Atmos. Meas. Tech. Discuss.,
doi:10.5194/amt-2019-143-RC2, 2019

Atmospheric Measurement Techniques Discussions

Interactive comment on "A neural network radiative transfer model approach applied to TROPOMI's aerosol height algorithm" *by* S. Nanda et al.

Anonymous Referee #2

Received and published: 3 July 2019

General Comments

Nanda et al. present a method to accelerate radiative transfer calculations based on neural networks (NN), to speed up an optimal-estimation based retrieval of aerosol layer height (ALH) from TROPOMI O2-A band observations. The neural network is trained/validated on a set of simulated TROPOMI spectra. The ALH retrieval using the NN-based forward model is compared to results using an explicit line-by-line forward model. The NN version of the retrieval produces results consistent with the explicit model for a synthetic and real test case, whilst improving the speed by three orders of magnitude.

Overall, I think this paper is interesting and well within the scope of AMT. In Its present form, it is missing some key details (see comments). Once these are addressed I will recommend publishing.

Specific Comments

Page 2 Line 18: "The bottleneck identified. . ."

There are a few other commonly used methods for accelerating RT simulations e.g. optical property PCA and low-streams interpolation (see cited review below). It may be worth mentioning why NN is being chosen over these.

Natraj V. (2013) A review of fast radiative transfer techniques. In: Kokhanovsky A. (eds) Light Scattering Reviews 8. Springer Praxis Books. Springer, Berlin, Heidelberg

Page 4 Line 30: From preliminary tests, the exclusion of RRS. . ..

Since there were preliminary tests for the sensitivity to RRS, it may be worth mentioning what these were and quantitatively how these impacted z_{aer}. Solar-induced chlorophyll fluorescence in the A-Band may also have a similar effect spectrally, and likely has a greater impact on the spectra. See

Frankenberg et al. (2011) Disentangling chlorophyll fluorescence from atmospheric scattering effects in O2 A-Band spectra of reflected sunlight, Geophysical Research Letters, vol 38, L03801 https://doi.org/10.1029/2010GL045896

Page 4 Line 31: The aerosol fraction is assumed as 1.0

Could you define what you mean by aerosol fraction?

Page 5, Line 1-5: Perhaps the largest simplification...

There are many assumptions here - The aerosol optical properties are fixed (0.95 SSA, g=0.9 using a simplified HG phase function). Either literature justifying why these assumptions are ok must be cited, or the authors need to test these how these assumptions impact the retrieval results e.g. by testing against synthetic data with realistic optical properties. I would be curious to see how the retrievals perform for cases with very different optical properties e.g. dust

Page 7, Line Line 19: The inputs for the NN are referred together as the feature vector...

The fixture of the aerosol optical properties in the optimal estimation approach seems quite restrictive. Have you performed an information content analysis of the TROPOMI O2-A band to check if retrieving some of these can reduce the uncertainty in aerosol height? E.g. allowing the SSA to vary may reduce the potential influence on its parameter error inducing a corresponding ALH error.

Page 7, Line 28: "whereas NN only uses the temperature at z_{aer}"

For the meteorological parameters, is there any rationale for excluding other potentially important predictors from being included e.g. PBL height or surface heating fluxes and wind speeds that may also provide prior information about the ALH.

Page 8 Section 3.2 First Paragraph

From my reading of this, the profiles are generated randomly after selecting a random set of tropomi solar-viewing geometry combinations. Naively, I would expect that the most representative way to create the training set would be to select the profiles from the ERA reanalysis corresponding to the randomly selected orbit geometries - the model probably doesn't need to reproduce a spectrum of Saharan dust for a typical Antarctic viewing geometry. Perhaps there is a heuristic argument that the way you are doing it could more reliably span the entire set of profile/viewing combinations, but if this is so you should state it in the manuscript. Ideally you would want to test the performance by comparing multiple approaches of generating the training data, though I am not sure of your computational resources

Page 9, Line 1: Finding the most optimal neural network configuration requires...

Using a single set for validation leads to questions about robustness in the validation.

Typically cross validation e.g. k-fold methods are used to derive more accurate estimates of model performance. Have you looked into how robust the single training set is?

Page 9, Line 6: The sigmoid function is chosen for activation...

Can you show evidence for the sigmoid function outperforming the other forms?

Page 9, Line 8: For each of the neural network models...

It would be useful showing a comparison of performance for the different models tested - I don't have an intuitive sense from the description of how different the actual results were for the different layer/node combinations, or how robust one instance of a 25000 iteration training is. For instance if I retrained the two layer model with 100+100 nodes again with a different initialization, would it still be the optimum configuration?

Page 11, Line 12: Although the retrieval algorithms have good agreements...

For the low aerosol loading scenes, what happens when you include both surface albedo values in the NN model?

Minor Corrections

Equation 2: Bold the x in the forward model to be consistent with notation for vectors.

Page 7, Line 17: Change "automatic differentiation which is a powerful algorithm that computes" to "automatic differentiation which computes"

Interactive comment on Atmos. Meas. Tech. Discuss., doi:10.5194/amt-2019-143, 2019.