Dear Andrew,

Thank you for your comments and the comments of the anonymous reviewer. The paper improved a lot since its initial submission. We addressed all comments in red below. One co-author was not sure whether it is ok to write "deg" instead of " ° ". If that is an issue, I can update that across the manuscript and figures.
Let me know if you have any remaining questions.

Thanks
Steffen and Co-authors

Comments from Anonymous Reviewer:
- P3, L11. Also mention in the introduction that you compared your typed AOT retrievals to CAMS.
P3, L11. Added CAMS to description of Section 5: "In Section 5 we apply the inverse model to AVIRIS-NG observations and compare results to AERONET and MODIS retrieved AOT **and the CAMS analysis product**."

- P8, L8. After "multilayer perceptrons", cite: Werbos, P. J.: Beyond Regression: New Tools for Prediction and Analysis in the Behavioral Sciences, PhD thesis, Harvard University, Cambridge, MA, United States, 1974
P8, L8. Citation added

- P8, L17. I would replace "operations" with "functional elements".
P8, L17. Changed to: "A multilayer perceptron is comprised of many individual **functional elements**, or *neurons*, that …"

- P10, L10. Aren't you actually multiplying by the square ratio between the actual Sun-Earth distance d and the mean Sun-Earth distance d0? Otherwise your modified radiance would be expressed in a strange unit (a radiance times the square of a length). Note that, even if d0=1 AU and thus numerically d/d0 equals d in AU, dimensionally what you are taking is still the ratio between the two quantities. [Editor's note: I think the reviewer is correct here, please check.]
P10, L10. Thanks for catching that! Changed to: "We then scale the radiance of a particular observation, L_j, by dividing through the cosine of the SZA_j and multiplying with the square of the **ratio between** Sun-Earth distance, d_j, **and mean Sun-Earth distance**, d_0, (Equation 10)."

Eq 10 updated: "$R_j = \dfrac{L_j * \left(\dfrac{d_j}{d_0}\right)^2}{\cos(SZA_j)}$ "

- P11, L5.Maybe you can emphasize that, if you look at Fig. 2, the spectral signature of sulphate and brown carbon look very similar (except for a scaling factor - which means that it may be possible to confuse sulfate with a certain AOT with brown carbon with a higher AOT) whereas dust has a markedly different spectrum thanks to its absorption at shorter wavelengths. Therefore, one may expect some degree of confusion between sulfate and carbon and a better capability of identifying dust.

P11, L6. Interpretation added: "Thus, the model has the smallest uncertainties for the retrieval of dust and sulfate, which are less strongly absorbing across most of the relevant spectral range, compared to carbon. Distinguishing the weaker signal from carbon from the surface contribution might be more challenging for the neural network and lead to the overall larger uncertainties for carbon. Additionally, the spectral signature of carbon and sulfate show high similarities at shorter wavelengths, except for a scaling factor (see Figure 2). This might further challenge the neural network to distinguishing the two. For wavelengths ranging from 500 nm to 600 nm, absorption of dust increases with decreasing wavelength, while the opposite is true for carbon and sulfate. This unique spectral signature of dust, compared to the other two aerosol types, might be another reason for the smaller standard error of its retrieval."

- Section 5.5. The sulfate AOT plot looks quite encouraging. It would be also interesting to see in how many of the 21 considered scenes CAMS and your NN agree on which is the aerosol type with the largest AOT.

P28, L7. Comparison added: "We further evaluate the agreement between AVIRIS-NG and CAMS for which aerosol type has the largest, second largest and smallest AOT in a given scene. For 16 out of the 21 compared scenes AVIRIS-NG and CAMS agree on which aerosol type dominates with the largest AOT (see Figure 18). 14 out of 21 compared scenes agree on which aerosol type has the second largest AOT and 16 out of 21 agree on which aerosol type has the smallest AOT. Thus, if one were to use the proposed neural network to classify the dominant aerosol type from AVIRIS-NG observations, the neural network would have an estimated accuracy of 76%.
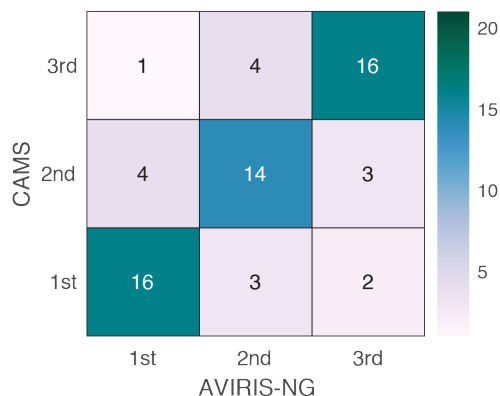

Figure 18: Number of scenes where CAMS and AVIRIS-NG agree and disagree for which aerosol type has the largest (1st), 2nd largest and 3rd largest AOT. The total number of analyzed scenes is 21. Thus, the 16 in the bottom left corresponds to 16 scenes in which AVIRIS-NG and CAMS identified the same aerosol type as having the largest AOT, out of 21 scenes."

Comments from Andrew Sayer:
Page 2 line 8: the reference should appear as "Pope et al.", not "Pope III et al". The surname is "Pope" and the author is the third of that name in the family. I am not sure if you are using BibTeX for references but if so a guide for typesetting names containing constructs like that is here: https://tex.stackexchange.com/questions/557/how-should-i-type-author-names-in-a-bib-file I think the correct syntax for the author name in the bibliography file is "Pope, III, C. A."

P2, L8: Citation changed to "Pope et al."

Page 2 line 18: it would be better to state that MODIS/MISR etc have "near-global" rather than "global" coverage, as there remain several systematic sampling gaps.
P2, L18: Changed to "near-global"

Equation 4: from this, it would be good to note somewhere in the manuscript that the surface is assumed to be Lambertian. (A familiar reader can infer this from the Equation, but it would be good to state it directly in the text too.)
P3, L31: Added information that surface is assumed to be Lambertian: "The three simulations are then used to calculate at sensor radiance for any given surface albedo, **which is assumed to be Lambertian,** utilizing …"

Equation 10: this quantity is commonly referred to as "top of atmosphere reflectance" in the aerosol remote sensing community (typically after also scaling by pi). I realise you may not wish to use the term reflectance due to the potential confusion with surface reflectance. I leave it up to you whether you wish to note this identity in the manuscript.
P10, L11: Added that the two are similar to each other: "This scales the magnitude of L_j while preserving its spectral shape **and is similar to deriving the *top of atmosphere reflectance***."

Page 11 line 5: the comment about SSA is wavelength-dependent, as dust is fairly strongly absorbing in the blue and UV. So I would add a qualifier about which wavelength range you are talking about here. One might instead right that dust and sulfate "are less strongly absorbing across most of the relevant spectral range" than carbon. Also, I think you mean "uncertainty" rather than "accuracy" here (since we're discussing total error, not bias).
P11, L6: Replaced "accuracy" with "uncertainty" and expanded possible explanation for differences in the standard errors for the three aerosol types: "Thus, the model has the smallest uncertainties for the retrieval of dust and sulfate, which are less strongly absorbing across most of the relevant spectral range, compared to carbon. Distinguishing the weaker signal from carbon from the surface contribution might be more challenging for the neural network and lead to the overall larger uncertainties for carbon. Additionally, the spectral signature of carbon and sulfate show high similarities at shorter wavelengths, except for a scaling factor (see Figure 2). This might further challenge the neural network to distinguishing the two. For wavelengths ranging from 500 nm to 600 nm, absorption of dust increases with decreasing wavelength, while the opposite is true for carbon and sulfate. This unique spectral signature of dust, compared to the other two aerosol types, might be another reason for the smaller standard error of its retrieval."

Figure 13: what is the background image here? I didn't see that stated; it should also have a colour bar if the colours are to be interpreted quantitatively.
P24, L2: Added information to Figure description. We did not add a color bar of this topographic map since it would distract from the information we try to show (spatial distribution of AERONET stations with respect to the AVIRIS-NG flights). "Figure 13: Location of AERONET stations and AVIRIS-NG flights on a topographic map of India."

Neural Network for Aerosol Retrieval from Hyperspectral Imagery

Steffen Mauceri[1,2], Bruce Kindel[2], Steven Massie[2], Peter Pilewskie[1,2]

[1] Department of Atmospheric and Oceanic Sciences, CU Boulder, Boulder, Colorado, USA
[2] Laboratory for Atmospheric and Space Physics, Boulder, Colorado, USA

5

Correspondence to: Steffen Mauceri (Steffen.Mauceri@lasp.colorado.edu)

Abstract. We retrieve aerosol optical thickness (AOT) independently for brown carbon-, dust and sulfate from hyperspectral image data. The model, a neural network, is trained on atmospheric radiative transfer calculations from MODTRAN 6.0 with

10 varying aerosol- concentration and type, surface albedo, water vapor and viewing geometries. From a set of test radiative transfer calculations, we are able to retrieve AOT with a standard error of better than ± 0.05. No a priori information of the surface albedo or atmospheric state is necessary for our model. We apply the model to AVIRIS-NG imagery from a recent campaign over India and demonstrate its performance under high and low aerosol loadings and different aerosol types.

1 Introduction

15 Remotely sensed surface spectral reflectance is used in many scientific disciplines including, geology, forestry, water studies and urban studies (Davis et al., 2002; Rencz and Ryerson, 1999). The surface reflectance can be either directly measured at the ground with portable field spectrometers or indirectly measured from air- and space-borne platforms. Observations at air- and space-borne instrument altitudes are sensitive not only to the signal from the surface but also the intervening atmosphere between surface and sensor. Thus, to derive surface reflectance from air- and space-borne observations, the data must be

20 corrected for atmospheric absorption and scattering effects. The main objective of atmospheric correction is the accurate removal of absorption and scattering by aerosols and gases. While absorption by water vapor and other gases is highly wavelength dependent, with relatively strong, discrete, absorption bands, aerosols extinction (the sum of absorption and scattering) is a smooth, continuous function of wavelength. This makes it challenging to separate it from the surface contribution. Most current atmospheric corrections ignore the aerosol variability within a scene. Instead, aerosol properties are

25 approximated from visibility (e.g., Gao, Heidebrecht and Goetz, 1993; Adler-Golden *et al.*, 1999) or derived from climatology. Such approximations can lead to large errors in retrieved surface reflectance, particularly for aerosol optical thickness (AOT) larger than 0.4, commonly found, for example, over east Asia (Bilal et al., 2014; Van Donkelaar et al., 2010). While instrument performance has steadily improved over the years, resulting in higher signal to noise ratios, improvements in the treatment of aerosols in atmospheric correction routines has not kept pace. To improve the retrieval of surface reflectance products from

30 air- or space-borne observations, the spatial variability of AOT and wavelength dependent single-scattering albedo and phase function or its moments within a scene have to be known.

1

Aerosols also pose a major uncertainty in climate predictions through their direct scattering and absorption of solar and thermal radiation and indirect effects on cloud albedo (Twomey, 1977) and clouds lifetime (Albrecht, 1989; Pincus and Baker, 1994). Their contribution to radiative forcing is now the biggest uncertainty to the total anthropogenic forcing between 1750 and 2011 (Myhre et al., 2013). The magnitude of the direct interaction of aerosol with radiation depends not only on their abundance, but also their single scattering properties and the spectral reflectance of the underlying surface (Haywood and Boucher, 2000; Nan and A., 2015). Better quantification of the global distribution and optical properties of aerosols is a top priority to further improve climate projections.

Finally, aerosols are an important health risk factor (Pope et al., 2009). For eastern Asia, the World Health Organization Air Quality PM$_{2.5}$ (amount of aerosols with a diameter less than 2.5 µm) Interim Target-1 (World Health Organization, 2006) is exceeded for 50% of the population (Van Donkelaar et al., 2010), leading to an increase in mortality of approximately 15%. On a global scale, an estimated 7 million deaths were attributed to air pollution in 2016 (World Health Organisation, 2018). A better understanding of aerosol sources and their mixing in urban areas can inform decision makers and perhaps mitigate these hazards.

Currently, aerosols are routinely retrieved from ground and space-borne platforms. Ground based aerosol retrievals from *AErosol RObotic NETwork* (AERONET) (Holben et al., 1998) have the lowest uncertainty in retrieved AOT of less than 0.02 (Eck et al., 1999) but are spatially restricted. Space-borne instruments like the *Moderate resolution Imaging Spectroradiometer* (MODIS) (Salomonson et al., 1989) and the *Multiangle Imaging SpectroRadiometer* (MISR) (Diner et al., 1998) provide near-global coverage but retrievals from their measurements require separating the aerosol signal from the surface contribution. This results in large differences between the derived aerosol products from different instruments (Chu et al., 2003; Levy et al., 2005, 2013; Prasad and Singh, 2007; Remer et al., 2005). Other approaches aim to use the vast information content from space-borne multiangle polarimetric observations that provide enhanced capability of separating aerosol signal from surface signal, and a better sensitivity to aerosol microphysical parameters. However, retrieving aerosol properties from such observations is highly complex and operational products have not yet reached the accuracy implied by theoretical calculations (Dubovik et al., 2019; Kokhanovsky et al., 2015). Hence, accurate aerosol retrieval from space-borne platforms is still an active research topic.

To increase accuracy of global aerosol retrievals, we propose a retrieval algorithm that will be applicable to current and future hyperspectral space-borne instruments, such as Hyperspectral Precursor and Application Mission (PRISMA) (Labate et al., 2009), EO-1 Hyperion (Folkman et al., 2001), *Climate Absolute Radiance and Refractive Observatory* (CLARREO) (Wielicki et al., 2013) and the *Hyperspectral Infrared Imager* (HyspIRI) (Lee et al., 2015). Exploiting the large data volumes and hundreds of spectral bands of these instruments requires new fast retrieval algorithms. To meet these needs, we propose using neural networks. In this study, we present a neural network that is used to independently retrieve dust, carbonaceous- and sulfate aerosols from hyperspectral imagery over land, with no a priori knowledge of the surface type or atmospheric state.

2

The neural network can retrieve multiple collocated aerosol types and their contribution to the total AOT within a given scene. After fitting the neural network parameters, also referred to as *training*, the model can be used to retrieve AOT in real time without further radiative transfer calculations. We apply the neural network to *Airborne Visible / Infrared Imaging Spectrometer Next Generation* (AVIRIS-NG) (Hamlin et al., 2011) imagery from a recent campaign over India and demonstrate its performance under high and low aerosol loadings and different aerosol types. AVIRIS-NG, a follow-on to the *Airborne Visible / Infrared Imaging Spectrometer* (AVIRIS) (Green et al., 1998), has a spectral range of 380 – 2510 nm, a spectral resolution of 5 nm and spatial resolution of 4 m to 20 m depending on flight altitude.

The structure of our paper is as follows: Section 2 describes the forward radiative transfer calculations used to train an inverse model, the neural network. Sections 3 and 4 detail the architecture, training procedure and performance of the inverse model. Furthermore, we explore how instrument noise and sampling resolution influence model performance. In Section 5 we apply the inverse model to AVIRIS-NG observations and compare results to AERONET and MODIS retrieved AOT and the CAMS analysis product. In section 6 we provide our conclusion.

2 Forward Model

To train a neural network for aerosol retrieval we need a dataset consisting of the model input – output pairs, or *samples*, of the inverse model. These samples need to span a wide variety of atmospheric states, viewing geometries and surface albedos. To generate such a dataset, we employ a forward model described in the following section.

2.1 Radiative Transfer Calculations

The forward model radiative transfer calculations were performed with the *MODerate spectral resolution atmospheric TRANSmittance algorithm and computer model* (MODTRAN) 6.0 (Berk et al., 2014) from 400 nm to 2500 nm. Multiple scattering was implemented with MODTRANs' DISORT algorithm (Stamnes et al., 1988), utilizing a conservative number of 32 streams. We chose the 'Tropical' atmospheric profile. The solar zenith angle (SZA) was varied between 25° and 50° and water vapor varied between 0.4 g cm^{-2} and 4.1 g cm^{-2}. The distance between ground and sensor (ground distance) was varied between 3 km and 6 km and the ground elevation between 0 m and 2000 m. The AOT at 550 nm varied between 0 and 1.0. Three types of aerosols, brown carbon, dust and sulfate (see Section 2.3) were modeled as an external mixture, with a fraction between 0 and 100%. For every parameter permutation we perform three radiative transfer calculations for a constant surface albedo of 0, 0.5 and 1. In the following, the calculated at sensor radiances for a surface albedo of 0, 0.5 and 1 are denoted as L_0, $L_{0.5}$ and L_1. The three simulations are then used to calculate at sensor radiance for any given surface albedo, which is

assumed to be Lambertian, utilizing the MODTRAN interrogation technique (Verhoef and Bach, 2003). We first extract three atmospheric parameters, namely the spherical albedo, ρ, two-way transmittance, τ, and path radiance, L_P:

$$\rho = \frac{2}{\frac{1}{f-1}+2} \quad with \quad f = \frac{L_1 - L_0}{2*(L_{0.5} - L_0)} \tag{1}$$

$$\tau = (L_1 - L_0)*(1-\rho) \tag{2}$$

$$L_P = L_0 \tag{3}$$

Afterwards, we calculate the at sensor radiance, L, for the generated surface spectra, r (see next Section):

$$L = L_P + \frac{\tau * r}{1 - r * \rho} \tag{4}$$

Finally, the radiance is convolved with a gaussian kernel with a full width half maximum (FWHM) of 5.6 nm in the UV and 5.8 nm in IR, similar to the AVIIRIS-NG spectral resolution.

2.2 Surface Spectra

To simulate a wide variety of surface types we need a multitude of surface spectra. However, the number of freely available surface spectra is limited. The risk of using too few surface spectra is that the model might not be able to extract general surface characteristics. Applied to scenes with a previously unseen surface spectra, the model would perform poorly. Furthermore, most catalogs provide pure surface spectra from pure surface materials, also referred to as endmember spectra. This case is not representative for most air- or space-borne observations over land where multiple surface types are present in a single instrument pixel. Therefore, we generate a catalog of mixed surface spectra by randomly combining a limited number of measured spectra from different sources. The combination is performed by taking the randomly weighted mean of two randomly chosen endmember spectra at a time, until we have a total of 100,000 mixed surface spectra.

Endmember spectra were obtained from https://ecosis.org/ and https://speclib.jpl.nasa.gov. The datasets include 844 vegetation reflectance spectra from Hawaii (Dennsion and Gardner, 2000), 173 vegetation spectra from Hawaii volcanoes national park

4

(Grimm, 2017), 1065 urban surfaces from Santa Barbara (Herold et al., 2004b) and 270 rock and soil spectra (Meerdink *et al.*, in prep.; Baldridge *et al.*, 2009). To remove high-frequency noise in the surface spectra due to low signal at some wavelengths (Herold et al., 2004a) we smooth the surface spectra with a Gaussian kernel as done by Thompson *et al.*, (2018).

An example of soil, sand and vegetation reflectances from the catalogs are shown on the left in Figure 1. The right side shows nine examples of how the three spectra are combined to generate mixed surface spectra.
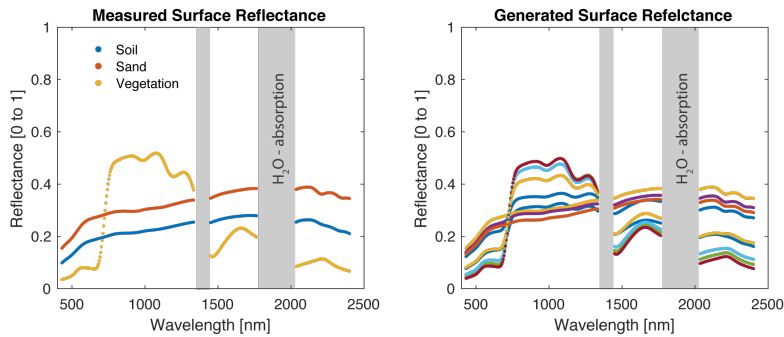
Figure 1: (Left): Surface reflectance for three different surface types (soil, sand and vegetation) from measured and smoothed surface spectra. (Right): Nine surface spectra, randomly generated from the three spectra on the left. Wavelengths with strong water vapor absorption are marked in grey.

2.3 Aerosol Parameterization

The optical properties of the three aerosols types that served as inputs to MODTRAN were calculated using three size distributions based upon Dubovik *et al.*, (2002) and the indices of refraction contained in HITRAN 2016 (Gordon et al., 2017). While aerosols cannot be strictly separated into types, we use these properties as representatives for dust, carbonaceous and sulfate aerosols. HITRAN 2016 includes H_2SO_4 indices at 300 K for sulfate aerosols and sand indices for dust aerosols from the AFCRL 1985 compilation (Fenn et al., 1985). The indices for sulfate and dust were selected because they cover the full wavelength range (0.2 to 40 µm) which is convenient for the use with MODTRAN. For brown carbon aerosols, from now on simply referred to as carbon, the indices are reported up to 1.2 µm by Alexander, Crozier and Anderson (2008). For longer wavelengths we extrapolated the real and imaginary parts.

Given the size distributions and the refractive indices, an extended version of the HITRAN-RI program (Massie and Hervig, 2013) was applied to calculate extinction, absorption, scattering spectra, and the Legendre moments of the phase function used in MODTRAN 6.0 calculations for carbon and sulfate. The calculations are based in Mie theory and thus assume homogenous

spherical aerosol particles. For dust we had to account for its non-spherical shape. We applied the T-matrix code of Mishchenko (Mishchenko and Travis, 1998), for randomly oriented particles, to generate the MODTRAN SAP files. The range of ratios of semi-major to semi-minor axes, or aspect ratios (AR), was varied between 1.01 and 1.8. This range contains the representative AR of 1.4 (Okada et al., 2001), while the aspect ratio of 1.01 corresponds to a nearly spherical particle. In our application of

5 the T-matrix code the second mode parameters (i.e. Rad_2=0.83 μm, σ_2=1.84, see Table 1) were used to specify the size distribution, and the AFCRL 1987 Sand indices are utilized.

Table 1 summarizes the inputs to the aerosol calculations, i.e. parameters for a size distribution with two log-normal distributions. Given the input size distribution and indices the resulting extinction spectra were distributed uniformly from the surface to 2 km altitude, with an additional stratospheric sulfate aerosol optical thickness of 0.006 distributed throughput the

10 stratosphere.

Table 1: Log-normal size distribution parameters for the three aerosol types considered

Aerosol Type	Den_1	Rad_1	σ_1	Den_2	Rad_2	σ_2	Indices
Sulfate	1.00	0.64	1.58	1.25	0.37	2.13	AFCRL 1987 H_2SO_4 300 K
Dust	-	-	-	-	0.83	1.84	AFCRL 1987 Sand
Brown Carbon	1.00	0.086	1.49	-	-	-	Alexander Brown Carbon

To highlight the optical properties of the simulated aerosols, Figure 2 shows the MODTRAN simulated radiances for the three

15 different aerosol types overlying a black surface. The observed radiance is simulated at an altitude of 3 km with a SZA of 25° and ground elevation at sea level. An AOT of 1.0 was selected for each aerosol type. The single scattering albedos close to unity of sulfate and dust have a larger effect on the simulated radiance, compared to the lower single scattering albedo of carbon. For a highly reflective surface, the effects would be reversed, and we would see the strongest deviation from the case of no aerosols for carbon.
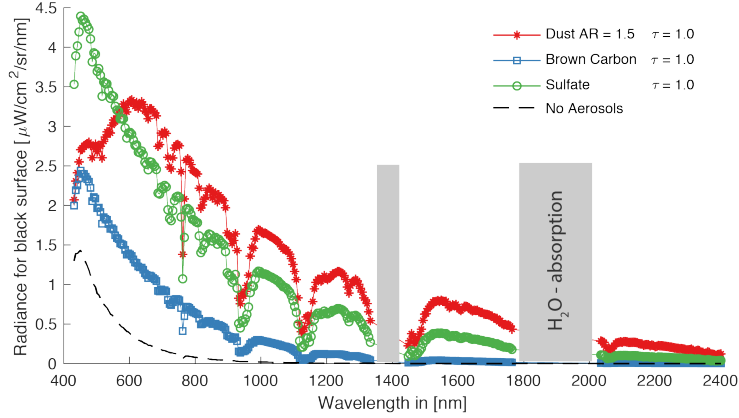
20

Figure 2: Radiance for a black surface, three different aerosol types and no aerosols, from an observed altitude of 3 km with a SZA of 25°. Every aerosol type (dust with an AR of 1.5, carbon and sulfate) is shown for an AOT of 1.0. No data are shown for the wavelengths of two water vapor absorption bands (grey bars).

5

2.4 Simulating Instrument Noise

Unlike radiative transfer calculations, the measured signal from real instruments contains noise. Therefore, we add noise to the radiative transfer calculations that is similar to the noise in the AVIRIS-NG instrument.

The noise is approximated by a three-parameter fit (see Equation 5) (Thompson et al., 2018) and derived from more complex

10 AVIRIR-NG noise models (Mouroulis et al., 2000, 2003; Tennant et al., 2008).

$$\sigma(L_{(\lambda)}, \lambda) = a(\lambda) * (b(\lambda) * L(\lambda))^{0.5} + c(\lambda) \tag{5}$$

The computed noise, $\sigma(L_{(\lambda)}, \lambda)$, is a function of three wavelength dependent parameters and observed radiance, $L(\lambda)$, which is a function of wavelength as well. Following a normal distribution, we randomly add the calculated noise to the radiative

15 transfer calculated radiances:

$$L_{noise}(\lambda) = L_{(\lambda)} + \mathcal{X} \quad \text{with} \quad \mathcal{X} \sim \mathcal{N}(0, \sigma(L_{(\lambda)}, \lambda)) \tag{6}$$

7

On average, the signal to noise level for a typical scene is about 100 in the ultraviolet, 200 in the visible, 300 in the near infrared and peaks at about 1600 nm with a signal to noise level of 700.

3 Inverse Model

After using the forward model to generate a dataset to train the neural network, we are now able to train an inverse model that
5   relates radiance spectra to AOT for the three aerosol types.

3.1 Model Architecture

A subclass of neural networks, called *multilayer perceptrons* (Werbos, 1974), have been shown to be able to approximate any linear or non-linear function (Hornik et al., 1989). After the training phase multilayer perceptrons can be used in real time at
10   low computational cost. This makes this model architecture ideal for our application. Neural networks have previously been used in many studies to extract information from remote sensing observations. For example to estimate cloud optical thickness and type (Minnis et al., 2016; Taravat et al., 2015), to un-mix surface types (Licciardi and Del Frate, 2011; Palsson et al., 2018) and to retrieve biophysical properties of vegetation (Verger et al., 2011; Xiao et al., 2014). Neural networks have also been applied to retrieve aerosol layer height from Ozone Monitoring Instrument (OMI) observations (Chimot et al., 2017), used to
15   estimate multiple aerosol parameters as the prior for an iterative Phillips-Tikhonov retrieval (Di Noia et al., 2017) and to estimate AOT from MODIS observations (Lary et al., 2009; Radosavljevic et al., 2010).

A multilayer perceptron is comprised of many individual ~~operations~~functional elements, or *neurons*, that multiply their inputs by a matrix, or *weight*s, sum the results and add an additional vector, called *bias*. A non-linear function, the *activation function*, is applied to the results, or *output*s, of these neurons, permitting non-linear projections from the input space to the output space.
20   In a network, the output of neurons can be used as the input to other neurons. Hence, neurons are organized in *layers*. In general, more layers, and more neurons per layer, allow for more complex information retrieval. However, if the network becomes too complex for a given dataset and task, it will perform poorly for new model inputs. The right number of neurons and layers as well as other parameters, or *hyperparameters*, have to be determined empirically for every application. Thus, we altered the hyperparameters, trained the neural network on the majority of the samples, the *training set*, and evaluated the
25   model performance with samples that are separate from the training set, the *validation set*. Once we could no further reduce a user-defined *cost function*, we froze the hyperparameters.

Our neural network consists of five layers between the input and output layer, or *hidden* layers, containing 128 neurons in the first four hidden layers each and 96 neurons in the last hidden layer. The input layer consists of 322 neurons and the output
30   layer consists of 3 neurons. The first five layers are *fully connected*, meaning that all layer outputs are used as layer inputs of the succeeding layer. The sixth layer is separated into three groups with 32 neurons each (see Figure 3). The inputs to the

neural network are the radiance at 319 wavelengths, the SZA, ground distance and ground elevation. The output of the network is the independently retrieved AOT of the three aerosol types.
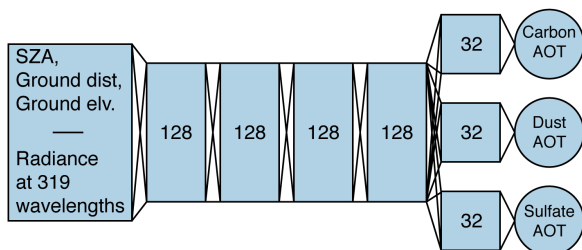
5   Figure 3: Model architecture of the neural network for aerosol retrieval. The inputs consist of the SZA, ground distance, ground elevation and the radiance at 319 individual wavelengths. The network has five hidden layers with 128 neurons in the first four layers and 96 neurons in the last hidden layer. The outputs of the network are the AOT for carbon-, dust- and sulfate aerosols.

10   To allow for non-linearities we add a rectified linear unit (ReLU) as the activation function, $g(x)$, which can be expressed as:

$$g(x) = \begin{cases} x, & if\ x \geq 0 \\ 0, & if\ x < 0 \end{cases} \tag{7}$$

where x is the output of a neuron. During training we minimize the cost function, given by:

$$cost = \alpha * R(\theta) + \frac{1}{2n} \sum_{j=1}^{n} \left(\bar{Y}_j - Y_j \right)^2 \tag{8}$$

15

$$R(\theta) = \|\theta\|_2 = \sqrt{\sum_{i=1}^{m} \theta_i^2} \tag{9}$$

For our network $\bar{Y}_j$ and Y_j are the n true and predicted AOT, respectively. We further add the L2 norm $\|\theta\|_2$ to the vector of the m neural network weights, θ, to our cost function (see Equation 9), also referred to as *L2 regularization* or *weight decay*.

This helps to avoid overfitting to the training set. The L2 regularization term, $R_{(\theta)}$, is weighted by α (see Equation 8) which is another hyperparameter that had to be determined empirically.

3.2 Pre-processing

From the 425 AVIRIS-NG channels we exclude calculated radiances at wavelengths with strong water vapor absorption. At these wavelengths, the majority of the surface spectra used in this study did not report or linear interpolate, surface reflectance. Furthermore, we exclude radiances at wavelengths that show strong signs of noise in the AVIRIS-NG data. A total of 319 wavelength channels remain. We then scale the radiance of a particular observation, L_j, by dividing through the cosine of the SZA_j~~SZA~~ and multiplying with the square of the ratio between Sun-Earth distance, d_j, and mean Sun-Earth distance, d_0~~d~~, (Equation 10). This scales the magnitude of L_j while preserving its spectral shape and is similar to deriving the *top of atmosphere reflectance*. Afterwards, we standardize the scaled observations, R_j, for the training process. During training, this results in a better conditioned cost function and allows the neural network to converge faster to a solution. Standardizing is performed by subtracting the mean, μ_λ, and dividing by the standard deviation, σ_λ, at every wavelength, (Equation 11). The mean and standard deviation was calculated from the complete set of radiative transfer calculations.

$$R_j = \frac{L_j * \left(\frac{d_j}{d_0}\right)^2 \cancel{d_j^2}}{\cos\left(SZA_j \cancel{SZA}\right)} \tag{10}$$

$$R_j = \frac{R_j - \mu_\lambda}{\sigma_\lambda} \quad with \quad \mu_\lambda = mean(R_j)_\lambda \quad and \quad \sigma_\lambda = std(R_j)_\lambda \tag{11}$$

3.3 Training, Validation and Test

The MODTRAN radiance samples were split into a trainings, validation and test set. The validation and test set contain 10,000 randomly chosen samples each and the training set consists of 280,000 samples. Training is performed with Googles' TensorFlow framework (Abadi et al., 2016). We gradually minimize the cost function by adjusting the randomly initialized weights and bias terms with the gradient-based optimizer Adam from Kingma and Ba (2014), at a learning rate of 0.001. During training we evaluate the neural network performance on the validation set and update the model architecture and training parameters. Once, the cost function cannot be further minimized, training is complete.

4 Results and Discussion

After training of the neural network is completed, we evaluate its performance on the test set. For the samples in the test set, that were not present during training, we find a linear correlation coefficient of 0.87, 0.98 and 0.96 for the AOT of carbon,
5 dust and sulfate, respectively (see Figure 4). The standard error for carbon-, dust- and sulfate aerosols is 0.05, 0.02 and 0.03, respectively. Thus, the model ~~has the smallest uncertainties~~ ~~accuracy is higher~~ for ~~the retrieval of~~ dust and sulfate, which ~~are less strongly absorbing across most of the relevant spectral range,~~ ~~have a larger single scattering albedo~~ compared to carbon. ~~Distinguishing the weaker signal from carbon from the surface contribution might be more challenging for the neural network and lead to the overall larger uncertainties for carbon. Additionally, the spectral signature of carbon and sulfate show high~~
10 ~~similarities at shorter wavelengths, except for a scaling factor (see Figure 2). This might further challenge the neural network to distinguishing the two. For wavelengths ranging from 500 nm to 600 nm, absorption of dust increases with decreasing wavelength, while the opposite is true for carbon and sulfate. This unique spectral signature of dust, compared to the other two aerosol types, might be another reason for the smaller standard error of its retrieval.~~
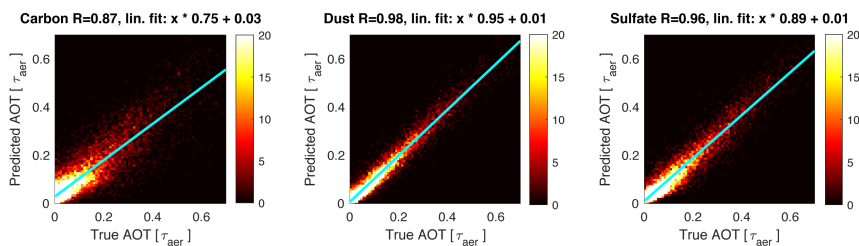

Figure 4: AOT for carbon, dust and sulfate aerosols, retrieved by the model vs true AOT from the test set. The cyan line shows the linear fit to the data with slope and y-intercept given in the respective titles.

We further investigate the model's performance for retrieved AOT under varying amounts of the three aerosol types. The
20 absolute error in retrieved AOT for the three aerosol types is shown in the top row of Figure 5. Horizontal gradients (vertical bands) indicate that the model's performance for the retrieval of a single aerosol type depends on the concentrations of the other aerosols in a given observation. Vertical gradients indicate that the model's performance is dependent on the AOT of the aerosol that we are trying to retrieve. For the error on retrieved carbon (Figure 5 a) and sulfate (Figure 5 c) we find dependencies on AOT while the error in the retrieval for dust (Figure 5 b) appears insensitive to its AOT. Examining the retrieval error in
25 percent of AOT (bottom row) we find that all three aerosol retrievals have higher relative errors for lower AOT and a standard

error of about 40% for an AOT of 0.1. We further analyzed the model's performance over a SZA range from 25° – 50°, ground elevation from 0 – 2000 m and ground distance from 3000 – 6000 m. No significant correlation between model error and the three parameters was found.
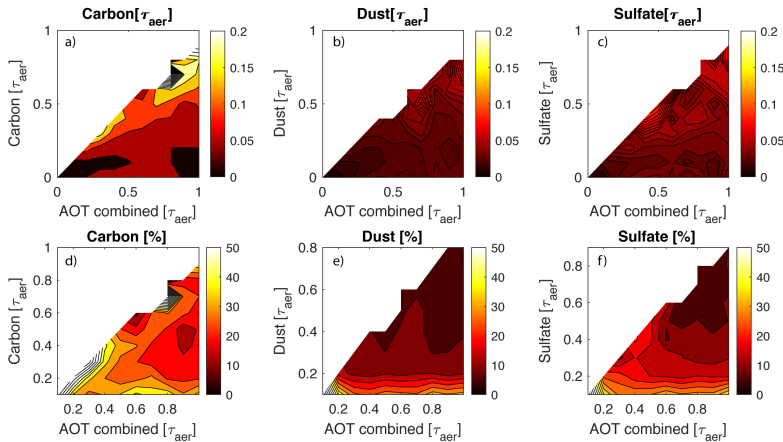

Figure 5: Error in retrieved AOT for carbon, dust and sulfate aerosols on the test set. (a, b, c) shows the absolute error while (d, e, f) shows the error in [%]. The color-mapping is held constant for each row and varies across the three columns.

4.1 Model Performance for Varying Surface Types

To investigate systematic, surface dependent biases in the model we derive AOT for the three aerosol types over various unmixed surface types. The data consist of 250,000 samples. The standard error and mean between true and predicted AOT for different surfaces types is summarized in Figure 6. For the retrieval of carbon, we find the largest standard error for asphalt with ± 0.08 and the largest systematic bias for grass of $+0.02$. For dust the largest systematic bias is less than +0.01 and occurs for scenes with concrete. The standard error is similar for all surface types and approximately ± 0.02. The systematic biases for the retrieval of sulfate aerosols are mostly negative with asphalt and concrete causing the largest bias of -0.01. Overall, the standard error for the retrieval of carbon over most surfaces is larger compared to the other two aerosol types. This is not surprising, considering the overall lower performance of the model for the retrieval of carbon aerosols. Note that the model's performance should be evaluated from the more realistic case of mixed surface spectra as was done in the previous section.
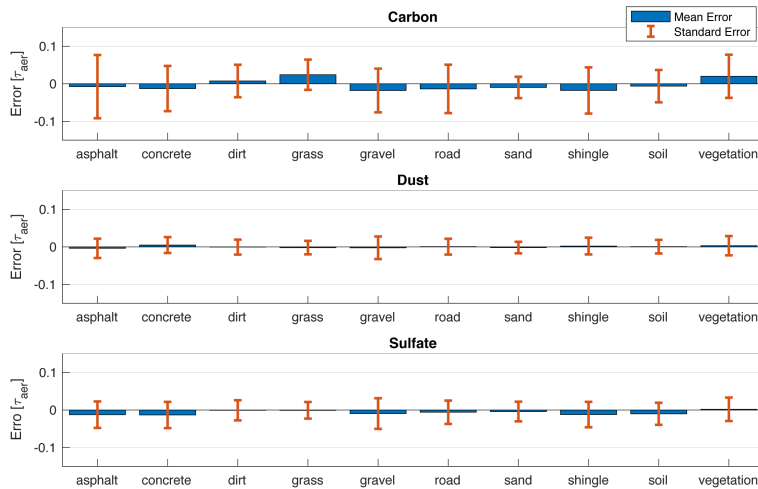
12

Figure 6: Mean error and standard error of retrieved AOT for carbon, dust and sulfate for various unmixed surface types from the test set.

5

4.2 Effect of Spectral Resolution, Sampling Resolution and Instrument Noise

Here we examine how spectral resolution, sampling resolution and instrument noise affect model performance. The underlying motivation is to estimate the model's performance for instruments other than AVIRIS-NG, which might have a 10 nm spectral resolution, fewer wavelength channels as well as a higher or lower signal to noise ratio. Hence, we train and analyze the

10   model's performance for an additional 23 networks with varying noise, spectral resolution and sampling resolution.

To simulate the fewer wavelength bands, the training-samples were reduced in sampling resolution, leaving 319, 107, 36 and 12 uniformly spaced, wavelengths per sample. Furthermore, to account for different signal to noise ratios, we changed the simulated AVIRIS-NG equivalent noise level (see Equation 5 and 6) by multiplying it with 0 (no noise), 1 and 3 before applying it to our training and test samples. Finally, we preformed all calculations once for the AVIRIS-NG spectral resolution

15 of approximately 5 nm and for a spectral resolution of 10 nm. All neural network parameters were kept constant, except the input layer, which had to be adapted to the reduced number of wavelengths. Training was stopped when the error on the validation set could not be reduced any further or we reached a maximum of 10,000 epochs, meaning that every training-sample was used during training 10,000 times. While we found dependencies of retrieval performance to varying amounts of

noise and number of wavelength channels, the spectral resolution had no significant effect. On average the models trained with a spectral resolution of 5 nm had a standard error in retrieved AOT that was only 0.001 smaller than for the cases with a spectral resolution of 10 nm. Therefore, we limit the following discussion to the results of the 12 neural networks trained on radiative transfer calculations with the AVIRIS-NG spectral resolution of approximately 5 nm and note that these values are
5   also representative for an instrument with a 10 nm spectral resolution.

The standard error on the test set of the respective 12 neural networks is shown in Figure 7. The left column shows the standard error for the complete test set (AOT is varied between 0 and 1) while the right column shows the standard error for low aerosol loadings, with AOT ranging between 0 and 0.3. As expected, we find a decrease in model accuracy for fewer wavelengths and more noise. This decrease in model accuracy, with respect to the idealized case of 319 wavelength bands and no noise, is
10   nearly symmetrical for our chosen test cases. Thus, if we reduce the number of wavelength bands by a factor of three the model has similar accuracy compared to if we add AVIRIS-NG equivalent noise and if we reduce the number of wavelength bands by a factor of nine the model has similar accuracy compared to applying three times AVIRIS-NG equivalent noise, and so on. This holds true for all aerosol types. Overall, the model has the highest accuracy for the retrieval of dust. To put the calculated standard errors in the left column into perspective: if the model would randomly guess the combined AOT of all three aerosols
15   between 0 and 1 and simply divide by three, the standard error would be ±0.10. Thus, all trained models show higher accuracy than guessing randomly. If we had a model that would be able to retrieve the combined AOT without error, and then simply divide by three, the standard error would be ±0.07. For the retrieval of carbon, the models with 12 wavelengths bands and 3 times AVIRIS-NG equivalent noise shows such a standard error. This is an indication that the AOT from carbon aerosols cannot be isolated from other aerosols for instruments with only 12 wavelengths and 3 times AVIRIS-NG equivalent noise.
20   The retrieval of dust and sulfate requires fewer wavelength bands and can tolerate more noise compared to the retrieval of carbon.

For aerosol retrieval under low AOT conditions (right column in Figure 7), a model that would guess the combined AOT randomly between 0 to 0.3 and divide by three, would have a standard error of ±0.06 and a model that can determine the combined AOT perfectly and then simply divides by three would have a standard error of ±0.03. Most combinations of
25   wavelength bands and instrument noise have standard errors that exceeds this threshold of ±0.03 for the retrieval of carbon. This highlights the limitations of the model for the separation of carbon aerosols for low levels of AOT. Additionally, it stresses the importance of low noise hyperspectral instruments, such as AVIRIS-NG.
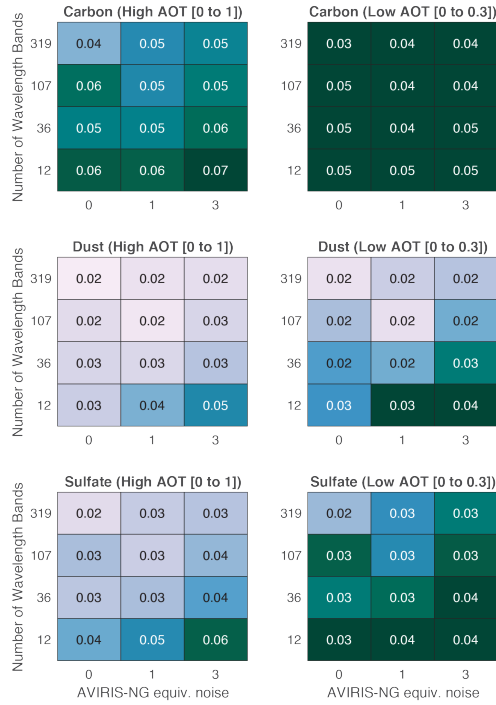
Figure 7: Standard error for retrieved AOT of 12 individually trained neural networks with (319, 107, 36, 12) wavelength bands and varying amount of simulated AVIRIS-NG equivalent noise (0, 1, 3) (see Equation 5 and 6) from the test set. Left column shows the standard error when AOT is varied between 0 and 1. Right column shows the standard error for AOT between 0 and 0.3.

4.3 Sensitivity Analysis

It is inherently difficult to interpret the inner workings of neural networks. However, by perturbing the inputs and observing the changes of the outputs one can infer the relative importance of an input for a given model (Blackwell, 2012). We perform such a sensitivity analysis by increasing one input at a time by 1%, while keeping the other 321 inputs unchanged. The model output is then calculated for the entire test set and compared to the retrieval without the perturbation. For example, AOT of carbon is derived while the model input, representing the observed radiance at 500 nm, is increased by 1%. All other model

inputs, for example radiance at 600 nm and 700 nm or SZA, are kept unchanged. We perform such a sensitivity analysis once for the model trained without noise (an ideal instrument) and once for the model trained with AVIRIS-NG equivalent noise. The sensitivity to every input is shown in Figure 8. For the model trained without noise (top, third and fifth row) we find more sensitivity at 687 nm and 762 – 767 nm for the retrieval of carbon and dust while sulfate shows more sensitivity to the latter.

5   These wavelengths correspond to the oxygen B-and A-band located at 685 nm – 695 nm and 759 - 771 nm, respectively. Multiple studies have suggested the use of these absorption bands for the retrieval of AOT and its vertical structure (Dubuisson et al., 2009; Heidinger and Stephens, 2000; Min et al., 2004). The sensitivity to small perturbations of SZA, ground distance and ground elevation is small compared to the radiances. From these three model inputs, surface elevation is indicated to be the most important for the retrieval of dust and sulfate.

10   For the model trained with AVIRIS-NG equivalent noise we find approximately an order of magnitude lower sensitivity at shorter wavelengths compared to their respective counterparts trained without noise (Note the different y-scales for the six sensitivity plots). This demonstrates how the model adapted to small perturbations (noise) at individual wavelengths by becoming less sensitive to these perturbations. For longer wavelengths, the change in sensitivity is less pronounced. In general, we observe a relative shift in sensitivity from shorter towards longer wavelengths when instrument noise is added. The shift

15   in sensitivity to longer wavelengths might be a direct effect of the noise distribution of AVIRIS-NG which allows for a higher signal to noise ratio at longer wavelengths. Additionally, there is an overall smoother shift in sensitivity between neighboring wavelengths. This can be interpreted as the model relying on multiple neighboring wavelengths to obtain their shared information content, rather than interpreting wavelengths individually.
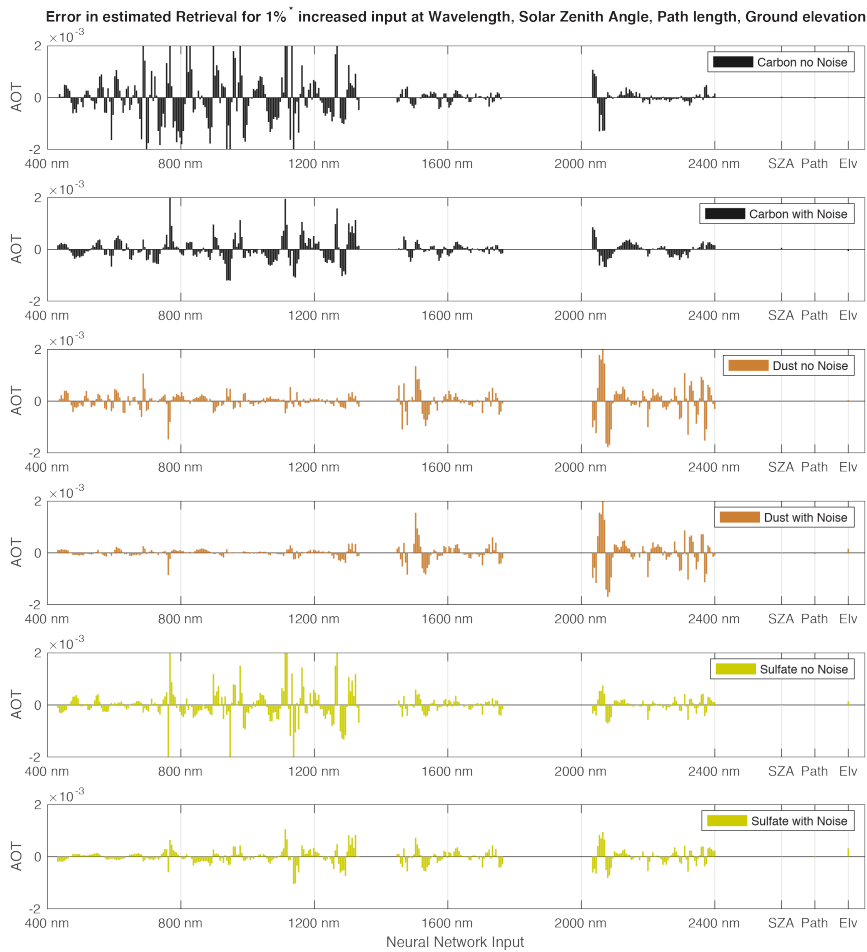
Figure 8: Sensitivity for retrieved AOT of carbon, dust and sulfate to all model inputs. The x-axis shows the model inputs (radiances at shown wavelength, SZA, ground distance and ground elevation). The y-axis shows the difference in retrieved AOT when increasing a given input by 1% while keeping all other inputs unchanged.

5

5 Applying the Model to Real Imagery

We apply the trained neural network to AVIRIS-NG observations from a flight campaign in 2016 over India in collaboration with Space Applications Centre, Indian Space Research Organization (SAC, ISRO). The results are compared to MODIS and
5   AERONET retrieved AOT and a reanalysis product.

5.1 Preprocessing of AVIRIS-NG Observations

To remove remaining noise in the AVIRIS-NG observations we use a principal component analysis (PCA) (Wold et al., 1987) and inspect the generated eigen-images manually. The PCA is only applied to the 319 wavelength channels that we used to
10   train the model on. As stated before, these channels were down selected from the 425 AVIRIS-NG channels to avoid wavelength bands with strong water absorption and instrument noise. The first 16 components explain approximately 99.9% of the variability and are dominated by image features (see Figure 9). Most higher principal components are dominated by systematic noise (vertical stripes along the flight path). We reconstruct the AVIRIS-NG observed radiances from these first 16 principal components. This effectively removes principal components higher than 16 from all analyzed AVIRIS-NG imagery.
15   Afterwards, the radiance for every pixel is treated as an independent observation and scaled and standardized (Equation 10 and 11) to match the training set. We acknowledge that the choice of retaining the first 16 principal components is rather arbitrary and should ideally be made on a per flight basis. However, for practical reasons we decided to use one threshold for all imagery considered in this study. The threshold is a tradeoff between removing valuable information and reducing noise. Experiments with more and fewer principal components indicated that the model was insensitive to the exact number of
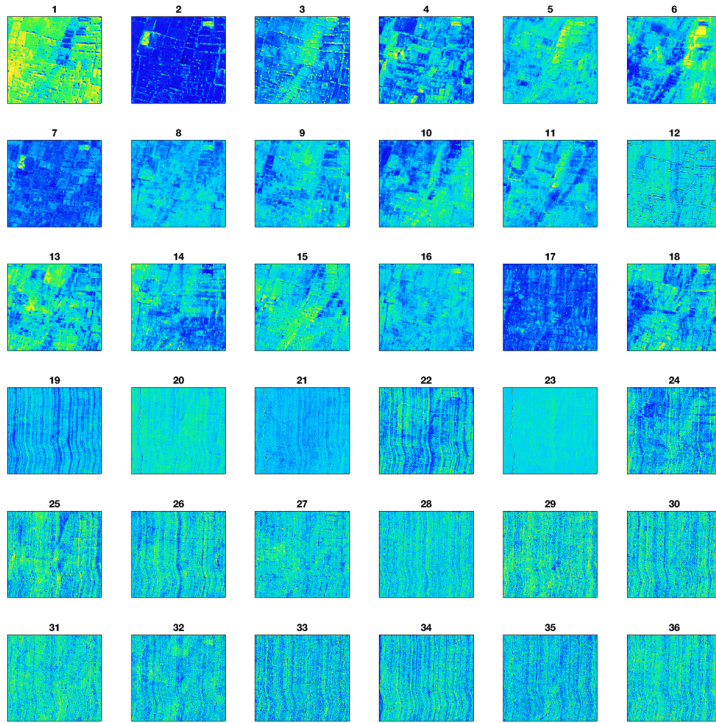20   remaining principal components.

Figure 9: First 36 eigen-images from an AVIRIS-NG flight on 10/01/2016 near Coimbatore, India. Shown is a spatially resolved scene of 100 x 100 ground pixel, approximately 500 x 500 m. Instrument artifacts (vertical stripes) are visible for eigen-images greater than 16 (for example 19 and 22).

5.2 Novelty Detection

Our model is trained on a limited set of training examples. The set of surface types available for training is not complete. Generally speaking, library spectra of surface materials vastly under-represent the spectral variability of surface materials found in nature. The variety of surface materials is just too great to include in any single library. Applying the model to scenes with new surface types, which have significant differences compared to the surface types in the training set can lead to false

aerosol retrieval by the model. Hence, it is important to measure the similarity of a given AVIRIS-NG scene to the training examples and discard individual image pixel that are far outside of the training space. This is referred to as *novelty detection*. For this purpose, we train a second neural network proposed by Japkowicz et al., (1995) on the training samples with AVIRIS-NG equivalent noise. The network architecture is an auto-associative multilayer perceptron (Kramer, 1992) with three hidden

5   layers and shown in Figure 10. All three hidden layers use a ReLU activation function and consist of 512, 32 and 512 neurons, each. The input- and output-layer consist of 322 neurons, each. The network takes 319 radiances at individual wavelengths (measurements of one image pixel), SZA, ground distance and ground elevation as input parameters and is trained to reproduce these parameters after some computation by the network. The network is trained in a manner similar to the model for aerosol retrieval and uses the same optimization algorithm and cost function (see Equation 8) with $n = 322$, and Y'_j and Y_j being the

10   original and reproduced radiances and SZA, ground distance and ground elevation. The first three layers (Input, Compression and Bottleneck) act similarly to deriving the first 32 principal components but are non-linear. The last two layers (Decompression and Output) can be interpreted as reproducing the radiances only from their first 32 principal components, but again, are non-linear. After the replication of the input parameters we compare those to the original inputs and calculate the mean square error between the two. During training, the neural network learns to minimize this error. For example, the

15   neural network learns that the radiance at 2100 nm is highly correlated with the radiance at 2300 nm. Thus, it can reconstruct (decompress) both radiances with only one value passed in from the Bottleneck layer with little error. Once the neural network is trained and applied to previously unseen features it will compress and decompress features that are similar to the training set (high correlation between 2100 nm and 2300 nm) with a smaller error than features that are different (low correlation between 2100 nm and 2300 nm). Finally, a threshold for the error is determined as a tradeoff between the number of remaining

20   aerosol retrieval and the number of remaining outliers. Samples above the determined threshold are considered new and not considered for the aerosol retrieval.
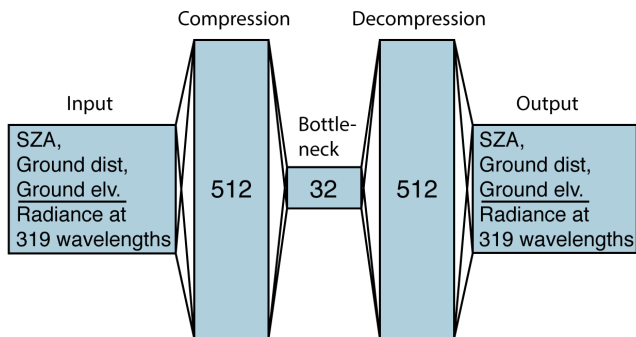
20

Figure 10: Auto-associative neural network for novelty detection used for novelty detection. The input and output layer consist of SZA, ground distance, ground elevation and radiances at 319 wavelengths. The network has three hidden layers with 512, 32, 512 neurons per layer.

5

5.3 Results

Figure 11 and Figure 12 show the aerosol retrieval for two of the 21 analyzed AVIRIS-NG scenes. The scene in Figure 11 was captured on 02/04/2016 near Kota, India. It shows a detail of the flight with 100 x 500 pixels and an approximate ground resolution of 5 m per pixel. The median and standard deviation of the retrieval is indicated at the top of the first four panels,

10 showing the combined AOT and un-mixed AOT for carbon, dust and sulfate. The normalized mean square error from the auto-associative neural network for novelty detection and a true color image is shown on the right as well. Image pixels that lie above a user defined threshold are highlighted in red and discarded. For the scene shown in Figure 11 the discarded image pixels consist of water features in the middle and bottom portions of the scene as well as some agricultural sites. The detection of water by the neural network for novelty detection is to be expected, since the spectral shape of water is very different to

15 most land surfaces and was not part of the training set. The aerosol retrieval still includes surface features. For example, it overestimates carbon aerosols over what appears to be a street (middle of second plot from the left). Some residual surface features are not entirely unexpected as less challenging atmospheric retrievals from imaging spectroscopy, for example water vapor (Thompson et al., 2015), often contain surface reflectance artifacts. The detail shown in Figure 12 is from an AVIRIS-NG flight near Gundlupet, India from 01/10/2016. The model for novelty detection excluded mostly individual fields with bare

20 soil. Similar to the figure above, we find some residual surface features in the retrieval. Both images show the limitation of the model in distinguishing small variations in AOT from different surface types. To minimize the residual surface features a median filter could be applied in post processing at the cost of lower spatial resolution.
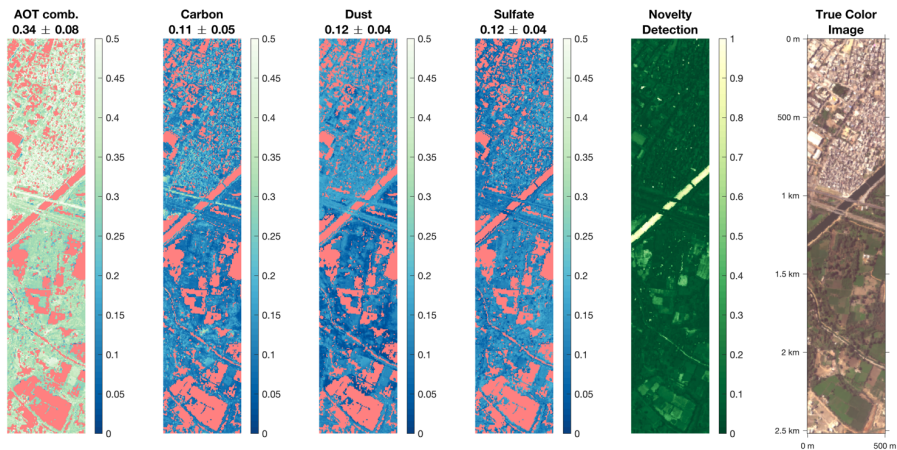
21

Figure 11: Aerosol retrieval with the model from AVIRIS-NG imagery near Kota, India, 02/04/2016. The median and standard deviation of the retrieval is indicated at the top of each panel. The normalized output of the neural network for novelty detection is shown in the panel, second from the right. Values above a chosen threshold are discarded from the aerosol retrievals and highlighted in red (e.g. a river in the middle of the images). A true color image of the scene is shown as well for reference.
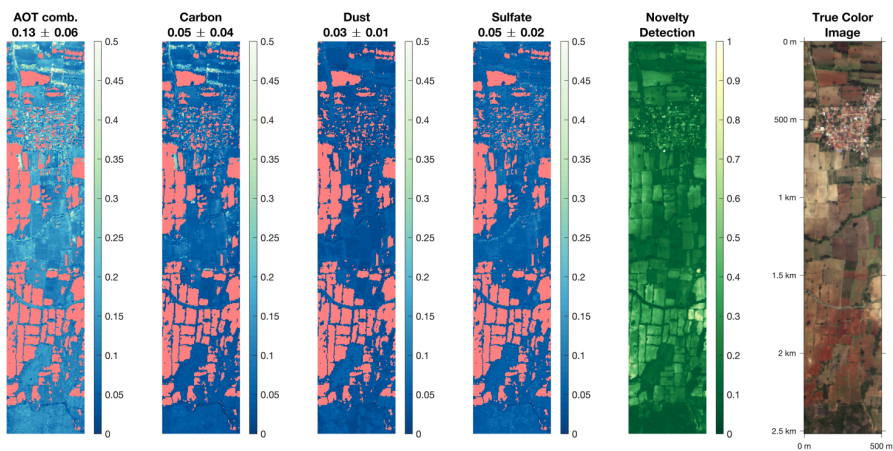
Figure 12: Aerosol retrieval with the model from AVIRIS-NG imagery near Gundlupet, India, 01/10/2016. The median and standard deviation of the retrieval is indicated at the top of each panel. The normalized output of the neural network for novelty detection is shown in the panel, second from the right. Values above a chosen threshold are discarded from the aerosol retrievals and highlighted in red. A true color image of the scene is shown as well for reference.

5

5.4 Comparison to AERONET and MODIS

We compare the combined aerosol retrievals from AVIRIS-NG to AERONET and MODIS retrievals. AERONET is a network of ground-based sun photometers distributed around the globe (Holben et al., 1998). AERONET instruments derive AOT at multiple wavelengths with an uncertainty of 0.01 to 0.02 (Eck et al., 1999). These low uncertainties make AERONET stations

10 a common source for validation of air- and space-borne AOT retrieval (Bilal et al., 2014; Chu et al., 2003; Levy et al., 2013). However, there are sparse AERONET locations in India. We, therefore, add a second source of AOT retrievals to the comparison from MODIS observations. MODIS makes daily and nearly global observations from two platforms, Aqua and Terra. MODIS has a spectral range from 410 nm to 14.5 μm over 36 discrete wavelength bands. Its ground resolution is better than 1 km, depending on the wavelength band (Salomonson et al., 1989). Two algorithms are utilized to derive AOT form

15 MODIS observations. The Dark Target (Kaufman et al., 1997) algorithm is used for dark ground targets such as vegetation and water. The Deep Blue (Hsu et al., 2004, 2006) algorithm is applied to measurements over dark and bright surfaces although it was originally developed for the aerosol retrieval over bright desert regions. Over land, MODIS retrieved AOT has an expected standard error of 0.05 + 15% of AOT (Levy et al., 2013). MODIS has larger uncertainties than AERONET, but the retrievals are in closer spatial and temporal proximity to the AVIRIS-NG flights.

20 For the period of the 21 AVIRIS-NG flights only three AERONET stations within India were operational. These are Gandhi College at 25.9°N 84.1°E, Jaipur at 26.9°N 75.8°E and Pune at 18.5°N 73.8°E. We make use of the daily means of their Level 2.0 data product, which is cloud-cleared and manually inspected. The locations of all three stations are shown in Figure 13 together with the location of all 21 AVIRIS-NG flights considered in the study. For a given flight we consider the AOT retrieved from all three AERONET stations within 1 and 2 days of the flight date. The time averaged, retrieved AOT of each

25 AERONET station, τ_{aer_i}, is weighted proportionally to the square of the distance, d_i, between station and flight:

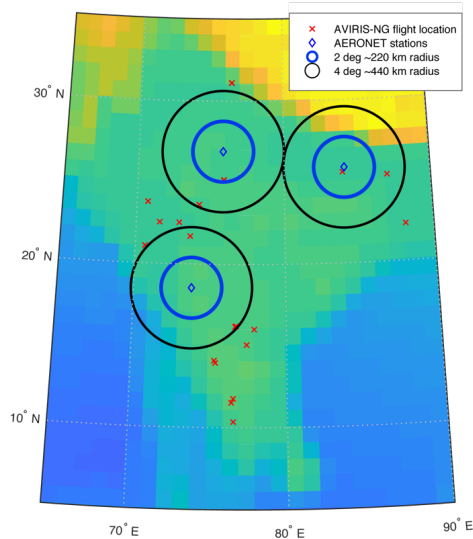$$\tau_{aer} = \frac{\sum_{i=1}^{3} \tau_{aer\_i} * d_i^{-2}}{\sum_{i=1}^{3} d_i^{-2}} \qquad (12)$$

23

Figure 13: Location of AERONET stations and AVIRIS-NG flights on a topographic map of India. The AERONET stations are marked with blue diamonds, a 4 deg and 2 deg (approximately 440 km and 220 km) radius around each station is indicated with a black and blue circle. The AVIRIS-NG flight locations are shown with red x. AVIRIS-NG flights outside the circles are not considered for the comparison to AERONET.

The comparison between AOT retrieved by AERONET and the AVIRIS-NG flights is shown in Figure 14 and Figure 15 for AERONET retrievals within 1 day and 2 days, respectively. For the comparison within 1 day of the AVIRIS-NG flights only four AERONET stations reported their measured AOT. Only one comparison falls within the specified 1-day window and is within 2 deg ($\approx$ 220 km) of the flight location (red circle). The three other comparisons are for flights with a distance ranging from 2 deg to 4 deg between the AERONET station and the AVIRIS-NG flight. The standard deviation of all considered AERONET retrievals that we compare to for a given flight is indicated by the vertical bars. The standard deviation within a scene for the analyzed AVIRIS-NG flights is shown with horizontal bars. For the 4 comparisons we find a root mean square difference (RMSD) of 0.09. However, due to the large spatial distance between AERONET stations and the considered AVIRIS-NG flights this value has to be interpreted with caution and comes with large uncertainties. Nevertheless, we included this comparison for completeness and hope to have more collocated flights of AVIRIS-NG and AERONET stations in the future.

24

Considering AERONET observations within 2 days of the flights we are able to compare eight flights in total with three flights within 2 deg and 5 flights within 2 to 4 deg. The RMSD for all eight comparisons is 0.08. Again, we caution that the distance between AERONET stations and AVIRIS-NG flights is significant. For the comparison within 2 days, the closest comparison has a distance of about 40 km and is shown in Figure 15 (circled red and furthest to the right).
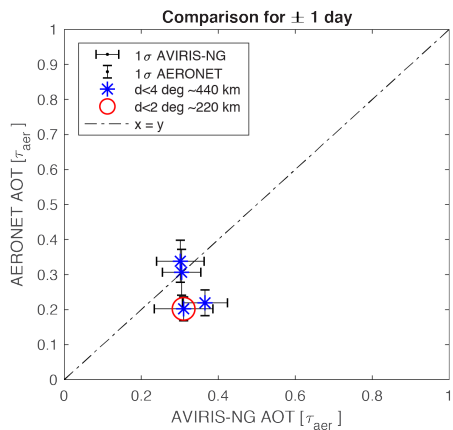
5

Figure 14: AOT retrieved by AERONET (see Equation 12) and the AOT retrieved from AVIRIS-NG with the model. The standard deviation of the considered AERONET measurements is shown with vertical bars and the standard deviation for the retrieval with AVIRIS-NG with horizontal bars. All comparisons between AERONET and AVIRIS-NG
10 flights are located within 4 deg ($\approx$ 440 km) and within 1 day from each other. The one comparison within 2 deg ($\approx$ 220 km) is circled in red.
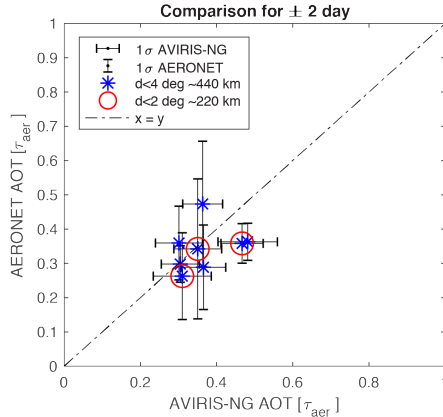
Figure 15: AOT retrieved by AERONET (see Equation 12) and the AOT retrieved from AVIRIS-NG with the model. The standard deviation for AERONET is shown with vertical bars and the standard deviation for the retrieval with horizontal bars. All comparisons between AERONET and AVIRIS-NG flights are located within 4 deg ($\approx$ 440 km) and within 2 day from each other. The three comparisons within 2 deg ($\approx$ 220 km) are circled in red.

For the comparison to MODIS we make use of the Collection 6, 'MODIS/Terra and MODIS/Aqua Level-2 (L2) Aerosol Product' (Levy et al., 2015, 2013). More specifically, we use the science data set 'AOD_550_Dark_Target_Deep_Blue_Combined' within the specified aerosol product. These data have a spatial resolution of 10 x 10 km and are derived utilizing the Dark Target and Deep Blue algorithm. All AOT retrievals come with a Quality Assurance Confidence (QAC), which is a measure of the algorithm performance. The QAC is determined by the number of examined pixel, fitting error and whether the solution falls into realistic physical conditions (Levy et al., 2013). In our study, we only consider derived AOT with the highest QAC = 3 and consider retrievals within 1 day and 0.2 deg $\approx$ 22 km of the AVIRIS-NG flights. The spatiotemporal cutoff is chosen as close in time and space as possible, while avoiding AVIRIS-NG flights with no collocated MODIS retrievals. This results in an average of 55 and minimum of 13 MODIS retrievals per AVIRIS-NG flight that we compare to. The comparison for the 21 AVIRIS-NG flights to the MODIS retrieved AOT is shown in Figure 16. The two AOT retrievals have a significant correlation of 0.81 and a RMSD of 0.12. However, the correlation might be mainly driven by the few high-AOT comparisons. The correlation and RMSD is similar to comparisons between AERONET and MODIS for India, with a correlation of 0.86 and RMSD of 0.19 (Gupta et al., 2018). Furthermore, AVIRIS-NG shows a positive bias of 0.07 compared to MODIS, which itself has a positive bias compared to AERONET (Gupta et al., 2018; Wang et al., 2019). This indicates that the AVIRIS-NG retrievals might overestimate combined AOT. Whether this bias holds true for a larger sample size and whether it is grounded in the model or the calibration of AVIRIS-NG warrants further

investigation. Interestingly, the two outliers at the bottom of Figure 16, where MODIS reports almost no aerosols are only 20 km and 1 day apart from each other. It has to be noted that the presented model was trained purely on radiative transfer calculations and not adjusted or calibrated to match the aerosol retrieval from MODIS or AERONET in any way. As with the comparison to AERONET, the comparison to MODIS comes with caveats. In essence, MODIS faces the same challenges as

5   our model, namely detecting the weak signal of aerosols in the presence of a strong signal from the underlying surface. Furthermore, MODIS AOT retrievals have a different spatial resolution and stem from observations recorded at different times than the AVIRIS-NG flight tracks. Nevertheless, in the absence of higher accuracy collocated measurements we included the comparison to MODIS.
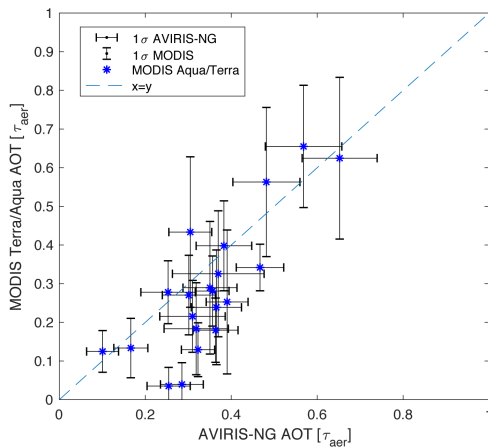

10

Figure 16: AOT retrieved by MODIS (y-axis) vs AOT retrieved by AVIRIS-NG with the model (x-axis). The standard deviation for MODIS is shown with vertical bars and the standard deviation for the retrieval with horizontal bars.

5.5 Comparison to CAMS

15   We further compare the retrieved AOT to the Copernicus Atmosphere Monitoring Service (CAMS) product. The CAMS system provides global analysis and forecasting of AOT for organic matter, dust and sulfate and is further described in (Benedetti et al., 2009; Morcrette et al., 2009). CAMS accounts for aerosol emissions, transport, sedimentation and deposition of various aerosol types. In contrast to MODIS and AERONET, one can directly compare the CAMS AOT for a specified aerosol type to the retrieved AOT. We make use of the CAMS 'near-real-time' product at a spatial resolution of 0.125°

20   available at: https://apps.ecmwf.int/datasets/data/cams-nrealtime/levtype=sfc/.

27

Figure 17 shows the comparison for the three considered aerosol types with the CAMS modeled AOT on the y-axis and AVIRIS-NG retrieved AOT on the x-axis. There seems to be general agreement between CAMS and AVIRIS-NG with AVIRIS-NG retrievals being on average 0.03 higher. The standard deviation of the difference between CAMS and AVIRIS-NG for the 21 analyzed scenes is 0.02, 0.04, 0.05 for carbon, dust and sulfate, respectively. For AOT below 0.1, CAMS and AVIRIS-NG differ significantly for carbon and dust with AVIRIS-NG retrieving higher AOT.

We further evaluate the agreement between AVIRIS-NG and CAMS for which aerosol type has the largest, second largest and smallest AOT in a given scene. For 16 out of the 21 compared scenes AVIRIS-NG and CAMS agree on which aerosol type dominates with the largest AOT (see Figure 18). 14 out of 21 compared scenes agree on which aerosol type has the second largest AOT and 16 out of 21 agree on which aerosol type has the smallest AOT. Thus, if one were to use the proposed neural network to classify the dominant aerosol type from AVIRIS-NG observations, the neural network would have an estimated accuracy of 76%.
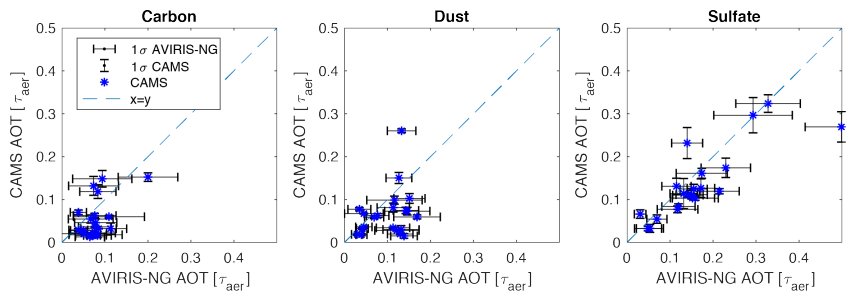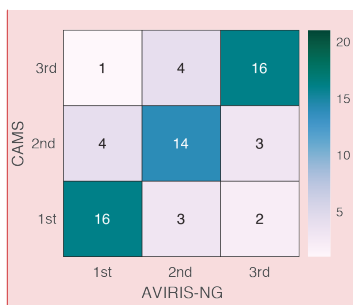

Figure 17: AOT modeled by CAMS (y-axis) vs AOT retrieved from AVIRIS-NG spectra with the neural network (x-axis). The standard deviation of the CAMS modeled AOT within 6 hours and 0.125° of the AVIRIS-NG observations are shown with vertical bars and the standard deviation for the AVIRIS-NG retrievals with horizontal bars.

28

Figure 18: Number of scenes where CAMS and AVIRIS-NG agree and disagree for which aerosol type has the largest (1st), 2nd largest and 3rd largest AOT. The total number of analyzed scenes is 21. Thus, the 16 in the bottom left corresponds to 16 scenes in which AVIRIS-NG and CAMS identified the same aerosol type as having the largest AOT, out of 21 scenes.

6 Conclusion

We demonstrated the retrieval of AOT from externally mixed dust, sulfate and carbonaceous aerosols from hyperspectral imagery with no a priori information of surface albedo or atmospheric state. We showed how sampling resolution and instrument noise influences the retrieval and, as expected, we find a decrease in model performance for fewer wavelengths and increased instrument noise. These results underline the need for low noise hyperspectral instruments. A sensitivity analysis gave insight in which wavelengths are important and how the neural network compensates for instrument noise; shifting sensitivity to multiple neighboring wavelengths and to longer wavelengths. We applied our model to AVIRIS-NG observations from a recent campaign over India and compared the retrieved AOT to AERONET and MODIS retrievals. The comparison to AERONET show a RMSD in AOT of 0.09 and 0.08 for collocated flights within 1 and 2 days, respectively. The comparison to MODIS finds a RMSD of 0.12. From a test set of radiative transfer calculations, we are able to retrieve AOT independently for dust, sulfate and brown carbon with a standard error of 0.03, 0.03 and 0.05, respectively. At execution time the presented neural network methodology can be executed at almost no computational cost. On a high-end consumer laptop (MacBook Pro CPU: i7 at 2.6 GHz) one can extract AOT, with the presented model, at about 250,000 spectra per second.

The results shown here are promising but also underline the difficulties of retrieving aerosol properties, especially over land: aerosol extinction is a weak, slowly varying spectral signal. Hyperspectral measurements can reduce uncertainty in aerosol

remote sensing, and we demonstrate that neural networks provide an efficient means for extracting information from large, multi-dimensional data sets, such as hyperspectral data cubes. As future satellite capabilities increase to acquire high spatial resolution hyperspectral data, there is a need to be able to process the large amount of data in a reasonable amount of time. Neural Networks can provide a solution for this task.

5

6.1 Future work

The current set of AVIRIS-NG flights in India has only a limited number of AERONET stations in close proximity to the various flight paths. To further validate our model, more collocated comparisons to AERONET observations are necessary. Deployed on a global platform, such as the upcoming CLARREO pathfinder or HyspIRI mission, many collocated observations with AERONET could systematically validate the retrieval and further improve the model performance through fine tuning. Furthermore, in situ microphysical measurements are necessary to validate the retrieved aerosol types. Finally, the presented methodology can be expanded in the future to retrieve other atmospheric and surface properties, such as water vapor, cloud properties and surface reflectance.

15

Code availability
The code is available at: https://github.com/SteffenMauceri/Aerosol-Retrieval-from-Hyperspectral-observations

Author contribution
Mauceri designed, implemented, analyzed and documented the research; Massie provided aerosol parameterization and optical properties; Massie, Kindel and Pilewskie provided guidance for research set-up, analysis and writing.

Competing interests
The authors declare that they have no conflict of interest.

25

References

Abadi, M., Barham, P., Chen, J., Chen, Z., Davis, A., Dean, J., Devin, M., Ghemawat, S., Irving, G., Isard, M. and others: Tensorflow: a system for large-scale machine learning., in OSDI, vol. 16, pp. 265–283., 2016.

Adler-Golden, S. M., Matthew, M. W., Bernstein, L. S., Levine, R. Y., Berk, A., Richtsmeier, S. C., Acharya, P. K., Anderson, G. P., Felde, J. W., Gardner, J. A. and others: Atmospheric correction for shortwave spectral imagery based on MODTRAN4, in Imaging Spectrometry V, vol. 3753, pp. 61–70., 1999.

Albrecht, B. A.: Aerosols, cloud microphysics, and fractional cloudiness, Science., 245(4923), 1227–1230,

doi:10.1126/science.245.4923.1227, 1989.

Alexander, D. T. L., Crozier, P. A. and Anderson, J. R.: Brown Carbon Spheres in East Asian Outflow and Their Optical Properties, Science., 321(5890), 833 LP – 836 [online] Available from: http://science.sciencemag.org/content/321/5890/833.abstract, 2008.

Baldridge, A. M., Hook, S. J., Grove, C. I. and Rivera, G.: The ASTER spectral library version 2.0, Remote Sens. Environ., 113(4), 711–715, 2009.

Benedetti, A., Morcrette, J.-J., Boucher, O., Dethof, A., Engelen, R. J., Fisher, M., Flentje, H., Huneeus, N., Jones, L., Kaiser, J. W. and others: Aerosol analysis and forecast in the European centre for medium-range weather forecasts integrated forecast system: 2. Data assimilation, J. Geophys. Res. Atmos., 114(D13), 2009.

Berk, A., Conforti, P., Kennett, R., Perkins, T., Hawes, F. and Van Den Bosch, J.: MODTRAN 6: A major upgrade of the MODTRAN radiative transfer code, in Proc.SPIE, vol. 90880H, pp. 9088-9088–7., 2014.

Bilal, M., Nichol, J. E. and Chan, P. W.: Validation and accuracy assessment of a Simplified Aerosol Retrieval Algorithm (SARA) over Beijing under low and high aerosol loadings and dust storms, Remote Sens. Environ., 153, 50–60, doi:10.1016/j.rse.2014.07.015, 2014.

Blackwell, W. J.: Neural network Jacobian analysis for high-resolution profiling of the atmosphere, EURASIP J. Adv. Signal Process., 2012(1), 71, 2012.

Chimot, J., Veefkind, J. P., Vlemmix, T., de Haan, J. F., Amiridis, V., Proestakis, E., Marinou, E. and Levelt, P. F.: An exploratory study on the aerosol height retrieval from OMI measurements of the 477 nm O_2-O_2 spectral band using a neural network approach, Atmos. Meas. Tech., 10(3), 783–809, doi:10.5194/amt-10-783-2017, 2017.

Chu, D. A., Kaufman, Y. J., Zibordi, G., Chern, J. D., Mao, J., Li, C. and Holben, B. N.: Global monitoring of air pollution over land from the Earth Observing System-Terra Moderate Resolution Imaging Spectroradiometer (MODIS), J. Geophys. Res. Atmos., 108(D21), 2003.

Davis, C. O., Bowles, J., Leathers, R. A., Korwan, D., Downes, T. V., Snyder, W. A., Rhea, W. J., Chen, W., Fisher, J., Bissett, W. P. and Reisse, R. A.: Ocean PHILLS hyperspectral imager: design, characterization, and calibration, Opt. Express, 10(4), 210–221, doi:10.1364/OE.10.000210, 2002.

Dennsion, P. and Gardner, M.: Hawaii 2000 vegetation species spectra, Available on-line [http://ecosis.org], from the Ecological Spectral Information System, doi:10.21232/C2HT0K, 2000.

Diner, D. J., Beckert, J. C., Reilly, T. H., Bruegge, C. J., Conel, J. E., Kahn, R. A., Martonchik, J. V, Ackerman, T. P., Davies, R., Gerstl, S. A. W. and others: Multi-angle Imaging SpectroRadiometer (MISR) instrument description and experiment overview, IEEE Trans. Geosci. Remote Sens., 36(4), 1072–1087, 1998.

Van Donkelaar, A., Martin, R. V, Brauer, M., Kahn, R., Levy, R., Verduzco, C. and Villeneuve, P. J.: Global estimates of ambient fine particulate matter concentrations from satellite-based aerosol optical depth: development and application, National Institute of Environmental Health Science., 2010.

Dubovik, O., Holben, B., Eck, T. F., Smirnov, A., Kaufman, Y. J., King, M. D., Tanré, D. and Slutsker, I.: Variability of Absorption and Optical Properties of Key Aerosol Types Observed in Worldwide Locations, J. Atmos. Sci., 59(3), 590–608, doi:10.1175/1520-0469(2002)059<0590:VOAAOP>2.0.CO;2, 2002.

Dubovik, O., Li, Z., Mishchenko, M. I., Tanré, D., Karol, Y., Bojkov, B., Cairns, B., Diner, D. J., Espinosa, W. R., Goloub, P. and others: Polarimetric remote sensing of atmospheric aerosols: instruments, methodologies, results, and perspectives, J. Quant. Spectrosc. Radiat. Transf., 224, 474–511, doi:https://doi.org/10.1016/j.jqsrt.2018.11.024, 2019.

Dubuisson, P., Frouin, R., Dessailly, D., Duforêt, L., Léon, J.-F., Voss, K. and Antoine, D.: Estimating the altitude of aerosol plumes over the ocean from reflectance ratio measurements in the O_2 A-band, Remote Sens. Environ., 113(9), 1899–1911, 2009.

Eck, T. F., Holben, B. N., Reid, J. S., Dubovik, O., Smirnov, A., O'neill, N. T., Slutsker, I. and Kinne, S.: Wavelength dependence of the optical depth of biomass burning, urban, and desert dust aerosols, J. Geophys. Res. Atmos., 104(D24), 31333–31349, 1999.

Fenn, R. W., Clough, S. A., Gallery, W. O., Good, R. E., Kneizys, F. X., Mill, J. D., Rothman, L. S., Shettle, E. P. and Volz, F. E.: Optical and infrared properties of the atmosphere, Handb. Geophys. Sp. Environ., 18, 1985.

Folkman, M. A., Pearlman, J., Liao, L. B. and Jarecke, P. J.: EO-1/Hyperion hyperspectral imager design, development, characterization, and calibration, in Hyperspectral Remote Sensing of the Land and Atmosphere, vol. 4151, pp. 40–51., 2001.

5   Gao, B.-C., Heidebrecht, K. B. and Goetz, A. F. H.: Derivation of scaled surface reflectances from AVIRIS data, Remote Sens. Environ., 44(2–3), 165–178, 1993.

Gordon, I. E., Rothman, L. S., Hill, C., Kochanov, R. V, Tan, Y., Bernath, P. F., Birk, M., Boudon, V., Campargue, A., Chance, K. V and others: The HITRAN2016 molecular spectroscopic database, J. Quant. Spectrosc. Radiat. Transf., 203, 3–69, 2017.

Green, R. O., Eastwood, M. L., Sarture, C. M., Chrien, T. G., Aronsson, M., Chippendale, B. J., Faust, J. A., Pavri, B. E., Chovit, C. J., Solis,
10   M. and others: Imaging spectroscopy and the airborne visible/infrared imaging spectrometer (AVIRIS), Remote Sens. Environ., 65(3), 227–248, 1998.

Grimm, J. M. A.: Hawaii Volcanoes National Park February 2017 Spectra. Data set, 2017.

Gupta, P., Remer, L. A., Levy, R. C. and Mattoo, S.: Validation of MODIS 3 km land aerosol optical depth from NASA's EOS Terra and Aqua missions, Atmos. Meas. Tech., 11(5), 3145–3159, doi:10.5194/amt-11-3145-2018, 2018.

15   Hamlin, L., Green, R. O., Mouroulis, P., Eastwood, M., Wilson, D., Dudik, M. and Paine, C.: Imaging spectrometer science measurements for terrestrial ecology: AVIRIS and new developments, in Aerospace Conference, 2011 IEEE, pp. 1–7., 2011.

Haywood, J. and Boucher, O.: Estimates of the direct and indirect radiative forcing due to tropospheric aerosols: A review, Rev. Geophys., 38(4), 513–543, 2000.

Heidinger, A. K. and Stephens, G. L.: Molecular line absorption in a scattering atmosphere. Part II: Application to remote sensing in the O_2
20   A band, J. Atmos. Sci., 57(10), 1615–1634, 2000.

Herold, M., Roberts, D. A., Gardner, M. E. and Dennison, P. E.: Spectrometry for urban area remote sensing—Development and analysis of a spectral library from 350 to 2400 nm, Remote Sens. Environ., 91(3–4), 304–319, 2004a.

Herold, M., Roberts, D. A., Gardner, M. and Dennison, P. E.: Urban Reflectance Spectra from Santa Barbara, CA. Data set, 2004b.

Holben, B. N., Eck, T. F., Slutsker, I., Tanré, D., Buis, J. P., Setzer, A., Vermote, E., Reagan, J. A., Kaufman, Y. J., Nakajima, T., Lavenu,
25   F., Jankowiak, I. and Smirnov, A.: AERONET—A Federated Instrument Network and Data Archive for Aerosol Characterization, Remote Sens. Environ., 66(1), 1–16, doi:10.1016/S0034-4257(98)00031-5, 1998.

Hornik, K., Stinchcombe, M. and White, H.: Multilayer feedforward networks are universal approximators, Neural Networks, 2(5), 359–366, doi:10.1016/0893-6080(89)90020-8, 1989.

Hsu, N. C., Tsay, S.-C., King, M. D. and Herman, J. R.: Aerosol properties over bright-reflecting source regions, IEEE Trans. Geosci.
30   Remote Sens., 42(3), 557–569, 2004.

Hsu, N. C., Tsay, S.-C., King, M. D. and Herman, J. R.: Deep blue retrievals of Asian aerosol properties during ACE-Asia, IEEE Trans. Geosci. Remote Sens., 44(11), 3180–3195, 2006.

Japkowicz, N., Myers, C., Gluck, M. and others: A novelty detection approach to classification, in IJCAI, vol. 1, pp. 518–523., 1995.

Kaufman, Y. J., Wald, A. E., Remer, L. A., Gao, B.-C., Li, R.-R. and Flynn, L.: The MODIS 2.1-/spl mu/m channel-correlation with visible
35   reflectance for use in remote sensing of aerosol, IEEE Trans. Geosci. Remote Sens., 35(5), 1286–1298, 1997.

Kingma, D. P. and Ba, J.: Adam: A Method for Stochastic Optimization, arXiv Prepr. arXiv1412.6980, 1–15, doi:http://doi.acm.org.ezproxy.lib.ucf.edu/10.1145/1830483.1830503, 2014.

Kokhanovsky, A. A., Davis, A. B., Cairns, B., Dubovik, O., Hasekamp, O. P., Sano, I., Mukai, S., Rozanov, V. V, Litvinov, P., Lapyonok,

T. and others: Space-based remote sensing of atmospheric aerosols: The multi-angle spectro-polarimetric frontier, Earth-science Rev., 145, 85–116, doi:https://doi.org/10.1016/j.earscirev.2015.01.012, 2015.

Kramer, M. A.: Autoassociative neural networks, Comput. Chem. Eng., 16(4), 313–328, 1992.

5 Labate, D., Ceccherini, M., Cisbani, A., De Cosmo, V., Galeazzi, C., Giunti, L., Melozzi, M., Pieraccini, S. and Stagi, M.: The PRISMA payload optomechanical design, a high performance instrument for a new hyperspectral mission, Acta Astronaut., 65(9–10), 1429–1436, 2009.

Lary, D. J., Remer, L. A., MacNeill, D., Roscoe, B. and Paradise, S.: Machine learning and bias correction of MODIS aerosol optical depth, IEEE Geosci. Remote Sens. Lett., 6(4), 694–698, 2009.

10 Lee, C. M., Cable, M. L., Hook, S. J., Green, R. O., Ustin, S. L., Mandl, D. J. and Middleton, E. M.: An introduction to the NASA Hyperspectral InfraRed Imager (HyspIRI) mission and preparatory activities, Remote Sens. Environ., 167, 6–19, doi:https://doi.org/10.1016/j.rse.2015.06.012, 2015.

Levy, R., Hsu, C., Sayer, A., Mattoo, S. and Lee, J.: MODIS Atmosphere L2 Aerosol Product, NASA MODIS Adapt. Process. Syst., doi:doi:10.5067/MODIS/MOD04_L2.006; doi:10.5067/MODIS/MYD04_L2.006, 2015.

Levy, R. C., Remer, L. A., Martins, J. V, Kaufman, Y. J., Plana-Fattori, A., Redemann, J. and Wenny, B.: Evaluation of the MODIS aerosol
15 retrievals over ocean and land during CLAMS, J. Atmos. Sci., 62(4), 974–992, 2005.

Levy, R. C., Mattoo, S., Munchak, L. A., Remer, L. A., Sayer, A. M., Patadia, F. and Hsu, N. C.: The Collection 6 MODIS aerosol products over land and ocean, Atmos. Meas. Tech., 6(11), 2989, 2013.

Licciardi, G. A. and Del Frate, F.: Pixel unmixing in hyperspectral data by means of neural networks, IEEE Trans. Geosci. Remote Sens., 49(11 PART 1), 4163–4172, doi:10.1109/TGRS.2011.2160950, 2011.

20 Massie, S. T. and Hervig, M.: HITRAN 2012 refractive indices, J. Quant. Spectrosc. Radiat. Transf., 130, 373–380, 2013.

Meerdink, S. K., Hook, S. J., Grove, C. I., Abbott, E. A. and Roberts, D. A.: The ECOSTRESS Spectral Library version 1.0., Remote Sens. Environ., 230(111196), 2019.

Min, Q.-L., Harrison, L. C., Kiedron, P., Berndt, J. and Joseph, E.: A high-resolution oxygen A-band and water vapor band spectrometer, J. Geophys. Res. Atmos., 109(D2), 2004.

25 Minnis, P., Hong, G., Sun-Mack, S., Smith Jr, W. L., Chen, Y. and Miller, S. D.: Estimating nocturnal opaque ice cloud optical depth from MODIS multispectral infrared radiances using a neural network method, J. Geophys. Res. Atmos., 121(9), 4907–4932, 2016.

Mishchenko, M. I. and Travis, L. D.: Capabilities and limitations of a current FORTRAN implementation of the T-matrix method for randomly oriented, rotationally symmetric scatterers, J. Quant. Spectrosc. Radiat. Transf., 60(3), 309–324, 1998.

Morcrette, J.-J., Boucher, O., Jones, L., Salmond, D., Bechtold, P., Beljaars, A., Benedetti, A., Bonet, A., Kaiser, J. W., Razinger, M. and
30 others: Aerosol analysis and forecast in the European Centre for medium-range weather forecasts integrated forecast system: Forward modeling, J. Geophys. Res. Atmos., 114(D6), 2009.

Mouroulis, P., Green, R. O. and Chrien, T. G.: Design of pushbroom imaging spectrometers for optimum recovery of spectroscopic and spatial information, Appl. Opt., 39(13), 2210–2220, 2000.

Mouroulis, P., Hartley, F. T., Wilson, D. W., White, V. E., Shori, A., Nguyen, S., Zhang, M. and Feldman, M.: Blazed grating fabrication
35 through gray-scale X-ray lithography, Opt. Express, 11(3), 270–281, 2003.

Myhre, G., Shindell, D., Bréon, F.-M., Collins, W., Fuglestvedt, J., Huang, J., Koch, D., Lamarque, J.-F., Lee, D., Mendoza, B. and others: Anthropogenic and natural radiative forcing, Clim. Chang., 423, 658–740, 2013.

Nan, F. and A., C. S.: Measurement-based estimates of direct radiative effects of absorbing aerosols above clouds, J. Geophys. Res. Atmos., 120(14), 6908–6921, doi:10.1002/2015JD023252, 2015.

33

Di Noia, A., Hasekamp, O. P., Wu, L., van Diedenhoven, B., Cairns, B. and Yorks, J. E.: Combined neural network/Phillips--Tikhonov approach to aerosol retrievals over land from the NASA Research Scanning Polarimeter, Atmos. Meas. Tech., 10(11), 4235–4252, doi:10.5194/amt-10-4235-2017, 2017.

Okada, K., Heintzenberg, J., Kai, K. and Qin, Y.: Shape of atmospheric mineral particles collected in three Chinese arid-regions, Geophys. Res. Lett., 28(16), 3123–3126, 2001.

Palsson, B., Sigurdsson, J., Sveinsson, J. R. and Ulfarsson, M. O.: Hyperspectral Unmixing Using a Neural Network Autoencoder, IEEE Access, 6, 25646–25656, 2018.

Pincus, R. and Baker, M. B.: Effect of precipitation on the albedo susceptibility of clouds in the marine boundary layer, Nature, 372(6503), 250, 1994.

Pope III, C. A., Ezzati, M. and Dockery, D. W.: Fine-particulate air pollution and life expectancy in the United States, N. Engl. J. Med., 360(4), 376–386, 2009.

Prasad, A. K. and Singh, R. P.: Comparison of MISR-MODIS aerosol optical depth over the Indo-Gangetic basin during the winter and summer seasons (2000-2005), Remote Sens. Environ., 107(1–2), 109–119, 2007.

Radosavljevic, V., Vucetic, S. and Obradovic, Z.: A data-mining technique for aerosol retrieval across multiple accuracy measures, IEEE Geosci. Remote Sens. Lett., 7(2), 411–415, 2010.

Remer, L. A., Kaufman, Y. J., Tanré, D., Mattoo, S., Chu, D. A., Martins, J. V., Li, R.-R., Ichoku, C., Levy, R. C., Kleidman, R. G. and others: The MODIS aerosol algorithm, products, and validation, J. Atmos. Sci., 62(4), 947–973, 2005.

Rencz, A. N. and Ryerson, R. A.: Manual of remote sensing, remote sensing for the earth sciences, John Wiley & Sons Incorporated., 1999.

Salomonson, V. V, Barnes, W. L., Maymon, P. W., Montgomery, H. E. and Ostrow, H.: MODIS: Advanced facility instrument for studies of the Earth as a system, IEEE Trans. Geosci. Remote Sens., 27(2), 145–153, 1989.

Stamnes, K., Tsay, S.-C., Wiscombe, W. and Jayaweera, K.: Numerically stable algorithm for discrete-ordinate-method radiative transfer in multiple scattering and emitting layered media, Appl. Opt., 27(12), 2502, doi:10.1364/AO.27.002502, 1988.

Taravat, A., Del Frate, F., Cornaro, C. and Vergari, S.: Neural networks and support vector machine algorithms for automatic cloud classification of whole-sky ground-based images, IEEE Geosci. Remote Sens. Lett., 12(3), 666–670, 2015.

Tennant, W. E., Lee, D., Zandian, M., Piquette, E. and Carmody, M.: MBE HgCdTe Technology: A very general solution to IR Detection, Described by "Rule 07", a very convenient Heuristic, J. Electron. Mater., 37(9), 1406–1410, 2008.

Thompson, D. R., Gao, B. C., Green, R. O., Roberts, D. A., Dennison, P. E. and Lundeen, S. R.: Atmospheric correction for global mapping spectroscopy: ATREM advances for the HyspIRI preparatory campaign, Remote Sens. Environ., 167, 64–77, doi:10.1016/j.rse.2015.02.010, 2015.

Thompson, D. R., Natraj, V., Green, R. O., Helmlinger, M. C., Gao, B.-C. and Eastwood, M. L.: Optimal estimation for imaging spectrometer atmospheric correction, Remote Sens. Environ., 216, 355–373, 2018.

Twomey, S.: The influence of pollution on the shortwave albedo of clouds, J. Atmos. Sci., 34(7), 1149–1152, 1977.

Verger, A., Baret, F. and Camacho, F.: Optimal modalities for radiative transfer-neural network estimation of canopy biophysical characteristics: Evaluation over an agricultural area with CHRIS/PROBA observations, Remote Sens. Environ., 115(2), 415–426, doi:10.1016/j.rse.2010.09.012, 2011.

Verhoef, W. and Bach, H.: Simulation of hyperspectral and directional radiance images using coupled biophysical and atmospheric radiative transfer models, Remote Sens. Environ., 87(1), 23–41, 2003.

Wang, Y., Yuan, Q., Li, T., Shen, H., Zheng, L. and Zhang, L.: Evaluation and comparison of MODIS Collection 6.1 aerosol optical depth against AERONET over regions in China with multifarious underlying surfaces, Atmos. Environ., 200, 280–301,

doi:https://doi.org/10.1016/j.atmosenv.2018.12.023, 2019.

Werbos, P.: Beyond Regression: New Tools for Prediction and Analysis in the Behavioral Sciences, Ph. D. Diss. Harvard Univ., 1974.

Wielicki, B. A., Young, D. F., Mlynczak, M. G., Thome, K. J., Leroy, S., Corliss, J., Anderson, J. G., Ao, C. O., Bantges, R., Best, F., Bowman, K., Brindley, H., Butler, J. J., Collins, W., Dykema, J. A., Doelling, D. R., Feldman, D. R., Fox, N., Huang, X., Holz, R., Huang, Y., Jin, Z., Jennings, D., Johnson, D. G., Jucks, K., Kato, S., Kirk-Davidoff, D. B., Knuteson, R., Kopp, G., Kratz, D. P., Liu, X., Lukashin, C., Mannucci, A. J., Phojanamongkolkij, N., Pilewskie, P., Ramaswamy, V., Revercomb, H., Rice, J., Roberts, Y., Roithmayr, C. M., Rose, F., Sandford, S., Shirley, E. L., Smith, W. L., Soden, B., Speth, P. W., Sun, W., Taylor, P. C., Tobin, D. and Xiong, X.: Achieving Climate Change Absolute Accuracy in Orbit, Bull. Am. Meteorol. Soc., 94(10), 1519–1539, doi:10.1175/BAMS-D-12-00149.1, 2013.

Wold, S., Esbensen, K. and Geladi, P.: Principal component analysis, Chemom. Intell. Lab. Syst., 2(1–3), 37–52, doi:10.1016/0169-7439(87)80084-9, 1987.

World Health Organisation: Burden of disease from the joint effects of household and ambient Air pollution for 2016, 2018.

World Health Organization: Air quality guidelines: global update 2005, World Health Organization., 2006.

Xiao, Z., Liang, S., Wang, J., Chen, P., Yin, X., Zhang, L. and Song, J.: Use of general regression neural networks for generating the GLASS leaf area index product from time-series MODIS surface reflectance, IEEE Trans. Geosci. Remote Sens., 52(1), 209–223, 2014.