explained in the text, the parameters used (a1,a2) have been determined by testing. However, it would be good to illustrate the impact of the (a1,a2) settings on the final results. Currently, it is hard to judge if the complexity is worth, compared to a more simple classification.

Section 2.3: the processing is not performed separately for each row. Why not? Would it improve/degrade the results?

Section 3.1: the SCDs over SAA are much better in the NN analysis, which is surprising. Any idea why?

Figure 7: the figure quality is not sufficient. When zooming over the subplots it is hard to see the emission patterns described in the text.

Figure 10b : the PCA-NN results seem to show striping features, although the analysis is performed separately for each row. Why?