



# Combined neural network/Phillips–Tikhonov approach to aerosol retrievals over land from the NASA Research Scanning Polarimeter

Antonio Di Noia<sup>1</sup>, Otto P. Hasekamp<sup>1</sup>, Lianghai Wu<sup>1</sup>, Bastiaan van Diedenhoven<sup>2,3</sup>, Brian Cairns<sup>3</sup>, and John E. Yorks<sup>4</sup>

<sup>1</sup>SRON Netherlands Institute for Space Research, Sorbonnelaan 2, 3584CA Utrecht, the Netherlands

<sup>2</sup>Columbia University, Center for Climate Systems Research, 2910 Broadway, New York, NY 10025, USA

<sup>3</sup>NASA Goddard Institute for Space Studies, 2880 Broadway, New York, NY 10025, USA

<sup>4</sup>NASA Goddard Space Flight Center, 8800 Greenbelt Rd, Greenbelt, MD 20771, USA

Correspondence to: Antonio Di Noia (a.di.noia@sron.nl)

Received: 8 May 2017 – Discussion started: 26 June 2017

Revised: 18 September 2017 – Accepted: 20 September 2017 – Published: 8 November 2017

**Abstract.** In this paper, an algorithm for the retrieval of aerosol and land surface properties from airborne spectropolarimetric measurements – combining neural networks and an iterative scheme based on Phillips–Tikhonov regularization – is described. The algorithm – which is an extension of a scheme previously designed for ground-based retrievals – is applied to measurements from the Research Scanning Polarimeter (RSP) on board the NASA ER-2 aircraft. A neural network, trained on a large data set of synthetic measurements, is applied to perform aerosol retrievals from real RSP data, and the neural network retrievals are subsequently used as a first guess for the Phillips–Tikhonov retrieval. The resulting algorithm appears capable of accurately retrieving aerosol optical thickness, fine-mode effective radius and aerosol layer height from RSP data. Among the advantages of using a neural network as initial guess for an iterative algorithm are a decrease in processing time and an increase in the number of converging retrievals.

## 1 Introduction

Multi-angular, multispectral measurements of intensity and linear polarization parameters of scattered solar radiation are a useful tool for the characterization of atmospheric aerosols (Mishchenko and Travis, 1997; Hasekamp and Landgraf, 2007). The recognition of this has led to the development of a number of remote sensing instruments with spectropolarimetric capability (Kokhanovsky et al., 2015). The Polarization and Directionality of the Earth's Reflectance (POLDER)

1, 2 and 3 satellite instruments (Deschamps et al., 1994), mounted on board the Japanese satellite Advanced Earth Observing Satellite (ADEOS) and on the French satellite Polarization & Anisotropy of Reflectances for Atmospheric Sciences coupled with Observations from a Lidar (PARASOL), have so far been the only instruments to perform in-orbit multi-angle spectropolarimetric measurements. Decommissioned in 2013, POLDER-3 will be followed by the Multi-viewing Multi-channel Multi-polarization Imaging (3MI) instrument (Marbach et al., 2013), expected for launch on board the EUMETSAT MetOp-SG satellite in 2021. While no satellite multi-angle spectropolarimeters are currently operating, research is being carried out in the development of innovative spectropolarimetric instruments, which are currently operated on aircrafts, in view of future satellite missions. These include the Research Scanning Polarimeter (RSP, Cairns et al., 1999), the Airborne Multi-angle SpectroPolarimetric Imager (AirMSPI, Diner et al., 2013) and the Dutch instrument SPEX (van Amerongen et al., 2016). Additional spectropolarimetric missions under development are the CubeSat HyperAngular Rainbow Polarimeter (HARP) mission (Martins et al., 2014) and the Ukrainian project Aerosol-UA (Milinevsky et al., 2016).

The aerosol properties are retrieved from multi-angle spectropolarimetric measurements by comparing these measurements to forward model simulations. Retrieval methods developed so far include look-up tables (LUTs, Deuzé et al., 2000, 2001) and iterative methods based on accurate forward modelling (Waquet et al., 2009; Dubovik et al., 2011; Hasekamp et al., 2011; Knobelspiesse et al., 2011; Wu et al.,

2015). In most of these methods a LUT is used to generate a first guess for the retrieved quantities, and an iterative algorithm (e.g. maximum a posteriori, Phillips–Tikhonov regularization) is then used in order to generate solutions that fit the measurements better than the first guess.

In an experiment performed on ground-based spectropolarimetric measurements, it has been shown (Di Noia et al., 2015) that the final result of an iterative aerosol retrieval may depend on the choice of the first guess, and that replacing a LUT-based first guess with a neural network algorithm may be beneficial for the algorithm convergence and computation time, as it is relatively simple to design neural networks that provide quicker and more accurate first guess retrievals than reasonably sized LUTs with modest computational effort. Extending this approach to aircraft and satellite measurements is possible, in principle, once a method is devised for taking the variability of the observation geometry into account when a training set for the neural network is generated. This problem is less important – possibly absent – when working with ground-based observations. In such a case, in fact, care may be taken so that a spectropolarimeter is operated only in the principal plane and at a predefined set of viewing angles. In this way, the solar zenith angle is the only variable determining the measurement geometry, and this can be easily taken into account in the neural network design process. When dealing with an airborne or satellite instrument, it is not possible to assume that the observation geometry will always be the same. This is especially true for instruments that have a swath (e.g. POLDER, AirMSPi), in which each pixel is observed at a different set of viewing zenith and azimuth angles, whereas instruments that do not possess a swath, such as the RSP, represent a situation of intermediate complexity. While it is safe to assume that an RSP measurement is always made at the same set of viewing zenith angles, the azimuth angles at which the scene is observed will be variable, and this needs to be taken into account when designing a neural network algorithm for processing this type of measurement.

In this paper we show an application of neural network retrievals to RSP measurements. We trained a neural network on a large set of simulated RSP measurements generated for a large number of combinations of aerosol and surface parameters and observation geometries. We then used the neural network retrievals as a first guess for an iterative algorithm based on the Phillips–Tikhonov method, described in Wu et al. (2015), to perform aerosol retrievals from RSP measurements acquired during the Polarimeter Definition Experiment (PODEX) and Studies of Emission and Atmospheric Composition, Clouds and Climate Coupling by Regional Surveys (SEAC<sup>4</sup>RS) measurement campaigns.

This paper is structured as follows. In Sect. 2 a brief description of the RSP instrument is given. In Sect. 3 some theoretical background is given on the use of neural networks as parameter estimation methods. Section 4 describes the design of the neural network algorithm for RSP and its valida-

tion on synthetic data. Section 5 describes the application of the Phillips–Tikhonov algorithm with neural-network-based initialization to a sample of synthetic data. In Sect. 6 the results of the application of the algorithm to real RSP measurements are discussed. In particular, the results of comparisons between RSP retrievals and correlative data from the Aerosol Robotic Network (AERONET, Holben et al., 1998) and from the Cloud Physics Lidar (CPL, McGill et al., 2002) are presented. Finally, in Sect. 7 conclusions are drawn.

## 2 The NASA Research Scanning Polarimeter

The Research Scanning Polarimeter (RSP) is an airborne multi-angle spectropolarimeter initially designed as a prototype for the Aerosol Polarimetry Sensor (APS) to be launched on the Glory satellite mission (Mishchenko et al., 2007) in 2011, aiming to observe aerosols and clouds with unprecedented accuracy by measuring intensity and linear polarization parameters of backscattered sunlight at a very high number of viewing angles (approximately 250) and in nine spectral bands (from 410 to 2250 nm). While the APS instrument has ultimately not reached its orbit as a consequence of the failure of the Glory satellite launch, the RSP instrument is still being operated on board the NASA ER-2 high-altitude aircraft. RSP performs its measurements at 152 viewing angles, at the same nine spectral bands as APS (410, 470, 550, 670, 865, 960, 1590, 1880 and 2250 nm). The bands from 410 to 875 nm are particularly sensitive to aerosols. The 960 and 1880 nm channels are used to characterize water vapour and cirrus clouds respectively, whereas the 1590 and 2250 nm channels can be used to characterize land surface and coarse-mode aerosols (Wu et al., 2015). RSP can cover an angular range from  $-60$  to  $60^\circ$  with respect to the aircraft vertical, where the minus sign indicates viewing directions pointed aftward with respect to the aircraft. However, because of the position of the instrument with respect to the aircraft, parts of the forward viewing directions from  $40$  to  $60^\circ$  are blocked by the aircraft body. The ground pixel size of an RSP measurement at nadir is 277 m. The radiometric uncertainty of RSP measurements is approximately 2% and its polarimetric uncertainty is around 0.5% (Cairns et al., 1999).

## 3 Neural network retrievals and their relationship to conventional retrievals

The retrieval of a vector of geophysical parameters  $\mathbf{x}$  (state vector) from a vector of measurements  $\mathbf{y}$  relies on the assumption that the measurements are related to the state vector by means of a forward model  $\mathbf{y} = \mathbf{F}(\mathbf{x}, \mathbf{b}) + \boldsymbol{\epsilon}$ , where  $\mathbf{F}$  is a vector-valued deterministic function,  $\mathbf{b}$  is a vector of parameters of  $\mathbf{F}$  that are not included in  $\mathbf{x}$  (e.g. observation angles, non-retrieved meteorological variables, etc.), and  $\boldsymbol{\epsilon}$  is the measurement noise, which is often treated as a multivari-

ate Gaussian variable with zero mean and covariance matrix  $S_\epsilon$ .

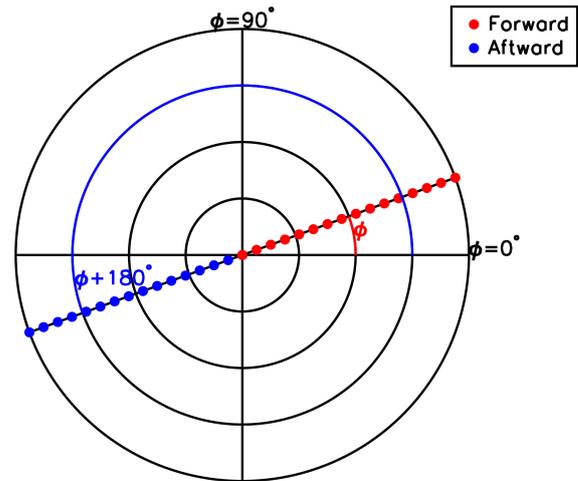
The conventional way of retrieving  $x$  from  $y$  consists of finding the value of  $x$  that yields the best agreement between  $y$  and  $F(x, b)$ . To this aim, a cost function depending on the discrepancy between  $y$  and  $F(x, b)$  is defined and minimized with respect to  $x$ . Because of the presence of measurement noise, the cost function is usually minimized in the least-square sense. Furthermore, given the non-uniqueness of the solution due to the fact that  $F$  is – in most cases – not a biunivocal relationship, the cost function may contain some regularization terms chosen in such a way to penalize solutions that are unphysical or unlikely (Rodgers, 2000). In general, the solution to a retrieval problem can be chosen as

$$\hat{x} = \arg \min_x J(x, y, b, c), \quad (1)$$

where  $J$  is the chosen cost function and  $c$  is a vector of parameters of the cost function (e.g. regularization parameters, regularization matrix, a priori error, etc.).

An alternative approach to the retrieval is based on the fact that, because of noise and non-uniqueness, an inverse relationship between  $x$  and  $y$  can be defined only in a statistical sense, and the conditional expectation  $E[x|y, b]$  can be used as an estimator for  $x$ . This conditional expectation is a function of  $y$  and  $b$ , but its functional form is not known in advance. However, under the reasonable assumption that such a function is smooth, it can be estimated by collecting a number of coincidences between  $x$ ,  $y$  and  $b$  (e.g. by multiple runs of a forward model) and using them to fit some smooth function  $g(y, b, w)$ , where  $w$  is a set of parameters of the chosen function  $g$  that must be determined during the fitting phase. If it is reasonable to assume that the sought relationship is linear, the problem simply reduces to a linear regression of  $x$  on  $y$  and  $b$ , an approach that is quite common – for instance – in atmospheric profile retrievals from infrared and microwave measurements (Smith et al., 1970; Jackson et al., 2006). If, instead, a linear relationship between  $E[x|y, b]$  and the other quantities cannot be assumed, a more general form for  $g$  must be chosen. For this purpose, neural network models are usually a good choice, for at least two reasons: (i) neural network functions with at least a non-linear hidden layer can approximate any continuous function on a compact set to an arbitrary accuracy (Hornik et al., 1989; Leshno et al., 1993); (ii) as the dimension of the training set tends to infinity, a neural network trained with the sum of squares error cost function tends to approximate the conditional expectation of the target quantity given the input vector, provided that the training samples are independent and identically distributed (Bishop, 1995a).

The latter concept can be used as a guiding principle when deciding which quantities should be used as inputs to a neural network method aimed at retrieving a certain set of geophysical parameters. These quantities do not only include the measurement vector  $y$ , but they also include the known pa-



**Figure 1.** Idealized viewing geometry assumed to model the angular dependence of RSP measurements. Each dot in the polar plot represents a single angular measurement. Forward-viewing measurements are represented in red, aftward-viewing measurements in blue.

rameters  $b$  affecting the value of  $y$ . For instance, in the case of a spectropolarimetric retrieval,  $y$  is a vector of reflectances and degrees of linear polarization (or eventually polarized reflectances, or Stokes parameters) measured at a number of angles and wavelengths. The value of each measurement is not only affected by the geophysical parameters we are trying to retrieve but is also affected by quantities such as the solar zenith angle, the viewing zenith angle and the relative azimuth angle (or, equivalently, the scattering angle). When the goal is to design a neural network capable of working for several combinations of these parameters, they should in principle be used as input quantities for the network. Formally, in this way the training set will be a sample drawn from the joint probability distribution of  $y$  and  $b$ , and the neural network retrieval will actually be a non-linear regression of the state vector  $x$  on  $y$  and  $b$ .

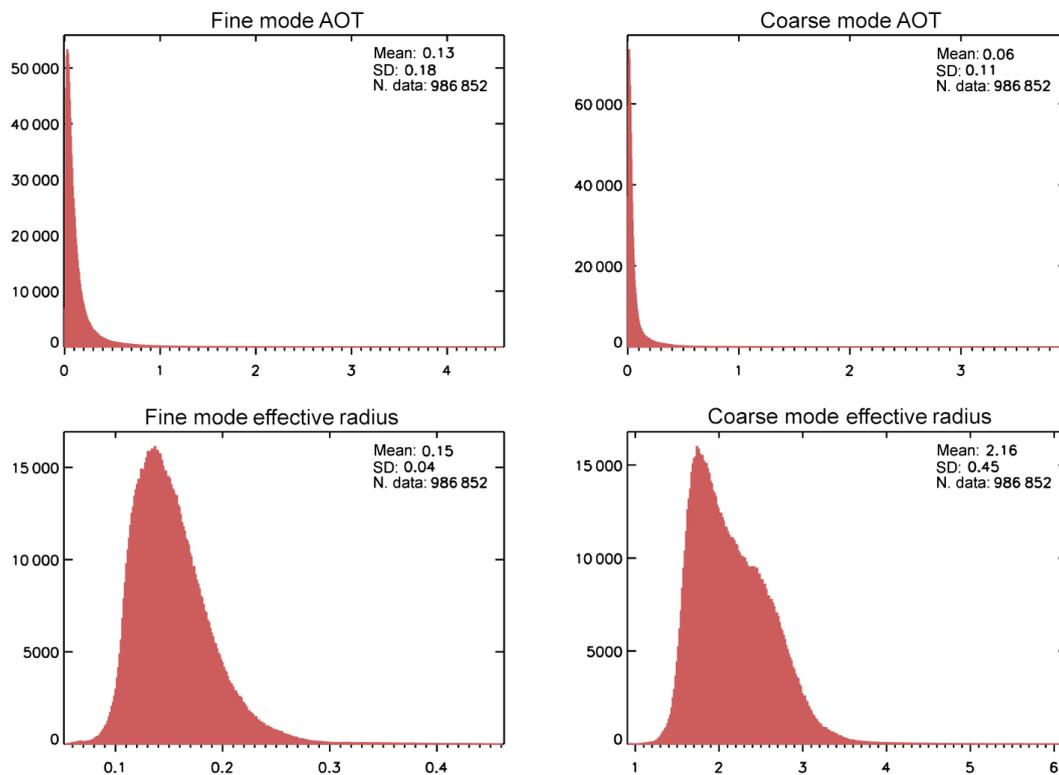
Several examples of application of neural networks in parameter retrieval algorithms from remotely sensed measurements exist (Krasnopolsky, 2007). The advantages of neural networks over traditional linear regression in statistical retrievals are discussed in Del Frate and Schiavon (1998a, b). Applications to aerosol retrievals are presented, for instance, in Radosavljevic et al. (2010), Taylor et al. (2014) and Chiotot et al. (2017).

#### 4 Neural network retrieval scheme for RSP: design and performance on synthetic data

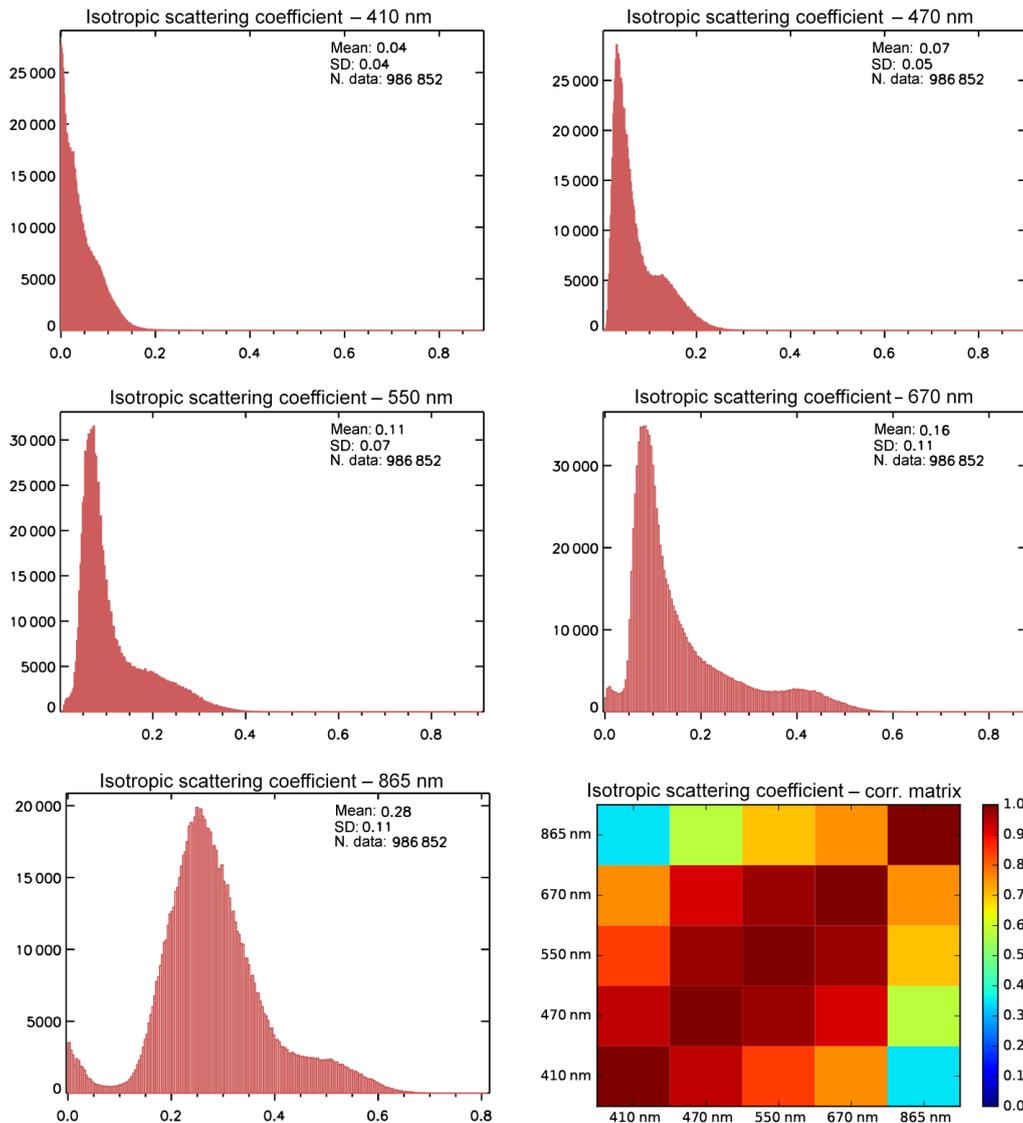
The procedure followed in the neural network retrieval scheme builds further on the work described in Di Noia et al. (2015) for groundSPEX measurements. The main differences are that in the algorithm presented in this paper sur-

**Table 1.** Details of the statistical distributions of the aerosol and surface parameters used to generate the training data set. Please refer to Figs. 2 and 3 for the histograms of the distributions indicated as “empirical”.

Parameter	Min	Max	Mean	SD	Distribution
Effective radius ( $\mu\text{m}$ ) – fine	0.05	0.46	0.15	0.04	Empirical
Effective variance – fine	0.1	0.3	0.2	0.06	Uniform
Refractive index (real) – fine	1.3	1.7	1.49	0.10	Uniform
Log. refractive index (im.) – fine	−11.5	−0.5	−5.75	3.03	Uniform
AOT (550 nm) – fine	0.0	4.58	0.13	0.18	Empirical
Effective radius ( $\mu\text{m}$ ) – coarse	0.92	6.12	2.16	0.45	Empirical
Effective variance – coarse	0.4	0.6	0.5	0.06	Uniform
Refractive index (real) – coarse	1.3	1.7	1.49	0.1	Uniform
Log. refractive index (im.) – coarse	−11.5	−0.5	−5.75	3.03	Uniform
AOT (550 nm) – coarse	0.0	3.95	0.06	0.11	Empirical
Spherical fraction – coarse	0.0	1.0	0.5	0.29	Uniform
Aerosol layer height (m)	250.0	8000.0	4125.0	2237.84	Uniform
Li-sparse BRDF parameter	0.0	0.25	0.12	0.07	Uniform
Ross-thick BRDF parameter	0.0	1.5	0.75	0.43	Uniform
Maignan BPDF parameter	0.02	10.0	4.83	2.86	Uniform
Isotropic scattering coefficient (410 nm)	0.0	0.89	0.04	0.04	Empirical
Isotropic scattering coefficient (470 nm)	0.0	0.90	0.07	0.05	Empirical
Isotropic scattering coefficient (550 nm)	0.0	0.90	0.11	0.07	Empirical
Isotropic scattering coefficient (670 nm)	0.0	0.87	0.16	0.11	Empirical
Isotropic scattering coefficient (865 nm)	0.0	0.80	0.28	0.11	Empirical



**Figure 2.** Histograms of the statistical distribution of fine- and coarse-mode AOT and effective radius used to generate the training data set.

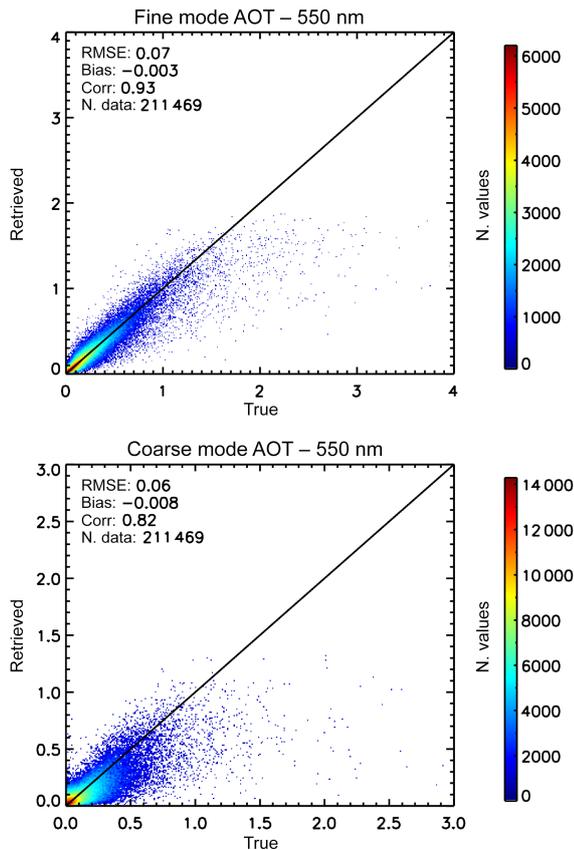


**Figure 3.** Histograms of the statistical distribution of the  $f_{\text{iso}}$  coefficient in the Ross–Li model at all the wavelengths used to generate the training data set and the correlation matrix between the values of the coefficient at the used wavelengths.

face properties and additional aerosol parameters (fine- and coarse-mode effective variance, fraction of spherical particles for coarse mode, aerosol layer height) have been included in the retrieval state vector, and the variability of RSP observing conditions has been taken into account in the input vector.

As mentioned in Sect. 2, the viewing zenith angle (VZA) range of RSP measurements goes from  $60^\circ$  aftward to  $40^\circ$  forward. This asymmetry is due to the blockage of part of the forward viewing directions by the ER-2 aircraft structure. In order to simplify the generation of a training set for the neural network scheme, we decided to use a symmetrical VZA range, going from  $40^\circ$  aftward to  $40^\circ$  forward. The remaining part of the VZA range is used, instead, in the iterative retrieval that uses the neural network output as a first guess,

as explained in Sect. 6. As for the relative azimuth angles of RSP measurements we have assumed that, for each multi-angular observation, the measurements at forward-viewing angles are made with a common relative azimuth angle  $\varphi$  and the corresponding aftward-viewing measurements are made at an angle  $\varphi + 180^\circ$  (see Fig. 1), with  $\varphi$  depending on the particular flight leg. We observed from real data that this is a good approximation for most RSP measurements. As a result of this assumption, the scalar value  $\varphi$  is the only variable we used in order to characterize the RSP viewing geometry in the neural network input vector. In order to avoid the issue of characterizing water vapour absorption in the training set, we only used the first five RSP channels (410, 470, 550, 670, 865 nm). We assumed the Ross-thick – Li-sparse reciprocal combination, hereinafter referred to as the “Ross–Li model”



**Figure 4.** Neural network retrieved versus true fine- and coarse-mode AOT from synthetic test data.

(Maignan et al., 2004, and references therein), for the surface bidirectional reflectance distribution function (BRDF) and the model by Maignan et al. (2009) for the bidirectional polarization distribution function (BPDF).

In the Ross–Li model, the surface BRDF  $R$  is expressed by means of the following linear combination (Lucht et al., 2000):

$$R(\theta_s, \theta_v, \varphi, \lambda) = f_{\text{iso}}(\lambda) + f_{\text{vol}}(\lambda)K_{\text{vol}}(\theta_s, \theta_v, \varphi) + f_{\text{geo}}(\lambda)K_{\text{geo}}(\theta_s, \theta_v, \varphi), \quad (2)$$

where  $\lambda$  is the wavelength,  $\theta_s$  and  $\theta_v$  are the solar and viewing zenith angles respectively, and  $\varphi$  is the relative azimuth angle. In Eq. (2),  $K_{\text{vol}}$  and  $K_{\text{geo}}$  are the Ross-thick and Li-sparse kernels respectively. These are functions of the viewing and solar geometry describing the angular dependence of the surface reflectance and are controlled by the leaf area index (LAI) of the underlying vegetation and by the surface roughness respectively (Roujean et al., 1992; Wanner et al., 1995). The wavelength-dependent coefficients  $f_{\text{vol}}$  and  $f_{\text{geo}}$  determine the relative weight of the two kernels, whereas the coefficient  $f_{\text{iso}}$  is an additional term describing isotropic scattering, i.e. scattering with a directionally constant amplitude.

In the Maignan model, the polarized reflectance  $R_p$  of a surface is expressed by the equation

$$R_p(\theta_s, \theta_v, \varphi) = \frac{C e^{-\nu} e^{-\tan(\Theta/2)} F_p(\frac{\Theta}{2}, n)}{4(\cos \theta_s + \cos \theta_v)}, \quad (3)$$

where  $\Theta$  is the scattering angle,  $F_p$  is a function of the scattering angle and of the refractive index  $n$  (Maignan et al., 2009), and  $C$  and  $\nu$  are two fitting parameters.

The neural network retrieves 20 quantities: 12 describing aerosols and 8 describing surface properties. The retrieved aerosol parameters are effective radius, effective variance, complex refractive index (assumed wavelength-independent) and aerosol optical thickness (AOT) at 550 nm for fine and coarse mode; fraction of spherical particles (FSP) for the coarse mode; and aerosol layer height, defined as the peak height of an assumed Gaussian aerosol profile. The fine-mode FSP and the full width at half maximum (FWHM) of the aerosol profile are not retrieved but are kept fixed at 100 % and 2 km respectively. The retrieved surface parameters are the  $f_{\text{iso}}$  coefficient of Eq. (2) at each of the five input wavelengths, the Ross-thick and Li-sparse coefficients  $f_{\text{vol}}$  and  $f_{\text{geo}}$  (assumed wavelength-independent), and the  $C$  parameter of the Maignan et al. (2009) BPDF model described in Eq. (3), whereas the  $\nu$  parameter of the Maignan model has been set to 0.1. The input quantities of the neural network are the reflectance and the degree of linear polarization (DoLP) measured at 33 angles (from 0 to 40° with a step of 2.5° forward and aftward) and at the 5 wavelengths mentioned above, plus the solar zenith angle, the relative azimuth angle  $\varphi$  described above and the surface pressure. The radiance and DoLP measurements were compressed using a principal component analysis (PCA) as explained in Di Noia et al. (2015). Twenty-five principal components were retained for radiance and 33 for DoLP.

Approximately one million data have been used to train the neural network. The data have been created by generating random combinations of the 20 retrieved quantities, solar zenith angle, relative azimuth angle and surface pressure and using these combinations as inputs for radiative transfer simulations, using a radiative transfer model described in Hasekamp and Landgraf (2002, 2005). During the training process, the synthetic measurements have been perturbed with additive Gaussian noise with zero mean and standard deviation of 2 % for reflectance and 0.2 % for DoLP. The angular correlation of noise in RSP measurements (Knobelspiesse et al., 2012) has not been taken into account.

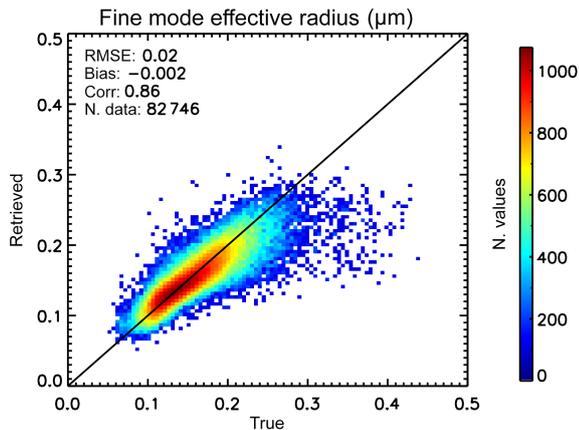
In order to generate statistical distributions that are as realistic as possible for most of the quantities to be retrieved, we have generated global data sets of fine- and coarse-mode AOT and effective radius by collecting data from all the AERONET stations located over land, and we have randomly sampled the MODIS MCD43C3 product (Schaaf and Wang, 2015) in order to generate a global data set for the  $f_{\text{iso}}$  coefficient in Eq. (2). For all the other parameters we have assumed

**Table 2.** Error statistics on the aerosol and surface parameters retrieved by the neural network, computed on a set of data not used during the training phase.

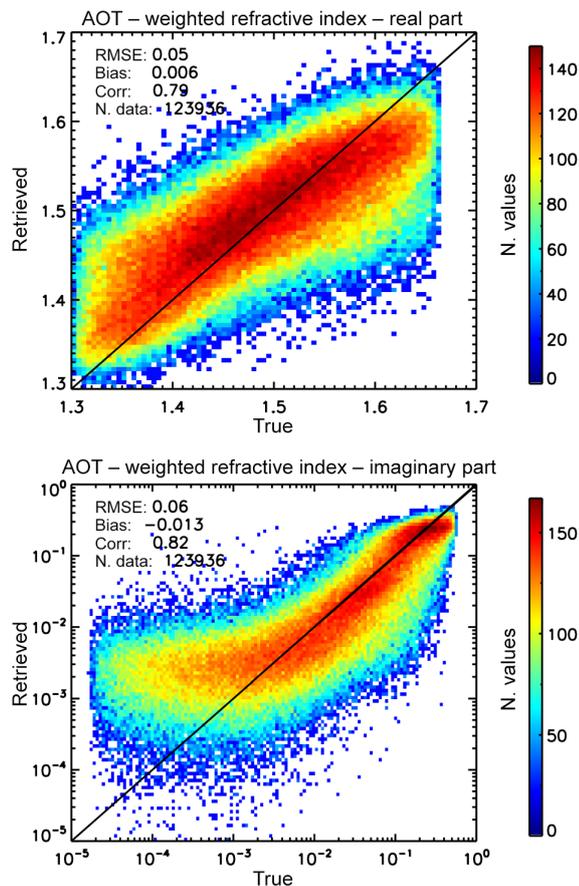
Parameter	Bias	RMSE	MAE	Corr.
Effective radius ( $\mu\text{m}$ ) – fine	−0.002	0.02	0.013	0.86
Effective variance – fine	−0.005	0.04	0.036	0.63
Refractive index (real) – fine	0.006	0.06	0.043	0.84
Refractive index (im.) – fine	−0.006	0.05	0.023	0.89
AOT (550 nm) – fine	−0.003	0.07	0.029	0.93
Effective radius ( $\mu\text{m}$ ) – coarse	−0.043	0.45	0.364	0.13
Effective variance – coarse	−0.001	0.06	0.050	0.02
Refractive index (real) – coarse	0.009	0.08	0.061	0.71
Refractive index (im.) – coarse	−0.022	0.09	0.038	0.63
AOT (550 nm) – coarse	−0.008	0.06	0.030	0.82
Spherical fraction – coarse	−0.016	0.25	0.207	0.50
Aerosol layer height (m)	42.424	1225.19	963.06	0.84
Li-sparse BRDF parameter	−0.002	0.02	0.015	0.95
Ross-thick BRDF parameter	−0.008	0.13	0.079	0.96
Maignan BPDF parameter	−0.086	0.79	0.46	0.96
Isotropic scattering coefficient (410 nm)	−0.001	0.01	0.006	0.96
Isotropic scattering coefficient (470 nm)	−0.001	0.01	0.006	0.98
Isotropic scattering coefficient (550 nm)	−0.001	0.02	0.009	0.98
Isotropic scattering coefficient (670 nm)	−0.001	0.02	0.010	0.99
Isotropic scattering coefficient (865 nm)	−0.001	0.02	0.014	0.98

**Table 3.** Rms error and correlation coefficient between true and retrieved parameters for the neural network retrieval alone (NN) and for the combined NN–Phillips–Tikhonov algorithm (NN–PT).

Parameter	NN		NN–PT	
	RMSE	Corr.	RMSE	Corr.
Effective radius ( $\mu\text{m}$ ) – fine	0.018	0.86	0.013	0.93
Effective variance – fine	0.047	0.59	0.044	0.67
Refractive index (real) – fine	0.059	0.83	0.050	0.88
Refractive index (im.) – fine	0.054	0.89	0.053	0.90
AOT (550 nm) – fine	0.062	0.94	0.032	0.98
Effective radius ( $\mu\text{m}$ ) – coarse	0.440	0.07	0.447	0.42
Effective variance – coarse	0.056	0.09	0.091	0.14
Refractive index (real) – coarse	0.083	0.69	0.074	0.76
Refractive index (im.) – coarse	0.093	0.66	0.091	0.67
AOT (550 nm) – coarse	0.057	0.82	0.039	0.92
Spherical fraction – coarse	0.254	0.46	0.230	0.62
Aerosol layer height (m)	1217.82	0.84	1135.06	0.86
SSA (410 nm)	0.054	0.94	0.034	0.97
Li-sparse BRDF parameter	0.027	0.93	0.021	0.96
Ross-thick BRDF parameter	0.137	0.95	0.101	0.97
Maignan BPDF parameter	0.843	0.96	0.597	0.98
Isotropic scattering coefficient (410 nm)	0.010	0.96	0.005	0.99
Isotropic scattering coefficient (470 nm)	0.010	0.98	0.006	0.99
Isotropic scattering coefficient (550 nm)	0.014	0.98	0.008	0.99
Isotropic scattering coefficient (670 nm)	0.017	0.99	0.011	0.99
Isotropic scattering coefficient (865 nm)	0.022	0.99	0.016	0.99



**Figure 5.** Neural network retrieved versus true fine-mode effective radius from synthetic test data. Only cases with fine-mode AOT larger than 0.1 are shown.



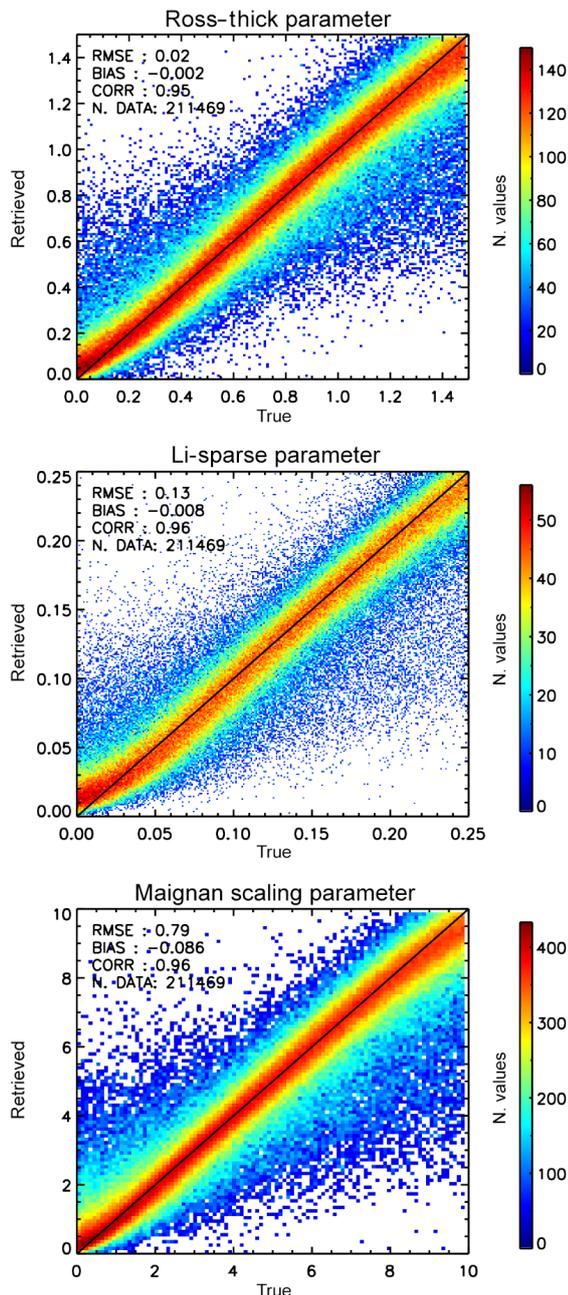
**Figure 6.** Neural network retrieved versus true real and imaginary part of the refractive index, weighted by mode AOT. Synthetic data. Only cases with total AOT larger than 0.1 are shown.

a uniform statistical distribution. The main features (maximum, minimum, mean and standard deviation and type of distribution) of the statistical distributions of all the aerosol

**Table 4.** List of the AERONET stations used for validation in this study.

Station name	Lat.	Lon.
Baskin	32.28	-91.74
Bozeman	45.66	-111.04
Caldwell Parish HS	32.06	-92.10
CART Site	36.61	-97.49
DRAGON Aldine	29.90	-95.33
DRAGON Arvin	35.24	-118.79
DRAGON Bakersfield	35.33	-119.00
DRAGON Channel View	29.80	-95.13
DRAGON Clinton	29.73	-95.26
DRAGON Clovis	36.82	-119.72
DRAGON Conroe	30.35	-95.43
DRAGON Corcoran	36.10	-119.57
DRAGON Deer Park	29.67	-95.13
DRAGON Drummond	36.71	-119.74
DRAGON Garland	36.78	-119.77
DRAGON Hanford	36.32	-119.64
DRAGON Huron	36.21	-120.11
DRAGON Madera City	36.95	-120.03
DRAGON ManvelCroix	29.52	-95.39
DRAGON NW Harris Co.	30.04	-95.67
DRAGON Parlier	36.60	-119.50
DRAGON Porterville	36.03	-119.06
DRAGON Shafter	35.50	-119.27
DRAGON Tranquility	36.63	-120.38
DRAGON UH W Liberty	30.06	-94.98
DRAGON Visalia	36.31	-119.39
DRAGON West Houston	29.83	-95.66
Fresno-2	36.78	-119.77
IMPROVE-MammothCave	37.13	-86.15
Leland HS	33.40	-90.89
Mingo	36.97	-90.14
Railroad Valley	38.50	-115.96
SEARCH Centreville	32.90	-87.25
SEARCH Centreville2	32.90	-87.25
University of Houston	29.72	-95.34
Upper Buffalo	35.83	-93.20

and surface parameters used to generate the training set are summarized in Table 1. The distributions of fine- and coarse-mode AOT and effective radius and those of the  $f_{iso}$  coefficient at RSP wavelengths are indicated in the table with the term “empirical”. Histograms of such empirical distributions are shown in Figs. 2 and 3 for aerosol and surface parameters respectively. It must be noted that the  $f_{iso}$  coefficients at the five RSP wavelengths are not statistically independent. Their correlation matrix is also plotted in Fig. 3. The decision to keep the spectral correlation of  $f_{iso}$  during the sampling of the MODIS data set was made consciously, with the aim of introducing a constraint in the neural network retrieval. In addition to the quantities listed in Table 1, solar zenith angle, relative azimuth angle and surface pressure for the radiative transfer simulations have also been chosen randomly,



**Figure 7.** Neural network retrieved versus true Ross-thick and Li-sparse BRDF parameters and Maignan BPDF parameter from synthetic test data.

by assuming uniform distributions in the intervals 20–85°, 0–60° and 850–1050 hPa respectively. A constant flight altitude of 19 km has been assumed in the simulations, as the actual flight altitude of the ER-2 aircraft hosting RSP is usually not far from this value.

The adopted neural network model is a multilayer perceptron (Werbos, 1974) with three hidden layers of 40 neurons each, which is the same as the groundSPEX network presented in Di Noia et al. (2015), and the training was carried

out using the standard backpropagation algorithm (Rumelhart et al., 1986) regularized through learning rate annealing (Böös and Amari, 1999) and noise injection during training (Bishop, 1995b).

The trained neural network has been tested on about  $2 \times 10^5$  random simulated data not included in the training data set. In Table 2 the mean error (bias), the root mean square (rms) error, the mean absolute error (MAE) and the Pearson correlation coefficient are reported for each of the retrieved aerosol and surface parameters. The statistics for the fine-mode microphysical parameters are computed using the test data for which the fine-mode AOT was larger than 0.1, and the same holds for the coarse-mode parameters. The aerosol layer height error statistics are computed only on the cases with total (fine plus coarse mode) AOT larger than 0.1. Fine-mode AOT and effective radius appear to be the most accurately retrieved aerosol parameters. Good accuracies are also observed for coarse mode AOT and for aerosol layer height. The surface parameters are also generally well retrieved. The retrieval accuracy for  $f_{\text{iso}}$  is about 0.01, which is not sufficient for land applications but appears to be within the requirements for use in climate models (Wang et al., 2004; He et al., 2014, and references therein). Coarse-mode effective radius and effective variance appear to be the most problematic parameters, and it can be said that no retrieval capability is displayed by the neural network for these two parameters. In general, the retrieval accuracy for coarse-mode aerosol parameters seems worse than for fine-mode parameters. This is possibly due to the particular choice of the statistical distribution of the training data, in which coarse-mode-dominated scenarios are underrepresented with respect to fine-mode-dominated cases. The prevalence of cases with a small total aerosol load in the training set also explains the good retrieval capability of the neural network for surface BRDF and BPDF parameters.

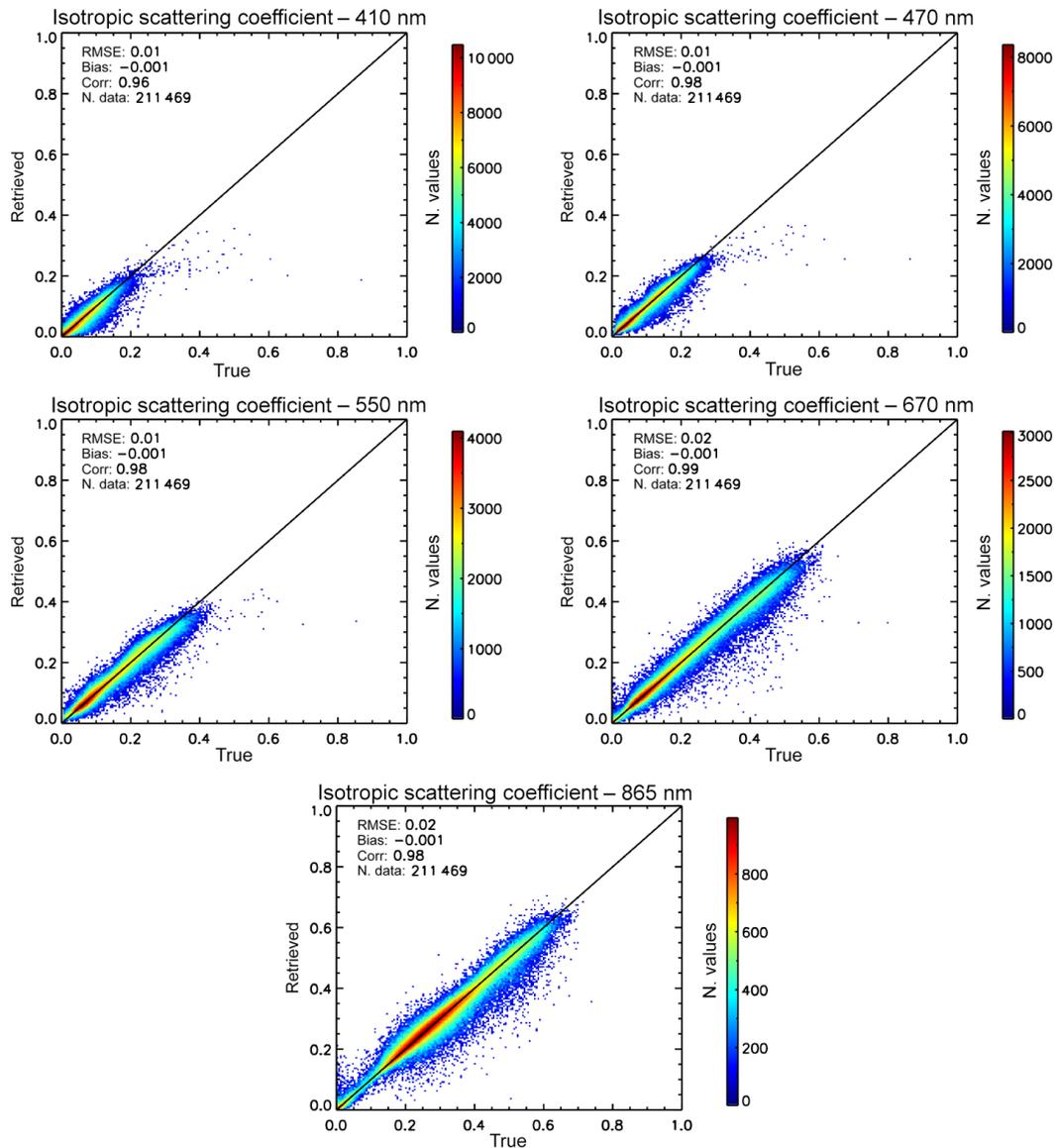
Scatter plots of the fine- and coarse-mode AOT and of the fine-mode effective radius are shown in Figs. 4 and 5. Figure 6 depicts scatter plots showing an average between fine- and coarse-mode complex refractive index weighted by the AOTs of the two modes. The plots are computed including data with total AOT larger than 0.1. Figures 7 and 8 show scatter plots of retrieved versus true BRDF and BPDF parameters and  $f_{\text{iso}}$  coefficients at all wavelengths respectively.

## 5 Use of the neural network outputs in a Phillips–Tikhonov retrieval scheme

The Phillips–Tikhonov retrieval scheme performs an iterative minimization of the following cost function:

$$C(x) = [y - F(x)]^T S_\epsilon^{-1} [y - F(x)] + \gamma (x - x_a) H(x - x_a). \quad (4)$$

In Eq. (4),  $x$  is the state vector,  $x_a$  is an a priori state vector,  $y$  is the measurement vector,  $F(x)$  is the simulated measurement vector corresponding to the state  $x$ ,  $S_\epsilon$  is the



**Figure 8.** Neural network retrieved versus true  $f_{\text{iso}}$  coefficients from synthetic test data, at the five wavelengths considered in this study.

measurement error covariance matrix,  $\mathbf{H}$  is a regularization matrix and  $\gamma$  is a regularization parameter, which is determined using the L-curve method (Hansen and O’Leary, 1993). The metric used to evaluate the convergence of a Phillips–Tikhonov retrieval is the so-called goodness-of-fit parameter

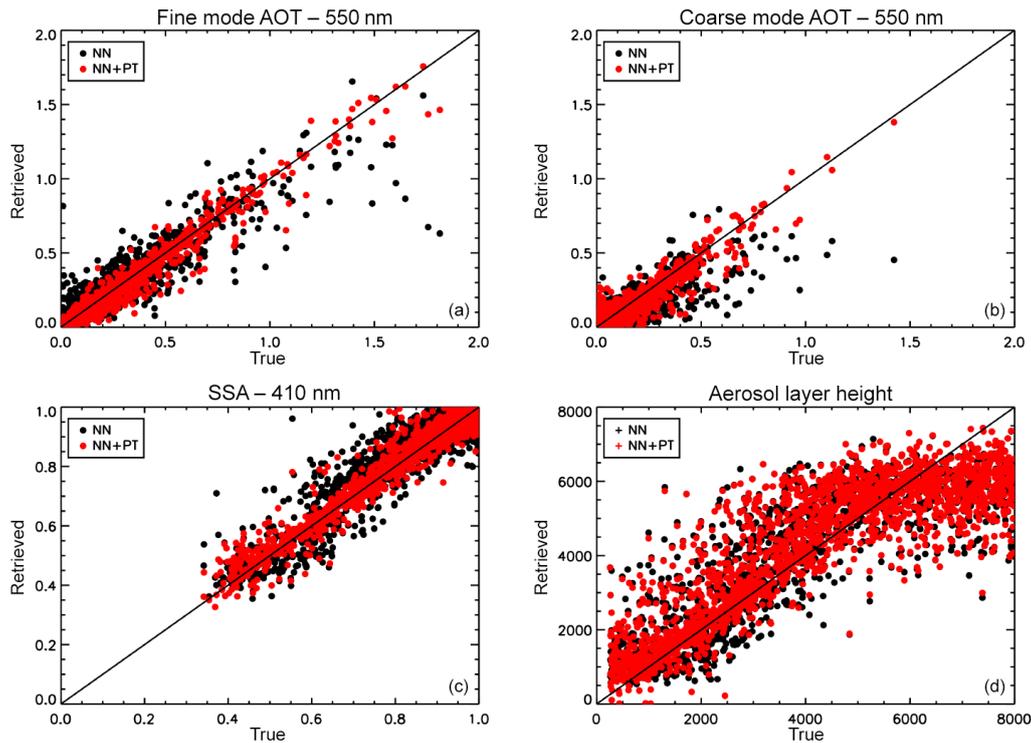
$$\chi^2 = \frac{1}{m} \sum_{i=1}^m \frac{[y_i - F_i(\mathbf{x})]^2}{\sigma_i^2}, \quad (5)$$

where  $m$  is the dimension of the measurement vector,  $y_i$  and  $F_i(\mathbf{x})$  are the  $i$ th components of  $\mathbf{y}$  and  $\mathbf{F}(\mathbf{x})$  respectively, and  $\sigma_i$  is the standard deviation of the measurement error for  $y_i$  ( $i = 1, \dots, m$ ). We empirically found that a retrieval can be

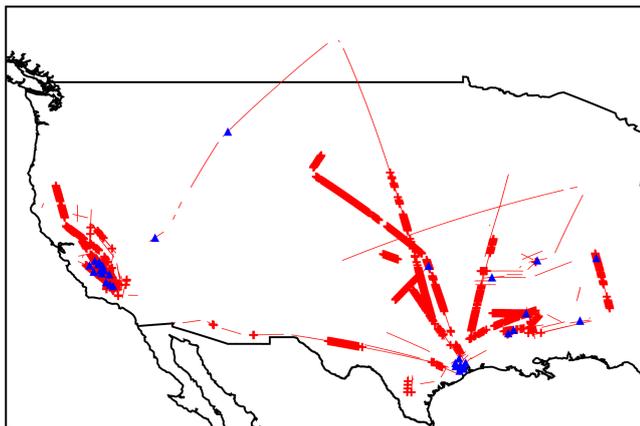
said to have converged successfully if it achieves a goodness-of-fit parameter smaller than 2.

In order to test the effect of updating the neural network retrievals through a Phillips–Tikhonov iterative scheme, we applied the combined algorithm to a sample of 5000 data, randomly picked from the approximately  $2 \times 10^5$  data used to test the neural network alone. The choice of using a subset of the data was made because applying the combined retrieval to the entire test data set would have required a much longer processing time. Out of the 5000 synthetic retrievals, 4798 converged to a  $\chi^2$  parameter smaller than 2.

The statistics of the comparison between the neural network retrieval and the combined neural network–Phillips–Tikhonov retrieval are shown in Table 3. It can be seen that the combined retrieval produces an appreciable improvement



**Figure 9.** Plots of retrieved versus true fine-mode AOT (a), coarse mode AOT (b), SSA at 410 nm (c) and aerosol layer height (d) for the neural network (NN) and the Phillips–Tikhonov algorithm using the neural network results as a first guess (NN–PT), on a sample of 4798 synthetic test cases. Statistics are given in Table 3.



**Figure 10.** Map of the RSP data used in this study. The red thin lines indicate RSP flights, the red crosses indicate the availability of CPL data and the blue triangles indicate the locations of the AERONET stations.

over the neural network retrieval on parameters such as the AOT, the single scattering albedo (SSA) and the  $f_{\text{iso}}$  coefficients, especially at the three shortest wavelengths. For these parameters, a reduction in the rms error by almost a factor two is observed. A noticeable improvement in the correlation coefficient is present for the fine-mode effective radius

and the coarse-mode FSP. For parameters such as aerosol refractive indices and aerosol layer height the improvement brought by the Phillips–Tikhonov retrieval appears less significant. Plots of retrieved versus true fine- and coarse-mode AOT, SSA and aerosol layer height are shown in Fig. 9 as examples.

## 6 Application to RSP measurements

The neural network scheme described in Sect. 4 has been used as initial guess for a Phillips–Tikhonov retrieval scheme described in Wu et al. (2015) in order to process a collection of RSP measurements carried out during the PODEX and SEAC<sup>4</sup>RS campaigns, which took place in 2012 and 2013 over several areas of the United States and part of Canada. A total of 7770 measurements have been collected, and the data of 36 AERONET stations have been used for validation. Furthermore, data from the Cloud Physics Lidar (CPL, McGill et al., 2002), hosted on the NASA ER-2 aircraft together with RSP, have been used to validate aerosol layer height retrievals and perform cloud screening. A map of the considered flights and AERONET stations used for validation is shown in Fig. 10 and the full list of AERONET stations providing validation data is given in Table 4.

In order to reduce the angular oscillations in RSP measurements due to imperfect co-registration, two preprocessing steps have been applied to RSP measurements before the application of the retrieval algorithm. First, the measurements have been averaged over a 5 km box, as in Wu et al. (2015). Second, each horizontally averaged multi-angular RSP measurement has been convolved with a moving-average Blackman filter of length 5. As explained in Sect. 3, only the measurements made at VZAs between 40° aftward and 40° forward have been used as input for the neural network. The measurement vector for the Phillips–Tikhonov retrieval described in Sect. 5, instead, covers the entire VZA range of RSP. The Phillips–Tikhonov scheme uses nine viewing angles, selected on a measurement-by-measurement basis in order to cover a scattering angle range that is as broad as possible, thereby maximizing the sensitivity of the spectropolarimetric measurements to the aerosol properties (Wu et al., 2015).

For the purpose of validating AOT retrievals, we considered each RSP measurement to be co-located with an AERONET measurement if the distance between the RSP and AERONET was not larger than 5 km and the measurements were taken no longer than 1 h apart. For the validation of aerosol properties other than AOT we had to relax the co-location criterion by allowing for a threshold of 20 km in distance and 1 day in time. This was necessary to ensure the availability of an adequate number of data points. Additional criteria imposed for validation were the following: (i) a minimum scattering angle of RSP measurements not larger than 85°, (ii) only RSP measurements flagged as cloud-free from CPL (cloud flag equal to 0) were considered, (iii) thresholds on AOT (explained later in this section) were applied when validating aerosol properties other than the AOT. Out of the original data set consisting of 7770 measurements, 2327 measurements satisfied the screening criteria (i) and (ii).

In total we found 95 AOT retrievals fulfilling the co-location criteria with a distance threshold set to 5 km and 825 retrievals with the looser threshold of 20 km. Twenty-five co-located retrievals satisfied the goodness-of-fit criterion mentioned earlier in this section ( $\chi^2 < 2$ ) with the 5 km threshold, whereas 103 retrievals met the same criterion with the 20 km threshold. The total number of converging retrievals (i.e. including non-co-located retrievals) was 223 out of 2327 retrievals which satisfy the criteria mentioned above on CPL cloud flag and minimum scattering angle. Thus, to summarize, approximately 10% of the retrievals achieved a  $\chi^2$  smaller than 2. A histogram of the  $\chi^2$  of all the retrievals is shown in Fig. 11. Approximately 75% of the retrievals reached a  $\chi^2$  less than 10, which is the convergence threshold used in Wu et al. (2015). Retrievals with a larger  $\chi^2$ , which form the tail of the distribution, are probably due to measurements that are still affected by angular oscillations after the filtering described earlier in this section.

Scatter plots of the retrieved AOTs versus those provided by the AERONET level 2 product based on the direct-sun algorithm (Holben et al., 1998) at four AERONET wavelengths are shown in Fig. 12. The results of the neural network alone as well as those of the Phillips–Tikhonov retrieval scheme are shown. The rms errors of neural network retrievals range between approximately 0.04 at 870 nm and 0.07 at 440 nm. The application of the Phillips–Tikhonov retrievals brings the rms error down to about 0.045 at all wavelengths (see Table 6). Looking at the plots, it appears that the Phillips–Tikhonov retrieval brings an improvement over the neural network retrieval in cases with AOT at 440 nm around 0.3 and larger. The neural network seems to underestimate the AOT in these cases, whereas the Phillips–Tikhonov seems to bring the retrieved values closer to the AERONET values. It is worth noting that the error statistics seem to depend critically on the distance threshold chosen for co-location, as shown in Table 5, from which a steady increase in the rms error is evident as the co-location criterion is relaxed.

A comparison between the RSP and AERONET 440–675 nm Ångström exponents, restricted to cases with retrieved AOT at 440 nm larger than 0.1, is shown in Fig. 13. From the figure it is evident that RSP retrieves systematically lower Ångström exponents than indicated by AERONET, which means that our retrieval algorithm may have a tendency to overestimate the size of aerosol particles.

A good agreement between the RSP and AERONET is observed for the fine-mode effective radius derived from the level 2 almucantar product (Dubovik and King, 2000; Dubovik et al., 2000). The results of the comparison – which is limited to cases with retrieved AOT at 440 nm larger than 0.2 – are shown in Fig. 14. The rms errors are slightly lower than 0.03  $\mu\text{m}$  for both the neural network retrieval and the final retrieval. As mentioned before, the presence of an additional filter on AOT required us to relax the co-location distance to 20 km in order to obtain a reasonable amount of data points. For this co-location criterion we obtained 46 data points, whereas by adopting a stricter criterion of 10 km the number of data would drastically drop to 16. Comparisons for complex refractive index and SSA are not shown because no AERONET data points satisfying the quality assurance criteria for level 2 products (Holben et al., 2006) were found.

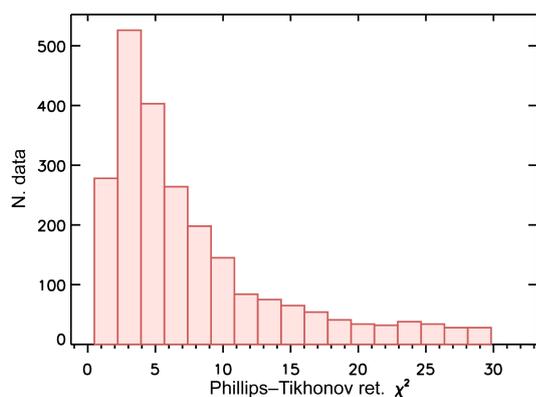
A comparison between the retrieved aerosol layer heights and CPL measurements has been made by taking into account all the cases in which the retrieved AOT at 440 nm was larger than 0.1. In Fig. 15 a plot of retrieved aerosol heights versus coincident CPL measurements is shown, with 141 coincident retrievals that passed the convergence and the other screening criteria. It is possible to see that the neural network retrievals alone are already in good agreement with the CPL, with a rms error around 1670 m and a correlation coefficient of 0.76. No clear improvement is observed over the neural network retrieval after the Phillips–Tikhonov algorithm is applied. On the contrary, a slight degradation of the valida-

**Table 5.** Rms difference between RSP and AERONET in the AOTs retrieved using the Phillips–Tikhonov scheme initialized by neural network at four wavelengths, as a function of the maximum tolerated distance for co-location.

Max. distance	N. data	440 nm	500 nm	675 nm	870 nm
5 km	25	0.044	0.044	0.045	0.045
10 km	46	0.047	0.047	0.051	0.052
15 km	71	0.067	0.068	0.070	0.069
20 km	103	0.071	0.072	0.074	0.074

**Table 6.** Error statistics from the comparisons shown from Figs. 12 to 15. for the neural network (NN) and for the Phillips–Tikhonov retrieval initialized using the neural network (NN–PT). N. data is the number of data points from which the reported statistics are computed.

Parameter	N.data	NN			NN–PT		
		Bias	RMSE	Corr.	Bias	RMSE	Corr.
AOT (440 nm)	25	−0.02	0.07	0.51	0.01	0.04	0.82
AOT (500 nm)	25	−0.01	0.06	0.48	0.02	0.04	0.80
AOT (675 nm)	25	0.01	0.05	0.35	0.02	0.04	0.69
AOT (870 nm)	25	0.01	0.04	0.18	0.02	0.04	0.50
Ångström exp. (440–675 nm)	20	−0.50	0.68	0.11	−0.44	0.61	0.14
Fine-mode effective radius ( $\mu\text{m}$ )	46	−0.02	0.03	0.66	−0.01	0.03	0.76
Aerosol layer height (km)	141	0.45	1.67	0.76	0.58	1.73	0.75

**Figure 11.** Histogram of the goodness-of-fit parameter ( $\chi^2$ ) of all the available retrievals.

tion statistics is observed (rms difference around 1735 m and correlation coefficient around 0.75), but it is doubtful that these differences bear any statistical significance. The results of the comparison also seem consistent with those shown in Wu et al. (2016), where an iterative algorithm with multiple LUT-based initializations was adopted.

The full statistics regarding bias, rms error and correlation coefficient for the comparisons shown in Figs. 12 to 15 are shown in Table 6.

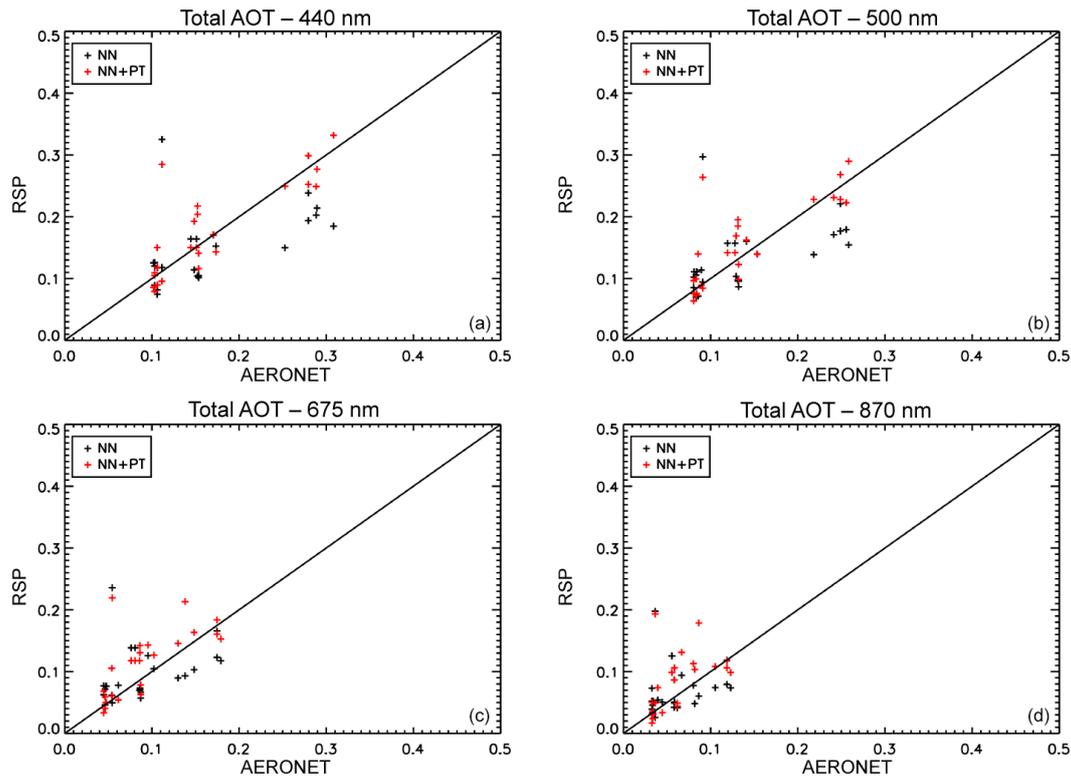
Compared to the traditional algorithm initialized by a LUT (Hasekamp et al., 2011; Wu et al., 2015), the algorithm initialized by means of the neural network is characterized by a larger number of converging retrievals and a lower processing time. By applying the LUT-based algorithm to the

data set used in this study, we obtained in total 47 converging retrievals (223 with the neural network). The time required in order to process the entire data set can be reduced by roughly a factor of 2 when using the neural network algorithm to provide the first guess. The differences in the retrieved aerosol parameters between the LUT-based and the NN-based algorithm do not seem significant. The same tendency to overestimate coarse-mode aerosol loads is also observed in LUT-based retrievals. Plots of retrieved versus AERONET (or CPL) aerosol parameters for the LUT-based algorithm are given in the Supplement.

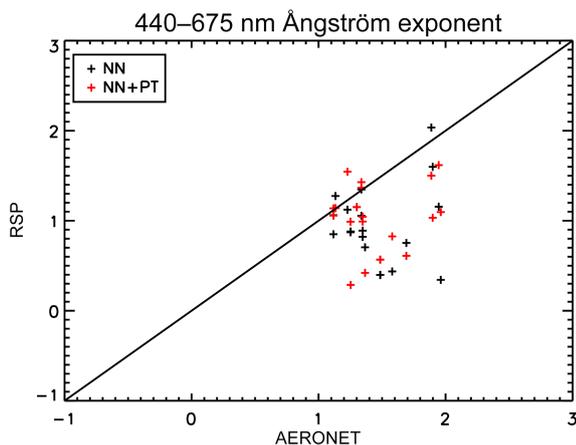
## 7 Conclusions

In this paper we have demonstrated the application of neural networks to aerosol retrievals from the RSP instrument. First, we trained a neural network to retrieve aerosol and surface properties from simulated multi-angle spectropolarimetric measurements, with an observation geometry that simulates that of RSP. Then we applied the neural network to real RSP measurements carried out over the United States in 2012 and 2013 and used the neural network retrieval as a first guess for an iterative algorithm based on the Phillips–Tikhonov method.

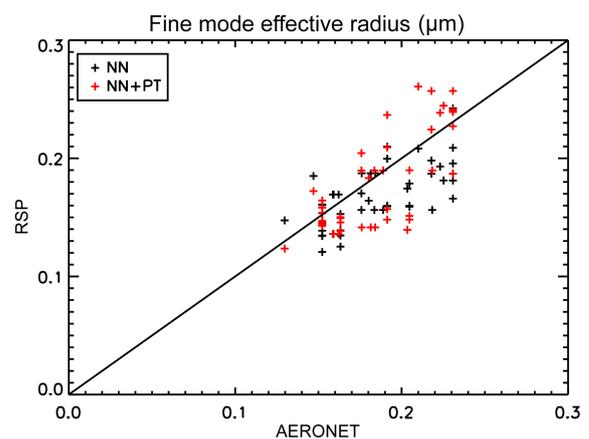
In order to assess the quality of our retrievals, we compared the retrieved aerosol parameters and aerosol layer heights to co-located AERONET and CPL lidar measurements respectively. We observed good retrieval capability for AOT (although with a possible overestimation of the coarse mode), fine-mode effective radius and aerosol layer height. A



**Figure 12.** RSP versus AERONET total AOT at the four AERONET wavelengths lying inside the spectral range of the RSP measurements used in this study. Black crosses: neural network retrievals. Red crosses: Phillips–Tikhonov with neural network first guess. Statistics shown in Table 6.



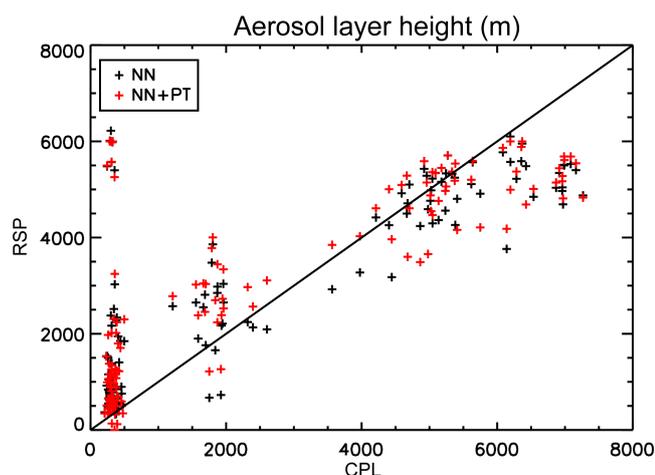
**Figure 13.** RSP versus AERONET 440–675 nm Ångström exponent. Black crosses: neural network retrievals. Red crosses: Phillips–Tikhonov with neural network first guess. Statistics shown in Table 6.



**Figure 14.** RSP versus AERONET fine-mode effective radius. Black crosses: neural network retrievals. Red crosses: Phillips–Tikhonov with neural network first guess. Statistics shown in Table 6.

comparison of the retrieval results with those obtained using an iterative algorithm initialized using a LUT confirms the finding that replacing the LUT with a neural network generally leads to a higher number of converging retrievals and to a reduced processing time.

Out of 2327 RSP measurements which passed screening criteria based on scattering angle and cloudiness, the application of the combined neural network/Phillips–Tikhonov scheme led to 223 retrievals which converged to a goodness-of-fit indicator lower than 2. Forty-two converging re-



**Figure 15.** RSP versus CPL aerosol layer height. Black crosses: neural network retrievals. Red crosses: Phillips–Tikhonov with neural network first guess. Statistics shown in Table 6.

trievals were obtained instead, using the previous version of the retrieval scheme, in which a LUT was used instead of the neural network to provide the first guess for the Phillips–Tikhonov algorithm. Thus, even with the neural network/Phillips–Tikhonov retrieval scheme, about 10 % of the retrievals satisfied the convergence criterion, which, in absolute terms, is a relatively low percentage. This behaviour may be due to several factors, ranging from residual co-registration errors in RSP to inadequacies in the assumptions in the surface model.

Training a neural network algorithm for multi-parameter aerosol retrievals from airborne or satellite multi-angle spectropolarimetric measurements remains challenging. In our case, for example, the use of a training set encompassing mostly fine-mode-dominated cases with moderate-to-low aerosol loads resulted in better retrieval capabilities for the fine-mode aerosol parameters than for the coarse-mode parameters and in good retrieval capabilities for all the surface parameters. Choosing a mixture of training cases for a single neural network that allows equal performances to be achieved for all the aerosol and surface parameters on fine mode as well as on coarse-mode-dominated situations, or with low as well as with high aerosol loads, may be a challenging task. A possible improvement on the present work may involve training multiple neural networks, each covering a certain range of situations, with possibly an additional neural network that performs a preliminary classification, thereby deciding which of the lower-level networks should be used for retrieval. Approaches of this kind have been already applied to simple inverse scattering problems involving spheres and spheroids (Ulanowski et al., 1998; Berdnik et al., 2004), and may be as well investigated – with due adaptations – for more complex retrievals. The aforementioned approach may be also useful for extending the method

presented in this paper to multi-angle spectropolarimeters in which the angular dependence of the measurements is more complicated, such as POLDER, AirMSPi and the forthcoming 3MI, where it is not possible to assume that all the measurements are made at the same relative azimuth angle apart from shifts of  $180^\circ$ .

*Data availability.* The data used or produced in this study are available upon request from the corresponding author. AERONET data are publicly available at the website [https://aeronet.gsfc.nasa.gov/new\\_web/data.html](https://aeronet.gsfc.nasa.gov/new_web/data.html).

**The Supplement related to this article is available online at <https://doi.org/10.5194/amt-10-4235-2017-supplement>.**

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

*Acknowledgements.* The RSP data used in this study are publicly available from the NASA Goddard Institute for Space Studies (<http://data.giss.nasa.gov/pub/rsp/>). The RSP data from the SEAC<sup>4</sup>RS and PODEX field experiments were funded by the NASA Radiation Sciences Program managed by Hal Maring and by the NASA Earth Science Division, as part of the pre-formulation study for the Aerosol Cloud and ocean Ecosystem (ACE) mission.

The CPL data are provided by the NASA Goddard Space Flight Center (<http://cpl.gsfc.nasa.gov/>).

We thank the principal investigators (PIs) and their staff for establishing and maintaining the AERONET sites, the data of which have been used in this paper.

The MODIS Terra+Aqua MCD43C3 BRDF/Albedo L3 data product was retrieved from the online data pool, courtesy of the NASA Land Processes Distributed Active Archive Center (LP DAAC), USGS/Earth Resources Observation and Science (EROS) Center, Sioux Falls, South Dakota, United States, [https://lpdaac.usgs.gov/data\\_access/data\\_pool](https://lpdaac.usgs.gov/data_access/data_pool).

Alexei Lyapustin and an anonymous reviewer are gratefully acknowledged for their constructive comments, which helped to improve this paper.

Edited by: Alexander Kokhanovsky

Reviewed by: Alexei Lyapustin and one anonymous referee

## References

- Berdnik, V. V., Mukhamedyarov, R. D., and Loiko, V. A.: Sizing of soft spheroidal particles by multiangle scattered light intensity data: application of neural networks, *J. Quant. Spectrosc. Ra.*, 89, 279–289, <https://doi.org/10.1016/j.jqsrt.2004.05.028>, 2004.
- Bishop, C. M.: *Neural Networks for Pattern Recognition*, Oxford University Press, New York, NY, United States, 1995a.

- Bishop, C. M.: Training with noise is equivalent to Tikhonov regularization, *Neural Comput.*, 7, 108–116, <https://doi.org/10.1162/neco.1995.7.1.108>, 1995b.
- Bös, S. and Amari, S.: Annealed online learning in multilayer neural networks, in: *On-Line Learning in Neural Networks*, edited by: Saad, D., chap. 10, 209–229, Cambridge University Press, New York, NY, United States, <https://doi.org/10.1017/CBO9780511569920.011>, 1999.
- Cairns, B., Russell, E. E., and Travis, L. D.: Research Scanning Polarimeter: calibration and ground-based measurements, in: *Proc. SPIE 3754, Polarization: Measurement, Analysis, and Remote Sensing II*, 186, <https://doi.org/10.1117/12.366329>, 1999.
- Chimot, J., Veeffkind, 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<sub>2</sub>-O<sub>2</sub> spectral band using a neural network approach, *Atmos. Meas. Tech.*, 10, 783–809, <https://doi.org/10.5194/amt-10-783-2017>, 2017.
- Del Frate, F. and Schiavon, G.: A combined natural orthogonal functions/neural network technique for the radiometric estimation of atmospheric profiles, *Radio Sci.*, 33, 405–410, <https://doi.org/10.1029/97RS02219>, 1998a.
- Del Frate, F. and Schiavon, G.: Neural networks for the retrieval of water vapor and liquid water from radiometric data, *Radio Sci.*, 33, 1373–1386, <https://doi.org/10.1029/98RS02133>, 1998b.
- Deschamps, P.-Y., Bréon, F.-M., Leroy, M., Podaire, A., Bricaud, A., Buriez, J.-C., and Sèze, G.: The POLDER mission: instrument characteristics and scientific objectives, *IEEE T. Geosci. Remote*, 32, 598–615, <https://doi.org/10.1109/36.297978>, 1994.
- Deuzé, J. L., Goloub, P., Herman, M., Marchand, A., Perry, G., Susana, S., and Tanré, D.: Estimate of the aerosol properties over the ocean with POLDER, *J. Geophys. Res.*, 105, 15329–15346, <https://doi.org/10.1029/2000JD900148>, 2000.
- Deuzé, J. L., Bréon, F., Devaux, C., Goloub, P., Herman, M., Lafrance, B., Maignan, F., Marchand, A., Nadal, F., Perry, G., and Tanré, D.: Remote sensing of aerosols over land surfaces from POLDER-ADEOS-1 polarized measurements, *J. Geophys. Res.*, 106, 4913–4926, <https://doi.org/10.1029/2000JD900364>, 2001.
- Diner, D. J., Xu, F., Garay, M. J., Martonchik, J. V., Rheingans, B. E., Geier, S., Davis, A., Hancock, B. R., Jovanovic, V. M., Bull, M. A., Capraro, K., Chipman, R. A., and McClain, S. C.: The Airborne Multiangle Spectropolarimetric Imager (AirMSPI): a new tool for aerosol and cloud remote sensing, *Atmos. Meas. Tech.*, 6, 2007–2025, <https://doi.org/10.5194/amt-6-2007-2013>, 2013.
- Di Noia, A., Hasekamp, O. P., van Harten, G., Rietjens, J. H. H., Smit, J. M., Snik, F., Henzing, J. S., de Boer, J., Keller, C. U., and Volten, H.: Use of neural networks in ground-based aerosol retrievals from multi-angle spectropolarimetric observations, *Atmos. Meas. Tech.*, 8, 281–299, <https://doi.org/10.5194/amt-8-281-2015>, 2015.
- Dubovik, O. and King, M. D.: A flexible inversion algorithm for retrieval of aerosol optical properties from Sun and sky radiance measurements, *J. Geophys. Res.*, 105, 20673–20796, <https://doi.org/10.1029/2000JD900282>, 2000.
- Dubovik, O., Smirnov, A., Holben, B. N., King, M. D., Kaufman, Y. J., Eck, T. F., and Slutsker, I.: Accuracy assessments of aerosol optical properties retrieved from Aerosol Robotic Network (AERONET) Sun and sky radiance measurements, *J. Geophys. Res.*, 105, 9791–9806, <https://doi.org/10.1029/2000JD900040>, 2000.
- Dubovik, O., Herman, M., Holdak, A., Lapyonok, T., Tanré, D., Deuzé, J. L., Ducos, F., Sinyuk, A., and Lopatin, A.: Statistically optimized inversion algorithm for enhanced retrieval of aerosol properties from spectral multi-angle polarimetric satellite observations, *Atmos. Meas. Tech.*, 4, 975–1018, <https://doi.org/10.5194/amt-4-975-2011>, 2011.
- Hansen, P. C. and O’Leary, D.: The use of the L-curve in the regularization of discrete ill-posed problems, *SIAM J. Sci. Comput.*, 14, 1487–1503, <https://doi.org/10.1137/0914086>, 1993.
- Hasekamp, O. P. and Landgraf, J.: A linearized vector radiative transfer model for atmospheric trace gas retrieval, *J. Quant. Spectrosc. Ra.*, 75, 221–238, [https://doi.org/10.1016/S0022-4073\(01\)00247-3](https://doi.org/10.1016/S0022-4073(01)00247-3), 2002.
- Hasekamp, O. P. and Landgraf, J.: Linearization of vector radiative transfer with respect to aerosol properties and its use in satellite remote sensing, *J. Geophys. Res.*, 110, D04203, <https://doi.org/10.1029/2004JD005260>, 2005.
- Hasekamp, O. P. and Landgraf, J.: Retrieval of aerosol properties over land surfaces: capabilities of multiple-viewing-angle intensity and polarization measurements, *Appl. Optics*, 46, 3332–3344, <https://doi.org/10.1364/AO.46.003332>, 2007.
- Hasekamp, O. P., Litvinov, P., and Butz, A.: Aerosol properties over the ocean from PARASOL multiangle photopolarimetric measurements, *J. Geophys. Res.*, 116, D14204, <https://doi.org/10.1029/2010JD015469>, 2011.
- He, T., Liang, S., and Song, D.-X.: Analysis of global land surface albedo climatology and spatial-temporal variation during 1981–2010 from multiple satellite products, *J. Geophys. Res.*, 119, 10281–10298, <https://doi.org/10.1002/2014JD021667>, 2014.
- Holben, B., Eck, T. F., Slutsker, I., Smirnov, A., Sinyuk, A., Schafer, J., Giles, D., and Dubovik, O.: AERONET’s Version 2.0 quality assurance criteria, in: *Proc. SPIE 6408, Remote Sensing of the Atmosphere and Clouds*, edited by: Tsay, S.-C., Nakajima, T., Singh, R. P., and Sridharan, R., 64080Q–64080Q–14, <https://doi.org/10.1117/12.706524>, 2006.
- 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., Lavenue, F., Jankowiak, I., and Smirnov, A.: AERONET – A federated instrument network and data archive for aerosol characterization, *Remote Sens. Environ.*, 66, 1–16, [https://doi.org/10.1016/S0034-4257\(98\)00031-5](https://doi.org/10.1016/S0034-4257(98)00031-5), 1998.
- Hornik, K., Stinchcombe, M., and White, H.: Multilayer feedforward networks are universal approximators, *Neural Networks*, 2, 359–366, [https://doi.org/10.1016/0893-6080\(89\)90020-8](https://doi.org/10.1016/0893-6080(89)90020-8), 1989.
- Jackson, D. L., Wick, G. A., and Bates, J. J.: Near-surface retrieval of air temperature and specific humidity using multisensor microwave satellite observations, *J. Geophys. Res.*, 111, D10306, <https://doi.org/10.1029/2005JD006431>, 2006.
- Knobelspiesse, K., Cairns, B., Redemann, J., Bergstrom, R. W., and Stohl, A.: Simultaneous retrieval of aerosol and cloud properties during the MILAGRO field campaign, *Atmos. Chem. Phys.*, 11, 6245–6263, <https://doi.org/10.5194/acp-11-6245-2011>, 2011.
- Knobelspiesse, K., Cairns, B., Mishchenko, M., Chowdhary, J., Tsigaridis, K., van Dierenhoven, B., Martin, W., Ottaviani, M., and Alexandrov, M.: Analysis of fine-mode aerosol retrieval capabilities by different passive remote sens-

- ing instrument designs, *Opt. Express*, 20, 21457–21484, <https://doi.org/10.1364/OE.20.021457>, 2012.
- Kokhanovsky, A. A., Davis, A. B., Cairns, B., Dubovik, O., Hasekamp, O. P., Sano, I., Mukai, S., Rozanov, V. V., Litvinov, P., Lapyonok, T., Kolomiets, I. S., Oberemok, Y. A., Savenkov, S. N., Martin, W., Wasilewski, A., Di Noia, A., Stap, F. A., Rietjens, J., Xu, F., Natraj, V., Duan, M., Cheng, T., and Munro, R.: Space-based remote sensing of atmospheric aerosols: The multi-angle spectro-polarimetric frontier, *Earth. Sci. Rev.*, 145, 85–116, <https://doi.org/10.1016/j.earscirev.2015.01.012>, 2015.
- Krasnopolsky, V. M.: Neural network emulations for complex multidimensional mappings: Applications of neural network techniques to atmospheric and oceanic satellite retrievals and numerical modeling, *Rev. Geophys.*, 45, RG3009, <https://doi.org/10.1029/2006RG000200>, 2007.
- Leshno, M., Lin, V. Y., Pinkus, A., and Schocken, S.: Multilayer feedforward networks with a nonpolynomial activation function can approximate any function, *Neural Networks*, 6, 861–867, [https://doi.org/10.1016/S0893-6080\(05\)80131-5](https://doi.org/10.1016/S0893-6080(05)80131-5), 1993.
- Lucht, W., Schaaf, C. B., and Strahler, A. H.: An Algorithm for the Retrieval of Albedo from Space Using Semiempirical BRDF Models, *IEEE T. Geosci. Remote Sens.*, 38, 977–998, <https://doi.org/10.1109/36.841980>, 2000.
- Maignan, F., Bréon, F.-M., and Lacaze, R.: Bidirectional reflectance of Earth targets: evaluation of analytical models using a large set of spaceborne measurements with emphasis on the Hot Spot, *Remote Sens. Environ.*, 90, 210–220, <https://doi.org/10.1016/j.rse.2003.12.006>, 2004.
- Maignan, F., Bréon, F.-M., Fédèle, E., and Bouvier, M.: Polarized reflectances of natural surfaces: Spaceborne measurements and analytical modeling, *Remote Sens. Environ.*, 113, 2642–2650, <https://doi.org/10.1016/j.rse.2009.07.022>, 2009.
- Marbach, T., Phillips, P., Lacan, A., and Schlüssel, P.: 3MI: The Multi-Viewing, -Channel, -Polarisation Imager of the EUMETSAT Polar System – Second Generation (EPS-SG) dedicated to aerosol characterization, in: *Proc. SPIE 8889, Sensors, Systems and Next-Generation Satellites XVIII*, <https://doi.org/10.1117/12.2028221>, 2013.
- Martins, J. V., Nielsen, T., Fish, C., Sparr, L., Fernandez-Borda, R., Schoeberl, M., and Remer, L.: HARP-CubeSat – An innovative hyperangular imaging polarimeter for Earth science applications, available at: <http://digitalcommons.usu.edu/cgi/viewcontent.cgi?filename=0&%article=3154&context=smallsat&type=additional>, presented at Small Sat Pre-Conference Workshop, Logan, UT, United States, 3 August 2014.
- McGill, M., Hlavka, D., Hart, W., Scott, V. S., Spinhirne, J., and Schmid, B.: Cloud Physics Lidar: instrument description and initial measurement results, *Appl. Optics*, 41, 3725–3734, <https://doi.org/10.1364/AO.41.003725>, 2002.
- Milinevsky, G., Yatskiv, Y., Degtyaryov, O., Syniavskiy, I., Mishchenko, M., Rosenbush, V., Ivanov, Y., Makarov, A., Bovchaliuk, A., Danylevsky, V., Sosonkin, M., Moskalov, S., Bovchaliuk, V., Lukenyuk, A., Shymkiv, A., and Udodov, E.: New satellite project Aerosol-UA: Remote sensing of aerosols in the terrestrial atmosphere, *Acta Astronaut.*, 123, 292–300, <https://doi.org/10.1016/j.actaastro.2016.02.027>, 2016.
- Mishchenko, M. I. and Travis, L. D.: Satellite retrievals of aerosol properties over the ocean using polarization as well as intensity of reflected sunlight, *J. Geophys. Res.*, 102, 16989–17013, <https://doi.org/10.1029/96JD02425>, 1997.
- Mishchenko, M. I., Cairns, B., Hansen, J. E., Travis, L. D., Kopp, G., Schuelern, C. F., Fafaul, B. A., Hooker, R. J., Maring, H. B., and Itchkawich, T.: Accurate Monitoring of Terrestrial Aerosols and Total Solar Irradiance: Introducing the Glory Mission, *B. Am. Meteorol. Soc.*, 88, 677–691, <https://doi.org/10.1175/BAMS-88-5-677>, 2007.
- Radosavljevic, V., Vucetic, S., and Obradovic, Z.: A data-mining technique for aerosol retrieval across multiple accuracy measures, *IEEE Geosci. Remote S.*, 7, 411–415, <https://doi.org/10.1109/LGRS.2009.2037720>, 2010.
- Rodgers, C. D.: *Inverse Methods for Atmospheric Sounding: Theory and Practice*, World Scientific, London, United Kingdom, 2000.
- Roujean, J.-L., Leroy, M., and Deschamps, P.-Y.: A bidirectional reflectance model of the Earth's surface for the correction of remote sensing data, *J. Geophys. Res.*, 97, 20455–20468, <https://doi.org/10.1029/92JD01411>, 1992.
- Rumelhart, D. E., Hinton, G. E., and Williams, R. J.: Learning representations by back-propagating errors, *Nature*, 323, 533–536, <https://doi.org/10.1038/323533a0>, 1986.
- Schaaf, C. and Wang, Z.: MCD43C3 MODIS/Terra+Aqua BRDF/Albedo Albedo Daily L3 Global 0.05Deg CMG V006, NASA EOSDIS Land Processes DAAC, <https://doi.org/10.5067/MODIS/MCD43C3.006>, 2015.
- Smith, W. L., Woolf, H. M., and Jacob, W. J.: A regression method for obtaining real-time temperature and geopotential height profiles from satellite spectrometer measurements and its application to Nimbus 3 “SIRS” observations, *Mon. Weather Rev.*, 98, 582–603, [https://doi.org/10.1175/1520-0493\(1970\)098<0582:ARMPFOR>2.3.CO;2](https://doi.org/10.1175/1520-0493(1970)098<0582:ARMPFOR>2.3.CO;2), 1970.
- Taylor, M., Kazadzis, S., Tsekeri, A., Gkikas, A., and Amiridis, V.: Satellite retrieval of aerosol microphysical and optical parameters using neural networks: a new methodology applied to the Sahara desert dust peak, *Atmos. Meas. Tech.*, 7, 3151–3175, <https://doi.org/10.5194/amt-7-3151-2014>, 2014.
- Ulanowski, Z., Wang, Z., Kaye, P., and Ludlow, I. K.: Application of neural networks to the inverse light scattering problem for spheres, *Appl. Optics*, 37, 4027–4033, <https://doi.org/10.1364/AO.37.004027>, 1998.
- van Amerongen, A., Rietjens, J., Smit, M., van Loon, D., van Brug, H., van der Meulen, W., Esposito, M., and Hasekamp, O.: SPEX: The Dutch roadmap towards aerosol measurement from space, in: *Proc. International Conference on Space Optics (ICSO) 2016*, Biarritz, France, 18–21 October 2016.
- Wang, K., Liu, J., Zhou, X., Sparrow, M., Ma, M., Sun, Z., and Jiang, W.: Validation of the MODIS global land surface albedo product using ground measurements in a semidesert region on the Tibetan Plateau, *J. Geophys. Res.*, 109, D05107, <https://doi.org/10.1029/2003JD004229>, 2004.
- Wanner, W., Li, X., and Strahler, A. H.: On the derivation of kernels for kernel-driven models of bidirectional reflectance, *J. Geophys. Res.*, 100, 21077–21089, <https://doi.org/10.1029/95JD02371>, 1995.
- Waquet, F., Cairns, B., Knobelspiesse, K., Chowdhary, J., Travis, L. D., Schmid, B., and Mishchenko, M. I.: Polarimetric remote sensing of aerosols over land, *J. Geophys. Res.*, 114, D01206, <https://doi.org/10.1029/2008JD010619>, 2009.

- Werbos, P. J.: Beyond Regression: New Tools for Prediction and Analysis in the Behavioral Sciences, PhD thesis, Harvard University, Cambridge, MA, United States, 1974.
- Wu, L., Hasekamp, O., van Diedenhoven, B., and Cairns, B.: Aerosol retrieval from multiangle, multispectral photopolarimetric measurements: importance of spectral range and angular resolution, *Atmos. Meas. Tech.*, 8, 2625–2638, <https://doi.org/10.5194/amt-8-2625-2015>, 2015.
- Wu, L., Hasekamp, O., van Diedenhoven, B., Cairns, B., Yorks, J. E., and Chowdhary, J.: Passive remote sensing of aerosol layer height using near-UV multiangle polarization measurements, *Geophys. Res. Lett.*, 43, 8783–8790, <https://doi.org/10.1002/2016GL069848>, 2016.