

## Retrieval of Aerosol Properties from TROPOMI Measurements

**Lanlan Rao**

Vollständiger Abdruck der von der TUM School of Engineering and Design der  
Technischen Universität München zur Erlangung des akademischen Grades einer

**Doktorin der Ingenieurwissenschaften (Dr.-Ing.)**

genehmigten Dissertation.

**Vorsitzender:**

Prof. Dr.-Ing. Florian Seitz

**Prüfende der Dissertation:**

1. Priv.-Doz. Dr.-Ing. habil. Adrian Doicu
2. Prof. Dr.-Ing. habil. Richard Hans Georg Bamler
3. Prof. Dr. Jason Blake Cohen

Die Dissertation wurde am 17.03.2022 bei der Technischen Universität München  
eingereicht und durch die TUM School of Engineering and Design am 10.10.2022  
angenommen.



# Abstract

Aerosols affect Earth's radiation budget by scattering and absorbing solar radiation (direct effect) and by influencing the cloud formation processes (indirect effect). Accurate assessments of aerosol properties, such as optical depth and layer height, are important for the global monitoring of air pollution in the lower atmosphere.

A number of passive satellite sensors enable to monitor aerosol properties on both regional and global scale using spectral information at various wavelengths. For example, measurements in the oxygen A-band from the Scanning Imaging Absorption Spectrometer for Atmospheric CHartographY (SCIAMACHY), the Global Ozone Mapping Experiment (GOME) and GOME-2, the Greenhouse Gases Observing Satellite (GOSAT), and the TROPospheric Monitoring Instrument (TROPOMI) are used to retrieve both aerosol optical depth and layer height. The ultimate generation of passive satellite sensors, as for example, the TROPOMI/S5P has an extraordinary spatial, temporal and spectral resolutions. The challenge of processing hyperspectral data is to increase the performance of the retrieval algorithms in order to achieve near-real-time requirements.

The goal of this thesis is the design of algorithms for retrieving aerosol parameters from TROPOMI/S5P measurements in the oxygen A-band. The designed algorithms can be grouped into two categories. The first category includes Bayesian-based retrieval algorithms for a specified aerosol model and a set of candidate models. In the latter case, two solutions estimates, namely (i) the maximum solution estimate, corresponding to the model with the highest evidence, and (ii) the mean solution estimate, representing a linear combination of solutions weighted by their evidences, are proposed. The algorithms use a linearized radiative transfer model relying on the discrete ordinate method with matrix exponential, and as acceleration approaches, the telescoping technique, the method of false discrete ordinate, the correlated k-distribution method, and the principal component analysis. The inverse problem is formulated as a least-squares problem and solved by means of the iteratively regularized Gauss-Newton method. The second category includes neural network retrieval algorithms. These are trained (i) to emulate the radiative transfer model, which is then used in conjunction with a Bayesian approach to solve the inverse problem, and (ii) to learn the inverse model using as input either the synthetic radiances or their principal components. The retrieval performances of the retrieval algorithms are analyzed on synthetic and real data.



# Zusammenfassung

Aerosole beeinflussen den Strahlungsbilanz der Erde, indem sie Sonnenstrahlung streuen und absorbieren (direkte Wirkung) und die Wolkenbildungsprozesse beeinflussen (indirekte Wirkung). Genaue Bewertungen der Aerosoleigenschaften wie optische Tiefe und Schichthöhe sind wichtig für die globale Überwachung der Luftverschmutzung in der unteren Atmosphäre.

Eine Reihe von passiven Satellitensensoren ermöglichen die Überwachung von Aerosoleigenschaften sowohl auf regionaler als auch auf globaler Ebene unter Verwendung von Spektralinformationen bei verschiedenen Wellenlängen. Zum Beispiel Messungen im Sauerstoff-A-Band vom Scanning Imaging Absorption Spectrometer for Atmospheric Chartography (SCIAMACHY), dem Global Ozone Mapping Experiment (GOME) und GOME-2, dem Greenhouse Gases Observing Satellite (GOSAT) und dem TROPospheric Monitoring Instrument (TROPOMI) werden verwendet, um sowohl die optische Tiefe als auch die Schichthöhe wiederherstellen. Die ultimative Generation passiver Satellitensensoren wie zum Beispiel der TROPOMI/S5P hat außergewöhnliche räumliche, zeitliche und spektrale Auflösungen erreicht. Die Herausforderung bei der Verarbeitung hyperspektraler Daten besteht darin, die Leistung der Retrievalalgorithmen zu steigern, um Anforderungen an die nahezu Echtzeit zu erfüllen.

Das Ziel dieser Arbeit ist die Entwicklung von Algorithmen zum Retrieval von Aerosolparametern aus TROPOMI/S5P-Messungen im Sauerstoff-A-Band. Die entworfenen Algorithmen werden in zwei Kategorien eingeteilt. Die erste Kategorie umfasst Bayes'sche Retrievalalgorithmen für ein bestimmtes Aerosolmodell und einen Satz von Kandidatenmodellen. Im letzteren Fall werden zwei Lösungsschätzungen vorgeschlagen, nämlich (i) die maximale Lösungsschätzung, die dem Modell mit der höchsten Evidenz entspricht, und (ii) die mittlere Lösungsschätzung, die eine lineare Kombination von Lösungen darstellt, die mit ihren Evidenzen gewichtet ist. Die Algorithmen verwenden ein linearisiertes Strahlungstransfermodell, das auf der diskreten Ordinatenmethode mit Matrixexponentialfunktion beruht, und als Beschleunigungsverfahren, die Teleskoptechnik, die Methode der falschen diskreten Ordinaten, die korrelierte k-Verteilungsmethode und die Hauptkomponentenanalyse benützt. Das inverse Problem wird als Kleinste-Quadrate-Problem formuliert und mittels des iterativ regularisierten Gauß-Newton-Verfahrens gelöst. Die zweite Kategorie umfasst Retrievalalgorithmen für neuronale Netze. Diese werden darauf trainiert, (i) das Strahlungstransfermodell zu emulieren, das dann in Verbindung mit einem Bayes'schen Verfahren verwendet wird, um das inverse Problem zu lösen, und (ii) das inverse Modell zu lernen, indem entweder die synthetischen Spektren oder ihre Hauptkomponenten als Eingabe verwendet werden. Die Leistungen der Retrievalalgorithmen werden anhand von synthetischen und realen Daten analysiert.



# Contents

<b>Abstract</b>	<b>iii</b>
<b>Zusammenfassung</b>	<b>v</b>
<b>1 Introduction</b>	<b>1</b>
1.1 Passive remote sensing techniques . . . . .	1
1.2 Objective of the thesis . . . . .	6
1.3 Thesis structure . . . . .	7
<b>2 Aerosols database</b>	<b>9</b>
2.1 Aerosol components . . . . .	9
2.2 Aerosol models . . . . .	11
2.3 Aerosol vertical profiles . . . . .	13
<b>3 Radiative transfer model</b>	<b>19</b>
3.1 Radiative transfer equation . . . . .	19
3.2 Discrete ordinate method with matrix exponential . . . . .	22
3.3 Acceleration techniques . . . . .	24
3.3.1 Telescoping technique . . . . .	24
3.3.2 False discrete ordinate . . . . .	25
3.3.3 The correlated $k$ -distribution method and the Principal Component Analysis (PCA) . . . . .	26
3.4 Linearized radiative transfer model . . . . .	28
3.4.1 Derivatives with respect to aerosol layer height . . . . .	28
3.4.2 Linearized forward approach . . . . .	29
3.4.3 Forward-adjoint approach . . . . .	30
3.5 Forward models . . . . .	35
<b>4 Bayesian-based retrieval algorithms</b>	<b>37</b>
4.1 Retrieval algorithm for a known aerosol model . . . . .	37
4.1.1 Algorithm description . . . . .	37
4.1.2 Sensitivity analysis . . . . .	40
4.2 Retrieval algorithm with aerosol model selection . . . . .	47
4.2.1 Algorithm description . . . . .	47
4.2.2 Numerical analysis . . . . .	50
<b>5 Neural network retrieval algorithms</b>	<b>65</b>
5.1 Algorithm descriptions . . . . .	66
5.1.1 Neural network for the forward operator . . . . .	66
5.1.2 Neural network for the inverse operator . . . . .	67
5.2 Numerical analysis . . . . .	68
5.2.1 Synthetic retrieval . . . . .	70

*Contents*

5.2.2	Retrieval from real data . . . . .	72
5.2.3	Conclusions . . . . .	76
<b>6</b>	<b>Conclusions</b>	<b>79</b>
6.1	Summary . . . . .	79
6.2	Outlook . . . . .	81
	<b>List of Figures</b>	<b>83</b>
	<b>List of Tables</b>	<b>87</b>
	<b>Bibliography</b>	<b>89</b>
	<b>Acknowledgements</b>	<b>99</b>
	<b>Appendices</b>	<b>101</b>
<b>A</b>	<b>Hyperspectral Satellite Remote Sensing of Aerosol Parameters</b>	<b>103</b>
<b>B</b>	<b>Optimization of Aerosol Model Selection for TROPOMI/S5P</b>	<b>121</b>
<b>C</b>	<b>Aerosol Parameters Retrieval From TROPOMI/S5P Using Physics-Based Neural Networks</b>	<b>141</b>



# 1 Introduction

Atmospheric aerosols from both natural and anthropogenic sources have a significant impact on Earth's climate and environment. They affect Earth's energy budget directly by scattering and absorbing solar radiation and indirectly by altering the lifetime and radiative properties of clouds. Moreover, highly absorbing aerosols have a warming effect on the atmosphere leading to the evaporation of cloud particles, and so, to a reduction of the cloud cover (semi-direct effect). Accurate assessments of aerosol properties, such as optical depth and layer height, are important for the global monitoring of air pollution in the lower atmosphere. In particular, the vertical distribution of aerosols determines the magnitude of the aerosol scattering effects [Peters et al., 2011; Wilcox, 2012; Zhang et al., 2013], as well as, the cloud cover and lifetime [Koch and Del Genio, 2010]. Smoke and dust aerosols can change the air temperature profile and modify the atmospheric stability in the free troposphere [Babu et al., 2011; Wendisch et al., 2008], while elevated dense aerosol plumes such as airborne volcanic ash, are a significant hazard to aviation, causing engine failure, damage to avionics systems and abrasion to exposed airframe parts [Sears et al., 2013]. Aerosol vertical distribution is essential in retrieving aerosol optical depth in the UV channels [Torres et al., 1998], and can affect the retrieval accuracy of aerosol microphysical properties from photopolarimetric measurements [Chowdhary et al., 2005; Waquet et al., 2009], the atmospheric correction for ocean color remote sensing [Duforêt et al., 2007], the retrieval of the thermal state of the atmosphere from infrared sounders [Maddy et al., 2012], and the remote sensing of the surface concentration of particulate matters from space [Wang and Christopher, 2003].

Because aerosol profiles simulated by current climate models can differ by up to an order of magnitude [Kipling et al., 2016; Koffi et al., 2012], satellite remote sensing techniques are used to infer the spatial and temporal variations of aerosol profiles across the globe. Aerosol profiles can be recovered either by (accurate) active remote sensing techniques using LIDAR (such as CALIOP equipped on the CALIPSO platform [Winker et al., 2009]), or by passive remote sensing techniques, which offer a better spatial coverage and more frequent measurements.

## 1.1 Passive remote sensing techniques

Several retrieval algorithms have been developed for deriving aerosol profile information from passive satellite sensors. These were mostly developed based on heritages of the remote sensing of cloud altitude [Kokhanovsky and Rozanov, 2004; Loyola et al., 2007; Wang et al., 2008; Yang et al., 2013]. However, the retrieval of aerosol height is a much more challenging task, because aerosols are less optically thick and have more complex optical properties (depending on the size distribution and chemical composition). Moreover, the heterogeneity of the underlying surface increases the complexity of the retrieval.

The passive remote sensing techniques can be divided into two main categories:

## 1 Introduction

1. the stereoscopic technique based on the principle of parallax [Fisher et al., 2014; Muller et al., 2007; Zakšek et al., 2013], and
2. the spectroscopic techniques based on radiative transfer methods.

The spectroscopic technique includes three approaches:

1. the polarimetric analysis of reflected sunlight at the near-ultraviolet spectrum [Wu et al., 2016],
2. the spectroscopy in the oxygen ( $O_2$ ) absorption bands [Corradini and Cervino, 2006; Ding et al., 2016; Dubuisson et al., 2009; Kokhanovsky and Rozanov, 2010; Sanders et al., 2015] and the  $O_2 - O_2$  band [Chimot et al., 2017; Park et al., 2016], and
3. the thermal infrared spectral radiance technique [Pierangelo et al., 2004; Vandenburg et al., 2013].

The first approach uses measurements at wavelengths with negligible gas absorption in the context of air molecular scattering, while the second approach uses measurements in the A and B absorption bands of molecular oxygen located in the spectral ranges of 755–775 nm and 685–695 nm, respectively.

The retrieval in the oxygen bands is based on the fact that the aerosol layer scatters photons back to space and reduces the amount of photons captured by the  $O_2$  molecules. A higher scattering layer provides a larger number of photons scattered back to space, and so, enhances the reflectivity in the oxygen A-band. Moreover, because  $O_2$  is a well-mixed gas in the atmosphere with a well-defined vertical structure, the absorption profile of molecular  $O_2$  can be accurately described. Thus, the spectral dynamics of the reflectance in the oxygen A-band characterize the interaction of aerosol particles with oxygen absorption through multiple scattering in different altitudes.

From the measurement perspective, the retrieval algorithms for aerosol vertical distribution based on  $O_2$  spectroscopy can be divided into two categories.

1. Algorithms of the first category derive the aerosol height based on the reflectance ratio of channels inside and outside the absorption band [Dubuisson et al., 2009; Duforêt et al., 2007; Xu et al., 2017]. This approach is well suited for narrow-band measurements in the  $O_2$  absorption spectrum, such as those from MERIS (Medium Resolution Imaging Spectrometer), POLDER (POLarization and Directionality of the Earth's Reflectances), and EPIC (Earth Polychromatic Imaging Camera). However, the reflectance ratio approach can only yield a single piece of aerosol height information, and requires appropriate assumptions on aerosol optical properties and surface reflectance. The errors in the retrieved aerosol height are larger for brighter surfaces, because the surface reflection dominates the apparent reflectance at the top-of-atmosphere. Therefore, it is difficult to retrieve aerosol height over land, for which the reflectances of the vegetation and soil surface are often larger than 0.3. To overcome this drawback, adding polarization measurements (the degree of polarization is less sensitive to the surface reflectance) [Boesche et al., 2009; Ding et al., 2016; Wang et al., 2014] and/or measurements in the oxygen B-band (the reflectance of a vegetation surface is much lower in the oxygen B-band) [Pflug and Ruppert, 1993] have been proposed.

2. Algorithms of the second category are based on the spectral fitting of the reflectance measured in the O<sub>2</sub> absorption bands by spectrometers at a moderate to high spectral resolution. The studies of Koppers and Murtagh [Koppers and Murtagh, 1997] for GOME (Global Ozone Mapping Experiment) data, Kokhanovsky and Rozanov [Kokhanovsky and Rozanov, 2010] and Sanghavi et al. [Sanghavi et al., 2012] for SCIAMACHY (SCanning Imaging Absorption Spectrometer for Atmospheric CHartography) data, and Sanders et al. [Sanders et al., 2015] for GOME-2 and TROPOMI (TROPOspheric Monitoring Instrument) data, are relevant in this context. The spectrally resolved measurements contain more pieces of information than a single reflectance ratio. As shown by Corradini and Cervino [Corradini and Cervino, 2006], SCIAMACHY measurements (with a spectral resolution of about 0.4 nm) can provide a degree of freedom of 3 (in the troposphere, a three-layer profile can be retrieved), though the retrieval depends strongly on the aerosol optical properties and the surface reflectivity. Higher spectral resolution measurements in the oxygen A-band, such as those from GOSAT (Greenhouse gases Observing SATellite), OCO-2 (Orbiting Carbon Observatory), and TanSat (Chinese Carbon Dioxide Observation Satellite Mission) with a spectral resolution as high as 0.04 nm, can improve the aerosol height information content by reaching a degree of freedom of 4–5; thus, the aerosol profile can be recovered [Colosimo et al., 2016; Geddes and Bösch, 2015; Hollstein and Fischer, 2014]. Moreover, if multiangle measurements in the oxygen A-band are used, a degree of freedom of about 7 can be reached [Frankenberg et al., 2012]. Note that for such an enhanced spectral resolution, the influence of aerosol optical properties and surface albedo is significantly reduced [Colosimo et al., 2016].

The Table 1.1 presents the main characteristics of the satellite instruments that perform measurements in O<sub>2</sub> absorption bands.

The retrieval algorithms use various assumptions on the aerosol profile and models, and employ several fast radiative transfer models.

1. *Aerosol profile.* The vertical distribution of aerosols is determined by the emission and deposition processes, aerosol microphysical properties (size and composition), and meteorological conditions (wind, atmospheric stability, planetary boundary layer evolution, and precipitation). Anthropogenic aerosols in industrial pollution source regions and sea salts over oceans are usually located below 1 km in the planetary boundary layer, while elevated aerosol layers are found over the dust belt and source regions of biomass-burning particles (e.g., southern Africa and South America). In retrieval algorithms, the common assumption is that aerosol is homogeneously distributed within a layer of given depth [Kokhanovsky and Rozanov, 2010; Pierangelo et al., 2004], which can extend from the surface or be an aloft layer in the atmosphere. In this case, the retrieved parameter is the top altitude [Kokhanovsky and Rozanov, 2010] or the central altitude of the aerosol layer [Pierangelo et al., 2004]. In other studies, it is assumed that the profile of the aerosol extinction coefficient follows a certain type of distribution function, as for example, an exponential-decay profile characterized by a scale height [Gordon, 1997], and a Gaussian [Ding et al., 2016; Dubovik et al., 2011; Wu et al., 2016; Xu et al., 2017] or a log-normal distribution function [Hollstein and Fischer, 2014; Sanghavi et al., 2012] characterized by a peak height and a half-width parameter.

**Table 1.1:** Satellite instruments that perform measurements in O<sub>2</sub> absorption bands.

Instrument/ Satellite	Spectral Channels	Spectral Resolution	Spatial Resolution
MERIS/ ENVISAT	753.75 nm; 761.75 nm	7.5 nm; 3.75 nm	ocean:1040 × 1200 m <sup>2</sup> land:260 × 260 m <sup>2</sup>
POLDER	763 nm; 765 nm	10 nm; 40 nm	
EPIC/ DISCOVER	680 nm; 688 nm; 764 nm	3 ± 0.6 nm; 0.8 ± 0.2 nm 1 ± 0.2 nm	12 × 12 km <sup>2</sup>
GOME-2/ METOP-A/B/C	240 – 790 nm	0.2 – 0.5 nm	80 × 40 km <sup>2</sup>
SCIAMACHY/ ENVISAT	240 – 2380 nm	0.48 nm	30 × 60 km <sup>2</sup>
TANSO-FTS/ GOSAT	758 – 775 nm	0.5 cm <sup>-1</sup>	10.5 × 10.5 km <sup>2</sup>
OCO-2	757 – 778 nm	0.042 nm	1.3 × 2.25 km <sup>2</sup>
CarbonSpec/ TanSat	758 – 778 nm	0.044 nm	2 × 2 km <sup>2</sup>
TROPOMI/ Sentinel-5 Precursor	675 – 775 nm	0.38 nm	3.5 × 7 km <sup>2</sup>

2. *Aerosol models.* The aerosol microphysical properties (size and composition) play an important role in the retrieval. To describe the wide range of possible compositions, the aerosol particles are modeled as components, which are externally mixed to form aerosol models. An aerosol component is characterized through the refractive index and the size distribution. The size distribution is assumed to be a log-normal distribution, and is specified as a number or a volume distribution. These are summarized below.
- a) The Optical Properties of Aerosols and Clouds (OPAC) database [Hess et al., 1998] includes as aerosol components: the water-insoluble part of aerosol particles, the water-soluble part of aerosols, the soot component, sea-salt, mineral aerosol, mineral transported aerosol, and the sulfate component, and as aerosol models: continental clean, continental polluted, urban, desert, maritime clean, maritime polluted, maritime tropical, arctic, and antarctic.
  - b) Ten aerosol models (sulfurous dusty smoke, marine sulfate, dusty sulfate, dust, sulfurous smoke, sulfurous dust, marine dusty sulfate, sulfate, sulfurous marine, and smokey sulfate), considered to be mixtures of sulfate, dust, sea salt, black carbon, and organic carbon, have been obtained by a cluster analysis using the Goddard Chemistry Aerosol Radiation and Transport (GOCART) model [Taylor et al., 2015].
  - c) The OMI multiwavelength aerosol retrieval algorithm [Torres et al., 2002] includes five major aerosol types, where each type consists of several aerosol models. On a global scale, five main aerosol types can be distinguished: (i) urban-industrial aerosols originating from fossil fuel combustion, (ii) carbonaceous aerosols generated from natural and anthropogenic biomass burning, (iii) desert dust aerosols, injected into the atmosphere by winds, (iv) naturally produced oceanic aerosols, and (v) volcanic aerosol.
  - d) The MODIS aerosol retrieval algorithm [Levy et al., 2007] includes a set of aerosol models derived from a cluster analysis on the entire time series of almucantur-derived aerosol properties from global AERONET [Holben et al., 1998] sites. There are three fine-dominated (spherical) and one coarse-dominated (spheroid) aerosol optical models that represent the range of likely and observable global aerosol conditions. The fine-dominated aerosol models, i.e., moderately absorbing, absorbing, and nonabsorbing, differ mainly in their values of the single scattering albedo.
  - e) At selected AERONET [Dubovik et al., 2002] sites, with well-known meteorological and environmental conditions, four aerosol models have been identified: (i) urban-industrial from fossil fuel combustion in populated industrial regions, (ii) biomass burning produced by forest and grassland fires, (iii) desert dust blown into the atmosphere by wind, and (iv) aerosol of marine origin.
3. *Radiative transfer models.* A variety of fast radiative transfer models for computing the spectral signal in the O<sub>2</sub> absorption bands have been designed. They rely on the discrete ordinate method [Kokhanovsky and Rozanov, 2010; Spurr, 2006], matrix operator method [Hollstein and Fischer, 2014; Hollstein and Lindstrot, 2014], and a layer-based orders of scattering method (a variant of the doubling-adding method [de Haan et al., 1987; Hovenier et al., 2004] in which the adding of different layers is

## 1 Introduction

replaced by orders of scattering for the atmospheric layers), The derivatives with respect to the fit parameters are calculated analytically (linearized forward approach) [Spurr, 2006], or by using the adjoint theory (linearized forward-adjoint approach) [Landgraf et al., 2001]. In order to reduce the computation time, several acceleration methods have been proposed. These include the correlated k-distribution method [Hasekamp and Butz, 2008], the principal component analysis applied on optical parameters [Natraj et al., 2005], the interpolation method within lookup table for monochromatic reflectances [Sanders and de Haan, 2013] and their principal components [Hollstein and Lindstrot, 2014], and neural networks to emulate the top-of-atmosphere radiances [Nanda et al.]. The aerosol parameters are usually retrieved with the optimal estimation method, an iterative retrieval scheme developed by Rodgers [Rodgers, 2000] that incorporates a priori knowledge of retrieval parameters into their estimation.

## 1.2 Objective of the thesis

The main objective of the thesis is the design of retrieval algorithms to recover the aerosol layer height and optical depth from TROPOMI/S5P measurements in the O<sub>2</sub>A-band.

Launched on 13 October 2017, TROPOMI onboard the Copernicus Sentinel-5 Precursor satellite [Veefkind et al., 2012] is the first of the satellite-based atmospheric composition monitoring instruments in the Sentinel mission of the European Space Agency. It was designed to be a push-broom grating spectrometer observing trace gas concentrations and aerosol/cloud properties that are associated with air quality, ozone layer, and climate forcing. The satellite flies in a sun-synchronous orbit at 824 km altitude with an Equator crossing time of 13:30 local solar time, and has a wide swath of 108° (~2600 km), which allows to achieve a full daily global surface coverage. TROPOMI can map global distributions of a broad range of air pollutants with a spatial resolution as high as  $5.5 \times 3.5 \text{ km}^2$  ( $7.0 \times 3.5 \text{ km}^2$  prior to 6 August 2019). The recorded TROPOMI spectra cover the ultraviolet–visible (UV–Vis, 270–500 nm), near-infrared (NIR, 675–775 nm), and short-wave infrared (SWIR, 2305–2385 nm). Band 6 of TROPOMI covers the oxygen A-band and records the radiances and solar irradiances with a spectral sampling of 0.125–0.126 nm and a spectral resolution of 0.34–0.35 nm. Further details of the instrument and measurement characteristics can be found in Refs. [Kleipool et al., 2018; Ludewig et al., 2020]. TROPOMI records approximately 1.4 million pixels within a single orbit where, on average, 50 000 pixels are typically identified as aerosol contaminated pixels. This places a steep requirement on the computational architecture with respect to processing all possible pixels from a single orbit. The operational TROPOMI/S5P aerosol layer height retrieval algorithm has the following key features [de Graaf et al., 2021; Nanda et al.].

1. The main fit parameters are the aerosol layer mid pressure and aerosol optical depth.
2. A single, average aerosol model, characterized by a (fixed) single scattering albedo of 0.95 and a Henyey-Greenstein phase function with an asymmetry parameter of 0.7, is considered.
3. Monochromatic (high-resolution) reflectances are computed with the layer-based orders-of-scattering method. Polarization and inelastic scattering (rotational Raman scattering) are not taken into account.

4. Derivatives of reflectance with respect to the fit parameters are calculated in a semi-analytical manner using reciprocity (equivalent to the adjoint method).
5. Line-by-line monochromatic calculation of top-of-atmosphere reflectance, and its derivatives with respect to the aerosol layer height and optical depth, are performed at 3980 absorption lines in the oxygen A-band (wavelengths  $\sim 758 - 770$  nm). The computation time is of about 30-40 seconds on a computer equipped with Intel R Xeon R CPU E3-1275 v5 at a clock speed of 3.60 GHz.
6. A neural network is trained to learn the (sun-normalized) top-of-atmosphere radiance, and other two neural networks to emulate the derivatives of the reflectance with respect to the aerosol layer height and optical depth.
7. The inverse problem is solved by using a Bayesian approach with a priori chosen error covariance matrices for the measurement and the a priori state vector.

In order to achieve the objective of the thesis, the following tasks will be performed.

1. Design of an aerosol database that includes several aerosol models and vertical profiles.
2. Design of a linearized radiative transfer model,
  - a) relying on the discrete ordinate method with matrix exponential,
  - b) using as acceleration approaches: the method of false discrete ordinate, the correlated k-distribution method, and principal component analysis, and
  - c) computing the partial derivatives of the spectral signal with respect to the aerosol optical depth and layer height by forward and forward-adjoint approaches.
3. Design of Bayesian-based retrieval algorithms for a specified aerosol model and a set of candidate models, using the iteratively regularized Gauss-Newton method as regularization tool.
4. Design of neural network algorithms
  - a) to emulate the radiative transfer model and to be used in conjunction with a Bayesian approach to solve the inverse problem,
  - b) to learn the inverse model using as input either the synthetic radiances or their principal components.

### 1.3 Thesis structure

This is a cumulative thesis which includes three full peer-reviewed research papers, namely

1. Rao, L.; Xu, J.; Efremenko, D.S.; Loyola, D.G.; Doicu, A. Hyperspectral Satellite Remote Sensing of Aerosol Parameters. *Front. Environ. Sci.*, 14 January 2022 — <https://doi.org/10.3389/fenvs.2021.770662>
2. Rao, L.; Xu, J.; Efremenko, D.S.; Loyola, D.G.; Doicu, A. Optimization of Aerosol Model Selection for TROPOMI/S5P. *Remote Sens.* 2021, 13, 2489. <https://doi.org/10.3390/rs13132489>

## 1 Introduction

3. Rao, L.; Xu, J.; Efremenko, D.S.; Loyola, D.G.; Doicu, A. Aerosol Parameters Retrieval From TROPOMI/S5P Using Physics-Based Neural Networks. *IEEE J. Sel. Top. Appl. Earth Obs.* 2022, 15, 6473-6484, doi: 10.1109/JSTARS.2022.3196843

The organization of the thesis is as follows. Chapter 2 describes the aerosol database, Chapter 3 provides a description of the radiative transfer models, Chapter 4 presents the Bayesian-based retrieval algorithm, Chapter 5 describes the neural network algorithms, and Chapter 6 contains some concluding remarks.



## 2 Aerosols database

The aerosol particles are modeled as components, each of them representing an internal mixture of all chemical substances that have a similar origin. These components can be externally mixed to form aerosol models, where external mixture means that there is no physical or chemical interaction between particles of different components.

### 2.1 Aerosol components

The size distribution of an aerosol component is a log-normal distribution. The log-normal mode can be described by

1. the number size distribution

$$\frac{dN(a)}{d \ln a} = \frac{N_0}{\sqrt{2\pi}\sigma} \exp\left[-\frac{(\ln a - \ln a_{\text{mod}})^2}{2\sigma^2}\right], \quad (2.1)$$

where  $a_{\text{mod}}$  is the modal or median radius of the number size distribution,  $\sigma$  the standard deviation, and

$$N_0 = \int_0^\infty \frac{dN(a)}{d \ln a} d \ln a \quad (2.2)$$

the total number of particles (per cross section of the atmospheric column), or

2. the volume size distribution

$$\frac{dV(a)}{d \ln a} = \frac{V_0}{\sqrt{2\pi}\sigma} \exp\left[-\frac{(\ln a - \ln a_v)^2}{2\sigma^2}\right], \quad (2.3)$$

where

$$a_v = a_{\text{mod}} \exp(-3\sigma^2) \quad (2.4)$$

is the median radius of the volume size distribution and

$$V_0 = \int_0^\infty \frac{4\pi a^3}{3} \frac{dN(a)}{d \ln a} d \ln a = N_0 \frac{4\pi a_{\text{mod}}^3}{3} \exp(4.5\sigma^2) \quad (2.5)$$

the volume of particles (per cross section of the atmospheric column).

Thus, the size distribution of an aerosol component is specified by

1. the modal radius  $a_{\text{mod}}$ , the standard deviation  $\sigma$ , and the total number of particles  $N_0$  for a number size distribution, or
2. the median radius of the volume size distribution  $a_v$ , the standard deviation  $\sigma$ , and the volume of particles  $V_0$  for a volume size distribution.

When these parameters together with the wavelength-dependent refractive index  $m_{\text{aer}}$  are known, the scattering characteristics of an aerosol component, i.e., the size averaged extinction and scattering cross sections  $\overline{C}_{\text{ext}}$  and  $\overline{C}_{\text{sct}}$ , respectively, as well as, the size averaged expansion coefficients  $\overline{\chi}_n$  of the phase matrix

$$P(a, \cos \Theta) = \sum_{n=0}^{N_{\text{rank}}} \sqrt{\frac{2n+1}{2}} \chi_n(a) P_n(\cos \Theta),$$

where  $P_n$  are the Legendre polynomials and  $N_{\text{rank}}$  the maximum expansion order, can be computed. In particular, if  $p(a)$  is the probability density function associated to the number size distribution, i.e.,

$$p(a) = \frac{1}{N_0} \frac{dN(a)}{da}, \quad (2.6)$$

implying

$$p(a) da = \frac{1}{N_0} dN(a) = \frac{1}{N_0} \frac{dN(a)}{d \ln a} d \ln a, \quad (2.7)$$

the size averaged quantities are computed by using the relations

$$\overline{C}_{\text{ext}} = \int_{a_{\text{min}}}^{a_{\text{max}}} C_{\text{ext}}(a) p(a) da, \quad (2.8)$$

$$\overline{C}_{\text{sct}} = \int_{a_{\text{min}}}^{a_{\text{max}}} C_{\text{sct}}(a) p(a) da, \quad (2.9)$$

$$\overline{\chi}_n = \int_{a_{\text{min}}}^{a_{\text{max}}} \chi_n(a) p(a) da, \quad n \geq 0, \quad (2.10)$$

where  $a_{\text{min}}$  and  $a_{\text{max}}$  are the lower and upper bounds of the size distribution, and  $C_{\text{ext}}(a)$ ,  $C_{\text{sct}}(a)$ , and  $\chi_n(a)$  are computed by an electromagnetic scattering code for a spherical particle of radius  $a$ .

Two databases for aerosol components are implemented in the retrieval algorithm.

**Database 1** The database, consisting of the aerosol components included in the OPAC model [Hess et al., 1998], i.e.,

1. water-insoluble (soil particles with a certain amount of organic material),
2. water-soluble (sulfates, nitrates, and water-soluble substances),
3. soot (absorbing black carbon),
4. sea-salt (accumulated and coarse),
5. mineral (accumulated and coarse) or desert dust,
6. mineral transported (desert dust transported over long distances with a reduced amount of large particles), and
7. sulfate (also used as stratospheric background aerosol),

provides the values of  $\overline{C}_{\text{ext}}$ ,  $\overline{C}_{\text{sct}}$ ,  $\overline{\omega} = \overline{C}_{\text{sct}}/\overline{C}_{\text{ext}}$ , and  $\overline{\chi}_n$ , computed by using a Mie electromagnetic scattering code, for

1. each aerosol component,
2. a size distribution with  $a_{\min} = 0.005 \mu\text{m}$  and  $a_{\max} = 10.0 \mu\text{m}$ ,
3. a set of 61 values of the wavelength in the range  $0.250 \mu\text{m} - 40 \mu\text{m}$ , and
4. a set of 8 values of the relative humidity: 0.00, 0.70, 0.80, 0.90, 0.95, 0.98, and 0.99 (note that mineral and water-insoluble aerosols have no relative humidity induced swelling).

**Database 2** The database, including the aerosol components:

1. black carbon,
2. dust,
3. organic carbon,
4. sea salt, and
5. sulfate,

provides the values of  $\bar{C}_{\text{ext}}$ ,  $\bar{C}_{\text{sct}}$ ,  $\bar{\omega} = \bar{C}_{\text{sct}}/\bar{C}_{\text{ext}}$ , and  $\bar{\chi}_n$  for

1. each aerosol components,
2. size distributions with  $a_{\min} = 0.005 \mu\text{m}$  and  $a_{\max} = 0.3 \mu\text{m}$  for black carbon, organic carbon and sulfate,  $a_{\min} = 0.1 \mu\text{m}$  and  $a_{\max} = 10.0 \mu\text{m}$  for dust, and  $a_{\min} = 0.03 \mu\text{m}$  and  $a_{\max} = 10.3 \mu\text{m}$  for sea salt [Gong, 2003],
3. a set of 61 values of the wavelength in the range  $0.250 \mu\text{m} - 40 \mu\text{m}$ , and
4. a set of 36 values of the relative humidity: 0.00, 0.05,  $\dots$ , 0.80, 0.81, 0.82,  $\dots$ , 0.99.

Here, the aerosol single scattering properties are calculated by using the formalism presented in Ref. [Chin et al., 2002], the growth factors for organic carbon and sulfate, as well as, the spectral complex refractive index are taken from the OPAC database [Hess et al., 1998], the empirical relationship of Gerber [Gerber, 1985] is used for determining the growth factors of sea salt, and the scattering characteristics are computed by using a Mie electromagnetic scattering code, except for dust, where a precomputed database for ellipsoidal particles [Meng et al., 2010] is used.

## 2.2 Aerosol models

The aerosol components are externally mixed to form aerosol models. For an aerosol model consisting of  $N_c$  aerosol components, the extinction and scattering cross sections, the single scattering albedo, and the expansion coefficients of the phase function are given

by

$$C_{\text{ext}}^{\text{aer}} = \sum_{i=1}^{N_c} w_i \bar{C}_{\text{ext}}^{(i)}, \quad (2.11)$$

$$C_{\text{sct}}^{\text{aer}} = \sum_{i=1}^{N_c} w_i \bar{C}_{\text{sct}}^{(i)}, \quad (2.12)$$

$$\omega_{\text{aer}} = C_{\text{sct}}^{\text{aer}} / C_{\text{ext}}^{\text{aer}}, \quad (2.13)$$

$$\chi_n^{\text{aer}} = \frac{1}{C_{\text{sct}}^{\text{aer}}} \sum_{i=1}^{N_c} w_i \bar{C}_{\text{sct}}^{(i)} \bar{\chi}_n^{(i)}, \quad (2.14)$$

respectively, where the weight

$$w_i = \frac{N_0^{(i)}}{\sum_{i=1}^{N_c} N_0^{(i)}} \quad (2.15)$$

is the number mixing ratio, and  $\bar{C}_{\text{ext}}^{(i)}$ ,  $\bar{C}_{\text{sct}}^{(i)}$ ,  $\bar{\chi}_n^{(i)}$ , and  $N_0^{(i)}$  correspond to the  $i$ th aerosol component. In this context, the extinction and scattering coefficients are calculated respectively, as

$$\sigma_{\text{ext}}^{\text{aer}} = n C_{\text{ext}}^{\text{aer}}, \quad (2.16)$$

$$\sigma_{\text{sct}}^{\text{aer}} = \omega_{\text{aer}} \sigma_{\text{ext}}^{\text{aer}} = n C_{\text{sct}}^{\text{aer}}, \quad (2.17)$$

where

$$n = \frac{\sum_{i=1}^{N_c} N_0^{(i)}}{V} = \sum_{i=1}^{N_c} \frac{N_0^{(i)}}{V} = \sum_{i=1}^{N_c} n_i, \quad (2.18)$$

is the number density of the aerosol particles, and  $n_i = N_0^{(i)}/V$  the number density of the  $i$ th aerosol component. Note that Eq. (2.15) can be expressed in terms of  $n_i$  as

$$w_i = \frac{N_0^{(i)}}{\sum_{i=1}^{N_c} N_0^{(i)}} = \frac{n_i}{n}. \quad (2.19)$$

By combining the aerosol components from the databases described in Section 2.1, different aerosol models can be defined. Using the optical parameters  $\bar{C}_{\text{ext}}^{(i)}$ ,  $\bar{C}_{\text{sct}}^{(i)}$ , and  $\bar{\chi}_n^{(i)}$  taken from one of the two databases, and specifying the weight  $w_i$  of each aerosol component  $i$ , the optical parameters of an aerosol model can be computed by means of Eqs. (2.11)–(2.14). For the second database, aerosol mixtures of sulfate, dust, sea salt, black carbon, and organic carbon, have been obtained by a cluster analysis using the Goddard Chemistry Aerosol Radiation and Transport (GOCART) model [Taylor et al., 2015]. The resulting aerosol models, i.e., smoke (sulfurous dusty, sulfurous), sulfate (marine, dusty, marine dusty, smokey), dust, and marine, together with the weights of their components (expressed in percent) are given in Table 2.1.

The aerosol databases found in the literature use different predefined aerosol models. In these databases, the aerosol components of each model are specified together with their microphysical properties, that is,

**Table 2.1:** Aerosol mixtures of sulfate (SS), dust (DU), sea salt (SS), black carbon (BC), and organic carbon (OC) obtained by a cluster analysis using the Goddard Chemistry Aerosol Radiation and Transport (GOCART) model.

Model	Component				
	SU	DU	SS	BC	OC
Sulfurous dusty SMOKE	27.4	30.7	5.9	5.9	30.1
Marine SULFATE	44.6	4.7	36.7	3.0	11.0
Dusty SULFATE	54.7	25.6	7.2	3.4	9.1
DUST	13.0	80.2	1.1	1.7	4.0
Sulfurous SMOKE	29.7	6.0	3.1	9.3	51.8
Sulfurous DUST	31.0	53.1	4.7	3.2	8.0
Marine dusty SULFATE	43.1	27.0	19.5	3.2	7.2
SULFATE	66.1	4.7	14.1	3.4	11.6
Sulfurous MARINE	28.8	3.8	58.4	1.7	7.3
Smokey SULFATE	45.0	6.8	14.0	6.7	27.5

1. size distribution parameters, e.g., the modal radius  $a_{\text{mod}}$ , the standard deviation  $\sigma$ , and the total number of particles  $N_0$ ,
2. the number mixing ratio  $w_i$ , and
3. the wavelength-dependent refractive index  $m_{\text{aer}}$ .

In the retrieval algorithm we included the aerosol models used in the OPAC database [Hess et al., 1998], the OMI multiwavelength aerosol retrieval algorithm [Torres et al., 2002], and the MODIS aerosol retrieval algorithm [Levy et al., 2007]. These are listed in Tables 2.2–2.4. The computational process is organized as follows:

1. knowing the microphysical properties of the aerosol components and using a Mie electromagnetic scattering code, we compute the optical parameters  $\overline{C}_{\text{ext}}^{(i)}$ ,  $\overline{C}_{\text{sct}}^{(i)}$ , and  $\overline{\chi}_n^{(i)}$  of each component  $i$  according to Eqs. (2.8)–(2.10), and then
2. we calculate the optical parameters  $\overline{C}_{\text{ext}}^{\text{aer}}$ ,  $\overline{C}_{\text{sct}}^{\text{aer}}$ , and  $\overline{\chi}_n^{\text{aer}}$  of the aerosol model by means of Eqs. (2.11)–(2.14).

It should be pointed out that the aerosol models included in the MODIS aerosol retrieval algorithm have dynamic (function of optical depth  $\tau_{\text{aer}}$ ) size parameters (radius, standard deviation, volume distribution) and complex refractive index. In this case, we compute  $\overline{C}_{\text{ext}}^{\text{aer}}$ ,  $\overline{C}_{\text{sct}}^{\text{aer}}$ , and  $\overline{\chi}_n^{\text{aer}}$  at a ten equidistant values of  $\log \tau_{\text{aer}}$ , with  $\tau_{\text{aer}}$  ranging from 0.01 to 2.0, and linearly interpolate on this data set during the retrieval.

## 2.3 Aerosol vertical profiles

To provide the aerosol optical depths, the altitude profile of the particle number density  $n(r)$  have to be specified. The number density is written as

$$n(r) = n_0 f(r, H_{\text{aer}}),$$

where  $n_0$  is the maximum value of the number density,  $f(r, H_{\text{aer}}) \leq 1$  the indicator function of the aerosols, and  $H_{\text{aer}}$  a parameter characterizing the altitude profile. For the indicator function, we have several choices, as for example,

**Table 2.2:** Table 3: Aerosol models from OPAC database. The relative refractive index corresponds to  $\lambda = 750$  nm and  $U = 0.8$ .

Model	Component	$r_{\text{mod}}(\mu\text{m})$	$s = e^\sigma$	$m = (\text{Re}(m), \text{Im}(m))$	$w_i$	$N_0(\text{total})$ ( $\text{cm}^{-3}$ )
Cont. clean	Water sol.	0.0212	2.24	$(1.40, 2.83 \times 10^{-3})$	1.0	2600
	Water insol.	0.4710	2.51	$(1.53, 8.0 \times 10^{-3})$	$0.577 \times 10^{-4}$	
Cont. avg.	Water sol.	0.0212	2.24	$(1.40, 2.83 \times 10^{-3})$	0.95	15300
	Soot	0.0118	2.00	$(1.75, 4.3 \times 10^{-1})$	0.05	
	Water insol.	0.4710	2.51	$(1.53, 8.0 \times 10^{-3})$	$0.261 \times 10^{-4}$	
Cont. pol.	Water sol.	0.0212	2.24	$(1.40, 2.83 \times 10^{-3})$	0.90	50000
	Soot	0.0118	2.00	$(1.75, 4.3 \times 10^{-1})$	0.10	
	Water insol.	0.4710	2.51	$(1.53, 8.0 \times 10^{-3})$	$0.12 \times 10^{-4}$	
Urban	Water sol.	0.0212	2.24	$(1.40, 2.83 \times 10^{-3})$	0.80	158000
	Soot	0.0118	2.00	$(1.75, 4.3 \times 10^{-1})$	0.20	
	Water insol.	0.4710	2.51	$(1.53, 8.0 \times 10^{-3})$	$0.949 \times 10^{-5}$	
Desert	Water sol.	0.0212	2.24	$(1.40, 2.83 \times 10^{-3})$	0.87	2300
	Mineral nuc.	0.0700	1.95	$(1.53, 4.00 \times 10^{-3})$	0.117	
	Mineral acc.	0.3900	2.00	$(1.53, 4.00 \times 10^{-3})$	$0.133 \times 10^{-1}$	
	Mineral coa.	1.9000	2.15	$(1.53, 4.00 \times 10^{-3})$	$0.617 \times 10^{-4}$	
Marit. clean	Water sol.	0.0212	2.24	$(1.40, 2.83 \times 10^{-3})$	0.987	1520
	Sea salt acc.	0.2090	2.03	$(1.35, 2.73 \times 10^{-7})$	$0.132 \times 10^{-1}$	
	Sea salt coa.	1.7500	2.03	$(1.35, 2.72 \times 10^{-7})$	$0.211 \times 10^{-5}$	
Marit. pol.	Water sol.	0.0212	2.24	$(1.40, 2.83 \times 10^{-3})$	0.422	9000
	Soot	0.0118	2.00	$(1.75, 4.3 \times 10^{-1})$	0.576	
	Sea salt acc.	0.2090	2.03	$(1.35, 2.73 \times 10^{-7})$	$0.222 \times 10^{-2}$	
	Sea salt coa.	1.7500	2.03	$(1.35, 2.72 \times 10^{-7})$	$0.356 \times 10^{-6}$	
Marit. trop.	Water sol.	0.0212	2.24	$(1.40, 2.83 \times 10^{-3})$	0.983	600
	Sea salt acc.	0.2090	2.03	$(1.35, 2.73 \times 10^{-7})$	$0.167 \times 10^{-1}$	
	Sea salt coa.	1.7500	2.03	$(1.35, 2.72 \times 10^{-7})$	$0.217 \times 10^{-5}$	
Arctic	Water sol.	0.0212	2.24	$(1.40, 2.83 \times 10^{-3})$	0.197	6600
	Soot	0.0118	2.00	$(1.75, 4.3 \times 10^{-1})$	0.803	
	Water insol.	0.4710	2.51	$(1.53, 8.0 \times 10^{-3})$	$0.152 \times 10^{-5}$	
	Sea salt acc.	0.2090	2.03	$(1.35, 2.73 \times 10^{-7})$	$0.288 \times 10^{-3}$	
Antarctic	Sulfate	0.0695	2.03	$(1.35, 1.39 \times 10^{-7})$	0.998	43
	Sea salt acc.	0.2090	2.03	$(1.35, 2.73 \times 10^{-7})$	$0.109 \times 10^{-2}$	
	Mineral tra.	0.5000	2.20	$(1.530, 4.0 \times 10^{-3})$	$0.123 \times 10^{-3}$	

**Table 2.3:** Aerosol models assumed in the OMI multiwavelength aerosol retrieval algorithm. For each model, the first and second lines are related to the accumulated and coarse modes, respectively.

Model	Sub-Model	$r_{\text{mod}}(\mu\text{m})$	$s = e^\sigma$	$m = (\text{Re}(m), \text{Im}(m))$	$w_{\text{coarse}}$
Weakly absorbing	WA1101	0.071	1.499	$(1.4, 5.0 \times 10^{-8})$	$4.36 \times 10^{-4}$
		0.369	2.160		
	WA1102	0.081	1.499	$(1.4, 5.0 \times 10^{-8})$	$4.04 \times 10^{-4}$
		0.378	2.160		
	WA1103	0.126	1.499	$(1.4, 5.0 \times 10^{-8})$	$8.10 \times 10^{-4}$
		0.421	2.160		
	WA1104	0.023	2.030	$(1.4, 5.0 \times 10^{-8})$	$1.53 \times 10^{-2}$
		0.186	2.030		
	WA1201	0.071	1.499	$(1.4, 4.0 \times 10^{-3})$	$4.36 \times 10^{-4}$
		0.369	2.160		
	WA1202	0.081	1.499	$(1.4, 4.0 \times 10^{-3})$	$4.04 \times 10^{-4}$
		0.378	2.160		
WA1203	0.126	1.499	$(1.4, 4.0 \times 10^{-3})$	$8.10 \times 10^{-4}$	
	0.421	2.160			
WA1301	0.071	1.499	$(1.4, 1.2 \times 10^{-2})$	$4.36 \times 10^{-4}$	
	0.369	2.160			
WA1302	0.081	1.499	$(1.4, 1.2 \times 10^{-2})$	$4.04 \times 10^{-4}$	
	0.378	2.160			
WA1303	0.126	1.499	$(1.4, 1.2 \times 10^{-2})$	$8.10 \times 10^{-4}$	
	0.421	2.160			
Biomass burning (carbonaceous)	BB2101	0.067	1.537	$(1.5, 1.0 \times 10^{-2})$	$1.70 \times 10^{-4}$
		0.374	2.203		
	BB2102	0.079	1.537	$(1.5, 1.0 \times 10^{-2})$	$2.06 \times 10^{-4}$
		0.415	2.203		
	BB2103	0.113	1.537	$(1.5, 1.0 \times 10^{-2})$	$2.94 \times 10^{-4}$
		0.526	2.203		
	BB2201	0.067	1.537	$(1.5, 2.0 \times 10^{-2})$	$1.70 \times 10^{-4}$
		0.374	2.203		
	BB2202	0.079	1.537	$(1.5, 2.0 \times 10^{-2})$	$2.06 \times 10^{-4}$
		0.415	2.203		
	BB2203	0.113	1.537	$(1.5, 2.0 \times 10^{-2})$	$2.94 \times 10^{-4}$
		0.526	2.203		
BB2301	0.067	1.537	$(1.5, 3.0 \times 10^{-2})$	$1.70 \times 10^{-4}$	
	0.374	2.203			
BB2302	0.079	1.537	$(1.5, 3.0 \times 10^{-2})$	$2.06 \times 10^{-4}$	
	0.415	2.203			
BB2303	0.113	1.537	$(1.5, 3.0 \times 10^{-2})$	$2.94 \times 10^{-4}$	
	0.526	2.203			
Desert dust (minerals)	DD3101	0.036	1.697	$(1.53, 4.0 \times 10^{-3})$	$4.35 \times 10^{-3}$
		0.562	1.806		
	DD3102	0.045	1.697	$(1.53, 4.0 \times 10^{-3})$	$4.35 \times 10^{-3}$
		0.562	1.806		
DD3201	0.036	1.697	$(1.53, 1.0 \times 10^{-2})$	$4.35 \times 10^{-3}$	
	0.562	1.806			
DD3202	0.036	1.697	$(1.53, 1.0 \times 10^{-2})$	$4.35 \times 10^{-3}$	
	0.562	1.806			
Maritime	Maritime mod. abs.	0.023	2.030	$(1.4, 4.0 \times 10^{-3})$	$1.55 \times 10^{-4}$
		0.186	2.030		
	Maritime abs.	0.023	2.030	$(1.4, 1.2 \times 10^{-2})$	$1.55 \times 10^{-4}$
		0.186	2.030		
Volcanic	VO4101	0.224	0.800	$(1.45, 7.5 \times 10^{-7})$	0.5
		0.224	0.800		

**Table 2.4:** Aerosol models considered in the MODIS aerosol retrieval algorithm. For each model, the first and second lines are related to the accumulated and coarse modes, respectively. The four values of the refractive index for dust, correspond to the wavelengths  $\lambda = 0.470, 0.550, 0.660, 2.100 \mu\text{m}$ .

Model	$r_v(\mu\text{m})$	$\sigma$	$m = (\text{Re}(m), \text{Im}(m))$	$V_0(\mu\text{m}^3/\mu\text{m}^2)$
Nonabs.	$0.160 + 0.0434\tau_{\text{aer}}$	$0.364 + 0.1529\tau_{\text{aer}}$	$(1.42, 0.004 - 0.0015\tau_{\text{aer}})$	$0.1718\tau_{\text{aer}}^{-0.821}$
	$3.325 + 0.1411\tau_{\text{aer}}$	$0.759 + 0.0168\tau_{\text{aer}}$		$0.0934\tau_{\text{aer}}^{-0.639}$
Modabs.	$0.145 + 0.0203\tau_{\text{aer}}$	$0.374 + 0.1365\tau_{\text{aer}}$	$(1.43, 0.008 - 0.002\tau_{\text{aer}})$	$0.1642\tau_{\text{aer}}^{-0.775}$
	$3.101 + 0.3364\tau_{\text{aer}}$	$0.729 + 0.098\tau_{\text{aer}}$		$0.1482\tau_{\text{aer}}^{-0.684}$
Abs.	$0.134 + 0.0096\tau_{\text{aer}}$	$0.383 + 0.0794\tau_{\text{aer}}$	$(1.51, 0.02)$	$0.1748\tau_{\text{aer}}^{-0.891}$
	$3.448 + 0.9489\tau_{\text{aer}}$	$0.743 + 0.0409\tau_{\text{aer}}$		$0.1043\tau_{\text{aer}}^{-0.682}$
Dust	$0.1416\tau_{\text{aer}}^{-0.052}$	$0.7561\tau_{\text{aer}}^{0.148}$	$(1.48\tau_{\text{aer}}^{-0.021}, 0.0025\tau_{\text{aer}}^{0.132})$ $(1.48\tau_{\text{aer}}^{-0.021}, 0.002)$	$0.0871\tau_{\text{aer}}^{-1.026}$
	2.2	$0.554\tau_{\text{aer}}^{-0.052}$		$(1.48\tau_{\text{aer}}^{-0.021}, 0.0018\tau_{\text{aer}}^{-0.08})$ $(1.46\tau_{\text{aer}}^{-0.021}, 0.0018\tau_{\text{aer}}^{-0.30})$



1. a ground box profile

$$f(r, H_{\text{aer}}) = \begin{cases} 1, & 0 \leq r \leq H_{\text{aer}} \\ 0, & \text{otherwise} \end{cases}, \quad (2.20)$$

2. an elevated box profile with a fixed thickness of 0.5 km ( $H_{\text{aer}} \geq 0.25$  km, treated as a ground box profile when  $H_{\text{aer}} < 0.25$  km)

$$f(r, H_{\text{aer}}) = \begin{cases} 1, & H_{\text{aer}} - 0.25 \text{ km} \leq r \leq H_{\text{aer}} + 0.25 \text{ km} \\ 0, & \text{otherwise} \end{cases}, \quad (2.21)$$

3. a Gaussian profile

$$f(r, H_{\text{aer}}) = \exp[-(r - H_{\text{aer}})^2 / (2R)^2], \quad (2.22)$$

where  $R$  is the profile standard deviation,

4. an exponential profile

$$f(r, H_{\text{aer}}) = \exp[-(r - H_{\text{aer}})/R], \quad (2.23)$$

where  $R$  is the scale height, and

5. a combination of a ground box and an exponential profile

$$f(r, H_{\text{aer}}) = \begin{cases} 1, & 0 \leq r \leq H_{\text{aer}} \\ \exp[-(r - H_{\text{aer}})/R], & \text{otherwise} \end{cases}. \quad (2.24)$$

As a result, the extinction coefficient profile and the optical depth of aerosols are given, respectively, by

$$\sigma_{\text{ext}}^{\text{aer}}(r) = f(r, H_{\text{aer}})n_0C_{\text{ext}}^{\text{aer}}$$

and

$$\tau_{\text{aer}} = \int_0^{\infty} \sigma_{\text{ext}}^{\text{aer}}(r)dr = n_0C_{\text{ext}}^{\text{aer}}F(H_{\text{aer}}),$$

where  $F(H_{\text{aer}}) = \int_0^{\infty} f(r, H_{\text{aer}})dr$ . It should be pointed out that if at some iteration step of the inversion process,  $\tau_{\text{aer}}$  and  $H_{\text{aer}}$  are known, a radiative transfer calculation is performed for the altitude extinction and scattering profiles

$$\sigma_{\text{ext}}^{\text{aer}}(r) = f(r, H_{\text{aer}})\sigma_{\text{ext}0}^{\text{aer}} \text{ and } \sigma_{\text{sct}}^{\text{aer}}(r) = \omega_{\text{aer}}\sigma_{\text{ext}}^{\text{aer}}(r),$$

respectively, where the maximum value of the extinction coefficient is computed as

$$\sigma_{\text{ext}0}^{\text{aer}} = n_0C_{\text{ext}}^{\text{aer}} = \tau_{\text{aer}}/F(H_{\text{aer}}).$$



## 3 Radiative transfer model

In atmospheric remote sensing, the state parameters  $\mathbf{x}$ , are related to collected measurements making up a set of data  $\mathbf{y}$  through a forward model  $\mathbf{F}$ , that is,

$$\mathbf{y} = \mathbf{F}(\mathbf{x}).$$

The task of computing the data  $\mathbf{y}$  given the state parameters  $\mathbf{x}$  is called the forward problem, while the task of computing  $\mathbf{x}$  given  $\mathbf{y}$  is called the inverse problem. The forward problem reflects the physics of the measurements and deals with the computation of the spectral signal measured by an instrument by using a radiative transfer model. In contrast, the inverse problem deals with the retrieval of some atmospheric parameters (given the measurement) by using a mathematical approach, which is to some extent independent of the physical process. Because the solution of the inverse problem requires the knowledge of the forward model function  $\mathbf{F}(\mathbf{x})$  and its Jacobian  $\mathbf{K}(\mathbf{x}) = \partial\mathbf{F}(\mathbf{x})/\partial\mathbf{x}$ , the radiative transfer model should not only deliver the spectral signal measured by an instrument but also its derivative with respect to an atmospheric parameter of interest.

In this section we present the linearized radiative transfer model which is used for aerosol retrieval.

### 3.1 Radiative transfer equation

In a pseudo-spherical atmosphere, the radiative transfer equation for the diffuse radiance  $I(r, \boldsymbol{\Omega})$  at point  $r$  in direction  $\boldsymbol{\Omega} = (\mu, \varphi)$  is given by

$$\begin{aligned} \mu \frac{dI}{dr}(r, \boldsymbol{\Omega}) &= -\sigma_{\text{ext}}(r)I(r, \boldsymbol{\Omega}) + F_0 \frac{\sigma_{\text{sct}}(r)}{4\pi} P(r, \boldsymbol{\Omega}, \boldsymbol{\Omega}_0) e^{-\tau_{\text{ext}}^0(|\mathbf{r} - \mathbf{r}_{\text{TOA}}|)} \\ &+ \frac{\sigma_{\text{sct}}(r)}{4\pi} \int_{4\pi} P(r, \boldsymbol{\Omega}, \boldsymbol{\Omega}') I(r, \boldsymbol{\Omega}') d\Omega', \end{aligned} \quad (3.1)$$

where,  $\sigma_{\text{ext}}$  and  $\sigma_{\text{sct}}$  are the extinction and scattering coefficients, respectively,  $F_0$  is the incident solar flux,  $P$  the scattering phase function,  $\boldsymbol{\Omega}_0 = (-\mu_0, \varphi_0)$  with  $\mu_0 > 0$  the incident solar direction, and  $\tau_{\text{ext}}^0(|\mathbf{r} - \mathbf{r}_{\text{TOA}}|)$  the solar optical depth between a generic point  $\mathbf{r}$  and the characteristic point at the top of the atmosphere  $\mathbf{r}_{\text{TOA}}$  in a spherical atmosphere. The boundary conditions associated to the radiative transfer equation (3.1) are (i) the top-of-atmosphere boundary condition ( $r = r_{\text{TOA}}$ ),

$$I(r_{\text{TOA}}, \boldsymbol{\Omega}^-) = 0, \quad (3.2)$$

and (ii) the surface boundary condition ( $r = r_s$ ),

$$\begin{aligned} I(r_s, \boldsymbol{\Omega}^+) &= F_0 \frac{A}{\pi} \mu_0 \rho(\boldsymbol{\Omega}^+, \boldsymbol{\Omega}_0) e^{-\tau_{\text{ext}}^0(|\mathbf{r}_s - \mathbf{r}_{\text{TOA}}|)} \\ &+ \frac{A}{\pi} \int_{2\pi} I(r_s, \boldsymbol{\Omega}^-) |\mu^-| \rho(\boldsymbol{\Omega}^+, \boldsymbol{\Omega}^-) d\boldsymbol{\Omega}^-, \end{aligned} \quad (3.3)$$

### 3 Radiative transfer model

where  $A$  and  $\rho$  are the surface albedo and the normalized bi-directional reflection function, respectively, and the notations  $\mathbf{\Omega}^+$  and  $\mathbf{\Omega}^-$  stand for an upward and a downward direction, respectively.

The phase function  $P$  is expanded in terms of normalized Legendre polynomials  $P_n$ , i.e.,

$$P(r, \mathbf{\Omega}, \mathbf{\Omega}') = P(r, \cos \Theta) = \sum_{n=0}^{\infty} c_n \chi_n(r) P_n(\cos \Theta),$$

where  $\chi_n$  are the expansion coefficients,  $c_n = \sqrt{(2n+1)/2}$ , and  $\cos \Theta = \mathbf{\Omega} \cdot \mathbf{\Omega}'$ . In the case of an atmosphere consisting of  $N_g$  gas molecules and aerosols, the extinction and scattering coefficients are computed respectively, as

$$\begin{aligned} \sigma_{\text{ext}}(r) &= \sigma_{\text{ext}}^{\text{aer}}(r) + \sigma_{\text{sct}}^{\text{mol}}(r) + \sum_{g=1}^{N_g} \sigma_{\text{abs}g}^{\text{gas}}(r), \\ \sigma_{\text{sct}}(r) &= \sigma_{\text{sct}}^{\text{aer}}(r) + \sigma_{\text{sct}}^{\text{mol}}(r), \end{aligned}$$

and the phase function as

$$\sigma_{\text{sct}}(r) P(r, \mathbf{\Omega}, \mathbf{\Omega}') = \sigma_{\text{sct}}^{\text{mol}}(r) P_{\text{mol}}(\mathbf{\Omega}, \mathbf{\Omega}') + \sigma_{\text{sct}}^{\text{aer}}(r) P_{\text{aer}}(\mathbf{\Omega}, \mathbf{\Omega}'),$$

where  $\sigma_{\text{ext}}^{\text{aer}}$ ,  $\sigma_{\text{sct}}^{\text{aer}}$ , and  $P_{\text{aer}}$  are the extinction coefficient, scattering coefficient, and the phase function of the aerosols, respectively,  $\sigma_{\text{sct}}^{\text{mol}}$  is the molecular scattering coefficient due to Rayleigh scattering,  $\sigma_{\text{abs}g}^{\text{gas}}$  the absorption coefficient of gas  $g$ , and  $P_{\text{mol}}$  the effective scattering phase function accounting for molecular scattering. For a specified aerosol model and under the assumption that the aerosol optical depth  $\tau_{\text{aer}}$  and the aerosol indicator function  $f(r, H_{\text{aer}})$  are known,

1. the expansion coefficients  $\chi_n^{\text{aer}}$  are computed by means of Eqs. (2.14), and
2. the aerosol extinction and scattering coefficients are calculated as  $\sigma_{\text{ext}}^{\text{aer}}(r) = f(r, H_{\text{aer}}) \sigma_{\text{ext}0}^{\text{aer}}$  and  $\sigma_{\text{sct}}^{\text{aer}}(r) = \omega_{\text{aer}} \sigma_{\text{ext}}^{\text{aer}}(r)$ , respectively, where  $\sigma_{\text{ext}0}^{\text{aer}} = \tau_{\text{aer}} / F(H_{\text{aer}})$  with  $F(H_{\text{aer}}) = \int_0^{\infty} f(r, H_{\text{aer}}) dr$ , is the maximum value of the extinction coefficient, and  $\omega_{\text{aer}}$ , computed by means of Eq. (2.13), is the aerosol single scattering albedo.

In the discrete ordinate method, we consider a cosine-azimuth expansion of the diffuse radiance ( $\varphi_0 = 0$ ),

$$I(r, \mathbf{\Omega}) = \sum_{m=0}^M I_m(r, \mu) \cos m\varphi,$$

and the phase function

$$P(r, \mathbf{\Omega}, \mathbf{\Omega}') = \sum_{m=0}^M (2 - \delta_{m0}) p_m(r, \mu, \mu') \cos[m(\varphi - \varphi')], \quad (3.4)$$

$$p_m(r, \mu, \mu') = \sum_{n=m}^M \chi_n(r) P_n^m(\mu) P_n^m(\mu'), \quad (3.5)$$

and discretize the radiative transfer equation in the angular domain by considering a set of Gauss-Legendre quadrature points and weights  $\{\mu_k, w_k\}_{k=1}^{N_{\text{DO}}}$  in the interval  $(0, 1)$ . For each azimuth mode  $m$ , we obtain the radiative transfer equation

$$\begin{aligned}
 & \pm\mu_k \frac{dI_m}{dr}(r, \pm\mu_k) \\
 & = -\sigma_{\text{ext}}(r)I_m(r, \pm\mu_k) \\
 & + (2 - \delta_{m0}) \frac{F_0}{4\pi} \sigma_{\text{sct}}(r) p_m(r, \pm\mu_k, -\mu_0) e^{-\tau_{\text{ext}}^0(|\mathbf{r}_s - \mathbf{r}_{\text{TOA}}|)} \\
 & + \frac{1}{2} \sum_{l=1}^{N_{\text{DO}}} w_l \sigma_{\text{sct}}(r) [p_m(r, \pm\mu_k, \mu_l) I_m(r, \mu_l) + p_m(r, \pm\mu_k, -\mu_l) I_m(r, -\mu_l)],
 \end{aligned} \tag{3.6}$$

where  $N_{\text{DO}}$  is the number of discrete ordinates per hemisphere and  $M = 2N_{\text{DO}} - 1$  the number of azimuth modes. Further, we discretize the atmosphere in  $N_{\text{lev}}$  levels:  $r_1 = r_{\text{TOA}} > r_2 > \dots > r_{N_{\text{lev}}} = r_s$ . The number of layers is  $N_{\text{lay}} = N_{\text{lev}} - 1$ , and a layer  $j$ , bounded above by the level  $r_j$  and below by the level  $r_{j+1}$ , has the geometrical thickness  $\Delta r_j = r_j - r_{j+1}$ . The extinction and scattering coefficients  $\sigma_{\text{ext}}(r)$  and  $\sigma_{\text{sct}}(r)$ , respectively, as well as the phase function coefficients  $\chi_n(r)$  and the azimuth phase function coefficients  $p_m(r, \mu, \mu')$  are assumed to be constant within each layer; their average values in layer  $j$  are  $\sigma_{\text{ext}j}$ ,  $\sigma_{\text{sct}j}$ ,  $\chi_{nj}$ , and  $p_{mj}(\mu, \mu') = \sum_{n=m}^M \chi_{nj} P_n^m(\mu) P_n^m(\mu')$ , respectively. By the discretization process we are led to the following linear system of differential equations in layer  $j$ ,

$$\frac{d\mathbf{i}_m}{dr}(r) = \mathbf{A}_{mj} \mathbf{i}_m(r) + e^{-\tau_{\text{ext}}^0(|\mathbf{r} - \mathbf{r}_{\text{TOA}}|)} \mathbf{b}_{mj}, \quad r_{j+1} \leq r \leq r_j, \tag{3.7}$$

where  $\mathbf{i}_m(r) = [\mathbf{i}_m^+(r); \mathbf{i}_m^-(r)]^T$  is the radiance vector in the discrete ordinate space,  $[\mathbf{i}_m^\pm(r)]_k = I_m(r, \pm\mu_k)$ ,  $k = 1, \dots, N_{\text{DO}}$ ,

$$\mathbf{A}_{mj} = \begin{bmatrix} \mathbf{A}_{mj}^{11} & \mathbf{A}_{mj}^{12} \\ -\mathbf{A}_{mj}^{12} & -\mathbf{A}_{mj}^{11} \end{bmatrix}, \tag{3.8}$$

with

$$[\mathbf{A}_{mj}^{11}]_{kl} = \frac{1}{2\mu_k} [w_l \sigma_{\text{sct}j} p_{mj}(\mu_k, \mu_l) - 2\sigma_{\text{ext}j} \delta_{kl}], \tag{3.9}$$

$$[\mathbf{A}_{mj}^{12}]_{kl} = \frac{1}{2\mu_k} w_l \sigma_{\text{sct}j} p_{mj}(\mu_k, -\mu_l), \tag{3.10}$$

the layer matrix, and  $\mathbf{b}_{mj} = [\mathbf{b}_{mj}^+; \mathbf{b}_{mj}^-]^T$  with

$$[\mathbf{b}_{mj}^\pm]_k = \pm \frac{1}{\mu_k} (2 - \delta_{m0}) \frac{F_0}{4\pi} \sigma_{\text{sct}j} p_{mj}(\pm\mu_k, -\mu_0),$$

the layer vector.

### 3 Radiative transfer model

For a practical implementation, it is useful to expand the product  $\sigma_{\text{sct}}P$  as (cf. Eqs. (3.4) and (3.5))

$$\sigma_{\text{sct}}(r)P(r, \boldsymbol{\Omega}, \boldsymbol{\Omega}') = \sum_{m=0}^M (2 - \delta_{m0}) s_m(r, \mu, \mu') \cos[m(\varphi - \varphi')], \quad (3.11)$$

$$\begin{aligned} s_m(r, \mu, \mu') &= \sigma_{\text{sct}}(r) p_m(r, \mu, \mu') \\ &= \sum_{n=m}^M \xi_n(r) P_n^m(\mu) P_n^m(\mu'), \end{aligned} \quad (3.12)$$

$$\xi_n(r) = \sigma_{\text{sct}}(r) \chi_n(r), \quad (3.13)$$

in which case, the product  $s_{mj}(\mu, \mu') = \sigma_{\text{sct}j} p_{mj}(\mu, \mu')$  in Eqs. (3.9) and (3.10) is computed as

$$s_{mj}(\mu, \mu') = \sigma_{\text{sct}j} p_{mj}(\mu, \mu') = \sum_{n=m}^M \xi_{nj} P_n^m(\mu) P_n^m(\mu'), \quad (3.14)$$

$$\xi_{nj} = \sigma_{\text{sct}j} \chi_{nj}. \quad (3.15)$$

By this construction, the optical input parameters of the radiative transfer model are  $\sigma_{\text{ext}j}$  and  $\xi_{nj}$  for  $j = 1, \dots, N_{\text{lay}}$ , and  $n = 0, \dots, M$ .

## 3.2 Discrete ordinate method with matrix exponential

The Discrete Ordinate method with Matrix Exponential (DOME) was described in Refs. [Doicu and Trautmann, 2009d,d]. In this section we provide a brief summary of the method.

Rewriting the layer equation (3.7) as

$$\frac{d\mathbf{i}_m}{d\rho}(\rho) = \mathbf{A}_{mj} \mathbf{i}_m(\rho) + e^{-\tau_{\text{ext}}^0(\rho)} \mathbf{b}_{mj}, \quad (3.16)$$

where  $\rho = r - r_{j+1}$  with  $0 \leq \rho \leq \Delta r_j$  is the layer coordinate, and integrating Eq. (3.16), yields the integral form of the layer equation

$$\mathbf{i}_{m,j+1} = e^{-\mathbf{A}_{mj} \Delta r_j} \mathbf{i}_{mj} - \int_0^{\Delta r_j} e^{-\mathbf{A}_{mj} \rho} e^{-\tau_{\text{ext}}^0(\rho)} \mathbf{b}_{mj} d\rho, \quad (3.17)$$

for the level values of the radiance field  $\mathbf{i}_{mj} = [\mathbf{i}_{mj}^+; \mathbf{i}_{mj}^-]^T$  with  $[\mathbf{i}_{mj}^\pm]_k = I_m(r_j, \pm\mu_k)$ ,  $k = 1, \dots, N_{\text{DO}}$ .

The matrix exponential in Eq. (3.17) can be computed by using the eigendecomposition method or Padé approximation.

1. The eigendecomposition method for computing the matrix exponential is based on a spectral decomposition of the matrix  $\mathbf{A}_{mj}$  given by Eq. (3.8). This can be obtained by one of the following methods: direct decomposition of an asymmetric matrix [Stamnes and Swanson, 1981], square-root decomposition [Nakajima and Tanaka, 1986], and Cholesky decomposition [Stamnes et al., 1988]. In the direct decomposition method, we use the block symmetry of  $\mathbf{A}_{mj}$ , to obtain

$$\mathbf{A}_{mj} = \mathbf{V}_{mj} \boldsymbol{\Lambda}_{mj} \mathbf{V}_{mj}^{-1}, \quad (3.18)$$

### 3.2 Discrete ordinate method with matrix exponential

where  $\Lambda_{mj} = \text{diag}[\lambda_k; -\lambda_k]$ , and

$$\mathbf{V}_{mj} = \begin{bmatrix} \mathbf{V}_{mj}^+ & \mathbf{V}_{mj}^- \\ \mathbf{V}_{mj}^- & \mathbf{V}_{mj}^+ \end{bmatrix}, \quad (3.19)$$

$$\mathbf{V}_{mj}^\pm = [\mathbf{v}_1^\pm, \dots, \mathbf{v}_{N_{\text{DO}}}^\pm]. \quad (3.20)$$

In Eqs. (3.19) and (3.20),  $\begin{bmatrix} \mathbf{v}_k^+ \\ \mathbf{v}_k^- \end{bmatrix}$  are the right eigenvectors of  $\mathbf{A}_{mj}$  corresponding to eigenvalue  $\lambda_k$ , and  $\begin{bmatrix} \mathbf{v}_k^- \\ \mathbf{v}_k^+ \end{bmatrix}$  the right eigenvectors of  $\mathbf{A}_{mj}$  corresponding to eigenvalue  $-\lambda_k$ . Using Eq. (3.18), we compute the matrix exponential as

$$e^{-\mathbf{A}_{mj}\Delta r_j} = \mathbf{V}_{mj}\Lambda_{mj}^0\mathbf{V}_{mj}^{-1},$$

where  $\Lambda_{mj}^0 = \text{diag}[a_0(\lambda_k\Delta r_j); a_0(-\lambda_k\Delta r_j)]$  and  $a_0(x) = e^{-x}$ , and derive the layer equation

$$\mathbf{A}_{mj}^1\mathbf{i}_{mj} + \mathbf{A}_{mj}^2\mathbf{i}_{mj+1} = \mathbf{b}_{mj}, \quad (3.21)$$

where

$$\begin{aligned} \mathbf{A}_{mj}^1 &= \mathbf{D}_{mj}^1\mathbf{V}_{mj}^{-1}, \\ \mathbf{A}_{mj}^2 &= -\mathbf{D}_{mj}^2\mathbf{V}_{mj}^{-1}, \\ \mathbf{b}_{mj} &= \mathbf{B}_{mj}\mathbf{V}_{mj}^{-1}(\Delta r_j\mathbf{b}_{mj}), \end{aligned}$$

and

$$\begin{aligned} \mathbf{D}_{mj}^1 &= \text{diag}[a_0(\lambda_k\Delta r_j); 1], \\ \mathbf{D}_{mj}^2 &= \text{diag}[1; a_0(\lambda_k\Delta r_j)], \\ \mathbf{B}_{mj} &= \text{diag}[b_1(\lambda_k\Delta r_j); b_2(\lambda_k\Delta r_j)], \\ b_1(x) &= \frac{e^{-(\tau_{\text{ext},j+1}^0+x)} - e^{-\tau_{\text{ext},j+1}^0}}{\tau_{\text{ext},j+1}^0 - \tau_{\text{ext},j}^0 - x}, \\ b_2(x) &= e^{-x}b_1(-x) = \frac{e^{-(\tau_{\text{ext},j+1}^0+x)} - e^{-\tau_{\text{ext},j}^0}}{\tau_{\text{ext},j}^0 - \tau_{\text{ext},j+1}^0 - x}, \end{aligned}$$

with  $\tau_{\text{ext},j}^0$  and  $\tau_{\text{ext},j+1}^0$  being the solar optical depths at the boundary levels  $j$  and  $j+1$ .

2. In the first-order Padé approximation to the matrix exponential, the layer equation reads as

$$\mathbf{A}_{mj}^1 = \mathbf{I} - \frac{\Delta r_j}{2}\mathbf{A}_{mj}, \quad (3.22)$$

$$\mathbf{A}_{mj}^2 = -\left(\mathbf{I} + \frac{\Delta r_j}{2}\mathbf{A}_{mj}\right), \quad (3.23)$$

$$\mathbf{b}_{mj} = \mathbf{B}_{mj}(\Delta r_j\mathbf{b}_{mj})e^{-\tau_{\text{ext},j}^0}, \quad (3.24)$$

where  $\mathbf{B}_{mj}$  is now given by

$$\mathbf{B}_{mj} = I_0(\tau_0)\mathbf{I} + \left[\frac{1}{2}I_0(\tau_0) - I_1(\tau_0)\right]\Delta r_j\mathbf{A}_{mj}, \quad (3.25)$$

### 3 Radiative transfer model

with  $\tau_0 = \tau_{\text{ext},j+1}^0 - \tau_{\text{ext},j}^0$  and

$$I_0(x) = \frac{1 - e^{-x}}{x}, \quad (3.26)$$

$$I_1(x) = \frac{1}{x} \left[ 1 - \frac{e^{-x}}{x} (1 - e^{-x}) \right]. \quad (3.27)$$

The layer equation (3.21) together with the boundary conditions at the top and bottom of the atmosphere, i.e., (cf. Eq. (3.2))

$$\mathbf{i}_{m1}^- = 0, \quad (3.28)$$

and (cf. Eq. (3.3))

$$\mathbf{i}_{mN_{\text{lev}}}^+ = \mathbf{R}_{mN_{\text{lev}}} \mathbf{i}_{mN_{\text{lev}}}^- + \mathbf{r}_{mN_{\text{lev}}}, \quad (3.29)$$

with

$$\begin{aligned} [\mathbf{R}_{mN_{\text{lev}}}]_{kl} &= 2Aw_l\mu_l\rho_m(\mu_k, -\mu_l), \\ [\mathbf{r}_{mN_{\text{lev}}}]_k &= A\frac{F_0}{\pi}\mu_0\rho_m(\mu_k, -\mu_0)e^{-\tau_{\text{ext}}^0(|\mathbf{r}_s - \mathbf{r}_{\text{TOA}}|)}, \end{aligned}$$

and

$$\rho(\boldsymbol{\Omega}^+, \boldsymbol{\Omega}^-) = \sum_{m=0}^M (2 - \delta_{m0}) \rho_m(\mu, -\mu') \cos[m(\varphi - \varphi')],$$

for  $\boldsymbol{\Omega} = (\mu, \varphi)$  and  $\boldsymbol{\Omega}' = (-\mu', \varphi')$  with  $\mu, \mu' > 0$ , respectively, are assembled into the global matrix  $\mathbf{A}_m$  of the entire atmosphere. The solution of the resulting system of equations  $\mathbf{A}_m \mathbf{i}_m = \mathbf{b}_m$  yields the level values of the radiance field. The matrix  $\mathbf{A}_m$  of dimension  $2N_{\text{DO}}N_{\text{lev}} \times 2N_{\text{DO}}N_{\text{lev}}$  has  $3N_{\text{DO}} - 1$  sub- and super-diagonals and is compressed into band-storage and then inverted using, for example, the LU factorization. For optically thin layers, for which the condition  $\|\mathbf{A}_{mj} \Delta r_j\| \leq 1$  is satisfied, the matrix exponential can be computed by means of the Padé approximation instead of the eigendecomposition method. Because for such layers, the Padé approximation is more efficient, the computation speed is enhanced.

## 3.3 Acceleration techniques

### 3.3.1 Telescoping technique

The telescoping technique relies on the following result [Spurr, 2008]: For an atmosphere consisting of gas molecules (Rayleigh layers) and any group of contiguous aerosol or cloud layers, the azimuth phase function coefficients  $p_{mj}$  vanish for all  $m > 2$ , and all Rayleigh layers  $j$ . As a result, the layer matrix  $\mathbf{A}_{mj}$  becomes a diagonal matrix, and the layer vector  $\mathbf{b}_{mj}$  vanishes. In practice, this technique is implemented as follows: solve a boundary-value problem for the (active) aerosol/cloud layers, and then compute the radiances at the remaining levels recursively.



### 3.3.2 False discrete ordinate

The radiance at the top-of-atmosphere in a measurement direction  $\mathbf{\Omega}_m = (\mu_m, \varphi_m)$  with  $\mu_m > 0$ ,

$$I_m = I(r_{\text{TOA}}, \mathbf{\Omega}_m), \quad (3.30)$$

is usually computed by the source function integration technique. An alternative approach is the method of false discrete ordinate. Specifically, the radiance at point  $r$  in the upward direction  $\mathbf{\Omega} = (\mu, \varphi)$  with  $\mu > 0$ , can be calculated by using the integral form of the radiative transfer equation, that is,

$$\begin{aligned} I(r, \mathbf{\Omega}) &= I(r_s, \mathbf{\Omega}) e^{-\frac{1}{\mu} \tau(r)} \\ &+ \frac{1}{\mu} \int_{r_s}^r J(r', \mathbf{\Omega}) e^{-\frac{1}{\mu} [\tau(r) - \tau(r')]} dr', \end{aligned} \quad (3.31)$$

where

$$\begin{aligned} J(r, \mathbf{\Omega}) &= F_0 \frac{\sigma_{\text{sct}}(r)}{4\pi} P(r, \mathbf{\Omega}, \Omega_0) e^{-\tau_{\text{ext}}^0(|\mathbf{r} - \mathbf{r}_{\text{TOA}}|)} \\ &+ \frac{\sigma_{\text{sct}}(r)}{4\pi} \int_{4\pi} P(r, \mathbf{\Omega}, \mathbf{\Omega}') I(r, \mathbf{\Omega}') d\mathbf{\Omega}' \end{aligned}$$

is the source function, and  $\tau(r) = \int_{r_s}^r \sigma_{\text{ext}}(r') dr'$  the vertical optical depth. As the signal received by the detector is  $I(r_{\text{TOA}}, \mathbf{\Omega}_m)$ , we can simulate this radiance in the framework of the discrete ordinate method by introducing an additional stream in the direction  $\mu_m$ . Because the integral  $\int_{4\pi} P(\cdot, \mathbf{\Omega}, \mathbf{\Omega}') I(r, \mathbf{\Omega}') d\mathbf{\Omega}'$ , which appears in the expression of the source function, involves only the contribution of the Gaussian quadrature points  $\mu_k$ ,  $k = 1, \dots, N_{\text{DO}}$ , we choose  $w_m = 0$ ; therefore,  $\mu_m$  is called a false discrete ordinate. In conclusion, the top-of-atmosphere radiance at the false discrete ordinate is exactly the top-of-atmosphere radiance in the measurement direction computed by the source function integration method.

For a highly anisotropic phase function, the delta-M method [Wiscombe, 1977] with the TMS correction of Nakajima and Tanaka [Nakajima and Tanaka, 1988] is very effective for computing the transmitted and reflected radiances [Rozanov and Lyapustin, 2010]. The TMS correction requires one additional computational step, which consists in the calculation of the single scattering correction term in the measurement direction  $\mathbf{\Omega}_m$ . As we intend to avoid any post-processing step, the implementation of the TMS correction in the framework of the false discrete ordinate method seems to be beneficial. For this purpose, we recall some fundamental results. In the delta-M method, the exact phase function  $P(r, \mu) = \sum_{n=0}^{N_{\text{rank}}} c_n \chi_n(r) P_n(\mu)$ , where  $\mu = \cos \Theta$  and  $N_{\text{rank}} > M$  is the maximum expansion order, is replaced by an approximate phase function  $\mathcal{P}(r, \mu)$ . The approximate phase function is chosen as  $\mathcal{P}(r, \mu) = 2f\delta(1 - \mu) + (1 - f)\bar{P}(r, \mu)$ , where  $\delta(1 - \mu)$  is the forward peak, and  $\bar{P}(r, \mu) = \sum_{n=0}^M c_n \bar{\chi}_n(r) P_n(\mu)$  the truncated phase function. The truncation factor is computed as  $f(r) = \chi_{M+1}(r)/2$ , while the expansion coefficients of the truncated phase functions are related to the expansion coefficients of the exact phase function by the relation  $\bar{\chi}_n(r) = [\chi_n(r) - 2f(r)]/[1 - f(r)]$ . In this context, the diffuse radiance is the solution of a radiative transfer equation with the scaled parameters  $\bar{\sigma}_{\text{ext}} = \sigma_{\text{ext}} - f\sigma_{\text{sct}}$ ,  $\bar{\sigma}_{\text{sct}} = (1 - f)\sigma_{\text{sct}}$  and  $\bar{P}$  substituting the original parameters  $\sigma_{\text{ext}}$ ,  $\sigma_{\text{sct}}$  and  $P$ , respectively. In the TMS method, the scaled single scattering term

$$J_{\text{ss}}(r, \mathbf{\Omega}_m) = F_0 \frac{\bar{\sigma}_{\text{sct}}(r)}{4\pi} \bar{P}(r, \mathbf{\Omega}_m, \Omega_0) e^{-\bar{\tau}_{\text{ext}}^0(|\mathbf{r} - \mathbf{r}_{\text{TOA}}|)}$$

### 3 Radiative transfer model

is expressed in terms of the exact phase functions  $P$  instead of the approximate phase functions  $\mathcal{P}$ . More precisely and taking into account that for  $\mu \neq 1$ ,  $\mathcal{P} = (1 - f)\bar{P}$ , the term  $\bar{\sigma}_{\text{sct}}\bar{P} = \bar{\sigma}_{\text{sct}}\mathcal{P}/(1 - f)$  is replaced by  $\bar{\sigma}_{\text{sct}}P/(1 - f) = \sigma_{\text{sct}}P$ ; thus,

$$J_{\text{ss}}(r, \mathbf{\Omega}_m) = F_0 \frac{\sigma_{\text{sct}}(r)}{4\pi} P(r, \mathbf{\Omega}_m, \mathbf{\Omega}_0) e^{-\bar{\tau}_{\text{ext}}^0(|\mathbf{r} - \mathbf{r}_{\text{TOA}}|)}.$$

In this regard, the TMS correction for a false discrete ordinate requires to compute the component of the layer vector  $\mathbf{b}_{mj}$  corresponding to  $\mu_m$  by using the unscaled scattering coefficient  $\sigma_{\text{sct}}$  and the azimuth phase function coefficients  $p_{mj}(\mu, \mu') = \sum_{n=m}^{N_{\text{rank}}} \chi_{nj} P_n^m(\mu) P_n^m(\mu')$  (involving all expansion coefficients of the phase function  $N_{\text{rank}}$ ).

#### 3.3.3 The correlated $k$ -distribution method and the Principal Component Analysis (PCA)

In hyper-spectral remote sensing, the monochromatic radiative transfer calculations are the most time consuming part of the computational process. In order to reduce the monochromatic computations, the spectral radiance is not computed in a line-by-line framework, but rather by taking into account the interdependency between spectral channels.

The spectral signal of an instrument that measures the radiance at the top of the atmosphere in direction  $\mathbf{\Omega}_m$  at wavelength  $\lambda$  in the spectral interval  $[\lambda_{\text{min}}, \lambda_{\text{max}}]$ , is given by

$$\mathcal{I}_m(\lambda) = \int_{\lambda_{\text{min}} - s/2}^{\lambda_{\text{max}} + s/2} g(\lambda - \lambda') I_m(\lambda') d\lambda', \quad (3.32)$$

where  $g(\lambda)$  is the slit function and  $s$  the slit width. In principle, the computation of  $\mathcal{I}_m(\lambda)$  requires the knowledge of  $I_m(\lambda)$  at a fine set of spectral points in the spectral interval  $[\lambda_{\text{min}} - s/2, \lambda_{\text{max}} + s/2]$ . To speed up the computations, DOME is equipped with a broadband acceleration approach that combines the correlated  $k$ -distribution method with dimensionality reduction techniques. These approaches are summarized below.

1. *The correlated  $k$ -distribution method.* Consider a discretization of the spectral interval  $[\lambda_{\text{min}} - s/2, \lambda_{\text{max}} + s/2]$  into a set  $\{\bar{\lambda}_k\}_{k=1}^{\bar{W}}$  of  $\bar{W}$  equally spaced wavelengths with the discretization step  $\Delta\lambda$ , and assume that the transmission within a spectral interval depends only on the distribution of the gas absorption coefficient  $\sigma_{\text{abs}}^{\text{gas}}(\lambda)$  within the spectral interval [Ambartzumian, 1936]. Let  $F = F(\sigma_{\text{abs}k}^{\text{gas}})$  be the cumulative density function of  $\sigma_{\text{abs}}^{\text{gas}}(\lambda)$  in the spectral interval  $[\bar{\lambda}_k - \Delta\lambda/2, \bar{\lambda}_k + \Delta\lambda/2]$ ,  $\sigma_{\text{abs}k}^{\text{gas}}(F)$  the inverse distribution function, and  $\{F_l, \varpi_l\}_{l=1}^{N_q}$  a set of  $N_q$  quadrature points and weights in the interval  $[0, 1]$ . The spectral signal (3.32) is then computed as

$$\mathcal{I}_m(\lambda) = \sum_{w=1}^W \omega_w g(\lambda - \lambda_w) I_m(\sigma_{\text{abs}}^{\text{gas}}(\lambda_w)), \quad (3.33)$$

where  $\lambda_w = \bar{\lambda}_k$ ,  $\omega_w = \Delta\lambda \varpi_l$ , and  $\sigma_{\text{abs}}^{\text{gas}}(\lambda_w) = \sigma_{\text{abs}k}^{\text{gas}}(F_l)$  for  $w = l + (k - 1)N_q$ ,  $k = 1, \dots, \bar{W}$ ,  $l = 1, \dots, N_q$ , and  $W = \bar{W}N_q$ . Thus, in the correlated  $k$ -distribution method,  $W$  monochromatic radiative transfer calculations are required for computing  $I_m(\lambda_w)$ , and so,  $\mathcal{I}_m(\lambda)$ . A further acceleration can be achieved when  $I_m(\lambda_w)$  is computed by using dimensionality reduction techniques, as for example, the principal component analysis [Natraj et al., 2005, 2010].

2. *Principal component analysis.* At wavelength  $\lambda_w$ , the signal  $I_m(\lambda_w)$  is related to the signal calculated by a simplified radiative transfer model  $I_m^s(\lambda_w)$  through the relation

$$\ln \frac{I_m(\lambda_w)}{I_m^s(\lambda_w)} = f(\lambda_w), \quad (3.34)$$

i.e.,

$$I_m(\lambda_w) = I_m^s(\lambda_w)e^{f(\lambda_w)}. \quad (3.35)$$

The simplified model used in this study is a two-stream version of DOME, in which the eigenvalues and the eigenvectors of the layer matrix are computed analytically, and the system of equations for the entire atmosphere is solved by means of a penta-diagonal solver as in Ref. [Spurr and Natraj, 2011]. The correction factor  $f(\lambda_w)$  in Eq. (3.34) is the quantity which is calculated by means of the principal component analysis. To summarize this approach, we assume that for each wavelength  $\lambda_w$ , the spectral variability of the optical parameters can be described by a vector  $\mathbf{x}_w \in \mathbb{R}^{\mathcal{N}}$ , defined by

$$\mathbf{x}_w^T = [\ln \sigma_{\text{abs}1}^{\text{gas}}(\lambda_w), \dots, \ln \sigma_{\text{abs}N_{\text{lay}}}^{\text{gas}}(\lambda_w), \ln \sigma_{\text{sct}1}^{\text{mol}}(\lambda_w), \dots, \ln \sigma_{\text{sct}N_{\text{lay}}}^{\text{mol}}(\lambda_w)], \quad (3.36)$$

where  $\sigma_{\text{abs}j}^{\text{gas}}$  and  $\sigma_{\text{sct}j}^{\text{mol}}$  are the optical coefficients in the  $j$ th layer, and  $\mathcal{N} = 2N_{\text{lay}}$ . Denoting by  $\bar{\mathbf{x}} = (1/W) \sum_{w=1}^W \mathbf{x}_w$  the sample mean of the data, the goal is to find an  $\mathcal{M}$ -dimensional subspace ( $\mathcal{M} < \mathcal{N}$ ) spanned by a set of linear independent vectors  $\{\mathbf{a}_k\}_{k=1}^{\mathcal{M}}$ , such that the centered data  $\mathbf{x}_w - \bar{\mathbf{x}}$  belong to this subspace, i.e.,  $\mathbf{x}_w \approx \bar{\mathbf{x}} + \sum_{k=1}^{\mathcal{M}} y_{wk} \mathbf{a}_k = \bar{\mathbf{x}} + \mathbf{A} \mathbf{y}_w$  yielding  $\mathbf{y}_w = \mathbf{A}^\dagger (\mathbf{x}_w - \bar{\mathbf{x}})$  for  $w = 1, \dots, W$ . Here,  $\mathbf{A} = [\mathbf{a}_k]_{k=1}^{\mathcal{M}} \in \mathbb{R}^{\mathcal{N} \times \mathcal{M}}$ ,  $\mathbf{A}^\dagger = (\mathbf{A}^T \mathbf{A})^{-1} \mathbf{A}^T \in \mathbb{R}^{\mathcal{M} \times \mathcal{N}}$  is the pseudoinverse of  $\mathbf{A}$ , and  $y_{wk}$  is the  $k$ th component of the vector of parameters  $\mathbf{y}_w \in \mathbb{R}^{\mathcal{M}}$ . In the principal component analysis, the original  $\mathcal{N}$ -dimensional centered data  $\mathbf{X} = [\mathbf{x}_w - \bar{\mathbf{x}}]_{w=1}^W \in \mathbb{R}^{\mathcal{N} \times W}$  are projected onto the  $\mathcal{M}$ -dimensional subspace spanned by the dominant singular vectors of the data covariance matrix  $\mathbf{C}_x = (1/W) \mathbf{X} \mathbf{X}^T \in \mathbb{R}^{\mathcal{N} \times \mathcal{N}}$ , that is, with  $\sigma_k$  and  $\mathbf{u}_k$  being the  $k$ th singular value and singular vector of the matrix  $\mathbf{C}_x$ , respectively, we choose  $\mathbf{A} = \mathbf{U}_{\mathcal{M}} \boldsymbol{\Sigma}_{\mathcal{M}}^{1/2}$  and  $\mathbf{A}^\dagger = \boldsymbol{\Sigma}_{\mathcal{M}}^{-1/2} \mathbf{U}_{\mathcal{M}}^T$ , where  $\mathbf{U}_{\mathcal{M}} = [\mathbf{u}_k]_{k=1}^{\mathcal{M}} \in \mathbb{R}^{\mathcal{N} \times \mathcal{M}}$  and  $\boldsymbol{\Sigma}_{\mathcal{M}} = \text{diag}[\sigma_k]_{k=1}^{\mathcal{M}} \in \mathbb{R}^{\mathcal{M} \times \mathcal{M}}$ . Furthermore, approximating the correction factor  $f(\mathbf{x}_w)$  by a second-order Taylor expansion around  $\bar{\mathbf{x}}$ , and the gradient and the Hessian of  $f$  by central differences, we are led to the computational formula

$$\begin{aligned} f(\mathbf{x}_w) &\approx f(\bar{\mathbf{x}}) + \frac{1}{2} \sum_{k=1}^{\mathcal{M}} [f(\bar{\mathbf{x}} + \mathbf{a}_k) - f(\bar{\mathbf{x}} - \mathbf{a}_k)] y_{wk} \\ &\quad + \frac{1}{2} \sum_{k=1}^{\mathcal{M}} [f(\bar{\mathbf{x}} + \mathbf{a}_k) - 2f(\bar{\mathbf{x}}) + f(\bar{\mathbf{x}} - \mathbf{a}_k)] y_{wk}^2. \end{aligned} \quad (3.37)$$

To compute the radiance correction factor we identify  $f(\mathbf{x}_w)$  with  $f(\lambda_w) = \ln[I_m(\lambda_w)/I_m^s(\lambda_w)]$ . From Eq. (3.37) it is apparent that the computation of the correction factor requires  $2\mathcal{M} + 1$  calls of the full- and two-stream models. As a result and taking into account that  $\mathcal{M} \ll W$ , we are led to a substantial reduction of the computational time. It should be pointed out that the principal component

### 3 Radiative transfer model

analysis can also be used to compute the derivative of  $I_m(\lambda_w)$  with respect to the atmospheric parameter  $\varsigma$ . In this case, we identify  $f(\mathbf{x}_w)$  with the derivative correction factor

$$f_\varsigma(\lambda_w) = \ln \left[ \frac{\partial I_m}{\partial \varsigma}(\lambda_w) / \frac{\partial I_m^s}{\partial \varsigma}(\lambda_w) \right]. \quad (3.38)$$

## 3.4 Linearized radiative transfer model

The radiance measured by an instrument at the top of the atmosphere in direction  $\mathbf{\Omega}_m = (\mu_m, \varphi_m)$ ,  $\mu_m > 0$ ,  $I_m = I(r_{\text{TOA}}, \mathbf{\Omega}_m)$  depends on various atmospheric parameters. The retrieval of an atmospheric parameter requires the knowledge of the partial derivatives of the measured radiance with respect to the atmospheric parameter of interest (the weighting functions). The process of computing the partial derivatives is commonly referred to as linearization. Two linearization methods can be found in the literature, namely, the linearized forward approach and the forward-adjoint approach.

1. In the linearized forward approach, the partial derivatives are computed analytically (see for example, Refs. [Spurr, 2002, 2008; Spurr and Christi, 2007; Spurr et al., 2001]).
2. In the forward-adjoint approach, the measured radiance is expressed as the scalar product of the solution of the adjoint problem and the source term of the forward problem. Applying the linearization technique to the forward and adjoint problems, analytical expressions for the weighting functions have been derived in Refs. [Box, 2002; Doicu and Trautmann, 2009a; Landgraf et al., 2001; Marchuk, 1964, 2013; Rozanov and Rozanov, 2007; Ustinov, 2001, 2005]. Note that the forward-adjoint approach is extremely efficient because only two radiative transfer calculations are required for derivative calculation.

In this section, we describe the main peculiarities of these approaches and their implementation in the framework of the discrete ordinate method with matrix exponential. In particular, we consider the computation of the partial derivatives of the measured radiance with respect to an atmospheric parameter  $\varsigma$ , where for the aerosol layer (2.20),  $\varsigma$  stands for the aerosol optical depth  $\tau_{\text{aer}}$  and layer height  $H_{\text{aer}}$ . Because the optical input parameters of the radiative transfer model are  $\sigma_{\text{ext}j}$  and  $\xi_{nj}$  for  $j = 1, \dots, N_{\text{lay}}$ , and  $n = 0, \dots, M$ , we assume that the partial derivatives  $\partial \sigma_{\text{ext}j} / \partial \varsigma$  and  $\partial \xi_{nj} / \partial \varsigma$  are computable quantities.

### 3.4.1 Derivatives with respect to aerosol layer height

For an aerosol layer, caution should be used when computing the derivatives with respect to the aerosol layer height  $H_{\text{aer}}$ . For doing this we consider an homogenization approach, that is, the atmosphere is discretized into a fixed grid, and the layer containing the aerosol height  $H_{\text{aer}}$  is homogenized. The homogenization approach is summarized below.

Let the optical properties of the homogeneous aerosol layer be described by the extinction coefficient  $\sigma_{\text{ext}}^{\text{aer}}$  and the expansion coefficients  $\xi_n^{\text{aer}}$ . Assume that the aerosol layer height  $H_{\text{aer}}$  lies between the levels  $r_{j_{\text{min}}+1}$  and  $r_{j_{\text{min}}}$ , i.e.,  $r_{j_{\text{min}}+1} < H_{\text{aer}} < r_{j_{\text{min}}}$ . In layer  $j_{\text{min}}$  containing the aerosol layer height  $H_{\text{aer}}$ , we compute the optical parameters  $\sigma_{\text{ext}j_{\text{min}}}$  and  $\xi_{n,j_{\text{min}}}$  by using the relations

$$\sigma_{\text{ext}j_{\min}} = \sigma_{\text{ext}j_{\min}}^0 + w(H_{\text{aer}})\sigma_{\text{ext}}^{\text{aer}}, \quad (3.39)$$

$$\xi_{n_{j_{\min}}} = \xi_{n_{j_{\min}}}^0 + w(H_{\text{aer}})\xi_n^{\text{aer}}, \quad (3.40)$$

where the superscript “0” refers to a clear sky atmosphere, and the weighting factor  $w$  is defined as

$$w(H_{\text{aer}}) = \frac{H_{\text{aer}} - r_{j_{\min}+1}}{\Delta r_j}. \quad (3.41)$$

By this homogenization procedure,  $\sigma_{\text{ext}j_{\min}}$  and  $\xi_{n_{j_{\min}}}$  become functions of  $H_{\text{aer}}$ , i.e.,

$$\sigma_{\text{ext}j_{\min}} = \sigma_{\text{ext}j_{\min}}(\sigma_{\text{ext}}^{\text{aer}}, H_{\text{aer}}) \text{ and } \xi_{n_{j_{\min}}} = \xi_{n_{j_{\min}}}(\sigma_{\text{ext}}^{\text{aer}}, H_{\text{aer}}),$$

so that the partial derivatives  $\partial\sigma_{\text{ext}j_{\min}}/\partial H_{\text{aer}}$  and  $\partial\xi_{n_{j_{\min}}}/\partial H_{\text{aer}}$  can be readily computed.

### 3.4.2 Linearized forward approach

In the linearized version of the discrete ordinate method with matrix exponential, we take the derivative of the layer equation (3.21) with respect to the atmospheric parameter  $\varsigma$ , and obtain [Doicu and Trautmann, 2009b]

$$A_{mj}^1 \frac{\partial \mathbf{i}_{mj}}{\partial \varsigma} + A_{mj}^2 \frac{\partial \mathbf{i}_{mj+1}}{\partial \varsigma} = \frac{\partial b_{mj}}{\partial \varsigma} - \frac{\partial A_{mj}^1}{\partial \varsigma} \mathbf{i}_{mj} - \frac{\partial A_{mj}^2}{\partial \varsigma} \mathbf{i}_{mj+1}. \quad (3.42)$$

Some comments can be made here.

1. The linearized layer equations (3.42) are assembled into a global system of equations, which coincides with the system matrix for radiance calculations; only the right-hand sides are different.
2. To compute  $\partial A_{mj}^1/\partial \varsigma$ ,  $\partial A_{mj}^2/\partial \varsigma$  and  $\partial b_{mj}/\partial \varsigma$ , we apply the chain rule. In the case of Padé approximation, the derivative calculations are trivial, but in the case of the eigendecomposition method we have to compute the partial derivatives of the inverse of the eigenvector matrix  $\mathbf{V}_{mj}^{-1}$  and of the eigenvalues  $\lambda_k$ . The procedure is based on the following idea. Let  $\mathbf{A}$  be an  $n \times n$  matrix with eigenvalues  $\mu_k$  and eigenvectors  $\mathbf{w}_k$ , that is,

$$\mathbf{A} \mathbf{w}_k = \mu_k \mathbf{w}_k, \quad k = 1, \dots, n. \quad (3.43)$$

Taking the derivative of this equation with respect to  $\varsigma$ , we obtain

$$\frac{\partial \mathbf{A}}{\partial \varsigma} \mathbf{w}_k + \mathbf{A} \frac{\partial \mathbf{w}_k}{\partial \varsigma} = \frac{\partial \mu_k}{\partial \varsigma} \mathbf{w}_k + \mu_k \frac{\partial \mathbf{w}_k}{\partial \varsigma}, \quad (3.44)$$

and we see that Eq. (3.44) is a system of  $n$  equations with  $n + 1$  unknowns: the scalar  $\partial \mu_k/\partial \varsigma$  and the vector  $\partial \mathbf{w}_k/\partial \varsigma$ . In order to obtain an additional equation, we take into account that the eigenvectors  $\mathbf{w}_k$  are normalized, i.e.,  $\mathbf{w}_k^T \mathbf{w}_k = 1$ ; hence, we get

$$\mathbf{w}_k^T \frac{\partial \mathbf{w}_k}{\partial \varsigma} = 0. \quad (3.45)$$

By Eqs. (3.44) and (3.45), we are led to the system of equations

$$\begin{bmatrix} \mathbf{w}_k & \mu_k \mathbf{I} - \mathbf{A} \\ 0 & \mathbf{w}_k^T \end{bmatrix} \begin{bmatrix} \frac{\partial \mu_k}{\partial \varsigma} \\ \frac{\partial \mathbf{w}_k}{\partial \varsigma} \end{bmatrix} = \begin{bmatrix} \frac{\partial \mathbf{A}}{\partial \varsigma} \mathbf{w}_k \\ 0 \end{bmatrix}, \quad (3.46)$$

which can be solved for  $\partial \mu_k/\partial \varsigma$  and  $\partial \mathbf{w}_k/\partial \varsigma$ .

### 3.4.3 Forward-adjoint approach

In order to describe the forward-adjoint approach we consider the boundary value problem for the total radiance  $I(r, \boldsymbol{\Omega})$ , which consists in the radiative transfer equation

$$\mu \frac{dI}{dr}(r, \boldsymbol{\Omega}) = -\sigma_{\text{ext}}(r)I(r, \boldsymbol{\Omega}) + \frac{\sigma_{\text{sct}}(r)}{4\pi} \int_{4\pi} P(r, \boldsymbol{\Omega}, \boldsymbol{\Omega}')I(r, \boldsymbol{\Omega}') d\boldsymbol{\Omega}', \quad (3.47)$$

and the top-of-atmosphere and surface boundary conditions

$$I(r_{\text{TOA}}, \boldsymbol{\Omega}^-) = F_0 \delta(\boldsymbol{\Omega}^- - \boldsymbol{\Omega}_0), \quad (3.48)$$

and

$$I(r_s, \boldsymbol{\Omega}^+) = \frac{A}{\pi} \int_{2\pi} \rho(\boldsymbol{\Omega}^+, \boldsymbol{\Omega}^-) |\mu^-| I(r_s, \boldsymbol{\Omega}^-) d\boldsymbol{\Omega}^-, \quad (3.49)$$

respectively. The total radiance is decomposed into the diffuse radiance  $I(r, \boldsymbol{\Omega})$  and the direct solar beam  $I_0(r, \boldsymbol{\Omega})$ , i.e.,

$$I(r, \boldsymbol{\Omega}) = I(r, \boldsymbol{\Omega}) + I_0(r, \boldsymbol{\Omega}), \quad (3.50)$$

where

$$I_0(r, \boldsymbol{\Omega}) = F_0 \delta(\boldsymbol{\Omega} - \boldsymbol{\Omega}_0) T(r), \quad (3.51)$$

and  $T(r) = \exp[-\tau_{\text{ext}}(|\mathbf{r} - \mathbf{r}_{\text{TOA}}|, \boldsymbol{\Omega}_0)]$  with  $\tau_{\text{ext}}(|\mathbf{r} - \mathbf{r}_{\text{TOA}}|, \boldsymbol{\Omega}_0) = \tau_{\text{ext}}^0(|\mathbf{r} - \mathbf{r}_{\text{TOA}}|)$ , is the solar transmission.

In the next step, we express the radiative transfer equation in an operator form. For this purpose, we define the forward transport operator  $\mathcal{L}$  and the forward source term  $Q$  by the relations

$$\begin{aligned} \mathcal{L}I(r, \boldsymbol{\Omega}) &= \mu \frac{dI}{dr}(r, \boldsymbol{\Omega}) + \sigma_{\text{ext}}(r)I(r, \boldsymbol{\Omega}) - \frac{\sigma_{\text{sct}}(r)}{4\pi} \int_{4\pi} P(r, \boldsymbol{\Omega}, \boldsymbol{\Omega}')I(r, \boldsymbol{\Omega}') d\boldsymbol{\Omega}' \\ &\quad - \frac{A}{\pi} \delta(r - r_s) H(\mu) \mu \int_{4\pi} \rho(\boldsymbol{\Omega}, \boldsymbol{\Omega}') H(-\mu') |\mu'| I(r, \boldsymbol{\Omega}') d\boldsymbol{\Omega}' \end{aligned} \quad (3.52)$$

and

$$Q(r, \boldsymbol{\Omega}) = F_0 \mu_0 \delta(r - r_{\text{TOA}}) \delta(\boldsymbol{\Omega} - \boldsymbol{\Omega}_0), \quad (3.53)$$

respectively, where  $\delta$  is the Dirac delta function and  $H$  is the Heaviside step function. In the adjoint radiative transfer theory, we consider the adjoint radiative transfer operator  $\mathcal{L}^\dagger$ , defined through the Lagrange identity

$$\langle \mathcal{L}I, I^\dagger \rangle = \langle I, \mathcal{L}^\dagger I^\dagger \rangle, \quad (3.54)$$

where the scalar product of the fields  $I_1$  and  $I_2$  is given by

$$\langle I_1, I_2 \rangle = \int_{r_s}^{r_{\text{TOA}}} \int_{4\pi} I_1(r, \boldsymbol{\Omega}) I_2(r, \boldsymbol{\Omega}) d\boldsymbol{\Omega} dr.$$

The expression of the adjoint operator  $\mathcal{L}^\dagger$ , under the assumptions that (i)  $\mathbf{I}$  satisfies the boundary conditions  $\mathbf{I}(r_{\text{TOA}}, \boldsymbol{\Omega}^-) = \mathbf{I}(r_s, \boldsymbol{\Omega}^+) = 0$ , and (ii)  $\mathbf{I}^\dagger$  satisfies the boundary conditions  $\mathbf{I}^\dagger(r_{\text{TOA}}, \boldsymbol{\Omega}^+) = \mathbf{I}^\dagger(r_s, \boldsymbol{\Omega}^-) = 0$ , is given by

$$\begin{aligned} & \mathcal{L}^\dagger \mathbf{I}^\dagger(r, \boldsymbol{\Omega}) \\ &= -\mu \frac{d\mathbf{I}^\dagger}{dr}(r, \boldsymbol{\Omega}) + \sigma_{\text{ext}}(r) \mathbf{I}^\dagger(r, \boldsymbol{\Omega}) - \frac{\sigma_{\text{sct}}(r)}{4\pi} \int_{4\pi} P(r, \boldsymbol{\Omega}', \boldsymbol{\Omega}) \mathbf{I}^\dagger(r, \boldsymbol{\Omega}') d\boldsymbol{\Omega}' \\ & - \frac{A}{\pi} \delta(r - r_s) H(-\mu) |\mu| \int_{4\pi} \rho(\boldsymbol{\Omega}', \boldsymbol{\Omega}) H(\mu') |\mu'| \mathbf{I}^\dagger(r, \boldsymbol{\Omega}') d\boldsymbol{\Omega}'. \end{aligned} \quad (3.55)$$

In view of Eq. (3.30) giving the expression of the measured radiance  $I_m$ , we define the adjoint source term by

$$Q^\dagger(r, \boldsymbol{\Omega}) = \delta(r - r_{\text{TOA}}) \delta(\boldsymbol{\Omega} - \boldsymbol{\Omega}_m), \quad (3.56)$$

in which case, the representation

$$I_m = \langle Q^\dagger, \mathbf{I} \rangle, \quad (3.57)$$

readily follows. The main result of the adjoint radiative transfer theory states that if (i) the radiance  $\mathbf{I}$  solves the forward problem consisting in the operator equation  $\mathcal{L}\mathbf{I} = Q$  and the boundary conditions  $\mathbf{I}(r_{\text{TOA}}, \boldsymbol{\Omega}^-) = \mathbf{I}(r_s, \boldsymbol{\Omega}^+) = 0$ , and (ii) the radiance  $\mathbf{I}^\dagger$  solves the adjoint problem consisting in the operator equation  $\mathcal{L}^\dagger \mathbf{I}^\dagger = Q^\dagger$  and the boundary conditions  $\mathbf{I}^\dagger(r_{\text{TOA}}, \boldsymbol{\Omega}^+) = \mathbf{I}^\dagger(r_s, \boldsymbol{\Omega}^-) = 0$ , then (cf. Eqs. (3.54) and (3.57))

$$I_m = \langle Q^\dagger, \mathbf{I} \rangle = \langle \mathcal{L}^\dagger \mathbf{I}^\dagger, \mathbf{I} \rangle = \langle \mathbf{I}^\dagger, \mathcal{L}\mathbf{I} \rangle = \langle \mathbf{I}^\dagger, Q \rangle. \quad (3.58)$$

Thus, the signal measured by the instrument is the scalar product between the adjoint radiance and the forward source term. The solution of the adjoint radiative transfer problem can be found by using the same solution method as for the forward problem with a modified source term. Actually, by using the symmetry properties of the phase function  $P(r, -\boldsymbol{\Omega}, -\boldsymbol{\Omega}') = P(r, \boldsymbol{\Omega}', \boldsymbol{\Omega})$  and of the normalized bi-directional reflection function  $\rho(-\boldsymbol{\Omega}, -\boldsymbol{\Omega}') = \rho(\boldsymbol{\Omega}', \boldsymbol{\Omega})$ , it can be shown that the conjugate adjoint radiance  $\widehat{\mathbf{I}}^\dagger$ , defined by  $\widehat{\mathbf{I}}^\dagger(r, \boldsymbol{\Omega}) = \mathbf{I}^\dagger(r, -\boldsymbol{\Omega})$  solves the conjugate adjoint problem consisting in the operator equation  $\widehat{\mathcal{L}}\widehat{\mathbf{I}}^\dagger = \widehat{Q}^\dagger$  and the boundary conditions  $\widehat{\mathbf{I}}^\dagger(r_{\text{TOA}}, \boldsymbol{\Omega}^-) = \widehat{\mathbf{I}}^\dagger(r_s, \boldsymbol{\Omega}^+) = 0$ , where  $\widehat{Q}^\dagger(r, \boldsymbol{\Omega}) = \widehat{F}_0 \mu_m \delta(r - r_{\text{TOA}}) \delta(\boldsymbol{\Omega} - \widehat{\boldsymbol{\Omega}}_m)$  is the conjugate adjoint source term,  $\widehat{\boldsymbol{\Omega}}_m = -\boldsymbol{\Omega}_m = (\widehat{\mu}_m, \widehat{\varphi}_m)$  is the conjugate adjoint direction ( $\widehat{\mu}_m = -\mu_m$  and  $\widehat{\varphi}_m = \varphi_m + \pi$ ), and  $\widehat{F}_0 = 1/\mu_m$ . Thus, the pseudo-forward radiance  $\widehat{\mathbf{I}}^\dagger$  solves the same type of boundary value problem as the forward radiance  $\mathbf{I}$ , excepting the source terms  $Q$  and  $\widehat{Q}^\dagger$ , which however, are of similar forms.

Coming to derivative calculations, we take the derivative of Eq. (3.58) with respect to  $\varsigma$ , and use the relation  $\partial Q^\dagger / \partial \varsigma = 0$ , to obtain

$$\frac{\partial I_m}{\partial \varsigma} = \langle Q^\dagger, \frac{\partial \mathbf{I}}{\partial \varsigma} \rangle = \langle \mathcal{L}^\dagger \mathbf{I}^\dagger, \frac{\partial \mathbf{I}}{\partial \varsigma} \rangle = \langle \mathbf{I}^\dagger, \mathcal{L} \frac{\partial \mathbf{I}}{\partial \varsigma} \rangle.$$

Further, by using the result

$$\mathcal{L} \frac{\partial \mathbf{I}}{\partial \varsigma} = \frac{\partial Q}{\partial \varsigma} - \frac{\partial \mathcal{L}}{\partial \varsigma} \mathbf{I},$$

### 3 Radiative transfer model

in conjunction with the relation  $\partial Q/\partial \varsigma = 0$ , we end up with

$$\frac{\partial I_m}{\partial \varsigma} = -\langle I^\dagger, \frac{\partial \mathcal{L}}{\partial \varsigma} I \rangle. \quad (3.59)$$

Next, we separate the total radiance  $I(r, \boldsymbol{\Omega})$  into a diffuse and a direct component  $I(r, \boldsymbol{\Omega})$  and  $I_0(r, \boldsymbol{\Omega})$ , respectively, (cf. Eq. (3.50)), and do the same for the conjugate adjoint total radiance  $\widehat{I}^\dagger(r, \boldsymbol{\Omega})$ , i.e.,

$$\widehat{I}^\dagger(r, \boldsymbol{\Omega}) = \widehat{I}^\dagger(r, \boldsymbol{\Omega}) + \widehat{I}_0^\dagger(r, \boldsymbol{\Omega}), \quad (3.60)$$

where

$$\widehat{I}_0^\dagger(r, \boldsymbol{\Omega}) = \widehat{F}_0 \delta(\boldsymbol{\Omega} - \widehat{\boldsymbol{\Omega}}_m) \widehat{T}^\dagger(r)$$

and  $\widehat{T}^\dagger(r) = \exp[-\tau_{\text{ext}}(|\mathbf{r} - \mathbf{r}_{\text{TOA}}|, \widehat{\boldsymbol{\Omega}}_m)]$ . Inserting the resulting expressions in Eq. (3.59) gives the computational formula

$$\frac{\partial I_m}{\partial \varsigma} = T_1 + T_2, \quad (3.61)$$

where the first term

$$\begin{aligned} T_1 = & - \int_{r_s}^{r_{\text{TOA}}} \int_{4\pi} \frac{\partial \sigma_{\text{ext}}}{\partial \varsigma}(r) \widehat{I}^\dagger(r, -\boldsymbol{\Omega}) I(r, \boldsymbol{\Omega}) \text{d}r \text{d}\boldsymbol{\Omega} \\ & - F_0 \int_{r_s}^{r_{\text{TOA}}} \frac{\partial \sigma_{\text{ext}}}{\partial \varsigma}(r) \widehat{I}^\dagger(r, -\boldsymbol{\Omega}_0) T(r) \text{d}r \\ & + \frac{1}{4\pi} \int_{r_s}^{r_{\text{TOA}}} \int_{4\pi} \widehat{I}^\dagger(r, -\boldsymbol{\Omega}) \text{d}r \text{d}\boldsymbol{\Omega} \int_{4\pi} \frac{\partial}{\partial \varsigma} [\sigma_{\text{sct}}(r) P(r, \boldsymbol{\Omega}, \boldsymbol{\Omega}')] I(r, \boldsymbol{\Omega}') \text{d}\boldsymbol{\Omega}' \\ & + \frac{F_0}{4\pi} \int_{r_s}^{r_{\text{TOA}}} \int_{4\pi} \frac{\partial}{\partial \varsigma} [\sigma_{\text{sct}}(r) P(r, \boldsymbol{\Omega}, \boldsymbol{\Omega}_0)] \widehat{I}^\dagger(r, -\boldsymbol{\Omega}) T(r) \text{d}r \text{d}\boldsymbol{\Omega}, \end{aligned} \quad (3.62)$$

involves integrals of the conjugate adjoint diffuse radiance  $\widehat{I}^\dagger$ , while the second term

$$\begin{aligned} T_2 = & -\widehat{F}_0 \int_{r_s}^{r_{\text{TOA}}} \frac{\partial \sigma_{\text{ext}}}{\partial \varsigma}(r) I(r, \boldsymbol{\Omega}_m) \widehat{T}^\dagger(r) \text{d}r \\ & + \frac{\widehat{F}_0}{4\pi} \int_{r_s}^{r_{\text{TOA}}} \widehat{T}^\dagger(r) \text{d}r \int_{4\pi} \frac{\partial}{\partial \varsigma} [\sigma_{\text{sct}}(r) P(r, \boldsymbol{\Omega}_m, \boldsymbol{\Omega}')] I(r, \boldsymbol{\Omega}') \text{d}\boldsymbol{\Omega}' \\ & + \frac{\widehat{F}_0 F_0}{4\pi} \int_{r_s}^{r_{\text{TOA}}} \frac{\partial}{\partial \varsigma} [\sigma_{\text{sct}}(r) P(r, \boldsymbol{\Omega}_m, \boldsymbol{\Omega}_0)] T(r) \widehat{T}^\dagger(r) \text{d}r, \end{aligned} \quad (3.63)$$

involves integrals of the (conjugate) adjoint transmission  $\widehat{T}^\dagger$ . The computation of the terms  $T_1$  and  $T_2$  involves the following steps.

1. Using the expansions (3.11) and (3.12), yielding

$$\begin{aligned} \frac{\partial}{\partial \varsigma} [\sigma_{\text{sct}}(r) P(r, \boldsymbol{\Omega}, \boldsymbol{\Omega}')] &= \sum_{m=0}^M (2 - \delta_{m0}) \frac{\partial s_m}{\partial \varsigma}(r, \mu, \mu') \cos[m(\varphi - \varphi')], \\ \frac{\partial s_m}{\partial \varsigma}(r, \mu, \mu') &= \sum_{n=m}^M \frac{\partial \xi_n}{\partial \varsigma}(r) P_n^m(\mu) P_n^m(\mu'), \end{aligned}$$



and assuming a cosine-azimuth expansion of the diffuse radiance, the integration over the azimuth angle in Eqs. (3.62) and (3.63) can be performed. The result is

$$\frac{\partial I_m}{\partial \zeta} = \sum_{m=0}^M (T_{1m} + T_{2m}) \cos[m(\varphi_m - \varphi_0)], \quad (3.64)$$

where

$$\begin{aligned} T_{1m} = & -(1 + \delta_{m0})\pi \left[ \int_{r_s}^{r_{\text{TOA}}} \int_{-1}^1 \frac{\partial \sigma_{\text{ext}}}{\partial \zeta}(r) \widehat{I}_m^\dagger(r, -\mu) I_m(r, \mu) dr d\mu \right] \\ & - F_0 \left[ \int_{r_s}^{r_{\text{TOA}}} \frac{\partial \sigma_{\text{ext}}}{\partial \zeta}(r) \widehat{I}_m^\dagger(r, \mu_0) T(r) dr \right] \\ & + \frac{1}{2}(1 + \delta_{m0})\pi \left[ \int_{r_s}^{r_{\text{TOA}}} \int_{-1}^1 \widehat{I}_m^\dagger(r, -\mu) dr d\mu \right. \\ & \times \left. \int_{-1}^1 \frac{\partial s_m}{\partial \zeta}(r, \mu, \mu') I_m(r, \mu') d\mu' \right] \\ & + \frac{F_0}{2} \left[ \int_{r_s}^{r_{\text{TOA}}} \int_{-1}^1 \frac{\partial s_m}{\partial \zeta}(r, \mu, -\mu_0) \widehat{I}_m^\dagger(r, -\mu) T(r) dr d\mu \right] \end{aligned} \quad (3.65)$$

and

$$\begin{aligned} T_{2m} = & -\widehat{F}_0 \left[ \int_{r_s}^{r_{\text{TOA}}} \frac{\partial \sigma_{\text{ext}}}{\partial \zeta}(r) I_m(r, \mu_m) \widehat{T}^\dagger(r) dr \right] \\ & + \frac{\widehat{F}_0}{2} \left[ \int_{r_s}^{r_{\text{TOA}}} \widehat{T}^\dagger(r) dr \int_{-1}^1 \frac{\partial s_m}{\partial \zeta}(r, \mu_m, \mu') I_m(r, \mu') d\mu' \right] \\ & + \frac{\widehat{F}_0 F_0}{4\pi} (2 - \delta_{m0}) \left[ \int_{r_s}^{r_{\text{TOA}}} \frac{\partial s_m}{\partial \zeta}(r, \mu_m, -\mu_0) T(r) \widehat{T}^\dagger(r) dr \right]. \end{aligned} \quad (3.66)$$

2. To integrate over the radial coordinate in Eqs. (3.65) and (3.66), we assume that the partial derivatives  $\partial \sigma_{\text{ext},j}/\partial \zeta$  and  $\partial \xi_{n,j}/\partial \zeta$  are nonzero on all layers  $j \in J_\zeta$ , where  $J_\zeta$  is a subset of  $\{1, \dots, N_{\text{lay}}\}$ . As a result, the integration with respect to  $r$  reduces to an integration over all layers in  $J_\zeta$ ; thus,  $\int_{r_s}^{r_{\text{TOA}}} dr = \sum_{j \in J_\zeta} \int_{r_{j+1}}^{r_j} dr$ . For the azimuthal mode  $m$ , let  $I_{mj}(\rho, \pm \mu_k)$  and  $\widehat{I}_{mj}^\dagger(\rho, \pm \mu_k)$  be the forward and conjugate adjoint radiances at an internal point  $\rho$  in the layer  $j$  ( $0 \leq \rho \leq \Delta r_j$ ), respectively. Using the analytic representation of the radiance at an internal layer point as given in Ref. [Doicu and Trautmann, 2009c], the integrals

$$\begin{aligned} \mathcal{I}_{mj}(\mu) &= \int_0^{\Delta r_j} I_{mj}(\rho, \mu) \widehat{T}^\dagger(\rho) d\rho \\ &= \int_0^{\Delta r_j} e^{-[\frac{\rho}{\Delta r_j} \tau_{\text{ext},j}^m + (1 - \frac{\rho}{\Delta r_j}) \tau_{\text{ext},j+1}^m]} I_{mj}(\rho, \mu) d\rho, \end{aligned} \quad (3.67)$$

$$\begin{aligned} \widehat{\mathcal{I}}_{mj}(\mu) &= \int_0^{\Delta r_j} \widehat{I}_{mj}^\dagger(\rho, \mu) T(\rho) d\rho \\ &= \int_0^{\Delta r_j} e^{-[\frac{\rho}{\Delta r_j} \tau_{\text{ext},j}^0 + (1 - \frac{\rho}{\Delta r_j}) \tau_{\text{ext},j+1}^0]} \widehat{I}_{mj}^\dagger(\rho, \mu) d\rho, \end{aligned} \quad (3.68)$$

$$\mathcal{J}_{mj}(\mu, \mu') = \int_0^{\Delta r_j} \widehat{I}_{mj}^\dagger(\rho, \mu) I_{mj}(\rho, \mu') d\rho, \quad (3.69)$$

### 3 Radiative transfer model

and

$$\begin{aligned} \mathcal{T}_j &= \int_0^{\Delta r_j} T(\rho) \widehat{T}^\dagger(\rho) d\rho \\ &= \int_0^{\Delta r_j} e^{-[\frac{\rho}{\Delta r_j}(\tau_{\text{ext}j}^0 + \tau_{\text{ext}j}^m) + (1 - \frac{\rho}{\Delta r_j})(\tau_{\text{ext}j+1}^0 + \tau_{\text{ext}j+1}^m)]} d\rho \end{aligned} \quad (3.70)$$

for  $\mu = \pm\mu_k$ ,  $\mu' = \pm\mu_l$ ,  $k, l = 1, \dots, N_{\text{DO}}$ , can be computed analytically. Here,  $\tau_{\text{ext}j}^0$  and  $\tau_{\text{ext}j+1}^0$  are the optical depths along the characteristic  $\boldsymbol{\Omega}_0 = (-\mu_0, \varphi_0)$  at the boundary levels  $j$  and  $j+1$ , respectively, while  $\tau_{\text{ext}j}^m$  and  $\tau_{\text{ext}j+1}^m$  are the optical depths along the characteristic  $\widehat{\boldsymbol{\Omega}}_m = (\widehat{\mu}_m, \widehat{\varphi}_m) = (-\mu_m, \varphi_m + \pi)$  at the boundary levels  $j$  and  $j+1$ , respectively. The final result is

$$\begin{aligned} T_{1m} &= -(1 + \delta_{m0})\pi \sum_{j \in J_\zeta} T_{1mj}^a - F_0 \sum_{j \in J_\zeta} T_{1mj}^b \\ &\quad - \frac{1}{2}(1 + \delta_{m0})\pi \sum_{j \in J_\zeta} T_{1mj}^c + \frac{F_0}{2} \sum_{j \in J_\zeta} T_{1mj}^d, \end{aligned} \quad (3.71)$$

$$T_{2m} = -\widehat{F}_0 \sum_{j \in J_\zeta} T_{2mj}^a + \frac{\widehat{F}_0}{2} \sum_{j \in J_\zeta} T_{2mj}^b + \frac{\widehat{F}_0 F_0}{4\pi} (2 - \delta_{m0}) \sum_{j \in J_\zeta} T_{2mj}^c, \quad (3.72)$$

where

$$T_{1mj}^a = \frac{\partial \sigma_{\text{ext}j}}{\partial \zeta} \sum_k w_k [\mathcal{J}_{mj}(-\mu_k, \mu_k) + \mathcal{J}_{mj}(\mu_k, -\mu_k)], \quad (3.73)$$

$$T_{1mj}^b = \frac{\partial \sigma_{\text{ext}j}}{\partial \zeta} \widehat{\mathcal{I}}_{mj}(\mu_0), \quad (3.74)$$

$$\begin{aligned} T_{1mj}^c &= \sum_l \sum_k w_l w_k \left[ \frac{\partial s_{mj}}{\partial \zeta}(\mu_l, \mu_k) \mathcal{J}_{mj}(-\mu_l, \mu_k) \right. \\ &\quad + \frac{\partial s_{mj}}{\partial \zeta}(\mu_l, -\mu_k) \mathcal{J}_{mj}(-\mu_l, -\mu_k) \\ &\quad \left. + \frac{\partial s_{mj}}{\partial \zeta}(-\mu_l, \mu_k) \mathcal{J}_{mj}(\mu_l, \mu_k) + \frac{\partial s_{mj}}{\partial \zeta}(-\mu_l, -\mu_k) \mathcal{J}_{mj}(\mu_l, -\mu_k) \right], \end{aligned} \quad (3.75)$$

$$T_{1mj}^d = \sum_k w_k \left[ \frac{\partial s_{mj}}{\partial \zeta}(\mu_k, -\mu_0) \widehat{\mathcal{I}}_{mj}(-\mu_k) + \frac{\partial s_{mj}}{\partial \zeta}(-\mu_k, -\mu_0) \widehat{\mathcal{I}}_{mj}(\mu_k) \right], \quad (3.76)$$

and

$$T_{2mj}^a = \frac{\partial \sigma_{\text{ext}j}}{\partial \zeta} \mathcal{I}_{mj}(\mu_m), \quad (3.77)$$

$$T_{2mj}^b = \sum_k w_k \left[ \frac{\partial s_{mj}}{\partial \zeta}(\mu_m, \mu_k) \mathcal{I}_{mj}(\mu_k) + \frac{\partial s_{mj}}{\partial \zeta}(\mu_m, -\mu_k) \mathcal{I}_{mj}(-\mu_k) \right], \quad (3.78)$$

$$T_{2mj}^c = \frac{\partial s_{mj}}{\partial \zeta}(\mu_m, -\mu_0) \mathcal{T}_j \quad (3.79)$$

with  $\partial s_{mj}(\mu, \mu')/\partial \zeta = \sum_{n=m}^M (\partial \xi_{nj}/\partial \zeta) P_n^m(\mu) P_n^m(\mu')$ . Note that the integrals  $\widehat{\mathcal{I}}_{mj}(\mu_0)$  and  $\mathcal{I}_{mj}(\mu_m)$ , which enter in the expressions of  $T_{1mj}^b$  and  $T_{2mj}^a$ , respectively, can be computed by using the method of false discrete ordinate (the false discrete ordinates  $\mu_0$  and  $\mu_m$  with zero weights are added to the set  $\{\mu_k\}_{k=1}^{N_{\text{DO}}}$ ).

### 3.5 Forward models

Consider the generic nonlinear equation  $\mathbf{y}^\delta = \mathbf{F}(\mathbf{x})$ , where  $\mathbf{F}$  is the forward model,  $\mathbf{x}$  the state vector, and  $\mathbf{y}^\delta$  the noisy data vector. The state vector  $\mathbf{x}$  encapsulates the atmospheric parameters to be retrieved. At first glance, the noisy data vector  $\mathbf{y}^\delta$  can be identified with the noisy spectral signals measured by the instrument at a set of  $N_{m\lambda}$  measurement wavelengths  $\{\lambda_{mk}\}_{k=1}^{N_{m\lambda}}$ , i.e.,

$$\mathbf{y}^\delta = \begin{bmatrix} \mathcal{I}_{\text{mes}}^\delta(\lambda_{m1}) \\ \vdots \\ \mathcal{I}_{\text{mes}}^\delta(\lambda_{mN_{m\lambda}}) \end{bmatrix}, \quad (3.80)$$

while the forward model can be identified with the spectral signals simulated by a radiative transfer model,

$$\mathbf{F}(\mathbf{x}) = \begin{bmatrix} \mathcal{I}_{\text{sim}}(\lambda_{m1}, \mathbf{x}) \\ \vdots \\ \mathcal{I}_{\text{sim}}(\lambda_{mN_{m\lambda}}, \mathbf{x}) \end{bmatrix}. \quad (3.81)$$

In the above equations, the subscript  $m$  stands for "measurement", and the dependency of the spectral signals on the measurement direction  $\boldsymbol{\Omega}_m$  is not indicated explicitly; this is tacitly assumed.

Unfortunately, the choices (3.80) and (3.81) lead to a nonlinear equation with a high degree of non-linearity. The following forward models, characterized by a lower degrees of non-linearity, are implemented in the retrieval algorithm:

1. the radiance model,

$$\ln \mathcal{I}_{\text{mes}}^\delta(\lambda_{mk}) = \ln \mathcal{I}_{\text{sim}}(\lambda_{mk}, \mathbf{x}) + \sum_{j=1}^{N_s} b_j S_j(\lambda_{mk}, \mathbf{x}_a), \quad (3.82)$$

2. the differential radiance model with internal smoothing (DRMI),

$$R_{\text{mes}}^\delta(\lambda_{mk}) = R_{\text{sim}}(\lambda_{mk}, \mathbf{x}) + \sum_{j=1}^{N_s} b_j S_j(\lambda_{mk}, \mathbf{x}_a), \quad k = 1, \dots, N_\lambda, \quad (3.83)$$

$$R_{\text{mes}}^\delta(\lambda_{mk}) = \ln \mathcal{I}_{\text{mes}}^\delta(\lambda_{mk}) - P_{\text{mes}}(\lambda_{mk}, \mathbf{c}_{\text{mes}}), \quad (3.84)$$

$$R_{\text{sim}}(\lambda_{mk}, \mathbf{x}) = \ln \mathcal{I}_{\text{sim}}(\lambda_{mk}, \mathbf{x}) - P_{\text{sim}}(\lambda_{mk}, \mathbf{c}_{\text{sim}}(\mathbf{x})), \quad (3.85)$$

and

3. the differential radiance model with external smoothing (DRME),

$$R_{\text{mes}}^\delta(\lambda_{mk}) = \ln \mathcal{I}_{\text{sim}}(\lambda_{mk}, \mathbf{x}) + \sum_{j=1}^{N_s} b_j S_j(\lambda_{mk}, \mathbf{x}_a) - P(\lambda_{mk}, \mathbf{c}), \quad k = 1, \dots, N_\lambda, \quad (3.86)$$

$$R_{\text{mes}}^\delta(\lambda_{mk}) = \ln \mathcal{I}_{\text{mes}}^\delta(\lambda_{mk}) - P_{\text{mes}}(\lambda_{mk}, \mathbf{c}_{\text{mes}}).$$

### 3 Radiative transfer model

Note that in DRMI, the measured and simulated differential spectral signals are fitted, while in DRME, the measured differential spectral signal and a simulated spectral signal with its smooth component extracted are fitted. In Eqs. (3.82)–(3.86),  $\mathbf{x}_a$  an a priori estimate of the state vector  $\mathbf{x}$ ,  $S_j(\lambda_{mk}, \mathbf{x}_a)$  with  $j = 1, \dots, N_s$ , the correction spectra including for example, the polarization correction spectrum and the Ring spectrum,  $N_s$  the number of correction spectra, and  $b_j$  with  $j = 1, \dots, N_s$ , the (wavelength independent) amplitudes of the correction spectra. The polynomials  $P_{\text{mes}}(\lambda, \mathbf{c}_{\text{mes}})$ ,  $P_{\text{sim}}(\lambda, \mathbf{c}_{\text{sim}}(\mathbf{x}))$ , and  $P(\lambda, \mathbf{c})$  account for the low-order spectral structure due to the scattering by clouds and aerosols. Specifically, in DRMI, the coefficients  $\mathbf{c}_{\text{mes}}$  and  $\mathbf{c}_{\text{sim}}(\mathbf{X})$  of the smoothing polynomials  $P_{\text{mes}}(\lambda, \mathbf{c}_{\text{mes}})$  and  $P_{\text{sim}}(\lambda, \mathbf{c}_{\text{sim}}(\mathbf{x}))$ , respectively, are computed as the solutions of the least-squares problems

$$\mathbf{c}_{\text{mes}} = \arg \min_{\mathbf{c}} \sum_{k=1}^{N_\lambda} [\ln \mathcal{I}_{\text{mes}}^\delta(\lambda_{mk}) - P_{\text{mes}}(\lambda_{mk}, \mathbf{c})]^2 \quad (3.87)$$

and

$$\mathbf{c}_{\text{sim}}(\mathbf{x}) = \arg \min_{\mathbf{c}} \sum_{k=1}^{N_\lambda} [\ln \mathcal{I}_{\text{sim}}(\lambda_{mk}, \mathbf{x}) - P_{\text{sim}}(\lambda_{mk}, \mathbf{c})]^2, \quad (3.88)$$

respectively. Thus, these coefficients, which are uniquely determined by  $\ln \mathcal{I}_{\text{mes}}^\delta(\lambda_{mk})$  and  $\ln \mathcal{I}_{\text{sim}}(\lambda_{mk}, \mathbf{x})$ , are not a part of the retrieval, while in DRME, the coefficients  $\mathbf{c}$  of the smoothing polynomial  $P(\lambda, \mathbf{c})$  are included in the retrieval.

## 4 Bayesian-based retrieval algorithms

For an aerosol layer with a box indicator function, we aim to retrieve the aerosol optical depth and layer height. Hereafter, in order to simplify the notation, these quantities will be denoted by  $\tau$  and  $H$ , respectively. The retrieval can be done by

1. assuming that the aerosol model is known, or
2. selecting the best aerosol model from a set of candidate models.

### 4.1 Retrieval algorithm for a known aerosol model

In this section, which is adapted from the journal article in Appendix 1, we present (i) the theoretical basis of a retrieval algorithm, designed under the assumption that the aerosol model is known, and (ii) its application to a sensitivity analysis.

#### 4.1.1 Algorithm description

Considering  $N_{\text{am}}$  aerosol models, we select an appropriate aerosol model  $m$  by using a prior seasonal or geographical information. The retrieval of the state vector  $\mathbf{x} = [\tau, H]$  is an inverse problem relying on the solution of the nonlinear equation

$$\mathbf{y}^\delta = \mathbf{F}_m(\mathbf{x}) + \boldsymbol{\delta}_m, \quad (4.1)$$

where  $\mathbf{y}^\delta$  is the measurement vector or the noisy data vector,  $\mathbf{F}_m(\mathbf{x})$  the forward model corresponding to the aerosol model  $m$ ,

$$\boldsymbol{\delta}_m = \boldsymbol{\delta}_{\text{mes}} + \boldsymbol{\delta}_{\text{aerm}}$$

the total data error vector,  $\boldsymbol{\delta}_{\text{mes}}$  the measurement error vector, and  $\boldsymbol{\delta}_{\text{aerm}}$  the aerosol model error vector, i.e., the error due to an inadequate aerosol model. Here,  $\mathbf{y}^\delta$  and  $\mathbf{F}_m(\mathbf{x})$  are the vectors of the log of the measured and simulated spectral signals, i.e.,  $[\mathbf{y}^\delta]_i = \ln \mathcal{I}_{\text{mes}}^\delta(\lambda_{mi}, \mathbf{x})$  and

$$[\mathbf{F}_m(\mathbf{x})]_i = \ln \mathcal{I}_{\text{sim}}^{(m)}(\lambda_{mi}, \mathbf{x}), \quad (4.2)$$

respectively, where  $\{\lambda_{mi}\}_{i=1}^{N_{m\lambda}}$  is a set of  $N_{m\lambda}$  measurement wavelengths in the considered spectral domain.

In a first step, the data model (4.1) is transformed into a model with white noise by using the prewhitening technique. The procedure is as follows. Assuming that

1.  $\boldsymbol{\delta}_{\text{mes}}$  is a Gaussian random vector with zero mean and covariance matrix

$$\mathbf{C}_{\text{mes}} = \text{diag}(\sigma_{\text{mesi}}^2)_{N_{m\lambda}} = \sigma_{\text{mes}}^2 \mathbf{C}_{\text{mes}},$$

$$\mathbf{C}_{\text{mes}} = \text{diag}\left(\frac{\sigma_{\text{mesi}}^2}{\sigma_{\text{mes}}^2}\right)_{N_{m\lambda}}, \quad \sigma_{\text{mes}}^2 = \frac{1}{N_{m\lambda}} \sum_{i=1}^{N_{m\lambda}} \sigma_{\text{mesi}}^2,$$

#### 4 Bayesian-based retrieval algorithms

where  $\sigma_{\text{mes}}^2$  is the measurement error variance and  $\mathbf{C}_{\text{mes}}$  a normalized measurement error covariance matrix;

2.  $\delta_{\text{aerm}}$  is a Gaussian random vector with zero mean and covariance matrix  $\mathbf{C}_{\text{aerm}} = \sigma_{\text{aerm}}^2 \mathbf{I}_{N_{m\lambda}}$ , where  $\sigma_{\text{aerm}}^2$  is the aerosol model error variance and  $\mathbf{I}_{N_{m\lambda}}$  the identity matrix; and
3.  $\delta_{\text{mes}}$  and  $\delta_{\text{aerm}}$  are independent random vectors,

we deduce that  $\delta_m$  is also a Gaussian random vector with zero mean and covariance matrix

$$\begin{aligned}\mathbf{C}_{\delta m} &= \mathbf{C}_{\text{mes}} + \mathbf{C}_{\text{aerm}} = \sigma_m^2 \mathbf{C}_{\delta m}, \\ \mathbf{C}_{\delta m} &= w_{\text{mes}m} \mathbf{C}_{\text{mes}} + (1 - w_{\text{mes}m}) \mathbf{I}_{N_{m\lambda}},\end{aligned}$$

where

$$\sigma_m^2 = \sigma_{\text{mes}}^2 + \sigma_{\text{aerm}}^2$$

is the data error variance,  $\mathbf{C}_{\delta m}$  a normalized data error covariance matrix, and

$$w_{\text{mes}m} = \sigma_{\text{mes}}^2 / \sigma_m^2$$

a weighting factor giving the contribution of  $\mathbf{C}_{\text{mes}}$  to the covariance matrix  $\mathbf{C}_{\delta m}$ . In this context, the scaled data model reads as

$$\bar{\mathbf{y}}^\delta = \bar{\mathbf{F}}_m(\mathbf{x}) + \bar{\delta}_m, \quad (4.3)$$

where  $\bar{\mathbf{y}}^\delta = \mathbf{P}\mathbf{y}^\delta$ ,  $\bar{\mathbf{F}}_m(\mathbf{x}) = \mathbf{P}\mathbf{F}_m(\mathbf{x})$ ,  $\bar{\delta}_m = \mathbf{P}\delta_m$ , and  $\mathbf{P} = \mathbf{C}_{\delta m}^{-1/2}$  is a scaling matrix. Because  $\bar{\mathbf{C}}_{\delta m} = \mathcal{E}\{\bar{\delta}_m \bar{\delta}_m^T\} = \sigma_m^2 \mathbf{I}_{N_{m\lambda}}$ , it is readily seen that  $\bar{\delta}_m \sim \mathbf{N}(\mathbf{0}, \bar{\mathbf{C}}_{\delta m} = \sigma_m^2 \mathbf{I}_{N_{m\lambda}})$ , where the notation  $\mathbf{N}(\mathbf{x}_{\text{mean}}, \mathbf{C}_x)$  stands for a normal distribution with mean  $\mathbf{x}_{\text{mean}}$  and covariance matrix  $\mathbf{C}_x$ .

Further, we assume that  $\mathbf{x} \sim \mathbf{N}(\mathbf{x}_a, \mathbf{C}_x)$ , where  $\mathbf{x}_a$  is the a priori state vector, the best beforehand estimate of the solution,  $\mathbf{C}_x = \sigma_x^2 \mathbf{C}_x$  the a priori covariance matrix, and  $\sigma_x^2$  the a priori state variance. Defining the regularization matrix  $\mathbf{L}$  and the regularization parameter  $\alpha$  through the relations  $\mathbf{C}_x^{-1} = \mathbf{L}^T \mathbf{L}$  and  $\alpha = \sigma_m^2 / \sigma_x^2$ , respectively, we express the a priori covariance matrix as  $\mathbf{C}_x = \sigma_m^2 (\alpha \mathbf{L}^T \mathbf{L})^{-1}$ .

The nonlinear equation (4.3) is solved by using a Bayesian approach. The key quantity in this approach is the conditional probability density of  $\mathbf{x}$  given the data  $\bar{\mathbf{y}}^\delta$  and the aerosol model  $m$ ,  $p(\mathbf{x} | \bar{\mathbf{y}}^\delta, m)$ , also called the a posteriori density. According to Bayes' theorem, this is given by

$$p(\mathbf{x} | \bar{\mathbf{y}}^\delta, m) = \frac{p(\bar{\mathbf{y}}^\delta | \mathbf{x}, m) p(\mathbf{x} | m)}{p(\bar{\mathbf{y}}^\delta | m)},$$

where  $p(\mathbf{x} | m)$  is the a priori density, i.e., the conditional probability density of  $\mathbf{x}$  given the aerosol model  $m$  before performing the measurement  $\bar{\mathbf{y}}^\delta$ ,  $p(\bar{\mathbf{y}}^\delta | \mathbf{x}, m)$  the likelihood density, i.e., the conditional probability density of  $\bar{\mathbf{y}}^\delta$  given the state  $\mathbf{x}$  and the aerosol model  $m$ , and

$$p(\bar{\mathbf{y}}^\delta | m) = \int p(\mathbf{x}, \bar{\mathbf{y}}^\delta | m) d\mathbf{x} = \int p(\bar{\mathbf{y}}^\delta | \mathbf{x}, m) p(\mathbf{x} | m) d\mathbf{x}, \quad (4.4)$$

the marginal likelihood density. For  $\mathbf{x} \sim \mathcal{N}(\mathbf{x}_a, \mathbf{C}_x = \sigma_m^2(\alpha \mathbf{L}^T \mathbf{L})^{-1})$  and  $\bar{\boldsymbol{\delta}}_m \sim \mathcal{N}(\mathbf{0}, \bar{\mathbf{C}}_{\delta m} = \sigma_m^2 \mathbf{I}_{N_{m\lambda}})$ , the Bayes' formula yields

$$p(\mathbf{x} | \bar{\mathbf{y}}^\delta, m) \propto \exp[-V_\alpha(\mathbf{x} | \bar{\mathbf{y}}^\delta, m)], \quad (4.5)$$

where

$$V_\alpha(\mathbf{x} | \bar{\mathbf{y}}^\delta, m) = \frac{1}{2\sigma_m^2} \left[ \|\bar{\mathbf{y}}^\delta - \bar{\mathbf{F}}_m(\mathbf{x})\|^2 + \alpha \|\mathbf{L}(\mathbf{x} - \mathbf{x}_a)\|^2 \right] \quad (4.6)$$

is the a posteriori potential, and the maximum a posteriori estimate  $\hat{\mathbf{x}}_{m\alpha}^\delta$  is computed as

$$\hat{\mathbf{x}}_{m\alpha}^\delta = \arg \min_{\mathbf{x}} V_\alpha(\mathbf{x} | \bar{\mathbf{y}}^\delta, m). \quad (4.7)$$

In a deterministic setting,

$$\mathcal{F}_{m\alpha}(\mathbf{x}) = \sigma_m^2 V_\alpha(\mathbf{x} | \bar{\mathbf{y}}^\delta, m) = \frac{1}{2\sigma_m^2} \left[ \|\bar{\mathbf{y}}^\delta - \bar{\mathbf{F}}_m(\mathbf{x})\|^2 + \alpha \|\mathbf{L}(\mathbf{x} - \mathbf{x}_a)\|^2 \right]$$

is the Tikhonov function for the nonlinear equation  $\bar{\mathbf{y}}^\delta = \bar{\mathbf{F}}_m(\mathbf{x})$  with the penalty term  $\alpha \|\mathbf{L}(\mathbf{x} - \mathbf{x}_a)\|^2$  and the regularization parameter  $\alpha$ . Thus, a regularized solution

$$\mathbf{x}_{m\alpha}^\delta = \arg \min_{\mathbf{x}} \mathcal{F}_{m\alpha}(\mathbf{x}). \quad (4.8)$$

coincides with the maximum a posteriori estimate, i.e.,  $\mathbf{x}_{m\alpha}^\delta = \hat{\mathbf{x}}_{m\alpha}^\delta$ . The computation of the regularized solution  $\mathbf{x}_{m\alpha}^\delta$  in the framework of the method of Tikhonov regularization requires the knowledge of the optimal value of the regularization parameter  $\hat{\alpha}$ , while in a Bayesian framework, the optimal regularization parameter is identified as the true ratio of the data error variance and the *a priori* state variance, i.e.,  $\hat{\alpha} = \sigma_m^2 / \sigma_x^2$ . Several regularization parameter choice methods have been discussed in Ref. [Doicu et al., 2010]. These include methods with constant regularization parameters, e.g., the maximum likelihood estimation, the generalized cross-validation, and the nonlinear L-curve method. Unfortunately, at present, there is no infallible method which guarantees an optimal choice of the regularization parameter.

An improvement of the problems associated with the regularization parameter selection is achieved in the framework of iterative regularization methods. These approaches are less sensitive to overestimates of the regularization parameter, but require more iteration steps to achieve convergence. A representative iterative approach is the iteratively regularized Gauss-Newton method. At the iteration step  $k$  of the iteratively regularized Gauss-Newton method, the new iterate minimizing the linearized function

$$\mathcal{F}_{m\alpha}(\mathbf{x}) = \|\bar{\mathbf{y}}_{mk}^\delta - \bar{\mathbf{K}}_{mk}(\mathbf{x} - \mathbf{x}_a)\|^2 + \alpha_k \|\mathbf{L}(\mathbf{x} - \mathbf{x}_a)\|^2, \quad (4.9)$$

is given by

$$\mathbf{x}_{mk+1}^\delta = \mathbf{x}_a + \bar{\mathbf{K}}_{mk}^\dagger \bar{\mathbf{y}}_{mk}^\delta, \quad (4.10)$$

where

$$\bar{\mathbf{K}}_{mk}^\dagger = (\bar{\mathbf{K}}_{mk}^T \bar{\mathbf{K}}_{mk} + \alpha_k \mathbf{L}^T \mathbf{L})^{-1} \bar{\mathbf{K}}_{mk}^T \quad (4.11)$$

is the regularized generalized inverse,  $\bar{\mathbf{K}}_{mk} = \bar{\mathbf{K}}(\mathbf{x}_{mk}^\delta) = \partial \bar{\mathbf{F}}(\mathbf{x}_{mk}^\delta) / \partial \mathbf{x}$  the Jacobian matrix evaluated at  $\mathbf{x}_{mk}^\delta$ , and

$$\bar{\mathbf{y}}_{mk}^\delta = \bar{\mathbf{y}}^\delta - \bar{\mathbf{F}}(\mathbf{x}_{mk}^\delta) + \bar{\mathbf{K}}_{mk}(\mathbf{x}_{mk}^\delta - \mathbf{x}_a) \quad (4.12)$$

the linearized data vector at the iteration step  $k$ . The following peculiarities of the iteratively regularized Gauss-Newton method deserve to be mentioned [Doicu et al., 2010].

1. In contrast to the method of Tikhonov regularization, the regularization parameter is not constant during the iterative process. Instead, the regularization parameters  $\alpha_k$  are the terms of a decreasing (geometric) sequence, i.e.,  $\alpha_k = q\alpha_{k-1}$  with  $q < 1$ . In this way, the amount of regularization is gradually decreased during the iterative process.
2. For iterative regularization methods, the number of iteration steps  $k$  plays the role of the regularization parameter, and the iterative process is stopped after an appropriate number of steps  $k^*$  in order to avoid an uncontrolled expansion of the errors in the data. The stopping rule used in this study is the discrepancy principle [Morozov, 1966], according to which, the iterative process is terminated after  $k^*$  steps such that

$$\|\mathbf{r}_{mk^*}^\delta\| \leq \tau\Delta < \|\mathbf{r}_{mk}^\delta\|, \quad 0 \leq k < k^*, \quad (4.13)$$

where  $\mathbf{r}_{mk}^\delta = \mathbf{y}^\delta - \mathbf{F}_m(\mathbf{x}_k^\delta)$  is the residual vector at  $\mathbf{x}_k^\delta$ ,  $\tau > 1$  a control parameter, and  $\Delta$  the noise level (an upper bound for the noise in the data). Because in practice, the noise level cannot be a priori estimated, we adopt a practical approach based on the observation that the residual  $\|\mathbf{r}_{mk}^\delta\|$  decreases during the iterative process and attains a plateau at approximately  $\Delta$ . Thus, if the nonlinear residuals  $\|\mathbf{r}_{mk}^\delta\|$  converge to  $\|\mathbf{r}_{m\infty}^\delta\|$  within a prescribed tolerance, we use the estimate  $\Delta \approx \|\mathbf{r}_{m\infty}^\delta\|$ .

3. The numerical experiments performed in Ref. [Doicu et al., 2010] showed that at the solution  $\mathbf{x}_{mk^*}^\delta$ , (i)  $\alpha_{k^*-1}$  is close to the optimal regularization parameter, and (ii)  $\mathbf{x}_{mk^*}^\delta$  is close to the Tikhonov solution corresponding to the optimal regularization parameter. Therefore, we assume that
  - a)  $\hat{\alpha} = \alpha_{k^*-1}$  is an estimate for the optimal regularization parameter, and
  - b)  $\mathbf{x}_{m\hat{\alpha}}^\delta = \mathbf{x}_{mk^*}^\delta$  is the minimizer of the Tikhonov function with regularization parameter  $\hat{\alpha}$ ,  $\mathcal{F}_{m\hat{\alpha}}(\mathbf{x})$ .

In conclusion, the iteratively regularized Gauss-Newton method provides an optimal value of the regularization parameter  $\hat{\alpha}$  (i.e., the ratio of the data error variance  $\sigma_m^2$  and the a priori state variance  $\sigma_x^2$ ) and the corresponding regularized solution  $\mathbf{x}_{m\hat{\alpha}}^\delta$ .

### 4.1.2 Sensitivity analysis

Deriving aerosol information from satellite measurements is a difficult task in terms of retrieval sensitivity and accuracy. Several critical factors affecting the accuracy of aerosol remote sensing includes the aerosol model, treatment of the underlying surface, sensor calibration, and cloud screening. In this section we evaluate the retrieval accuracy of the aerosol parameters by performing a sensitivity analysis. By studying the impact of the forward model and instrument uncertainties on the retrieval, the expected retrieval performance using real measurements can be estimated.

The retrieval is performed by using the Bayesian-based retrieval algorithm for a known aerosol model and an aerosol layer that extends from the surface to the height  $H$  (box vertical profile). The absorption cross sections of gas molecules are computed by using LBL calculations [Schreier et al., 2014] with optimized rational approximations for the Voigt line profile [Schreier, 2011], the wavenumber grid point spacing is a fraction (e.g., 1/4) of the minimum half-width of the Voigt lines taken from HITRAN database [Rothman et al., 2009], and the Rayleigh cross-section and depolarization ratios are calculated as in Ref. [Bodhaine et al., 1999]. In particular, the uncertainties related to



1. different aerosol models,
2. surface properties,
3. solar and viewing zenith angles, and
4. the wavelength shift

are analyzed.

**Aerosol model** Three sets of aerosol models are used in the numerical analysis. The first set (Set I) is taken from the MODIS aerosol retrieval algorithm, the second set (Set II) from the OMI aerosol retrieval algorithm, and the third set (Set III) includes the aerosol models obtained by a cluster analysis using the GOCART model. In Fig. 4.1 we illustrate the reflectances and the phase functions for the moderately absorbing (MODABS) aerosol model in Set I, weakly absorbing (WA1202) model in Set II, and a mixture of organic carbon and sulfate in Set III. For sets II and III, we considered three values for the weight of the coarse mode, i.e.,  $w_{\text{coarse}} = 0.0, 0.5, \text{ and } 1.0$ , and organic carbon i.e.,  $w_{\text{OC}} = 0.0, 0.5, \text{ and } 1.0$ , respectively

The plots show that

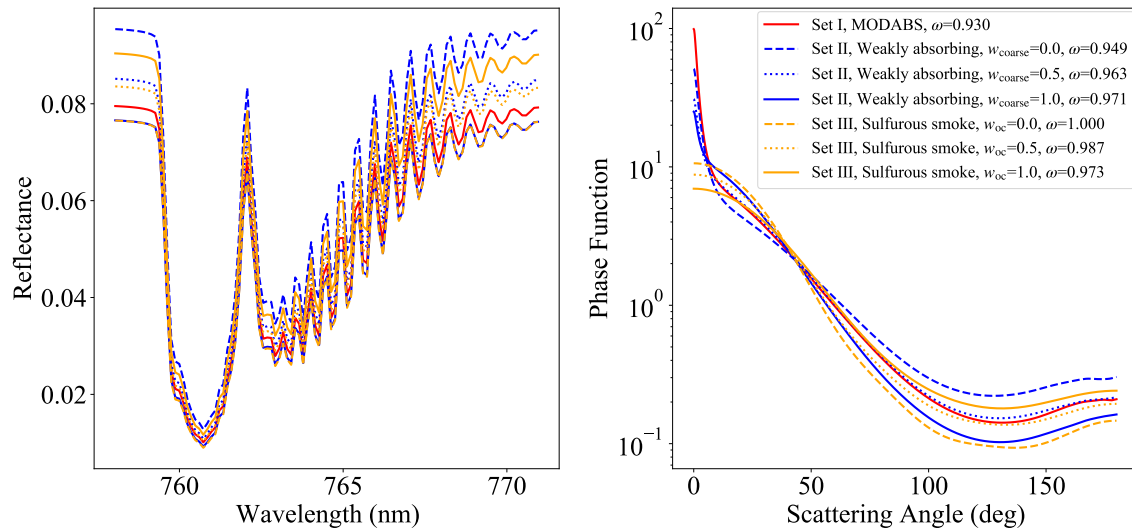
1. the reflectance for Set I lies between the highest and lowest reflectances for Sets II and III,
2. the reflectances for Set II with  $w_{\text{coarse}} = 1.0$  and Set III with  $w_{\text{OC}} = 0.0$  are similar,
3. the spectra agree well between 760 and 762 nm (where the main height information is contained), and poorly elsewhere,
4. significant differences in the phase functions are visible for small scattering angles.

In order to analyze the influence of the aerosol models on the retrieval results, we generate synthetic measurement spectra with the aerosol model MODABS in Set I, and perform the retrieval with the other two models. The plots in Figs. 4.2 and 4.3 show that an inappropriate aerosol model may result in significant retrieval errors. Specifically,

1. the error in the aerosol layer height reaches up to 1.9 km, when the true value is 9.5 km, while
2. the largest negative error in the aerosol optical depth is about  $-0.5$ , when the true value is 2.0.

**Surface albedo** Previous sensitivity studies [Nanda et al., 2018; Sanders et al., 2015] demonstrated that the accuracy in describing the surface properties can significantly influence the aerosol retrieval.

In general, the sensitivity of the signal with respect a retrieved parameter can be described by its partial derivative with respect to this parameter. The partial derivatives of the reflectance with respect to the aerosol layer height and optical depth as functions of surface albedo are illustrated in Fig. 4.4. The left panel in Fig. 4.4 shows that the derivatives decrease faster with decreasing the aerosol layer height (for the same values of the aerosol optical depth), while the right panel shows that the derivative curves do



**Figure 4.1:** Simulated reflectances versus the wavelength in the TROPOMI oxygen A-band (left), and the phase functions versus the scattering angle (right). The results correspond to the aerosol models from Sets I, II, and III, the aerosol optical depth  $\tau = 0.5$ , aerosol layer height  $H = 3.5$  km, surface albedo  $A = 0.05$ , solar zenith angle  $\theta_0 = 30^\circ$ , viewing zenith angle  $\theta = 0^\circ$ , and relative azimuth angle  $\Delta\varphi = 180^\circ$ . For sets II and III three values of 0.0, 0.5, and 1.0 for the weights of the coarse mode and organic carbon, respectively, were chosen.

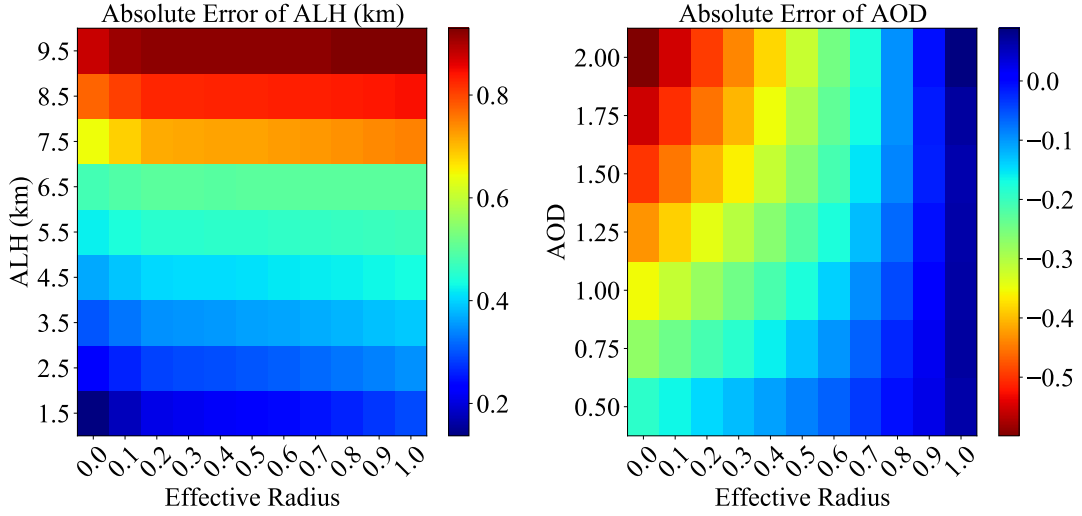
not differ significantly for the three values of the aerosol optical depth. A value of the surface albedo which is critical for the retrieval of the aerosol layer height and optical depth is the value for which the partial derivative of the reflectance with respect to the aerosol parameters is zero (dotted line). The plots show that the critical surface albedo is between 0.2 and 0.3 for both the aerosol layer height (orange line in the left panel of Fig. 4.4) and aerosol optical depth (blue line in the right panel of Fig. 4.4).

The relative retrieval errors resulting from the uncertainty in the surface albedo are shown in Fig. 4.5. The following conclusions can be drawn.

1. Over darker or less bright surface (with albedo values of 0.05 and 0.15), the uncertainty of the surface albedo produces less impact on the aerosol retrieval, while over a brighter surface (with albedo values of 0.5 and 0.9), the retrieval error is significantly higher.
2. When the value of the surface albedo is around the critical surface albedo, an error of 5% in the surface albedo can yield errors of about 180% (green line in the left panel of Fig. 4.5) and 80% (green line in the right panel of Fig. 4.5) in the retrieved aerosol layer height and optical depth respectively.

**Geometry** In Figs. 4.6 and 4.7 we illustrate the relative retrieval errors in the aerosol layer height and optical depth due to the relative errors in the solar and viewing zenith angles, respectively. The plots show that the relative errors increase with increasing

1. the solar zenith angle (for an error of 5% in the solar zenith angle at  $75^\circ$ , the relative errors in the retrieved aerosol layer height and optical depth reach up to 20% and 90%, respectively), and



**Figure 4.2:** Absolute retrieval errors in the aerosol layer height (ALH) and aerosol optical depth (AOD) for the weakly absorbing model in Set II. The synthetic measurement spectra are generated with the aerosol model MODABS in Set I. The parameters of the simulation are as in Fig. 4.1.

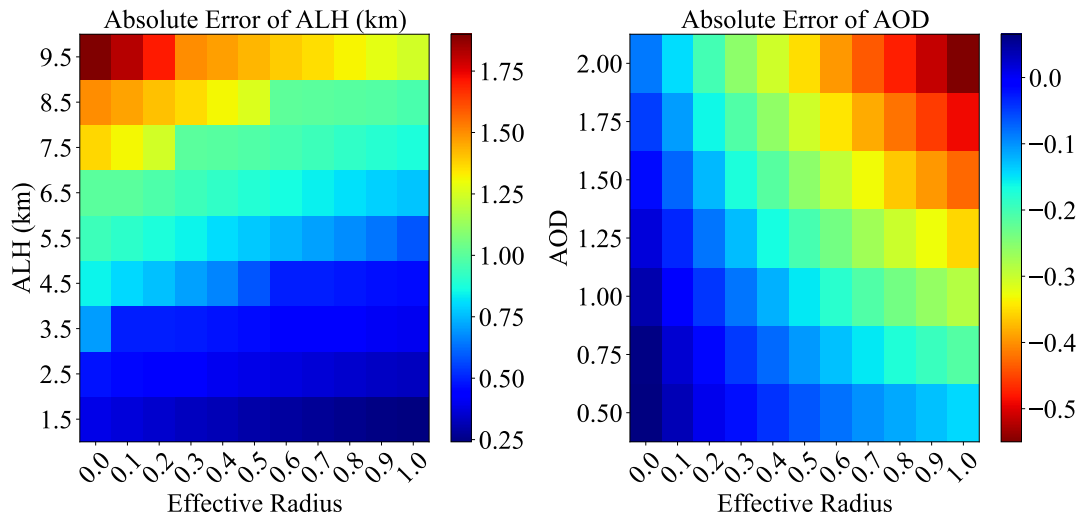
2. the viewing zenith angle (for an error of 5% in the viewing zenith angle at  $75^\circ$ , the relative errors in the retrieved aerosol layer height and optical depth reach up to 30% and 20%, respectively).

However, because the accuracy of the measured solar and viewing zenith angles is within 1%, we expect that they have a minor impact on the retrieval results.

**Wavelength calibration** An accurate wavelength calibration of the radiance and solar irradiance measurements is required during the Level-1b and Level-2 data processing. The reason is that wavelength uncertainties can introduce significant errors in the retrieval. Figs. 4.8 and 4.9 depict the relative retrieval errors in the aerosol layer height and optical depth together with the fit residuals versus the wavelength shift (an offset in the spectral position of a measured signal). The results correspond to two values of the surface albedo, namely 0.05 and 0.15, respectively. The plots show that

1. the residual increases monotonically with increasing the wavelength shift, and
2. the wavelength shift has a greater impact on
  - a) the retrieval over brighter surfaces, and
  - b) the layer height retrieval than the optical depth retrieval.

**Application to real data** In this section, we present the retrieval results from two real TROPOMI/S5P measurements taken on 22 June 2018 and 6 June 2020, respectively. Fig. 4.10 displays the corresponding true-color images from the Visible Infrared Imaging Radiometer (VIIRS) on the Suomi National Polar-orbiting Partnership (Suomi NPP) satellite, where the red rectangular region indicates the chosen TROPOMI scene. The first scene observes a part of Atlantic ocean near West Africa with latitudes between  $10^\circ$



**Figure 4.3:** The same as in Fig. 4.2 but for the sulfurous smoke model in Set III.

**Table 4.1:** Spectral characteristics of TROPOMI oxygen A-band measurements and the main input retrieval parameters. The aerosol models are the weakly absorbing (WA1202) and dust (DD3102) models in Set II.

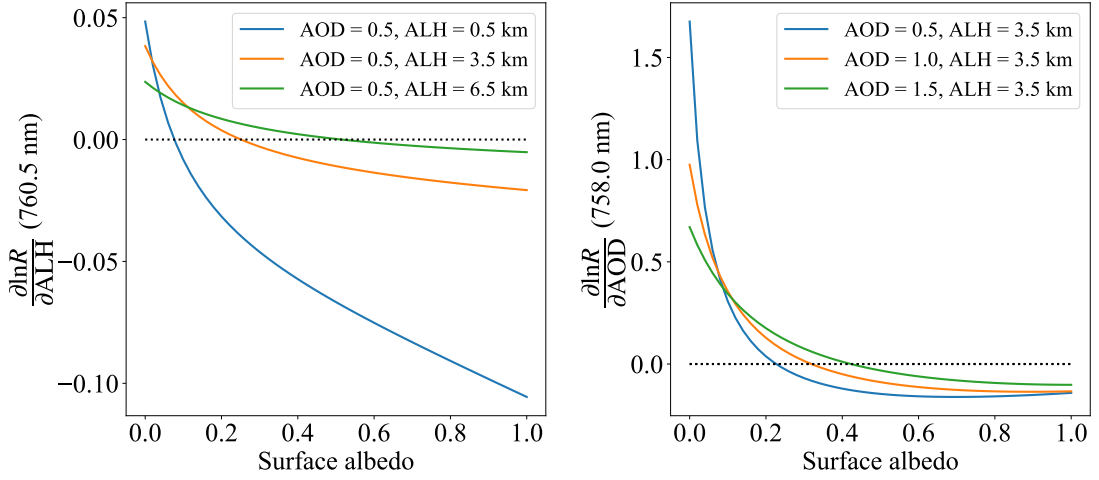
Parameter	Description
Band ID	6
Spectral range	758-771 nm
Spectral sampling	0.126 nm
Aerosol model	weakly absorbing and dust
Surface albedo	GE_LER
Cloud parameters	OCRA/ROCINN

and  $12^\circ$  N and longitudes between  $22^\circ$  and  $24^\circ$  W, while the second scene detects a desert dust aerosol case over the Sahara with latitudes between  $12^\circ$  and  $21^\circ$  N and longitudes between  $16^\circ$  and  $20^\circ$  W.

Table 4.1 describes the TROPOMI oxygen A-band measurements and the input retrieval parameters. In the retrieval, pixels with a cloud fraction greater than 0.15 are removed, and pixels with an aerosol index below 0.0 and a solar zenith angle larger than  $75^\circ$  are not processed. The cloud parameters and the surface albedo are taken from the operational TROPOMI cloud products (OCRA/ROCINN)[Loyola et al., 2018] and the geometry-dependent effective Lambertian equivalent reflectivity (GE\_LER) products [Loyola et al., 2020], respectively.

The retrieval results for the two scenes, plotted in Figs. 4.11 and 4.12, respectively, seem to capture the spatial patterns seen in Fig. 4.10 under different aerosol loading scenarios. In particular,

1. the retrieved aerosol layer height on 22 June 2018 is higher in the south than in the north region, and the lowest values of the aerosol layer height is in the northeast region, where the aerosol optical depth has high values;



**Figure 4.4:** Derivative of the reflectance spectrum with respect to the aerosol layer height (ALH) and aerosol optical depth (AOD) as functions of surface albedo.

2. although, due to a heavy cloud contamination, a small number of valid pixels are processed on 6 June 2020, the spatial distribution of desert dust aerosols is well described by the retrieved aerosol parameters.

For reference, the operational retrieval products processed by KNMI are plotted in the bottom panel of Fig. 4.11 and 4.12. As compared to the operational products,

1. the retrieved aerosol parameters capture nearly the same spatial patterns,
2. the aerosol layer heights are slightly underestimated, and
3. the aerosol optical depths are evidently lower.

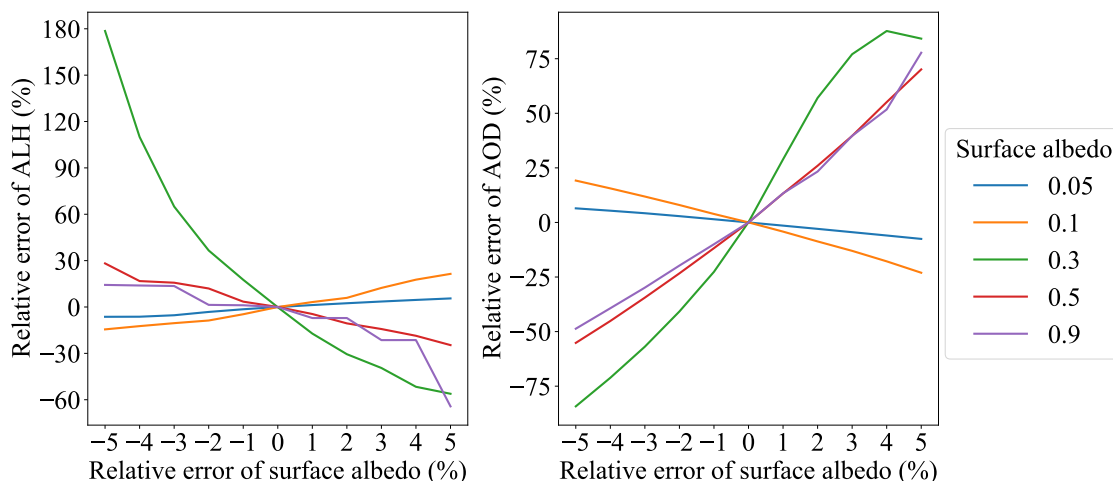
Possible reasons for these discrepancies are the different surface albedo data and aerosol microphysical properties used in the two retrieval algorithms.

Fig. 4.13 compares the observed and modeled reflectance spectra at the solution. From the plots, it can be seen that

1. the relative residuals are high near 760 nm, where the reflectance is rather low, and
2. for all pixels, the simulated and observed reflectance spectra at the solution are in a good agreement; this is an evidence for an overall good fit.

Finally, in Table 4.2, the propagated retrieval error due to errors in the surface albedo and wavelength calibration are listed. From these results we conclude that the uncertainty in the surface albedo has a significant impact on the retrieval.

**Conclusions** The objective of this section was to investigate the impact of forward and instrument model parameters on the retrieval. From our numerical analysis, the following conclusions can be drawn



**Figure 4.5:** Relative retrieval errors in the aerosol layer height (ALH) and aerosol optical depth (AOD) due to errors in the surface albedo. The results correspond to the aerosol model MODABS in Set I, the aerosol optical depth  $\tau = 0.5$ , and the aerosol layer height  $H = 3.5$  km.

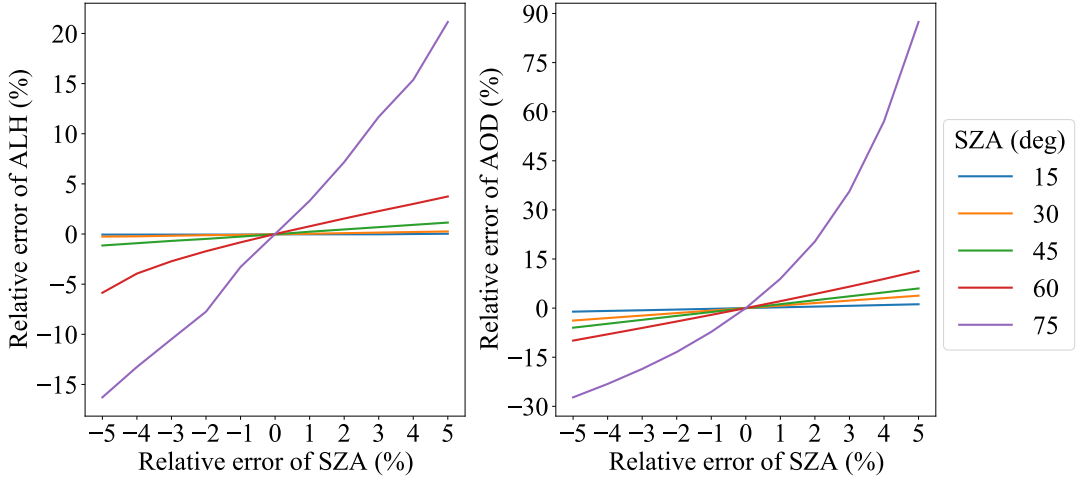
**Table 4.2:** Bias in the aerosol layer height (ALH) and aerosol optical depth (AOD) due to the errors in the surface albedo and wavelength calibration.

Parameter	Error	ALH bias	AOD bias
Surface albedo	0.01	0.1303 km	0.0918
Wavelength grid	0.007 nm	0.0793 km	0.0030

1. An inaccurate aerosol model can lead to a moderate loss of the retrieval accuracy. However, choosing an appropriate aerosol model is beneficial in the operational data processing.
2. Uncertainties in the surface albedo influence significantly the retrieval accuracy, particularly if the surface albedo is around the critical surface albedo.
3. Uncertainties in the solar and viewing zenith angles within 1% do not have a significant impact on the retrieval accuracy, provided that the zenith angles are below  $75^\circ$ .
4. The wavelength shift can potentially deteriorate the quality of the retrieved parameters (particularly, of the aerosol layer height), although it may have a minor impact on the fit residuals.

Retrievals using real TROPOMI/S5P data recorded on 22 June 2018 and 6 June 2020 have shown that

1. the retrieved aerosol parameters resemble the aerosol loading scenarios identified in the VIIRS images,
2. the simulated spectra at the solution approximate sufficiently well the observed ones,



**Figure 4.6:** Relative retrieval errors in the aerosol layer height (ALH) and aerosol optical depth (AOD) due to errors in the solar zenith angle (SZA).

3. an inaccurate surface albedo is supposed to be the most important error source in practice.

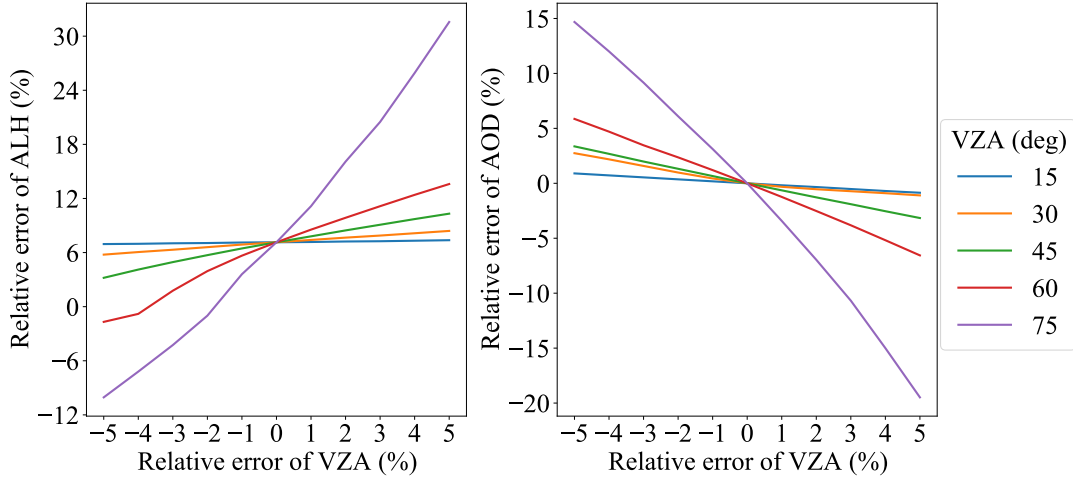
## 4.2 Retrieval algorithm with aerosol model selection

In the above retrieval algorithms, an aerosol model, chosen from a set of candidate models, is assumed to reflect the real scenario. However, model selection is not a trivial task because several models may fit the data equally well. The Bayesian approach offers the possibility to select the best fitting models from a set of candidate models without any prior seasonal or geographical information. Specifically, in the Bayesian model selection, the model with the highest evidence (i.e., the a posteriori probability densities of a model) is selected, while in the Bayesian model averaging [Hoeting et al., 1999], the retrieval results, corresponding to different candidate models weighted by their evidences, are combined. Note that Määttä et al. [Määttä et al., 2014] introduced the Bayesian approach into the aerosol model selection of the Ozone Monitoring instrument AEROSol (OMAERO) algorithm, Kauppi et al. [Kauppi et al., 2017] applied the Bayesian approach to real data of OMI, while Sasi et al. [Sasi et al., 2020a,b] applied the Bayesian approach to EPIC [Marshak et al., 2018] synthetic measurements. **The content of this section is adapted from the journal article in Appendix 2.**

### 4.2.1 Algorithm description

For model comparison, the so-called relative evidence of the  $m$ th aerosol model  $p(m | \bar{\mathbf{y}}^\delta)$  plays an important role. In view of Bayes' theorem, this conditional probability density is defined by

$$p(m | \bar{\mathbf{y}}^\delta) = \frac{p(\bar{\mathbf{y}}^\delta | m)p(m)}{p(\bar{\mathbf{y}}^\delta)} = \frac{p(\bar{\mathbf{y}}^\delta | m)}{\sum_{m=1}^{N_{\text{am}}} p(\bar{\mathbf{y}}^\delta | m)}, \quad (4.14)$$



**Figure 4.7:** The same as in Fig. 4.6 but for the errors in the viewing zenith angle (VZA).

where  $p(\bar{\mathbf{y}}^\delta | m)$  is the marginal likelihood density, and the last equality follows from the mean formula  $p(\bar{\mathbf{y}}^\delta) = \sum_{m=1}^{N_{\text{am}}} p(\bar{\mathbf{y}}^\delta | m)p(m)$  and the assumption that all aerosol models are equally like, i.e.,  $p(m) = 1/N_{\text{am}}$ . In terms of  $p(m | \bar{\mathbf{y}}^\delta)$ , the mean a posteriori density is given by

$$p_{\text{mean}}(\mathbf{x} | \bar{\mathbf{y}}^\delta) = \sum_{m=1}^{N_{\text{am}}} p(\mathbf{x} | \bar{\mathbf{y}}^\delta, m)p(m | \bar{\mathbf{y}}^\delta), \quad (4.15)$$

while the mean and the maximum solution estimates are defined by

$$\hat{\mathbf{x}}_{\text{mean}}^\delta = \sum_{m=1}^{N_{\text{am}}} \mathbf{x}_{m\hat{\alpha}}^\delta p(m | \bar{\mathbf{y}}^\delta), \quad (4.16)$$

and

$$\hat{\mathbf{x}}_{\text{max}}^\delta = \mathbf{x}_{m^*\hat{\alpha}}^\delta, \quad m^* = \arg \max_m p(m | \bar{\mathbf{y}}^\delta), \quad (4.17)$$

respectively. In Eq. (4.16), the Bayesian model averaging technique is used to combine the individual solutions weighted by their evidences, while in Eq. (4.15), the aerosol model with the highest evidence is considered to be the best among all aerosol models involved. Intuitively, we expect that in practice, a linear combination of the retrieved parameters corresponding to different candidate models will better reproduce the real scenario than an a priori selected model. The relative evidence  $p(m | \bar{\mathbf{y}}^\delta)$  can be computed as follows.

1. In a stochastic setting,  $p(m | \bar{\mathbf{y}}^\delta)$  can be calculated via Eq. (4.14) by using a linearization of the forward model around the solution, i.e.,

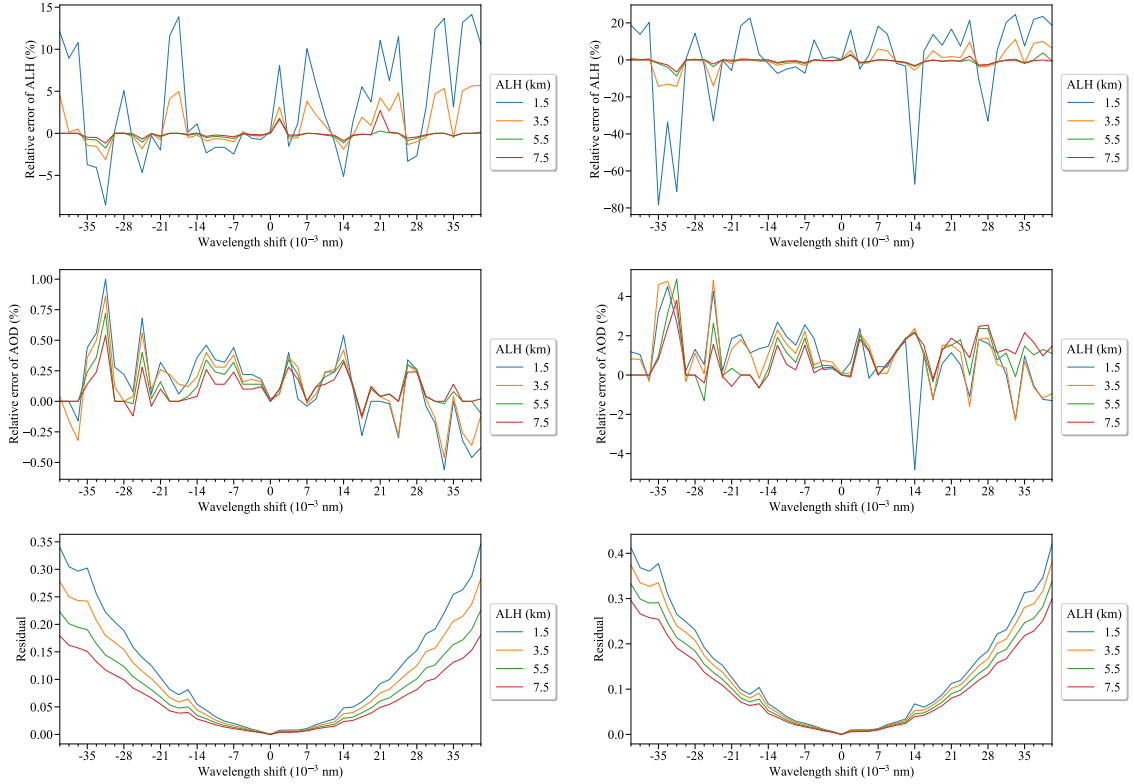
$$\bar{\mathbf{F}}_m(\mathbf{x}) = \bar{\mathbf{F}}_m(\mathbf{x}_{m\hat{\alpha}}^\delta) + \bar{\mathbf{K}}_{m\hat{\alpha}}(\mathbf{x} - \mathbf{x}_{m\hat{\alpha}}^\delta),$$

in which case, the marginal likelihood density computes as

$$p(\bar{\mathbf{y}}^\delta | m) = \sqrt{\frac{\det(\mathbf{I}_{N_{m\lambda}} - \hat{\mathbf{A}}_{m\hat{\alpha}})}{(2\pi\sigma_m^2)^{N_{m\lambda}}}} \times \exp\left[-\frac{1}{2\sigma_m^2} \bar{\mathbf{y}}_{m\hat{\alpha}}^{\delta\text{T}} (\mathbf{I}_{N_{m\lambda}} - \hat{\mathbf{A}}_{m\hat{\alpha}}) \bar{\mathbf{y}}_{m\hat{\alpha}}^\delta\right], \quad (4.18)$$



## 4.2 Retrieval algorithm with aerosol model selection



**Figure 4.8:** Relative retrieval errors in the aerosol layer height (ALH) and aerosol optical depth (AOD), as well as, the corresponding fit residuals, versus the wavelength shift. In the retrieval, the true values of the aerosol layer height are 1.5, 3.5, 5.5, and 7.5 km, while the true value of the aerosol optical depth is 0.5. The plots in the left panel correspond to a surface albedo of 0.05, while the plots in the right panel to a surface albedo of 0.15.

where

$$\bar{\mathbf{y}}_{m\hat{\alpha}}^\delta = \bar{\mathbf{y}}^\delta - \bar{\mathbf{F}}(\mathbf{x}_{m\hat{\alpha}}^\delta) + \bar{\mathbf{K}}_{m\hat{\alpha}}(\mathbf{x}_{m\hat{\alpha}}^\delta - \mathbf{x}_a),$$

and  $\hat{\mathbf{A}}_{m\hat{\alpha}} = \bar{\mathbf{K}}_{m\hat{\alpha}} \bar{\mathbf{K}}_{m\hat{\alpha}}^\dagger$ ,  $\bar{\mathbf{K}}_{m\hat{\alpha}}$  and  $\bar{\mathbf{K}}_{m\hat{\alpha}}^\dagger$  are the influence matrix, the Jacobian matrix, and the generalized inverse at the solution  $\mathbf{x}_{m\hat{\alpha}}^\delta$ , respectively. In Eq. (4.18),  $\sigma_m^2$  can be estimated in the framework of the maximum marginal likelihood estimation [Kendall et al., 1999; Patterson and Thompson, 1971; Smyth and Verbyla, 1996] as

$$\sigma_m^2 = \frac{1}{N_{m\lambda}} \bar{\mathbf{y}}_{m\hat{\alpha}}^{\delta T} (\mathbf{I}_{N_{m\lambda}} - \hat{\mathbf{A}}_{m\hat{\alpha}}) \bar{\mathbf{y}}_{m\hat{\alpha}}^\delta, \quad (4.19)$$

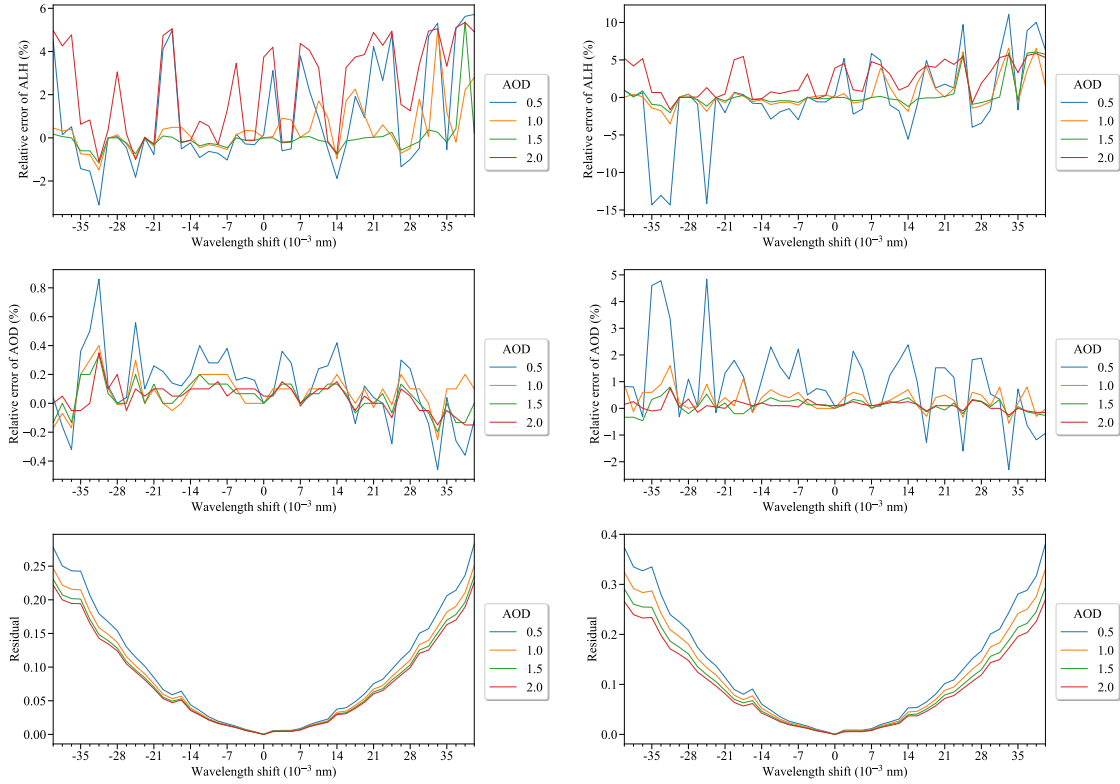
or in the framework of generalized cross-validation [Wahba, 1977, 1990] as

$$\sigma_m^2 = \frac{\|\mathbf{r}_{m\hat{\alpha}}^\delta\|^2}{[\text{trace}(\mathbf{I}_{N_{m\lambda}} - \hat{\mathbf{A}}_{m\hat{\alpha}})]}, \quad (4.20)$$

where  $\mathbf{r}_{m\hat{\alpha}}^\delta = \bar{\mathbf{y}}^\delta - \bar{\mathbf{F}}_m(\mathbf{x}_{m\hat{\alpha}}^\delta)$  the nonlinear residual vector.

2. In a deterministic setting,  $p(m | \bar{\mathbf{y}}^\delta)$ , regarded as a merit function characterizing the solution  $\mathbf{x}_{m\hat{\alpha}}^\delta$ , can be defined in terms of the marginal likelihood function or the

## 4 Bayesian-based retrieval algorithms



**Figure 4.9:** The same as in Fig. 4.8, but for the true aerosol optical depth values of 0.5, 1.0, 1.5, and 2.0, and a true aerosol layer height value of 3.5 km.

generalized cross-validation function. In the second case,  $p(m | \bar{\mathbf{y}}^\delta)$  is computed as

$$p(m | \mathbf{y}^\delta) = \frac{1/v(m)}{\sum_{m=1}^{N_{\text{am}}} 1/v(m)}, \quad (4.21)$$

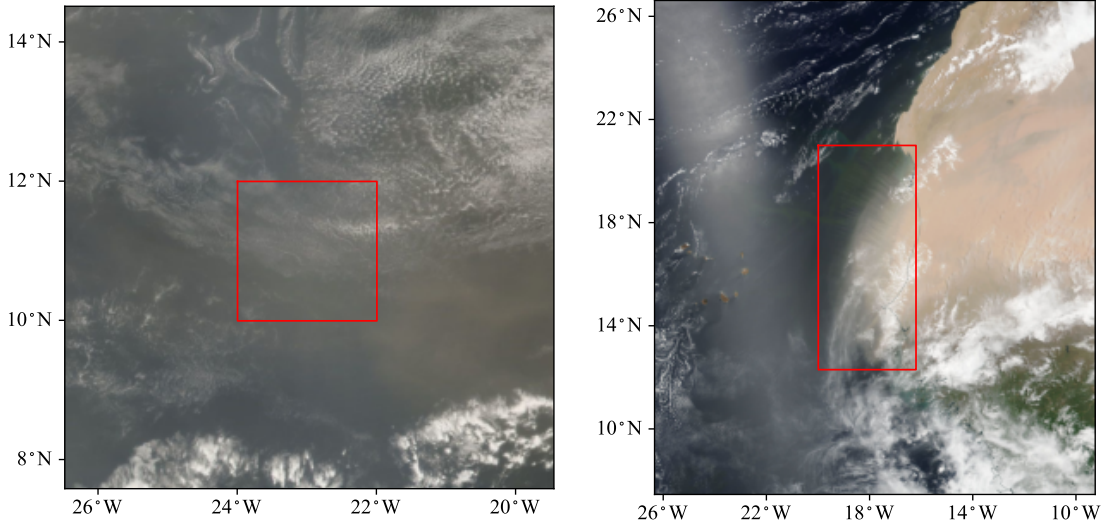
where

$$v(m) = \frac{\|\mathbf{r}_{m\hat{\alpha}}^\delta\|^2}{[\text{trace}(\mathbf{I} - \hat{\mathbf{A}}_{m\hat{\alpha}})]^2} \quad (4.22)$$

is the generalized cross-validation function.

### 4.2.2 Numerical analysis

Two sets of aerosol models are used in the numerical analysis. The first set (Set 1) is taken from the MODIS aerosol retrieval algorithm and includes the following four aerosol models: non-absorbing (NONABS), moderately absorbing (MODABS), absorbing (ABS), and desert dust (DUST), while the second set (Set 2) is taken from the OMI aerosol retrieval algorithm and includes the following five major aerosol types: weakly absorbing (WA), biomass burning (BB), desert dust (DD), marine (MA), and volcanic (VO). In the following, if not stated otherwise, the surface albedo, solar zenith angle, viewing zenith angle, and relative azimuth angle are  $A = 0.05$ ,  $\theta_0 = 60^\circ$ ,  $\theta = 0^\circ$ , and  $\Delta\varphi = 180^\circ$ , respectively.



**Figure 4.10:** VIIRS true-color images on 22 June 2018 (left) and 6 June 2020 (right). The red rectangular region displays the chosen TROPOMI scene.

**Test 1** In the first test example, synthetic measurement spectra are computed for each aerosol model included in Set 1 ( $m_e = \text{NONABS, MODABS, ABS, DUST}$ ), and for each measurement, all aerosol models from Set 1 are considered in the retrieval. Thus, the retrieval algorithm has the possibility of identifying the correct aerosol model. The exact aerosol optical depths and layer heights to be retrieved are  $\tau_e = 0.25, 0.5, 0.75, 1, 1.25, 1.75,$  and  $2$ , and  $H_e = 1.5, 2.5, 3.5, 4.5, 5.5, 6.5, 7.5, 8.5,$  and  $9.5$  km, respectively. The accuracy of the solution estimates is quantified through the relative errors

$$\varepsilon_{\text{mean}}^{\tau}(\tau_e, H_e) = \frac{|\tau_{\text{mean}} - \tau_e|}{\tau_e} \quad \text{and} \quad \varepsilon_{\text{mean}}^H(\tau_e, H_e) = \frac{|H_{\text{mean}} - H_e|}{H_e} \quad (4.23)$$

corresponding to (cf. Eq. (4.16))  $\hat{\mathbf{x}}_{\text{mean}}^{\delta} = [\tau_{\text{mean}}, H_{\text{mean}}]$  and

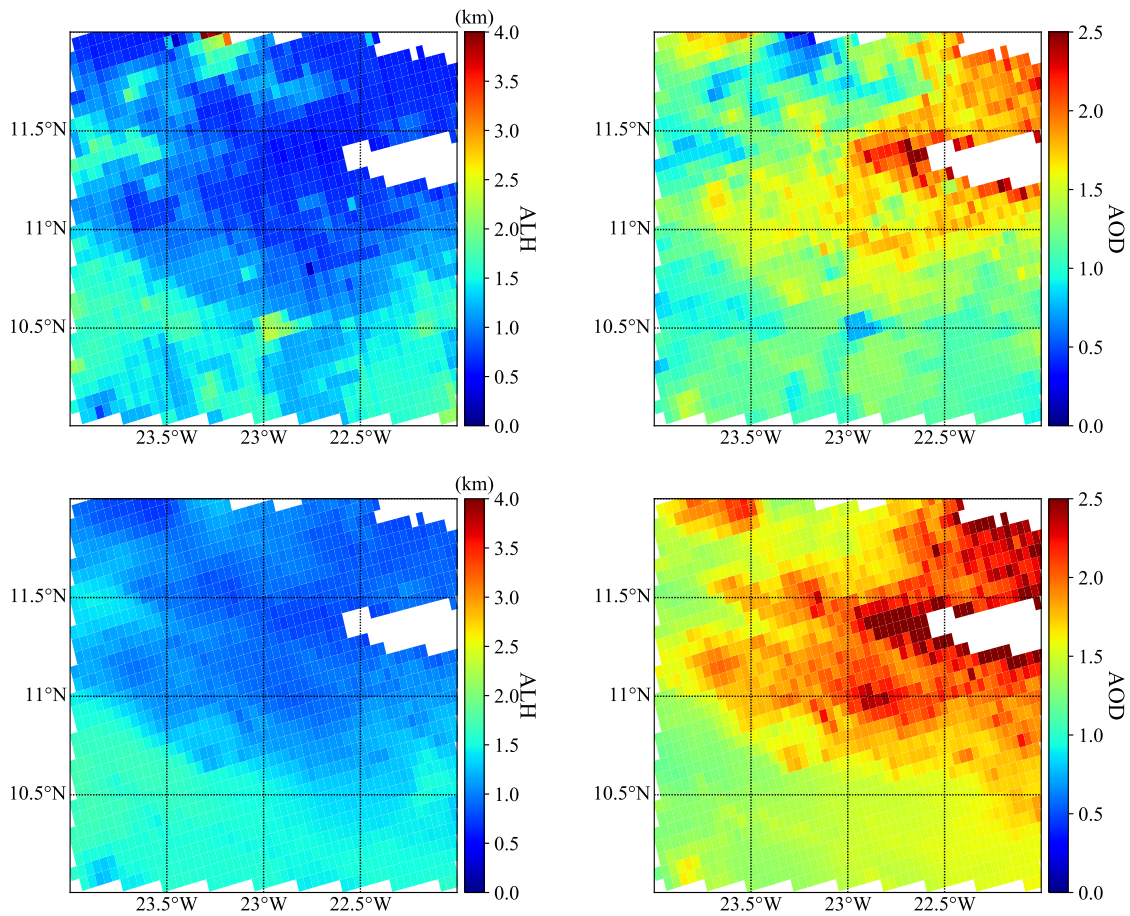
$$\varepsilon_{\text{max}}^{\tau}(\tau_e, H_e) = \frac{|\tau_{\text{max}} - \tau_e|}{\tau_e} \quad \text{and} \quad \varepsilon_{\text{max}}^H(\tau_e, H_e) = \frac{|H_{\text{max}} - H_e|}{H_e} \quad (4.24)$$

corresponding to (cf. Eq. (4.17))  $\hat{\mathbf{x}}_{\text{max}}^{\delta} = [\tau_{\text{max}}, H_{\text{max}}]$ , where  $\mathbf{x}_e = [\tau_e, H_e]$  is the exact solution.

In Fig. 4.14 and 4.15, we illustrate the variations of the relative errors with respect to the exact aerosol layer height  $H_e$  for  $\tau_e = 0.5$  (i.e.,  $\varepsilon_{\text{mean,max}}^{\tau,H}(\tau_e = 0.5, H_e)$ ), and the aerosol optical depth  $\tau_e$  for  $H_e = 3.5$  km (i.e.,  $\varepsilon_{\text{mean,max}}^{\tau,H}(\tau_e, H_e = 3.5 \text{ km})$ ), respectively. The following conclusions can be drawn:

1. The relative errors corresponding to the maximum solution estimate ( $\varepsilon_{\text{max}}^{\tau}$  and  $\varepsilon_{\text{max}}^H$ ) are significantly smaller than those corresponding to the mean solution estimate ( $\varepsilon_{\text{mean}}^{\tau}$  and  $\varepsilon_{\text{mean}}^H$ ). Because the errors are acceptable according to the scientific requirements defined in the pre-launch characterization tests, we conclude that the retrieval algorithm recognize correctly the exact aerosol model.
2. The maximum solution estimate of the aerosol optical depth is more accurate than the maximum solution estimate of the aerosol layer height.

#### 4 Bayesian-based retrieval algorithms

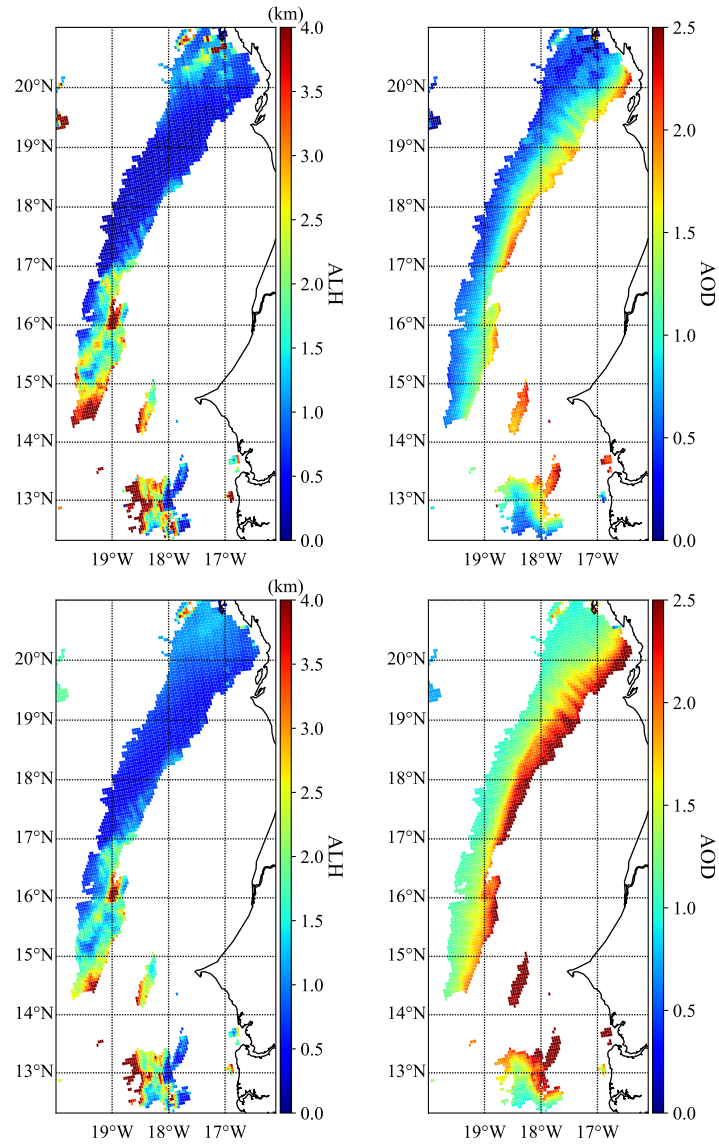


**Figure 4.11:** Retrieved aerosol layer height (ALH) and aerosol optical depth (AOD) for the TROPOMI scene on 22 June 2018. The results are obtained with the present retrieval algorithm (top panel) and the operational retrieval algorithm (bottom panel).

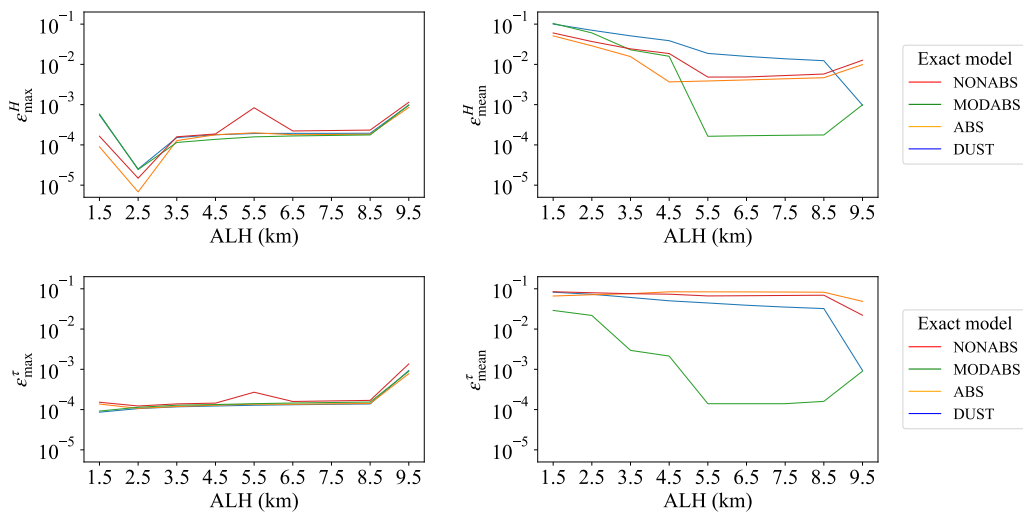
3. Because the inversion process is not perfect, an inappropriate aerosol model with unexpected large errors can occasionally be identified ( $\tau_e = 1, 1.25$ ).

The relative errors  $\varepsilon_{\max}^{\tau}(\tau_e, H_e = 3.5 \text{ km})$  and  $\varepsilon_{\max}^H(\tau_e, H_e = 3.5 \text{ km})$  attain their highest values (of about 0.22 and 0.016, respectively) for  $m_e = \text{MODABS}$  and  $\tau_e = 1, 1.25$ . To explain this result, we plot in Fig. 4.16 the a posteriori densities  $p(\mathbf{x} = [\tau, H] | \bar{\mathbf{y}}^\delta, m)$  for  $m = \text{NONABS}, \text{MODABS}, \text{ABS}, \text{DUST}$ , and the mean a posteriori densities  $p_{\text{mean}}(\mathbf{x} = [\tau, H] | \bar{\mathbf{y}}^\delta)$  in the case  $m_e = \text{MODABS}$ ,  $\tau_e = 1.25$  and  $H_e = 3.5 \text{ km}$ . The plots indicate that the a posteriori density for  $m = \text{DUST}$  is of similar height and width to that for  $m = \text{MODABS}$ ; the maximum solution is achieved at  $m = \text{MODABS}$ , i.e., the DUST appears to be the model with the highest evidence.

## 4.2 Retrieval algorithm with aerosol model selection

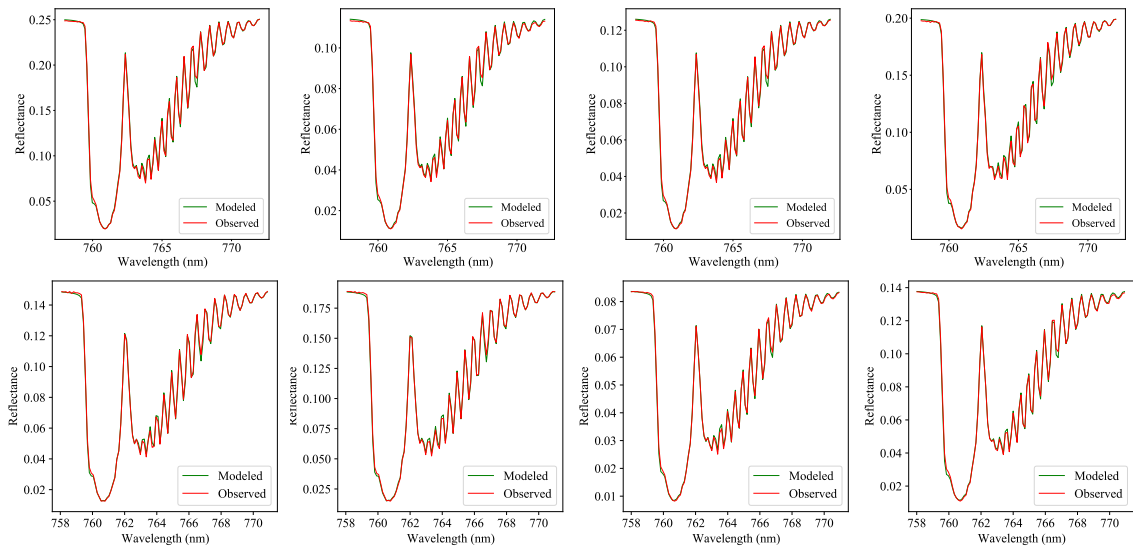


**Figure 4.12:** The same as in Fig. 4.11 but for the TROPOMI scene on 6 June 2020.

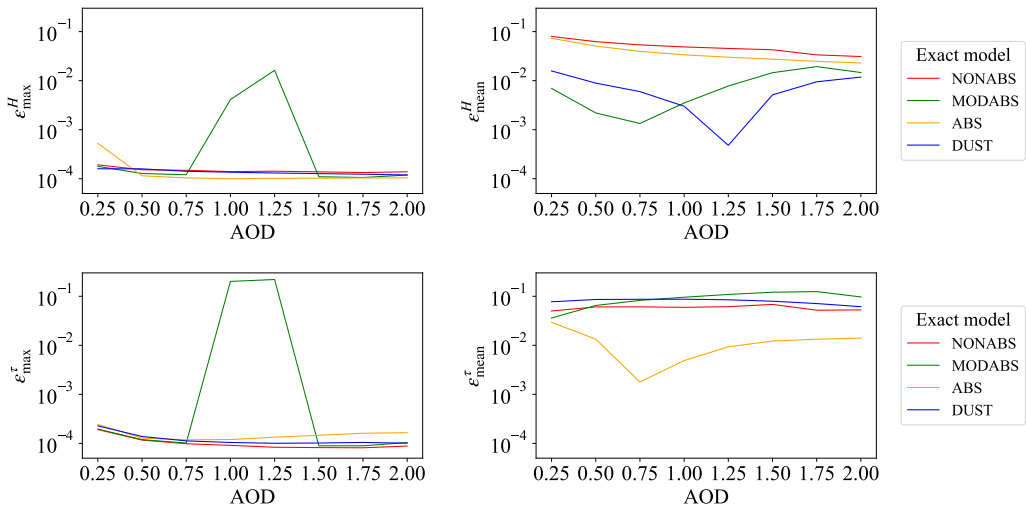


**Figure 4.14:** Relative errors  $\varepsilon_{\text{mean,max}}^{\tau,H}(\tau_e = 0.5, H_e)$  versus the aerosol layer height (ALH) for the aerosol models from Set 1.

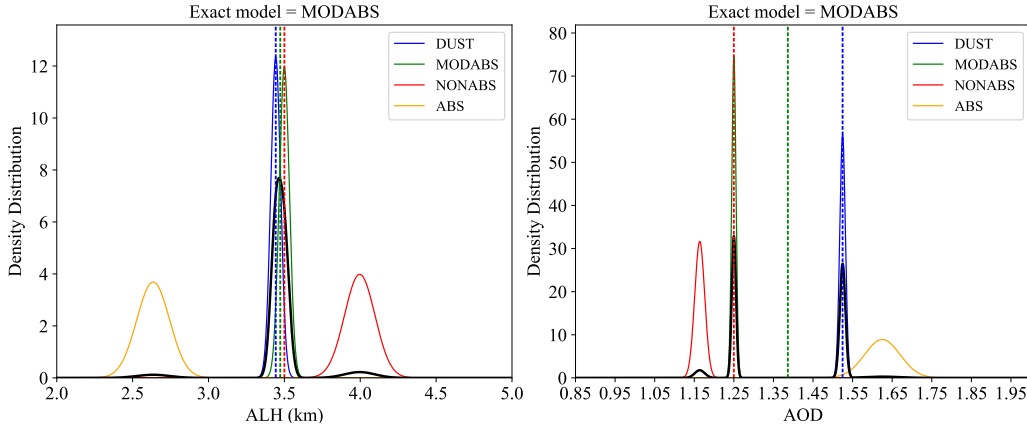
#### 4 Bayesian-based retrieval algorithms



**Figure 4.13:** Observed and modeled reflectance spectra at the solution for four (randomly chosen) TROPOMI ground pixels on 22 June 2018 (top panel) and 6 June 2020 (bottom panel).



**Figure 4.15:** Relative errors  $\epsilon_{\text{mean,max}}^{\tau,H}(\tau_e, H_e = 3.5 \text{ km})$  versus the aerosol optical depth (AOD) for the aerosol models from Set 1.



**Figure 4.16:** The a posteriori densities  $p(\mathbf{x} = [\tau, H] | \bar{\mathbf{y}}^\delta, m)$  for  $m = \text{NONABS}, \text{MODABS}, \text{ABS}$ , and  $\text{DUST}$ , and the mean a posteriori densities  $p_{\text{mean}}(\mathbf{x} = [\tau, H] | \bar{\mathbf{y}}^\delta)$  in the case  $m_e = \text{MODABS}$ ,  $\tau_e = 1.25$ , and  $H_e = 3.5$  km. The black curve indicates the mean a posteriori density. In each plot, the red vertical dashed line corresponds to the exact values to be retrieved  $(\tau_e, H_e)$ , the blue vertical dashed line to the maximum solution estimate  $(\tau_{\text{max}}, H_{\text{max}})$ , and the green dashed line to the mean solution estimates  $(\tau_{\text{mean}}, H_{\text{mean}})$ .

**Test 2** In the second test example, synthetic measurement spectra are computed for each aerosol model included in Set 1 ( $m_e = \text{NONABS}, \text{MODABS}, \text{ABS}, \text{DUST}$ ), and for each measurement, all aerosol models from Set 2 are considered in the retrieval. The mean solution estimate and the mean a posteriori density are computed for the first 10 aerosol models with the highest evidence.

The variations of the relative errors with respect to the exact aerosol layer height  $H_e$  for  $\tau_e = 0.5$  and the aerosol optical depth  $\tau_e$  for  $H_e = 3.5$  km are illustrated in Fig. 4.17 and 4.18, respectively. The plots show that

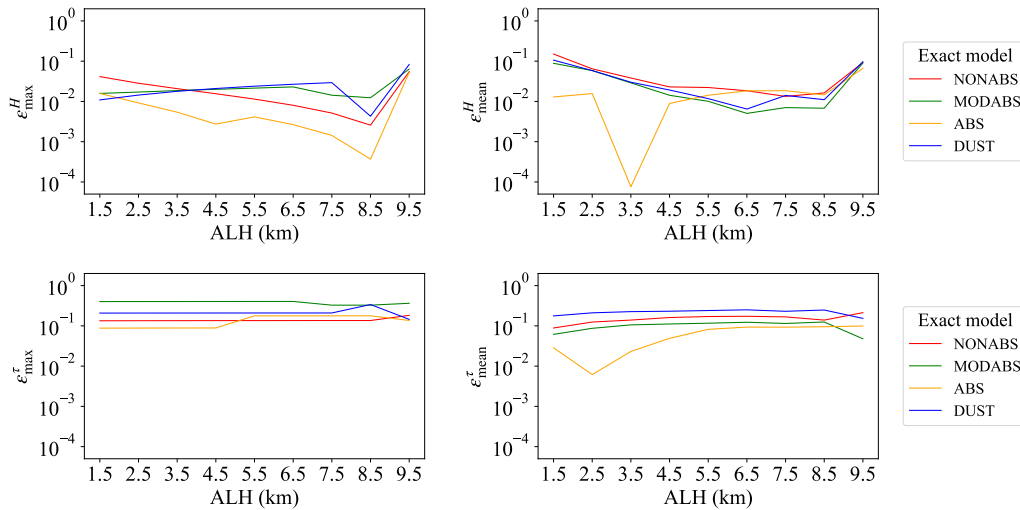
1. the relative errors are larger than those in the first test example,
2. the relative errors corresponding to the maximum solution estimate ( $\varepsilon_{\text{max}}^\tau$  and  $\varepsilon_{\text{max}}^H$ ) and the mean solution estimate ( $\varepsilon_{\text{mean}}^\tau$  and  $\varepsilon_{\text{mean}}^H$ ) are comparable, and
3. on average, the retrieved aerosol layer height has a higher accuracy than the retrieved aerosol optical depth.

The mean a posteriori densities  $p_{\text{mean}}(\mathbf{x} = [\tau, H] | \bar{\mathbf{y}}^\delta)$  are shown in Fig. 4.19 and 4.20 for  $\tau_e = 0.5$ ,  $H_e = 3.5$  km, and all exact aerosol models  $m_e = \text{NONABS}, \text{MODABS}, \text{ABS}$ , and  $\text{DUST}$ . The following conclusions can be drawn:

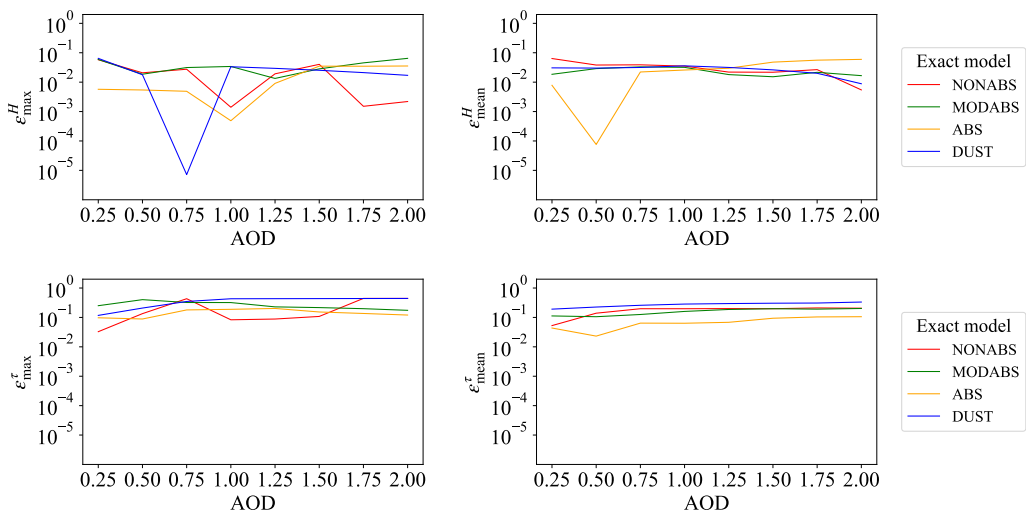
1.  $H_{\text{mean}}$  and  $H_{\text{max}}$  are both not too far from  $H_e$ ; thus, for aerosol layer height retrieval, the maximum solution estimate and the mean solution estimate ( $\varepsilon_{\text{mean}}^\tau$  and  $\varepsilon_{\text{mean}}^H$ ) have similar accuracies;
2.  $\tau_{\text{mean}}$  is relatively closer to  $\tau_e$  than  $\tau_{\text{max}}$ ; thus, for aerosol optical depth retrieval, the mean solution estimate performs better than the maximum solution estimate;
3. aerosol layer height retrievals have wide a posteriori densities that cover the exact layer height; and

#### 4 Bayesian-based retrieval algorithms

- aerosol optical depth retrievals have multi-peaked densities, in which the exact optical depth does not have the highest probability.

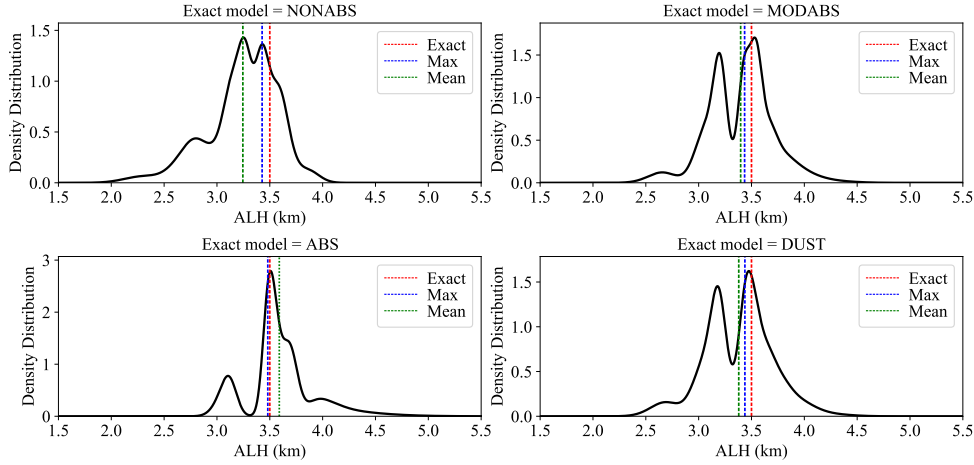


**Figure 4.17:** Relative errors  $\varepsilon_{\text{mean,max}}^{\tau,H}(\tau_e = 0.5, H_e)$  for the aerosol models from Set 2.

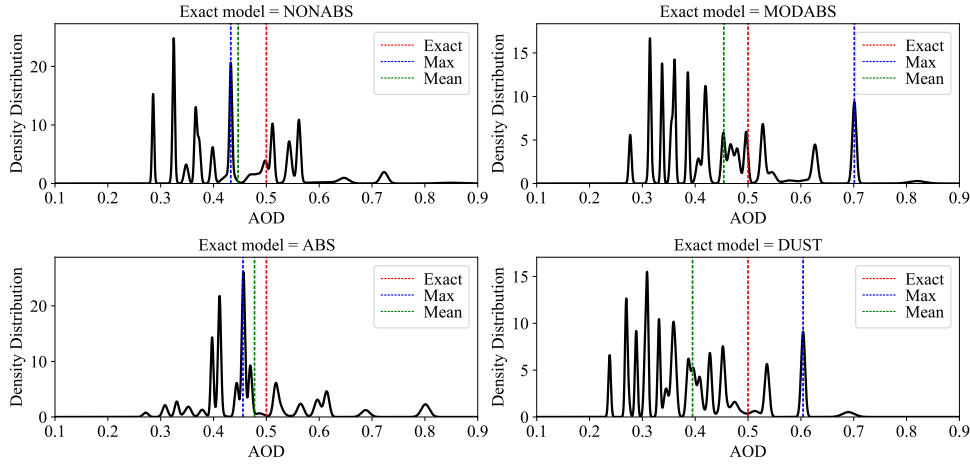


**Figure 4.18:** Relative errors  $\varepsilon_{\text{mean,max}}^{\tau,H}(\tau_e, H_e = 3.5 \text{ km})$  for the aerosol models from Set 2.





**Figure 4.19:** The mean a posteriori densities  $p_{\text{mean}}(H | \bar{\mathbf{y}}^\delta)$  for  $m_e = \text{NONABS}, \text{MODABS}, \text{ABS}, \text{DUST}$ ,  $\tau_e = 0.5$ , and  $H_e = 3.5$  km. In each plot, the red vertical dashed line correspond to the exact values to be retrieved  $(\tau_e, H_e)$ , the blue vertical dashed line to the maximum solution estimate  $(\tau_{\text{max}}, H_{\text{max}})$ , and the green dashed line to the mean solution estimates  $(\tau_{\text{mean}}, H_{\text{mean}})$ .



**Figure 4.20:** The same as in Figure 4.19 but for the mean a posteriori densities  $p_{\text{mean}}(\tau | \bar{\mathbf{y}}^\delta)$ .

**Application to real data** To test the performance of the retrieval algorithm on real TROPOMI/S5P data, we choose a wild fire scene in South Africa and consider the measurements recorded on 4 and 5 July 2019. As it can be seen from the corresponding VIIRS images in Fig. 4.21, the wild fire smoke on 4 July 2019 traveled beyond the coastline and extended over the ocean, so that the smoke on 5 July 2019 was thinner.

In the retrieval, the aerosol models included in Sets 1 and 2 are considered, the ground pixels with cloud fraction larger than 0.15 are excluded, and the geometry-dependent effective Lambertian equivalent reflectivity (GE\_LER) products are used.

A synthetic presentation of the results of our numerical simulation is given below.

#### 4 Bayesian-based retrieval algorithms

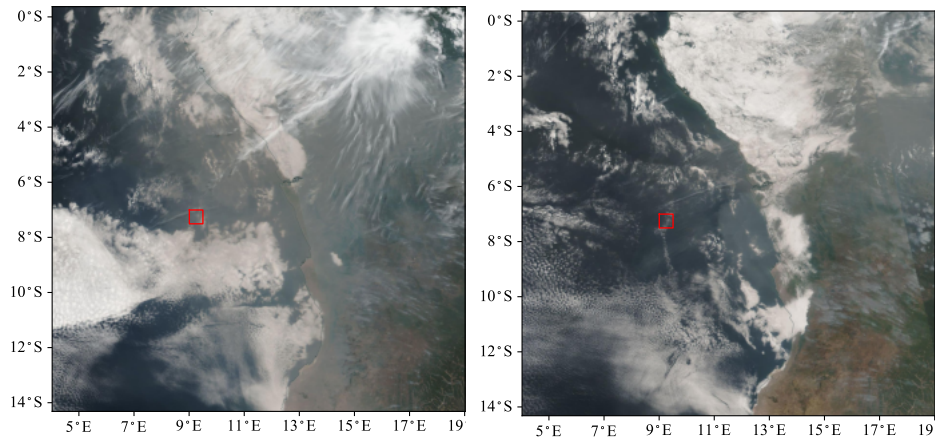
1. For the retrieval on 4 July 2019, we show in Fig. 4.22 the aerosol model with the highest evidence from Set 1, as well as, the aerosol type containing the aerosol model with the highest evidence from Set 2. The most likely models are absorbing (ABS) from Set 1 and biomass burning (BB) from Set 2. The model evidence for each aerosol model from Set 1 is shown in Figure 4.23. Observe that (i) the differences between the model evidences for the four aerosol models are not very large, and (ii) the model evidence of absorbing aerosols was slightly larger than those of the other models. In Set 2, there are 26 aerosol models and five aerosol types. The sum of the first 10 best aerosol model evidences for each aerosol type from Set 2 are illustrated in Fig. 4.24. The most probable type is biomass burning. In conclusion, the most plausible aerosol models identified by the algorithm are of biomass burning type, and this strongly absorbing aerosol type is consistent with the thick smoke observed in the true-color image.
2. For the retrieval on 5 July 2019, the predominant models are absorbing (ABS) and dust (DUST) among Set 1 (Fig. 4.25a), and biomass burning (BB) and desert dust (DD) among Set 2 (Fig. 4.25b). Thus, in addition to the aerosol models identified on 4 July 2019, the dust aerosol model comes into play. The model evidence for each aerosol model from Set 1 and each aerosol type from Set 2 are displayed in Figs. 4.26 and 4.27, respectively. In conclusion, compared with the result on 4 July 2019, the dominance of biomass burning aerosol type (ABS and BB) is less obvious. Taking into account the thinner smoke on 5 July 2019 and the long traveling distance from the origin, the presence of a less absorbing mixture of aerosol types (biomass burning and dust) seems to be plausible.
3. The maximum and mean solution estimates are shown in Figs. 4.28 and 4.30 for Set 1, and Figs. 4.29 and 4.31 for Set 2.

Based on the above results, we see that

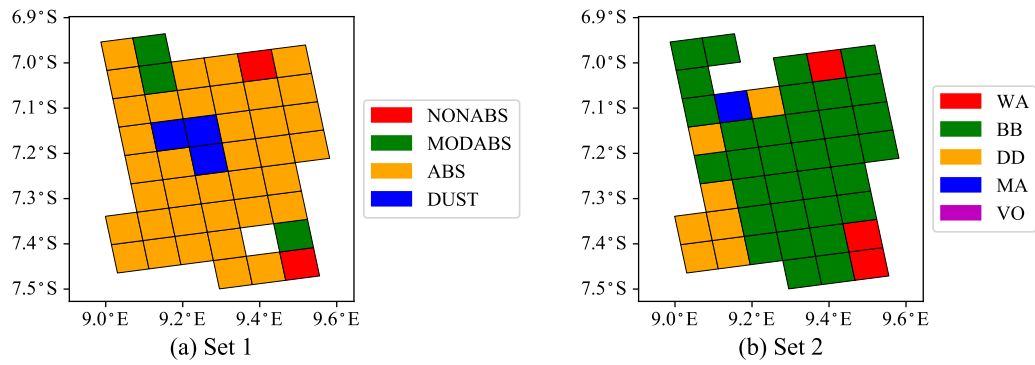
1. the most likely model varies from pixel to pixel, indicating that in some cases a "wrong" model can be chosen (this is consistent with the findings using synthetic data),
2. the mean solution estimates show a slightly smoother spatial pattern than the maximum solution estimates, and
3. despite the differences in the micro-physical properties of the aerosol models from Sets 1 and 2, the spatial distributions of the mean solution estimate are comparable.

Note that on a Dell desktop (with 12 processors at 3.2 GHz, 31.2 GB of RAM) the retrieval took less than 10 min (10–15 iterations in total) for all models included in Set 1, and longer than 60 min (approximately 100 iterations in total) for all models included in Set 2.

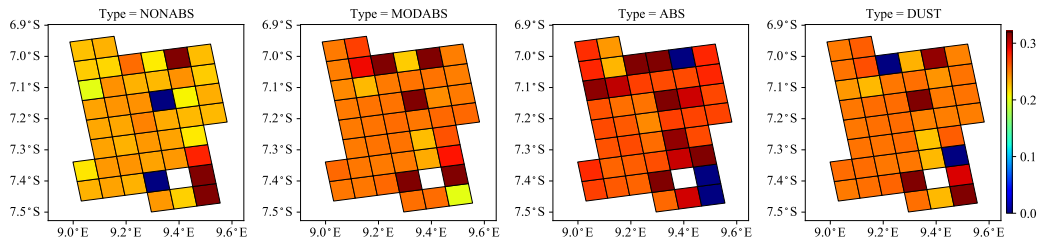
## 4.2 Retrieval algorithm with aerosol model selection



**Figure 4.21:** True-color VIIRS images recorded on 4 (left) and 5 (right) July 2019.

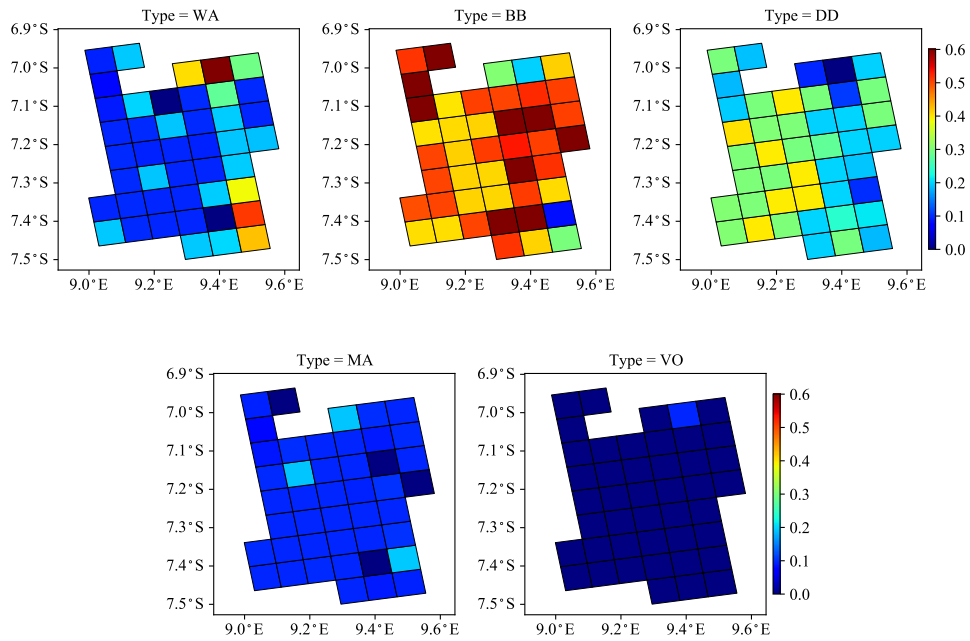


**Figure 4.22:** The aerosol model with the highest evidence from Set 1 (a), and the aerosol type containing the aerosol model with the highest evidence from Set 2 (b). The TROPOMI spectra were recorded on 4 July 2019.

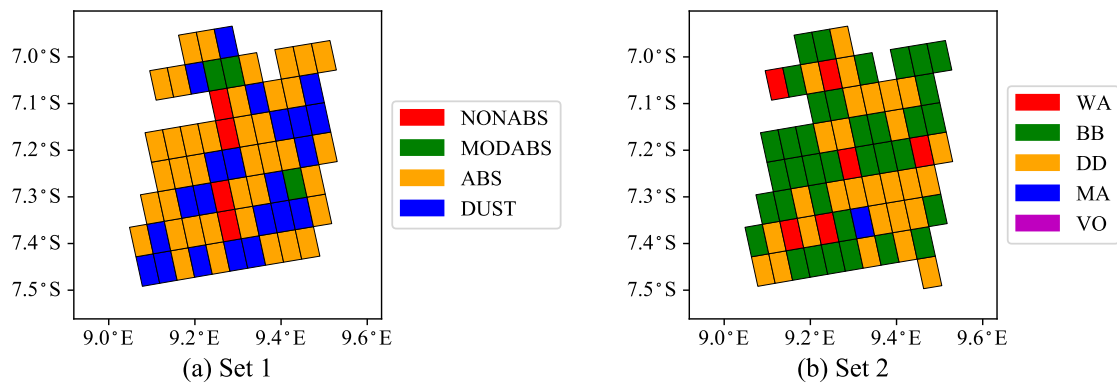


**Figure 4.23:** The model evidence for each aerosol model from Set 1. The TROPOMI spectra were recorded on 4 July 2019.

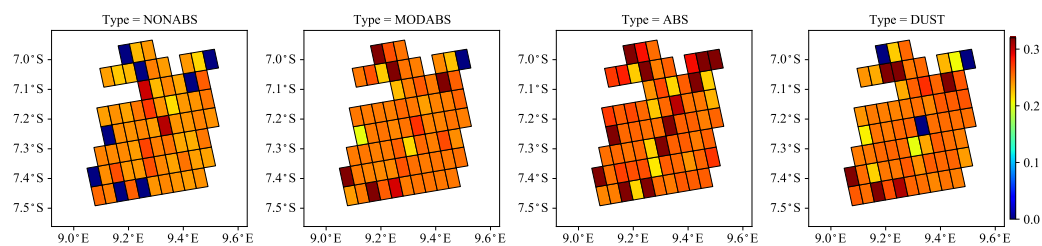
#### 4 Bayesian-based retrieval algorithms



**Figure 4.24:** The sum of the first 10 best aerosol model evidences for each aerosol type from Set 2. The TROPOMI spectra were recorded on 4 July 2019.

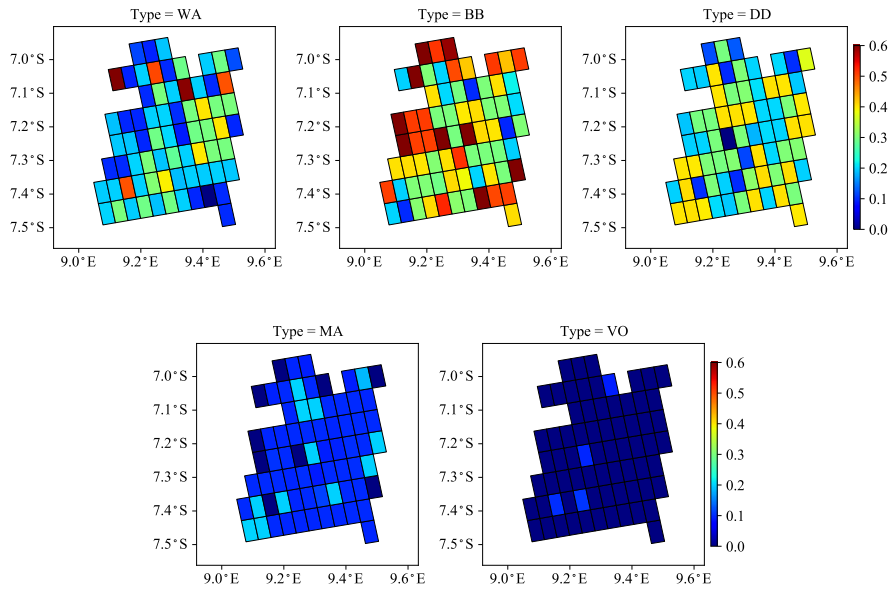


**Figure 4.25:** The same as in Fig. 4.22 but for the data on 5 July 2019.

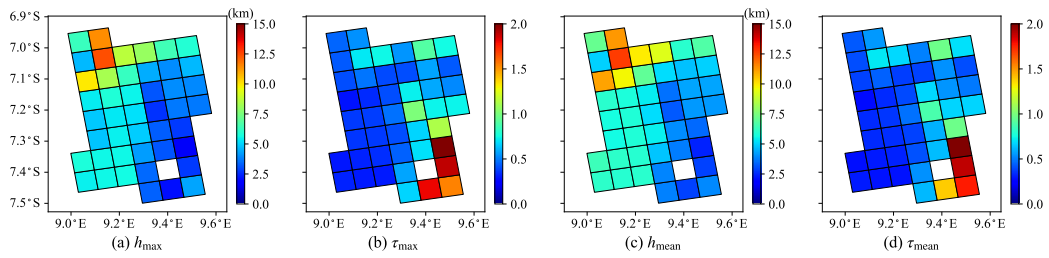


**Figure 4.26:** The same as in Fig. 4.23 but for the data on 5 July 2019.

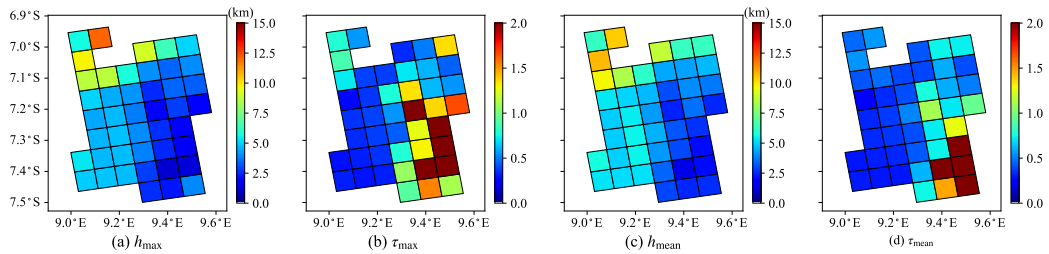
## 4.2 Retrieval algorithm with aerosol model selection



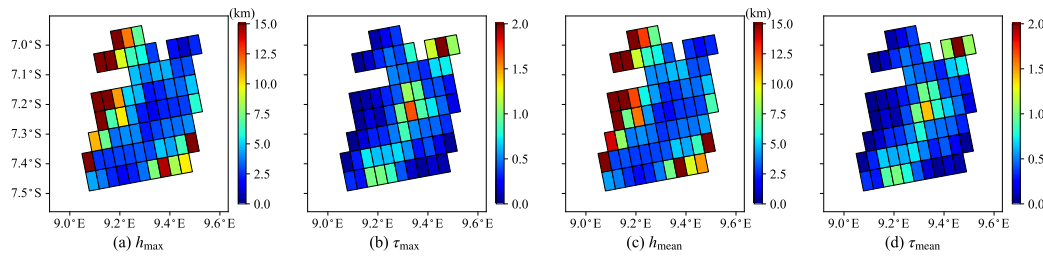
**Figure 4.27:** The same as in Fig. 4.24 but for the data on 5 July 2019.



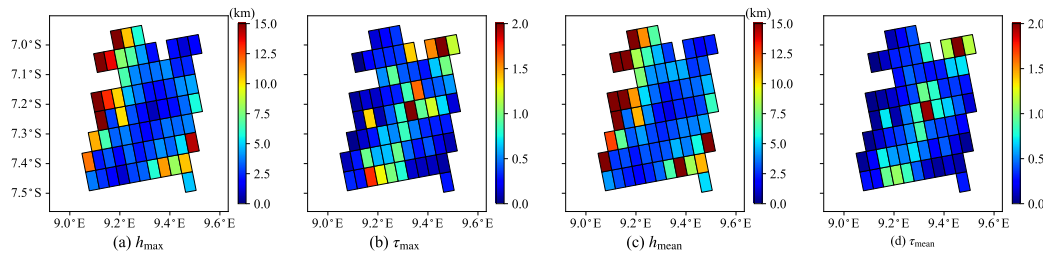
**Figure 4.28:** The maximum solution estimates ( $H_{\max}$ ,  $\tau_{\max}$ ) and the mean solution estimates ( $H_{\text{mean}}$ ,  $\tau_{\text{mean}}$ ) for Set 1 and data on 4 July 2019.



**Figure 4.29:** The maximum solution estimates ( $H_{\max}$ ,  $\tau_{\max}$ ) and the mean solution estimates ( $H_{\text{mean}}$ ,  $\tau_{\text{mean}}$ ) for Set 2 and data on 4 July 2019.



**Figure 4.30:** The same as in Fig. 4.28 but for the data on 5 July 2019.



**Figure 4.31:** The same as in Fig. 4.29 but for the data on 5 July 2019.

**Conclusions** In this section, the results of aerosol retrieval computed by means of a Bayesian algorithm that takes into account the uncertainty in aerosol model selection are presented. The algorithm is applied to the retrieval of aerosol optical depth and layer height from synthetic and real TROPOMI/S5P data by considering two sets of aerosol models; these are taken from the MODIS aerosol retrieval algorithm the OMI Multiwavelength aerosol retrieval algorithm.

Through a numerical analysis, we come to the following findings.

1. When the exact aerosol model, for which synthetic data are generated, is included in the set of candidate models, the relative errors corresponding to the maximum solution estimate are relatively small. When this is not the case, it is likely that several aerosol models are able to fit the data equally well. In such situations, the mean solution estimate has a smaller bias than the maximum solution estimate.
2. For the real measurements on 4 July 2019 taken over a wild fire scene in South Africa, the absorbing aerosol model from Set 1 and the biomass burning aerosol type from Set 2 are found to be the most plausible. This result is in agreement with the thick smoke observed in the true-color image. For the thinner smoke scenario on 5 July 2019, the above models together with the dust aerosol model are found to be the most probable aerosol models. Actually, no dominant aerosol model, but rather a less absorbing mixture of different aerosol types, is identified in this case. The mean and maximum solution estimates have a similar spatial distribution, but the mean solution estimates have a more continuous spatial pattern.
3. A definitive choice between Sets 1 and 2 as possible candidate models cannot be made, and a suitable aerosol model seems to be problem dependent.

Note that when applying the Bayesian approach, we have to perform a retrieval for each candidate model. For this reason, the retrieval algorithm is computationally expensive,

## 4.2 *Retrieval algorithm with aerosol model selection*

especially for Set 2 which contains a larger number of aerosol models. To enhance the retrieval efficiency, a machine learning-based retrieval algorithm will be designed.





## 5 Neural network retrieval algorithms

In Bayesian-based retrieval algorithms, the computation of the forward model and its Jacobian are computational expensive. Therefore, these approaches are not efficient for the operational processing of large volumes of data. In contrast, an inversion method based on neural networks is very efficient because it can provide accurate predictions of the forward model and its Jacobian in a fraction of millisecond. **The content of this section is adapted from the journal article in Appendix 3.**

In atmospheric remote sensing, neural networks have been designed to

1. emulate a radiative transfer model, which is then used in conjunction with a Bayesian approach to solve the inverse problem [Fan et al., 2019, 2017; Gao et al., 2021; Shi et al., 2020],
2. learn the retrieval mappings directly from data [Brath et al., 2018; Håkansson et al., 2018; Holl et al., 2014; Jiménez et al., 2003; Strandgren et al., 2017; Wang et al., 2017], and
3. recover the atmospheric parameters of interest, which are then used as initial guess in an optimization algorithm [Di Noia et al., 2015; Noia et al., 2017].

The TROPOMI/S5P operational retrieval algorithm in the O<sub>2</sub> A-band uses a neural network based forward model and a Bayesian inversion approach [Nanda et al., 2018; Sanders and de Haan, 2013; Sanders et al., 2015]. For this instrument, the difficulty in designing neural networks for learning the inverse model consists in the fact that each swath row (angle)  $r$  with  $r = 1, \dots, \mathcal{R}$ , is characterized by its own measurement wavelength grid  $\{\lambda_{mk}^r\}_{k=1}^{N_{m\lambda}}$  and slit function  $g_r$ , where  $\mathcal{R}$  is the number of swath rows and  $N_{m\lambda}$  the number of spectral grid points. This is due to the point spread function, aberrations, and defocussing of the optical system and small changes in the width of the slit. Specifically, for an aerosol model  $m$ , the noisy and synthetic spectral signals measured by the instrument at a wavelength  $\lambda_{mk}^r$  are given respectively, by

$$\mathcal{I}_{\text{mes}}^\delta(\lambda_{mk}^r) = \mathcal{I}_{\text{sim}}^{(m)}(\lambda_{mk}^r) + \delta_{mk}, \quad (5.1)$$

$$\mathcal{I}_{\text{sim}}^{(m)}(\lambda_{mk}^r) = \int g_r(\lambda_{mk}^r - \lambda) I_{\text{sim}}^{(m)}(\lambda) d\lambda, \quad (5.2)$$

where  $I_{\text{sim}}^{(m)}(\lambda)$  is the radiance computed by a radiative transfer model at a forward wavelength  $\lambda$  (before convolution) and  $\delta_{mk}$  the (row-independent) measurement noise.

To explain the neural network technique, we consider a generic model  $\mathbf{y} = \mathbf{F}(\mathbf{x})$ , where  $\mathbf{x} \in \mathbb{R}^{N_x}$  is the input vector,  $\mathbf{F}$  a deterministic function, and  $\mathbf{y} \in \mathbb{R}^{N_y}$  the output vector, and intend to approximate  $\mathbf{F}(\mathbf{x})$  by a neural network model  $\mathbf{f}(\mathbf{x}, \boldsymbol{\omega})$  characterized by a set of parameters  $\boldsymbol{\omega}$ . Deep neural networks consist of units (nodes) arranged in an input layer, an output layer, and a number of hidden layers. For a neural network with  $L + 1$

layers and  $N_l$  units in layer  $l$ , the feed-forward operations can be written in matrix form as

$$\mathbf{y}_0 = \mathbf{x}, \quad (5.3)$$

$$\mathbf{u}_l = \mathbf{W}_l \mathbf{y}_{l-1} + \mathbf{b}_l, \quad (5.4)$$

$$\mathbf{y}_l = \phi_l(\mathbf{u}_l), \quad l = 1, \dots, L, \quad (5.5)$$

$$\mathbf{f}(\mathbf{x}, \boldsymbol{\omega}) = \mathbf{y}_L, \quad (5.6)$$

where  $l = 0$  is the input layer,  $l = L$  the output layer,  $\phi_l$  the layer activation function,  $\mathbf{W}_l \in \mathbb{R}^{N_l \times N_{l-1}}$  the matrix of weights connecting the layers  $l-1$  and  $l$ ,  $\mathbf{b}_l \in \mathbb{R}^{N_l}$  the vector of biases corresponding to layer  $l$ , and  $\boldsymbol{\omega} = \{\mathbf{W}_l, \mathbf{b}_l\}_{l=1}^L$  the set of network parameters. The process of computing the network parameters  $\boldsymbol{\omega}$  on a data set  $\mathcal{D} = \{(\mathbf{x}^{(n)}, \mathbf{y}^{(n)})\}_{n=1}^N$ , where  $\mathbf{y}^{(n)} = \mathbf{F}(\mathbf{x}^{(n)})$  and  $N$  is the number of samples, represents the so-called deep learning process. Actually, a point estimate  $\hat{\boldsymbol{\omega}}$  is computed as the minimizer of a loss function, defined as the log likelihood of the data set, with eventual a regularization term to penalize the network parameters, i.e.,

$$E(\boldsymbol{\omega}) = \frac{1}{2} \sum_{n=1}^N \|\mathbf{y}^{(n)} - \mathbf{f}(\mathbf{x}^{(n)}, \boldsymbol{\omega})\|^2 + \alpha \|\boldsymbol{\omega}\|^2,$$

where  $\alpha$  is the regularization parameter. In order to train a neural network, we split the optical and geometrical parameters in (i) retrieval parameters, i.e., the aerosol optical depth  $\tau$  and layer height  $H$ , and (ii) forward model parameters (parameters that are known to some accuracy, but are not included in the retrieval), as for example, the solar zenith angle  $\theta_0$ , viewing zenith angle  $\theta$ , relative azimuth angle  $\Delta\varphi$ , surface height  $H_s$ , and surface albedo  $A$ .

In this section we design neural networks for learning

1. the radiative transfer model (neural network for the forward operator), and
2. the inverse model (neural network for the inverse operator).

## 5.1 Algorithm descriptions

### 5.1.1 Neural network for the forward operator

For emulating the radiative transfer model, we consider a neural network in which, the input  $\mathbf{x}$  is the set of optical and geometrical parameters, while the output  $\mathbf{y}$  is the set of synthetic radiances  $I_{\text{sim}}^{(m)}(\lambda_k)$  computed on the forward wavelength grid  $\{\lambda_k\}_{k=1}^{N_\lambda}$ , i.e.,

$$\mathbf{x} = \begin{bmatrix} [\tau, H]^T \\ [\theta_0, \theta, \Delta\varphi, H_s, A_s]^T \end{bmatrix} \mapsto \mathbf{y} = [I_{\text{sim}}^{(m)}(\lambda_k)]_{k=1}^{N_\lambda}. \quad (5.7)$$

In this case, the dimensions of the input and output vectors are  $N_x = 7$  and  $N_y = N_\lambda$ , respectively.

After the radiative transfer model is learned, the synthetic radiances computed on the forward wavelength grid  $I_{\text{sim}}^{(m)}(\lambda_k)$  are convolved with a slit function  $g_r$  to obtain the synthetic spectral signals on the measurement wavelength grid  $\mathcal{I}_{\text{sim}}^{(m)}(\lambda_{mk}^r)$  (cf. Eq. (5.2)). The aerosol parameters encapsulated now in the state vector  $\mathbf{x} = [\tau, H]^T$  are obtained

as the solution of the nonlinear equation (cf. Eq. (4.1))  $\mathbf{y}^\delta = \mathbf{F}_m(\mathbf{x}) + \boldsymbol{\delta}_m$ , where, for a given swath row  $r$ ,  $\mathbf{y}^\delta = [\ln \mathcal{I}_{\text{mes}}^\delta(\lambda_{mk}^r)]_{k=1}^{N_{m\lambda}}$  is the measurement vector, and  $\mathbf{F}_m(\mathbf{x}) = [\ln \mathcal{I}_{\text{sim}}^{(m)}(\lambda_{mk}^r, \mathbf{x})]_{k=1}^{N_{m\lambda}}$  the forward model. This equation can be solved by using the inversion algorithm described in Section 4.1.

### 5.1.2 Neural network for the inverse operator

For solving the inverse problem, we designed two types of neural networks, which use as input

1. the synthetic radiances computed on the measurement wavelength grid, and
2. the principal-component transform of synthetic radiances.

**Neural network for the inverse operator with synthetic radiances** For emulating the inverse model, we may use a neural network in which, the input  $\mathbf{x}$  includes the spectral signals measured by the instrument and the forward model parameters, while the output  $\mathbf{y}$  includes the parameters to be retrieved, i.e.,

$$\mathbf{x} = \begin{bmatrix} [\mathcal{I}_{\text{sim}}^{(m)}(\lambda_{mk}^r) + \delta_{mk}]_{k=1}^{N_{m\lambda}} \\ [\theta_0, \theta, \Delta\varphi, H_s, A_s]^T \end{bmatrix} \mapsto \mathbf{y} = [\tau, H]^T. \quad (5.8)$$

The measurement noise vector  $\boldsymbol{\delta}_m$  is assumed to be a Gaussian random vector with zero mean and noise covariance matrix  $\mathbf{C}_{\delta m}$ . For simplicity, we approximate  $\mathbf{C}_{\delta m}$  by a diagonal matrix, i.e.,  $\mathbf{C}_{\delta m} = \text{diag}[\sigma_{mk}^2]_{k=1}^{N_{m\lambda}}$ , where  $\sigma_{mk}^2$  are the diagonal elements of  $\mathbf{C}_{\delta m}$ , and assume that estimates of  $\sigma_{mk}^2$  are available.

From Eq. (5.8), we see that the same output  $[\tau, H]^T$  corresponds to different realization of the random noise  $\delta_{mk}$ , as well as, to different wavelength grids  $\lambda_{mk}^r$ . This problem, which leads to a huge dimension of the data set, can be overcome by using the jitter approach, under the assumption that the measurement wavelength grid  $\{\lambda_{mk}^r\}_{k=1}^{N_{m\lambda}}$  is a discrete random variable which can take the values  $\{\lambda_{mk}^1\}_{k=1}^{N_{m\lambda}}, \dots, \{\lambda_{mk}^{\mathcal{R}}\}_{k=1}^{N_{m\lambda}}$ . According to this approach, at each forward pass through the network, a measurement wavelength grid  $\{\lambda_{mk}^r\}_{k=1}^{N_{m\lambda}}$  is randomly selected from the  $\mathcal{R}$  wavelength grids, and a new random noise  $\delta_{mk} \sim \mathcal{N}(0, \sigma_{mk}^2)$  is added to the synthetic spectral signal  $\mathcal{I}_{\text{sim}}^{(m)}(\lambda_{mk}^r)$ . In other words, the input sample is different every time it is passed through the network.

**Neural network for the inverse operator with the principal-component transform of synthetic radiances** In Section 3.3.3 the principal component analysis has been applied on the optical properties of the medium. Here, this approach is applied on the spectral data, and in particular, on the synthetic radiance vector  $\mathbf{i}_m = [\mathcal{I}_{\text{sim}}^{(m)}(\lambda_{mk}^r)]_{k=1}^{N_{m\lambda}} \in \mathbb{R}^{N_{m\lambda}}$  with the purpose of reducing its dimension (note that the dependency of  $\mathbf{i}_m$  on the swath row  $r$  is implicitly assumed). Considering the  $N_t$ -dimensional data set  $\{\mathbf{i}_m^{(n)}\}_{n=1}^{N_t}$ , the principal component analysis involves the following steps.

1. compute the sample mean of the data  $\langle \mathbf{i}_m \rangle = (1/N_t) \sum_{n=1}^{N_t} \mathbf{i}_m^{(n)}$ ;
2. stack all centered data  $\mathbf{i}_m^{(n)} - \langle \mathbf{i}_m \rangle$  into the columns of the matrix  $\mathcal{I}$ , i.e.,  $\mathcal{I} = [\mathbf{i}_m^{(1)} - \langle \mathbf{i}_m \rangle, \dots, \mathbf{i}_m^{(N_t)} - \langle \mathbf{i}_m \rangle] \in \mathbb{R}^{N_{m\lambda} \times N_t}$ ;

**Table 5.1:** Intervals of variation of the optical and geometrical parameters for generating the data set.

Parameter	Interval of variation
$\tau$	0.05 – 5
$H$	0.1 – 15.75 km
$\theta_0$	0 – 75°
$\theta$	0 – 70°
$\Delta\varphi$	0 – 180°
$H_s$	0 – 2.61 km
$A$	0 – 0.4

3. compute the covariance matrix  $\mathbf{C} = (1/N_t)\mathbf{I}\mathbf{I}^T \in \mathbb{R}^{N_{m\lambda} \times N_{m\lambda}}$  and a singular value decomposition of  $\mathbf{C}$ , i.e.,  $\mathbf{C} = \mathbf{U}\mathbf{S}\mathbf{U}^T$ , where  $\mathbf{S} = \text{diag}[s_k]_{k=1}^{N_{m\lambda}}$  is the diagonal matrix of the singular values  $s_1 > s_2 > \dots > s_{N_{m\lambda}} > 0$  and  $\mathbf{U} = [\mathbf{u}_1, \dots, \mathbf{u}_{N_{m\lambda}}] \in \mathbb{R}^{N_{m\lambda} \times N_{m\lambda}}$  is the orthogonal matrix of the singular vectors;
4. take the inverse transformation matrix as  $\mathbf{U}_{\mathcal{M}} = [\mathbf{u}_1, \dots, \mathbf{u}_{\mathcal{M}}] \in \mathbb{R}^{N_{m\lambda} \times \mathcal{M}}$ , in which case, the forward transformation matrix is  $\mathbf{U}_{\mathcal{M}}^\dagger = \mathbf{U}_{\mathcal{M}}^T \in \mathbb{R}^{\mathcal{M} \times N_{m\lambda}}$  and the dimensionality-reduced input is  $\hat{\mathbf{i}}_m^{(n)} = \mathbf{U}_{\mathcal{M}}^\dagger(\mathbf{i}_m^{(n)} - \langle \mathbf{i}_m \rangle) \in \mathbb{R}^{\mathcal{M}}$ , and
5. compute the number of principal components  $\mathcal{M}$  by monitoring the reconstruction error  $E_{\mathcal{M}} = \frac{1}{N_t} \sum_{n=1}^{N_t} \|(\mathbf{i}_m^{(n)} - \langle \mathbf{i}_m \rangle) - \mathbf{U}_{\mathcal{M}}\hat{\mathbf{i}}_m^{(n)}\|_2$  as function of  $\mathcal{M}$ , that is, determine  $\mathcal{M}$ , for which  $E_{\mathcal{M}}$  is below a prescribed tolerance.

Note that for the noisy radiance vector  $\mathbf{i}_m^\delta = \mathbf{i}_m + \boldsymbol{\delta}_m$ , where  $\boldsymbol{\delta}_m \sim \mathbf{N}(\mathbf{0}, \mathbf{C}_{\delta m})$  is the measurement noise vector, we find  $\langle \mathbf{i}_m^\delta \rangle = \langle \mathbf{i}_m \rangle$ , yielding  $\hat{\mathbf{i}}_m^\delta = \mathbf{U}_{\mathcal{M}}^T(\mathbf{i}_m^\delta - \langle \mathbf{i}_m \rangle) = \hat{\mathbf{i}}_m + \hat{\boldsymbol{\delta}}_m$  with  $\hat{\boldsymbol{\delta}}_m = \mathbf{U}_{\mathcal{M}}^T \boldsymbol{\delta}_m \sim \mathbf{N}(\mathbf{0}, \hat{\mathbf{C}}_{\delta m})$  and  $\hat{\mathbf{C}}_{\delta m} = \mathbf{U}_{\mathcal{M}}^T \mathbf{C}_{\delta m} \mathbf{U}_{\mathcal{M}} \in \mathbb{R}^{\mathcal{M} \times \mathcal{M}}$ .

Thus, instead of the synthetic radiances  $\mathbf{i}_m \in \mathbb{R}^{N_{m\lambda}}$ , the input of the neural network is the principal-component transform of synthetic radiances  $\hat{\mathbf{i}}_m = \mathbf{U}_{\mathcal{M}}^\dagger(\mathbf{i}_m - \langle \mathbf{i}_m \rangle) \in \mathbb{R}^{\mathcal{M}}$ , and during each forward pass through the network, the random noise  $\hat{\boldsymbol{\delta}}_m \sim \mathbf{N}(\mathbf{0}, \hat{\mathbf{C}}_{\delta m})$  is added to  $\hat{\mathbf{i}}_m$ .

## 5.2 Numerical analysis

The data set is obtained, by generating samples of optical and geometrical parameters by means of the smart sampling technique [Loyola et al., 2016]. The intervals of variation of these parameters are given in Table 5.1. In the numerical analysis, the moderately absorbing aerosol model taken from the MODIS aerosol retrieval algorithm is considered, and the aerosol layer is assumed to be homogeneous with a fixed thickness of 0.5 km, spreading evenly from  $H - 0.25$  km to  $H + 0.25$  km.

For the design of neural network architectures, the Python PyTorch machine learning library is used. The following features of the neural networks deserve to be mentioned.

1. The numbers of hidden layers and units in each layer are optimized by using 10% of the samples from the training set for validation. In the validation stage, the holdout cross-validation together with a grid search over a set of 3 values for the

number of hidden layers, i.e.,  $\{2, 3, 4\}$ , and a set of 5 values for the number of units, i.e.,  $\{20, 40, 60, 80, 100\}$ , are used. Through a numerical analysis, we found that a network architecture with 4 hidden layers and 40 nodes in each layer yields the lowest root-mean-square error on the validation data set.

2. A hyperbolic tangent activation function is assumed.
3. The mini-batch gradient descent in conjunction with ADaptive Moment Estimation (ADAM) is used as optimization tool.

Other peculiarities of the neural networks are listed below.

1. In the case of the neural network for the forward operator
  - a) the forward wavelength grid consists of  $N_\lambda = 485$  equidistant spectral points ranging from 758 to 771 nm,
  - b) the number of samples in the training set is  $N_t = 151\,423$ ,
  - c) in the inversion step, the noise covariance matrix is chosen as  $\mathbf{C}_{\delta m} = \text{diag}[\sigma_{mk}^2]_{k=1}^{N_{m\lambda}}$  with  $\sigma_{mk} = 0.1 \times \bar{\mathcal{I}}_{\text{sim}k}^{(m)}$  and

$$\bar{\mathcal{I}}_{\text{sim}k}^{(m)} = \frac{1}{\mathcal{R}} \sum_{r=1}^{\mathcal{R}} \mathcal{I}_{\text{sim}}^{(m)}(\lambda_{mk}^r, \mathbf{x}_a)$$

for all  $k = 1, \dots, N_{m\lambda}$ , and the a priori covariance matrix as  $\mathbf{C}_x = \text{diag}[\sigma_{xk}^2]_{k=1}^2$  with  $\sigma_{xk} = 0.2 \times x_{ak}$  and  $x_{ak}$  standing for the a priori values  $\tau_a = 1$  and  $H_a = 2$  km.

2. In the case of the neural network for the inverse operator
  - a) the number of swath rows is  $\mathcal{R} = 448$ , the number of points in each measurement wavelength grid is  $N_{m\lambda} = 131$ , and the measurement wavelength grids are chosen from the TROPOMI Level-1 products, e.g.,

$$\begin{aligned} \{\lambda_{mk}^1\}_{k=1}^{N_{m\lambda}} &= \{755.120, \dots, 770.929 \text{ nm}\}, \\ \{\lambda_{mk}^2\}_{k=1}^{N_{m\lambda}} &= \{755.133, \dots, 770.942 \text{ nm}\} \\ &\vdots \\ \{\lambda_{mk}^{\mathcal{R}}\}_{k=1}^{N_{m\lambda}} &= \{755.264, \dots, 771.071 \text{ nm}\}, \end{aligned}$$

- b) the noisy spectra are generated as in the case of the neural network for the forward operator,
- c) the number of samples in the training set is  $N_t = 404\,901$ , where each sample consists of a set of optical and geometrical parameters and the corresponding synthetic radiances computed on all measurement wavelength grids.
- d) for the inverse operator with the principal-component transform of synthetic radiances, we
  - i. approximate the dimensionality-reduced noise covariance matrix  $\hat{\mathbf{C}}_{\delta m} = \mathbf{U}_{\mathcal{M}}^T \mathbf{C}_{\delta m} \mathbf{U}_{\mathcal{M}}$  by a diagonal matrix, i.e.,  $\hat{\mathbf{C}}_{\delta m} \approx \text{diag}[\hat{C}_{m,kk}]_{k=1}^M$ , where  $\hat{C}_{m,ij}$  are the entries of  $\hat{\mathbf{C}}_{\delta m}$ , and
  - ii. use  $\mathcal{M} = 14$  principal components, for which the reconstruction error  $E_{\mathcal{M}}$  is below  $1.6 \times 10^{-3}$ .

### 5.2.1 Synthetic retrieval

The performances of the retrieval algorithms are tested on a prediction set consisting of  $N_p = 10\,000$  samples

$$(\tau^{(n)}, H^{(n)}, \theta_0^{(n)}, \theta^{(n)}, \Delta\varphi^{(n)}, H_s^{(n)}, A^{(n)}),$$

chosen randomly in their intervals of variation (see Table 5.1). For a statistical interpretation of the results, we split the interval of variation of  $x$ ,  $[x_{\min}, x_{\max}]$ , where  $x$  stands for  $\tau$  and  $H$ , into  $N_b = 40$  equidistant bins, i.e.,  $[x_{\min}, x_{\max}] = \cup_{j=1}^{N_b} B_{xj}$ , and compute the (bin) mean

$$\mathbb{E}_j(x_{\text{pred}}) = \frac{1}{N_{xj}} \sum_{n, \text{ s.t. } x^{(n)} \in B_{xj}} x_{\text{pred}}^{(n)}$$

and standard deviation

$$\sqrt{\mathbb{E}_j([x_{\text{pred}} - \mathbb{E}_j(x_{\text{pred}})]^2)} = \sqrt{\frac{1}{N_{xj}} \sum_{n, \text{ s.t. } x^{(n)} \in B_{xj}} [x_{\text{pred}}^{(n)} - \mathbb{E}_j(x_{\text{pred}})]^2},$$

over all  $N_{xj}$  samples  $x_{\text{pred}}^{(n)}$ , whose corresponding  $x^{(n)}$  are in  $B_{xj}$ . To quantify the retrieval accuracy we use the first two moments of the absolute error over the prediction set  $\Delta_x = x_{\text{pred}} - x$ , where  $x_{\text{pred}}$  and  $x$  are the predicted and true values, respectively. These are the mean absolute error

$$\mathbb{E}(|\Delta_x|) = \frac{1}{N_p} \sum_{n=1}^{N_p} |\Delta_x^{(n)}|$$

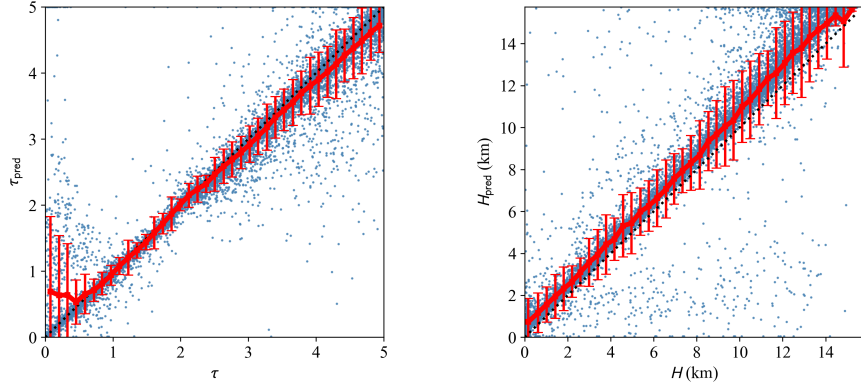
and the standard deviation of the absolute error

$$\sqrt{\mathbb{E}([\Delta_x - \mathbb{E}(\Delta_x)]^2)} = \sqrt{\frac{1}{N_p} \sum_{n=1}^{N_p} [\Delta_x^{(n)} - \mathbb{E}(\Delta_x)]^2}$$

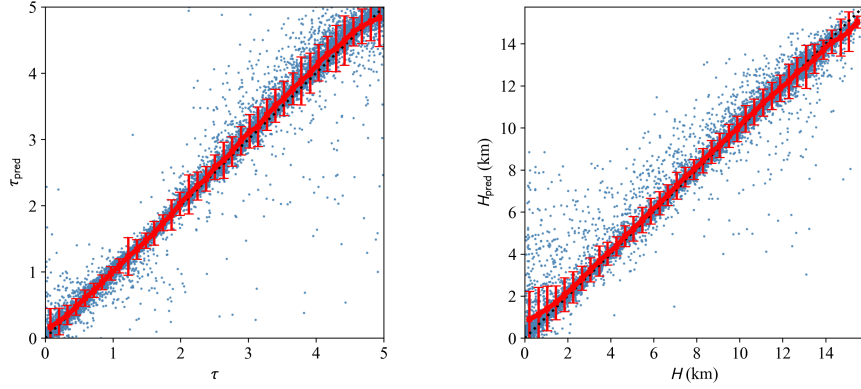
computed over the prediction set. In Figs. 5.1, 5.2, and 5.3, we illustrate the mean  $\mathbb{E}_j(x_{\text{pred}})$  and standard deviation  $\sqrt{\mathbb{E}_j([x_{\text{pred}} - \mathbb{E}_j(x_{\text{pred}})]^2)}$  versus the midpoint  $\bar{x}_j$  of the  $j$ th bin, while in Table 5.2 we show the mean absolute error  $\mathbb{E}(|\Delta_x|)$  and the standard deviation of the absolute error  $\sqrt{\mathbb{E}([\Delta_x - \mathbb{E}(\Delta_x)]^2)}$  over the prediction set. The following conclusions can be drawn.

1. The accuracy is low for small values of the aerosol optical depth  $\tau$  and layer height  $H$ .
2. The inverse-operator neural networks with synthetic radiances and their principal components have comparable accuracies; these are higher than that of the forward-operator neural network.

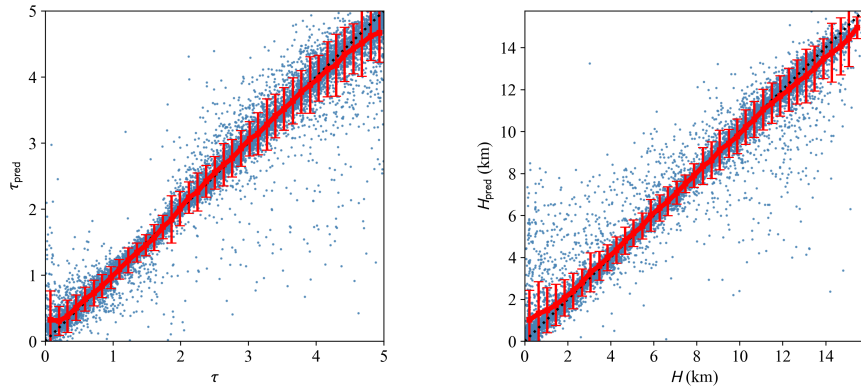
The low accuracy for small values of  $\tau$  and  $H$  is due to the fact that in this domain, the residual of the measured and synthetic radiances has several local minima, while the low accuracy of the forward-operator neural network is due to the inherent errors of the inversion algorithm.



**Figure 5.1:** Predictions of the forward-operator neural network. The plots show the predicted values  $x_{\text{pred}}^{(n)}$  (blue points) together with the mean  $\mathbb{E}_j(x_{\text{pred}})$  (red points) and standard deviation  $\sqrt{\mathbb{E}_j(x_{\text{pred}} - \mathbb{E}_j(x_{\text{pred}}))^2}$  (red error bars) over all samples  $x_{\text{pred}}^{(n)}$ , whose corresponding  $x^{(n)}$  are in the  $j$ th bin. The interval of variation of  $x$  is split into  $N_b = 40$  bins, where  $x$  stands for the aerosol optical depth  $\tau$  and aerosol layer height  $H$ .



**Figure 5.2:** The same as in Fig. 5.1 but for the inverse-operator neural network with synthetic radiances.



**Figure 5.3:** The same as in Fig. 5.1 but for the inverse-operator neural network with the principal-component transform of synthetic radiances.

**Table 5.2:** Mean absolute error  $\mathbb{E}(|\Delta_x|)$ , and the standard deviation of the absolute error  $\sqrt{\mathbb{E}([\Delta_x - \mathbb{E}(\Delta_x)]^2)}$  over the prediction set. The results correspond to the forward-operator neural network (Method 1), and the inverse-operator neural networks with synthetic radiances (Method 2) and the principal-component transform of synthetic radiances (Method 3).

Method	$x$	$\mathbb{E}( \Delta_x )$	$\sqrt{\mathbb{E}([\Delta_x - \mathbb{E}(\Delta_x)]^2)}$
1	$\tau$	0.169	0.410
	$H$	0.879	1.749
2	$\tau$	0.115	0.243
	$H$	0.336	0.740
3	$\tau$	0.136	0.316
	$H$	0.437	0.951

The variations of the absolute error  $\Delta_x = x_{\text{pred}} - x$  with respect to the optical and geometrical parameters are illustrated in Figs. 5.4, 5.5, and 5.6. Here, the interval of variation of a parameter  $b$ ,  $[b_{\min}, b_{\max}]$  is split into  $N_b = 40$  equidistant bins, i.e.,  $[b_{\min}, b_{\max}] = \cup_{j=1}^{N_b} B_{bj}$ , and the mean  $\mathbb{E}_j(\Delta_x)$  and standard deviation  $\sqrt{\mathbb{E}_j([\Delta_x - \mathbb{E}_j(\Delta_x)]^2)}$  over all samples  $\Delta_x^{(n)}$ , whose corresponding  $b^{(n)}$  are in the  $j$ th bin  $B_{bj}$ , are plotted versus the midpoint  $\bar{b}_j$  of the bin. From Figs. 5.4, 5.5, and 5.6 we infer that

1. the standard deviation of the absolute error in the aerosol optical depth  $\tau$  is large when the solar zenith angle  $\theta_0$ , viewing angle  $\theta$ , and surface albedo  $A$  are large,
2. the standard deviation of the absolute error in the aerosol layer height  $H$  is large when the aerosol optical thickness  $\tau$  is small and the surface albedo  $A$  is large, and
3. the inverse-operator neural network with synthetic radiances yields the smallest standard deviations, and the forward-operator neural network the largest ones.

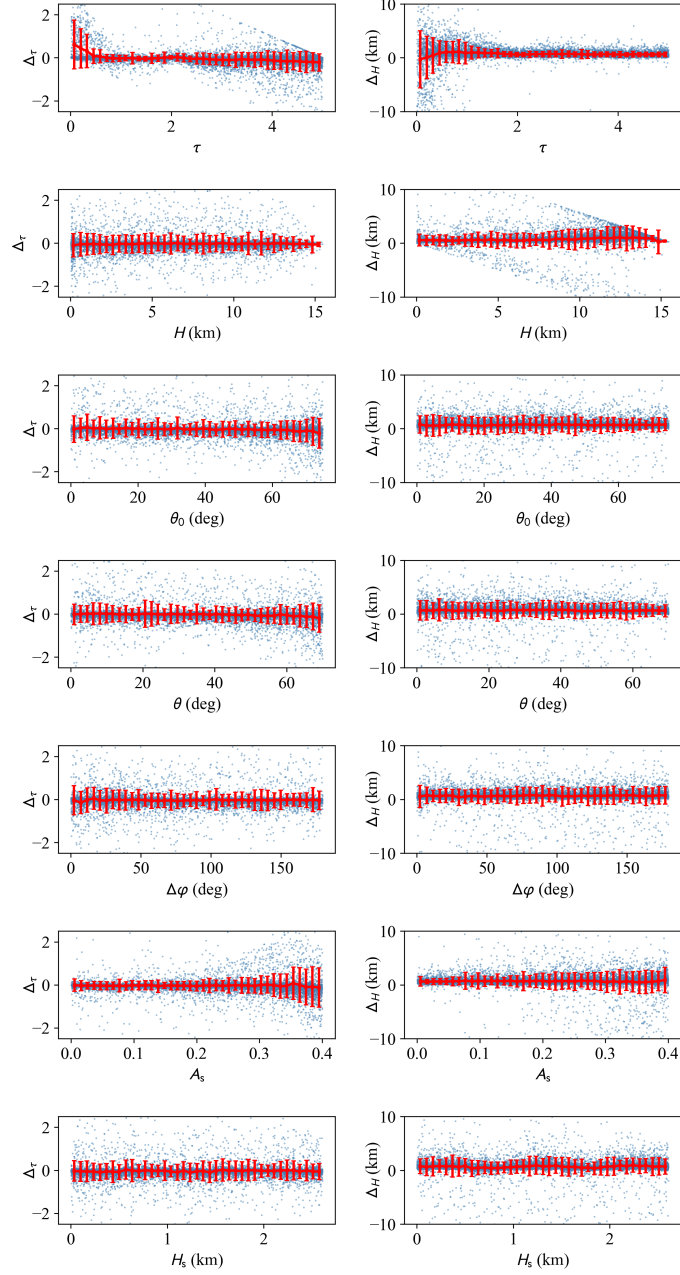
## 5.2.2 Retrieval from real data

To test the performances of the retrieval algorithms on real TROPOMI/S5P data, we choose a wild fire scene in California and consider the measurements recorded on 12 December 2017.

The retrieval results for the aerosol optical depth and layer height are illustrated in Fig. 5.7. In Fig. 5.8, we illustrate the absolute errors in the retrieved aerosol optical depth  $\delta_\tau = \tau_{\text{pred}} - \tau_{\text{pred}}^{\text{ref}}$  and aerosol layer height  $\delta_H = H_{\text{pred}} - H_{\text{pred}}^{\text{ref}}$  corresponding to the forward-operator neural network and the inverse-operator neural network with the principal-component transform of synthetic radiances, where the results provided by the inverse-operator neural network with synthetic radiances are taken as a reference. The plots show that

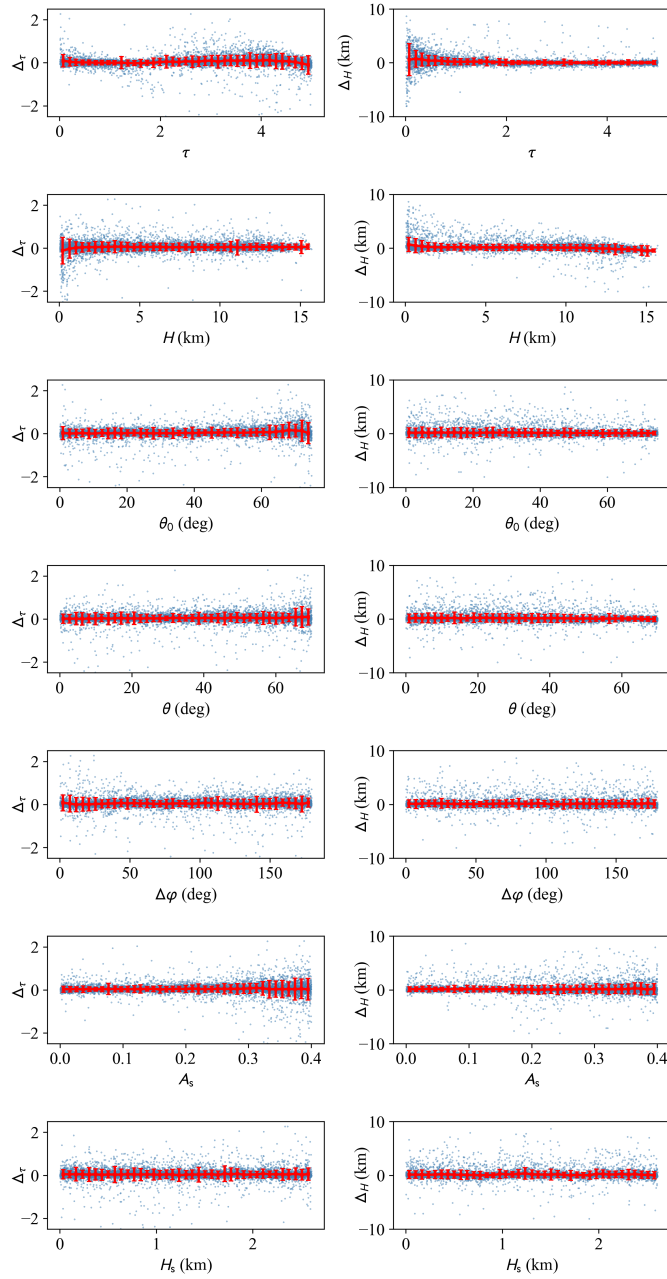
1. the absolute errors in the retrieved aerosol optical depth  $\delta_\tau$  are smaller than 0.1 over the entire scene,



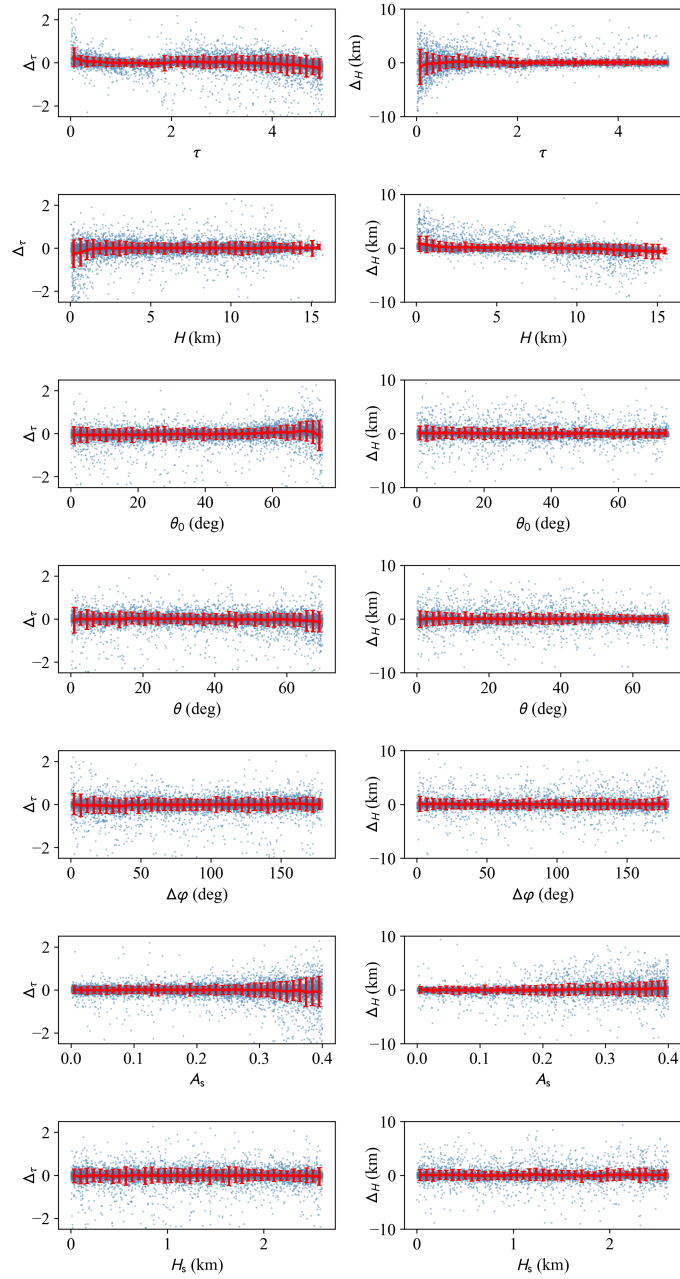


**Figure 5.4:** Absolute error in the retrieved aerosol optical depth  $\Delta_\tau$  and aerosol layer height  $\Delta_H$  versus the optical and geometrical parameters  $b$ , where  $b$  stands for the aerosol optical depth  $\tau$ , aerosol layer height  $H$ , solar zenith angle  $\theta_0$ , viewing zenith angle  $\theta$ , relative azimuth angle  $\Delta\varphi$ , surface height  $H_s$ , and surface albedo  $A$ . The plots show the absolute error  $\Delta_x^{(n)}$  (blue points) together with the mean  $\mathbb{E}_j(\Delta_x)$  (red points) and standard deviation  $\sqrt{\mathbb{E}_j([\Delta_x - \mathbb{E}_j(\Delta_x)]^2)}$  (red error bars) over all samples  $\Delta_x^{(n)}$ , whose corresponding  $b^{(n)}$  are in the  $j$ th bin. The results correspond to the forward-operator neural network.

5 Neural network retrieval algorithms



**Figure 5.5:** The same as in Fig. 5.4 but for the inverse-operator neural network with synthetic radiances.



**Figure 5.6:** The same as in Fig. 5.4 but for the inverse-operator neural network with the principal-component transform of synthetic radiances.

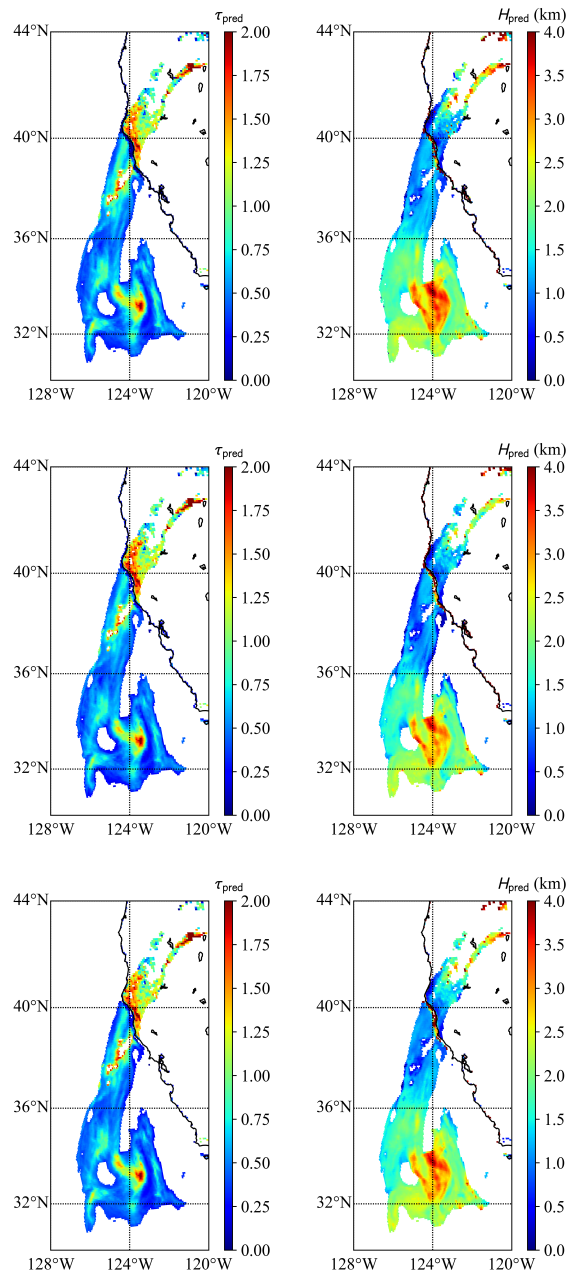
## 5 Neural network retrieval algorithms

2. the absolute errors in the retrieved aerosol layer height  $\delta_H$  are smaller than 0.4 km over the entire scene.

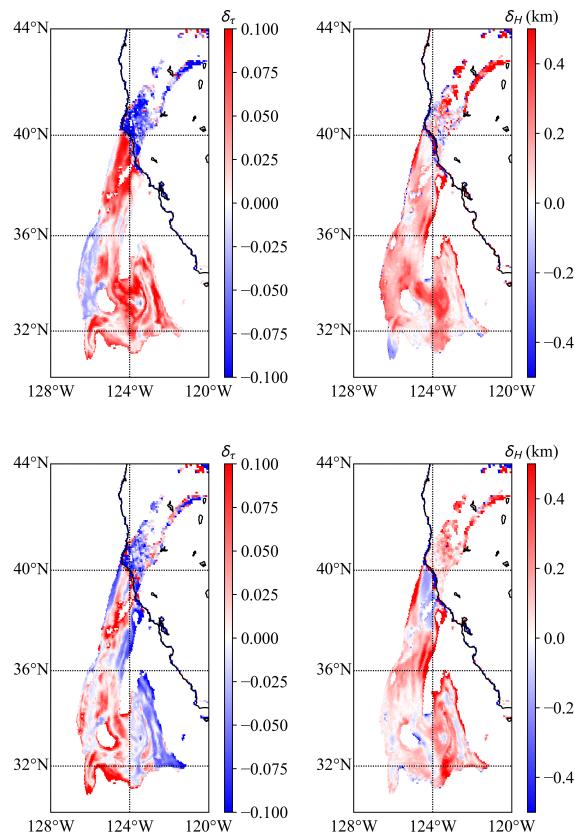
The computational time of a Bayesian-based retrieval algorithm is 2 – 3 minutes for one pixel on a computer Intel Core i7-4770M CPU 3.40GHz with 16 GB RAM, while the computational times of a forward- and an inverse-operator neural networks are 0.4 and 0.003 seconds, respectively.

### 5.2.3 Conclusions

In this section, we analyzed the numerical performances of three neural network algorithms for aerosol retrieval from TROPOMI/S5P measurements in the oxygen A-band. They use neural networks (i) to emulate the radiative transfer model and a Bayesian approach to solve the inverse problem, (ii) to learn the inverse model from the synthetic radiances, and (iii) to learn the inverse model from the principal-component transform of synthetic radiances. The neural networks were trained for moderately absorbing aerosols taken from the MODIS aerosol retrieval algorithm. Our numerical analysis, performed on synthetic and real data, has shown that inverse-operator neural networks, which use the jitter approach in order to deal with the row-dependent measurement wavelength grid and the random measurement noise, are more accurate and efficient than a forward-operator neural network. However, all neural network retrieval algorithms have incomparably higher efficiency than a Bayesian-based retrieval algorithm.



**Figure 5.7:** Retrieved aerosol optical depth  $\tau_{\text{pred}}$  and aerosol layer height  $H_{\text{pred}}$  from TROPOMI/S5P measurements recorded on 12 December **2017** in California. The results correspond to the forward-operator neural network (upper panels), the inverse-operator neural network with synthetic radiances (middle panels), and the inverse-operator neural network with the principal-component transform of synthetic radiances (lower panels).



**Figure 5.8:** Absolute errors in the retrieved aerosol optical depth  $\delta_\tau$  and aerosol layer height  $\delta_H$  corresponding to the forward-operator neural network (upper panels) and the inverse-operator neural network with the principal-component transform of synthetic radiances (lower panels). The results provided by the inverse-operator neural network with synthetic radiances are taken as a reference.

# 6 Conclusions

## 6.1 Summary

This thesis aims to design algorithms for retrieving aerosol parameters (layer height and optical depth) from TROPOMI/S5P measurements in the oxygen A-band. In summary, the major achievements of this thesis are outlined:

1. A database that provides the single scattering albedo and the phase function coefficients of different aerosol models has been created. The aerosol models included in the database are either user defined or taken from the GOCART model, OPAC database, OMI multiwavelength aerosol retrieval algorithm, and MODIS aerosol retrieval algorithm.
2. A Bayesian-based retrieval algorithm has been designed under the assumption that the aerosol model is known. The algorithm uses a linearized radiative transfer model relying on the discrete ordinate method with matrix exponential. To speed up the computation, the telescoping technique, the method of false discrete ordinate, the correlated k-distribution method, and principal component analysis are used. For aerosol retrieval, the partial derivatives of the spectral signal with respect to the aerosol optical depth and layer height are computed by a forward-adjoint approach in conjunction with an homogenization procedure for the aerosol layer, while the forward model used in the inversion algorithm is the radiance model without spectral corrections. The inverse problem is formulated as a least-squares problem and solved by means of the iteratively regularized Gauss-Newton method. This approach is less sensitive to overestimates of the regularization parameter and provides an optimal value of the regularization parameter (the ratio of the data error variance and the a priori state variance ) and the corresponding regularized solution.
3. A Bayesian-based retrieval algorithm, that take into account the uncertainty in model selection, has been designed. In this case, two solutions estimates can be computed
  - a) the maximum solution estimate, corresponding to the model with the highest evidence, and
  - b) the mean solution estimate, representing a linear combination of solutions weighted by their evidences.
4. Three neural network retrieval algorithms have been developed. The first one uses a neural network to emulate the radiative transfer model and a Bayesian approach to solve the inverse problem. The second and third ones employ a neural network to learn the inverse model and use as input either the synthetic radiances computed on the measurement wavelength grid or the principal-component transform of synthetic radiances.

A comparison of the operational and the designed retrieval algorithm is given in Table 6.1.

**Table 6.1:** Special features of the operational and designed retrieval algorithm.

Feature	Operational Retrieval Algorithm	Designed Retrieval Algorithm
aerosol model	average aerosol model with fixed single scattering albedo and asymmetry parameter	aerosol database with (i) user defined aerosol models, and (ii) aerosol models included, in the GOCART, OPAC, OMI, and MODIS databases.
radiative transfer model	layer based orders-of-scattering method	discrete ordinate method equipped with acceleration methods (telescoping technique, method of false discrete ordinate method, k-distribution method, and principal component analysis)
linearization approach	semi-analytic approach (equivalent to the adjoint method)	forward and forward-adjoint approaches relying on an analytic computation of the derivatives and the adjoint theory, respectively
retrieval algorithm	neural network to emulate the radiative transfer model, and a Bayesian approach (with a priori chosen covariance matrices) to solve the inverse problem	<ol style="list-style-type: none"> <li>1. Bayesian-based retrieval algorithms for (i) a specified aerosol model and (ii) a set of candidate models, using a Bayesian approach and the iteratively regularized Gauss-Newton method to solve the inverse problem;</li> <li>2. Neural network algorithms (i) to emulate the radiative transfer model and (ii) to learn the inverse model (using as input either the synthetic radiances or their principal components).</li> </ol>



## 6.2 Outlook

The main conclusion of this thesis is that inverse-operator neural networks are the best choice for retrieving the aerosol parameters from TROPOMI/S5P measurements in oxygen A-band. However, the design of neural networks for atmospheric retrieval is a very complex research field that requires more developments. In fact, the inverse-operator neural networks presented in this thesis need to be extended by including additional features. Along this line, the inverse-operator neural networks should be

1. trained on the remaining aerosol models considered in the MODIS aerosol retrieval algorithm, i.e., non-absorbing, absorbing, and desert dust,
2. trained to learn the relative evidences of different aerosol models, so that, a mean solution estimate, representing a linear combination of candidate solutions weighted by their evidences, can be computed, and
3. redesigned in a Bayesian deep learning framework in order to predict input aleatoric and model uncertainties [Doicu et al., 2021].



# List of Figures

4.1	Simulated reflectances versus the wavelength in the TROPOMI oxygen A-band (left), and the phase functions versus the scattering angle (right). The results correspond to the aerosol models from Sets I, II, and III, the aerosol optical depth $\tau = 0.5$ , aerosol layer height $H = 3.5$ km, surface albedo $A = 0.05$ , solar zenith angle $\theta_0 = 30^\circ$ , viewing zenith angle $\theta = 0^\circ$ , and relative azimuth angle $\Delta\varphi = 180^\circ$ . For sets II and III three values of 0.0, 0.5, and 1.0 for the weights of the coarse mode and organic carbon, respectively, were chosen. . . . .	42
4.2	Absolute retrieval errors in the aerosol layer height (ALH) and aerosol optical depth (AOD) for the weakly absorbing model in Set II. The synthetic measurement spectra are generated with the aerosol model MODABS in Set I. The parameters of the simulation are as in Fig. 4.1. . . . .	43
4.3	The same as in Fig. 4.2 but for the sulfurous smoke model in Set III. . . . .	44
4.4	Derivative of the reflectance spectrum with respect to the aerosol layer height (ALH) and aerosol optical depth (AOD) as functions of surface albedo. . . . .	45
4.5	Relative retrieval errors in the aerosol layer height (ALH) and aerosol optical depth (AOD) due to errors in the surface albedo. The results correspond to the aerosol model MODABS in Set I, the aerosol optical depth $\tau = 0.5$ , and the aerosol layer height $H = 3.5$ km. . . . .	46
4.6	Relative retrieval errors in the aerosol layer height (ALH) and aerosol optical depth (AOD) due to errors in the solar zenith angle (SZA). . . . .	47
4.7	The same as in Fig. 4.6 but for the errors in the viewing zenith angle (VZA). . . . .	48
4.8	Relative retrieval errors in the aerosol layer height (ALH) and aerosol optical depth (AOD), as well as, the corresponding fit residuals, versus the wavelength shift. In the retrieval, the true values of the aerosol layer height are 1.5, 3.5, 5.5, and 7.5 km, while the true value of the aerosol optical depth is 0.5. The plots in the left panel correspond to a surface albedo of 0.05, while the plots in the right panel to a surface albedo of 0.15. . . . .	49
4.9	The same as in Fig. 4.8, but for the true aerosol optical depth values of 0.5, 1.0, 1.5, and 2.0, and a true aerosol layer height value of 3.5 km. . . . .	50
4.10	VIIRS true-color images on 22 June 2018 (left) and 6 June 2020 (right). The red rectangular region displays the chosen TROPOMI scene. . . . .	51
4.11	Retrieved aerosol layer height (ALH) and aerosol optical depth (AOD) for the TROPOMI scene on 22 June 2018. The results are obtained with the present retrieval algorithm (top panel) and the operational retrieval algorithm (bottom panel). . . . .	52
4.12	The same as in Fig. 4.11 but for the TROPOMI scene on 6 June 2020. . . . .	53
4.14	Relative errors $\varepsilon_{\text{mean,max}}^{\tau,H}(\tau_e = 0.5, H_e)$ versus the aerosol layer height (ALH) for the aerosol models from Set 1. . . . .	53

List of Figures

4.13	Observed and modeled reflectance spectra at the solution for four (randomly chosen) TROPOMI ground pixels on 22 June 2018 (top panel) and 6 June 2020 (bottom panel). . . . .	54
4.15	Relative errors $\varepsilon_{\text{mean,max}}^{\tau,H}(\tau_e, H_e = 3.5 \text{ km})$ versus the aerosol optical depth (AOD) for the aerosol models from Set 1. . . . .	54
4.16	The a posteriori densities $p(\mathbf{x} = [\tau, H]   \bar{\mathbf{y}}^\delta, m)$ for $m = \text{NONABS, MOD-ABS, ABS, and DUST}$ , and the mean a posteriori densities $p_{\text{mean}}(\mathbf{x} = [\tau, H]   \bar{\mathbf{y}}^\delta)$ in the case $m_e = \text{MODABS}$ , $\tau_e = 1.25$ , and $H_e = 3.5 \text{ km}$ . The black curve indicates the mean a posterior density. In each plot, the red vertical dashed line corresponds to the exact values to be retrieved $(\tau_e, H_e)$ , the blue vertical dashed line to the maximum solution estimate $(\tau_{\text{max}}, H_{\text{max}})$ , and the green dashed line to the mean solution estimates $(\tau_{\text{mean}}, H_{\text{mean}})$ . . . . .	55
4.17	Relative errors $\varepsilon_{\text{mean,max}}^{\tau,H}(\tau_e = 0.5, H_e)$ for the aerosol models from Set 2. . . . .	56
4.18	Relative errors $\varepsilon_{\text{mean,max}}^{\tau,H}(\tau_e, H_e = 3.5 \text{ km})$ for the aerosol models from Set 2. . . . .	56
4.19	The mean a posteriori densities $p_{\text{mean}}(H   \bar{\mathbf{y}}^\delta)$ for $m_e = \text{NONABS, MOD-ABS, ABS, DUST}$ , $\tau_e = 0.5$ , and $H_e = 3.5 \text{ km}$ . In each plot, the red vertical dashed line correspond to the exact values to be retrieved $(\tau_e, H_e)$ , the blue vertical dashed line to the maximum solution estimate $(\tau_{\text{max}}, H_{\text{max}})$ , and the green dashed line to the mean solution estimates $(\tau_{\text{mean}}, H_{\text{mean}})$ . . . . .	57
4.20	The same as in Figure 4.19 but for the mean a posteriori densities $p_{\text{mean}}(\tau   \bar{\mathbf{y}}^\delta)$ . . . . .	57
4.21	True-color VIIRS images recorded on 4 (left) and 5 (right) July 2019. . . . .	59
4.22	The aerosol model with the highest evidence from Set 1 (a), and the aerosol type containing the aerosol model with the highest evidence from Set 2 (b). The TROPOMI spectra were recorded on 4 July 2019. . . . .	59
4.23	The model evidence for each aerosol model from Set 1. The TROPOMI spectra were recorded on 4 July 2019. . . . .	59
4.24	The sum of the first 10 best aerosol model evidences for each aerosol type from Set 2. The TROPOMI spectra were recorded on 4 July 2019. . . . .	60
4.25	The same as in Fig. 4.22 but for the data on 5 July 2019. . . . .	60
4.26	The same as in Fig. 4.23 but for the data on 5 July 2019. . . . .	60
4.27	The same as in Fig. 4.24 but for the data on 5 July 2019. . . . .	61
4.28	The maximum solution estimates $(H_{\text{max}}, \tau_{\text{max}})$ and the mean solution estimates $(H_{\text{mean}}, \tau_{\text{mean}})$ for Set 1 and data on 4 July 2019. . . . .	61
4.29	The maximum solution estimates $(H_{\text{max}}, \tau_{\text{max}})$ and the mean solution estimates $(H_{\text{mean}}, \tau_{\text{mean}})$ for Set 2 and data on 4 July 2019. . . . .	61
4.30	The same as in Fig. 4.28 but for the data on 5 July 2019. . . . .	62
4.31	The same as in Fig. 4.29 but for the data on 5 July 2019. . . . .	62
5.1	Predictions of the forward-operator neural network. The plots show the predicted values $x_{\text{pred}}^{(n)}$ (blue points) together with the mean $\mathbb{E}_j(x_{\text{pred}})$ (red points) and standard deviation $\sqrt{\mathbb{E}_j(x_{\text{pred}} - \mathbb{E}_j(x_{\text{pred}}))^2}$ (red error bars) over all samples $x_{\text{pred}}^{(n)}$ , whose corresponding $x^{(n)}$ are in the $j$ th bin. The interval of variation of $x$ is split into $N_b = 40$ bins, where $x$ stands for the aerosol optical depth $\tau$ and aerosol layer height $H$ . . . . .	71
5.2	The same as in Fig. 5.1 but for the inverse-operator neural network with synthetic radiances. . . . .	71

5.3	The same as in Fig. 5.1 but for the inverse-operator neural network with the principal-component transform of synthetic radiances. . . . .	71
5.4	Absolute error in the retrieved aerosol optical depth $\Delta_\tau$ and aerosol layer height $\Delta_H$ versus the optical and geometrical parameters $b$ , where $b$ stands for the aerosol optical depth $\tau$ , aerosol layer height $H$ , solar zenith angle $\theta_0$ , viewing zenith angle $\theta$ , relative azimuth angle $\Delta\varphi$ , surface height $H_s$ , and surface albedo $A$ . The plots show the absolute error $\Delta_x^{(n)}$ (blue points) together with the mean $\mathbb{E}_j(\Delta_x)$ (red points) and standard deviation $\sqrt{\mathbb{E}_j([\Delta_x - \mathbb{E}_j(\Delta_x)]^2)}$ (red error bars) over all samples $\Delta_x^{(n)}$ , whose corresponding $b^{(n)}$ are in the $j$ th bin. The results correspond to the forward-operator neural network. . . . .	73
5.5	The same as in Fig. 5.4 but for the inverse-operator neural network with synthetic radiances. . . . .	74
5.6	The same as in Fig. 5.4 but for the inverse-operator neural network with the principal-component transform of synthetic radiances. . . . .	75
5.7	Retrieved aerosol optical depth $\tau_{\text{pred}}$ and aerosol layer height $H_{\text{pred}}$ from TROPOMI/S5P measurements recorded on 12 December <b>2017</b> in California. The results correspond to the forward-operator neural network (upper panels), the inverse-operator neural network with synthetic radiances (middle panels), and the inverse-operator neural network with the principal-component transform of synthetic radiances (lower panels). . . . .	77
5.8	Absolute errors in the retrieved aerosol optical depth $\delta_\tau$ and aerosol layer height $\delta_H$ corresponding to the forward-operator neural network (upper panels) and the inverse-operator neural network with the principal-component transform of synthetic radiances (lower panels). The results provided by the inverse-operator neural network with synthetic radiances are taken as a reference. . . . .	78



# List of Tables

1.1	Satellite instruments that perform measurements in O <sub>2</sub> absorption bands.	4
2.1	Aerosol mixtures of sulfate (SS), dust (DU), sea salt (SS), black carbon (BC), and organic carbon (OC) obtained by a cluster analysis using the Goddard Chemistry Aerosol Radiation and Transport (GOCART) model.	13
2.2	Table 3: Aerosol models from OPAC database. The relative refractive index corresponds to $\lambda = 750$ nm and $U = 0.8$ .	14
2.3	Aerosol models assumed in the OMI multiwavelength aerosol retrieval algorithm. For each model, the first and second lines are related to the accumulated and coarse modes, respectively.	15
2.4	Aerosol models considered in the MODIS aerosol retrieval algorithm. For each model, the first and second lines are related to the accumulated and coarse modes, respectively. The four values of the refractive index for dust, correspond to the wavelengths $\lambda = 0.470, 0.550, 0.660, 2.100$ $\mu\text{m}$ .	16
4.1	Spectral characteristics of TROPOMI oxygen A-band measurements and the main input retrieval parameters. The aerosol models are the weakly absorbing (WA1202) and dust (DD3102) models in Set II.	44
4.2	Bias in the aerosol layer height (ALH) and aerosol optical depth (AOD) due to the errors in the surface albedo and wavelength calibration.	46
5.1	Intervals of variation of the optical and geometrical parameters for generating the data set.	68
5.2	Mean absolute error $\mathbb{E}( \Delta_x )$ , and the standard deviation of the absolute error $\sqrt{\mathbb{E}([\Delta_x - \mathbb{E}(\Delta_x)]^2)}$ over the prediction set. The results correspond to the forward-operator neural network (Method 1), and the inverse-operator neural networks with synthetic radiances (Method 2) and the principal-component transform of synthetic radiances (Method 3).	72
6.1	Special features of the operational and designed retrieval algorithm.	80





# Bibliography

- Ambartzumian, V. The effect of the absorption lines on the radiative equilibrium of the outer layers of the stars. *Publ. Obs. Astron. Univ., Leningrad*, 6:7–18, 1936.
- Babu, S. S., Moorthy, K. K., Manchanda, R. K., Sinha, P. R., Satheesh, S., Vajja, D. P., Srinivasan, S., and Kumar, V. A. Free tropospheric black carbon aerosol measurements using high altitude balloon: Do BC layers build “their own homes” up in the atmosphere. *Geophysical research letters*, 38(8), doi: 10.1029/2011GL046654, 2011.
- Bodhaine, B. A., Wood, N. B., Dutton, E. G., and Slusser, J. R. On rayleigh optical depth calculations. *Journal of Atmospheric and Oceanic Technology*, 16(11):1854–1861, 1999.
- Boesche, E., Stammes, P., and Bennartz, R. Aerosol influence on polarization and intensity in near-infrared O<sub>2</sub> and CO<sub>2</sub> absorption bands observed from space. *Journal of Quantitative Spectroscopy and Radiative Transfer*, 110(3):223–239, doi: 10.1016/j.jqsrt.2008.09.019, 2009.
- Box, M. A. Radiative perturbation theory: A review. *Environmental Modelling & Software*, 17(1): 95–106, doi: 10.1016/S1364-8152(01)00056-1, 2002.
- Brath, M., Fox, S., Eriksson, P., Harlow, R. C., Burgdorf, M., and Buehler, S. A. Retrieval of an ice water path over the ocean from ISMAR and MARSS millimeter and submillimeter brightness temperatures. *Atmospheric Measurement Techniques*, 11(1):611–632, doi: 10.5194/amt-11-611-2018, 2018.
- 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<sub>2</sub> – O<sub>2</sub> spectral band using a neural network approach. *Atmospheric Measurement Techniques*, 10(3):783–809, doi: 10.5194/amt-10-783-2017, 2017.
- Chin, M., Ginoux, P., Kinne, S., Torres, O., Holben, B. N., Duncan, B. N., Martin, R. V., Logan, J. A., Higurashi, A., and Nakajima, T. Tropospheric aerosol optical thickness from the GOCART model and comparisons with satellite and Sun photometer measurements. *J. Atmos. Sci.*, 59(3):461–483, doi: 10.1175/1520-0469(2002)059<0461:TAOTFT>2.0.CO;2, 2002.
- Chowdhary, J., Cairns, B., Mishchenko, M. I., Hobbs, P. V., Cota, G. F., Redemann, J., Rutledge, K., Holben, B. N., and Russell, E. Retrieval of aerosol scattering and absorption properties from photopolarimetric observations over the ocean during the CLAMS experiment. *Journal of the Atmospheric Sciences*, 62(4):1093–1117, doi: 10.1175/JAS3389.1, 2005.
- Colosimo, S. F., Natraj, V., Sander, S. P., and Stutz, J. A sensitivity study on the retrieval of aerosol vertical profiles using the oxygen A-band. *Atmospheric Measurement Techniques*, 9(4): 1889–1905, doi: 10.5194/amt-9-1889-2016, 2016.
- Corradini, S. and Cervino, M. Aerosol extinction coefficient profile retrieval in the oxygen A-band considering multiple scattering atmosphere. Test case: SCIAMACHY nadir simulated measurements. *J. Quant. Spectrosc. Radiat. Transf.*, 97(3):354–380, doi: 10.1016/j.jqsrt.2005.05.061, 2006.
- de Graaf, M., de Haan, J., and Sanders, A. TROPOMI ATBD of the aerosol layer height. <https://sentinel.esa.int/documents/247904/2476257/Sentinel-5P-TROPOMI-ATBD-Aerosol-Height>, 2021.

## Bibliography

- de Haan, J. F., Bosma, P., and Hovenier, J. The adding method for multiple scattering calculations of polarized light. *Astronomy and astrophysics*, 183:371–391, 1987.
- Di Noia, A., Hasekamp, O., Van Harten, G., Rietjens, J., Smit, J., Snik, F., Henzing, J., De Boer, J., Keller, C., and Volten, H. Use of neural networks in ground-based aerosol retrievals from multi-angle spectropolarimetric observations. *Atmospheric Measurement Techniques*, 8(1):281, doi: 10.5194/amt-8-281-2015, 2015.
- Ding, S., Wang, J., and Xu, X. Polarimetric remote sensing in oxygen A and B bands: sensitivity study and information content analysis for vertical profile of aerosols. *Atmospheric Measurement Techniques*, 9(5):2077–2092, doi: 10.5194/amt-9-2077-2016, 2016.
- Doicu, A. and Trautmann, T. Adjoint problem of radiative transfer for a pseudo-spherical atmosphere and general viewing geometries. *Journal of Quantitative Spectroscopy and Radiative Transfer*, 110(8):464–476, doi: 10.1016/j.jqsrt.2009.01.027, 2009a.
- Doicu, A. and Trautmann, T. Two linearization methods for atmospheric remote sensing. *Journal of Quantitative Spectroscopy and Radiative Transfer*, 110(8):477–490, doi: 10.1016/j.jqsrt.2009.02.001, 2009b.
- Doicu, A. and Trautmann, T. Discrete-ordinate method with matrix exponential for a pseudo-spherical atmosphere: Scalar case. *J. Quant. Spectrosc. Radiat. Transf.*, 110:146–158, doi: 10.1016/j.jqsrt.2008.09.014, 2009c.
- Doicu, A. and Trautmann, T. Discrete-ordinate method with matrix exponential for a pseudo-spherical atmosphere: Vectors case. *J. Quant. Spectrosc. Radiat. Transf.*, 110:159–172, doi: 10.1016/j.jqsrt.2008.09.013, 2009d.
- Doicu, A., Trautmann, T., and Schreier, F. *Numerical regularization for atmospheric inverse problems*. Springer Science & Business Media, 2010.
- Doicu, A., Doicu, A., Efremenko, D. S., Loyola, D., and Trautmann, T. An overview of neural network methods for predicting uncertainty in atmospheric remote sensing. *Remote Sensing*, 13(24):5061, doi: 10.3390/rs13245061, 2021.
- 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. *Atmospheric Measurement Techniques*, 4(5):975–1018, doi: 10.5194/amt-4-975-2011, 2011.
- 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. *Journal of the atmospheric sciences*, 59(3):590–608, doi: 10.1175/1520-0469(2002)059<0590:VOAAOP>2.0.CO;2, 2002.
- 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  $\text{o}_2\text{-a}$ -band. *Remote Sensing of Environment*, 113(9):1899–1911, doi: 10.1016/j.rse.2009.04.018, 2009.
- Duforêt, L., Frouin, R., and Dubuisson, P. Importance and estimation of aerosol vertical structure in satellite ocean-color remote sensing. *Appl. Opt.*, 46(7):1107–1119, doi: 10.1364/AO.46.001107, Mar 2007.
- Fan, C., Fu, G., Di Noia, A., Smit, M., HH Rietjens, J., A Ferrare, R., Burton, S., Li, Z., and P Hasekamp, O. Use of a neural network-based ocean body radiative transfer model for aerosol retrievals from multi-angle polarimetric measurements. *Remote Sensing*, 11(23):2877, doi: 10.3390/rs11232877, 2019.

- Fan, Y., Li, W., Gatebe, C. K., Jamet, C., Zibordi, G., Schroeder, T., and Stamnes, K. Atmospheric correction over coastal waters using multilayer neural networks. *Remote Sensing of Environment*, 199:218–240, doi: 10.1016/j.rse.2017.07.016, 2017.
- Fisher, D., Muller, J.-P., and Yershov, V. N. Automated stereo retrieval of smoke plume injection heights and retrieval of smoke plume masks from AATSR and their assessment with CALIPSO and MISR. *IEEE Transactions on Geoscience and Remote Sensing*, 52(2):1249–1258, doi: 10.1109/TGRS.2013.2249073, 2014.
- Frankenberg, C., Hasekamp, O., O’Dell, C., Sanghavi, S., Butz, A., and Worden, J. Aerosol information content analysis of multi-angle high spectral resolution measurements and its benefit for high accuracy greenhouse gas retrievals. *Atmos. Meas. Tech.*, 5(7):1809–1821, doi: 10.5194/amt-5-1809-2012, 2012.
- Gao, M., Franz, B. A., Knobelspiesse, K., Zhai, P.-W., Martins, V., Burton, S., Cairns, B., Ferrare, R., Gales, J., Hasekamp, O., et al. Efficient multi-angle polarimetric inversion of aerosols and ocean color powered by a deep neural network forward model. *Atmospheric Measurement Techniques*, 14(6):4083–4110, doi: 10.5194/amt-14-4083-2021, 2021.
- Geddes, A. and Bösch, H. Tropospheric aerosol profile information from high-resolution oxygen A-band measurements from space. *Atmospheric Measurement Techniques*, 8(2):859–874, doi: 10.5194/amt-8-859-2015, 2015.
- Gerber, H. E. Relative-humidity parameterization of the Navy Aerosol Model (NAM). Technical report, Naval Research Lab Washington DC, 1985.
- Gong, S. A parameterization of sea-salt aerosol source function for sub-and super-micron particles. *Global biogeochemical cycles*, 17(4), doi: 10.1029/2003GB002079, 2003.
- Gordon, H. R. Atmospheric correction of ocean color imagery in the Earth Observing System era. *Journal of Geophysical Research: Atmospheres*, 102(D14):17081–17106, doi: 10.1029/96JD02443, 1997.
- Håkansson, N., Adok, C., Thoss, A., Scheirer, R., and Hörnquist, S. Neural network cloud top pressure and height for MODIS. *Atmospheric Measurement Techniques*, 11(5):3177–3196, doi: 10.5194/amt-11-3177-2018, 2018.
- Hasekamp, O. P. and Butz, A. Efficient calculation of intensity and polarization spectra in vertically inhomogeneous scattering and absorbing atmospheres. *Journal of Geophysical Research: Atmospheres*, 113(D20), doi: 10.1029/2008JD010379, 2008.
- Hess, M., Koepke, P., and Schult, I. Optical Properties of Aerosols and Clouds: The software package OPAC. *Bull. Am. Met. Soc.*, 79:831–844, doi: 10.1175/1520-0477(1998)079<0831:OPOAAC>2.0.CO;2, 1998.
- Hoeting, J. A., Madigan, D., Raftery, A. E., and Volinsky, C. T. Bayesian model averaging: a tutorial (with comments by M. Clyde, David Draper and EI George, and a rejoinder by the authors. *Statistical science*, 14(4):382–417, doi: 10.1214/ss/1009212519, 1999.
- Holben, B., Eck, T., Slutsker, I., Tanré, D., Buis, J., Setzer, A., Vermote, E., Reagan, J., Kaufman, Y., Nakajima, T., Lavenu, F., Jankowiak, I., and Smirnov, A. AERONET—A federated instrument network and data archive for aerosol characterization. *Remote Sensing of Environment*, 66(1):1–16, doi: 10.1016/S0034-4257(98)00031-5, 1998.
- Holl, G., Eliasson, S., Mendrok, J., and Buehler, S. SPARE-ICE: Synergistic ice water path from passive operational sensors. *Journal of Geophysical Research: Atmospheres*, 119(3):1504–1523, doi: 10.1002/2013JD020759, 2014.

## Bibliography

- Hollstein, A. and Fischer, J. Retrieving aerosol height from the oxygen A band: a fast forward operator and sensitivity study concerning spectral resolution, instrumental noise, and surface inhomogeneity. *Atmospheric Measurement Techniques*, 7(5):1429–1441, doi: 10.5194/amt-7-1429-2014, 2014.
- Hollstein, A. and Lindstrot, R. Fast reconstruction of hyperspectral radiative transfer simulations by using small spectral subsets: application to the oxygen A band. *Atmospheric Measurement Techniques*, 7(2):599–607, doi: 10.5194/amt-7-599-2014, 2014.
- Hovenier, J. W., Van der Mee, C. V., and Domke, H. *Transfer of polarized light in planetary atmospheres: basic concepts and practical methods*, volume 318. Springer Science & Business Media, doi: 10.1007/978-1-4020-2856-4, 2004.
- Jiménez, C., Eriksson, P., and Murtagh, D. Inversion of Odin limb sounding submillimeter observations by a neural network technique. *Radio Science*, 38(4):27–1, doi: 10.1029/2002RS002644, 2003.
- Kauppi, A., Kolmonen, P., Laine, M., and Tamminen, J. Aerosol-type retrieval and uncertainty quantification from OMI data. *Atmospheric Measurement Techniques*, 10(11):4079–4098, doi: 10.5194/amt-10-4079-2017, 2017.
- Kendall, M., Stuart, A., Ord, J., and Arnold, S. Vol. 2A: Classical inference and the linear model. *London [etc.]: Arnold [etc.]*, 1999.
- Kipling, Z., Stier, P., Johnson, C. E., Mann, G. W., Bellouin, N., Bauer, S. E., Bergman, T., Chin, M., Diehl, T., Ghan, S. J., et al. What controls the vertical distribution of aerosol relationships between process sensitivity in HadGEM3–UKCA and inter-model variation from AeroCom Phase II. *Atmospheric Chemistry and Physics*, 16(4):2221–2241, doi: 10.5194/acp-16-2221-2016, 2016.
- Kleipool, Q., Ludewig, A., Babić, L., Bartstra, R., Braak, R., Dierssen, W., Dewitte, P.-J., Kenter, P., Landzaat, R., Leloux, J., Loots, E., Meijering, P., van der Plas, E., Rozemeijer, N., Schepers, D., Schiavini, D., Smeets, J., Vacanti, G., Vonk, F., and Veeffkind, P. Pre-launch calibration results of the TROPOMI payload on-board the Sentinel-5 Precursor satellite. *Atmospheric Measurement Techniques*, 11(12):6439–6479, doi: 10.5194/amt-11-6439-2018, 2018.
- Koch, D. and Del Genio, A. D. Black carbon semi-direct effects on cloud cover: review and synthesis. *Atmospheric Chemistry and Physics*, 10(16):7685–7696, doi: 10.5194/acp-10-7685-2010, 2010.
- Koffi, B., Schulz, M., Bréon, F.-M., Griesfeller, J., Winker, D., Balkanski, Y., Bauer, S., Berntsen, T., Chin, M., Collins, W. D., et al. Application of the CALIOP layer product to evaluate the vertical distribution of aerosols estimated by global models: AeroCom phase I results. *Journal of Geophysical Research: Atmospheres*, 117(D10), doi: 10.1029/2011JD016858, 2012.
- Kokhanovsky, A. A. and Rozanov, V. V. The physical parameterization of the top-of-atmosphere reflection function for a cloudy atmosphere—underlying surface system: the oxygen A-band case study. *Journal of Quantitative Spectroscopy and Radiative Transfer*, 85(1):35–55, doi: 10.1016/S0022-4073(03)00193-6, 2004.
- Kokhanovsky, A. A. and Rozanov, V. V. The determination of dust cloud altitudes from a satellite using hyperspectral measurements in the gaseous absorption band. *Int. J. Remote Sensing*, 31(10):2729–2744, doi: 10.1080/01431160903085644, 2010.
- Koppers, G. A. A. and Murtagh, D. P. Retrieval of height resolved aerosol optical thickness in the atmospheric band. Stockholm University, Stockholm, Sweden, 1997.

- Landgraf, J., Hasekamp, O. P., Box, M. A., and Trautmann, T. A linearized radiative transfer model for ozone profile retrieval using the analytical forward-adjoint perturbation theory approach. *Journal of Geophysical Research: Atmospheres*, 106(D21):27291–27305, doi: 10.1029/2001JD000636, 2001.
- Levy, R. C., Remer, L. A., and Dubovik, O. Global aerosol optical properties and application to Moderate Resolution Imaging Spectroradiometer aerosol retrieval over land. *J. Geophys. Res. Atmos.*, 112(D13):n/a–n/a, doi: 10.1029/2006JD007815, 2007.
- Loyola, D. G., Thomas, W., Livschitz, Y., Ruppert, T., Albert, P., and Hollmann, R. Cloud properties derived from GOME/ERS-2 backscatter data for trace gas retrieval. *IEEE Trans. Geosci. Remote Sens.*, 45(9):2747–2758, doi: 10.1109/TGRS.2007.901043, 2007.
- Loyola, D. G., Pedernana, M., and Gimeno Garcia, S. Smart sampling and incremental function learning for very large high dimensional data. *Neural Networks*, 78:75–87, doi: 10.1016/j.neunet.2015.09.001, 2016.
- Loyola, D. G., Gimeno García, S., Lutz, R., Argyrouli, A., Romahn, F., Spurr, R. J. D., Pedernana, M., Doicu, A., Molina García, V., and Schüssler, O. The operational cloud retrieval algorithms from TROPOMI on board Sentinel-5 Precursor. *Atmos. Meas. Tech.*, 11(1):409–427, doi: 10.5194/amt-11-409-2018, 2018.
- Loyola, D. G., Xu, J., Heue, K.-P., and Zimmer, W. Applying FP\_ILM to the retrieval of Geometry-dependent Effective Lambertian Equivalent Reflectivity (GELER) daily maps from UVN satellite measurements. *Atmos. Meas. Tech.*, 13(2):985–999, doi: 10.5194/amt-13-985-2020, 2020.
- Ludewig, A., Kleipool, Q., Bartstra, R., Landzaat, R., Leloux, J., Loots, E., Meijering, P., van der Plas, E., Rozemeijer, N., Vonk, F., and Veeffkind, P. In-flight calibration results of the TROPOMI payload on board the Sentinel-5 Precursor satellite. *Atmospheric Measurement Techniques*, 13(7):3561–3580, doi: 10.5194/amt-13-3561-2020, 2020.
- Määttä, A., Laine, M., Tamminen, J., and Veeffkind, J. P. Quantification of uncertainty in aerosol optical thickness retrieval arising from aerosol microphysical model and other sources, applied to Ozone Monitoring Instrument (OMI) measurements. *Atmospheric Measurement Techniques*, 7(5):1185–1199, doi: 10.5194/amt-7-1185-2014, 2014.
- Maddy, E., DeSouza-Machado, S., Nalli, N., Barnet, C., L Strow, L., Wolf, W., Xie, H., Gambacorta, A., King, T., Joseph, E., et al. On the effect of dust aerosols on AIRS and IASI operational level 2 products. *Geophysical research letters*, 39(10), doi: 10.1029/2012GL052070, 2012.
- Marchuk, G. Equation for the value of information from weather satellites and formulation of inverse problems. *Cosmic Res*, 2:394–409, 1964.
- Marchuk, G. I. *Adjoint equations and analysis of complex systems*, volume 295. Springer Science & Business Media, 2013.
- Marshak, A., Herman, J., Adam, S., Karin, B., Carn, S., Cede, A., Geogdzhayev, I., Huang, D., Huang, L.-K., Knyazikhin, Y., et al. Earth observations from DSCOVR EPIC instrument. *Bulletin of the American Meteorological Society*, 99(9):1829–1850, doi: 10.1175/BAMS-D-17-0223.1, 2018.
- Meng, Z., Yang, P., Kattawar, G. W., Bi, L., Liou, K., and Laszlo, I. Single-scattering properties of tri-axial ellipsoidal mineral dust aerosols: A database for application to radiative transfer calculations. *Journal of Aerosol Science*, 41(5):501–512, doi: http://dx.doi.org/10.1016/j.jaerosci.2010.02.008, 2010.
- Morozov, V. A. On the solution of functional equations by the method of regularization. volume 167, pages 510–512. Russian Academy of Sciences, 1966.

## Bibliography

- Muller, J., Denis, M., Dundas, R. D., Mitchell, K. L., Naud, C., and Mannstein, H. On the solution of functional equations by the method of regularization. *International Journal of Remote Sensing*, 28(9):1921–1938, doi: 10.1080/01431160601030975, 2007.
- Nakajima, T. and Tanaka, M. Matrix formulations for the transfer of solar radiation in a plane-parallel scattering atmosphere. *Journal of Quantitative Spectroscopy and Radiative Transfer*, 35(1):13–21, doi: 10.1016/0022-4073(86)90088-9, 1986.
- Nakajima, T. and Tanaka, M. Algorithms for radiative intensity calculations in moderately thick atmospheres using a truncation approximation. *Journal of Quantitative Spectroscopy and Radiative Transfer*, 40(1):51–69, doi: 10.1016/0022-4073(88)90031-3, July 1988.
- Nanda, S., de Graaf, M., Veefkind, J. P., ter Linden, M., Sneep, M., de Haan, J., and Levelt, P. F. A neural network radiative transfer model approach applied to the TROPospheric Monitoring instrument.
- Nanda, S., de Graaf, M., Sneep, M., de Haan, J. F., Stammes, P., Sanders, A. F. J., Tuinder, O., Veefkind, J. P., and Levelt, P. F. Error sources in the retrieval of aerosol information over bright surfaces from satellite measurements in the oxygen A band. *Atmos. Meas. Tech.*, 11(1):161–175, doi: 10.5194/amt-11-161-2018, 2018.
- Natraj, V., Shia, R.-L., Huang, X., Margolis, J. S., and Yung, Y. L. Application of principal component analysis to high spectral resolution radiative transfer: A case study of the  $\text{o}_2\text{A}$  band. *J. Quant. Spectrosc. Radiat. Transf.*, 95(4):539–556, doi: 10.1016/j.jqsrt.2004.12.024, 2005.
- Natraj, V., Shia, R.-L., and Yung, Y. L. On the use of principal component analysis to speed up radiative transfer calculations. *J. Quant. Spectrosc. Radiat. Transf.*, 111(3):810–816, doi: 10.1016/j.jqsrt.2009.11.004, 2010.
- Noia, A. D., Hasekamp, O. P., Wu, L., Diedenhoven, B. v., Cairns, B., and Yorks, J. E. Combined neural network/Phillips–Tikhonov approach to aerosol retrievals over land from the NASA Research Scanning Polarimeter. *Atmospheric Measurement Techniques*, 10(11):4235–4252, doi: 10.5194/amt-10-4235-2017, 2017.
- Park, S. S., Kim, J., Lee, H., Torres, O., Lee, K.-M., and Lee, S. D. Utilization of  $\text{O}_4$  slant column density to derive aerosol layer height from a space-borne uv–visible hyperspectral sensor: sensitivity and case study. *Atmospheric Chemistry and Physics*, 16(4):1987–2006, doi: 10.5194/acp-16-1987-2016, 2016.
- Patterson, H. D. and Thompson, R. Recovery of inter-block information when block sizes are unequal. *Biometrika*, 58(3):545–554, doi: 10.1093/biomet/58.3.545, 1971.
- Peters, K., Quaas, J., and Bellouin, N. Effects of absorbing aerosols in cloudy skies: a satellite study over the Atlantic Ocean. *Atmospheric Chemistry and Physics*, 11(4):1393–1404, doi: 10.5194/acp-11-1393-2011, 2011.
- Pflug, B. M. and Ruppert, T. Information content of measurements in the  $\text{o}_2\text{A}$ - and  $\text{o}_2\text{B}$ -bands for monitoring of aerosols from space. volume 1968, pages 533–544. SPIE, doi: 10.1117/12.154856, 1993.
- Pierangelo, C., Chédin, A., Heilliette, S., Jacquinet-Husson, N., and Armante, R. Dust altitude and infrared optical depth from AIRS. *Atmospheric Chemistry and Physics*, 4(7):1813–1822, doi: 10.5194/acp-4-1813-2004, 2004.
- Rodgers, C. D. *Inverse methods for atmospheric sounding: Theory and practice*, volume 2. World scientific, doi: 10.1142/3171, 2000.

- Rothman, L. S., Gordon, I. E., Barbe, A., Benner, D. C., Bernath, P. F., Birk, M., Boudon, V., Brown, L. R., Campargue, A., Champion, J.-P., et al. The HITRAN 2008 molecular spectroscopic database. *Journal of Quantitative Spectroscopy and Radiative Transfer*, 110(9-10): 533–572, doi: 10.1016/j.jqsrt.2009.02.013, 2009.
- Rozanov, V. V. and Lyapustin, A. I. Similarity of radiative transfer equation: Error analysis of phase function truncation techniques. *Journal of Quantitative Spectroscopy and Radiative Transfer*, 111(12):1964–1979, doi: 10.1016/j.jqsrt.2010.03.018, 2010.
- Rozanov, V. V. and Rozanov, A. V. Relationship between different approaches to derive weighting functions related to atmospheric remote sensing problems. *Journal of Quantitative Spectroscopy and Radiative Transfer*, 105(2):217–242, doi: 10.1016/j.jqsrt.2006.12.006, 2007.
- Sanders, A. F. J. and de Haan, J. F. Retrieval of aerosol parameters from the oxygen A band in the presence of chlorophyll fluorescence. *Atmos. Meas. Tech.*, 6(10):2725–2740, doi: 10.5194/amt-6-2725-2013, 2013.
- Sanders, A. F. J., de Haan, J. F., Sneep, M., Apituley, A., Stammes, P., Viteitez, M. O., Tilstra, L. G., Tuinder, O. N. E., Koning, C. E., and Veefkind, J. P. Evaluation of the operational Aerosol Layer Height retrieval algorithm for Sentinel-5 Precursor: application to  $\text{O}_2$  A band observations from GOME-2A. *Atmos. Meas. Tech.*, 8(11):4947–4977, doi: 10.5194/amt-8-4947-2015, 2015.
- Sanghavi, S., Martonchik, J. V., Landgraf, J., and Platt, U. Retrieval of the optical depth and vertical distribution of particulate scatterers in the atmosphere using  $\text{O}_2$  A- and B-band SCIAMACHY observations over Kanpur: a case study. *Atmos. Meas. Tech.*, 5(5):1099–1119, doi: 10.5194/amt-5-1099-2012, 2012.
- Sasi, S., Natraj, V., Molina García, V., Efremenko, D. S., Loyola, D., and Doicu, A. Model selection in atmospheric remote sensing with an application to aerosol retrieval from DSCOVR/EPIC, Part 1: Theory. *Remote Sens.*, 12(22), doi: 10.3390/rs12223724, 2020a.
- Sasi, S., Natraj, V., Molina García, V., Efremenko, D. S., Loyola, D., and Doicu, A. Model selection in atmospheric remote sensing with application to aerosol retrieval from DSCOVR/EPIC. Part 2: Numerical analysis. *Remote Sens.*, 12(21), doi: 10.3390/rs12213656, 2020b.
- Schreier, F. Optimized implementations of rational approximations for the voigt and complex error function. *Journal of Quantitative Spectroscopy and Radiative Transfer*, 112(6):1010–1025, doi: 10.1016/j.jqsrt.2010.12.010, 2011.
- Schreier, F., Garcia, S. G., Hedelt, P., Hess, M., Mendrok, J., Vasquez, M., and Xu, J. GARLIC–A general purpose atmospheric radiative transfer line-by-line infrared-microwave code: Implementation and evaluation. *Journal of Quantitative Spectroscopy and Radiative Transfer*, 137:29–50, doi: 10.1016/j.jqsrt.2013.11.018, 2014.
- Sears, T., Thomas, G., Carboni, E., A. Smith, A., and Grainger, R.  $\text{SO}_2$  as a possible proxy for volcanic ash in aviation hazard avoidance. *Journal of Geophysical Research: Atmospheres*, 118(11):5698–5709, doi: 10.1002/jgrd.50505, 2013.
- Shi, C., Hashimoto, M., Shiomi, K., and Nakajima, T. Development of an algorithm to retrieve aerosol optical properties over water using an artificial neural network radiative transfer scheme: First result from GOSAT-2/CAI-2. *IEEE Transactions on Geoscience and Remote Sensing*, doi: 10.1109/TGRS.2020.3038892, 2020.
- Smyth, G. K. and Verbyla, A. P. A conditional likelihood approach to residual maximum likelihood estimation in generalized linear models. *Journal of the Royal Statistical Society: Series B (Methodological)*, 58(3):565–572, doi: 10.1111/j.2517-6161.1996.tb02101.x, 1996.

## Bibliography

- Spurr, R. and Natraj, V. A linearized two-stream radiative transfer code for fast approximation of multiple-scatter fields. *Journal of Quantitative Spectroscopy and Radiative Transfer*, 112(16): 2630–2637, doi: 10.1016/j.jqsrt.2011.06.014, 2011.
- Spurr, R. Simultaneous derivation of intensities and weighting functions in a general pseudo-spherical discrete ordinate radiative transfer treatment. *Journal of Quantitative Spectroscopy and Radiative Transfer*, 75(2):129–175, doi: 10.1016/S0022-4073(01)00245-X, 2002.
- Spurr, R. LIDORT and VLIDORT: Linearized pseudo-spherical scalar and vector discrete ordinate radiative transfer models for use in remote sensing retrieval problems. volume 3, pages 229–271. Springer Verlag, Berlin, doi: 10.1007/978-3-540-48546-9\_7, 2008.
- Spurr, R. and Christi, M. Linearization of the interaction principle: Analytic jacobians in the “Radiants” model. *Journal of Quantitative Spectroscopy and Radiative Transfer*, 103(3):431–446, doi: 10.1016/j.jqsrt.2006.05.001, 2007.
- Spurr, R., Kurosu, T., and Chance, K. A linearized discrete ordinate radiative transfer model for atmospheric remote-sensing retrieval. *Journal of Quantitative Spectroscopy and Radiative Transfer*, 68(6):689–735, doi: 10.1016/S0022-4073(00)00055-8, 2001.
- Spurr, R. J. VLIDORT: A linearized pseudo-spherical vector discrete ordinate radiative transfer code for forward model and retrieval studies in multilayer multiple scattering media. *Journal of Quantitative Spectroscopy and Radiative Transfer*, 102(2):316–342, doi: 10.1016/j.jqsrt.2006.05.005, 2006.
- Stamnes, K. and Swanson, R. A. A new look at the discrete ordinate method for radiative transfer calculations in anisotropically scattering atmospheres. *Journal of Atmospheric sciences*, 38(2): 387–399, doi: 10.1175/1520-0469(1981)038<0387:ANLATD>2.0.CO;2, 1981.
- Stamnes, K., Tsay, S.-C., and Nakajima, T. Computation of eigenvalues and eigenvectors for the discrete ordinate and matrix operator methods in radiative transfer. *Journal of Quantitative Spectroscopy and Radiative Transfer*, 39(5):415–419, doi: 10.1016/0022-4073(88)90107-0, 1988.
- Strandgren, J., Bugliaro, L., Sehnke, F., and Schröder, L. Cirrus cloud retrieval with MSG/SE-VIRIs using artificial neural networks. *Atmospheric Measurement Techniques*, 10(9):3547–3573, doi: 10.5194/amt-10-3547-2017, 2017.
- Taylor, M., Kazadzis, S., Amiridis, V., and Kahn, R. Global aerosol mixtures and their multiyear and seasonal characteristics. *Atmos. Environ.*, 116:112–129, doi: 10.1016/j.atmosenv.2015.06.029, 2015.
- Torres, O., Bhartia, P. K., Herman, J. R., Ahmad, Z., and Gleason, J. Derivation of aerosol properties from satellite measurements of backscattered ultraviolet radiation: Theoretical basis. *J. Geophys. Res. Atmos.*, 103(D14):17099–17110, doi: 10.1029/98JD00900, 1998.
- Torres, O., Decae, R., Veefkind, J., and de Leeuw, G. OMI aerosol retrieval algorithm, OMI algorithm theoretical basis document: Clouds, aerosols and surface UV irradiance, Vol. 3, version 2, OMI-ATBD-03. <http://eospsso.gsfc.nasa.gov/eoshomepage/forscientists/atbd/docs/OMI/ATBD-OMI-03.pdf>, 2002.
- Ustinov, E. A. Adjoint sensitivity analysis of radiative transfer equation: temperature and gas mixing ratio weighting functions for remote sensing of scattering atmospheres in thermal IR. *Journal of Quantitative Spectroscopy and Radiative Transfer*, 68(2):195–211, doi: 10.1016/S0022-4073(00)00022-4, 2001.



- Ustinov, E. A. Atmospheric weighting functions and surface partial derivatives for remote sensing of scattering planetary atmospheres in thermal spectral region: general adjoint approach. *Journal of Quantitative Spectroscopy and Radiative Transfer*, 92(3):351–371, doi: 10.1016/j.jqsrt.2004.08.003, 2005.
- Vandenbussche, S., Kochenova, S., Vandaele, A. C., Kumps, N., and De Mazière, M. Retrieval of desert dust aerosol vertical profiles from IASI measurements in the TIR atmospheric window. *Atmospheric Measurement Techniques*, 6(10):2577–2591, doi: 10.5194/amt-6-2577-2013, 2013.
- Veefkind, J., Aben, I., McMullan, K., Förster, H., de Vries, J., Otter, G., Claas, J., Eskes, H., de Haan, J., Kleipool, Q., van Weele, M., Hasekamp, O., Hoogeveen, R., Landgraf, J., Snel, R., Tol, P., Ingmann, P., Voors, R., Kruizinga, B., Vink, R., Visser, H., and Levelt, P. TROPOMI on the ESA Sentinel-5 Precursor: A GMES mission for global observations of the atmospheric composition for climate, air quality and ozone layer applications. *Remote Sens. Environ.*, 120: 70–83, doi: 10.1016/j.rse.2011.09.027, 2012.
- Wahba, G. Practical approximate solutions to linear operator equations when the data are noisy. *SIAM Journal on Numerical Analysis*, 14(4):651–667, doi: 10.1137/0714044, 1977.
- Wahba, G. *Spline models for observational data*. SIAM, doi: 10.1137/1.9781611970128, 1990.
- Wang, D., Prigent, C., Aires, F., and Jimenez, C. A statistical retrieval of cloud parameters for the millimeter wave Ice Cloud Imager on board MetOp-SGs. *IEEE Access*, 5:4057–4076, doi: 10.1109/ACCESS.2016.2625742, 2017.
- Wang, J. and Christopher, S. A. Intercomparison between satellite-derived aerosol optical thickness and PM<sub>2.5</sub> mass: Implications for air quality studies. *Geophysical research letters*, 30(21), doi: 10.1029/2003GL018174, 2003.
- Wang, J., Xu, X., Ding, S., Zeng, J., Spurr, R., Liu, X., Chance, K., and Mishchenko, M. A numerical testbed for remote sensing of aerosols, and its demonstration for evaluating retrieval synergy from a geostationary satellite constellation of GEO-CAPE and GOES-R. *Journal of Quantitative Spectroscopy and Radiative Transfer*, 146:510–528, doi: 10.1016/j.jqsrt.2014.03.020, 2014.
- Wang, P., Stammes, P., van der A, R., Pinardi, G., and van Roozendaal, M. FRESCO+: an improved O<sub>2</sub> a-band cloud retrieval algorithm for tropospheric trace gas retrievals. *Atmos. . Phys.*, 8(21):6565–6576, doi: 10.5194/acp-8-6565-2008, 2008.
- Waquet, F., Riedi, J., Labonnote, L., Goloub, P., Cairns, B., Deuzé, J., and Tanré, D. Aerosol remote sensing over clouds using A-Train observations. *Journal of the atmospheric sciences*, 66 (8):2468–2480, doi: 10.1175/2009JAS3026.1, 2009.
- Wendisch, M., Hellmuth, O., Ansmann, A., Heintzenberg, J., Engelmann, R., Althausen, D., Eichler, H., Müller, D., Hu, M., Zhang, Y., and Mao, J. Radiative and dynamic effects of absorbing aerosol particles over the Pearl River Delta, China. *Atmospheric Environment*, 42 (25):6405–6416, doi: 10.1016/j.atmosenv.2008.02.033, 2008.
- Wilcox, E. M. Direct and semi-direct radiative forcing of smoke aerosols over clouds. *Atmospheric Chemistry and Physics*, 12(1):139–149, doi: 10.5194/acp-12-139-2012, 2012.
- Winker, D. M., Vaughan, M. A., Omar, A., Hu, Y., Powell, K. A., Liu, Z., Hunt, W. H., and Young, S. A. Overview of the CALIPSO mission and CALIOP data processing algorithms. *Journal of Atmospheric and Oceanic Technology*, 26(11):2310–2323, doi: 10.1175/2009JTECHA1281.1, 2009.

## Bibliography

- Wiscombe, W. The delta-M method: Rapid yet accurate radiative flux calculations for strongly asymmetric phase functions. *Journal of Atmospheric Sciences*, 34(9):1408–1422, doi: 10.1175/1520-0469(1977)034<1408:TDMRYA>2.0.CO;2, 1977.
- 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. *Geophysical research letters*, 43(16):8783–8790, doi: 10.1002/2016GL069848, 2016.
- Xu, J., Schüssler, O., Loyola R, D., Romahn, F., and Doicu, A. A novel ozone profile shape retrieval using Full-Physics Inverse Learning Machine (FP-ILM). *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, 10(12):5442–5457, doi: 10.1109/JSTARS.2017.2740168, 2017.
- Yang, Y., Marshak, A., Mao, J., Lyapustin, A., and Herman, J. A method of retrieving cloud top height and cloud geometrical thickness with oxygen A and B bands for the Deep Space Climate Observatory (DSCOVR) mission: Radiative transfer simulations. *Journal of Quantitative Spectroscopy and Radiative Transfer*, 122:141–149, doi: 10.1016/j.jqsrt.2012.09.017, 2013.
- Zakšek, K., Hort, M., Zaletelj, J., and Langmann, B. Monitoring volcanic ash cloud top height through simultaneous retrieval of optical data from polar orbiting and geostationary satellites. *Atmospheric Chemistry and Physics*, 13(5):2589–2606, doi: 10.5194/acp-13-2589-2013, 2013.
- Zhang, L., Li, Q. B., Gu, Y., Liou, K. N., and Meland, B. Dust vertical profile impact on global radiative forcing estimation using a coupled chemical-transport–radiative-transfer model. *Atmospheric Chemistry and Physics*, 13(14):7097–7114, doi: 10.5194/acp-13-7097-2013, 2013.

# Acknowledgments

Foremost, I would like to thank my supervisor Adrian Doicu for his consistent guidance throughout my PhD study. The dissertation would not have been completed without his efforts and help. He has given me a lot of security. I look up to him as my role model and thank him for his supervision and care. The same gratitude goes to Prof. Dr. Richard Bamler and Prof. Dr. Florian Seitz for their kind support.

A special thank goes to my mentor Jian Xu for his countless efforts, patience and support from the very beginning till the end of my PhD. I am very grateful for his affirmation and encouragement. The dissertation would not have been possible without him.

The gratitude also goes to Prof. Dr. Trautmann who has given me continuous support and care during my stay in the department of Atmospheric Processors. I am also thankful for Dr. Diego Lolya and Dr. Dmitry Efremenko who have given me a lot of professional advice and strong support for my work. I would also like to thank all the colleagues in the department of Atmospheric Processors, specially Nathalie, Bernd, Walter, Zhuoru, Lok, Song, Sruthy, Victor, Matthias, Ana, Zhibin. Because of them, my work has been easier and more enjoyable. The experience in the department has left me with many good memories that I will cherish for a life time.

I would like to thank the China Scholarship Council and the Remote Sensing Technology Institute (IMF) at DLR for their financial support.

I would like to take the the opportunity to acknowledge Prof. Dr. Kai Qin, my Master thesis supervisor, who made my PhD study in Germany possible and has helped me a lot over the past years.

Finally and most importantly, I would like to express my gratitude to my family for their unceasing love, understanding and support that made me who I am today.



# Appendices



# A Hyperspectral Satellite Remote Sensing of Aerosol Parameters

Rao, L.; Xu, J.; Efremenko, D.S.; Loyola, D.G.; Doicu, A. Hyperspectral Satellite Remote Sensing of Aerosol Parameters. *Front. Environ. Sci.* 2022, 9, doi: 10.3389/fenvs.2021.770662



# Hyperspectral Satellite Remote Sensing of Aerosol Parameters: Sensitivity Analysis and Application to TROPOMI/S5P

Lanlan Rao<sup>1,2</sup>, Jian Xu<sup>3,1\*</sup>, Dmitry S. Efremenko<sup>1</sup>, Diego G. Loyola<sup>1</sup> and Adrian Doicu<sup>1</sup>

<sup>1</sup>German Aerospace Center (DLR), Remote Sensing Technology Institute, Oberpfaffenhofen, Germany, <sup>2</sup>Department of Aerospace and Geodesy, Technische Universität München, Munich, Germany, <sup>3</sup>National Space Science Center, Chinese Academy of Sciences, Beijing, China

## OPEN ACCESS

### Edited by:

Tianjie Zhao,  
Aerospace Information Research  
Institute (CAS), China

### Reviewed by:

Husi Letu,  
Aerospace Information Research  
Institute (CAS), China  
Yerong Wu,  
Hunan University of Science and  
Technology, China

### \*Correspondence:

Jian Xu  
xujian@nssc.ac.cn

### Specialty section:

This article was submitted to  
Environmental Informatics and Remote  
Sensing,  
a section of the journal  
Frontiers in Environmental Science

**Received:** 04 September 2021

**Accepted:** 27 December 2021

**Published:** 14 January 2022

### Citation:

Rao L, Xu J, Efremenko DS, Loyola DG  
and Doicu A (2022) Hyperspectral  
Satellite Remote Sensing of Aerosol  
Parameters: Sensitivity Analysis and  
Application to TROPOMI/S5P.  
*Front. Environ. Sci.* 9:770662.  
doi: 10.3389/fevs.2021.770662

Precise knowledge about aerosols in the lower atmosphere (optical properties and vertical distribution) is particularly important for studying the Earth's climatic and weather conditions. Measurements from satellite sensors in sun-synchronous and geostationary orbits can be used to map distributions of aerosol parameters in global or regional scales. The new-generation sensor Tropospheric Monitoring Instrument (TROPOMI) onboard the Copernicus Sentinel-5 Precursor (S5P) measures a wide variety of atmospheric trace gases and aerosols that are associated with climate change and air quality using a number of spectral bands between the ultraviolet and the shortwave infrared. In this study, we perform a sensitivity analysis of the forward model parameters and instrument information that are associated with the retrieval accuracy of aerosol layer height (ALH) and optical depth (AOD) using the oxygen (O<sub>2</sub>) A-band. Retrieval of aerosol parameters from hyperspectral satellite measurements requires accurate surface representation and parameterization of aerosol microphysical properties and precise radiative transfer calculations. Most potential error sources arising from satellite retrievals of aerosol parameters, including uncertainties in aerosol models, surface properties, solar/satellite viewing geometry, and wavelength shift, are analyzed. The impact of surface albedo accuracy on retrieval results can be dramatic when surface albedo values are close to the critical surface albedo. An application to the real measurements of two scenes indicates that the retrieval works reasonably in terms of retrieved quantities and fit residuals.

**Keywords:** aerosol retrievals, aerosol layer height, O<sub>2</sub> A-band, radiative transfer, TROPOMI/S5P

## 1 INTRODUCTION

Estimating aerosol optical properties and vertical distribution appears to be a challenging task because of real-time variations in aerosol microphysical properties. Remote sensing techniques for measuring aerosol properties from space have been developing rapidly and can be classified into two major groups. Active remote sensors such as the Cloud-Aerosol Lidar with Orthogonal Polarization (CALIOP) instrument measure backscattered signal and have the capability to accurately profile the scattering/absorption owing to aerosols/clouds in the atmosphere, whereas passive remote sensors can by far not offer the same level of details, but provide a global mapping of aerosol properties. Although aerosol height information with a vertical resolution as fine as 30 m can be obtained, the



CALIPSO observations possess a limited spatial coverage (Winker et al., 2009). In regard to passive sensors, considerable effort has been put into derive aerosol vertical information by employing the O<sub>2</sub>-O<sub>2</sub> absorption band (~477 nm), e.g., from the Ozone Monitoring Instrument (OMI) (Park et al., 2016; Chimot et al., 2017, 2018; Choi et al., 2019). Absorption of reflected sunlight by O<sub>2</sub> in its A band (~760 nm) has been extensively used to derive cloud height information, and the relevant studies can be found in Kokhanovsky et al. (2006a), Kokhanovsky et al. (2006b), Wang et al. (2008), Lelli et al. (2014), Loyola Rodriguez et al. (2007), Loyola et al. (2018). A number of passive satellite sensors have been launched to monitor aerosol properties on a global or regional scale using spectral information at various wavelengths. Atmospheric absorption in the O<sub>2</sub> A-band provides an opportunity to derive vertical distributions of aerosols as a result of the dynamic range of optical depth in this spectral domain. Recently, a great amount of efforts have been made to retrieve the aerosol height information from the O<sub>2</sub> A-band, e.g., the Scanning Imaging Absorption Spectrometer for Atmospheric Chartography (SCIAMACHY) (Corradini and Cervino, 2006; Kokhanovsky and Rozanov, 2010; Sanghavi et al., 2012), the Global Ozone Mapping Experiment (GOME) (Koppers and Murtagh, 1997) and GOME-2 (Tilstra et al., 2019), the Greenhouse Gases Observing Satellite (GOSAT) (Frankenberg et al., 2012), the Orbiting Carbon Observatory-2 (OCO-2) (Zeng et al., 2020). Some studies also focused on the joint use of O<sub>2</sub> A and B bands for vertical profiling of aerosols (Ding et al., 2016; Xu et al., 2017b).

As a new generation of hyperspectral sensor, the Tropospheric Monitoring Instrument onboard the Copernicus Sentinel-5 Precursor satellite (TROPOMI/S5P, hereafter referred to as TROPOMI) (Veefkind et al., 2012) was designed to be a push-broom grating spectrometer observing trace gas concentrations and aerosol/cloud properties that are associated with air quality, ozone layer, and climate forcing. The satellite flies in a sun-synchronous orbit at 824 km altitude with an Equator crossing time of 13:30 local solar time, allowing to achieve a full daily global surface coverage thanks to a wide swath of 108° (~2,600 km). The recorded TROPOMI spectra cover the ultraviolet-visible (UV-Vis, 270–500 nm), near-infrared (NIR, 675–775 nm), and shortwave infrared (SWIR, 2,305–2,385 nm). TROPOMI is the first Copernicus mission for atmospheric monitoring, launched on October 13, 2017, for a nominal lifetime of 7 years. In addition to the broad spectral coverage, TROPOMI can map global distributions of a broad range of air pollutants with a spatial resolution as high as  $5.5 \times 3.5 \text{ km}^2$  ( $7.0 \times 3.5 \text{ km}^2$  prior to August 6, 2019). Band 6 of TROPOMI covers the O<sub>2</sub> A-band and records the radiances and solar irradiances with a spectral sampling of 0.125–0.126 nm and a spectral resolution of 0.34–0.35 nm. The first calibration observations showed 3,000–5,000 and 250–700 for the high- and low-albedo signal-to-noise ratios, respectively. The main products of Band 6 are aerosols (height) and clouds (height and optical thickness). Further details of the instrument and measurement characteristics can be found in Kleipool et al. (2018), Ludewig et al. (2020).

Aerosol parameters like UV aerosol index, aerosol layer height (ALH) and optical depth (AOD) are useful to the global monitoring of air pollution in the lower atmosphere. The TROPOMI operational ALH retrieval algorithms in the O<sub>2</sub> A-band were developed by the Royal Netherlands Meteorological Institute (KNMI) and use a neural network based forward model and the optimal estimation method for inversion (Rodgers, 2000). For more details about the operational retrieval algorithms see (Sanders and de Haan, 2013; Sanders et al., 2015; Nanda et al., 2018a; Nanda et al., 2018b).

Deriving aerosol information from satellite measurements remains a critical challenge in terms of retrieval sensitivity and accuracy. This is in general an underdetermined task and often requires several assumptions to be made with respect to the properties of aerosol and surface (Kokhanovsky and Rozanov, 2010). Li et al. (2009) also discussed several critical factors affecting the accuracy of aerosol remote sensing, including the assumptions in the aerosol model, treatment of the underlying surface, sensor calibration, and cloud screening. In a conventional retrieval framework, an operational retrieval handles the minimization of the objective function, which should include sufficiently fast radiative transfer computations, and is capable of dealing with large amount of satellite measurements and needs to converge robustly. Retrievals from synthetic measurements are necessary and important for analyzing the impact of forward and instrument model parameters on the retrieval output and exploring the expected retrieval performance using real measurements. Based on these experiments, an optimal retrieval setup for realistic measurement conditions and a better understanding of instrument characteristics could be achieved.

We have developed a conventional retrieval framework dedicated to estimating aerosol and cloud parameters from satellite measurements. Retrieval applications to the Earth Polychromatic Imaging Camera (EPIC) onboard the Deep Space Climate Observatory (DSCOVER) satellite were reported (Molina García et al., 2018a; Molina García et al., 2018b; Sasi et al., 2020a; Sasi et al., 2020b). In this work, we adapt the framework to the TROPOMI measurements, and the primary objective is to evaluate the retrieval feasibility and accuracy of aerosol parameters (ALH and AOD) using the O<sub>2</sub> A-band of TROPOMI. Concerning the associated retrieval error characterization for the O<sub>2</sub> A-band, only a few sensitivity studies were carried out (Hollstein and Fischer, 2014; Sanders et al., 2015). In this study, we extend the sensitivity analysis by taking into account more inputs during the inversion, i.e., different models for aerosol microphysical parameterization, surface properties, solar/viewing geometry, and wavelength shift. These inputs and information are considered to likely affect the retrieval accuracy and this sensitivity analysis aims to quantify the impact and importance of each input. Additionally, an application with real TROPOMI data can help us to better understand the measurement characteristics and retrieval performance in reality. We seek a characterization of the associated retrieval error by reasonably assuming uncertainties on the crucial inputs

identified in the sensitivity analysis. Accordingly, the retrieval could be further optimized by refining these inputs.

The remainder of the article is formulated as follows: A brief description of the retrieval algorithm is given in **Section 2**. **Section 3** analyzes the sensitivity of retrieved aerosol parameters to different parameters and information associated with the instrument itself and radiative transfer calculations. A retrieval application using real TROPOMI measurements is given in **Section 4**. **Section 5** concludes the study.

## 2 THEORY

We have developed an algorithm dedicated to aerosol parameters retrieval from hyperspectral satellite sensors like TROPOMI. The theoretical concepts of atmospheric retrieval are presented in this section. This physical retrieval algorithm consists of a forward model in which radiative transfer of electromagnetic radiation through the atmosphere is calculated, and an inversion process in which a nonlinear minimization problem is solved. The purpose of the forward model is to simulate the signal received by the sensor as a function of atmospheric parameters and surface properties of interest, employing the discrete ordinate method to solve the radiative transfer equation. Inverse problems arising in atmospheric retrieval are typically ill-posed in the sense that perturbations in the data can cause large errors in the retrieval result. Our retrieval problem is formulated as a nonlinear least squares problem and can be solved by the Gauss–Newton method with the aid of Tikhonov regularization. In this study, the retrieval relies on the TROPOMI O<sub>2</sub> A-band (758–771 nm). The recorded radiances and solar irradiances are converted to the reflectances. The inversion returns the best estimates of the retrieval target by approximating the measured reflectances with the simulated ones. In the forward model, aerosols are assumed as a single atmospheric layer with a fixed thickness of 0.5 km. The retrieval target ALH is defined as the middle height of this aerosol layer. **Section 2.1** describes the physical and mathematical fundamentals of radiative transfer and different models for characterizing aerosol microphysical properties. **Section 2.2** presents the inversion procedure and associated approaches.

### 2.1 Radiative Transfer

In a pseudo-spherical atmosphere, the radiative transfer equation for the diffuse radiance  $I(r, \Omega)$  at point  $r$  in direction  $\Omega = (\mu, \varphi)$  is given by

$$\mu \frac{dI}{dr}(r, \Omega) = -\sigma_{\text{ext}}(r)I(r, \Omega) + J(r, \Omega), \quad (1)$$

where

$$\begin{aligned} J(r, \Omega) &= J_{\text{ss}}(r, \Omega) + J_{\text{ms}}(r, \Omega), \\ J_{\text{ss}}(r, \Omega) &= F_0 \frac{\sigma_{\text{sct}}(r)}{4\pi} P(r, \Omega, \Omega_0) e^{-\tau_{\text{ext}}^0(|r-r_{\text{TOA}}(r)|)}, \\ J_{\text{ms}}(r, \Omega) &= \frac{\sigma_{\text{sct}}(r)}{4\pi} \int_{4\pi} P(r, \Omega, \Omega') I(r, \Omega') d\Omega', \end{aligned}$$

is the source function summing the contributions of the single and multiple scattering terms  $J_{\text{ss}}(r, \Omega)$  and  $J_{\text{ms}}(r, \Omega)$ , respectively,  $\sigma_{\text{ext}}$  and  $\sigma_{\text{sct}}$  are the extinction and scattering coefficients, respectively,  $F_0$  is the incident solar flux,  $P$  the scattering phase function,  $\Omega_0 = (-\mu_0, \varphi_0)$  with  $\mu_0 > 0$  the incident solar direction, and  $\tau_{\text{ext}}^0(|r-r_{\text{TOA}}(r)|)$  the solar optical depth between point  $r$  and the characteristic point at the top of the atmosphere  $r_{\text{TOA}}$  in a spherical atmosphere. For the phase function  $P$ , we assume an expansion in terms of normalized Legendre polynomials  $P_n$ , i.e.,

$$P(r, \Omega, \Omega') = P(r, \cos \Theta) = \sum_{n=0}^{\infty} \sqrt{\frac{2n+1}{2}} \chi_n(r) P_n(\cos \Theta), \quad (2)$$

where  $\cos \Theta = \Omega \cdot \Omega'$ . The boundary conditions associated to the radiative transfer (**Eq. 1**) consist in the top-of-atmosphere boundary condition ( $r = r_{\text{TOA}}$ ),

$$I(r_{\text{TOA}}, \Omega^-) = 0, \quad (3)$$

and the surface boundary condition ( $r = r_s$ ),

$$I(r_s, \Omega^+) = F_0 \frac{A}{\pi} \mu_0 e^{-\tau_{\text{ext}}^0(|r_s-r_{\text{TOA}}(r_s)|)} + \frac{A}{\pi} \int_{2\pi} I(r_s, \Omega^-) |\mu^-| d\Omega^-, \quad (4)$$

where for a Lambertian surface,  $A$  is the surface albedo, and the notations  $\Omega^+$  and  $\Omega^-$  stand for upward and downward directions, respectively.

The aerosol particles are modeled as components, while the size distribution  $dN(a)/d \ln a$  of an aerosol component is chosen a log-normal distribution, characterized by the modal radius  $r_{\text{mod}}$ , the standard deviation  $\sigma$ , and the total number of particles  $N_0$ . If these parameters together with the (wavelength-dependent) refractive index  $m_{\text{aer}}$  are specified, the size averaged extinction and scattering cross sections, as well as the expansion coefficients of the size averaged phase function are computed as

$$\begin{aligned} \bar{C}_{\text{ext}} &= \int_{a_{\text{min}}}^{a_{\text{max}}} C_{\text{ext}}(a) p(a) da, \\ \bar{C}_{\text{sct}} &= \int_{a_{\text{min}}}^{a_{\text{max}}} C_{\text{sct}}(a) p(a) da, \\ \bar{\chi}_n &= \int_{a_{\text{min}}}^{a_{\text{max}}} \chi_n(a) p(a) da, \quad n \geq 0, \end{aligned}$$

respectively, where  $a_{\text{min}}$  and  $a_{\text{max}}$  are the lower and upper bounds of the size distribution,  $p(a) = (1/N_0)dN(a)/da$  is the probability density function associated to the number size distribution, and  $C_{\text{ext}}(a)$ ,  $C_{\text{sct}}(a)$ , and  $\chi_n(a)$  are computed by an electromagnetic scattering code for a spherical particle of radius  $a$ . The aerosol components are externally mixed to form aerosol models. For an aerosol model  $m$  consisting of  $N$  aerosol components, the extinction and scattering cross sections, and the expansion coefficients of the phase function are computed as

$$\begin{aligned} \bar{C}_{\text{ext}}^{(m)} &= \sum_{i=1}^N w_i \bar{C}_{\text{ext}}^{(i)}, \\ \bar{C}_{\text{sct}}^{(m)} &= \sum_{i=1}^N w_i \bar{C}_{\text{sct}}^{(i)}, \\ \bar{\chi}_n^{(m)} &= \frac{1}{\bar{C}_{\text{sct}}^{(m)}} \sum_{i=1}^N w_i \bar{C}_{\text{sct}}^{(i)} \bar{\chi}_n^{(i)}, \end{aligned}$$

respectively, where the weight  $w_i = N_0^{(i)} / \sum_{i=1}^N N_0^{(i)}$  is number mixing ratio, and  $\bar{C}_{\text{ext}}^{(i)}$ ,  $\bar{C}_{\text{sct}}^{(i)}$ ,  $\bar{\chi}_n^{(i)}$ , and  $N_0^{(i)}$  correspond to the  $i$ th aerosol component. In this context, the extinction and scattering coefficients that enter into the radiative transfer model are calculated as  $\sigma_{\text{ext}}^{(m)} = n_0 \bar{C}_{\text{ext}}^{(m)}$  and  $\sigma_{\text{sct}}^{(m)} = n_0 \bar{C}_{\text{sct}}^{(m)}$ , respectively, where  $n_0 = \sum_{i=1}^N n_0^{(i)}$  is the total number density of the aerosol particles, and  $n_0^{(i)}$  the number density of the  $i$ th aerosol component.

A set of aerosol models is an ensemble of a variety of aerosol models corresponding to different aerosol types. In the forward model, the following sets of aerosol models are taken into account:

- **Set I** The aerosol models employed in the Moderate Resolution Imaging Spectroradiometer (MODIS) aerosol retrieval algorithm (Levy et al., 2007a; Levy et al., 2007b). There are three fine-dominated (spherical) and one coarse-dominated (spheroid) aerosol models that differ by the single scattering albedo  $\bar{\omega}^{(m)} = \bar{C}_{\text{sct}}^{(m)} / \bar{C}_{\text{ext}}^{(m)}$ ; we distinguish moderately absorbing ( $\bar{\omega}_m = 0.90$ ), absorbing ( $\bar{\omega}_m = 0.85$ ), and nonabsorbing ( $\bar{\omega}_m = 0.95$ ) aerosols. For each aerosol model, the parameters of the size distribution and the refractive index depend on the aerosol optical depth.
- **Set II** The aerosol models employed in the Ozone Monitoring Instrument (OMI) Multiwavelength aerosol retrieval algorithm (Torres et al., 1998). There are five major aerosol types, i.e., weakly absorbing, biomass burning, desert dust, marine, and volcanic, whereby each type consists of several aerosol models depending on their optical properties and particle size distribution.
- **Set III** The aerosol models (mixtures of sulfate, dust, sea salt, black carbon, and organic carbon components) obtained by a cluster analysis using the Goddard Chemistry Aerosol Radiation and Transport (GOCART) model (Chin et al., 2002; Taylor et al., 2015).
- **Set IV** The aerosol models (mixtures of water-insoluble, water-soluble, soot, sea-salt, mineral, mineral transported, and sulfate components) included in the Optical Properties of Aerosols and Clouds (OPAC) dataset (Hess et al., 1998; Thomas et al., 2009).

The radiative transfer computation relies on the discrete ordinate method with matrix exponential (Doicu and Trautmann, 2009a; Doicu and Trautmann, 2009b). To deal with computationally expensive radiative transfer calculations in the TROPOMI O<sub>2</sub> A-band absorption channel, several acceleration techniques, as for example, the telescoping technique (Spurr, 2008; Efremenko et al., 2013), the method of false discrete ordinate, the correlated  $k$ -distribution method (Goody et al., 1989), and principal component analysis (Natraj et al., 2005, 2010) are implemented.

In Section 3.1 we investigate the impact of different aerosol models on the retrieval performance.

## 2.2 Inversion

The retrieval is performed by using the method of Tikhonov regularization (Tikhonov, 1963). Essentially, the inverse problem is solved by minimizing the objective function,

$$\mathcal{F}(\mathbf{x}) = \frac{1}{2} \left[ \|\mathbf{F}(\mathbf{x}) - \mathbf{y}^\delta\|^2 + \lambda \|\mathbf{L}(\mathbf{x} - \mathbf{x}_a)\|^2 \right], \quad (5)$$

where  $\mathbf{F}: \mathbb{R}^n \rightarrow \mathbb{R}^m$  and  $\mathbf{y}^\delta \in \mathbb{R}^m$  are the vector-valued forward model and the noisy measurement vector, respectively,  $\lambda$  is the regularization parameter,  $\mathbf{L}$  the regularization matrix, and  $\mathbf{x}_a$  the *a priori* state vector.

The goal of minimizing the Tikhonov function (Eq. 5) is to search for a solution providing a compromise between the residual term  $\|\mathbf{F}(\mathbf{x}) - \mathbf{y}^\delta\|$  and the penalty term  $\|\mathbf{L}(\mathbf{x} - \mathbf{x}_a)\|$ . A global minimizer  $\mathbf{x}_\lambda$  is called a regularized solution. The minimization procedure can be performed by means of nonlinear optimization algorithms like Newton-type methods. At the iteration step  $i$ , the objective function is approximated by its linearization around the current iterate  $\mathbf{x}_{\lambda,i}$ . The regularized solution (the new iterate  $\mathbf{x}_{\lambda,i+1}$ ) is found by an iterative process:

$$\mathbf{x}_{\lambda,i+1} = \mathbf{x}_a + \mathbf{K}_{\lambda,i}^\dagger (\mathbf{y}^\delta - \mathbf{F}(\mathbf{x}_{\lambda,i}) + \mathbf{K}_i (\mathbf{x}_{\lambda,i} - \mathbf{x}_a)), \quad (6)$$

where

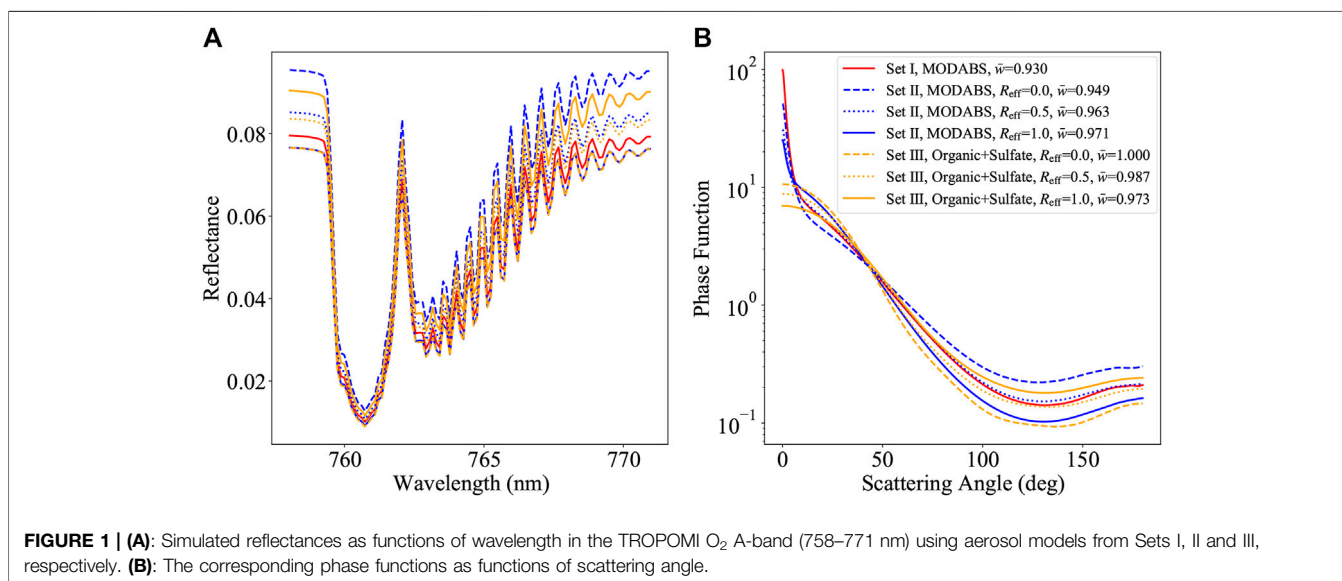
$$\mathbf{K}_{\lambda,i}^\dagger = (\mathbf{K}_i^T \mathbf{K}_i + \lambda \mathbf{L}^T \mathbf{L})^{-1} \mathbf{K}_i^T, \quad (7)$$

is the regularized generalized inverse matrix, and  $\mathbf{K}_i$  the Jacobian matrix of  $\mathbf{F}$  at  $\mathbf{x}_{\lambda,i}$ .

The penalty term  $\|\mathbf{L}(\mathbf{x} - \mathbf{x}_a)\|$  directly influences the inversion result, for instance, constraining the solution to be within a range. In fact, the regularization parameter  $\lambda$  controls the relative weight of the residual term and the penalty term, whereas the regularization matrix  $\mathbf{L}$  influences the magnitude or smoothness of the solution. An appropriate value of  $\lambda$  (constant or changeable during iterations) and a proper choice of  $\mathbf{L}$  (e.g., the identity matrix, discrete approximations of the first and second order derivative operators, the Cholesky factor of the *a priori* profile covariance matrix) can help to obtain a solution with a well-defined physical sense. Xu et al. (2016), Xu et al. (2020) compared a variety of approaches for choosing  $\lambda$  and  $\mathbf{L}$ , and suggested optimal strategies for practical problems. Note that as an alternative, iterative regularization methods (e.g., the iteratively regularized Gauss-Newton method) that employ a monotonically decreasing sequence of  $\lambda$  and an *a posteriori* stopping criterion, have been proved to be effective in practice.

The iterative process is terminated when a favorable convergence is reached in accordance with predefined stopping tolerances. The favorable convergence is based on two tests:

1. the  $\mathbf{x}$ -convergence test, which checks if the sequence  $\mathbf{x}_{\lambda,i}$  is converging and the change in  $\mathbf{x}_\lambda$  satisfies a predefined criteria, and
2. the relative-function-convergence test, which checks if the relative change in  $\mathcal{F}$  is within a predefined value.



### 3 SENSITIVITY TO EXPECTED ERROR SOURCES

In practice, several assumptions in the forward model (e.g., aerosol model, instrument parameters etc.) are required. Nevertheless, it is of importance to investigate how imperfect forward model inputs and instrumental knowledge contribute to the retrieval performance. In this section, we analyze the impact of different inputs used in the forward and instrument models on the retrieval of aerosol properties from the O<sub>2</sub> A-band spectral measurements. Two types of uncertainty are discussed: 1) model uncertainty (aerosol model) and 2) parameter uncertainty (surface albedo, solar/sensor viewing geometry, and wavelength calibration). Only the most representative error sources identified in the consortium for satellite remote sensing of aerosols (also from previous studies, e.g., (Sanders and de Haan, 2013; Sanders et al., 2015) were considered in this study. The sensitivity analysis was based on simulated reflectance spectra that largely resemble typical TROPOMI measurements in the O<sub>2</sub> A-band.

Here, the state vector  $\mathbf{x}$  consists of two components, i.e., AOD  $\tau_{\text{aer}}$  and ALH  $h_{\text{aer}}$ . The principle of AOD retrieval lies on the aerosol scattering and absorption features in the O<sub>2</sub> A-band. The ALH retrieval depends mainly on a narrow oxygen absorption band (between 760 and 762 nm) where aerosol layer will attenuate the oxygen absorption below. The inversion results in this section are represented as retrieval errors (with respect to the true state) due to various inputs in the forward and instrument models.

#### 3.1 Aerosol Model

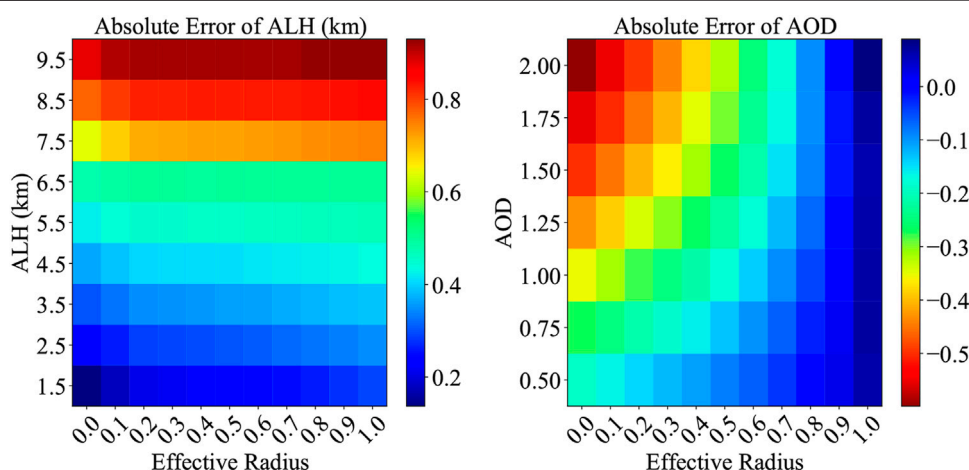
Each set of aerosol models described in **Section 2.1** is a collection of aerosol models that are employed to parameterize the aerosol microphysical properties for specific aerosol types, including the corresponding scattering and absorption properties, which plays an important role in an accurate retrieval. In this section, its

influence on the retrieval of AOD and ALH is discussed. For an assumption of aerosol microphysical properties, three sets of aerosol models were considered in this analysis: Sets I, II, and III that have been used in satellite retrievals and weather/climate model simulations, respectively.

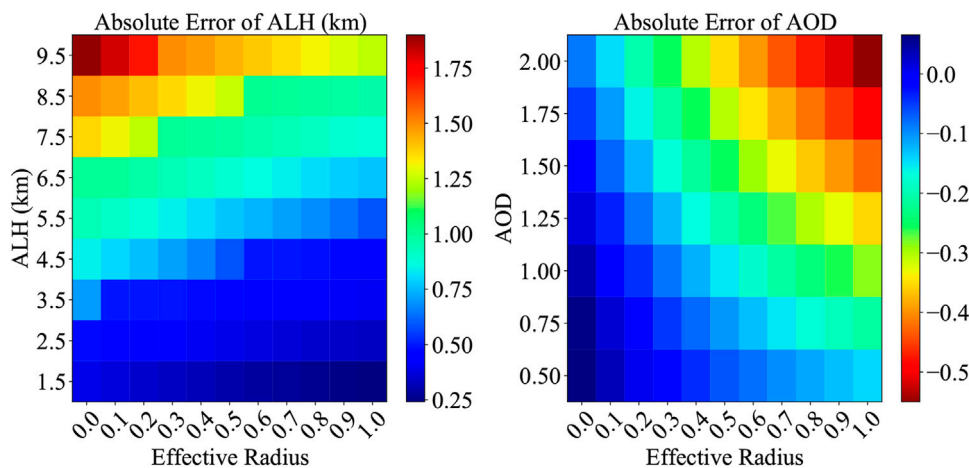
The simulated reflectance spectra (758–771 nm) using three sets of aerosol models are compared in **Figure 1**. “MODABS” from Set I, “MODABS” from Set II, and “Organic + Sulfate” from Set III were considered as the three models should represent the same aerosol characteristics. For Sets II and III, three values of 0.0, 0.5, and 1.0 for the effective radius  $R_{\text{eff}}$  were chosen. An aerosol loading scenario with 0.5 and 3.5 km for AOD and ALH, respectively, was considered. 0.05, 30°, 0°, and 180° were chosen for the surface albedo, solar and viewing zenith, and relative azimuth angles, respectively.

The left panel of **Figure 1** shows that the reflectance using Set I lies between the highest and lowest reflectances using Sets II and III. By using  $R_{\text{eff}} = 1.0$  and  $R_{\text{eff}} = 0.0$  for Sets II and III, respectively, the simulated reflectances are almost equivalent. Nevertheless, the spectra do not match perfectly between the three sets of aerosol models. For example, the blue solid line (“Organic + Sulfate” from Set III,  $R_{\text{eff}} = 0.0$ ) and the yellow dotted line (“MODABS” from Set II,  $R_{\text{eff}} = 1.0$ ) are close. Both spectra agree well between 760 and 762 nm where contains majority of aerosol height information, while slight discrepancies can be found elsewhere. The phase functions ( $P$ ) and single scattering albedos ( $\bar{\omega}$ ) computed with the parameters of the size distribution and the refractive index are compared as well in the right panel of **Figure 1**. Significant differences between different phase functions can be noticed when scattering angles are close to zero. These discrepancies should not be overlooked because it may produce a noticeable error in the retrieved ALH.

To study the influence of aerosol models on the retrieval results, an experiment using synthetic spectra simulated with



**FIGURE 2** | Retrieval (absolute) errors of ALH and AOD using “MODABS” from Set II. The synthetic spectra were simulated using “MODABS” from Set I.



**FIGURE 3** | Retrieval (absolute) errors of ALH and AOD using “Organic + Sulfate” from Set III. The synthetic spectra were simulated using “MODABS” from Set I.

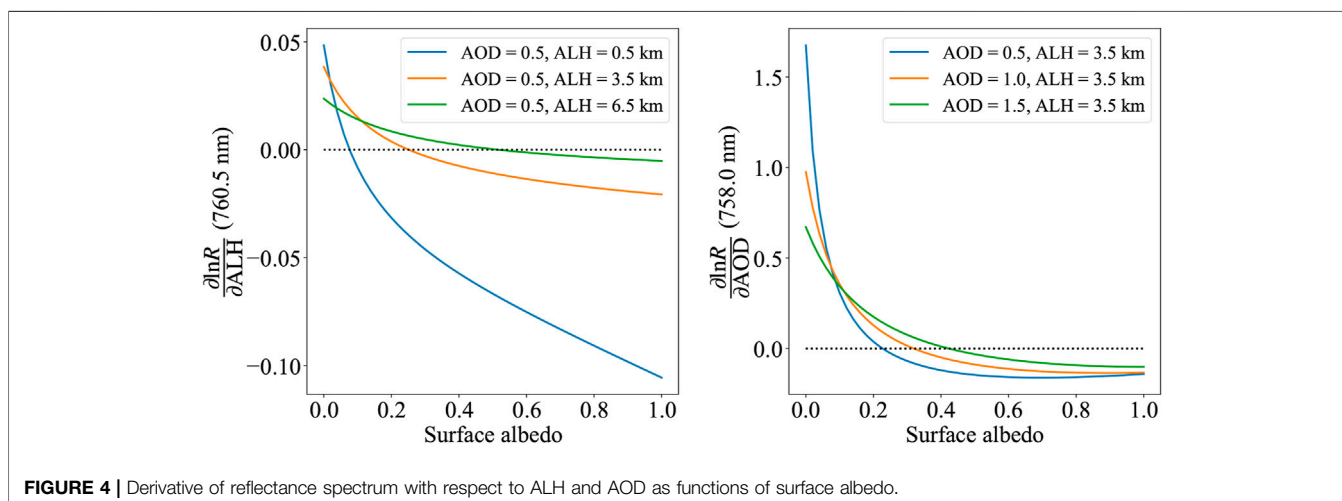
“MODABS” from Set I was carried out. In the subsequent retrieval procedure, “MODABS” from Set II and “Organic + Sulfate” from Set III were used, respectively. In this case, the underlying surface was assumed to be a dark surface (with an albedo of 0.05). The retrieval errors of ALH and AOD using Sets II and III (with  $R_{\text{eff}} = 0.0, 0.1, 0.2, \dots, 1.0$ ) are shown in **Figures 2, 3**, respectively. **Figures 2, 3** reveal that applying an inappropriate aerosol model may result in significant retrieval errors. In both figures, the error of ALH reaches up to 1.9 km when the true value is 9.5 km, and the largest negative error of AOD is about  $-0.5$  when the true value is 2.0. Interestingly, the retrieval error gradually increases in **Figure 2** with increasing  $R_{\text{eff}}$ , while the error in **Figure 3** behaves in an opposite way. Although the differences in the simulated reflectance spectra between the models characterizing the same aerosol type found in **Figure 1** can result in different retrieval outputs, its impact is estimated to be moderate.

### 3.2 Surface Albedo

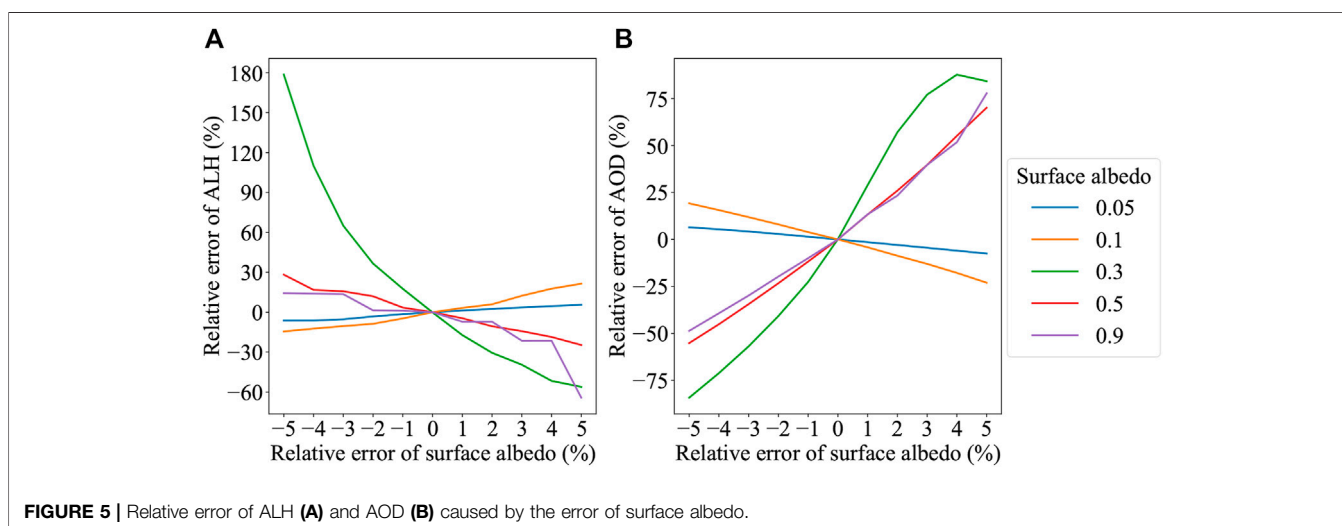
Apart from the aerosol parameterization, the previous sensitivity studies (Sanders et al., 2015; Nanda et al., 2018a) demonstrated that the accuracy of surface properties could greatly influence the aerosol retrieval from the  $\text{O}_2$  A-band measurements.

The sensitivity of the reflectance spectrum to the retrieved parameter can be described by its partial derivative with respect to this parameter. The partial derivatives of the reflectance (in terms of the natural logarithm) with respect to ALH and AOD as functions of surface albedo are illustrated in **Figure 4**. The left panel of **Figure 4** indicates that the derivatives decrease faster with the decreasing ALH when the value of AOD was assumed to be identical. As can be seen from the right panel of **Figure 4**, the derivative curves do not differ significantly between the three values of AOD.

A surface albedo at the turning point where the partial derivative of the reflectance with respect to ALH or AOD is



**FIGURE 4** | Derivative of reflectance spectrum with respect to ALH and AOD as functions of surface albedo.



**FIGURE 5** | Relative error of ALH (A) and AOD (B) caused by the error of surface albedo.

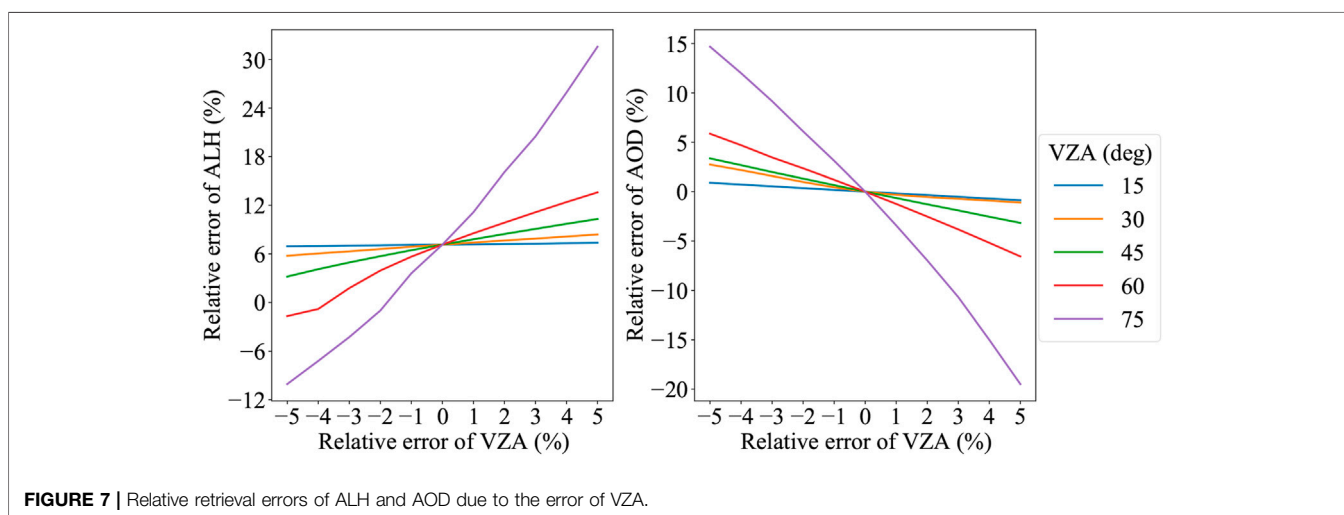
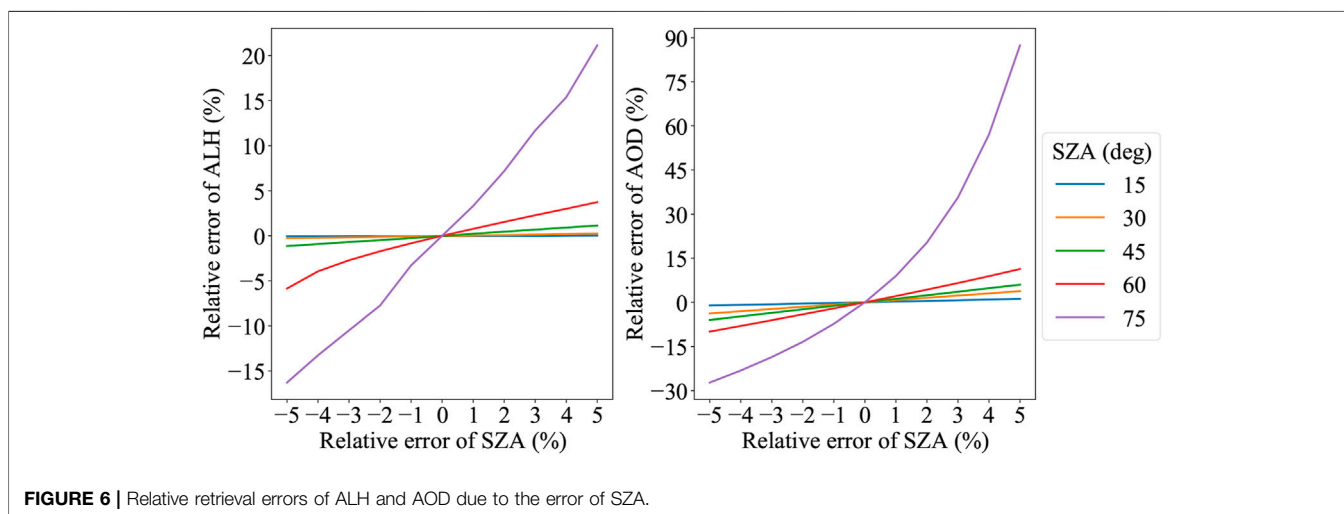
zero (dotted line), is called the critical surface albedo (Seidel and Popp, 2012), namely, this special surface albedo is “critical” for retrieval of ALH or AOD. It is noteworthy that the critical surface albedo varies with the value of ALH or AOD. **Figure 4** illustrates that the value of the critical surface albedo increases substantially with the increasing ALH (left panel for a fixed AOD), while the value of the critical surface albedo increases gradually when AOD increases (right panel for a fixed ALH).

The relative retrieval errors resulting from the uncertainty in the surface albedo is shown in **Figure 5**. In this case, 0.5 and 3.5 km were used as the true values of AOD and ALH, respectively, “MODABS” from Set I was used as the aerosol model. Over darker or less bright surface (with albedo values of 0.05 and 0.15), the uncertainty of the surface albedo seems to produce less impact on the aerosol retrieval. Apparently, the retrieval error is more pronounced over a brighter surface (with albedo values of 0.5 and 0.9). When the value of surface albedo is around the critical surface albedo, an error of 5% in the surface albedo could yield errors of about 180% (green line in the left

panel of **Figure 5**) and 80% (green line in the right panel of **Figure 5**) in the retrieved ALH and AOD, respectively.

The critical surface albedo is between 0.2 and 0.3 for ALH (orange line in the left panel of **Figure 4**) and AOD (blue line in the right panel of **Figure 4**). Along with **Figure 5**, an indication is that an overestimated surface albedo tends to introduce an underestimation of ALH and an overestimation of AOD for surface albedos higher than the critical surface albedo, whereas for surface albedos lower than the critical surface albedo, an underestimated surface albedo is inclined to cause an overestimation of ALH and an underestimation of AOD. When the surface albedo is close to the critical surface albedo, the retrieval can be exceptionally challenging even though the error of the surface albedo is small. In **Figure 5**, the largest retrieval errors of ALH and AOD correspond to the case with surface albedo of 0.3. These findings are consistent with the implications from (Seidel and Popp, 2012).

The consequence of using inaccurate surface albedos in retrieval can be serious, and a joint-fitting of the surface



albedo cannot guarantee an improved retrieval accuracy, albeit with more computational effort (Sanders et al., 2015). In practice, the geometry-dependent effective Lambertian equivalent reflectivity (GE\_LER) (Loyola et al., 2020) can be considered to be a reliable alternative. Its retrieval algorithm is based on the framework called the “Full-Physics Inverse Learning Machine” (FP-ILM) (Xu et al., 2017a; Efremenko et al., 2017; Hedelt et al., 2019). In contrast to LER climatologies, GE\_LER takes into account the drastically improved spatial resolution of TROPOMI and represents actual surface conditions accurately. The retrieved GE\_LER data has been included in the operational TROPOMI products for cloud and UV/VIS trace gases O<sub>3</sub>, SO<sub>2</sub>, and HCHO.

### 3.3 Geometry

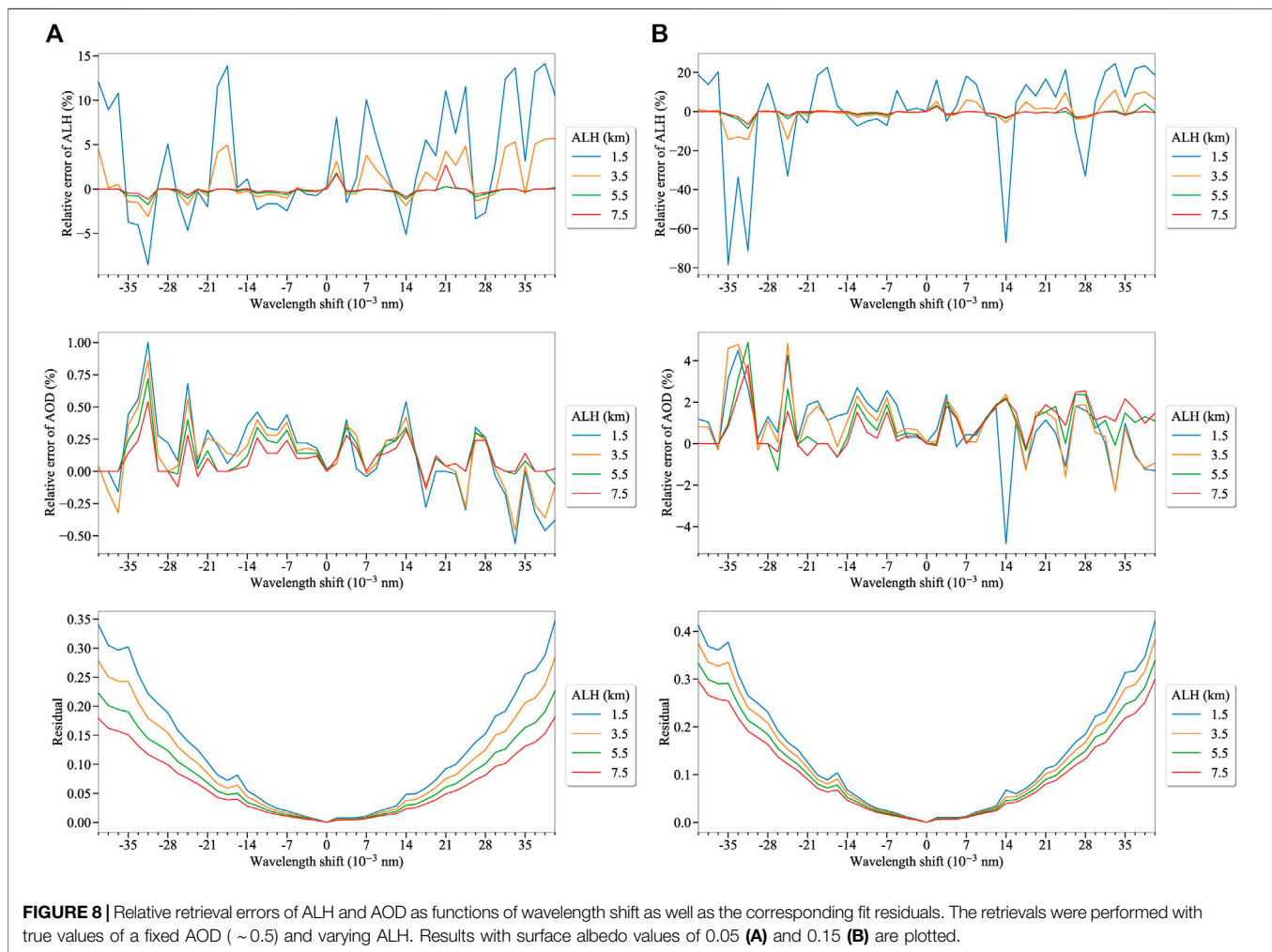
In this section, we discuss the influence of solar and sensor viewing geometry parameters on the aerosol retrieval. The simulation was performed by assuming an error in the solar zenith angle (SZA) and viewing zenith angle (VZA), respectively.

**Figure 6** shows the relative retrieval errors of AOD and ALH due to the relative error of SZA. As expected, when dealing with measurements at higher SZAs, the retrieval accuracy is more sensitive to the error of SZA. With an error of 5% in SZA at 75°, the relative error of retrieved ALH and AOD reaches up to 20 and 90%, respectively.

**Figure 7** shows the relative retrieval errors of AOD and ALH due to the relative error of VZA. Likewise, the retrieval error of both parameters increases with the increasing VZA. The relative error of retrieved ALH and AOD is up to 30 and 20%, respectively. As compared to **Figure 6**, VZA seems to impose less impact on the retrieved aerosol parameters. In reality, the accuracy of measured SZA and VZA is within 1% and estimated to make a minor impact on retrieval results.

### 3.4 Wavelength Calibration

An accurate wavelength calibration of the radiance and solar irradiance measurements is required during the Level-1b and Level-2 data processing. A wavelength shift is an offset found in



the spectral position of a measured signal. It is anticipated that uncertainty in the wavelength can introduce an error in the retrieval output. Here, we performed retrievals by assuming a wavelength shift with different combinations of aerosol parameters over two surface types (with albedo values of 0.05 and 0.15). **Figures 8, 9** depicts the retrieval errors of ALH and AOD as functions of the wavelength shift as well as the fit residuals. In **Figure 8**, the true value of ALH was assumed to be 1.5, 3.5, 5.5, and 7.5 km, respectively, and the true value of AOD was fixed to be 0.5; whereas in **Figure 9**, the true value of AOD was 0.5, 1.0, 1.5, and 2.0, respectively, and the true value of ALH was fixed to be 3.5 km.

The wavelength shift has a seemingly greater impact on the retrieval results over the brighter surface. The retrieval errors of ALH and AOD and the corresponding residuals are “augmented” by the larger surface albedo. The residual plots show a monotonic increase with the increasing wavelength shift, whereas the retrieval errors appear to increase oscillatingly with the increasing wavelength shift. The ALH retrieval only relies on the information from a narrow range (760–762 nm), and therefore, the error plots indicate a more pronounced impact on the ALH retrieval when

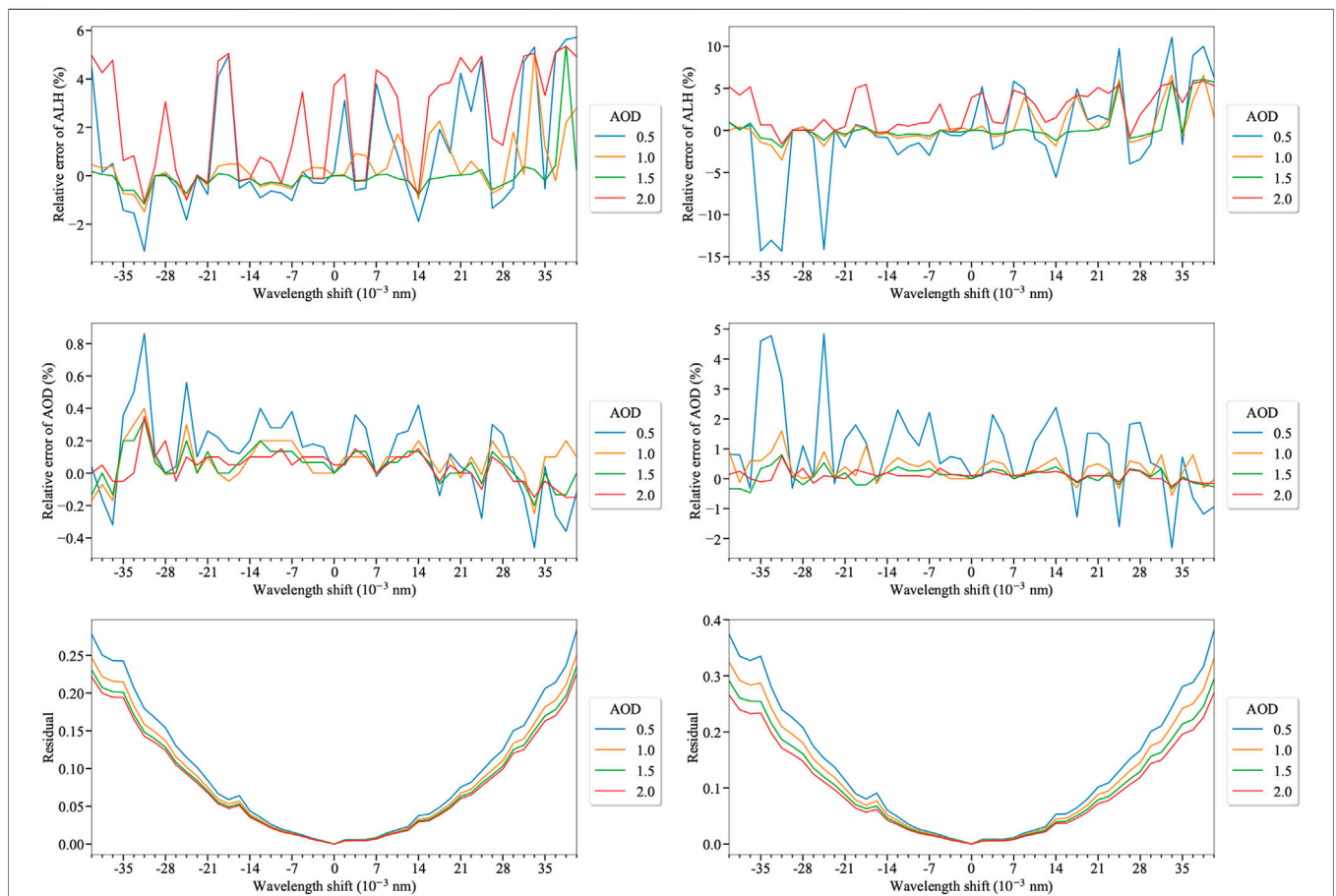
the wavelength shift increases, as compared to the AOD retrieval.

Based on the synthetic analysis, uncertainties in surface albedo and wavelength calibration could cause significant effects on the retrieval output and should not be neglected when dealing with retrievals from real measurements.

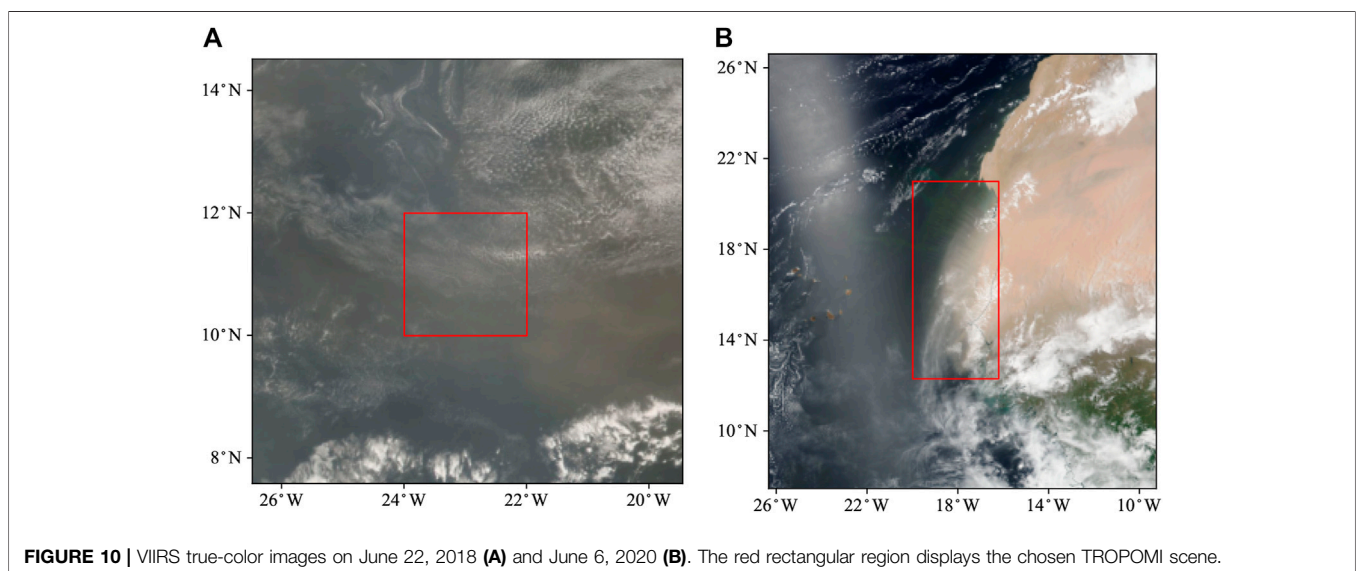
## 4 APPLICATION TO REAL DATA

In this section, we present the retrieval results using the real TROPOMI measurements. We considered two scenes on June 22, 2018 and June 6, 2020, respectively. **Figure 10** displays the corresponding true-color images from the Visible Infrared Imaging Radiometer (VIIRS) on the Suomi National Polar-orbiting Partnership (Suomi NPP) satellite. The red rectangular region indicates the chosen TROPOMI scene. The first TROPOMI scene observed a part of Atlantic ocean near West Africa with latitudes between 10.0 and 12.0°N and longitudes between 22.0 and 24.0°W. The second TROPOMI scene detected a desert dust aerosol case over the Sahara with latitudes between 12.0 and 21.0°N and longitudes between 16.0 and 20.0°W.





**FIGURE 9 |** The same as in **Figure 8** but with true values of a fixed ALH (~3.5 km) and varying AOD.



**FIGURE 10 |** VIIRS true-color images on June 22, 2018 (**A**) and June 6, 2020 (**B**). The red rectangular region displays the chosen TROPOMI scene.

**Table 1** describes the TROPOMI O<sub>2</sub> A-band measurements and the input parameters for retrieval. The aerosol models “MODABS” and “DUST” from Set II were

selected for the retrieval of the two scenes, respectively. The cloud parameters were taken from the operational TROPOMI cloud products (OCRA/ROCINN) whose

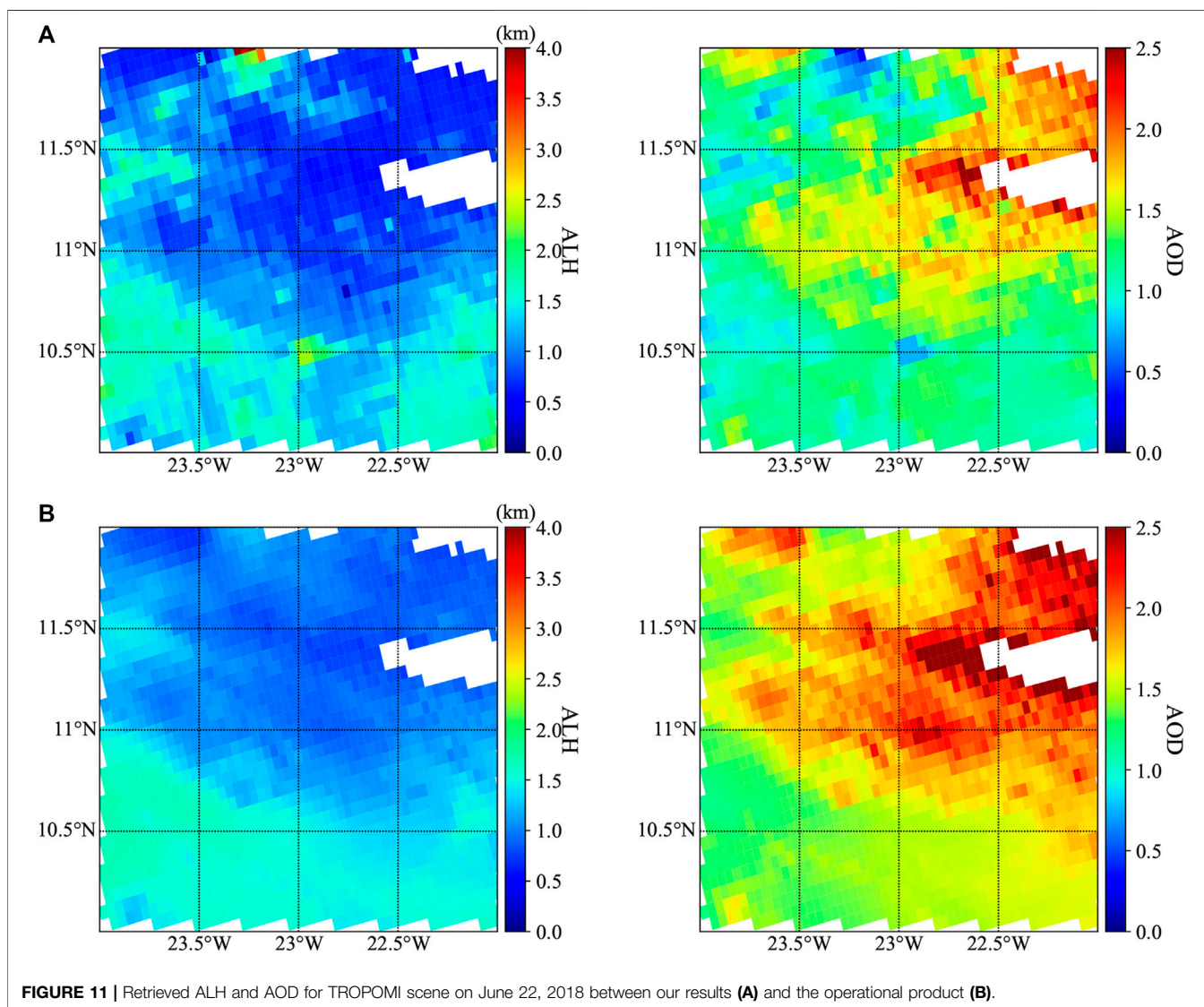
**TABLE 1** | Spectral characteristics of TROPOMI O<sub>2</sub> A-band measurements and the main input parameters for retrieval.

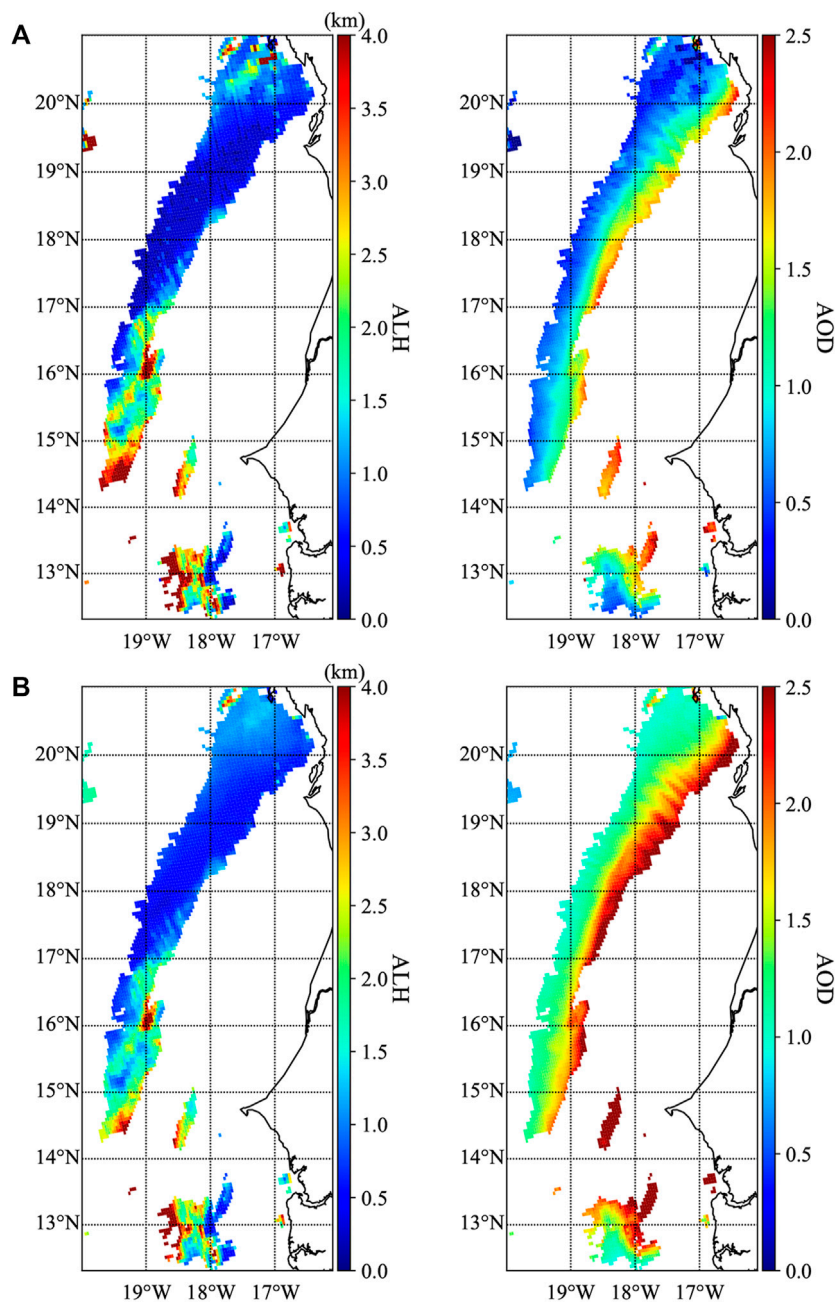
Parameter	Description
Band ID	6
Spectral range	758–771 nm
Spectral sampling	0.126 nm
Aerosol model	"MODABS" and "DUST" (Set II)
Surface albedo	GE_LER
Cloud parameters	OCRA/ROCINN

retrieval algorithms were described by Loyola et al. (2018). Inaccuracy in surface properties could play a crucial role in retrieval, as discussed in Section 3.2. Instead of fitting it as an additional parameter in the state vector or using LER climatologies, the surface albedo was taken from the retrieved GE\_LER product. The *a priori* state  $x_a$  was updated with the retrieval from the previous processed pixel.

We removed pixels with a cloud fraction greater than 0.15 that ensures a sufficient number of valid retrievals without significant cloud contamination. The pixels with the TROPOMI UV aerosol index below 0.0 were not processed. The maximum SZA for processing was 75°.

The inversion calculation carried out by the computer code (yet to be optimized) typically converges in less than five iterations. This translates to a 2–3 min processing time of one pixel on an up-to-date desktop. The retrieval results of ALH and AOD for the two selected TROPOMI scenes are plotted in Figures 11, 12, respectively, which seem to capture the spatial patterns seen from Figure 10 under different aerosol loading scenarios. The retrieved ALH on June 22, 2018 looks higher in the south, and the lowest values of ALH can be found in the northeast where the highest values of AOD are located. Due to heavy cloud contamination on June 6, 2020, a small number of valid pixels were processed. Nevertheless, the spatial distribution of desert dust aerosols is well described by the retrieved aerosol





**FIGURE 12 |** Retrieved ALH and AOD for TROPOMI scene on June 6, 2020 between our results (A) and the operational product (B).

parameters. For reference, the operational retrieval products processed by KNMI were plotted in the bottom row of **Figures 11, 12**. As compared to the operational products, our retrieved aerosol parameters capture nearly the same spatial pattern and the ALH results are slightly underestimated. However, our retrieved AOD values seem evidently lower than the operational ones. Different surface albedo data and aerosol microphysical properties used in the two retrieval algorithms are possibly the main factors explaining these discrepancies. A comprehensive validation is needed in the future.

**Table 2** lists the propagated retrieval error with taking into account the most important error sources that are figured out in **Section 3**. According to a thorough comparison with the OMI LER for clear sky scenarios (Loyola et al., 2020), an error of 0.01 was added to the original values of surface albedo from GE\_LER, causing a mean bias of 0.130 3 km and 0.091 8 on the retrieved ALH and AOD, respectively. For wavelength calibration, we computed the effect of a shift of 0.007 nm in the nominal wavelength grid for the radiance (with the spectral bin size is 0.10 nm). A mean bias of 0.079 3 km and

**TABLE 2** | Bias of ALH and AOD due to errors in surface albedo and wavelength calibration.

Parameter	Error	ALH bias	AOD bias
Surface albedo	0.01	0.1303 km	0.0918
Wavelength grid	0.007 nm	0.0793 km	0.0030

0.0030 on the retrieved ALH and AOD was achieved. Please note that a wavelength shift of 0.007 nm in the case of TROPOMI is already quite large and used here as a conservative estimate.

**Figure 13** compares the observed and modeled reflectance spectra on June 22, 2018 (top row) and June 6, 2020 (bottom row). The modeled reflectance spectra were simulated with the retrieved aerosol quantities after convergence. We randomly chosen four pixels with converged retrieval runs from each scene. The relative residuals turn out to be higher near 760 nm where the reflectance is rather low. For all pixels, the simulated and observed reflectance spectra reach a good agreement, providing an evidence of an overall good fit.

## 5 CONCLUSION

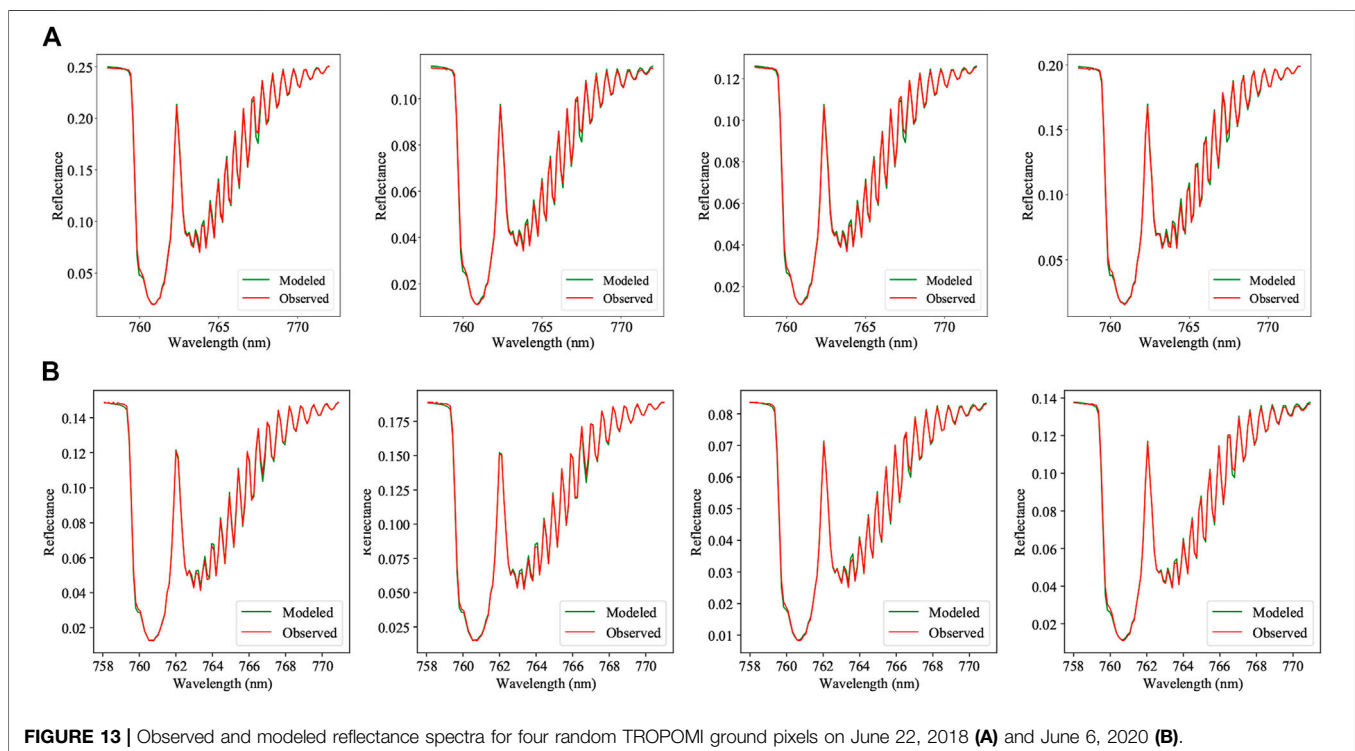
A conventional retrieval algorithm for hyperspectral satellite remote sensing that meets the scientific requirements should enable efficient radiative transfer calculations and reliable inversion computation. We have developed an aerosol retrieval algorithm for estimating aerosol parameters (ALH and AOD) from the O<sub>2</sub> A-band of TROPOMI

onboard the S5P satellite. The key objective of this study was to investigate the impact of forward and instrument model parameters on the retrieval result. The aerosol model for microphysical properties, solar/viewing geometry, surface properties, wavelength calibration have been taken into account.

Aerosol models play an important role in accurately describing aerosol microphysical properties under various measurement conditions. The forward and retrieval simulations using aerosol models from three sets have been compared for the same aerosol type. An inaccurate aerosol model could have a moderate loss of accuracy of retrieved aerosol parameters. Choosing an appropriate aerosol model would be useful in the operational data processing. However, this is not an easy job since for a given measurement it is likely that not just one aerosol model delivers the good fit. Rao et al. (2021) have developed an optimized model selection scheme based on the Bayesian approach and are currently validating its applicability to the real TROPOMI measurements.

As expected, an error of the surface albedo can contribute largely to the accuracy of aerosol retrievals, particularly if the surface albedo is around the critical surface albedo. We suggest that the GE\_LER product can be employed instead of LER climatologies based on low-resolution measurements or fitting the surface albedo simultaneously.

The solar and viewing zenith angles represent the solar and viewing geometry and their accuracy are important to the retrieval accuracy as well. An enhancement in the retrieval algorithm is needed particularly when dealing with satellite measurements at higher SZAs. Another instrument parameter worthy of attention is the wavelength uncertainty. The

**FIGURE 13** | Observed and modeled reflectance spectra for four random TROPOMI ground pixels on June 22, 2018 (A) and June 6, 2020 (B).

wavelength shift can potentially deteriorate the quality of the retrieved parameters (particularly of ALH), although it may have a minor impact on the fit residuals.

Retrievals using real TROPOMI O<sub>2</sub> A-band data recorded on June 22, 2018 and June 6, 2020 have been performed. The retrieved aerosol parameters resemble both aerosol loading scenarios identified in the VIIRS images and the simulated spectra well approximate the observed ones, that have proved the application feasibility of the algorithm itself. Inaccurate surface albedo was supposed to be the most important error sources in practice and reliable measurements of surface albedo are required.

The future work will focus on a comprehensive global/regional validation with other satellite-based (e.g., CALIPO) and ground-based measurements. For efficiency purposes, the development of a retrieval framework using machine learning techniques is ongoing.

## DATA AVAILABILITY STATEMENT

The raw data supporting the conclusion of this article will be made available by the authors, without undue reservation.

## REFERENCES

- Chimot, J., Veeffkind, J. P., Vlemmix, T., de Haan, J. F., Amiridis, V., Proestakis, E., et al. (2017). 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. doi:10.5194/amt-10-783-2017
- Chimot, J., Veeffkind, J. P., Vlemmix, T., and Levelt, P. F. (2018). Spatial Distribution Analysis of the OMI Aerosol Layer Height: a Pixel-By-Pixel Comparison to CALIOP Observations. *Atmos. Meas. Tech.* 11, 2257–2277. doi:10.5194/amt-11-2257-2018
- Chin, M., Ginoux, P., Kinne, S., Torres, O., Holben, B. N., Duncan, B. N., et al. (2002). Tropospheric Aerosol Optical Thickness from the GOCART Model and Comparisons with Satellite and Sun Photometer Measurements. *J. Atmos. Sci.* 59, 461–483. doi:10.1175/1520-0469(2002)059<0461:TAOTFT>2.0.CO;2
- Choi, W., Lee, H., Kim, J., Ryu, J.-Y., Park, S. S., Park, J., et al. (2019). Effects of Spatiotemporal O<sub>4</sub> Column Densities and Temperature-dependent O<sub>4</sub> Absorption Cross-Section on an Aerosol Effective Height Retrieval Algorithm Using the O<sub>4</sub> Air Mass Factor from the Ozone Monitoring Instrument. *Remote Sensing Environ.* 229, 223–233. doi:10.1016/j.rse.2019.05.001
- Corradini, S., and Cervino, M. (2006). Aerosol Extinction Coefficient Profile Retrieval in the Oxygen A-Band Considering Multiple Scattering Atmosphere. Test Case: SCIAMACHY Nadir Simulated Measurements. *J. Quantitative Spectrosc. Radiative Transfer* 97, 354–380. doi:10.1016/j.jqsrt.2005.05.061
- Ding, S., Wang, J., and Xu, X. (2016). Polarimetric Remote Sensing in Oxygen A and B Bands: Sensitivity Study and Information Content Analysis for Vertical Profile of Aerosols. *Atmos. Meas. Tech.* 9, 2077–2092. doi:10.5194/amt-9-2077-2016
- Doicu, A., and Trautmann, T. (2009a). Discrete-ordinate Method with Matrix Exponential for a Pseudo-spherical Atmosphere: Scalar Case. *J. Quantitative Spectrosc. Radiative Transfer* 110, 146–158. doi:10.1016/j.jqsrt.2008.09.014
- Doicu, A., and Trautmann, T. (2009b). Discrete-ordinate Method with Matrix Exponential for a Pseudo-spherical Atmosphere: Vector Case. *J. Quantitative Spectrosc. Radiative Transfer* 110, 159–172. doi:10.1016/j.jqsrt.2008.09.013
- Efremenko, D., Doicu, A., Loyola, D., and Trautmann, T. (2013). Acceleration Techniques for the Discrete Ordinate Method. *J. Quantitative Spectrosc. Radiative Transfer* 114, 73–81. doi:10.1016/j.jqsrt.2012.08.014
- Efremenko, D. S., Loyola R., D. G., Hedelt, P., and Spurr, R. J. D. (2017). Volcanic SO<sub>2</sub> Plume Height Retrieval from UV Sensors Using a Full-Physics Inverse Learning Machine Algorithm. *Int. J. Remote Sensing* 38, 1–27. doi:10.1080/01431161.2017.1348644
- Frankenberg, C., Hasekamp, O., O'Dell, C., Sanghavi, S., Butz, A., and Worden, J. (2012). Aerosol Information Content Analysis of Multi-Angle High Spectral Resolution Measurements and its Benefit for High Accuracy Greenhouse Gas Retrievals. *Atmos. Meas. Tech.* 5, 1809–1821. doi:10.5194/amt-5-1809-2012
- Goody, R., West, R., Chen, L., and Crisp, D. (1989). The Correlated-K Method for Radiation Calculations in Nonhomogeneous Atmospheres. *J. Quant. Spectrosc. Radiat. Transf.* 42, 437–651. doi:10.1016/0022-4073(89)90044-7
- Hedelt, P., Efremenko, D. S., Loyola, D. G., Spurr, R., and Clarisse, L. (2019). Sulfur Dioxide Layer Height Retrieval from Sentinel-5 Precursor/TROPOMI Using FP\_ILM. *Atmos. Meas. Tech.* 12, 5503–5517. doi:10.5194/amt-12-5503-2019
- Hess, M., Koepke, P., and Schult, I. (1998). Optical Properties of Aerosols and Clouds: The Software Package OPAC. *Bull. Amer. Meteorol. Soc.* 79, 831–844. doi:10.1175/1520-0477(1998)079<0831:opoac>2.0.co;2
- Hollstein, A., and Fischer, J. (2014). Retrieving Aerosol Height from the Oxygen a Band: a Fast Forward Operator and Sensitivity Study Concerning Spectral Resolution, Instrumental Noise, and Surface Inhomogeneity. *Atmos. Meas. Tech.* 7, 1429–1441. doi:10.5194/amt-7-1429-2014
- Kleipool, Q., Ludewig, A., Babić, L., Bartstra, R., Braak, R., Dierssen, W., et al. (2018). Pre-launch Calibration Results of the TROPOMI Payload On-Board the Sentinel-5 Precursor Satellite. *Atmos. Meas. Tech.* 11, 6439–6479. doi:10.5194/amt-11-6439-2018
- Kokhanovsky, A. A., Rozanov, V. V., Naus, T., Reudenbach, C., Daniel, J. S., Miller, H. L., et al. (2006a). The Semianalytical Cloud Retrieval Algorithm for SCIAMACHY I. The Validation. *Atmos. Chem. Phys.* 6, 1905–1911. doi:10.5194/acp-6-1905-2006
- Kokhanovsky, A. A., and Rozanov, V. V. (2010). The Determination of Dust Cloud Altitudes from a Satellite Using Hyperspectral Measurements in the Gaseous Absorption Band. *Int. J. Remote Sensing* 31, 2729–2744. doi:10.1080/01431160903085644
- Kokhanovsky, A. A., von Hoyningen-Huene, W., Rozanov, V. V., Noël, S., Gerilowski, K., Bovensmann, H., et al. (2006b). The Semianalytical Cloud Retrieval Algorithm for SCIAMACHY II. The Application to MERIS and SCIAMACHY Data. *Atmos. Chem. Phys.* 6, 4129–4136. doi:10.5194/acp-6-4129-2006
- Koppers, G. A. A., and Murtagh, D. P. (1997). “Retrieval of Height Resolved Aerosol Optical Thickness in the Atmospheric Band,” in *Radiative Transfer in*

## AUTHOR CONTRIBUTIONS

Conceptualization, JX and LR; methodology, JX and AD; software and visualization, LR; investigation, DE and DL; data curation, LR; writing—original draft preparation, LR and JX; writing—review and editing, DE, DL, and AD; supervision, JX and AD; project administration, DE and DL.

## FUNDING

This research was supported by the DLR programmatic (Nachwuchsgruppe “Retrieval der nächsten Generation”, 2472469) and the CAS “Pioneering Initiative Talents Program” under Grant E1RC2WB2. The work of LR was partly funded by the Chinese Scholarship Council.

## ACKNOWLEDGMENTS

The authors are grateful to the DLR’s UPAS team for processing the cloud and GE\_LER products from TROPOMI.

- the Absorption Bands of Oxygen: Studies of Their Significance in Ozone Chemistry and Potential for Aerosol Remote Sensing* (Stockholm, Sweden: Stockholm University).
- Lelli, L., Kokhanovsky, A. A., Rozanov, V. V., Vountas, M., and Burrows, J. P. (2014). Linear Trends in Cloud Top Height from Passive Observations in the Oxygen A-Band. *Atmos. Chem. Phys.* 14, 5679–5692. doi:10.5194/acp-14-5679-2014
- Levy, R. C., Remer, L. A., and Dubovik, O. (2007a). Global Aerosol Optical Properties and Application to Moderate Resolution Imaging Spectroradiometer Aerosol Retrieval over Land. *J. Geophys. Res.* 112. doi:10.1029/2006JD007815
- Levy, R. C., Remer, L. A., Mattoo, S., Vermote, E. F., and Kaufman, Y. J. (2007b). Second-generation Operational Algorithm: Retrieval of Aerosol Properties over Land from Inversion of Moderate Resolution Imaging Spectroradiometer Spectral Reflectance. *J. Geophys. Res. Atmos.* 112. doi:10.1029/2006JD007811
- Li, Z., Zhao, X., Kahn, R., Mishchenko, M., Remer, L., Lee, K.-H., et al. (2009). Uncertainties in Satellite Remote Sensing of Aerosols and Impact on Monitoring its Long-Term Trend: a Review and Perspective. *Ann. Geophys.* 27, 2755–2770. doi:10.5194/angeo-27-2755-2009
- Loyola, D. G., Gimeno García, S., Lutz, R., Argyrouli, A., Romahn, F., Spurr, R. J. D., et al. (2018). The Operational Cloud Retrieval Algorithms from TROPOMI on Board Sentinel-5 Precursor. *Atmos. Meas. Tech.* 11, 409–427. doi:10.5194/amt-11-409-2018
- Loyola, D. G., Xu, J., Heue, K.-P., and Zimmer, W. (2020). Applying FP\_ILM to the Retrieval of Geometry-dependent Effective Lambertian Equivalent Reflectivity (GE\_LER) Daily Maps from UVN Satellite Measurements. *Atmos. Meas. Tech.* 13, 985–999. doi:10.5194/amt-13-985-2020
- Ludewig, A., Kleipool, Q., Bartstra, R., Landzaat, R., Leloux, J., Loots, E., et al. (2020). In-flight Calibration Results of the TROPOMI Payload on Board the Sentinel-5 Precursor Satellite. *Atmos. Meas. Tech.* 13, 3561–3580. doi:10.5194/amt-13-3561-2020
- Molina García, V., Sasi, S., Efremenko, D. S., Doicu, A., and Loyola, D. (2018a). Linearized Radiative Transfer Models for Retrieval of Cloud Parameters from EPIC/DSCOVR Measurements. *J. Quantitative Spectrosc. Radiative Transfer* 213, 241–251. doi:10.1016/j.jqsrt.2018.03.008
- Molina García, V., Sasi, S., Efremenko, D. S., Doicu, A., and Loyola, D. (2018b). Radiative Transfer Models for Retrieval of Cloud Parameters from EPIC/DSCOVR Measurements. *J. Quantitative Spectrosc. Radiative Transfer* 213, 228–240. doi:10.1016/j.jqsrt.2018.03.014
- Nanda, S., de Graaf, M., Sneep, M., de Haan, J. F., Stammes, P., Sanders, A. F. J., et al. (2018a). Error Sources in the Retrieval of Aerosol Information over Bright Surfaces from Satellite Measurements in the Oxygen A Band. *Atmos. Meas. Tech.* 11, 161–175. doi:10.5194/amt-11-161-2018
- Nanda, S., Veeffkind, J. P., de Graaf, M., Sneep, M., Stammes, P., de Haan, J. F., et al. (2018b). A Weighted Least Squares Approach to Retrieve Aerosol Layer Height over Bright Surfaces Applied to GOME-2 Measurements of the Oxygen A Band for forest Fire Cases over Europe. *Atmos. Meas. Tech.* 11, 3263–3280. doi:10.5194/amt-11-3263-2018
- Natraj, V., Jiang, X., Shia, R.-I., Huang, X., Margolis, J. S., and Yung, Y. L. (2005). Application of Principal Component Analysis to High Spectral Resolution Radiative Transfer: A Case Study of the Band. *J. Quantitative Spectrosc. Radiative Transfer* 95, 539–556. doi:10.1016/j.jqsrt.2004.12.024
- Natraj, V., Shia, R.-L., and Yung, Y. L. (2010). On the Use of Principal Component Analysis to Speed up Radiative Transfer Calculations. *J. Quantitative Spectrosc. Radiative Transfer* 111, 810–816. doi:10.1016/j.jqsrt.2009.11.004
- Park, S. S., Kim, J., Lee, H., Torres, O., Lee, K.-M., and Lee, S. D. (2016). Utilization of O4 Slant Column Density to Derive Aerosol Layer Height from a Space-Borne UV-Visible Hyperspectral Sensor: Sensitivity and Case Study. *Atmos. Chem. Phys.* 16, 1987–2006. doi:10.5194/acp-16-1987-2016
- Rao, L., Xu, J., Efremenko, D. S., Loyola, D. G., and Doicu, A. (2021). Optimization of Aerosol Model Selection for TROPOMI/S5P. *Remote Sens* 13. doi:10.3390/rs13132489
- Rodgers, C. (2000). *Inverse Methods For Atmospheric Sounding: Theory And Practise*. Singapore: World Scientific.
- Rodriguez, D. G. L., Thomas, W., Livschitz, Y., Ruppert, T., Albert, P., and Hollmann, R. (2007). Cloud Properties Derived from GOME/ERS-2 Backscatter Data for Trace Gas Retrieval. *IEEE Trans. Geosci. Remote Sensing* 45, 2747–2758. doi:10.1109/TGRS.2007.901043
- Sanders, A. F. J., and de Haan, J. F. (2013). Retrieval of Aerosol Parameters from the Oxygen A Band in the Presence of Chlorophyll Fluorescence. *Atmos. Meas. Tech.* 6, 2725–2740. doi:10.5194/amt-6-2725-2013
- Sanders, A. F. J., de Haan, J. F., Sneep, M., Apituley, A., Stammes, P., Vieitez, M. O., et al. (2015). Evaluation of the Operational Aerosol Layer Height Retrieval Algorithm for Sentinel-5 Precursor: Application to O2 A Band Observations from GOME-2A. *Atmos. Meas. Tech.* 8, 4947–4977. doi:10.5194/amt-8-4947-2015
- Sanghavi, S., Martonchik, J. V., Landgraf, J., and Platt, U. (2012). Retrieval of the Optical Depth and Vertical Distribution of Particulate Scatterers in the Atmosphere Using O2 A- and B-Band SCIAMACHY Observations over Kanpur: a Case Study. *Atmos. Meas. Tech.* 5, 1099–1119. doi:10.5194/amt-5-1099-2012
- Sasi, S., Natraj, V., Molina García, V., Efremenko, D. S., Loyola, D., and Doicu, A. (2020a). Model Selection in Atmospheric Remote Sensing with an Application to Aerosol Retrieval from DSCOVR/EPIC, Part 1: Theory. *Remote Sens* 12, 3724. doi:10.3390/rs12223724
- Sasi, S., Natraj, V., Molina García, V., Efremenko, D. S., Loyola, D., and Doicu, A. (2020b). Model Selection in Atmospheric Remote Sensing with Application to Aerosol Retrieval from DSCOVR/EPIC. Part 2: Numerical Analysis. *Remote Sens* 12, 3656. doi:10.3390/rs12213656
- Seidel, F. C., and Popp, C. (2012). Critical Surface Albedo and its Implications to Aerosol Remote Sensing. *Atmos. Meas. Tech.* 5, 1653–1665. doi:10.5194/amt-5-1653-2012
- Spurr, R. (2008). “LIDORT and VLIDORT. Linearized Pseudo-spherical Scalar and Vector Discrete Ordinate Radiative Transfer Models for Use in Remote Sensing Retrieval Problems.” *Light Scattering Reviews*. Editor A. Kokhanovsky (Berlin: Springer-Verlag), 3, 229–271.
- Taylor, M., Kazadzis, S., Amiridis, V., and Kahn, R. A. (2015). Global Aerosol Mixtures and Their Multiyear and Seasonal Characteristics. *Atmos. Environ.* 116, 112–129. doi:10.1016/j.atmosenv.2015.06.029
- Thomas, G. E., Poulsen, C. A., Sayer, A. M., Marsh, S. H., Dean, S. M., Carboni, E., et al. (2009). The GRAPE Aerosol Retrieval Algorithm. *Atmos. Meas. Tech.* 2, 679–701. doi:10.5194/amt-2-679-2009
- Tikhonov, A. (1963). On the Solution of Incorrectly Stated Problems and a Method of Regularization. *Dokl. Acad. Nauk SSSR* 151, 501–504.
- Tilstra, L., Tuinder, O., Wang, P., and Stammes, P. (2019). GOME-2 Absorbing Aerosol Height Algorithm Theoretical Basis Document. Tech. rep. De Bilt, Netherlands: Royal Netherlands Meteorological Institute KNMI.
- Torres, O., Bhartia, P. K., Herman, J. R., Ahmad, Z., and Gleason, J. (1998). Derivation of Aerosol Properties from Satellite Measurements of Backscattered Ultraviolet Radiation: Theoretical Basis. *J. Geophys. Res.* 103, 17099–17110. doi:10.1029/98JD00900
- Veeffkind, J. P., Aben, I., McMullan, K., Förster, H., de Vries, J., Otter, G., et al. (2012). TROPOMI on the ESA Sentinel-5 Precursor: A GMES mission for Global Observations of the Atmospheric Composition for Climate, Air Quality and Ozone Layer Applications. *Remote Sensing Environ.* 120, 70–83. doi:10.1016/j.rse.2011.09.027
- Wang, P., Stammes, P., van der A, R., Pinardi, G., and van Roozendael, M. (2008). FRESCO+: an Improved O2 A-Band Cloud Retrieval Algorithm for Tropospheric Trace Gas Retrievals. *Atmos. Chem. Phys.* 8, 6565–6576. doi:10.5194/acp-8-6565-2008
- Winker, D. M., Vaughan, M. A., Omar, A., Hu, Y., Powell, K. A., Liu, Z., et al. (2009). Overview of the CALIPSO mission and CALIOP Data Processing Algorithms. *J. Atmos. Oceanic Techn.* 26, 2310–2323. doi:10.1175/2009JTECHA1281.1
- Xu, J., Rao, L., Schreier, F., Efremenko, D. S., Doicu, A., and Trautmann, T. (2020). Insight into Construction of Tikhonov-type Regularization for Atmospheric Retrievals. *Atmosphere* 11, 1052. doi:10.3390/atmos11101052
- Xu, J., Schreier, F., Doicu, A., and Trautmann, T. (2016). Assessment of Tikhonov-type Regularization Methods for Solving Atmospheric Inverse Problems. *J. Quantitative Spectrosc. Radiative Transfer* 184, 274–286. doi:10.1016/j.jqsrt.2016.08.003
- Xu, J., Schussler, O., Rodriguez, D. G. L. D., Romahn, F., and Doicu, A. (2017a). A Novel Ozone Profile Shape Retrieval Using Full-Physics Inverse Learning Machine (FP-ILM). *IEEE J. Sel. Top. Appl. Earth Observations Remote Sensing* 10, 5442–5457. doi:10.1109/JSTARS.2017.2740168

- Xu, X., Wang, J., Wang, Y., Zeng, J., Torres, O., Yang, Y., et al. (2017b). Passive Remote Sensing of Altitude and Optical Depth of Dust Plumes Using the Oxygen A and B Bands: First Results from EPIC/DSCOVR at Lagrange-1 point. *Geophys. Res. Lett.* 44, 7544–7554. doi:10.1002/2017GL073939
- Zeng, Z.-C., Chen, S., Natraj, V., Le, T., Xu, F., Merrelli, A., et al. (2020). Constraining the Vertical Distribution of Coastal Dust Aerosol Using OCO-2 O<sub>2</sub> A-Band Measurements. *Remote Sensing Environ.* 236, 111494. doi:10.1016/j.rse.2019.111494

**Conflict of Interest:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

**Publisher's Note:** All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2022 Rao, Xu, Efremenko, Loyola and Doicu. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.





## **B Optimization of Aerosol Model Selection for TROPOMI/S5P**

Rao, L.; Xu, J.; Efremenko, D.S.; Loyola, D.G.; Doicu, A. Optimization of Aerosol Model Selection for TROPOMI/S5P. *Remote Sens.* 2021, 13(13), 2489, doi: 10.3390/rs13132489



Article

# Optimization of Aerosol Model Selection for TROPOMI/S5P

Lanlan Rao <sup>1,2</sup> , Jian Xu <sup>1,\*</sup> , Dmitry S. Efremenko <sup>1</sup> , Diego G. Loyola <sup>1</sup> and Adrian Doicu <sup>1</sup>

<sup>1</sup> Remote Sensing Technology Institute, German Aerospace Center (DLR), 82234 Oberpfaffenhofen, Germany; lanlan.rao@dlr.de (L.R.); dmitry.efremenko@dlr.de (D.S.E.); diego.loyola@dlr.de (D.G.L.); adrian.doicu@dlr.de (A.D.)

<sup>2</sup> Department of Aerospace and Geodesy, Technische Universität München, 80331 Munich, Germany

\* Correspondence: jian.xu@dlr.de; Tel.: +49-8153-28-3353

† Current address: National Space Science Center, Chinese Academy of Sciences, Beijing 100190, China.

**Abstract:** To retrieve aerosol properties from satellite measurements, micro-physical aerosol models have to be assumed. Due to the spatial and temporal inhomogeneity of aerosols, choosing an appropriate aerosol model is an important task. In this paper, we use a Bayesian algorithm that takes into account model uncertainties to retrieve the aerosol optical depth and layer height from synthetic and real TROPOMI O<sub>2</sub>A band measurements. The results show that in case of insufficient information for an appropriate micro-physical model selection, the Bayesian algorithm improves the accuracy of the solution.

**Keywords:** model selection; aerosol retrievals; TROPOMI/S5P



**Citation:** Rao, L.; Xu, J.; Efremenko, D.S.; Loyola, D.G.; Doicu, A. Aerosol Model Selection for TROPOMI/S5P. *Remote Sens.* **2021**, *13*, 2489. <https://doi.org/10.3390/rs13132489>

Academic Editor: Luke Knibbs

Received: 14 May 2021  
Accepted: 22 June 2021  
Published: 25 June 2021

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Aerosols affect the Earth's climate directly by disturbing the Earth's radiation budget and indirectly by altering cloud processes. To better understand the role of aerosols in the Earth's climate, it is important to observe concentrations and properties of aerosols. Satellite sensors provide long-term measurements that can effectively monitor aerosol information on both regional and global scales.

The information on the aerosol optical depth can be retrieved from the data provided by satellite sensors, such as the Advanced Very High Resolution Radiometer (AVHRR) [1], the Moderate Resolution Imaging Spectroradiometer (MODIS) [2], the Visible Infrared Imaging Radiometer (VIIRS) [3], and the Advanced Himawari Imager (AHI) [4], helping to understand the temporal and spatial distribution characteristics of atmospheric aerosols. Aerosol height information can be retrieved from (i) multi-angle instruments, e.g., the Multi-angle Imaging Spectroradiometer (MISR) [5], and the Advanced Along-Track Scanning Radiometer (AATSR) [6]; (ii) polarization measurements, e.g., the Polarization and Directionality of the Earth's Reflectances (POLDER) [7]; and (iii) measurements in the oxygen absorption band, e.g., the TROPospheric Monitoring Instrument (TROPOMI) [8]. A combination of multi-angle and polarization observations [9] can also provide information of micro-physical parameters such as particle size distribution and refractive index.

However, the information that can be retrieved from space is quite limited. To retrieve the aerosol parameters, aerosol models characterizing the micro-physical properties have to be assumed. Aerosol properties exhibit high spatial inhomogeneity because of various origins and complex processes during transportation in the atmosphere. Aerosol particles are originated from both natural processes (such as wind-blown desert dust and sea salt, wild forest fire, and volcano eruption) and anthropogenic activities (such as industrial activities, artificial vegetation fire, and fossil fuel combustion). The selection of a suitable aerosol model in the retrieval algorithm relies on the knowledge of emission sources.

There are several databases and sets of aerosol models portraying the aerosol micro-physical properties on a global scale. The Optical Properties of Aerosols and Clouds (OPAC) database [10] describes the size distribution and spectral refractive index of 10 aerosol

components under different humidities. These components can form various aerosol types through internal mixture. The dark-target algorithm of the Moderate-Resolution Imaging Spectroradiometer (MODIS) characterizes a set of aerosol models and provides global distributions of aerosol types for different seasons based on a cluster analysis of the AERONET climatology [2,11]. The OMI near-UV (OMAERUV) algorithm and the multi-wavelength algorithm (OMAERO) consider several major aerosol types which are split into different aerosol models. The selection of an aerosol model is based on spectral and geographic considerations [12]. A chemical transport model, such as the Goddard Chemistry Aerosol Radiation and Transport (GOCART) model also supplies distributions of different aerosol types [13,14]. Besides, a number of studies coping with classification of aerosol types based on satellite remote sensing were carried out, see, e.g., in [15–19].

In standard retrieval algorithms, an aerosol model is chosen from a set of candidate models, and the retrieval is performed as if the selected model reflects the real scenario. In general, model selection is not a trivial task because for a given measurement, several models may fit the data equally well. The Bayesian approach and, in particular, the Bayesian model selection and model averaging Hoeting et al. [20], is a statistical method using measurement data to select the best fitting models from a set of candidate models without any prior seasonal or geographical information. The Bayesian method provides a posteriori probability densities for given models, also known as model evidences. In the Bayesian model selection, we select the model with the highest evidence, while in the Bayesian model averaging, we combine the retrieval results corresponding to different candidate models weighted by their evidences. Määttä et al. [21] introduced the Bayesian approach into the aerosol model selection of the OMAERO algorithm, Kauppi et al. [22] applied the Bayesian approach to real data of OMI, while Sasi et al. [23,24] applied the Bayesian approach to EPIC (Earth Polychromatic Imaging Camera) [25] synthetic measurements.

In this paper, for the first time, we use the Bayesian approach to jointly retrieve the aerosol optical depth and aerosol layer height from TROPOMI/S5P (Sentinel-5 Precursor) [26] measurements in the O<sub>2</sub>A band (758–771 nm). TROPOMI is a hyperspectral instrument on board the Copernicus Sentinel-5 Precursor satellite launched on 13 October 2017, measuring the solar radiance backscattered by atmosphere and Earth's surface in the ultraviolet (UV), visible (VIS), near-infrared (NIR), and shortwave infrared (SWIR) spectral ranges. The aerosol parameters are retrieved using NIR measurements with a spectral resolution of ~0.45 nm. As the first atmospheric monitoring mission within the Copernicus program, TROPOMI has a very high spatial resolution of 3.5 × 7 km<sup>2</sup> (3.5 × 5.5 km<sup>2</sup> since 6 August 2019), as compared with its predecessors. In particular, the spectra in O<sub>2</sub>A band (758–771 nm) provides a way to retrieve the aerosol height information. The physical principle of aerosol height detection in O<sub>2</sub>A band lies on the fact that the aerosol layer attenuate the reflection of solar radiance by the lower atmosphere at high oxygen absorption wavelengths. This attenuation decreases as the decline of oxygen absorption coefficient. To our best knowledge, currently no satellite passive sensor except TROPOMI can provide official product of aerosol height information.

The paper is organized as follows. In Section 2, we review the Bayesian model selection approach and discuss its application to aerosol retrievals. Section 3 describes the sets of aerosol models used in our numerical analysis. The accuracy of the Bayesian model selection approach is analyzed in Section 4 for synthetic measurements and in Section 5 for real data over a wild fire scene in South Africa.

## 2. Methodology

We have developed a retrieval algorithm dedicated to satellite remote sensing of aerosol and cloud parameters. The physics-based retrieval algorithm comprises a forward model calculating radiative transfer of electromagnetic radiation through a planetary atmosphere and an inversion module solving a nonlinear minimization problem. In the forward model, the radiative transfer calculation depends on the discrete ordinate method with

matrix exponential. To speed up the computation in the oxygen absorption band from sensors (e.g., TROPOMI) with very high spectral resolution, we have implemented acceleration techniques like the telescoping technique [27,28], the false discrete ordinate approach, the correlated  $k$ -distribution method [29], and the principal components analysis [30,31]. The inversion is performed by the means of Tikhonov regularization with optimal strategies for constructing the regularization parameter and matrix [32,33]. For further details about the retrieval algorithm and its forward model, we refer to the works in [34–36].

In this study, the aerosol optical depth  $\tau$  and the layer height  $h$  are retrieved in the O<sub>2</sub>A band (758–771 nm). The retrieval algorithm can deal with four types of aerosol profiles: Gaussian, exponential decay, elevated box, and a combination of exponential decay and ground box. To simplify the analysis, the aerosol layer is assumed to be homogeneous, spreading evenly from near surface to the top aerosol layer height  $h$ . Considering  $N_m$  aerosol models, the retrieval of the state vector  $\mathbf{x} = [\tau, h]$  is an inverse problem relying on the solution of the nonlinear equation

$$\mathbf{y}^\delta = \mathbf{F}_m(\mathbf{x}) + \delta_m, \quad (1)$$

where  $\mathbf{y}^\delta$  is the measurement vector,  $\mathbf{F}_m(\mathbf{x})$  is the forward model corresponding to the aerosol model  $m$  with  $m = 1, \dots, N_m$ ,  $\delta_m = \delta_{\text{mes}} + \delta_{\text{aerm}}$  the total data error vector,  $\delta_{\text{mes}}$  the measurement error vector, and  $\delta_{\text{aerm}}$  the aerosol model error vector, i.e., the error due to an inadequate aerosol model. In our analysis,  $\mathbf{F}_m(\mathbf{x})$  is the vector of the log of the simulated radiances corresponding to aerosol model  $m$ , i.e.,  $[\mathbf{F}_m(\mathbf{x})]_i = \ln I_m(\lambda_i, \mathbf{x})$ , where  $\{\lambda_i\}_{i=1}^M$  is a set of  $M$  wavelengths in the considered spectral domain.

The data model (1) is transformed into a model with white noise by using the prewhitening technique. The procedure is as follows. Assuming that

1.  $\delta_{\text{mes}}$  is a Gaussian random vector with zero mean and covariance matrix  $\mathbf{C}_{\text{mes}} = \sigma_{\text{mes}}^2 \mathbf{C}_{\text{mes}}$ , where  $\sigma_{\text{mes}}^2$  is the measurement error variance and  $\mathbf{C}_{\text{mes}}$  a normalized measurement error covariance matrix;
2.  $\delta_{\text{aerm}}$  is a Gaussian random vector with zero mean and covariance matrix  $\mathbf{C}_{\text{aerm}} = \sigma_{\text{aerm}}^2 \mathbf{I}_M$ , where  $\sigma_{\text{aerm}}^2$  is the aerosol model error variance and  $\mathbf{I}_M$  the identity matrix; and
3.  $\delta_{\text{mes}}$  and  $\delta_{\text{aerm}}$  are independent random vectors,

we deduce that  $\delta_m$  is also a Gaussian random vector with zero mean and covariance matrix  $\mathbf{C}_{\delta m} = \mathbf{C}_{\text{mes}} + \mathbf{C}_{\text{aerm}} = \sigma_m^2 \mathbf{C}_{\delta m}$ , where  $\sigma_m^2 = \sigma_{\text{mes}}^2 + \sigma_{\text{aerm}}^2$  is the data error variance and  $\mathbf{C}_{\delta m} = w_{\text{mes}m} \mathbf{C}_{\text{mes}} + (1 - w_{\text{mes}m}) \mathbf{I}_M$  with  $w_{\text{mes}m} = \sigma_{\text{mes}}^2 / \sigma_m^2$ , a normalized data error covariance matrix. In this context, the scaled data model reads as

$$\bar{\mathbf{y}}^\delta = \bar{\mathbf{F}}_m(\mathbf{x}) + \bar{\delta}_m, \quad (2)$$

where  $\bar{\mathbf{y}}^\delta = \mathbf{P} \mathbf{y}^\delta$ ,  $\bar{\mathbf{F}}_m(\mathbf{x}) = \mathbf{P} \mathbf{F}_m(\mathbf{x})$ ,  $\bar{\delta}_m = \mathbf{P} \delta_m$ , and  $\mathbf{P} = \mathbf{C}_{\delta m}^{-1/2}$  is a scaling matrix. As  $\bar{\mathbf{C}}_{\delta m} = \mathcal{E}\{\bar{\delta}_m \bar{\delta}_m^T\} = \sigma_m^2 \mathbf{I}_M$ , it is readily seen that  $\bar{\delta}_m \sim \mathcal{N}(\mathbf{0}, \bar{\mathbf{C}}_{\delta m} = \sigma_m^2 \mathbf{I}_M)$ , where the notation  $\mathcal{N}(\mathbf{x}_{\text{mean}}, \mathbf{C}_x)$  stands for a normal distribution with mean  $\mathbf{x}_{\text{mean}}$  and covariance matrix  $\mathbf{C}_x$ . In a stochastic setting, we assume that  $\mathbf{x} \sim \mathcal{N}(\mathbf{x}_a, \mathbf{C}_x)$ , where  $\mathbf{x}_a$  is the a priori state vector, the best beforehand estimate of the solution,  $\mathbf{C}_x = \sigma_x^2 \mathbf{C}_x$  is the a priori covariance matrix, and  $\sigma_x^2$  the a priori state variance. Defining the regularization matrix  $\mathbf{L}$  and the regularization parameter  $\alpha$  through the relations  $\mathbf{C}_x^{-1} = \mathbf{L}^T \mathbf{L}$  and  $\alpha = \sigma_m^2 / \sigma_x^2$ , respectively, we express the a priori covariance matrix as  $\mathbf{C}_x = \sigma_m^2 (\alpha \mathbf{L}^T \mathbf{L})^{-1}$ .

The scaled nonlinear Equation (2) is solved by means of a Bayesian approach. The key quantity in this approach is the a posteriori density  $p(\mathbf{x} | \bar{\mathbf{y}}^\delta, m)$ , which represents the conditional probability density of  $\mathbf{x}$  given the data  $\bar{\mathbf{y}}^\delta$  and the aerosol model  $m$ . According to Bayes' theorem, the a posteriori density is given by

$$p(\mathbf{x} | \bar{\mathbf{y}}^\delta, m) = \frac{p(\bar{\mathbf{y}}^\delta | \mathbf{x}, m) p(\mathbf{x} | m)}{p(\bar{\mathbf{y}}^\delta | m)},$$

where  $p(\mathbf{x} | m)$  is the a priori density, i.e., the conditional probability density of  $\mathbf{x}$  given the aerosol model  $m$  before performing the measurement  $\bar{\mathbf{y}}^\delta$ ,  $p(\bar{\mathbf{y}}^\delta | \mathbf{x}, m)$  the likelihood density, i.e., the conditional probability density of  $\bar{\mathbf{y}}^\delta$  given the state  $\mathbf{x}$  and the aerosol model  $m$ , and

$$p(\bar{\mathbf{y}}^\delta | m) = \int p(\mathbf{x}, \bar{\mathbf{y}}^\delta | m) d\mathbf{x} = \int p(\bar{\mathbf{y}}^\delta | \mathbf{x}, m)p(\mathbf{x} | m) d\mathbf{x}, \quad (3)$$

the marginal likelihood density. Although, in the Bayesian parameter estimation, the marginal likelihood density  $p(\bar{\mathbf{y}}^\delta | m)$  plays the role of a normalization constant and is usually ignored, this probability density is of particular importance in the Bayesian model selection. For  $\mathbf{x} \sim N(\mathbf{x}_a, \mathbf{C}_x = \sigma_m^2(\alpha \mathbf{L}^T \mathbf{L})^{-1})$  and  $\bar{\delta}_m \sim N(\mathbf{0}, \bar{\mathbf{C}}_{\delta m} = \sigma_m^2 \mathbf{I}_M)$ , the Bayes' formula gives

$$p(\mathbf{x} | \bar{\mathbf{y}}^\delta, m) \propto \exp\left[-\frac{1}{2}V_\alpha(\mathbf{x} | \bar{\mathbf{y}}^\delta, m)\right], \quad (4)$$

where

$$V_\alpha(\mathbf{x} | \bar{\mathbf{y}}^\delta, m) = \frac{1}{\sigma_m^2} \left[ \|\bar{\mathbf{y}}^\delta - \bar{\mathbf{F}}_m(\mathbf{x})\|^2 + \alpha \|\mathbf{L}(\mathbf{x} - \mathbf{x}_a)\|^2 \right] \quad (5)$$

is the a posteriori potential. Consequently, the maximum a posteriori estimate  $\hat{\mathbf{x}}_{m\alpha}^\delta$  is computed as

$$\hat{\mathbf{x}}_{m\alpha}^\delta = \arg \min_{\mathbf{x}} V_\alpha(\mathbf{x} | \bar{\mathbf{y}}^\delta, m). \quad (6)$$

In a deterministic setting,  $\mathcal{F}_{m\alpha}(\mathbf{x}) = \sigma_m^2 V_\alpha(\mathbf{x} | \bar{\mathbf{y}}^\delta, m)$  is the Tikhonov function for the nonlinear equation  $\bar{\mathbf{y}}^\delta = \bar{\mathbf{F}}_m(\mathbf{x})$  with the penalty term  $\alpha \|\mathbf{L}(\mathbf{x} - \mathbf{x}_a)\|^2$  and the regularization parameter  $\alpha$ . Thus, a regularized solution  $\mathbf{x}_{m\alpha}^\delta$ , which minimizes  $\mathcal{F}_{m\alpha}(\mathbf{x})$ , coincides with the maximum a posteriori estimate, i.e.,  $\mathbf{x}_{m\alpha}^\delta = \hat{\mathbf{x}}_{m\alpha}^\delta$ . The computation of the regularized solution  $\mathbf{x}_{m\alpha}^\delta$  in the framework of the method of Tikhonov regularization requires the knowledge of the optimal value of the regularization parameter  $\hat{\alpha}$ . Because in practice, this is a very challenging task, the nonlinear equation  $\bar{\mathbf{y}}^\delta = \bar{\mathbf{F}}_m(\mathbf{x})$  is solved by means of the iteratively regularized Gauss–Newton method. This method provides an optimal value of the regularization parameter  $\hat{\alpha}$  (i.e., the ratio of the data error variance  $\sigma_m^2$  and the a priori state variance  $\sigma_x^2$ ) and the corresponding regularized solution  $\mathbf{x}_{m\hat{\alpha}}^\delta$ .

For model comparison, the so-called relative evidence of the  $m$ th aerosol model  $p(m | \bar{\mathbf{y}}^\delta)$  plays an important role. In view of Bayes' theorem, this conditional probability density is defined by

$$p(m | \bar{\mathbf{y}}^\delta) = \frac{p(\bar{\mathbf{y}}^\delta | m)p(m)}{p(\bar{\mathbf{y}}^\delta)} = \frac{p(\bar{\mathbf{y}}^\delta | m)}{\sum_{m=1}^{N_m} p(\bar{\mathbf{y}}^\delta | m)}, \quad (7)$$

where the last equality follows from the mean formula  $p(\bar{\mathbf{y}}^\delta) = \sum_{m=1}^{N_m} p(\bar{\mathbf{y}}^\delta | m)p(m)$  and the assumption that all aerosol models are equally like, i.e.,  $p(m) = 1/N_m$ . In terms of  $p(m | \bar{\mathbf{y}}^\delta)$ , the mean a posteriori density reads as

$$p_{\text{mean}}(\mathbf{x} | \bar{\mathbf{y}}^\delta) = \sum_{m=1}^{N_m} p(\mathbf{x} | \bar{\mathbf{y}}^\delta, m)p(m | \bar{\mathbf{y}}^\delta), \quad (8)$$

while the mean and the maximum solution estimates are defined by

$$\hat{\mathbf{x}}_{\text{mean}}^\delta = \sum_{m=1}^{N_m} \mathbf{x}_{m\hat{\alpha}}^\delta p(m | \bar{\mathbf{y}}^\delta), \quad (9)$$

and

$$\hat{\mathbf{x}}_{\text{max}}^\delta = \mathbf{x}_{m^* \hat{\alpha}}^\delta, \quad m^* = \arg \max_m p(m | \bar{\mathbf{y}}^\delta), \quad (10)$$

respectively. In Equation (9), the Bayesian model averaging technique is used to combine the individual solutions weighted by their evidences, while in Equation (8), the aerosol

model with the highest evidence is considered to be the best among all aerosol models involved. Intuitively, we expect that in practice, a linear combination of the retrieved parameters corresponding to different candidate models will better reproduce the real scenario than the retrieved parameters corresponding to an a priori selected model. In a stochastic setting, the relative evidence  $p(m | \bar{\mathbf{y}}^\delta)$  can be computed via Equation (7) by using a linearization of the forward model around the solution and under the assumption that the data error variance  $\sigma_m^2$  is known. In [23,24], estimates for  $\sigma_m^2$  were obtained in the framework of the maximum marginal likelihood estimation [37–39] and the generalized cross-validation method [40,41]. In a deterministic setting,  $p(m | \bar{\mathbf{y}}^\delta)$ , regarded as a merit function characterizing the solution  $\mathbf{x}_{m\hat{\alpha}}^\delta$ , can be defined in terms of the marginal likelihood function or the generalized cross-validation function. In the latter case, the computational formula is

$$p(m | \mathbf{y}^\delta) = \frac{1/v(m)}{\sum_{m=1}^{N_m} 1/v(m)}. \quad (11)$$

where

$$v(m) = \frac{\|\mathbf{r}_{m\hat{\alpha}}^\delta\|^2}{[\text{trace}(\mathbf{I} - \hat{\mathbf{A}}_{m\hat{\alpha}})]^2} \quad (12)$$

is the generalized cross-validation function,  $\mathbf{r}_{m\hat{\alpha}}^\delta = \bar{\mathbf{y}}^\delta - \bar{\mathbf{F}}_m(\mathbf{x}_{m\hat{\alpha}}^\delta)$  the nonlinear residual vector,  $\hat{\mathbf{A}}_{m\hat{\alpha}} = \bar{\mathbf{K}}_{m\hat{\alpha}} \bar{\mathbf{K}}_{m\hat{\alpha}}^\dagger$  the influence matrix,  $\bar{\mathbf{K}}_{m\hat{\alpha}}$  the Jacobian matrix, and  $\bar{\mathbf{K}}_{m\hat{\alpha}}^\dagger$  the generalized inverse at the solution  $\mathbf{x}_{m\hat{\alpha}}^\delta$ .

### 3. Aerosol Models

Two sets of aerosol micro-physical models are used in our numerical analysis. The first set (Set 1) is taken from the MODIS dark-target algorithm [11] and includes the following four aerosol models:

1. non-absorbing (NONABS) aerosols, generated from fossil fuel combustion in urban-industrial areas;
2. moderately absorbing (MODABS) aerosols;
3. absorbing (ABS) aerosols, generated from biomass burning; and
4. desert dust (DUST), originated from desert and transported by wind.

The volume size distribution of each aerosol model is a bimodal log-normal distribution consisting of a fine and a coarse mode. The parameters of the size distribution (median radius, standard deviation, and volume of particles) and the complex refractive index, which depend on the aerosol optical depth, are illustrated in Table 1. The second set (Set 2) is taken from the OMAERO algorithm and includes the following five major aerosol types:

1. weakly absorbing (WA),
2. biomass burning (BB),
3. desert dust (DD),
4. marine (MA), and
5. volcanic (VO).

Each type is split into several aerosol models depending on their optical properties and particle size distribution. The parameters of the size distribution (median radius, standard deviation, and fraction of coarse mode) and the complex refractive index are shown in Table 2.

In the forward model, the scattering characteristics (e.g., the single scattering albedo, the phase function, and the asymmetry parameter) can be computed by the Mie theory in the case of spherical particles, and the null-field method with discrete sources in the case of spheroidal particles with a size parameter smaller than 50. For spheroidal particles with large size parameter, we use a precomputed database as given in [42]. In this study, the aerosol particles are assumed to be spherical for simplicity.

**Table 1.** Micro-physical properties of aerosols models of Set 1. Each model is composed of a fine and coarse mode. The median radius of the volume distribution  $r_v$ , standard deviation  $\sigma$ , the volume of particles  $V_0$ , and complex refractive index  $m$  of each mode are listed in the table.

Model	Mode	$r_v$ ( $\mu\text{m}$ )	$\sigma$	$m = (\text{Re}(m), \text{Im}(m))$	$V_0$ ( $\mu\text{m}^3/\mu\text{m}^2$ )
NONABS	fine	$0.160 + 0.0434\tau$	$0.364 + 0.1529\tau$	$(1.42, 0.004 - 0.0015\tau)$	$0.1718\tau^{0.821}$
	coarse	$3.325 + 0.1411\tau$	$0.759 + 0.0168\tau$		$0.0934\tau^{0.639}$
MODABS	fine	$0.145 + 0.0203\tau$	$0.374 + 0.1365\tau$	$(1.43, 0.008 - 0.002\tau)$	$0.1642\tau^{0.775}$
	coarse	$3.101 + 0.3364\tau$	$0.729 + 0.098\tau$		$0.1482\tau^{0.684}$
ABS	fine	$0.134 + 0.0096\tau$	$0.383 + 0.0794\tau$	$(1.51, 0.02)$	$0.1748\tau^{0.891}$
	coarse	$3.448 + 0.9489\tau$	$0.743 + 0.0409\tau$		$0.1043\tau^{0.682}$
DUST	fine	$0.1416\tau^{-0.052}$	$0.7561\tau^{0.148}$	$(1.48\tau^{-0.021}, 0.0018\tau^{-0.08})$	$0.087\tau^{1.026}$
	coarse	2.2	$0.554\tau^{-0.052}$		$0.6786\tau^{1.057}$

**Table 2.** Micro-physical properties of aerosols models of Set 2. The median radius of the number size distribution  $r_{\text{mod}}$ , stand deviation  $\sigma$ , and complex refractive index  $m$  of two modes are listed in the table.  $w_{\text{coarse}}$  is the fraction of coarse mode in number concentration. The two lines of  $m$  for MA mod. abs. and MA abs. aerosol model are the refraction index of fine and coarse modes, respectively.

Type	Model	Fine Mode		Coarse Mode		$m = (\text{Re}, \text{Im})$	$w_{\text{coarse}}$
		$r_{\text{mod}}$	$e^\sigma$	$r_{\text{mod}}$	$e^\sigma$		
WA	WA1101	0.078	1.499	0.497	2.160	$(1.4, 5.0 \times 10^{-8})$	$(4.36 \times 10^{-4})$
	WA1102	0.088	1.499	0.509	2.160	$(1.4, 5.0 \times 10^{-8})$	$(4.04 \times 10^{-4})$
	WA1103	0.137	1.499	0.567	2.160	$(1.4, 5.0 \times 10^{-8})$	$(8.10 \times 10^{-4})$
	WA1104	0.030	2.030	0.240	2.030	$(1.4, 5.0 \times 10^{-8})$	$(1.53 \times 10^{-2})$
	WA1201	0.078	1.499	0.497	2.160	$(1.4, 4.0 \times 10^{-3})$	$(4.36 \times 10^{-4})$
	WA1202	0.088	1.499	0.509	2.160	$(1.4, 4.0 \times 10^{-3})$	$(4.04 \times 10^{-4})$
	WA1203	0.137	1.499	0.567	2.160	$(1.4, 4.0 \times 10^{-3})$	$(8.10 \times 10^{-4})$
	WA1301	0.078	1.499	0.497	2.160	$(1.4, 1.2 \times 10^{-2})$	$(4.36 \times 10^{-4})$
	WA1302	0.088	1.499	0.509	2.160	$(1.4, 1.2 \times 10^{-2})$	$(4.04 \times 10^{-4})$
	WA1303	0.137	1.499	0.567	2.160	$(1.4, 1.2 \times 10^{-2})$	$(8.10 \times 10^{-4})$
BB	BB2101	0.074	1.537	0.511	2.203	$(1.5, 1.0 \times 10^{-2})$	$(1.70 \times 10^{-4})$
	BB2102	0.087	1.537	0.567	2.203	$(1.5, 1.0 \times 10^{-2})$	$(2.06 \times 10^{-4})$
	BB2103	0.124	1.537	0.719	2.203	$(1.5, 1.0 \times 10^{-2})$	$(2.94 \times 10^{-4})$
	BB2201	0.074	1.537	0.511	2.203	$(1.5, 2.0 \times 10^{-2})$	$(1.70 \times 10^{-4})$
	BB2202	0.087	1.537	0.509	2.203	$(1.5, 2.0 \times 10^{-2})$	$(2.06 \times 10^{-4})$
	BB2203	0.124	1.537	0.719	2.203	$(1.5, 2.0 \times 10^{-2})$	$(2.94 \times 10^{-4})$
	BB2102	0.087	1.537	0.509	2.203	$(1.5, 3.0 \times 10^{-2})$	$(2.06 \times 10^{-4})$
	BB2103	0.124	1.537	0.719	2.203	$(1.5, 3.0 \times 10^{-2})$	$(2.94 \times 10^{-4})$
DD	BB2101	0.074	1.537	0.511	2.203	$(1.5, 3.0 \times 10^{-2})$	$(1.70 \times 10^{-4})$
	DD3101	0.042	1.697	0.670	1.806	$(1.53, 4.0 \times 10^{-3})$	$(4.35 \times 10^{-3})$
	DD3102	0.052	1.697	0.670	1.806	$(1.53, 4.0 \times 10^{-3})$	$(4.35 \times 10^{-3})$
	DD3201	0.042	1.697	0.670	1.806	$(1.53, 1.0 \times 10^{-2})$	$(4.35 \times 10^{-3})$
	DD3202	0.052	1.697	0.670	1.806	$(1.53, 1.0 \times 10^{-2})$	$(4.35 \times 10^{-3})$

Table 2. Cont.

Type	Model	Fine Mode		Coarse Mode		$m = (\text{Re}, \text{Im})$	$w_{\text{coarse}}$
		$r_{\text{mod}}$	$e^{\sigma}$	$r_{\text{mod}}$	$e^{\sigma}$		
MA	MA mod. abs.	0.030	2.030	0.240	2.030	$(1.4, 4.0 \times 10^{-3})$ $(1.4, 5.0 \times 10^{-8})$	$(1.55 \times 10^{-4})$
	MA abs.	0.030	2.030	0.240	2.030	$(1.4, 1.2 \times 10^{-2})$ $(1.4, 5.0 \times 10^{-8})$	$(1.55 \times 10^{-4})$
VO	VO4101	0.230	0.800	0.240	2.030	$(1.45, 7.5 \times 10^{-8})$	0.5

Both sets of aerosol models have been widely used in satellite remote sensing of aerosol properties and are representative for characterizing aerosol microphysical properties. According to the EPIC experiment in [24], Set 2 slightly outperformed in the retrieval outcome and can be suggested as a proper choice.

#### 4. Tests with Synthetic Data

In this section, we analyze the accuracy of the Bayesian model selection algorithm for synthetic measurements.

##### 4.1. Test 1

In the first test example, synthetic measurement spectra are simulated for each aerosol model included in Set 1 ( $m_e = \text{NONABS}, \text{MODABS}, \text{ABS}, \text{DUST}$ ), and for each measurement, all aerosol models from Set 1 are considered in the retrieval. Thus, the retrieval algorithm has the possibility of identifying the correct aerosol model. The exact aerosol optical depths and layer heights to be retrieved are

$$\tau_e = 0.25, 0.5, 0.75, 1, 1.25, 1.5, 1.75, 2 \quad (13)$$

and

$$h_e = 1.5, 2.5, 3.5, 4.5, 5.5, 6.5, 7.5, 8.5, 9.5 \text{ km}, \quad (14)$$

respectively. The a priori values, which coincide with the initial guesses, are  $\tau_a = 2.0$  and  $h_a = 2 \text{ km}$ , and a Lambertian surface with an albedo of 0.05 being assumed. The solar zenith, viewing zenith, and relative azimuth angles are  $\theta_o = 60^\circ$ ,  $\theta_v = 0^\circ$ , and  $\Delta\phi = 180^\circ$ . For the exact solution  $\mathbf{x}_e = [\tau_e, h_e]$ , the accuracy of the solution estimates is quantified through the relative errors

$$\varepsilon_{\text{mean}}^{\tau}(\tau_e, h_e) = \frac{|\tau_{\text{mean}} - \tau_e|}{\tau_e} \text{ and } \varepsilon_{\text{mean}}^h(\tau_e, h_e) = \frac{|h_{\text{mean}} - h_e|}{h_e} \quad (15)$$

corresponding to (cf. Equation (9))  $\hat{\mathbf{x}}_{\text{mean}}^{\delta} = [\tau_{\text{mean}}, h_{\text{mean}}]$  and

$$\varepsilon_{\text{max}}^{\tau}(\tau_e, h_e) = \frac{|\tau_{\text{max}} - \tau_e|}{\tau_e} \text{ and } \varepsilon_{\text{max}}^h(\tau_e, h_e) = \frac{|h_{\text{max}} - h_e|}{h_e} \quad (16)$$

corresponding to (cf. Equation (10))  $\hat{\mathbf{x}}_{\text{max}}^{\delta} = [\tau_{\text{max}}, h_{\text{max}}]$ .

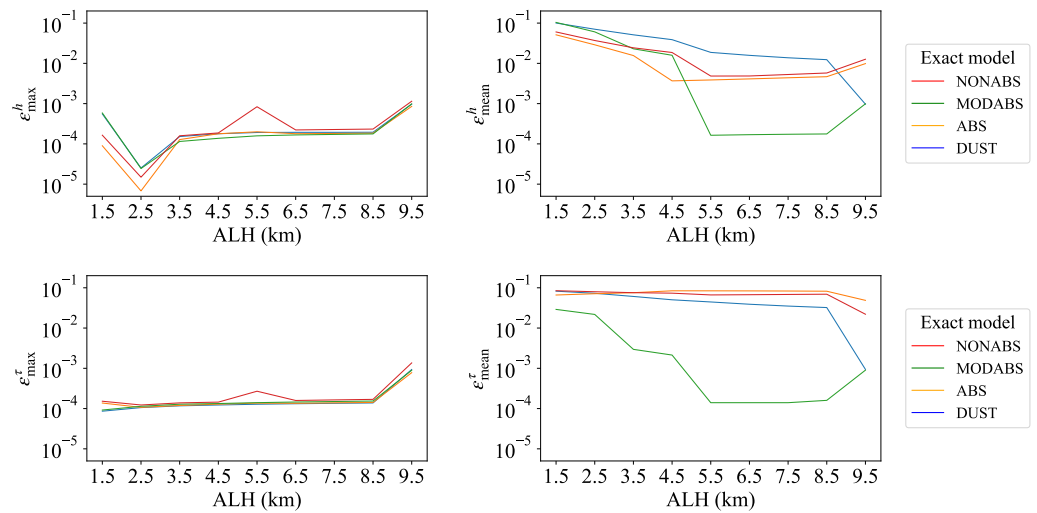
In Figures 1 and 2, we illustrate the variations of the relative errors with respect to the exact aerosol layer height  $h_e$  for  $\tau_e = 0.5$  (i.e.,  $\varepsilon_{\text{mean,max}}^{\tau,h}(\tau_e = 0.5, h_e)$ ), and the aerosol optical depth  $\tau_e$  for  $h_e = 3.5 \text{ km}$  (i.e.,  $\varepsilon_{\text{mean,max}}^{\tau,h}(\tau_e, h_e = 3.5 \text{ km})$ ), respectively. The following conclusions can be drawn:

1. The relative errors corresponding to the maximum solution estimate ( $\varepsilon_{\text{max}}^{\tau}$  and  $\varepsilon_{\text{max}}^h$ ) are considered to be acceptable according to the scientific requirements defined in the pre-launch characterization tests and significantly smaller than those corresponding to the mean solution estimate ( $\varepsilon_{\text{mean}}^{\tau}$  and  $\varepsilon_{\text{mean}}^h$ ). Thus, the retrieval algorithm can recognize correctly the exact aerosol model.

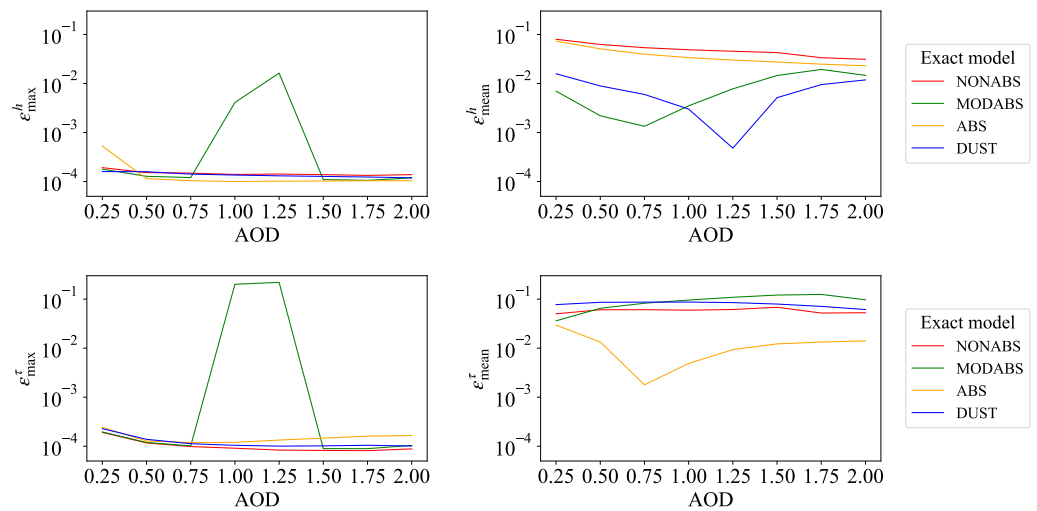


2. For the maximum solution estimate, the retrieved aerosol optical depth achieves a higher accuracy than the retrieved aerosol layer height.
3. Different aerosol models could have similar a posteriori densities as the inversion process is not ideally perfect. An inappropriate aerosol model may occasionally be identified, which can result in unexpected errors ( $\tau_e = 1, 1.25$ ).

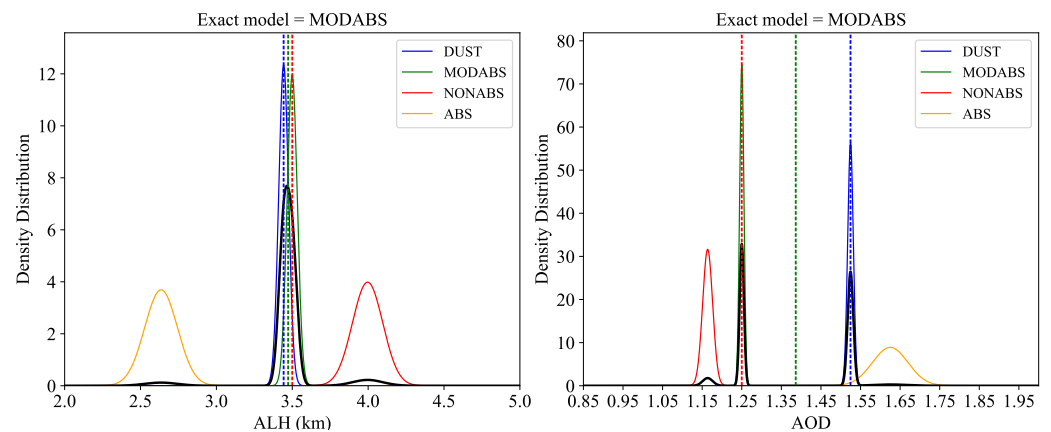
The relative errors  $\varepsilon_{\max}^{\tau}(\tau_e, h_e = 3.5 \text{ km})$  and  $\varepsilon_{\max}^h(\tau_e, h_e = 3.5 \text{ km})$  attain their highest values (of about 0.22 and 0.016, respectively) for  $m_e = \text{MODABS}$  and  $\tau_e = 1, 1.25$ . To explain this result, in Figure 3 we plot the a posteriori densities  $p(\mathbf{x} = [\tau, h] | \bar{\mathbf{y}}^\delta, m)$  for  $m = \text{NONABS}, \text{MODABS}, \text{ABS}, \text{DUST}$ , and the mean a posteriori densities  $p_{\text{mean}}(\mathbf{x} = [\tau, h] | \bar{\mathbf{y}}^\delta)$  in the case  $m_e = \text{MODABS}, \tau_e = 1.25$  and  $h_e = 3.5 \text{ km}$ . The plots indicate that a posterior density for  $m = \text{DUST}$  is of similar height and width to that for  $m = \text{MODABS}$ ; the maximum solution is achieved at  $m = \text{MODABS}$ , i.e., the DUST appears to be the model with the highest evidence.



**Figure 1.** Relative errors  $\varepsilon_{\text{mean,max}}^{\tau,h}(\tau_e = 0.5, h_e)$  for the aerosol models from Set 1. ALH represents the aerosol layer height.



**Figure 2.** Relative errors  $\varepsilon_{\text{mean,max}}^{\tau,h}(\tau_e, h_e = 3.5 \text{ km})$  for the aerosol models from Set 1. AOD represents the aerosol optical depth.



**Figure 3.** The a posteriori densities  $p(\mathbf{x} = [\tau, h] | \bar{\mathbf{y}}^\delta, m)$  for  $m = \text{NONABS}, \text{MODABS}, \text{ABS},$  and  $\text{DUST}$ , and the mean a posteriori densities  $p_{\text{mean}}(\mathbf{x} = [\tau, h] | \bar{\mathbf{y}}^\delta)$  in the case  $m_e = \text{MODABS}$ ,  $\tau_e = 1.25$ , and  $h_e = 3.5$  km. The black curve indicates the mean a posteriori density. In each plot, the red vertical dashed line corresponds to the exact values to be retrieved  $(\tau_e, h_e)$ , the blue vertical dashed line to the maximum solution estimate  $(\tau_{\text{max}}, h_{\text{max}})$ , and the green dashed line to the mean solution estimates  $(\tau_{\text{mean}}, h_{\text{mean}})$ .

#### 4.2. Test 2

In the second test example, synthetic measurement spectra are produced for each aerosol model included in Set 1 ( $m_e = \text{NONABS}, \text{MODABS}, \text{ABS}, \text{DUST}$ ), and for each measurement all aerosol models from Set 2 are considered in the retrieval. The mean solution estimate and the mean a posteriori density are computed for the first 10 aerosol models with the highest evidence.

The variations of the relative errors with respect to the exact aerosol layer height  $h_e$  for  $\tau_e = 0.5$  and the aerosol optical depth  $\tau_e$  for  $h_e = 3.5$  km are illustrated in Figures 4 and 5, respectively. The plots indicate that

1. the relative errors are larger than those in the first test example,
2. the relative errors corresponding to the maximum solution estimate ( $\varepsilon_{\text{max}}^\tau$  and  $\varepsilon_{\text{max}}^h$ ) and the mean solution estimate ( $\varepsilon_{\text{mean}}^\tau$  and  $\varepsilon_{\text{mean}}^h$ ) are comparable, and
3. on average, the retrieved aerosol layer height obtains a higher accuracy than the retrieved aerosol optical depth.

The mean a posteriori densities  $p_{\text{mean}}(\mathbf{x} = [\tau, h] | \bar{\mathbf{y}}^\delta)$  are shown in Figures 6 and 7 for  $\tau_e = 0.5$ ,  $h_e = 3.5$  km, and all exact aerosol models  $m_e = \text{NONABS}, \text{MODABS}, \text{ABS},$  and  $\text{DUST}$ . The following conclusions could be made:

1.  $h_{\text{mean}}$  and  $h_{\text{max}}$  are both not too far from  $h_e$ ; thus, for aerosol layer height retrieval, the maximum solution estimate and the mean solution estimate ( $\varepsilon_{\text{mean}}^\tau$  and  $\varepsilon_{\text{mean}}^h$ ) have similar accuracies;
2.  $\tau_{\text{mean}}$  is relatively closer to  $\tau_e$  than  $\tau_{\text{max}}$ ; thus, for aerosol optical depth retrieval, the mean solution estimate performs better than the maximum solution estimate;
3. aerosol layer height retrievals have wide a posteriori densities that cover the exact layer height; and
4. aerosol optical depth retrievals have multi-peaked densities, in which the exact optical depth does not have the highest probability.

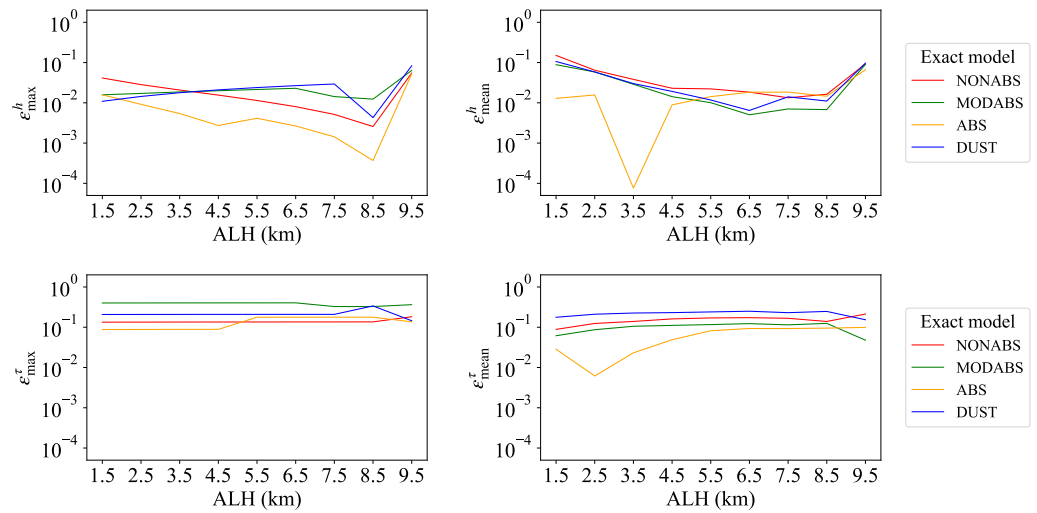


Figure 4. Relative errors  $\epsilon_{\text{mean,max}}^{\tau,h}(\tau_e = 0.5, h_e)$  for the aerosol models from Set 2.

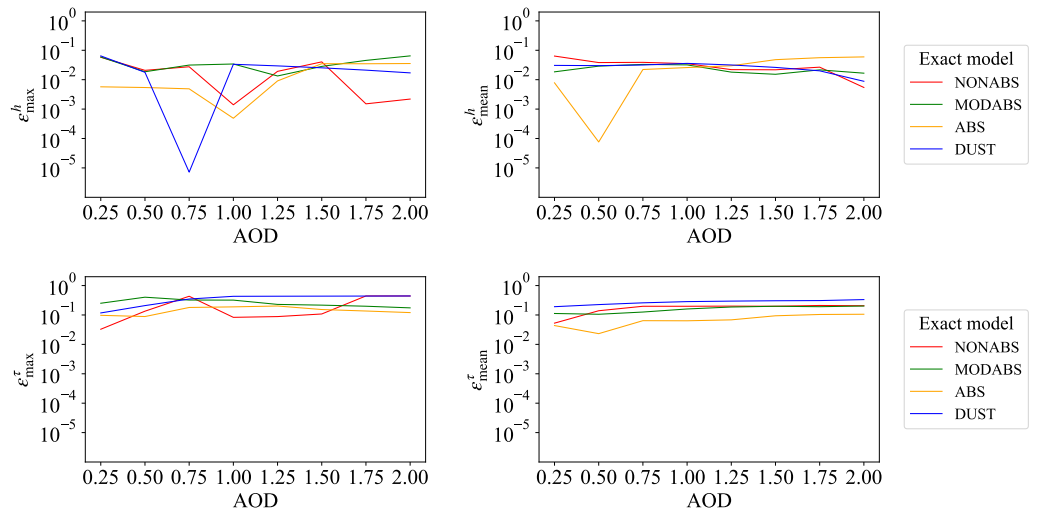


Figure 5. Relative errors  $\epsilon_{\text{mean,max}}^{\tau,h}(\tau_e, h_e = 3.5 \text{ km})$  for the aerosol models from Set 2.

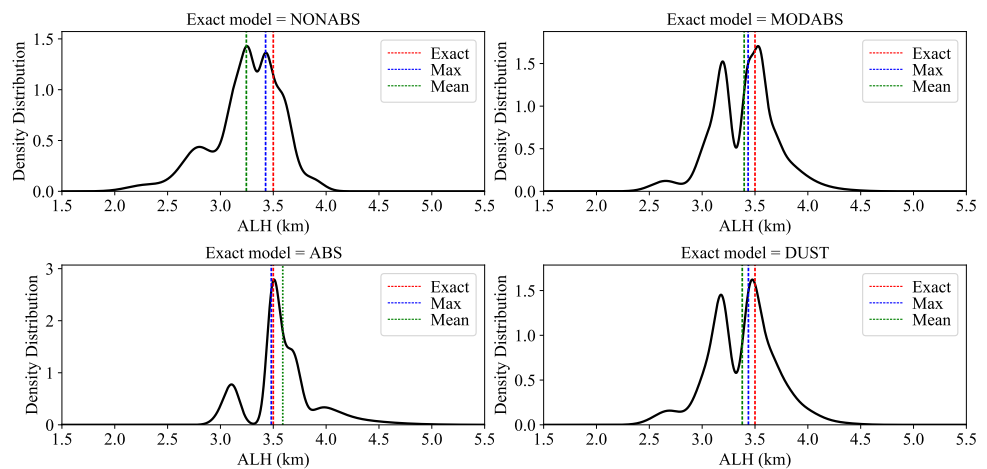


Figure 6. The mean a posteriori densities  $p_{\text{mean}}(h | \bar{y}^\delta)$  for  $m_e = \text{NONABS, MODABS, ABS, DUST}$ ,  $\tau_e = 0.5$ , and  $h_e = 3.5 \text{ km}$ . In each plot, the red vertical dashed line correspond to the exact values to be retrieved  $(\tau_e, h_e)$ , the blue vertical dashed line to the maximum solution estimate  $(\tau_{\max}, h_{\max})$ , and the green dashed line to the mean solution estimates  $(\tau_{\text{mean}}, h_{\text{mean}})$ .

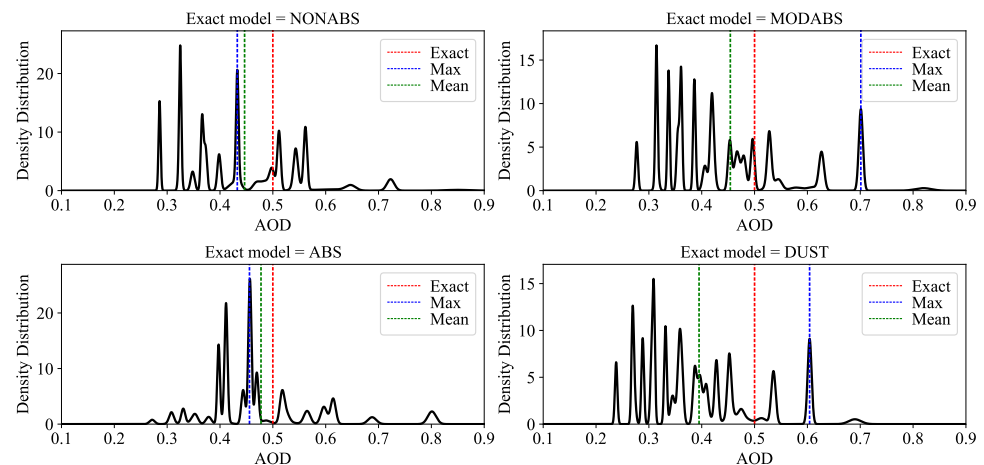


Figure 7. The same as in Figure 6 but for the mean a posteriori densities  $p_{\text{mean}}(\tau | \bar{y}^{\delta})$ .

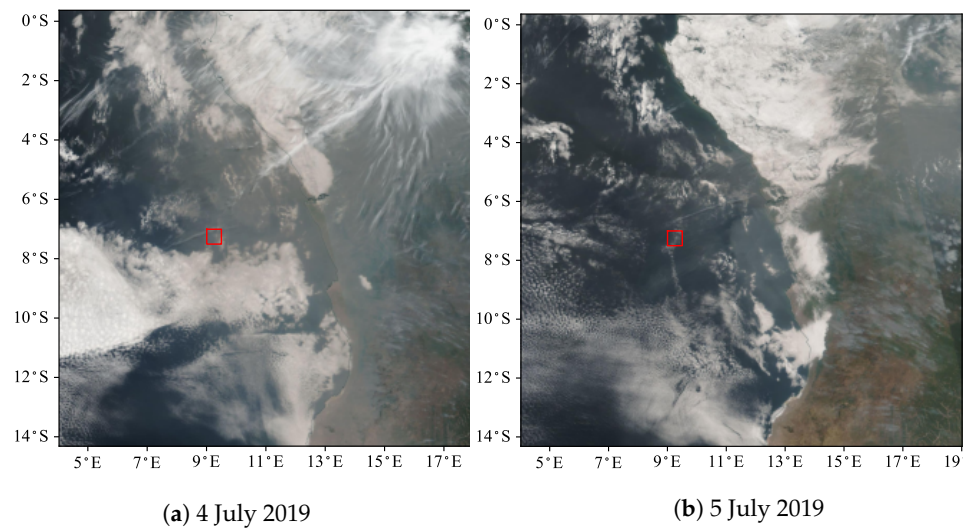
### 5. Case Study with TROPOMI Data

To test the performance of the retrieval algorithm on real TROPOMI data, we chose a wild fire scene in South Africa and considered the measurements recorded on 4–5 July 2019. As can be seen from the respective VIIRS images, the wild fire smoke on 4 July 2019 (Figure 8a) traveled beyond the coastline and extended over the ocean, so that the smoke on 5 July 2019 (Figure 8b) was thinner. Regional studies of aerosol optical/microphysical properties during biomass burning can be found in [43–45].

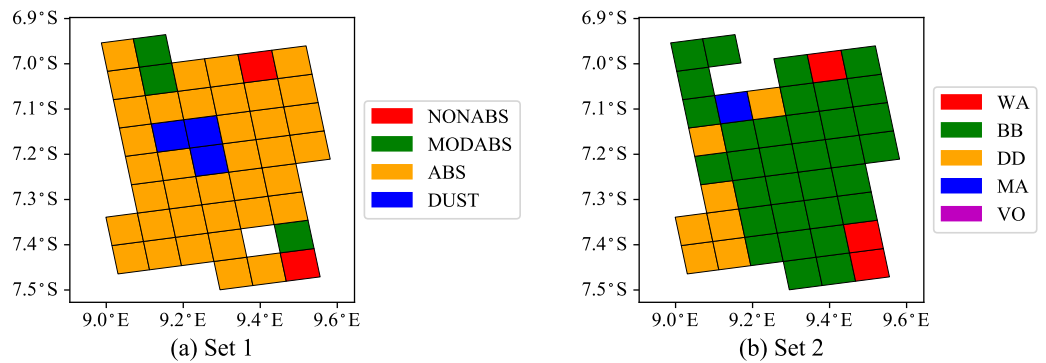
The aerosol models included in Sets 1 and 2 are used in the retrieval. To decrease the retrieval uncertainty caused by unrealistic surface properties, the geometry-dependent effective Lambertian equivalent reflectivity (GE\_LER) products [46] are used. The ground pixels with cloud fraction larger than 0.15 are excluded for this analysis, in which case, the scene is assumed to be cloud free so that we can retrieve valid aerosol information on sufficient number of pixels without significant impact by cloud contamination.

Figure 9 shows the aerosol model with the highest evidence from Set 1 as well as the aerosol type containing the aerosol model with the highest evidence from Set 2. The most likely models are ABS from Set 1 and BB type from Set 2. The model evidence for each aerosol model from Set 1 is shown in Figure 10. Note that the differences between the model evidences for the four aerosol models are not very large, and the model evidence of ABS was slightly larger than those of the other models. In Set 2, there are 26 aerosol models and five aerosol types. The sum of the first 10 best aerosol model evidences for each aerosol type from Set 2 are illustrated in Figure 11. The most probable type is BB. In conclusion, the most plausible aerosol models identified by the algorithm, that is, ABS from Set 1 and BB from Set 2, are of the biomass burning aerosol type. This strongly absorbing aerosol type is consistent with the thick smoke observed in the true-color image.

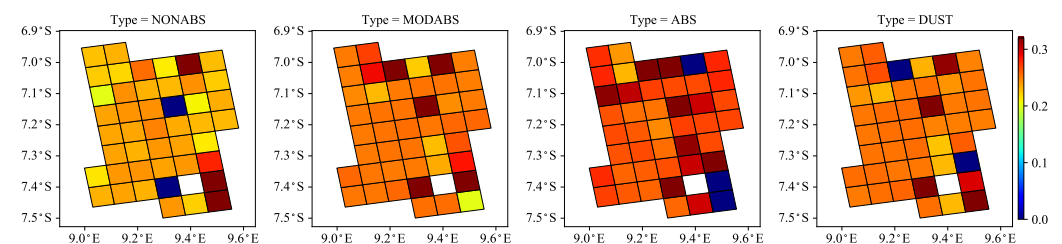
The predominant models for retrieval on 5 July 2019 are ABS and DUST among Set 1 (Figure 12a), and BB and DD among Set 2 (Figure 12b). Thus, in addition to the aerosol models identified for 4 July 2019, the dust aerosol model comes into play. The model evidence for each aerosol model from Set 1 and each aerosol type from Set 2 are displayed in Figures 13 and 14, respectively. In conclusion, compared with that on 4 July 2019, the dominance of biomass burning aerosol type (ABS and BB) is less obvious. Taking into account the thinner smoke on 5 July 2019 and the long traveling distance from the origin, the presence of a less absorbing mixture of different aerosol types (biomass burning and dust) seems to be plausible. The corresponding maximum and mean solution estimates are shown in Figures 15–18 (Figures 15 and 17 for Set 1, and Figures 16 and 18 for Set 2).



**Figure 8.** True-color VIIRS images recorded on (a) 4 and (b) 5 July 2019.

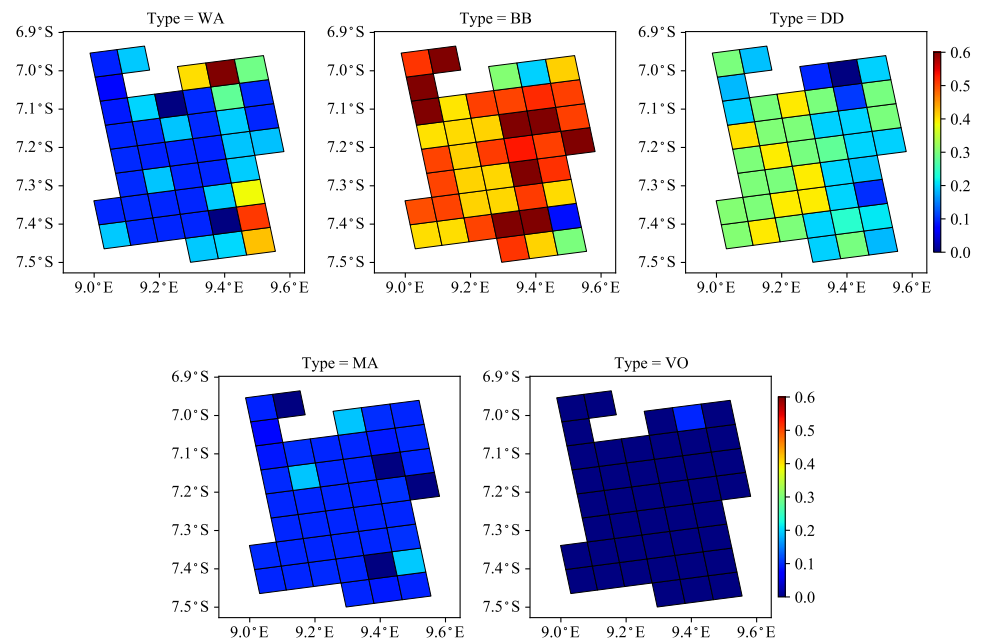


**Figure 9.** (a) The aerosol model with the highest evidence from Set 1, and (b) the aerosol type containing the aerosol model with the highest evidence from Set 2. The TROPOMI spectra were recorded on 4 July 2019.



**Figure 10.** The model evidence for each aerosol model from Set 1. The TROPOMI spectra were recorded on 4 July 2019.

To demonstrate the performance under various aerosol loading scenarios, we performed retrievals for another two cases from TROPOMI. The first case focused on a desert dust aerosol case in Sahara on 6 June 2020 (see Figure S1 for the VIIRS image). The model evidence for the aerosol models in Set 1 and the aerosol types of Set 2 are given in Figures S2–S4. The prevailing aerosol model and aerosol type are DUST from Set 1 and DD from Set 2, given the fact that both models represent desert dust aerosols. The second case was for a urban aerosol case on 10 February 2020 over eastern China (see Figure S7 for the VIIRS image) where many industrial cities are located. As shown in Figures S8–S10, the NONABS model in Set 1 and the WA aerosol type in Set 2 are the most plausible choices, as both stands for industrial aerosols.



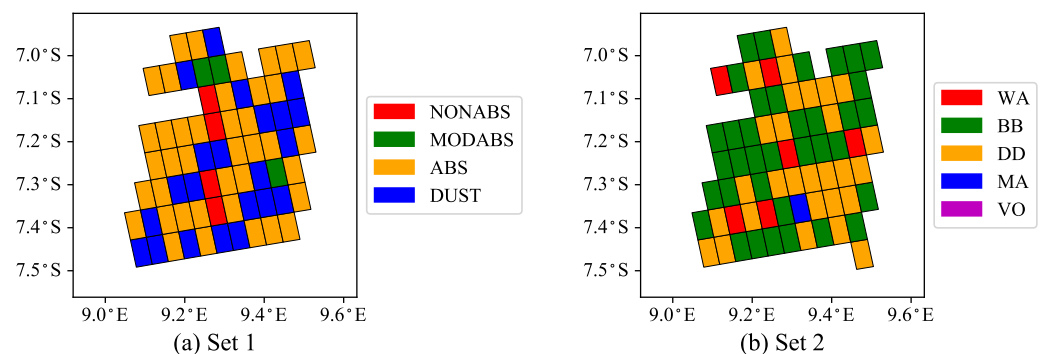
**Figure 11.** The sum of the first 10 best aerosol model evidences for each aerosol type from Set 2. The TROPOMI spectra were recorded on 4 July 2019.

Figures S5 and S6 illustrate the maximum and mean solution estimates for the first case, respectively. Figures S11 and S12 depict the corresponding solution estimates for the second case.

The dominant aerosol type or aerosol model for each study can be found from the above analysis. However, the most likely model varies from pixel to pixel, indicating that sometimes a “wrong” model may be chosen, which is consistent with the findings using the synthetic data. Based on the results of the retrieval solutions, we can see that

1. the mean solution estimates show a slightly smoother spatial pattern than the maximum solution estimates, and
2. despite the differences in the micro-physical properties of the aerosol models from Sets 1 and 2, the spatial distributions of the mean retrieval results are comparable.

In this study, the state vector was a two-element vector (aerosol optical depth and layer height) by considering the box profile for simplicity. The degree of freedom was estimated to be 2 in most cases. From the practical point of view, the retrieval on a Dell desktop (with 12 processors at 3.2 GHz, 31.2 GB of RAM) took less than 10 min (10–15 iterations in total) by running the program with all the models included in Set 1 and longer than 60 min (approximately 100 iterations in total) by considering all the models included in Set 2.



**Figure 12.** The same as in Figure 9 but for the data on 5 July 2019.

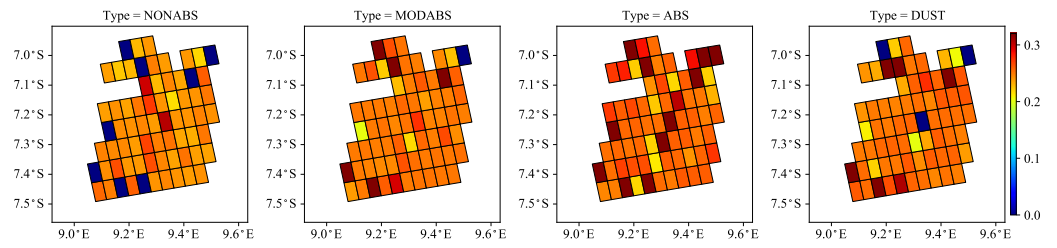


Figure 13. The same as in Figure 10 but for the data on 5 July 2019.

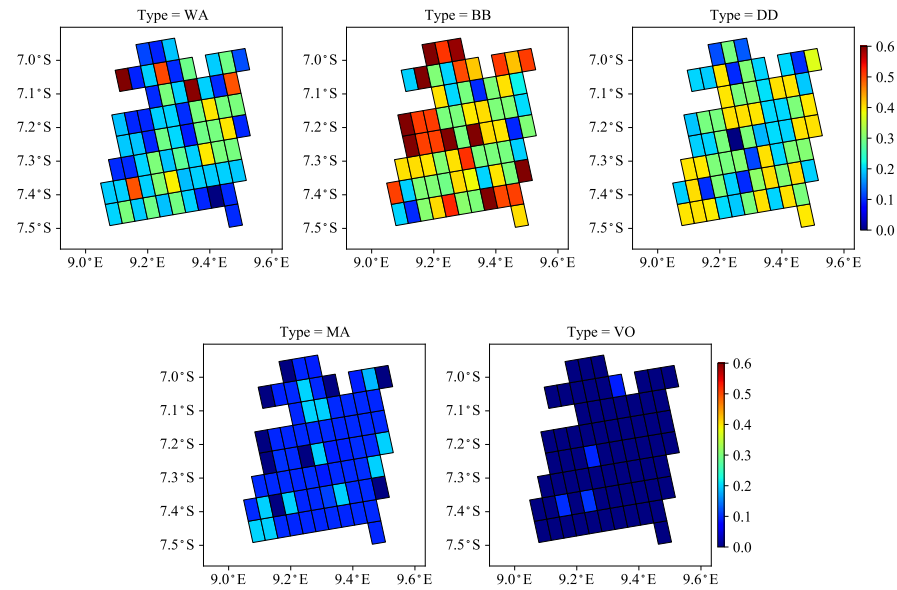


Figure 14. The same as in Figure 11 but for the data on 5 July 2019.

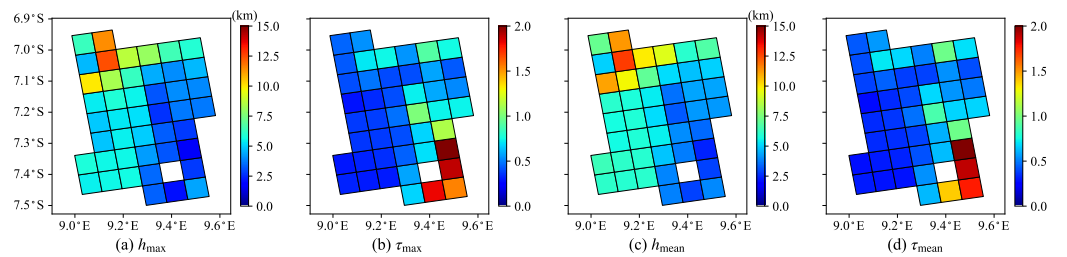


Figure 15. The maximum solution estimates ( $h_{max}$ ,  $\tau_{max}$ ) and the mean solution estimates ( $h_{mean}$ ,  $\tau_{mean}$ ) for Set 1 and data on 4 July 2019.

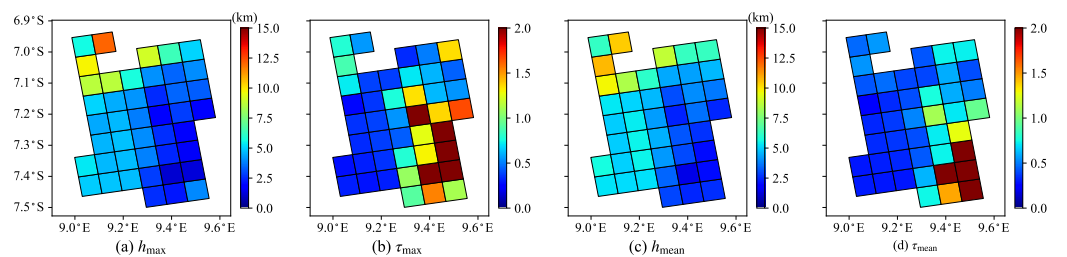


Figure 16. The maximum solution estimates ( $h_{max}$ ,  $\tau_{max}$ ) and the mean solution estimates ( $h_{mean}$ ,  $\tau_{mean}$ ) for Set 2 and data on 4 July 2019.

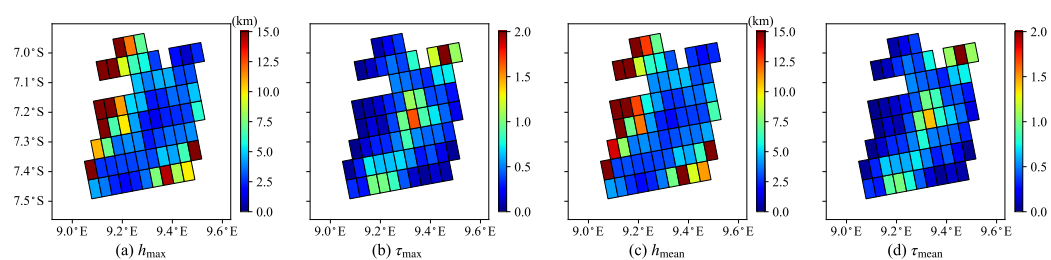


Figure 17. The same as in Figure 15 but for the data on 5 July 2019.

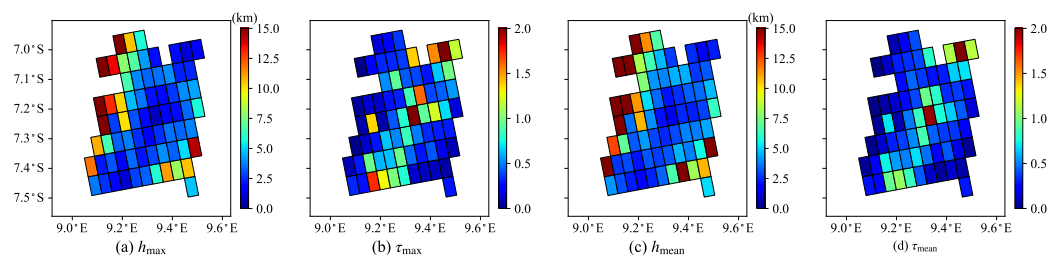


Figure 18. The same as in Figure 16 but for the data on 5 July 2019.

## 6. Conclusions

In this paper, the results of aerosol retrieval computed by means of a Bayesian algorithm that takes into account the uncertainty in aerosol model selection are presented. The solution corresponding to a specific aerosol model is characterized by a relative evidence which is used to construct (i) the maximum solution estimate, corresponding to the aerosol model with the highest evidence, and (ii) the mean solution estimate, representing a linear combination of solutions weighted by their evidences. The algorithm is applied to the retrieval of aerosol optical depth and layer height from synthetic and real TROPOMI data. The real TROPOMI data were taken on 4–5 July 2019 over a wild fire scene in South Africa. In the retrieval, two sets of aerosol models are taken into account; these correspond to the MODIS dark-target and OMAERO algorithms. The following conclusions are drawn.

1. When the exact aerosol model, for which synthetic data are generated, is included in the set of candidate models, the relative errors corresponding to the maximum solution estimate are relatively small. When this is not the case, it is likely that several aerosol models are able to fit the data equally well. In such situations, the mean solution estimate has a smaller bias than the maximum solution estimate.
2. For the real measurements on 4 July 2019, the absorbing aerosol model from Set 1 and the biomass burning aerosol type from Set 2 are found to be the most plausible. This result is in agreement with the thick smoke observed in the true-color image. For the thinner smoke scenario on 5 July 2019, the above models together with the dust aerosol model are found to be the most probable aerosol models. Actually, no dominant aerosol model, but rather a less absorbing mixture of different aerosol types, is identified in this case. The mean and maximum solution estimates have a similar spatial distribution, but the mean solution estimates have a more continuous spatial pattern.
3. The two TROPOMI cases on 6 June 2020 and 10 February 2020 for desert dust and urban aerosols, respectively, have demonstrated the promising performance of the proposed algorithm under various aerosol loading scenarios.
4. A definite choice between Sets 1 and 2 for possible candidate models may not exist and a suitable one could be problem dependent.

Note that when applying the Bayesian approach, we have to perform a retrieval for each candidate model. For this reason, the retrieval algorithm is computationally expensive, especially a set like Set 2 contains a larger number of aerosol models. To enhance its efficiency, development of a machine learning-based scheme is currently ongoing.



**Supplementary Materials:** The following are available online at <https://www.mdpi.com/article/10.3390/rs13132489/s1>, Figure S1: True-color VIIRS image recorded on 6 June 2020. Figure S2: (a) The aerosol model with the highest evidence from Set 1, and (b) the aerosol type containing the aerosol model with the highest evidence from Set 2. The TROPOMI spectra were recorded on 6 June 2020. Figure S3: The model evidence for each aerosol model from Set 1. The TROPOMI spectra were recorded on 6 June 2020. Figure S4: The sum of the first 10 best aerosol model evidences for each aerosol type from Set 2. The TROPOMI spectra were recorded on 6 June 2020. Figure S5: The maximum solution estimates ( $h_{\max}$ ,  $\tau_{\max}$ ) and the mean solution estimates ( $h_{\text{mean}}$ ,  $\tau_{\text{mean}}$ ) for Set 1 and data on 6 June 2020. Figure S6: The same as in Figure S5 but for Set 2. Figure S7: True-color VIIRS image recorded on 10 February 2020. Figure S8: (a) The aerosol model with the highest evidence from Set 1, and (b) the aerosol type containing the aerosol model with the highest evidence from Set 2. The TROPOMI spectra were recorded on 10 February 2020. Figure S9: The model evidence for each aerosol model from Set 1. The TROPOMI spectra were recorded on 10 February 2020. Figure S10: The sum of the first 10 best aerosol model evidences for each aerosol type from Set 2. The TROPOMI spectra were recorded on 10 February 2020. Figure S11: The maximum solution estimates ( $h_{\max}$ ,  $\tau_{\max}$ ) and the mean solution estimates ( $h_{\text{mean}}$ ,  $\tau_{\text{mean}}$ ) for Set 1 and data on 10 February 2020. Figure S12: The same as in Figure S11 but for Set 2.

**Author Contributions:** Conceptualization, L.R. and J.X.; methodology, J.X. and A.D.; software, L.R.; investigation, D.S.E. and D.G.L.; data curation, L.R. and J.X.; writing—original draft preparation, L.R., J.X. and A.D.; writing—review and editing, D.S.E., D.G.L., and A.D.; visualization, L.R.; supervision, D.G.L. and A.D.; project administration, D.S.E. and D.G.L. All authors have read and agreed to the published version of the manuscript

**Funding:** This research was supported by the DLR programmatic [Nachwuchsgruppe “Retrieval der nächsten Generation”, 2,472,469]. The work of L. Rao was funded by the Chinese Scholarship Council.

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** Not applicable.

**Acknowledgments:** The authors are grateful to DLR and KNMI for sharing the related information of TROPOMI/S5P.

**Conflicts of Interest:** The authors declare no conflict of interest.

## Abbreviations

The following abbreviations are used in this manuscript:

ALH	Aerosol Layer Height
AOD	Aerosol Optical Depth
GE_LER	Geometry-dependent Effective Lambertian Equivalent Reflectivity
MODIS	Moderate-resolution Imaging Spectroradiometer
NIR	Near-infrared
OMI	Ozone Monitoring Instrument
OMAERO	OMI Multi-wavelength
S5P	Sentinel-5 Precursor
TROPOMI	Tropospheric Monitoring Instrument
ABS	Absorbing (Set 1)
DUST	Desert dust (Set 1)
MODABS	Moderately absorbing (Set 1)
NONABS	Non-absorbing (Set 1)
BB	Biomass Burning (Set 2)
DD	Desert Dust (Set 2)
MA	Marine (Set 2)
VO	Volcanic (Set 2)
WA	Weakly absorbing (Set 2)

## References

1. Holben, B.; Vermote, E.; Kaufman, Y.J.; Tanré, D.; Kalb, V. Aerosol retrieval over land from AVHRR data-application for atmospheric correction. *IEEE Trans. Geosci. Remote Sens.* **1992**, *30*, 212–222. [[CrossRef](#)]
2. Levy, R.; Mattoo, S.; Munchak, L.; Remer, L.; Sayer, A.; Patadia, F.; Hsu, N. The Collection 6 MODIS aerosol products over land and ocean. *Atmos. Meas. Tech.* **2013**, *6*, 2989–3034. [[CrossRef](#)]
3. Jackson, J.M.; Liu, H.; Laszlo, I.; Kondragunta, S.; Remer, L.A.; Huang, J.; Huang, H.C. Suomi-NPP VIIRS aerosol algorithms and data products. *J. Geophys. Res. Atmos.* **2013**, *118*, 12673–12689. [[CrossRef](#)]
4. Wei, J.; Li, Z.; Sun, L.; Peng, Y.; Zhang, Z.; Li, Z.; Su, T.; Feng, L.; Cai, Z.; Wu, H. Evaluation and uncertainty estimate of next-generation geostationary meteorological Himawari-8/AHI aerosol products. *Sci. Total Environ.* **2019**, *692*, 879–891. [[CrossRef](#)] [[PubMed](#)]
5. Kahn, R.A.; Li, W.H.; Moroney, C.; Diner, D.J.; Martonchik, J.V.; Fishbein, E. Aerosol source plume physical characteristics from space-based multiangle imaging. *J. Geophys. Res. Atmos.* **2007**, *112*. [[CrossRef](#)]
6. Virtanen, T.; Kolmonen, P.; Rodríguez, E.; Sogacheva, L.; Sundström, A.M.; de Leeuw, G. Ash plume top height estimation using AATSR. *Atmos. Meas. Tech.* **2014**, *7*, 2437–2456. [[CrossRef](#)]
7. Dubovik, O.; Herman, M.; Holdak, A.; Lapyonok, T.; Tanré, D.; Deuzé, J.; Ducos, F.; Sinyuk, A.; Lopatin, A. Statistically optimized inversion algorithm for enhanced retrieval of aerosol properties from spectral multi-angle polarimetric satellite observations. *Atmos. Meas. Tech.* **2011**, *4*, 975–1018. [[CrossRef](#)]
8. Nanda, S.; Graaf, M.d.; Veefkind, J.P.; Sneep, M.; Linden, M.t.; Sun, J.; Levelt, P.F. A first comparison of TROPOMI aerosol layer height (ALH) to CALIOP data. *Atmos. Meas. Tech.* **2020**, *13*, 3043–3059. [[CrossRef](#)]
9. Kokhanovsky, A.; Davis, A.; Cairns, B.; Dubovik, O.; Hasekamp, O.; Sano, I.; Mukai, S.; Rozanov, V.; Litvinov, P.; Lapyonok, T.; et al. Space-based remote sensing of atmospheric aerosols: The multi-angle spectro-polarimetric frontier. *Earth Sci. Rev.* **2015**, *145*, 85–116. [[CrossRef](#)]
10. Hess, M.; Koepke, P.; Schult, I. Optical properties of aerosols and clouds: The software package OPAC. *Bull. Am. Meteorol. Soc.* **1998**, *79*, 831–844. [[CrossRef](#)]
11. Levy, R.C.; Remer, L.A.; Dubovik, O. Global aerosol optical properties and application to Moderate Resolution Imaging Spectroradiometer aerosol retrieval over land. *J. Geophys. Res. Atmos.* **2007**, *112*. [[CrossRef](#)]
12. Torres, O.; Tanskanen, A.; Veihelmann, B.; Ahn, C.; Braak, R.; Bhartia, P.K.; Veefkind, P.; Levelt, P. Aerosols and surface UV products from Ozone Monitoring Instrument observations: An overview. *J. Geophys. Res. Atmos.* **2007**, *112*. [[CrossRef](#)]
13. Chin, M.; Ginoux, P.; Kinne, S.; Torres, O.; Holben, B.N.; Duncan, B.N.; Martin, R.V.; Logan, J.A.; Higurashi, A.; Nakajima, T. Tropospheric aerosol optical thickness from the GOCART model and comparisons with satellite and Sun photometer measurements. *J. Atmos. Sci.* **2002**, *59*, 461–483. [[CrossRef](#)]
14. Taylor, M.; Kazadzis, S.; Amiridis, V.; Kahn, R. Global aerosol mixtures and their multiyear and seasonal characteristics. *Atmos. Environ.* **2015**, *116*, 112–129. [[CrossRef](#)]
15. Martonchik, J.; Diner, D.; Kahn, R.; Ackerman, T.; Verstraete, M.; Pinty, B.; Gordon, H. Techniques for the retrieval of aerosol properties over land and ocean using multiangle imaging. *IEEE Trans. Geosci. Remote Sens.* **1998**, *36*, 1212–1227. [[CrossRef](#)]
16. Bellouin, N.; Boucher, O.; Tanré, D.; Dubovik, O. Aerosol absorption over the clear-sky oceans deduced from POLDER-1 and AERONET observations. *Geophys. Res. Lett.* **2003**, *30*. [[CrossRef](#)]
17. Kim, J.; Lee, J.; Lee, H.C.; Higurashi, A.; Takemura, T.; Song, C.H. Consistency of the aerosol type classification from satellite remote sensing during the Atmospheric Brown Cloud–East Asia Regional Experiment campaign. *J. Geophys. Res. Atmos.* **2007**, *112*. [[CrossRef](#)]
18. Penning de Vries, M.J.M.; Beirle, S.; Hörmann, C.; Kaiser, J.W.; Stammes, P.; Tilstra, L.G.; Tuinder, O.N.E.; Wagner, T. A global aerosol classification algorithm incorporating multiple satellite data sets of aerosol and trace gas abundances. *Atmos. Chem. Phys.* **2015**, *15*, 10597–10618. [[CrossRef](#)]
19. Kumar, K.R.; Kang, N.; Yin, Y. Classification of key aerosol types and their frequency distributions based on satellite remote sensing data at an industrially polluted city in the Yangtze River Delta, China. *Int. J. Climatol.* **2018**, *38*, 320–336. [[CrossRef](#)]
20. Hoeting, J.A.; Madigan, D.; Raftery, A.E.; Volinsky, C.T. Bayesian model averaging: A tutorial. *Stat. Sci.* **1999**, *14*, 382–401.
21. Määttä, A.; Laine, M.; Tamminen, J.; Veefkind, J. Quantification of uncertainty in aerosol optical thickness retrieval arising from aerosol microphysical model and other sources, applied to Ozone Monitoring Instrument (OMI) measurements. *Atmos. Meas. Tech.* **2014**, *7*, 1185–1199. [[CrossRef](#)]
22. Kauppi, A.; Kolmonen, P.; Laine, M.; Tamminen, J. Aerosol-type retrieval and uncertainty quantification from OMI data. *Atmos. Meas. Tech.* **2017**, *10*. [[CrossRef](#)]
23. Sasi, S.; Natraj, V.; Molina García, V.; Efremenko, D.S.; Loyola, D.; Doicu, A. Model Selection in Atmospheric Remote Sensing with an Application to Aerosol Retrieval from DSCOVR/EPIC, Part 1: Theory. *Remote Sens.* **2020**, *12*, 3724. [[CrossRef](#)]
24. Sasi, S.; Natraj, V.; Molina García, V.; Efremenko, D.S.; Loyola, D.; Doicu, A. Model Selection in Atmospheric Remote Sensing with Application to Aerosol Retrieval from DSCOVR/EPIC. Part 2: Numerical Analysis. *Remote Sens.* **2020**, *12*, 3656. [[CrossRef](#)]
25. Marshak, A.; Herman, J.; Adam, S.; Karin, B.; Carn, S.; Cede, A.; Geogdzhayev, I.; Huang, D.; Huang, L.K.; Knyazikhin, Y.; et al. Earth observations from DSCOVR EPIC instrument. *Bull. Am. Meteorol. Soc.* **2018**, *99*, 1829–1850. [[CrossRef](#)]

26. Veefkind, J.; Aben, I.; McMullan, K.; Förster, H.; de Vries, J.; Otter, G.; Claas, J.; Eskes, H.; de Haan, J.; Kleipool, Q.; et al. TROPOMI on the ESA Sentinel-5 Precursor: A GMES mission for global observations of the atmospheric composition for climate, air quality and ozone layer applications. *Remote Sens. Environ.* **2012**, *120*, 70–83. [[CrossRef](#)]
27. Spurr, R. LIDORT and VLIDORT. Linearized pseudo-spherical scalar and vector discrete ordinate radiative transfer models for use in remote sensing retrieval problems. In *Light Scattering Reviews*; Kokhanovsky, A., Ed.; Springer: Berlin/Heidelberg, Germany, 2008; Volume 3, pp. 229–271.
28. Efremenko, D.; Doicu, A.; Loyola, D.; Trautmann, T. Acceleration techniques for the discrete ordinate method. *J. Quant. Spectrosc. Radiat. Transf.* **2013**, *114*, 73–81. [[CrossRef](#)]
29. Goody, R.; West, R.; Chen, L.; Crisp, D. The correlated-k method for radiation calculations in nonhomogeneous atmospheres. *J. Quant. Spectrosc. Radiat. Transf.* **1989**, *42*, 437–651. [[CrossRef](#)]
30. Natraj, V.; Shia, R.L.; Huang, X.; Margolis, J.S.; Yung, Y.L. Application of principal component analysis to high spectral resolution radiative transfer: A case study of the O<sub>2</sub> A band. *J. Quant. Spectrosc. Radiat. Transf.* **2005**, *95*, 539–556. [[CrossRef](#)]
31. Natraj, V.; Shia, R.L.; Yung, Y.L. On the use of principal component analysis to speed up radiative transfer calculations. *J. Quant. Spectrosc. Radiat. Transf.* **2010**, *111*, 810–816. [[CrossRef](#)]
32. Xu, J.; Schreier, F.; Doicu, A.; Trautmann, T. Assessment of Tikhonov-type regularization methods for solving atmospheric inverse problems. *J. Quant. Spectrosc. Radiat. Transf.* **2016**, *184*, 274–286. [[CrossRef](#)]
33. Xu, J.; Rao, L.; Schreier, F.; Efremenko, D.S.; Doicu, A.; Trautmann, T. Insight into Construction of Tikhonov-Type Regularization for Atmospheric Retrievals. *Atmosphere* **2020**, *11*, 1052. [[CrossRef](#)]
34. Rao, L.; Xu, J.; Efremenko, D.S.; Loyola, D.G.; Doicu, A. Hyperspectral Satellite Remote Sensing of Aerosol Parameters. In Preparation.
35. Molina García, V.; Sasi, S.; Efremenko, D.S.; Doicu, A.; Loyola, D. Radiative transfer models for retrieval of cloud parameters from EPIC/DSCOVER measurements. *J. Quant. Spectrosc. Radiat. Transf.* **2018**, *213*, 228–240. [[CrossRef](#)]
36. Molina García, V.; Sasi, S.; Efremenko, D.S.; Doicu, A.; Loyola, D. Linearized radiative transfer models for retrieval of cloud parameters from EPIC/DSCOVER measurements. *J. Quant. Spectrosc. Radiat. Transf.* **2018**, *213*, 241–251. [[CrossRef](#)]
37. Patterson, H.D.; Thompson, R. Recovery of inter-block information when block sizes are unequal. *Biometrika* **1971**, *58*, 545–554. [[CrossRef](#)]
38. Smyth, G.K.; Verbyla, A.P. A conditional likelihood approach to residual maximum likelihood estimation in generalized linear models. *J. R. Stat. Soc. Ser. B Methodol.* **1996**, *58*, 565–572. [[CrossRef](#)]
39. Stuart, A.; Ord, K.; Arnold, S. *Kendall's Advanced Theory of Statistics. Volume 2A: Classical Inference and the Linear Model*; John Wiley & Sons, Ltd.: Hoboken, NJ, USA, 2000.
40. Wahba, G. Practical approximate solutions to linear operator equations when the data are noisy. *SIAM J. Numer. Anal.* **1977**, *14*, 651–667. [[CrossRef](#)]
41. Wahba, G. *Spline Models for Observational Data*; SIAM: Philadelphia, PA, USA, 1990.
42. Meng, Z.; Yang, P.; Kattawar, G.W.; Bi, L.; Liou, K.; Laszlo, I. Single-scattering properties of tri-axial ellipsoidal mineral dust aerosols: A database for application to radiative transfer calculations. *J. Aerosol Sci.* **2010**, *41*, 501–512. [[CrossRef](#)]
43. Alados-Arboledas, L.; Müller, D.; Guerrero-Rascado, J.L.; Navas-Guzmán, F.; Pérez-Ramírez, D.; Olmo, F.J. Optical and microphysical properties of fresh biomass burning aerosol retrieved by Raman lidar, and star-and sun-photometry. *Geophys. Res. Lett.* **2011**, *38*. [[CrossRef](#)]
44. Waquet, F.; Cornet, C.; Deuzé, J.L.; Dubovik, O.; Ducos, F.; Goloub, P.; Herman, M.; Lapyonok, T.; Labonnote, L.C.; Riedi, J.; et al. Retrieval of aerosol microphysical and optical properties above liquid clouds from POLDER/PARASOL polarization measurements. *Atmos. Meas. Tech.* **2013**, *6*, 991–1016. [[CrossRef](#)]
45. Yu, X.; Shi, C.; Ma, J.; Zhu, B.; Li, M.; Wang, J.; Yang, S.; Kang, N. Aerosol optical properties during firework, biomass burning and dust episodes in Beijing. *Atmos. Environ.* **2013**, *81*, 475–484. [[CrossRef](#)]
46. Loyola, D.G.; Xu, J.; Heue, K.P.; Zimmer, W. Applying FP\_ILM to the retrieval of geometry-dependent effective Lambertian equivalent reflectivity (GE\_LER) daily maps from UVN satellite measurements. *Atmos. Meas. Tech.* **2020**, *13*, 985–999. [[CrossRef](#)]



# **C Aerosol Parameters Retrieval From TROPOMI/S5P Using Physics-Based Neural Networks**

Rao, L.; Xu, J.; Efremenko, D.S.; Loyola, D.G.; Doicu, A. Aerosol Parameters Retrieval From TROPOMI/S5P Using Physics-Based Neural Networks. *IEEE J. Sel. Top. Appl. Earth Obs.* 2022, 15, 6473-6484, doi: 10.1109/JSTARS.2022.3196843

# Aerosol Parameters Retrieval From TROPOMI/S5P Using Physics-Based Neural Networks

Lanlan Rao , Jian Xu , *Member, IEEE*, Dmitry S. Efremenko , Diego G. Loyola , *Senior Member, IEEE*, and Adrian Doicu

**Abstract**— In this article, we present three algorithms for aerosol parameters retrieval from TROPospheric Monitoring Instrument measurements in the O<sub>2</sub> A-band. These algorithms use neural networks 1) to emulate the radiative transfer model and a Bayesian approach to solve the inverse problem, 2) to learn the inverse model from the synthetic radiances, and 3) to learn the inverse model from the principal-component transform of synthetic radiances. The training process is based on full-physics radiative transfer simulations. The accuracy and efficiency of the neural network based retrieval algorithms are analyzed with synthetic and real data.

**Index Terms**—Aerosol information retrieval, neural networks, TROPospheric Monitoring Instrument/Sentinel-5 Precursor (TROPOMI/S5P).

## I. INTRODUCTION

AEROSOLS affect Earth's radiation budget by scattering and absorbing solar radiation (direct effect) and by influencing the cloud formation processes (indirect effect). Highly absorbing aerosols also have a warming effect on the atmosphere leading to the evaporation of cloud particles, which results in a reduction of the cloud cover (semidirect effect). Accurate assessments of aerosol properties, such as optical depth and layer height, are important for the global monitoring of air pollution in the lower atmosphere.

A number of passive satellite sensors enable to monitor aerosol properties on both regional and global scale using spectral information at various wavelengths. For instance, measurements in the O<sub>2</sub> A-band from the Global Ozone Mapping Experiment (GOME) [1] and GOME-2 [2], the Scanning Imaging Absorption Spectrometer for Atmospheric Chartography [3]–[5], the Greenhouse Gases Observing Satellite [6], and the TROPospheric Monitoring Instrument (TROPOMI) onboard

the Sentinel-5 Precursor (S5P) [7], [8] are used to retrieve aerosol optical depth and height information.

The inversion methods used in atmospheric remote sensing aim to recover atmospheric parameters by minimizing the residual between the measurements and the radiative transfer model simulations. The solution of the minimization problem can be found by using deterministic (e.g., Tikhonov-type regularization methods [9], [10]) or stochastic approaches (e.g., Bayesian methods [11]). In both cases, the computations of the forward model and the Jacobian matrix impose the performance bottleneck in the whole processing chain. Therefore, it would be problematic to adopt these approaches for the operational processing of remote sensing data from new-generation sensors. To tackle this problem, artificial neural networks, which are able to approximate very quickly any continuous function with a sufficiently high accuracy [12], [13] and to estimate the derivatives of the function with respect to the model inputs, can be used. Actually, a trained neural network may provide accurate estimates of the forward model and its Jacobian, in a fraction of time compared to classical retrieval algorithms. In atmospheric remote sensing, neural networks have already been widely applied. These techniques have been used 1) to approximate a radiative transfer model (or a part of it) [14]–[19], 2) to learn the inverse mappings [20]–[28], and 3) to recover some atmospheric retrieval parameters, which are then taken as initial guesses in an optimization approach [29], [30]. In this context, it should be pointed out that the two TROPOMI operational retrieval algorithms based on the O<sub>2</sub> A-band measurements use neural network based forward models together with a Bayesian approach for the retrieval of cloud properties [31], [32] and the aerosol layer height [33]–[35].

In this study, we present three types of neural networks for aerosol retrieval from TROPOMI measurements. The first one uses a neural network to emulate the radiative transfer model and a Bayesian approach to solve the inverse problem, the second one uses a neural network to learn the inverse model from the synthetic radiances, and the third one uses a neural network to learn the inverse model from the principal-component transform of synthetic radiances following the full-physics inverse learning machine method [18]. The major goal of this study is to incorporate the three neural network algorithms into a common tool and to analyze and compare their retrieval performances. To the best of our knowledge, such a comparison study had not been done before. This article is organized as follows. In Section II, we summarize the main features of the adopted radiative transfer

Manuscript received 7 May 2022; revised 29 June 2022; accepted 21 July 2022. Date of publication 5 August 2022; date of current version 17 August 2022. This work was supported in part by the German Aerospace Center (DLR) programmatic [Nachwuchsgruppe "Retrieval der nächsten Generation," 2 472 469] and in part by the Chinese Academy of Sciences (CAS) "Pioneering Initiative Talents Program" under Grant E1RC2WB2. The work of Lanlan Rao was supported by the Chinese Scholarship Council. (*Corresponding author: Jian Xu.*)

Lanlan Rao, Dmitry S. Efremenko, Diego G. Loyola, and Adrian Doicu are with the Remote Sensing Technology Institute, German Aerospace Center (DLR), 82234 Oberpfaffenhofen, Germany (e-mail: Lanlan.rao@dlr.de; Dmitry.Efremenko@dlr.de; diego.loyola@dlr.de; Adrian.Doicu@dlr.de).

Jian Xu is with the National Space Science Center, Chinese Academy of Sciences, Beijing 100190, China (e-mail: xujian@nssc.ac.cn).

Digital Object Identifier 10.1109/JSTARS.2022.3196843

model, while in Section III, we provide a detailed description of three physics-based algorithms using neural networks. In Section IV, the corresponding retrieval performances of the neural networks are analyzed using synthetic and real TROPOMI data.

## II. RADIATIVE TRANSFER MODEL

Any physics-based retrieval algorithm uses a model for computing the radiative transfer in a planetary atmosphere. We begin our analysis by summarizing the peculiarities of the radiative transfer model used in this study.

Regarding the TROPOMI instrument, each swath row (angle)  $r$  with  $r = 1, \dots, R$  is characterized by its own measurement wavelength grid  $\{\lambda_{mk}^r\}_{k=1}^{N_{m\lambda}}$  and slit function  $g_r$ , where  $R$  is the number of swath rows and  $N_{m\lambda}$  is the number of spectral grid points. This is due to the optics of the spectrometer (point spread function, aberrations, and defocusing) and small changes in the width of the slit. The noisy and synthetic radiances measured by the instrument at a wavelength  $\lambda_{mk}^r$  are given, respectively, by

$$I^\delta(\lambda_{mk}^r) = I(\lambda_{mk}^r) + \delta_{mk}, \quad (1)$$

$$I(\lambda_{mk}^r) = \int_{-\infty}^{\infty} g_r(\lambda_{mk}^r - \lambda) I(\lambda) d\lambda \quad (2)$$

where  $I(\lambda)$  is the radiance computed by a radiative transfer model at a forward wavelength  $\lambda$  (before convolution) and  $\delta_{mk}$  is the measurement noise. Note that because  $\{\lambda_{mk}^r\}_{k=1}^{N_{m\lambda}}$  changes slightly with  $r$ , we assumed that the noise is row independent.

The synthetic radiances  $I(\lambda)$  are computed on the forward wavelength grid by a radiative transfer model relying on the discrete ordinate method with matrix exponential [36], [37]. To speed up the computation, standard acceleration methods, i.e., the telescoping technique [38], [39] and the method of false discrete ordinate [40], as well as hyperspectral acceleration methods, i.e., the correlated  $k$ -distribution method [41] and principal component analysis [42]–[46], are employed. Line-by-line calculations [47] with optimized rational approximations for the Voigt line profile [48] (taken from HITRAN database [49]) are used to compute the absorption cross sections of gas molecules; the methodology described in [50] is used to compute the Rayleigh cross section and depolarization ratios, and the delta-M approximation [51] and truncated multiple and single scattering approximation correction [52], [53] are used in the radiative transfer calculations. The radiative transfer model includes several types of bidirectional reflectance distribution functions over land and water. However, in order to simplify the analysis, we use the geometry-dependent effective Lambertian equivalent reflectivity (GE\_LER) product that accounts for satellite viewing dependencies and improves the accuracy for actual snow/ice conditions [28].

The aerosol optical depth  $\tau$  and layer height  $H$  are retrieved in the oxygen absorption A-band (758–771 nm). To simplify the analysis, the aerosol layer is assumed to be homogeneous with a fixed thickness of 0.5 km, spreading evenly from  $H - 0.25$  km to the  $H + 0.25$  km. As the considered spectral range is narrow, the aerosol optical depth is assumed to be constant within this range. There are five sets of aerosol microphysical models [54], [55] included in the radiative transfer model under the assumption

of spherical aerosol particles. For example, the set of aerosol models taken from the MODIS dark-target (DT) algorithm [56] includes nonabsorbing, moderately absorbing, and absorbing aerosols, as well as desert dust. Each aerosol model is characterized by a bi-mode log-normal volume size distribution comprising a fine and a coarse mode. Specifically, in the case of moderately absorbing aerosols, the median radius of the fine mode  $r_{vf}$  and the coarse mode  $r_{vc}$  are given, respectively, by

$$r_{vf} = 0.145 + 0.0203\tau \text{ (}\mu\text{m)}, \quad (3)$$

$$r_{vc} = 3.101 + 0.3364\tau \text{ (}\mu\text{m)} \quad (4)$$

the standard deviations of the fine mode  $s_f$  and the coarse mode  $s_c$  by

$$s_f = 0.374 + 0.1365\tau, \quad (5)$$

$$s_c = 0.729 + 0.098\tau \quad (6)$$

the volumes of the fine mode particles  $V_{of}$  and the coarse mode particles  $V_{oc}$  by

$$V_{of} = 0.1642\tau^{0.775} \text{ (}\mu\text{m}^3/\mu\text{m}^2), \quad (7)$$

$$V_{oc} = 0.1482\tau^{0.684} \text{ (}\mu\text{m}^3/\mu\text{m}^2) \quad (8)$$

and the complex refractive index of the aerosol particles by  $m = (1.43, 0.008 - 0.002\tau)$ . Note that the parameters of the size distribution and the refractive index depend on the aerosol optical depth  $\tau$ . In addition, the radiative transfer model can deal with several types of aerosol profiles, such as, for example, elevated box, exponential decay, a combination of exponential decay and ground box, and Gaussian.

## III. NEURAL NETWORK ALGORITHMS

In this study, neural networks are employed 1) to emulate the radiative transfer model and 2) to learn the inverse model. The first one is referred to as a neural network for the forward operator, whereas the second one as a neural network for the inverse operator.

The neural network approach can be briefly summarized as follows. Let us consider the model  $\mathbf{y} = \mathbf{F}(\mathbf{x})$ , where  $\mathbf{x} \in \mathbb{R}^{N_x}$  and  $\mathbf{y} \in \mathbb{R}^{N_y}$  are the input and output vectors, respectively, and  $\mathbf{F}$  is a deterministic function. In order to approximate  $\mathbf{F}(\mathbf{x})$  by a neural network model  $\mathbf{f}(\mathbf{x}, \boldsymbol{\omega})$  with parameters  $\boldsymbol{\omega}$ , we consider a deep neural network consisting of units (nodes) arranged in an input, output, and several hidden layers. For a neural network with  $L + 1$  layers and  $N_l$  units in layer  $l$ , the feed-forward operations read as

$$\mathbf{u}_l = \mathbf{W}_l \mathbf{y}_{l-1} + \mathbf{b}_l, \quad (9)$$

$$\mathbf{y}_l = \phi_l(\mathbf{u}_l), \quad l = 1, \dots, L \quad (10)$$

where  $l = 0$  and  $l = L$  are the input and output layers, respectively,  $\phi_l$  is the layer activation function,  $\mathbf{W}_l \in \mathbb{R}^{N_l \times N_{l-1}}$  is the matrix of weights connecting the layers  $l - 1$  and  $l$ ,  $\mathbf{b}_l \in \mathbb{R}^{N_l}$  is the vector of biases corresponding to layer  $l$ , and  $\boldsymbol{\omega} = \{\mathbf{W}_l, \mathbf{b}_l\}_{l=1}^L$  is the set of network parameters. In the input and output layers, we have  $\mathbf{y}_0 = \mathbf{x}$  and  $\mathbf{y}_L = \mathbf{f}(\mathbf{x}, \boldsymbol{\omega})$ , respectively, so that  $N_0 = N_x$  and  $N_L = N_y$ . Deep learning

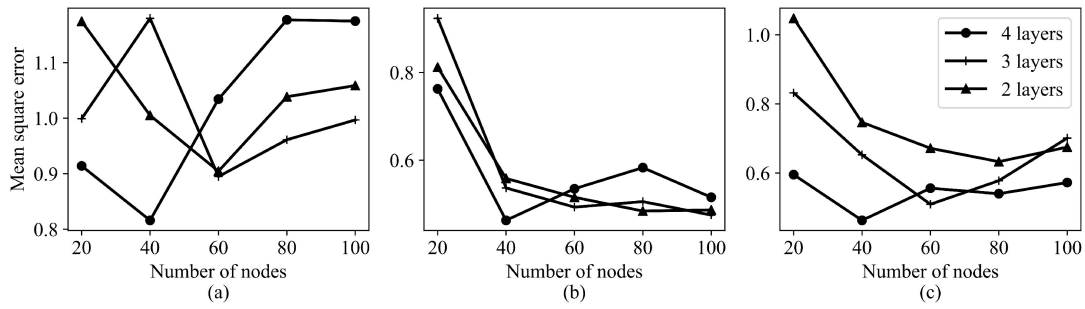


Fig. 1. Mean square error for the validation data set. The results correspond to all neural networks considered in this study, i.e., (a) the forward-operator neural network, (b) the inverse-operator neural network with synthetic radiances, and (c) inverse-operator neural network with the principal-component transform of synthetic radiances.

TABLE I  
VALUE RANGES OF THE OPTICAL AND GEOMETRICAL PARAMETERS  
FOR GENERATING THE DATA SET

Parameter	Value range
$\tau$	0.05 – 5
$H$	0.1 – 15.75 km
$\theta_0$	0 – 75°
$\theta$	0 – 70°
$\Delta\varphi$	0 – 180°
$H_s$	0 – 2.61 km
$A_s$	0 – 0.4

is the process of computing the network parameters  $\omega$  on a data set  $\mathcal{D} = \{(\mathbf{x}^{(n)}, \mathbf{y}^{(n)})\}_{n=1}^N$ , where  $\mathbf{y}^{(n)} = \mathbf{F}(\mathbf{x}^{(n)})$  and  $N$  is the number of samples. This is done by computing a point estimate  $\hat{\omega}$  as the minimizer of a loss function with a penalty term controlling the amplitudes of the network parameters, i.e.,  $E(\omega) = \frac{1}{2} \sum_{n=1}^N \|\mathbf{y}^{(n)} - \mathbf{f}(\mathbf{x}^{(n)}, \omega)\|^2 + \gamma \|\omega\|^2$ , where  $\gamma$  is the regularization parameter.

To formulate the retrieval problem, we group the optical and geometrical parameters into the following:

- 1) retrieval parameters, which include the aerosol optical depth  $\tau$  at 760 nm and aerosol layer height  $H$  defined as the middle height of an aerosol layer with a fixed thickness of 0.5 km;
- 2) forward model parameters, which include the solar zenith angle  $\theta_0$ , viewing zenith angle  $\theta$ , relative azimuth angle  $\Delta\varphi$ , surface height  $H_s$ , and surface albedo  $A_s$  (note that the forward model parameters are not part of the retrieval).

For generating the data set, samples of optical and geometrical parameters are produced by means of a smart sampling technique [57] based on Halton sequences [58]; their ranges of variations are shown in Table I. The neural networks are trained for the moderately absorbing aerosol model from the MODIS DT algorithm.

The hyperparameters of the neural network, i.e., the number of hidden layers and the number of units in each layer, are optimized by using 10% of the samples from the training set for validation. In the validation stage, the holdout cross-validation and a grid search procedure are used; the grid search is performed over the sets  $\{2, 3, 4\}$  of hidden layers and  $\{20, 40, 60, 80, 100\}$  of layer units. For all neural network considered in this study, a network architecture with 4 hidden layers and 40 units in each layer yields the lowest mean-square error on the validation data set

(see Fig. 1), and no overfitting has been observed. A hyperbolic tangent activation function is taken, and as optimization tool, the mini-batch gradient descent with adaptive moment estimation [59] is utilized.

#### A. Neural Network for the Forward Operator

For emulating the radiative transfer model, we consider a neural network in which the input  $\mathbf{x}$  is the set of optical and geometrical parameters, while the output  $\mathbf{y}$  is the set of synthetic radiances  $I(\lambda_k)$  computed on the forward wavelength grid  $\{\lambda_k\}_{k=1}^{N_\lambda}$ , i.e.,

$$\mathbf{x} = \begin{bmatrix} [\tau, H]^T \\ [\theta_0, \theta, \Delta\varphi, H_s, A_s]^T \end{bmatrix} \mapsto \mathbf{y} = [I(\lambda_k)]_{k=1}^{N_\lambda}. \quad (11)$$

Thus, the dimensions of the input and output vectors are  $N_x = 7$  and  $N_y = N_\lambda$ , respectively.

The forward wavelength grid consists of  $N_\lambda = 465$  equidistant spectral points ranging from 757.4 to 771.6 nm, while the number of combinations of optical and geometrical parameters defined by Table I. After the radiative transfer model is learned, the synthetic radiances computed at a high spectral resolution on the forward wavelength grid  $I(\lambda_k)$  are convolved with a slit function  $g_r$  to obtain the synthetic radiances on the measurement wavelength grid  $I(\lambda_{mk}^r)$  [cf. (2)].

The retrieval of aerosol parameters encapsulated now in the state vector  $\mathbf{x} = [\tau, H]^T$  requires the solution of the nonlinear equation

$$\mathbf{y}^\delta = \mathbf{F}(\mathbf{x}) + \delta_m \quad (12)$$

where, for a given swath row  $r$ ,  $\mathbf{y}^\delta = [I_{\text{mes}}(\lambda_{mk}^r)]_{k=1}^{N_{m\lambda}}$  is the measurement vector,  $\mathbf{F}(\mathbf{x}) = [I(\lambda_{mk}^r)]_{k=1}^{N_{m\lambda}}$  is the forward model, and  $\delta_m = [\delta_{mk}]_{k=1}^{N_{m\lambda}}$  is the measurement noise vector. The nonlinear equation (12) is solved by using a Bayesian approach [11]. In this approach, the *a posteriori* density  $p(\mathbf{x} | \mathbf{y}^\delta)$  representing the conditional probability density of the state vector  $\mathbf{x}$  given the data  $\mathbf{y}^\delta$  is the quantity of interest. Assuming that

- 1) the measurement noise vector  $\delta_m$  is a Gaussian random vector with zero mean and noise covariance matrix  $\mathbf{C}_m = \text{diag}[\sigma_{mk}^2]_{k=1}^{N_{m\lambda}} = \sigma_m^2 \mathbf{C}_m$ , i.e.,  $\delta_m \sim \mathcal{N}(\mathbf{0}, \mathbf{C}_m)$ , where  $\sigma_m^2 = \sum_{k=1}^{N_{m\lambda}} \sigma_{mk}^2$  is the noise variance, and



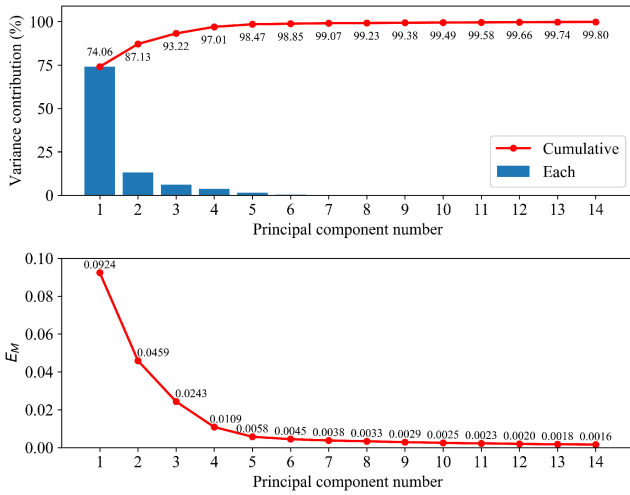


Fig. 2. Contribution of the first 14 principal components to the total variance (top) and reconstruction error with respect to the first 14 principal components (bottom).

2) the state vector  $\mathbf{x}$  is a Gaussian random vector with mean  $\mathbf{x}_a$  and *a priori* covariance matrix  $\mathbf{C}_x = \text{diag}[\sigma_{xk}^2]_{k=1}^{N_x} = \sigma_x^2 \mathbf{C}_x$ , i.e.,  $\mathbf{x} \sim \mathcal{N}(\mathbf{x}_a, \mathbf{C}_x)$ , where  $\mathbf{x}_a$  is the *a priori* state vector,  $\sigma_x^2 = \sum_{k=1}^{N_x} \sigma_{xk}^2$  is the *a priori* state variance, and the notation  $\mathcal{N}(\mathbf{x}_{\text{mean}}, \mathbf{C}_x)$  symbolizes a normal distribution with mean  $\mathbf{x}_{\text{mean}}$  and covariance matrix  $\mathbf{C}_x$ ,

we find

$$p(\mathbf{x} | \mathbf{y}^\delta) \propto \exp\left[-\frac{1}{2} V_\alpha(\mathbf{x} | \mathbf{y}^\delta)\right] \quad (13)$$

where

$$V_\alpha(\mathbf{x} | \mathbf{y}^\delta) = \frac{1}{\sigma_m^2} \left\{ [\mathbf{y}^\delta - \mathbf{F}(\mathbf{x})] \mathbf{C}_m^{-1} [\mathbf{y}^\delta - \mathbf{F}(\mathbf{x})]^T + \alpha (\mathbf{x} - \mathbf{x}_a) \mathbf{C}_x^{-1} (\mathbf{x} - \mathbf{x}_a)^T \right\} \quad (14)$$

is the *a posteriori* potential and  $\alpha = \sigma_m^2 / \sigma_x^2$  is the regularization parameter. Here, the notation  $\mathcal{N}(\mathbf{x}_{\text{mean}}, \mathbf{C}_x)$  stands for a normal distribution with mean  $\mathbf{x}_{\text{mean}}$  and covariance matrix  $\mathbf{C}_x$ . The maximum *a posteriori* estimate  $\hat{\mathbf{x}}_\alpha^\delta$  is then computed as

$$\hat{\mathbf{x}}_\alpha^\delta = \arg \min_{\mathbf{x}} V_\alpha(\mathbf{x} | \mathbf{y}^\delta). \quad (15)$$

After scaling the data model (12) with the matrix  $\mathbf{P} = \mathbf{C}_m^{-1/2}$ , i.e., after transforming the data model into a model with white noise and introducing the regularization matrix  $\mathbf{L}$  through the Cholesky factorization  $\mathbf{C}_x^{-1} = \mathbf{L}^T \mathbf{L}$ , we are led to a least-squares problem which is solved by using the iteratively regularized Gauss–Newton method [60]. This method supplies the optimal value of the regularization parameter and the corresponding regularized solution.

In the inversion step, the noise covariance matrix is chosen as  $\mathbf{C}_m = \text{diag}[\sigma_{mk}^2]_{k=1}^{N_{m\lambda}}$  with  $\sigma_{mk} = 0.02 \times \bar{I}(\lambda_{mk})$  and  $\bar{I}(\lambda_{mk}) = (1/R) \sum_{r=1}^R I(\lambda_{mk}^r)$  for all  $k = 1, \dots, N_{m\lambda}$ , and the *a priori* covariance matrix as  $\mathbf{C}_x = \text{diag}[\sigma_{xk}^2]_{k=1}^2$  with  $\sigma_{xk} = 0.2 \times x_k$  and  $x_k$  standing for  $\tau$  and  $H$ .

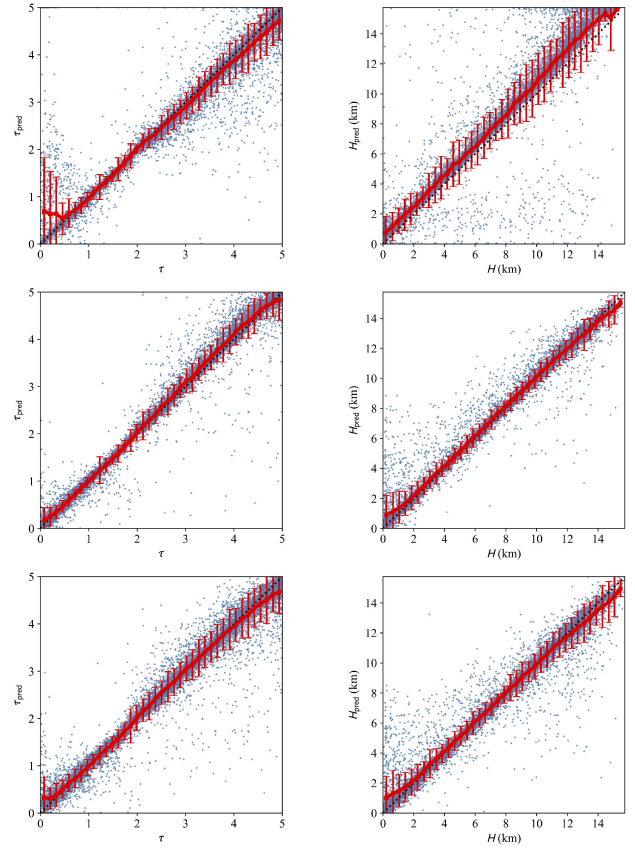


Fig. 3. Predictions of the forward-operator neural network (top row), the inverse-operator neural networks with synthetic radiances (middle row), and the principal-component transform of synthetic radiances (bottom row), respectively. The plots show the predicted values  $x_{\text{pred}}^{(n)}$  (blue points) together with the mean  $\mathbb{E}_j(x_{\text{pred}})$  (red points) and standard deviation  $\sqrt{\mathbb{E}_j(x_{\text{pred}} - \mathbb{E}_j(x_{\text{pred}}))^2}$  (red error bars) over all samples  $x_{\text{pred}}^{(n)}$ , whose corresponding  $x^{(n)}$  are in the  $j$ th bin. The value range of  $x$  is split into  $N_b = 40$  bins, where  $x$  stands for the aerosol optical depth  $\tau$  and aerosol layer height  $H$ .

## B. Neural Network for the Inverse Operator

For solving the inverse problem, we designed two types of neural networks following the full-physics inverse learning machine method. The first one uses as input the synthetic radiances computed on the measurement wavelength grid, while the second one uses as input the principal-component transform of synthetic radiances.

1) *Neural Network for the Inverse Operator With Synthetic Radiances*: In principle, for emulating the inverse model, we may use a neural network in which the input  $\mathbf{x}$  includes the noisy radiances on a measurement wavelength grid and the forward model parameters, while the output  $\mathbf{y}$  includes the set of parameters to be retrieved, i.e.,

$$\mathbf{x} = \left[ [I(\lambda_{mk}^r) + \delta_{mk}]_{k=1}^{N_{m\lambda}}, [\theta_0, \theta, \Delta\varphi, H_s, A_s]^T \right] \mapsto \mathbf{y} = [\tau, H]^T. \quad (16)$$

In this case, the dimensions of the input and output vectors are  $N_x = N_{m\lambda} + 5$  and  $N_y = 2$ , respectively.

The problem which appears is that because we are dealing with a random measurement noise and a set of measurement

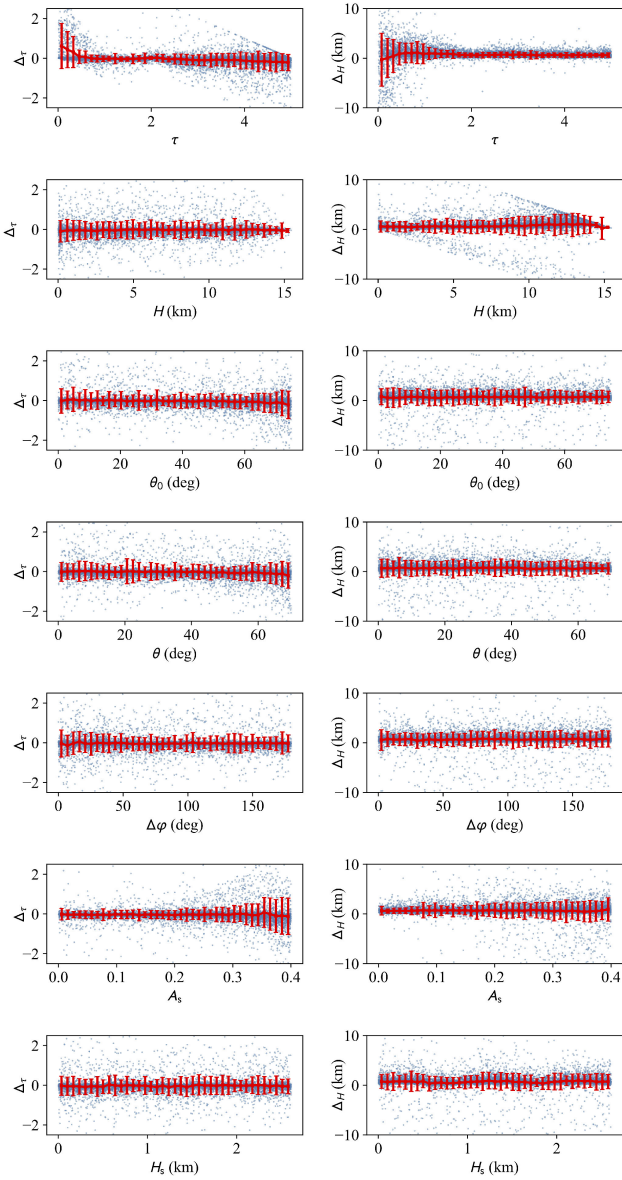


Fig. 4. Absolute error in the retrieved aerosol optical depth  $\Delta\tau$  and aerosol layer height  $\Delta H$  versus the optical and geometrical parameters  $b$ , where  $b$  stands for the aerosol optical depth  $\tau$ , aerosol layer height  $H$ , solar zenith angle  $\theta_0$ , viewing zenith angle  $\theta$ , relative azimuth angle  $\Delta\varphi$ , surface albedo  $A_s$ , and surface height  $H_s$ . The results correspond to the forward-operator neural network. The plots show the absolute error  $\Delta x^{(n)}$  (blue points) together with the mean  $\mathbb{E}_j(\Delta x)$  (red points) and standard deviation  $\sqrt{\mathbb{E}_j([\Delta x - \mathbb{E}_j(\Delta x)]^2)}$  (red error bars) over all samples  $\Delta x^{(n)}$ , whose corresponding  $b^{(n)}$  are in the  $j$ th bin. The interval of variation of each parameter  $b$  is split into  $N_b = 40$  bins.

wavelength grids, the same output corresponds to different realization of the random noise as well as to different wavelength grids. To reduce the dimension of the data set, we use the jitter approach under the assumption that the measurement wavelength grid  $\{\lambda_{mk}^r\}_{k=1}^{N_{m\lambda}}$  is a discrete random variable which can take the values  $\{\lambda_{mk}^1\}_{k=1}^{N_{m\lambda}}, \dots, \{\lambda_{mk}^R\}_{k=1}^{N_{m\lambda}}$ . According to this approach, at each forward pass through the network, a measurement wavelength grid  $\{\lambda_{mk}^r\}_{k=1}^{N_{m\lambda}}$  is randomly selected from the  $R$  wavelength grids, and a new random noise  $\delta_{mk} \sim \mathcal{N}(0, \sigma_{mk}^2)$

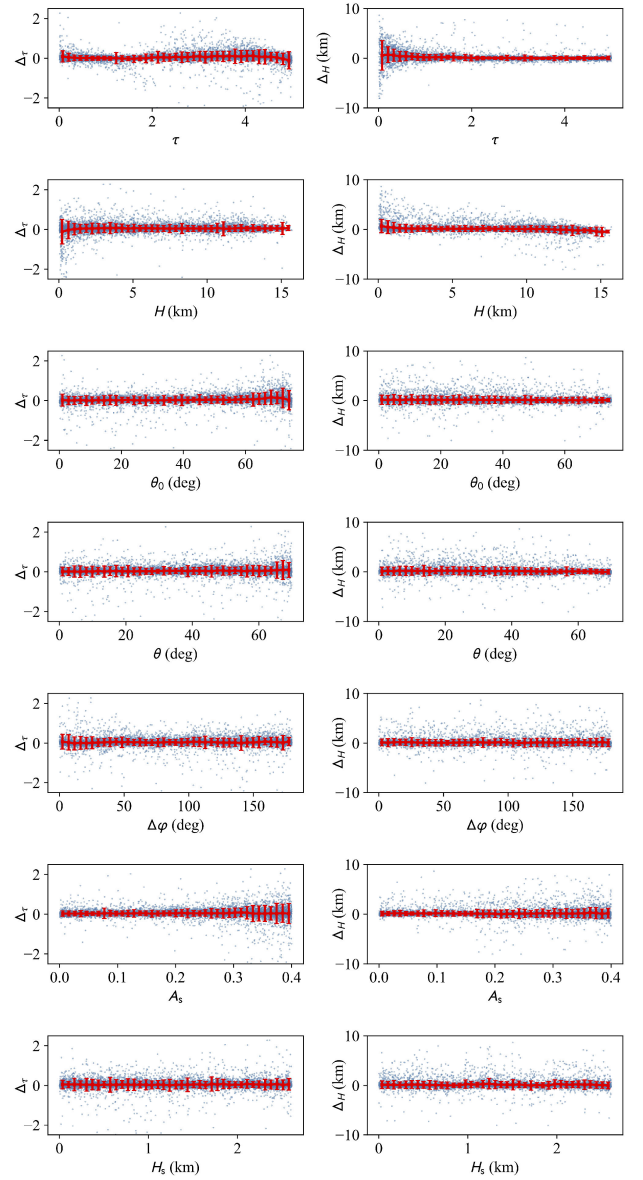


Fig. 5. Same as in Fig. 4 but for the inverse-operator neural network with synthetic radiances.

is added to the synthetic radiance  $I(\lambda_{mk}^r)$ . In other words, the input sample is different every time it is passed through the network.

In the training stage, the number of swath rows is  $R = 448$ , the number of points in each measurement wavelength grid is  $N_{m\lambda} = 131$ , and the measurement wavelength grids are chosen from the TROPOMI Level-1 product, e.g.,

$$\begin{aligned} \{\lambda_{mk}^1\}_{k=1}^{N_{m\lambda}} &= \{755.120, \dots, 770.929 \text{ nm}\}, \\ \{\lambda_{mk}^2\}_{k=1}^{N_{m\lambda}} &= \{755.133, \dots, 770.942 \text{ nm}\} \\ &\vdots \\ \{\lambda_{mk}^R\}_{k=1}^{N_{m\lambda}} &= \{755.264, \dots, 771.071 \text{ nm}\}. \end{aligned} \quad (17)$$

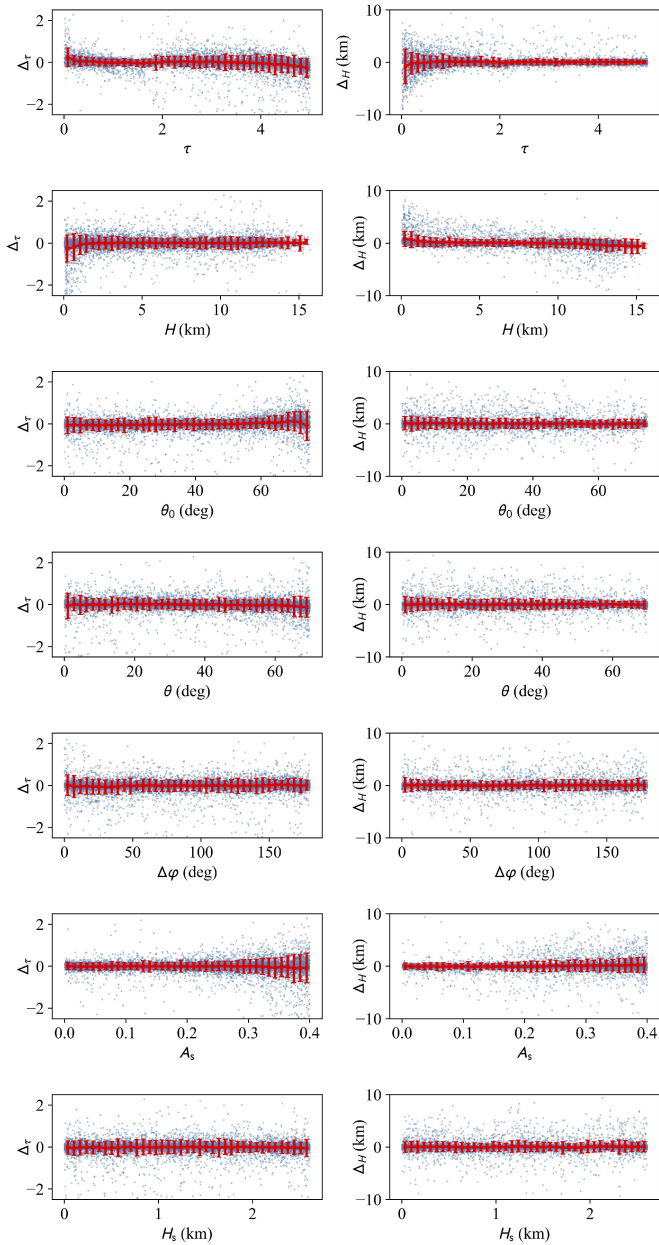


Fig. 6. Same as in Fig. 4 but for the inverse-operator neural network with the principal-component transform of synthetic radiances.

As before, the noisy spectra are generated with the noise standard deviation  $\sigma_{mk} = 0.02 \times \bar{I}(\lambda_{mk})$ , where  $\bar{I}(\lambda_{mk}) = (1/R) \sum_{r=1}^R I(\lambda_{mk}^r)$ . The number of samples in the training set is  $N_t = 404901$ , where each sample consists of a set of optical and geometrical parameters and the corresponding synthetic radiances computed on all measurement wavelength grids  $\{\lambda_{mk}^r\}_{k=1}^{N_{m\lambda}}, r = 1, \dots, R$ .

2) *Neural Network for the Inverse Operator With the Principal-Component Transform of Synthetic Radiances:* To reduce the dimension of the synthetic radiance vector  $\mathbf{i}_m = [I(\lambda_{mk}^r)]_{k=1}^{N_{m\lambda}} \in \mathbb{R}^{N_{m\lambda}}$ , the principal-component analysis is applied. Here, the dependency of  $\mathbf{i}_m$  on the swath row  $r$  is implicitly assumed. For the  $N_t$ -dimensional data set  $\{\mathbf{i}_m^{(n)}\}_{n=1}^{N_t}$ , let  $\langle \mathbf{i}_m \rangle =$

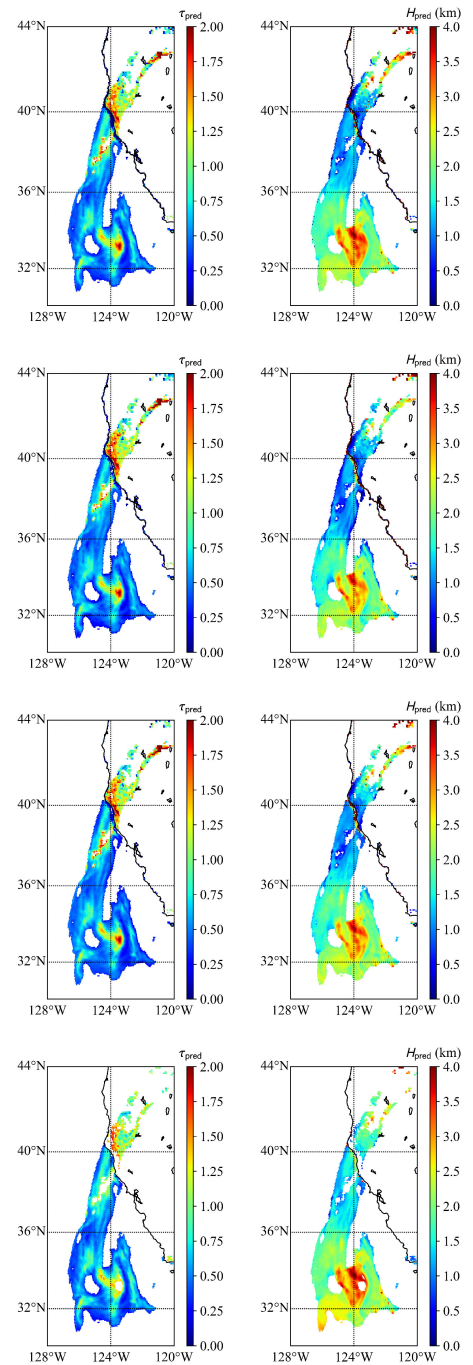


Fig. 7. Retrieved aerosol optical depth  $\tau_{\text{pred}}$  and aerosol layer height  $H_{\text{pred}}$  from TROPOMI measurements recorded on December 12, 2017 in California. The results correspond to the forward-operator neural network (the first row), the inverse-operator neural network with synthetic radiances (the second row), the inverse-operator neural network with the principal-component transform of synthetic radiances (the third row), and the TROPOMI official operational product (the fourth row).

$(1/N_t) \sum_{n=1}^{N_t} \mathbf{i}_m^{(n)}$  be the sample mean of the data. The goal of a linear embedding method is to find an  $M$ -dimensional subspace ( $M < N_{m\lambda}$ ) spanned by a set of linear independent vectors  $\{\mathbf{u}_k\}_{k=1}^M$ , such that the centered data  $\mathbf{i}_m^{(n)} - \langle \mathbf{i}_m \rangle$  belongs mainly to this subspace, i.e.,  $\mathbf{i}_m^{(n)} \approx \langle \mathbf{i}_m \rangle + \mathbf{U}_M \hat{\mathbf{i}}_m^{(n)}$ , where  $\mathbf{U}_M =$

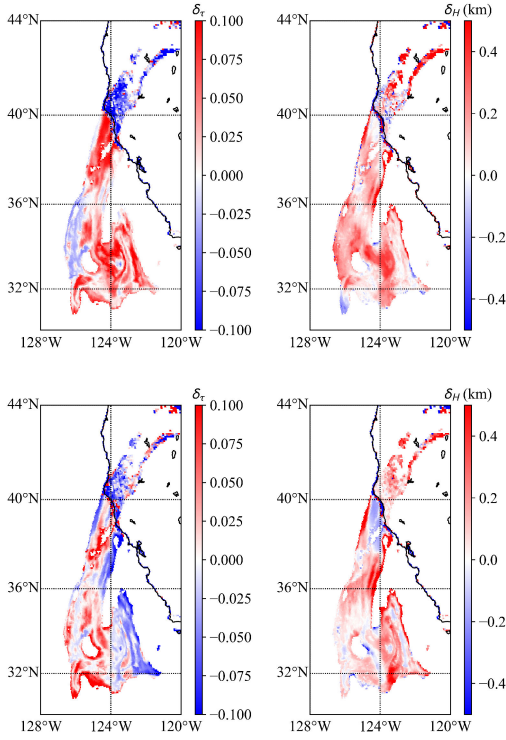


Fig. 8. Absolute differences in the retrieved aerosol optical depth  $\delta_\tau$  and aerosol layer height  $\delta_H$  corresponding to the forward-operator neural network (upper panels) and the inverse-operator neural network with the principal-component transform of synthetic radiances (lower panels). The results provided by the inverse-operator neural network with synthetic radiances are taken as a reference.

$\{\mathbf{u}_1, \dots, \mathbf{u}_M\} \in \mathbb{R}^{N_{m\lambda} \times M}$  is the (inverse) mapping from the low- to the high-dimensional space. The dimensionality-reduced input is then  $\hat{\mathbf{i}}_m^{(n)} = \mathbf{U}_M^\dagger (\mathbf{i}_m^{(n)} - \langle \mathbf{i}_m \rangle) \in \mathbb{R}^M$ , where the (forward) mapping from the high- to the low-dimensional space  $\mathbf{U}_M^\dagger \in \mathbb{R}^{M \times N_{m\lambda}}$  is the pseudoinverse of  $\mathbf{U}_M$ , i.e.,  $\mathbf{U}_M^\dagger \mathbf{U}_M = \mathbf{I}_M$ . In the principal-component analysis, the transformation matrix  $\mathbf{U}_M$  is computed as follows: 1) stack all centered data  $\mathbf{i}_m^{(n)} - \langle \mathbf{i}_m \rangle$  into the columns of the matrix  $\mathcal{I}$ , i.e.,  $\mathcal{I} = [\mathbf{i}_m^{(1)} - \langle \mathbf{i}_m \rangle, \dots, \mathbf{i}_m^{(N_t)} - \langle \mathbf{i}_m \rangle] \in \mathbb{R}^{N_{m\lambda} \times N_t}$ , 2) compute the covariance matrix  $\mathbf{C} = (1/N_t) \mathcal{I} \mathcal{I}^T \in \mathbb{R}^{N_{m\lambda} \times N_{m\lambda}}$  and a singular value decomposition of  $\mathbf{C}$ , i.e.,  $\mathbf{C} = \mathbf{U} \mathbf{\Sigma} \mathbf{U}^T$ , where  $\mathbf{\Sigma} = \text{diag}[\sigma_k]_{k=1}^{N_{m\lambda}}$  is the diagonal matrix of the singular values  $\sigma_1 > \sigma_2 > \dots > \sigma_{N_{m\lambda}} > 0$  and  $\mathbf{U} = [\mathbf{u}_1, \dots, \mathbf{u}_{N_{m\lambda}}] \in \mathbb{R}^{N_{m\lambda} \times N_{m\lambda}}$  is the orthogonal matrix of the singular vectors, and 3) take the inverse transformation matrix as  $\mathbf{U}_M = [\mathbf{u}_1, \dots, \mathbf{u}_M] \in \mathbb{R}^{N_{m\lambda} \times M}$ . In this case, the forward transformation matrix is  $\mathbf{U}_M^\dagger = \mathbf{U}_M^T \in \mathbb{R}^{M \times N_{m\lambda}}$ . The number of principal components  $M$  is determined by monitoring the reconstruction error

$$E_M = \sum_{n=1}^{N_t} \|(\mathbf{i}_m^{(n)} - \langle \mathbf{i}_m \rangle) - \mathbf{U}_M \hat{\mathbf{i}}_m^{(n)}\|_2^2 = \sum_{n=1}^{N_t} \|(\mathbf{i}_m^{(n)} - \langle \mathbf{i}_m \rangle) - \mathbf{U}_M \mathbf{U}_M^\dagger (\mathbf{i}_m^{(n)} - \langle \mathbf{i}_m \rangle)\|_2^2 \quad (18)$$

as a function of  $M$ , and the  $M$ , for which  $E_M$  is below a prescribed tolerance, is chosen. For the noisy radiance vector  $\mathbf{i}_m^\delta = \mathbf{i}_m + \delta_m$ , where  $\delta_m \sim \mathcal{N}(\mathbf{0}, \mathbf{C}_m)$  is the measurement noise vector, we find  $\langle \mathbf{i}_m^\delta \rangle = \langle \mathbf{i}_m \rangle$ , yielding  $\hat{\mathbf{i}}_m^\delta = \mathbf{U}_M^T (\mathbf{i}_m^\delta - \langle \mathbf{i}_m \rangle) = \hat{\mathbf{i}}_m + \hat{\delta}_m$  with  $\hat{\delta}_m = \mathbf{U}_M^T \delta_m \sim \mathcal{N}(\mathbf{0}, \hat{\mathbf{C}}_m)$  and  $\hat{\mathbf{C}}_m = \mathbf{U}_M^T \mathbf{C}_m \mathbf{U}_M \in \mathbb{R}^{M \times M}$ .

Thus, instead of the synthetic radiances  $\mathbf{i}_m \in \mathbb{R}^{N_{m\lambda}}$ , the input of the neural network is the principal-component transform of synthetic radiances  $\hat{\mathbf{i}}_m \in \mathbb{R}^M$ , and during each forward pass through the network, the random noise  $\hat{\delta}_m \sim \mathcal{N}(\mathbf{0}, \hat{\mathbf{C}}_m)$  is added to  $\hat{\mathbf{i}}_m$ . To simplify the computation, we approximate the dimensionality-reduced noise covariance matrix  $\hat{\mathbf{C}}_m$  by a diagonal matrix, i.e.,  $\hat{\mathbf{C}}_m \approx \text{diag}[\hat{C}_{mkk}]_{k=1}^M$ , where  $\hat{C}_{mij}$  are the entries of  $\hat{\mathbf{C}}_m$ . Through a numerical analysis, we found that for  $M = 14$ ,  $E_M < 1.6 \times 10^{-3}$ ; thus, 14 principal components appear to be sufficient for aerosol retrieval. The corresponding analysis results for the first 14 principal components are shown in Fig. 2.

It should be pointed out that the number of principal components  $M$  can be also determined by reducing the measurement noise [29]. In this case, the reconstruction error  $E_M = \sum_{n=1}^{N_t} \|(\mathbf{i}_m^{(n)} - \langle \mathbf{i}_m \rangle) - \mathbf{U}_M \mathbf{U}_M^\dagger (\mathbf{i}_m^{(n)} - \langle \mathbf{i}_m \rangle)\|_2^2$  is monitored, and the  $M$  that minimizes  $E_M$  is chosen.

#### IV. RESULTS AND DISCUSSION

In this section, we analyze the retrieval performances of the neural network retrieval algorithms using the synthetic and real TROPOMI data.

##### A. Synthetic Retrieval

To test the performances of the retrieval algorithms on synthetic data, we consider a prediction or a test set consisting of  $N_p = 11\,868$  samples

$$(\tau^{(n)}, H^{(n)}, \theta_0^{(n)}, \theta^{(n)}, \Delta\varphi^{(n)}, H_s^{(n)}, A_s^{(n)}) \quad (19)$$

chosen randomly in their assumed intervals of variation. For the forward-operator neural network equipped with a Bayesian approach, the initial and *a priori* values for the aerosol optical depth and layer height are 2 and 2 km, respectively.

To interpret the results, we split the interval of variation of  $x$ ,  $[x_{\min}, x_{\max}]$ , where  $x$  stands for  $\tau$  and  $H$ , into  $N_b = 40$  equidistant bins, i.e.,  $[x_{\min}, x_{\max}] = \cup_{j=1}^{N_b} B_{xj}$ , and compute the (bin) mean

$$\mathbb{E}_j(x_{\text{pred}}) = \frac{1}{N_{xj}} \sum_{n, \text{ s.t. } x^{(n)} \in B_{xj}} x_{\text{pred}}^{(n)} \quad (20)$$

and standard deviation

$$\begin{aligned} & \sqrt{\mathbb{E}_j([x_{\text{pred}} - \mathbb{E}_j(x_{\text{pred}})]^2)} \\ &= \sqrt{\frac{1}{N_{xj}} \sum_{n, \text{ s.t. } x^{(n)} \in B_{xj}} [x_{\text{pred}}^{(n)} - \mathbb{E}_j(x_{\text{pred}})]^2} \end{aligned} \quad (21)$$

over all  $N_{xj}$  samples  $x_{\text{pred}}^{(n)}$ , whose corresponding  $x^{(n)}$  are in  $B_{xj}$ . To quantify the retrieval accuracy, we use the first two moments of the absolute error over the prediction set  $\Delta_x = x_{\text{pred}} - x$ ,

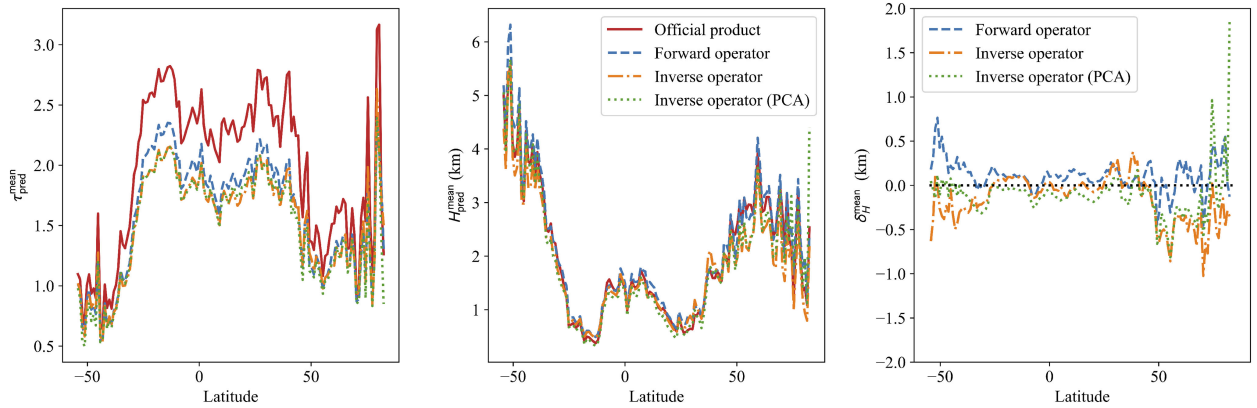


Fig. 9. Mean values of aerosol optical depth  $\tau^{\text{mean}}_{\text{pred}}$  and aerosol layer height  $H^{\text{mean}}_{\text{pred}}$  for the three neural network based algorithms and the official operational product. The mean differences in aerosol layer height against the official operational product are plotted as  $\delta_H^{\text{mean}}$ . The comparison is done for the entire year of 2019.

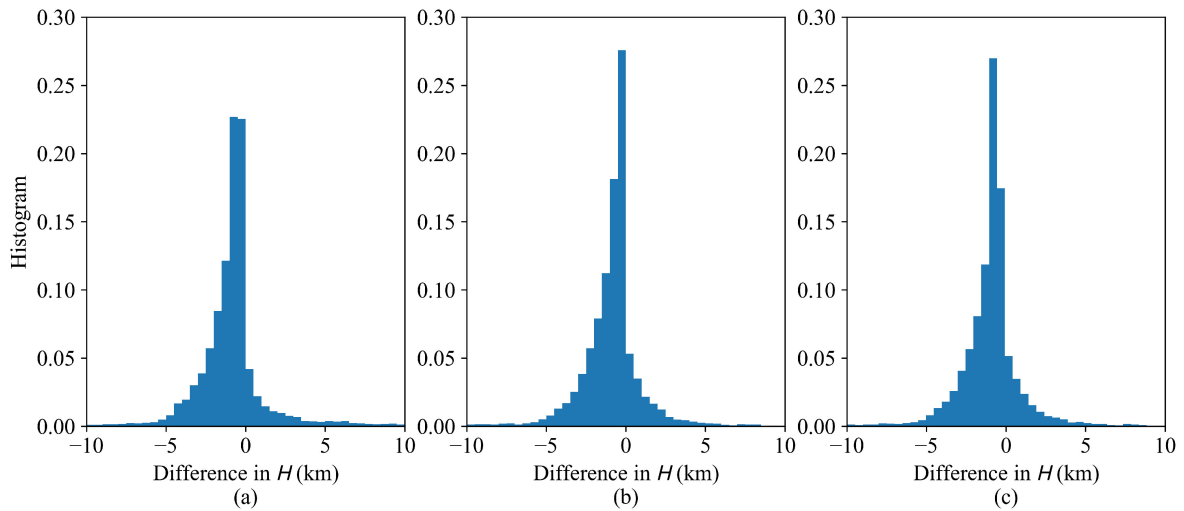


Fig. 10. Histogram of the differences in aerosol layer height  $H$  between the TROPOMI and CALIPSO products. The TROPOMI retrievals are generated by (a) forward-operator neural network, (b) inverse-operator neural networks with synthetic radiances, and (c) principal-component transform of synthetic radiances.

where  $x_{\text{pred}}$  and  $x$  are the predicted and true values, respectively. These are the mean absolute error

$$\mathbb{E}(|\Delta_x|) = \frac{1}{N_p} \sum_{n=1}^{N_p} |\Delta_x^{(n)}| \quad (22)$$

and the standard deviation of the absolute error

$$\sqrt{\mathbb{E}([\Delta_x - \mathbb{E}(\Delta_x)]^2)} = \sqrt{\frac{1}{N_p} \sum_{n=1}^{N_p} [\Delta_x^{(n)} - \mathbb{E}(\Delta_x)]^2}. \quad (23)$$

In Fig. 3, the mean  $\mathbb{E}_j(x_{\text{pred}})$  and standard deviation  $\sqrt{\mathbb{E}_j([x_{\text{pred}} - \mathbb{E}_j(x_{\text{pred}})]^2)}$  are plotted versus the midpoint  $\bar{x}_j$  of the  $j$ th bin, while in Table II, we show the mean absolute error  $\mathbb{E}(|\Delta_x|)$  and the standard deviation of the absolute error  $\sqrt{\mathbb{E}([\Delta_x - \mathbb{E}(\Delta_x)]^2)}$  over the prediction set. Note that 1)  $\sqrt{\mathbb{E}([\Delta_x - \mathbb{E}(\Delta_x)]^2)}$  reproduces the square root of the diagonal elements of the so-called epistemic covariance matrix of all

TABLE II  
MEAN ABSOLUTE ERROR  $\mathbb{E}(|\Delta_x|)$  AND THE STANDARD DEVIATION OF THE ABSOLUTE ERROR  $\sqrt{\mathbb{E}([\Delta_x - \mathbb{E}(\Delta_x)]^2)}$  OVER THE PREDICTION SET

Method	$x$	$\mathbb{E}( \Delta_x )$	$\sqrt{\mathbb{E}([\Delta_x - \mathbb{E}(\Delta_x)]^2)}$
1	$\tau$	0.169	0.410
	$H$	0.879	1.749
2	$\tau$	0.115	0.243
	$H$	0.336	0.740
3	$\tau$	0.136	0.316
	$H$	0.437	0.951

The results correspond to the forward-operator neural network (Method 1), and the inverse-operator neural networks with synthetic radiances (Method 2) and the principal-component transform of synthetic radiances (Method 3).

network errors over the prediction set, and 2) the epistemic uncertainties are large if there are large variations around the

mean, e.g., if  $\sqrt{\mathbb{E}_j([x_{\text{pred}} - \mathbb{E}_j(x_{\text{pred}})]^2)}$  are large. Also note that nonoptimal hyperparameters and training parameters as well as a nonoptimal optimization algorithm are the main sources of epistemic or model uncertainty [61]. The following conclusions can be drawn.

- 1) In general, the accuracy is low for small values of the aerosol optical depth  $\tau$  and layer height  $H$ .
- 2) The inverse-operator neural networks with synthetic radiances and the principal components of synthetic radiances have comparable accuracies; these are higher than that of the forward-operator neural network. The low accuracy of the forward-operator neural network, especially for small value of the aerosol optical depth and layer height, is due to the fact that in this domain, the residual function has several local minima and the global minimum cannot be found by the iteratively regularized Gaussian method (which is a local optimization method).

The variations of the absolute error  $\Delta_x = x_{\text{pred}} - x$  with respect to the optical and geometrical parameters are illustrated in Figs. 4–6. As before, each interval of variation of a parameter  $b$ ,  $[b_{\min}, b_{\max}]$  is split into  $N_b = 40$  equidistant bins, i.e.,  $[b_{\min}, b_{\max}] = \cup_{j=1}^{N_b} B_{bj}$ , and the mean  $\mathbb{E}_j(\Delta_x)$  and standard deviation  $\sqrt{\mathbb{E}_j([\Delta_x - \mathbb{E}_j(\Delta_x)]^2)}$  over all samples  $\Delta_x^{(n)}$ , whose corresponding  $b^{(n)}$  are in the  $j$ th bin  $B_{bj}$ , are plotted versus the midpoint  $\bar{b}_j$  of the bin. The plots show the following.

- 1) The standard deviation of the absolute error in the aerosol optical depth  $\tau$  is large when  $\tau$  is small, and the solar zenith angle  $\theta_0$ , viewing angle  $\theta$ , and surface albedo  $A_s$  are large.
- 2) The standard deviation of the absolute error in the aerosol layer height  $H$  is large for small values of the aerosol optical thickness  $\tau$  and large values of the surface albedo  $A_s$ .
- 3) The smallest standard deviations correspond to the inverse-operator neural network with synthetic radiances, while the largest correspond to the forward-operator neural network.

For instance, in Fig. 4 (the forward-operator neural network), the standard deviation in  $\Delta_\tau$  can be of 1.14 for small values of  $\tau$ , and of 0.69, 0.64, and 0.92 for large values of  $\theta_0$ ,  $\theta$ , and  $A_s$ , respectively, whereas the standard deviation in  $\Delta_H$  can be of 5.32 km for small values of  $\tau$  and of 2.36 km for large values of  $A_s$ .

## B. Retrieval from Real Data

To investigate the performances of the retrieval algorithms on real TROPOMI data, we first choose a wild fire scene in California on December 12, 2017. In this case, the surface albedo is given by the GE\_LER product [28], and pixels with

- 1) a cloud fraction [taken from the operational TROPOMI cloud product (OCRA/ROCINN) [31]] greater than 0.25, or
- 2) an aerosol absorbing index (taken from the TROPOMI Level-2 AAI product) lower than 1, are not considered in the retrieval.

The retrieval results for the aerosol optical depth and layer height are illustrated in Fig. 7. With the results by the inverse-operator neural network with synthetic radiances as a reference, Fig. 8 shows the absolute differences in the retrieved aerosol optical depth  $\delta_\tau = \tau_{\text{pred}} - \tau_{\text{pred}}^{\text{ref}}$  and aerosol layer height  $\delta_H = H_{\text{pred}} - H_{\text{pred}}^{\text{ref}}$  corresponding to the forward-operator neural network and the inverse-operator neural network with the principal-component transform of synthetic radiances. The plots demonstrate that the differences in the retrieved aerosol optical depth  $\delta_\tau$  are smaller than 0.1 over the entire scene, while the absolute differences in the retrieved aerosol layer height  $\delta_H$  are smaller than 0.4 km.

Finally, we compare the retrieval results between the three neural network based algorithms and the official operational algorithm for the entire year of 2019. From Fig. 9, the following features are apparent.

- 1) The mean values of the aerosol optical depth delivered by the three neural network algorithms are in general underestimated. The reason for this discrepancy is that the official retrieval algorithm uses different aerosol microphysical properties (a fact also seen in the previous study [8]).
- 2) The mean values of the aerosol layer height delivered by the three neural network algorithms agree well with those by the operational retrieval algorithm. However, as expected, a better agreement can be seen between the forward-operator neural network and the operational algorithm whose forward model adopts the neural network approach [35].

It should be pointed out that the computational time of a physics-based retrieval algorithm using online radiative transfer calculations is approximately 3 min for one ground pixel on a computer Intel Core i7-4770 CPU 3.40 GHz with 16 GB RAM, while the computational time of a forward- and an inverse-operator neural networks are 0.4 and 0.003 s, respectively. For these calculations, the Bayesian approach typically converges in less than five iterations.

Likewise, we perform a comparison with the weighted aerosol heights derived from the Cloud-Aerosol LIDAR Infrared Pathfinder Satellite Observations (CALIPSO) Level-2 aerosol extinction profile product for the entire year of 2019. The weighted aerosol height is computed as a linear combination of the heights above the sea level weighted by the corresponding extinction coefficients. Fig. 10 depicts the histograms of the differences in aerosol layer height (TROPOMI minus CALIPSO), where a TROPOMI value is the mean value over all TROPOMI pixels within a distance of 100 km to a single CALIPSO pixel. The difference  $E(\Delta_H) \pm \sqrt{E([\Delta_H - E(\Delta_H)]^2)}$  is highest for the forward operator neural network ( $1.54 \pm 1.62$  km), whereas a better agreement with the CALIPSO product is apparently achieved by the inverse operator with synthetic radiances ( $1.38 \pm 1.47$  km) and the inverse operator with the principal-component transform of synthetic radiances ( $1.48 \pm 1.43$  km).

## V. CONCLUSION

In this article, we have developed three neural network algorithms for aerosol retrieval from TROPOMI measurements in the O<sub>2</sub> A-band.

- 1) The first algorithm uses a neural network to emulate the radiative transfer model and a Bayesian approach to solve the inverse problem. To speed up the computation, the radiative transfer model combines standard acceleration methods (the telescoping technique and the method of false discrete ordinate) with hyperspectral acceleration methods (the correlated  $k$ -distribution method and the principal-component analysis). The inverse problem is solved by using the iteratively regularized Gauss–Newton method, which provides at the same time the optimal value of the regularization parameter and the corresponding regularized solution.
- 2) The second and third algorithms employ a neural network to learn the inverse model and use as input either the synthetic radiances computed on the measurement wavelength grid or the principal-component transform of synthetic radiances. The design of an inverse-operator neural network for TROPOMI/S5P is not a trivial task because, on the one hand, we are dealing with random measurement noise and, on the other hand, there are a large number of measurement wavelength grids corresponding to each swath row. To solve this problem, we used the jitter approach. More precisely, in the training stage and at each forward pass through the network, a measurement wavelength grid is randomly selected from a set of possible wavelength grids, and a new random noise is added to the synthetic radiance. Note that in the algorithm relying on the principal-component analysis, the random noise is described through an analytic dimensionality-reduced noise covariance matrix.

The neural networks are incorporated into a common tool and their performances are analyzed and compared with synthetic and real data. Our numerical analysis has shown that the inverse-operator neural networks are more accurate and efficient than a forward-operator neural network. The reason is that for small values of the aerosol optical depth and layer height, the residual function has several local minima, and in the latter case, the inversion method, which is a local optimization method, cannot determine the global minimum. These results may suggest that a TROPOMI operational retrieval algorithm can be built on an inverse-operator neural network rather than on a forward-operator neural network.

The design and refinement of neural networks for atmospheric retrieval is a very complicated research field that requires more developments that consist of the following:

- 1) application of the inverse-operator neural networks to the remaining aerosol models considered in the MODIS algorithm, i.e., nonabsorbing, absorbing, and desert dust (the selection of an appropriate aerosol model is then based on a combination of spectral and geographic information);
- 2) training the neural networks to learn the relative evidences of different aerosol models, so that a mean solution estimate, representing a linear combination of candidate solutions weighted by their evidences, can be computed [62];
- 3) redesign of the neural networks in a Bayesian deep learning framework in order to predict input aleatoric and model uncertainties [61].

#### ACKNOWLEDGMENT

The authors are grateful to DLR and KNMI colleagues for sharing the relevant data and information of TROPOMI/S5P. The authors would like to thank EU/ESA/KNMI/DLR for providing the TROPOMI/S5P Level-1b data.

#### REFERENCES

- [1] G. A. A. Koppers and D. P. Murtagh, *Retrieval of Height Resolved Aerosol Optical Thickness in the Atmospheric Band*. Stockholm, Sweden: Stockholm Univ., 1997.
- [2] L. Tilstra, O. Tuinder, P. Wang, and P. Stammes, "GOME-2 absorbing aerosol height algorithm theoretical basis document," Roy. Netherlands Meteorological Inst. (KNMI), De Bilt, Netherlands, Tech. Rep. SAF/AC/KNMI/ATBD/005, 2019.
- [3] S. Corradini and M. Cervino, "Aerosol extinction coefficient profile retrieval in the oxygen a-band considering multiple scattering atmosphere. Test case: SCIAMACHY nadir simulated measurements," *J. Quant. Spectrosc. Radiative Transfer*, vol. 97, no. 3, pp. 354–380, 2006.
- [4] A. A. Kokhanovsky and V. V. Rozanov, "The determination of dust cloud altitudes from a satellite using hyperspectral measurements in the gaseous absorption band," *Int. J. Remote Sens.*, vol. 31, no. 10, pp. 2729–2744, 2010.
- [5] S. Sanghavi, J. V. Martonchik, J. Landgraf, and U. Platt, "Retrieval of the optical depth and vertical distribution of particulate scatterers in the atmosphere using  $\text{O}_2$  a- and b-band SCIAMACHY observations over Kanpur: A case study," *Atmos. Meas. Techn.*, vol. 5, no. 5, pp. 1099–1119, 2012.
- [6] C. Frankenberg, O. Hasekamp, C. O'Dell, S. Sanghavi, A. Butz, and J. Worden, "Aerosol information content analysis of multi-angle high spectral resolution measurements and its benefit for high accuracy greenhouse gas retrievals," *Atmos. Meas. Techn.*, vol. 5, no. 7, pp. 1809–1821, 2012.
- [7] S. Nanda et al., "A first comparison of TROPOMI aerosol layer height (ALH) to CALIOP data," *Atmos. Meas. Techn.*, vol. 13, no. 6, pp. 3043–3059, 2020.
- [8] L. Rao, J. Xu, D. S. Efremenko, D. G. Loyola, and A. Doicu, "Hyperspectral satellite remote sensing of aerosol parameters: Sensitivity analysis and application to TROPOMI/S5P," *Front. Environ. Sci.*, vol. 9, 2022, Art. no. 770662.
- [9] J. Xu, F. Schreier, A. Doicu, and T. Trautmann, "Assessment of Tikhonov-type regularization methods for solving atmospheric inverse problems," *J. Quant. Spectrosc. Radiative Transfer*, vol. 184, pp. 274–286, 2016.
- [10] J. Xu, L. Rao, F. Schreier, D. S. Efremenko, A. Doicu, and T. Trautmann, "Insight into construction of Tikhonov-type regularization for atmospheric retrievals," *Atmosphere*, vol. 11, no. 10, 2020, Art. no. 1052.
- [11] C. D. Rodgers, *Inverse Methods for Atmospheric Sounding: Theory and Practice*, vol. 2. Singapore: World Scientific, 2000.
- [12] V. Arnold, "On the representation of functions of several variables as a superposition of functions of a smaller number of variables," *Math., Teaching, Appl. History, Matem. Prosv.*, Ser. 2, vol. 3, pp. 41–61, 1958.
- [13] G. Cybenko, "Approximation by superpositions of a sigmoidal function," *Math. Control, Signals, Syst.*, vol. 2, no. 4, pp. 303–314, 1989.
- [14] Y. Fan et al., "Atmospheric correction over coastal waters using multilayer neural networks," *Remote Sens. Environ.*, vol. 199, pp. 218–240, 2017.
- [15] C. Fan et al., "Use of a neural network-based ocean body radiative transfer model for aerosol retrievals from multi-angle polarimetric measurements," *Remote Sens.*, vol. 11, no. 23, 2019, Art. no. 2877.
- [16] C. Shi, M. Hashimoto, K. Shiomi, and T. Nakajima, "Development of an algorithm to retrieve aerosol optical properties over water using an artificial neural network radiative transfer scheme: First result from GOSAT-2/CAL-2," *IEEE Trans. Geosci. Remote Sens.*, vol. 59, no. 12, pp. 9861–9872, Dec. 2021.
- [17] M. Gao et al., "Efficient multi-angle polarimetric inversion of aerosols and ocean color powered by a deep neural network forward model," *Atmos. Meas. Techn.*, vol. 14, no. 6, pp. 4083–4110, 2021.
- [18] D. G. R. Loyola, "Applications of neural network methods to the processing of earth observation satellite data," *Neural Netw.*, vol. 19, no. 2, pp. 168–177, 2006.
- [19] D. S. Efremenko, "Discrete ordinate radiative transfer model with the neural network based eigenvalue solver: Proof of concept," *Light Eng.*, vol. 01, pp. 56–62, 2021.

- [20] C. Jiménez, P. Eriksson, and D. Murtagh, "Inversion of odin limb sounding submillimeter observations by a neural network technique," *Radio Sci.*, vol. 38, no. 4, 2003, Art. no. 8062.
- [21] G. Holl, S. Eliasson, J. Mendrok, and S. Buehler, "SPARE-ICE: Synergistic ice water path from passive operational sensors," *J. Geophysical Res.: Atmos.*, vol. 119, no. 3, pp. 1504–1523, 2014.
- [22] J. Strandgren, L. Bugliaro, F. Sehnke, and L. Schröder, "Cirrus cloud retrieval with MSG/SEVIRIs using artificial neural networks," *Atmos. Meas. Techn.*, vol. 10, no. 9, pp. 3547–3573, 2017.
- [23] D. Wang, C. Prigent, F. Aires, and C. Jimenez, "A statistical retrieval of cloud parameters for the millimeter wave ice cloud imager on board MetOp-SGs," *IEEE Access*, vol. 5, pp. 4057–4076, 2017.
- [24] M. Brath, S. Fox, P. Eriksson, R. C. Harlow, M. Burgdorf, and S. A. Buehler, "Retrieval of an ice water path over the ocean from ISMAR and MARSS millimeter and submillimeter brightness temperatures," *Atmos. Meas. Techn.*, vol. 11, no. 1, pp. 611–632, 2018.
- [25] N. Håkansson, C. Adok, A. Thoss, R. Scheirer, and S. Hörnquist, "Neural network cloud top pressure and height for MODIS," *Atmos. Meas. Techn.*, vol. 11, no. 5, pp. 3177–3196, 2018.
- [26] D. S. Efremenko, D. G. Loyola R., P. Hedelt, and R. J. D. Spurr, "Volcanic SO<sub>2</sub> plume height retrieval from UV sensors using a full-physics inverse learning machine algorithm," *Int. J. Remote Sens.*, vol. 38, no. suppl, pp. 1–27, 2017.
- [27] J. Xu, O. Schüssler, D. Loyola R., F. Romahn, and A. Doicu, "A novel ozone profile shape retrieval using full-physics inverse learning machine (FP-ILM)," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 10, no. 12, pp. 5442–5457, Dec. 2017.
- [28] D. G. Loyola, J. Xu, K.-P. Heue, and W. Zimmer, "Applying FP\_ILM to the retrieval of geometry-dependent effective Lambertian equivalent reflectivity (GE\_LER) daily maps from UVN satellite measurements," *Atmos. Meas. Techn.*, vol. 13, no. 2, pp. 985–999, 2020.
- [29] A. Di Noia et al., "Use of neural networks in ground-based aerosol retrievals from multi-angle spectropolarimetric observations," *Atmos. Meas. Techn.*, vol. 8, no. 1, pp. 281–299, 2015.
- [30] A. D. Noia, O. P. Hasekamp, L. Wu, B. v. Diedenhoven, B. Cairns, and J. E. Yorks, "Combined neural network/Phillips-Tikhonov approach to aerosol retrievals over land from the NASA research scanning polarimeter," *Atmos. Meas. Techn.*, vol. 10, no. 11, pp. 4235–4252, 2017.
- [31] D. G. Loyola et al., "The operational cloud retrieval algorithms from TROPOMI on board Sentinel-5 precursor," *Atmos. Meas. Techn.*, vol. 11, no. 1, pp. 409–427, 2018.
- [32] S. Compennolle et al., "Validation of the Sentinel-5 precursor TROPOMI cloud data with cloudnet, aura OMI O 2-O 2, MODIS, and suomi-NPP VIIRS," *Atmos. Meas. Techn.*, vol. 14, pp. 2451–2476, 2021.
- [33] A. F. J. Sanders and J. F. de Haan, "Retrieval of aerosol parameters from the oxygen a band in the presence of chlorophyll fluorescence," *Atmos. Meas. Techn.*, vol. 6, no. 10, pp. 2725–2740, 2013.
- [34] A. F. J. Sanders et al., "Evaluation of the operational aerosol layer height retrieval algorithm for Sentinel-5 precursor: Application to O<sub>2</sub> a band observations from GOME-2A," *Atmos. Meas. Techn.*, vol. 8, no. 11, pp. 4947–4977, 2015.
- [35] S. Nanda et al., "A neural network radiative transfer model approach applied to the TROPospheric monitoring instrument aerosol height algorithm," *Atmos. Meas. Techn.*, vol. 12, no. 12, pp. 6619–6634, 2019.
- [36] A. Doicu and T. Trautmann, "Discrete-ordinate method with matrix exponential for a pseudo-spherical atmosphere: Scalar case," *J. Quant. Spectrosc. Radiative Transfer*, vol. 110, pp. 146–158, 2009.
- [37] D. S. Efremenko, V. Molina García, S. Gimeno García, and A. Doicu, "A review of the matrix-exponential formalism in radiative transfer," *J. Quant. Spectrosc. Radiative Transfer*, vol. 196, pp. 17–45, 2017.
- [38] R. Spurr, *LIDORT and VLIDORT: Linearized Pseudo-Spherical Scalar and Vector Discrete Ordinate Radiative Transfer Models for Use in Remote Sensing Retrieval Problems*, vol. 3. A. Kokhanovsky, Ed. Berlin: Springer Verlag, 2008, pp. 229–271.
- [39] D. Efremenko, A. Doicu, D. Loyola, and T. Trautmann, "Acceleration techniques for the discrete ordinate method," *J. Quant. Spectrosc. Radiative Transfer*, vol. 114, pp. 73–81, 2013.
- [40] E. Chalhoub and R. Garcia, "The equivalence between two techniques of angular interpolation for the discrete-ordinates method," *J. Quant. Spectrosc. Radiative Transfer*, vol. 64, no. 5, pp. 517–535, 2000.
- [41] R. Goody, R. West, L. Chen, and D. Crisp, "The correlated-k method for radiation calculations in nonhomogeneous atmospheres," *J. Quant. Spectrosc. Radiative Transfer*, vol. 42, no. 6, pp. 437–651, 1989.
- [42] V. Natraj, R.-L. Shia, X. Huang, J. S. Margolis, and Y. L. Yung, "Application of principal component analysis to high spectral resolution radiative transfer: A case study of the O<sub>2</sub>A band," *J. Quant. Spectrosc. Radiative Transfer*, vol. 95, no. 4, pp. 539–556, 2005.
- [43] V. Natraj, R.-L. Shia, and Y. L. Yung, "On the use of principal component analysis to speed up radiative transfer calculations," *J. Quant. Spectrosc. Radiative Transfer*, vol. 111, no. 3, pp. 810–816, 2010.
- [44] D. Efremenko, A. Doicu, D. Loyola, and T. Trautmann, "Optical property dimensionality reduction techniques for accelerated radiative transfer performance: Application to remote sensing total ozone retrievals," *J. Quant. Spectrosc. Radiative Transfer*, vol. 133, pp. 128–135, 2014.
- [45] W. Hou, J. Wang, X. Xu, J. S. Reid, and D. Han, "An algorithm for hyperspectral remote sensing of aerosols: 1 development of theoretical framework," *J. Quant. Spectrosc. Radiative Transfer*, vol. 178, pp. 400–415, 2016.
- [46] V. Molina García, S. Sasi, D. S. Efremenko, A. Doicu, and D. Loyola, "Radiative transfer models for retrieval of cloud parameters from EPIC/DSCOVR measurements," *J. Quant. Spectrosc. Radiative Transfer*, vol. 213, pp. 228–240, 2018.
- [47] F. Schreier et al., "GARLIC-A general purpose atmospheric radiative transfer line-by-line infrared-microwave code: Implementation and evaluation," *J. Quant. Spectrosc. Radiative Transfer*, vol. 137, pp. 29–50, 2014.
- [48] F. Schreier, "Optimized implementations of rational approximations for the voigt and complex error function," *J. Quant. Spectrosc. Radiative Transfer*, vol. 112, no. 6, pp. 1010–1025, 2011.
- [49] L. S. Rothman et al., "The HITRAN 2008 molecular spectroscopic database," *J. Quant. Spectrosc. Radiative Transfer*, vol. 110, no. 9–10, pp. 533–572, 2009.
- [50] B. A. Bodhaine, N. B. Wood, E. G. Dutton, and J. R. Slusser, "On rayleigh optical depth calculations," *J. Atmos. Ocean. Technol.*, vol. 16, no. 11, pp. 1854–1861, 1999.
- [51] W. Wiscombe, "The delta-m method: Rapid yet accurate radiative flux calculations for strongly asymmetric phase functions," *J. Atmos. Sci.*, vol. 34, no. 9, pp. 1408–1422, 1977.
- [52] T. Nakajima and M. Tanaka, "Algorithms for radiative intensity calculations in moderately thick atmospheres using a truncation approximation," *J. Quant. Spectrosc. Radiative Transfer*, vol. 40, no. 1, pp. 51–69, 1988.
- [53] M. Momoi, H. Irie, T. Nakajima, and M. Sekiguchi, "Efficient calculation of radiative intensity including the polarization effect in moderately thick atmospheres using a truncation approximation," *J. Quant. Spectrosc. Radiative Transfer*, vol. 277, 2022, Art. no. 107976.
- [54] S. Sasi, V. Natraj, V. M. García, D. S. Efremenko, D. Loyola, and A. Doicu, "Model selection in atmospheric remote sensing with an application to aerosol retrieval from DSCOVR/EPIC, part 1: Theory," *Remote Sens.*, vol. 12, no. 22, 2020, Art. no. 3724.
- [55] S. Sasi, V. Natraj, V. Molina García, D. S. Efremenko, D. Loyola, and A. Doicu, "Model selection in atmospheric remote sensing with application to aerosol retrieval from DSCOVR/EPIC. part 2: Numerical analysis," *Remote Sens.*, vol. 12, no. 21, 2020, Art. no. 3656.
- [56] R. C. Levy, L. A. Remer, and O. Dubovik, "Global aerosol optical properties and application to moderate resolution imaging spectroradiometer aerosol retrieval over land," *J. Geophys. Res. Atmos.*, vol. 112, no. D13, pp. n/a–n/a, 2007.
- [57] D. G. Loyola, M. Pedergrana, and S. Gimeno Garcia, "Smart sampling and incremental function learning for very large high dimensional data," *Neural Netw.*, vol. 78, pp. 75–87, 2016.
- [58] J. H. Halton, "Algorithm 247: Radical-inverse quasi-random point sequence," *Commun. ACM*, vol. 7, no. 12, pp. 701–702, 1964.
- [59] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," 2017, *arXiv:1412.6980*.
- [60] A. Doicu, T. Trautmann, and F. Schreier, *Numerical Regularization for Atmospheric Inverse Problems*. Berlin, Germany: Springer Science & Business Media, 2010.
- [61] A. Doicu, A. Doicu, D. S. Efremenko, D. Loyola, and T. Trautmann, "An overview of neural network methods for predicting uncertainty in atmospheric remote sensing," *Remote Sens.*, vol. 13, no. 24, 2021, Art. no. 5061.
- [62] L. Rao, J. Xu, D. S. Efremenko, D. G. Loyola, and A. Doicu, "Optimization of aerosol model selection for TROPOMI/S5P," *Remote Sens.*, vol. 13, no. 13, 2021, Art. no. 2489.





**Lanlan Rao** received the master's degree in cartography and geographic information engineering from the China University of Mining and Technology, Xuzhou, China, in 2017. She is currently working toward the Ph.D. degree in satellite remote sensing of aerosols with the German Aerospace Center (DLR), Remote Sensing Technology Institute (IMF), Weßling, Germany and the Technische Universität München (TUM), Munich, Germany.

Her research interests focus on remote sensing of aerosol properties from satellite data.



**Jian Xu** (Member, IEEE) received the B.E. degree in geographic information systems from Hohai University, Nanjing, China, in 2004, and the M.S. degree in Earth-oriented space science and technology and the Dr.-Ing. degree in atmospheric remote sensing from the Technische Universität München, Munich, Germany, in 2009 and 2015, respectively.

From 2010 to 2021, he was with the Remote Sensing Technology Institute (IMF), German Aerospace Center (DLR), Oberpfaffenhofen, Weßling, Germany. His was involved in the development/refinement of atmospheric retrieval algorithms for ESA's Sentinel-5P and Sentinel-4 satellite missions. Since 2021, he has been with the National Space Science Center (NSSC), Chinese Academy of Sciences (CAS), Beijing, China. His research interests include remote sensing of planetary atmosphere, radiative transfer modeling, and ill-posed inverse problems.



**Dmitry S. Efremenko** received the graduation degree in technical physics of fission reactors and nuclear setups from the Moscow Power Engineering Institute (MPEI), Moscow, Russia, in 2009, the Ph.D. degree in physics and mathematics from the Moscow State University, Moscow, Russia, in 2011, and the Habilitation degree in technical sciences from MPEI in 2017.

Since 2011, he has been working as a Research Scientist with the German Aerospace Center (DLR), Cologne (Köln), Germany. He gives academic courses "Radiative Transfer" and "Non-linear Optimization" at the Technische Universität München, Munich, Germany. His scientific interests include radiative transfer, remote sensing, deconvolution problems, and machine learning. Since 2018, he has been leading a junior research group "Atmospheric Retrieval of the Next Generation."

Dr. Efremenko was the recipient of the Elsevier/JQSRT Goody Award in 2017.



**Diego G. Loyola** (Senior Member, IEEE) received the Lic. Inf. degree in computer science from ESLAI, Universidad de Lujan, Lujan, Argentina, in 1988, and the Dr.-Ing. degree in remote sensing and machine learning from the Technische Universität München, Munich, Germany, in 2013.

Since 1990, he has been a Scientist and Project Manager with the German Aerospace Center (Deutsches Zentrum für Luft- und Raumfahrt - DLR), Oberpfaffenhofen, Germany, and a Visiting Scientist with the NASA Goddard Space Flight Center, Greenbelt, MD, USA, in 2000, 2016/2017, and 2022/2023. Since 2000, he has been with Remote Sensing Technology Institute, DLR, where he leads the passive sensing group responsible for the operational atmospheric composition and cloud products from the European missions GOME on ERS-2, GOME-2 on MetOp-A/-B and -C, TROPOMI on Sentinel-5 Precursor, and Sentinel-4 on MTG-S. Additionally, he is a Visiting Professor with the School of Physics, Aristotle University of Thessaloniki, Thessaloniki, Greece. His research interests include remote sensing of atmospheric trace gases, cloud and aerosol properties, as well as computational intelligence, and machine learning techniques.

Dr. Loyola has been holding the DLR title Senior Scientist since 2015. He is the winner of the DLR Scientific Award 2011. He is a senior member of the IEEE Geoscience and Remote Sensing Society, IEEE Computational Intelligence Society, and International Neural Network Society.



**Adrian Doicu** received the Engineering and Ph.D. degrees in mechanical engineering and optics from the Technical University of Bucharest, Bucharest, Romania, in 1986 and 1996, respectively.

He is currently a Senior Scientist with the Remote Sensing Technology Institute (IMF), German Aerospace Center (DLR), Oberpfaffenhofen, Germany. Since 2012, he has been appointed as an Associate Professor (German: Privatdozent) in atmospheric remote sensing at the Technische Universität München, Munich, Germany. His

research interests include electromagnetic scattering theory, inverse problems, and radiative transfer.