EGUsphere, author comment AC1
https://doi.org/10.5194/egusphere-2022-806-AC1, 2022

Reply on RC2

Joanna Joiner et al.

Author comment on "Use of machine learning and principal component analysis to retrieve nitrogen dioxide (NO_2) with hyperspectral imagers and reduce noise in spectral fitting" by Joanna Joiner et al., EGUsphere, https://doi.org/10.5194/egusphere-2022-806-AC1, 2022

We thank the reviewer for their careful review that has helped to improve the manuscript and reply to each comment (repeated below) in bold face.

Joiner et al. presented a new method based on machine learning to retrieve NO2 from low spectral resolution instruments. This is an interesting study and the authors demonstrated well the potential of this approach for forthcoming missions with high spatial resolution. The paper is very well written and structured and there is little to say on the content. It should be published in AMT, after few (mostly minor) corrections.

Overall, the quality of the figures is quite poor, and it would be needed to improve the resolution of all figures.

We tried to improve the quality of the figures. We think the issue was in conversion from .eps to another that format would work within the LaTeX template. We have tried another method that does not degrade the quality of the figures so that the text and content are sharper. Note: the flow diagrams were saved directly in pdf format, so they have not been altered.

Comments

My main reservation is that the paper covers two things: (1) the retrieval of NO2 from instruments with lower spectral resolution, (2) the retrieval of NO2 using machine learning. What I would like to see is a DOAS-type retrieval of NO2 for the lower spectral resolution data so that one can evaluate the benefit of the machine learning approach directly. I realize that this is probably quite some work but it would be nice to understand if machine learning can improve the retrievals or not. It is understood that machine learning is interesting in terms of computational time but from the results shown here it is not clear if it is the only advantage.

We thank the reviewer for this question. We have modified the last paragraph of the introduction to make it clearer that we attempt to answer two questions related to the two points above. We also changed the title to reflect these two aspects.

We agree that it would be quite some work to do the DOAS-type retrieval, and we were after a timely answer to the question of whether or not we could retrieve NO_2 with the lower resolution OCI sensor before its launch, so we turned

to machine learning. We did find relevant references on DOAS type retrievals using the APEX airborne sensor whose resolution is ~2.4-3.4 nm across the NO_2 fitting window and have included them in the revised manuscript. The authors were only able to use a limited fitting window of 470–510 nm for DOAS NO_2 retrievals due to "interference with unidentified instrumental artefacts or features prevents us from extending the fitting window to wavelengths lower than 470 nm," so such retrievals appear to be difficult and beyond the scope of the present work.

Please see answers below for further clarification on the potential benefits of the neural network approach.

-Section 2:

*A small section introducing OCI and GLIMR is missing here. Details on instruments (spectral range, sampling, performance, etc) should be added (e.g., as a form of a table).

We have added such a table (Table 1 in the revised manuscript) with instrument characteristics as suggested and moved the SNR figures and some other text to this section.

*It would be good to refer to past studies of low spectral resolution NO2 retrievals attempts.

We found a few works with lower spectral resolution NO_2 retrievals: the APEX aircraft instrument (Tack et al., 2017; Kuhlmann et al., 2022) and studies with the Russian Resurs-P satellite (Postylyakov et al., 2017, 2019; Zakharova et al., 2021) and have referenced them as appropriate. We hope we haven't missed any other references.

*line 89 reads '..a retrieval would likely need to make use of the broad continuum absorption..'. Not sure what is meant here. Is the author meaning a large wavelength range to better constrain the fit or is it really the broadband contribution of the NO2 absorption which is targeted? Please clarify.

We changed this sentence to "At GLIMR spectral resolution, a retrieval would likely need to make use of the broad NO_2 absorption feature peaking at around 400 nm rather than the finer spectral features used in DOAS retrievals."

*p8, line 179 '…to half the number of spectral elements was sufficient to capture the spectral information associated with NO2 while providing some noise reduction'. I don't understand why this should provide a noise reduction. Could you elaborate?

We have added text to elaborate: "Since the trailing PCs typically express random spectral noise, eliminating these modes can lead to noise reduction."

*p9, 275: the use of the logarithm of the NO2 SCD does not correspond to anything physical. Could you clarify why this was used? Is there a justification for this, other than it gives good results.

We added: ", likely because the distribution of SCDs is more normally distributed in log space, which is desirable for neural network training."

*p10, line 237: about the NN ability to capture the wavelength dependence of the SNR. What about wavelengths cross-correlations? It is this information available from the OCI and GLIMR? I guess not but could it affect the performance of the NN approach?

We added "Here, we assume that errors are not correlated with wavelength as Information on correlated errors was not provided. Correlated errors could possibly degrade the performance of the retrieval if the neural network is not able to effectively account for them."

-Section 3

***** for figs 5,6,9, it is not always clear to what settings they correspond. E.g., what wavelength range was used, noise added or not, etc. I propose to detail this in all figure captions.

Details are now listed in the figure captions as suggested.

* Table 1: why is the bias larger for the case where the wavelength range is close to that of OMI NO2 (400-470 nm)? In general, the SCDs in different windows should not be identical (because of the AMF wavelength dependence). I find this aspect is not discussed enough.

While we are using different wavelength ranges for the NN training, the training target output is the OMNO2 slant columns so that the output slant columns always pertain to the OMNO2 fitting window. The differences in biases between the different runs are fairly small (insignificant) and are probably related to uncertainty in the training. We added "Note that in all cases, the training target is SCD from the OMNO2 algorithm that corresponds to the OMNO2 fitting window, and all statistics are computed with respect to the OMNO2 SCDs."

*P14, l 313: it is mentioned that the 'largest errors occur over highly polluted areas'. Could it be because the NN is not sufficiently trained for these conditions? Could the situation improve by adding such highly polluted scenes in the training set? If yes, what would be the weight of such scenes in the training set?

As described in the section "Simulated OCI and GLIMR data", we added extra days with high pollution to provide more samples under these conditions and this improved the results a bit as compared with not including the extra days. Any imperfections in the OMNO2 data could also contribute, and these may be amplified in polluted conditions. We tried many different approaches to try to improve the results, e.g., separate training for land and ocean. We were not able to improve the results. It is possible that a separate training could be done and applied for cases of high pollution if we are to completely trust the OMNO2 results. We revised this paragraph to replace the words "error" and "bias" with "differences" as OMNO2 may not represent the absolute truth. We added a sentence discussing that we tested other training schemes: "We tried alternate training scenarios such as training and applying NNs separately over land and water, but this failed to remove all of the differences."

*p14,l 328 : '..but rather the VCD'. Why? SCD is the actual signal, not the VCD.

We modified the sentence to "The desired retrieved quantity for atmospheric correction in ocean color algorithms is not the NO_2 SCD for a particular fitting window, but rather the VCD such that the appropriate absorption can then be accurately computed at any wavelength for atmospheric correction (Ahmad et al., 2007)."

* In general, how frequent should the NN be trained? Have you tried the algorithm for periods affected by the row anomaly of older data in the OMI mission? How is the data quality affected?

We did not try the algorithm for periods affected by the row anomaly. Regarding how often to train, the algorithm would need to be retrained whenever there are instrumental changes, for example, drifting wavelength-dependent calibration or increased noise or other artifacts. This would need to be assessed on a case-by-case basis. We added a paragraph in the section on practical implementation issues:

"Another consideration is how often a NN would need to be retrained. If the instrument were spectrally stable, retraining might not be necessary or might be infrequent. However, destriping may still be necessary to correct for transient spectral artifacts. Retraining should be done whenever there is a substantial change in the instrument spectral characteristics. Since the OMNO2 algorithm uses monthly-averaged solar irradiances, it may be more optimal to similarly normalize with respect to the same set of solar irradiances before training than to use only the radiances as we have done here as the solar data may help to account for instrumental changes."

*p19, line 392: it would be good to add these results in SI.

We added a table showing the statistics in the main part of the paper.

Typos

-P1, line4: S5P resolution is 3.5 km x 5.5 km (not 3.5 km x 5 km)

Fixed, thank you.

-p4, line94: 'gloyoxal'-> 'glyoxal'

Fixed, thank you.

-p4, line94: 'spectral imprint' -> unclear what is meant by 'imprint'

We changed "imprint" to "signature". We hope that is clearer.