Atmos. Meas. Tech. Discuss.,
doi:10.5194/amt-2019-228-RC1, 2019

Atmospheric Measurement Techniques
Discussions

Interactive comment on "Neural Network for Aerosol Retrieval from Hyperspectral Imagery" *by* Steffen Mauceri et al.

Anonymous Referee #1

Received and published: 5 July 2019

This paper describes the design of a neural network algorithm uses imaging spectrometer measurements to separately retrieve aerosol optical depth (AOD) for three aerosol types (dust, sulphate, brown carbon). The neural network is trained using synthetic spectra and applied to two images from the airborne instrument AVIRIS-NG over India. Overall, the methodology is clearly explained and the paper fits the scope of AMT, but in my opinion the study has a number of shortcomings which should be properly addressed before the paper can be published. The main problem is that, while the NN results are satisfactory on synthetic data, the retrievals on real measurements display large surface features. Other shortcomings are the lack of validation for the per type AOD retrievals and the assumption of spherical dust particles made in the creation of the training set, which may be inaccurate. Below are my detailed comments.

MAIN COMMENTS

- The surface features in the AOD retrieval are really striking, and it would be important to investigate whether the problem may be mitigated by changing the design of the NN. For example, I have the impression that, if you attempt a simple classification of the original image (e.g., into vegetation, bare soil, urban, water classes) and correlate the AOD retrieved on each pixel to the class the pixel belongs to, you may see a strong correlation. If this is the case, then it may mean that you need to pass the result of this simple classification as an additional input to the AOD-retrieval NN, or to train multiple NNs, one per class. This may reduce this effect in your retrievals, which in my opinion is too large at the moment.

- The use of a spherical model for dust aerosols in the generation of the training set may also lead to inaccurate results when the NN is applied to real data (Kalashnikova and Sokolik, 2004, Dubovik et al., 2006, Lee et al., 2017). For this reason, I would recommend to retrain the NN by using a nonspherical model. References:

Kalashnikova, O., and Sokolik, I. N., "Modeling the radiative properties of nonspherical soil-derived mineral aerosols", J. Quant. Spectrosc. Rad. Transfer, 87, 137-166, doi: 10.1016/j.jqsrt.2003.12.026

Dubovik, O. et al. (2006), "Application of spheroid models to account for aerosol particle nonsphericity in remote sensing of desert dust", J. Geophys. Res., 111, D11208, doi: 10.1029/2005JD006619

Lee, J. et al. (2017), "AERONET-based nonspherical dust optical models and effects on the VIIRS Deep Blue/SOAR over water aerosol product", J. Geophys. Res., 122, 10384-10401, doi: 10.1002/2017JD027258

- An additional problem I see is that you try to estimate the AOT for each aerosol type, but you do not provide any indication of the credibility of the per-type retrievals performed with real measurements. While direct observations of "typed" AODs are probably not available, it would be important to at least have a look of what are the outputs of some reanalysis models (e.g. MERRA-2, CAMS) around the locations you considered and in the same dates. I imagine that these models may have a much coarser spatial grid than your images, but I guess they would be your only possible source of verification.

MINOR COMMENTS

- P1, L21. "... absorption ... are ..." -> "... is ...".

- P1, L24-25. "Instead, a common practice ...". Maybe you mean that one takes the darkest pixel in the image, assumes that the observed radiance over that pixel only comes from the atmosphere and subtracts that radiance from all the other pixels in the image. If so, make that explicit in the paper. If not, clarify what you mean instead.

- P1, L28, "great" -> "larger".

- P2, L24. Please also mention the advances made possible by multiangle polarimetry, which provides an enhanced capability of separating the aerosol signal from the surface signal, and a better sensitivity to the aerosol microphysical parameters (Kokhanovsky et al., 2015, Dubovik et al., 2019). References:

Kokhanovsky, A.A. et al. (2015), "Space-based remote sensing of atmospheric aerosols: The multi-angle spectro-polarimetric frontier", Earth Sci. Rev., 145, 85-116, doi: 10.1016/j.earscirev.2015.01.012

Dubovik, O. et al. (2019), "Polarimetric remote sensing of atmospheric aerosols: Instruments, methodologies, results, and perspectives", J. Quant. Spectrosc. Rad. Transfer, 224, 474-511, doi: 10.1016/j.jqsrt.2018.11.024

- P2, L28. Do you also foresee applying the proposed method to existing satellite imagers such as EO-1 Hyperion, or the recently launched PRISMA?

- P6, Fig.2. I would suggest to change "t_aer" to "tau_aer" in the legend.

- P6. The problem with this study of the sensitivity of the TOA radiance to the aerosol type is that the microphysical properties of the aerosol types are prescribed. Thus, your capability of distinguishing them in your simulation may be greatly overestimated compared to what happens in nature, where I don't think you will see dust, brown carbon etc. always with the same size distribution. Even the refractive index of certain aerosol types (dust in particular) is highly variable, so it would be better to incorporate this variability in the training set (as you already tried to do with the surface properties) in order to have a better hope of making your NN scheme more robust. Note that the aerosol size, in particular, mainly influences the spectral slope of your radiance. Thus, it may well be that your retrieval just tries to distinguish between three "size classes" of aerosols, which you map to "aerosol types" through a rather arbitrary 1:1 correspondence. Also for this reason, it is really important to compare your retrieved aerosol speciation on real data to the outputs of some reanalyses. This would be the only way to obtain at least a preliminary indication that your AOD retrieval distinguished into types actually works in reality.

- P7, L26. I guess you mean "biophysical properties of vegetation".

- P7, L26. Given that your application concerns aerosols, you should also mention previous work on NNs for aerosol retrievals (Radosavljevic et al., 2010, Chimot et al., 2017, Di Noia et al., 2017). References:

Radosavljevic, V. et al. (2010), "A data-mining technique for aerosol retrieval across multiple accuracy measures", IEEE Geosci. Remote Sens. Lett., 7, 411-415, doi: 10.1109/LGRS.2009.2037720

Chimot, J. et al. (2017), "An exploratory study on the aerosol height retrieval from OMI measurements of the 477 nm O2-O2 spectral band using a neural network approach", Atmos. Meas. Tech. 10, 783-809, doi: 10.5194/amt-10-783-2017

Di Noia, A. et al. (2017), "Combined neural network/Phillips-Tikhonov approach to aerosol retrievals over land from the NASA Research Scanning Polarimeter", Atmos.

Meas. Tech., 10, 4235-4252, doi: 10.5194/amt-10-4235-2017

- P8, L8-9, “training-set” -> “training set”, “validation-set” -> “validation set”.

- P8, L16. Why are you using radiance and not reflectance as an input? Why are you using ground distance and ground elevation (which should have a relatively minor effect on the top-of-atmosphere radiance or reflectance), but are not using viewing zenith angle and viewing azimuth angle, which may have a greater effect?

- P9, Eqs. 8 and 9. I have the impression that the "n" in Eq. 9 is not the same "n" as in Eq. 8. Please adopt an unambiguous notation and explain the meaning of any symbol you use.

- P9, L6. Add that theta is a vector containing all the weights of the NN (right?). Furthermore, in the next sentence I don’t think theta should be the subscript “i” in the L2 norm.

- P9, L7, add “or weight decay” after “L2 regularization”.

- P9, L8. “The L2 regularization is weighted” -> “The L2 regularization term R(theta) is weighted”

- P11. Consider splitting Figure 4 into three plots (one per aerosol type). The plot for carbonaceous aerosol looks completely hidden.

- P13, L11-12. In addition to just reducing the number of sampling points, it would be more interesting to also change the spectral resolution of the instrument (I mean, the width of a slit function you may convolve your synthetic spectra with). This would make your setup more similar to that of existing satellite imagers, which typically have a spectral resolution of $\sim$10 nm.

- P14, L7. Since you are using synthetic data with a spherical dust model, it would be important to repeat your experiment with a more realistic model for dust. Otherwise the numbers you provide for the retrieval accuracy are not really meaningful, as they

cannot be really taken as an indication of what would happen in a real scenario.

- L15. You say, “It is inherently difficult to interpret the inner workings of neural networks”. Actually, the derivative of the NN output with respect to its input can be computed analytically (Blackwell, 2012, Di Noia et al., 2013). This may enable more systematic sensitivity analyses, as it means that the NN retrieval can be rigorously linearized around its actual input (spectrum + viewing geometry). It may be also useful to feed the values retrieved by the NN back to a radiative transfer model. Combined with the NN input Jacobian mentioned above, this may enable estimating the sensitivity of the NN retrieval to the true state vector (Jiménez and Eriksson, 2001). References:

Blackwell, W. (2012), “Neural network Jacobian analysis for high-resolution profiling of the atmosphere”, EURASIP J. Adv. Sig. Proc., 2012, 1-11, doi: 10.1186/1687-6180-2012-71

Di Noia, A. et al. (2013), “Global tropospheric ozone column retrievals from OMI data by means of neural networks”, Atmos. Meas. Tech., 6, 895-915, doi: 10.5194/amt-6-895-2013

Jiménez, C. and Eriksson, P. (2001), “A neural network technique for inversion of atmospheric observations from microwave limb sounders”, Radio Sci., 36, 941-953, doi: 10.1029/2000RS002561

- P17, L9. “To apply the model to real imagery, one would ideally train the model further on real observations”. I don’t think this is necessarily true. If your forward simulations and your knowledge of the instrument are realistic enough, using synthetic data should be feasible (again, you can look at Chimot et al. (2017) or Di Noia et al. (2017) for examples). Furthermore, training on real data is guaranteed to introduce sampling biases and co-location errors that may counterbalance the advantage of implicitly incorporating the real instrument characteristics in the training set. Furthermore, for the particular task of retrieving AOD separated into types it may be even impossible to find a training dataset with real observations.

- P18, Section 5.1. There is one aspect that is not totally clear to me. You perform the PCA of AVIRIS-NG observations and retain 16 principal components. Do I correctly understand that what you then pass to the NN are not directly the principal components but the radiance spectra reconstructed from the 16 principal components? If so, please add a sentence somewhere in the section to make this clear.

Apart from this, I have a more fundamental question. You use the PCA as a tool to denoise AVIRIS-NG imagery, which is fine to do, and derive the PCA coefficients from the AVIRIS imagery itself. However, you trained your NN with synthetic spectra, and in order to apply your NN to real observations it is important to make sure that your real data look as similar as possible to the data you used to train the NN. How confident are you that your PCA-based denoising does not change the statistical distribution of the reconstructed radiances compared to that assumed in the training set? It would be interesting to check what happens to the synthetic spectra you used to train the NN if you compress them and reconstruct them with the PCA transformation you derived from the AVIRIS imagery. If they change significantly, then this may be a warning flag that there may be problems when you apply your NN to real measurements pre-processed with your PCA transformation.

- P18, L17. Are you sure components 17 and 18 in Fig. 9 do not contain useful information? They seem to display some "structured" spatial patterns.

- P25, L17. The correlation value looks misleading, as it looks mainly driven by the two high-AOT data points in the upper right.

Interactive comment on Atmos. Meas. Tech. Discuss., doi:10.5194/amt-2019-228, 2019.