Aerosol Absorption from Global Satellite Measurements in the Ultra-Violet

From Qualitative Aerosol Index to Quantitative Aerosol Absorptive Properties

Sun, J.

DOI
10.4233/uuid:724acbc5-64b6-41d9-8142-dea2b137b8af

Publication date
2022

Document Version
Final published version

Citation (APA)
Sun, J. (2022). *Aerosol Absorption from Global Satellite Measurements in the Ultra-Violet: From Qualitative Aerosol Index to Quantitative Aerosol Absorptive Properties*. [Dissertation (TU Delft), Delft University of Technology]. https://doi.org/10.4233/uuid:724acbc5-64b6-41d9-8142-dea2b137b8af

Aerosol Absorption from Global Satellite Measurements in the Ultra-Violet

From Qualitative Aerosol Index to Quantitative Aerosol Absorptive Properties

AEROSOL ABSORPTION FROM GLOBAL SATELLITE MEASUREMENTS IN THE ULTRA-VIOLET

FROM QUALITATIVE AEROSOL INDEX TO QUANTITATIVE AEROSOL ABSORPTIVE PROPERTIES

Dissertation

ter verkrijging van de graad van doctor
aan de Technische Universiteit Delft,
op gezag van de Rector Magnificus prof. ir. T.H.J.J. van der Hagen,
voorzitter van het College voor Promoties,
in het openbaar te verdedigen op **maandag 14 maart 2022 om 17:30 uur**

door

Jiyunting SUN

Master of Science in Applied Earth Sciences,
Technische Universiteit Delft, Delft, the Netherlands,
geboren te Nanjing, China.

Dit proefschrift is goedgekeurd door de promotoren.

Samenstelling promotiecommissie bestaat uit:

Rector Magnificus,	voorzitter
Prof. dr. P. F. Levelt,	Technische Universiteit Delft, National Center for Atmospheric Research, promoter
Dr. ir. J. P. Veefkind,	Technische Universiteit Delft, Koninklijk Nederlands Meteorologisch Instituut, copromoter

Onafhankelijke leden:

Prof. dr. ir. H. W. J. Russchenberg,	Technische Universiteit Delft
Prof. dr. A. P. Siebesma,	Technische Universiteit Delft, Koninklijk Nederlands Meteorologisch Instituut
Prof. dr. G. de Leeuw,	Koninklijk Nederlands Meteorologisch Instituut
Dr. J. Joiner,	National Aeronautics and Space Administration

Overige leden:

Dr. P. F. J. van Velthoven,	Koninklijk Nederlands Meteorologisch Instituut

Keywords:, Absorbing Aerosols, Ultra-Violet Aerosol Index, Absorbing Aerosol Optical Depth, Single Scattering Albedo, Aerosol Vertical Distribution, Machine Learning, Support Vector Regression, Deep Neural Network, Satellite Remote Sensing, Radiative Transfer

Printed by: **Print & Bind**

Front & Back: Designed by Jiyunting Sun.

ISBN 978-94-6384-306-5

An electronic version of this dissertation is available at
http://repository.tudelft.nl/.

CONTENTS

SUMMARY

Atmospheric aerosols are solid or liquid particles suspended in the air. The majority of them are produced by natural processes, including sea salt from oceans, mineral dust from (semi-)arid regions, carbon containing particles from wildfires, and sulfates and ash from volcanic activities. Anthropogenic aerosols are produced by industrial activities, power generation, transportation, agriculture, and human-induced biomass burning events. Depending on the meteorological conditions, aerosol particles can stay in the atmosphere for several hours to several months and can be transported over long distances, causing adverse effects on human health, visibility and climate.

This thesis focuses on the aerosol optical properties, particularly the light absorption of the aerosol particles that has significant effects on the Earth's climate system.

This thesis starts with a general introduction of atmospheric aerosols, including its sources, categories, physical properties and measurement techniques (Chapter 1). Next, the Ultra-Violet Aerosol Index (UVAI) is introduced, which is calculated from satellite measurements of the radiance at two wavelengths in the UV. UVAI contains information of aerosol absorption, and it has a very long and almost continuous data record starting in 1978. Direct use of UVAI is challenging because it is not a geophysical quantity, but a numerical index. The objective of this thesis is to derive quantitative properties on aerosol absorption from the UVAI (e.g. single scattering albedo, absorption aerosol optical depth) that can be directly used in aerosol radiative transfer assessments. Two types of methods have been developed, i.e. physically-based methods and statistically-based methods. The first compares the observed UVAI to the one simulated by radiative transfer models. The second uses Machine Learning algorithms trained by existing data sets.

The physically-based methods have been applied to quantify aerosol absorption of several large scale wildfires (Chapter 2 and 3). An important challenge of these method is that assumptions have to be made on the aerosol micro-physical properties, leading to significant uncertainties in the results, whereas the Machine Learning-based methods can avoid this kind of assumptions. Chapter 3 investigates the feasibility to quantify aerosol absorption from UVAI using a Machine Learning algorithm. Despite the higher computational efficiency and better results, the application of such data-driven methods is still restricted by the limited data on the aerosol vertical distribution. Therefore, in Chapter 4, a database of aerosol height is created from a chemistry transport model. This database is applied in Chapter 5, where a Deep Neural Network method is used to derive the quantitative aerosol absorptive properties from the OMI/Aura UVAI for the period from 2006 to 2019. In comparison to ground-based observations, the results of the Deep Neural Network agree better than satellite retrievals and also better than chemistry transport model simulations.

This thesis demonstrates the feasibility of deriving quantitative aerosol absorptive properties from the satellite retrieved UVAI. We use traditional radiative transfer simulations meanwhile investigating the new possibilities of data-driven methods in aerosol remote sensing. Although the retrieval results are encouraging, there remain limitations and challenges which need to be addressed. These are discussed in Chapter 6 with corresponding suggestions and prospects. Despite the challenges, it is expected that a synthetic database of global aerosol absorption can be derived from UVAI observations provided by multiple satellite products. Such a data set will make great contributions to quantify the effect of absorbing aerosols on the climate system.

SAMENVATTING

Atmosferische aërosolen zijn vaste of vloeibare deeltjes die in de lucht zweven. De meeste daarvan worden geproduceerd door natuurlijke processen, waaronder zeezout afkomstig van oceanen, mineraalstof afkomstig van (semi-)aride gebieden, koolstofhoudende deeltjes geproduceerd door bosbranden en sulfaat en as van vulkanische activiteit. Antropogene aërosoldeeltjes worden geproduceerd door industriële activiteiten, energieopwekking, transport, landbouw, en door de mens veroorzaakte verbranding van biomassa. Afhankelijk van de meteorologische omstandigheden kunnen aërosoldeeltjes meerdere dagen in de atmosfeer verblijven en over lange afstanden worden getransporteerd, met nadelige gevolgen voor de menselijke gezondheid, het zicht en het klimaat.

Dit proefschrift richt zich op de optische eigenschappen van aërosolen, met name van lichtabsorptie, welke een significant effect heeft op het klimaatsysteem van de aarde. Dit proefschrift begint met een algemene introductie over atmosferische aërosolen, inclusief de bronnen, categorieën, fysische eigenschappen en metingen (hoofdstuk 1). Vervolgens wordt de Ultraviolet Aerosol Index (UVAI) geïntroduceerd, welke wordt bepaald uit satellietmetingen van de UV-straling bij twee golflengtes. De UVAI bevat informatie over absorptie aërosol deeltjes en heeft een zeer lange en bijna continue meetreeks vanaf 1978. Direct gebruik van UVAI is beperkt doordat het geen geofysische grootheid is, maar een numerieke index. Het doel van dit proefschrift is om kwantitatieve grootheden over aërosolabsorptie uit de UVAI af te leiden (bijv. de enkelvoudige verstrooiingsalbedo en de absorptie optische diepte van het aërosol), die direct kunnen worden gebruikt om de effecten daarvan op het stralingstransport in de atmosfeer te bepalen. Er zijn twee soorten methoden gebruikt, namelijk fysisch-gebaseerde methodes en statistisch-gebaseerde methodes. De eerste vergelijken de waargenomen UVAI met die gesimuleerd door stralingstransportmodellen. De tweede maken gebruik van Machine Learning-algoritmen getraind op bestaande datasets.

De fysische methodes zijn toegepast om aërosolabsorptie van verschillende grootschalige bosbranden te kwantificeren (Hoofdstuk 2 en 3). Een belangrijke uitdaging voor deze methodes is dat er aannames gemaakt moeten worden over de microfysische eigenschappen van de aërosoldeeltjes, wat leidt tot aanzienlijke onzekerheden in de resultaten, terwijl de op Machine Learning gebaseerde methodes dit soort aannames kunnen vermijden. Hoofdstuk 3 onderzoekt de haalbaarheid om aërosolabsorptie uit de UVAI te kwantificeren met behulp van een Machine Learning-algoritme. Ondanks de hogere rekenefficiëntie en goede resultaten, wordt de toepassing van dergelijke methoden beperkt door de beperkte gegevens over de verticale verdeling van de aërosoldeeltjes. Daarom is een database van aërosolhoogtes afgeleid uit simulaties met een chemietransportmodel (Hoofdstuk 4). De verkregen database is toegepast in Hoofdstuk 5, waar een Deep Neural Network-methode wordt gebruikt om kwantitatieve eigenschappen van aërosolabsorptie af te leiden uit de OMI/Aura UVAI metingen voor de periode van 2006 tot 2019. De resultaten van het Deep Neural Network vergelijken beter met waarnemingen vanaf de grond dan bepalingen met satellietwaarnemingen gebaseerd op een fysische methode, en ook beter dan simulaties met een chemietransportmodel.

Dit proefschrift demonstreert de haalbaarheid om kwantitatieve aërosolabsorptie eigenschappen af te leiden uit satellietmetingen van de UVAI. Daarbij is zowel gebruik gemaakt van traditionele methodes m.b.v. stralingstransport, als van nieuwe mogelijkheden van data-gestuurde methoden. Hoewel deze bemoedigend zijn, blijven er beperkingen en uitdagingen die verder moeten worden onderzocht. Deze worden in hoofdstuk 6 besproken met bijbehorende suggesties en vooruitzichten. Ondanks de uitdagingen is de verwachting dat een dataset van de wereldwijde aërosolabsorptie eigenschappen kan worden afgeleid uit de UVAI datareeks vanaf 1978, die gemeten is door meerdere satellieten. Een dergelijke dataset zal een belangrijke bijdrage leveren om het effect van absorberende aërosolen op het klimaatsysteem te kwantificeren.

1

INTRODUCTION

1

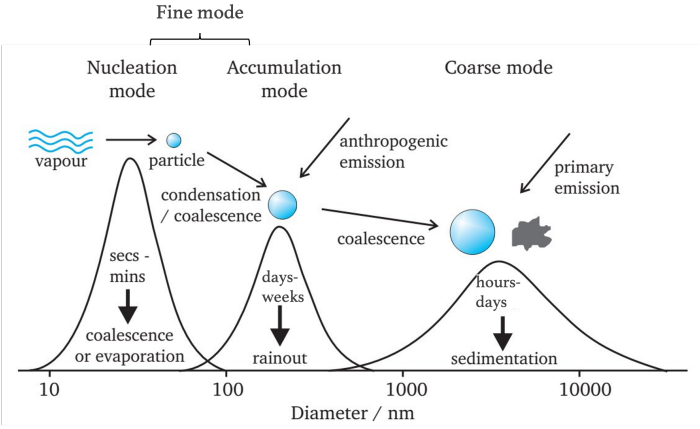
1.1. ATMOSPHERIC AEROSOLS

The Earth's atmosphere contains tens of millions of liquid droplets and solid particles that may have significant impacts on the Earth's climate and our health, the so-called atmospheric aerosols. The largest contribution to the mass of aerosols is from natural processes, with a best estimate emission rate of natural aerosols is 3062 Tg/yr, while the best estimate of anthropogenic particles emission rate is only 386 Tg/yr [Hidy, 2003].

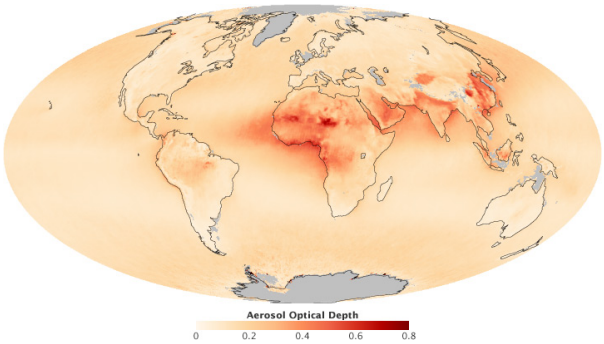
Directly emitted particles are also called **primary aerosols**, including sea salt, mineral dust, volcanic debris (ash and sulfates), biological materials, carbonaceous particles due to combustion processes. Materials generated from atmospheric chemical processes (gas-to-particle conversion), are termed **secondary aerosols**. Pollutant gases from either natural or anthropogenic processes, such as sulfur dioxide, nitrogen oxides, ammonia and reactive volatile organic compounds (VOCs), are common precursors for the secondary aerosols [Boucher et al., 2013].

The emitted particles are (internally or externally) mixed and aged in the atmosphere, becoming complicated mixtures. The internal mixture contains only homogeneous aerosol composition (including water-coated form), while the external mixture contains more than 2 types of aerosols. Figure 1.1 shows the phases of aerosol growth. Particles converted from gases (sulfates, ammonium salt, nitrates, organic carbon, etc.) usually have the smallest size (**nucleation mode**, diameter of 10-100 nm). The lifetime of these particles is short as they grow rapidly into **accumulation mode** (diameter of 100-1000 nm) through condensation and coagulation of gases and water vapour. Accumulation mode aerosols are usually removed from the atmosphere via scavenging by cloud droplets and raindrops. The representative aerosols of this kind are coagulated nuclei-mode aerosols, smog, soot and smoke particles from combustion processes. The nucleation mode and accumulation mode are together termed **fine mode**. Fine aerosols have a much longer lifetime of days to weeks (in the stratosphere volcanic sulfate aerosols can even stay months to years [Holzer-Popp et al., 2013]). Suspended in the atmosphere for longer time, they are the major contributor to the atmospheric visibility reduction. Particles larger than 1000 nm are termed **coarse mode**, and usually composed of mechanical-forming sea salt, mineral dust, volcanic ash, and those from agriculture and surface mining. Due to their large size, the coarse aerosols sediment at a significant rate, adding another pathway for deposition apart from rain-outs [Seinfeld and Pandis, 1998]. As a result, the lifetime of coarse aerosols is typically hours to days (fine volcanic ash can stay for months in stratosphere [Vernier et al., 2016]).

Geographically localized sources and sinks give aerosols an extreme spatial and temporal inhomogeneity in the atmosphere. The state-of-the-art satellite remote sensing techniques can provide maps



Figuur 1.1: Aerosol dynamics adapted from https://allaboutaerosol.com/aerosol-in-the-atmosphere/.



Figuur 1.2: Global average distribution of aerosols from June 2000 through May 2010, based on measurements of MISR. Source: https://earthobservatory.nasa.gov/features/Aerosols

of the global aerosol distribution. For example as shown in Figure 1.2, the aerosol optical depth (AOD, a measure of columnar aerosol amount) is largest in Asia and Africa. The major sources in eastern Asia are anthropogenic sulfates and carbonaceous aerosols produced by economic activities (e.g. industry, agriculture, transportation, daily life, etc.). The largest natural aerosol source over land is the Sahara Desert. The dust particles can even cross the Atlantic Ocean and reach the Amazon Basin via the northeasterly winds [Bristow et al., 2010]. The seasonal biomass burning events in tropical Africa and southeast Asia due to agricultural activities also present significant signals in the annual climatology map. On a vertical scale, the majority of aerosols are well-mixed within the boundary layer and decay rapidly with altitude in the troposphere. But sometimes, aerosol particles can also be elevated with a peak concentration at significant height above the surface, e.g. soil particles can be easily injected into free troposphere by large-scale vertical advection; smoke plumes generated by forest fires can lift into the upper troposphere by thermal radiative energy [Xu et al., 2018a]. The stratospheric aerosol loading is mainly due to volcanic activities. In 1991, the eruption of Mount Pinatubo in the Philippines ejected more than 20 million tons of sulfur dioxide to an altitude over 30 km, enhancing the aerosol concentration from an negligible ambient value to near 20-100 $\mu g/m^3$ [Hidy, 2003, Zhao et al., 1995].

1.2. EFFECTS OF AEROSOLS

Atmospheric aerosols have profound effects on air quality and public health, atmospheric visibility, local weather and global climate.

1.2.1. AEROSOL EFFECTS ON AIR QUALITY

In the context of environmental sciences, aerosol particles are one of the major pollutants and well-known as particulate matter (PM). Depending on the aerodynamic diameter (d_a), aerosols particles are normally divided into $PM_{2.5}$ ($d_a \leq 2.5\mu m$) and PM_{10} ($d_a \leq 10\mu m$) [Davidson et al., 2005]. Smaller particles consist of the gas-converted aerosols, combustion-produced particles and re-condensed organic and metal vapours. The larger particles also contain earth crust materials and fugitive dust from roads and industries. Both $PM_{2.5}$ and PM_{10} are inhalable particles that can penetrate and lodge deep inside lungs, while $PM_{2.5}$ is more health-damaging as it can penetrate into the alveolus and enter the blood system. A chronic exposure to these particles contributes to the risk of developing cardiovascular and respiratory diseases, as well as of lung cancer [Davidson et al., 2005, WHO, 2003, 2006]. The WHO air quality guideline levels of $PM_{2.5}$ are 10 $\mu g/m^3$ (annual mean) and 25 $\mu g/m^3$ (24 hour mean), and that of PM_{10} are 20 $\mu g/m^3$ (annual mean) and 50 $\mu g/m^3$ (24 hour mean) [WHO, 2006].

Moreover, aerosols are also an important pathway for viruses transmission. Studies have shown

that there are many respiratory diseases spread by airborne aerosols such as tuberculosis, measles, chickenpox and influenza [Davidson et al., 2005, Escombe et al., 2007, Hammond et al., 1989, Leclair et al., 1980, Tellier, 2006, 2009]. It is also suggested that the coronavirus disease pandemic in 2019 (COVID)-19 could be transmitted via airborne aerosols [Liu et al., 2020, van Doremalen et al., 2020].

1.2.2. AEROSOL EFFECTS ON VISIBILITY

The suspended particles in the atmosphere interact with electromagnetic radiation, leading to light extinction and visibility reduction. Atmospheric visibility is defined by the ability of our eyes to distinguish an object from the surrounding background [Horvath, 1981]. It is usually quantified by the visual range (unit: km), which expresses how far one can see in a given direction. Objects at distances over hundreds of kilometers can be observed under high visibility conditions, while targets only a few kilometers ahead can be detected under low visibility conditions [Hyslop, 2009]. Visibility lower than 100 m is considered as 0.

Figure 1.3 show two cases of impaired visibility due to the high loading of particulate matter. Particles with diameters in the range of 0.1-1.0 μm are the most effective at degrading visibility [WHO, 2006], i.e. accumulation mode in Figure 1.1, as the lifetime of this aerosol regime is longer than the nucleation and coarse modes. Atmospheric visibility is important to traffic safety in all forms (roads, sailing and aviation).

Moreover, visibility is also an indicator of air quality. It is found that clear sky visibility has decreased globally over land from 1973 to 2007 [Wang et al., 2009]. One of the most famous efforts to improve the visibility is the 'Blue Sky Project' in 1998 that investigated by how much emissions should be reduced to increase the blue sky frequency in Beijing, where the 2008 Summer Olympic Games was hosted. Many actions had been taken (some of them continues until now), including mitigating high-polluted industries outside Beijing, temporally shutting down some industries, replacing coal fuels and leaded gasoline, and road space rationing (a travel demand management strategy, also called no-drive days), etc. The 'Blue Sky Days' continued increasing yearly from 100 days in 1998 to 246 in 2007 [Andrews, 2008].

1.2.3. AEROSOL EFFECTS ON CLIMATE

The aerosol-radiation interactions also have impacts on the Earth's climate system. Aerosol particles can **directly scatter or absorb** solar radiation and absorb thermal infrared radiation. Most aerosols are non-absorbing, such as sulfates, nitrates, and sea salt, etc. They reflect nearly all the incoming sunlight into the outer space, resulting in a cooling effect on the Earth's climate. Dark-colored aerosols (black

(a) 24-Feb-2019 in Nanjing

(b) 13-May-2017 in Shanghai

Figuur 1.3: Two cases of visibility reduction due to high loading of particulate matter. The dominant pollutant in both cases is $PM_{2.5}$. (a) $PM_{2.5}$ was 101 $\mu g/m^3$ in Nanjing on 24-Feb-2019, 06:00 a.m.; (b) $PM_{2.5}$ was 64 $\mu g/m^3$ in Shanghai on 13-May-2017, 05:30 a.m. (Data source: https://www.zq12369.com).

carbon), mineral dust and some organic aerosols can directly absorb the solar radiation, the solar radiation reflected by the surface-atmosphere-cloud system and thermal radiation, posing a positive radiative forcing to the climate [Haywood and Boucher, 2000, Takemura et al., 2005]. Absorption of long-wave thermal infrared radiation is negligible except for large particles like mineral dust [Highwood et al., 2003].

Aerosols also have an **in-direct effect** on clouds. Some particles, such as sulfates [Hegg et al., 1993], organic aerosols [Rivera-Carpio et al., 1996] and sea salt [Feingold et al., 1999], work as cloud condensation nuclei (CCN), affecting the cloud albedo and lifetime [Haywood and Boucher, 2000]. The increased particle number concentration can decrease the cloud droplet effective radius, leading to higher cloud albedo and cloud optical depth [Twomey et al., 1974]. Moreover, the reduced cloud droplet effective radius is less likely to form precipitation, which increases the liquid water content, the cloud lifetime and the cloud thickness [Albrecht, 1989, Pincus and Baker, 1994].

Like scattering aerosols, absorbing particles also have indirect effects on climate by changing the number of liquid cloud droplets, enhancing precipitation in mixed-phase clouds, and changing ice particle number and cloud extent. Besides, absorbing aerosols heat up the atmosphere at the same time reducing the sunlight absorption by surface, leading to higher static stability at low-level and lower surface moisture fluxes. This further reduces the low-cloud coverage, or the probability of cloud formation [Hansen et al., 1997, Ramanathan et al., 2001]. The above is considered as the so-called **semi-direct effect** of absorbing aerosols on clouds. Soot inside the clouds droplets and ice crystals can also decrease the albedo of clouds by increasing the absorption by droplets and ice crystals [Bond et al., 2013, Ramanathan and Carmichael, 2008]. Moreover, dark-colored aerosols dim the surface albedo by depositing on ice and snow, and the heat produced by light absorption speeds up the ice/snow melting process [Hansen and Nazarenko, 2004, Nair et al., 2013].

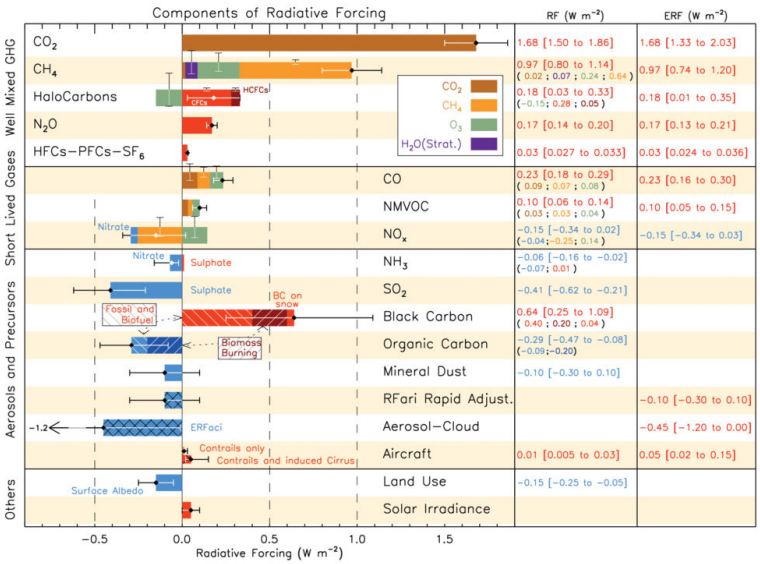
According to IPCC, atmospheric aerosols have a dominating cooling effect on climate with a partially warming offset [Pachauri et al., 2014] (Figure 1.4). The anthropogenic aerosol effective radiative forcing due to interaction with radiation and clouds from 1750 to 2011 is -0.9 [-1.9 to -0.1] W/m^2 (medium confidence), among which the aerosol direct effects on climate is accessed to be –0.45 [–0.95 to +0.05] W/m^2. The contribution of absorbing aerosols deposited on bright surfaces is +0.04 [+0.02 to +0.09] W/m^2. Despite improvements have been achieved with observational and modelling breakthroughs, aerosols are still regarded as the largest uncertainty to the total radiative forcing assessment [Pachauri et al., 2014] (for comparison, the radiative forcing of well-mixed greenhouse gases is 2.83 [2.54 to 3.12] W/m^2). The aerosol effect on climate is not fully understood, due to the complex aerosol optical properties, high variability in space and time, and the poor understanding of aerosol emissions and dynamic processes (particularly aerosol-cloud interactions) [Carslaw et al., 2013, Loeb and Su, 2010, McComiskey and Feingold, 2008, McComiskey et al., 2008, Penner et al., 2001, Watson-Parris et al., 2020].

1.3. AEROSOL OPTICAL PROPERTIES

Aerosol optical properties are one of the major error sources in the aerosol radiative forcing assessment. Depending on the chemical composition, the particle size distribution and the morphology (spherical or non-spherical, regular or irregular), aerosols present various levels of capability in absorption and scattering. Besides, the environmental relative humidity can also influence the aerosol chemical composition and size distribution via multiple processes (e.g.aqueous-phase reactions, gas-particle partitioning, etc.) [Sun et al., 2013].

1.3.1. SCATTERING REGIMES

The type and magnitude of scattering depend on the relationship between the particle size (represented in diameter, D, unit: usually expressed in nm or μm) and the incident radiative wavelength (λ, unit: usually expressed in nm or μm), represented by a dimensionless size parameter X [Feingold et al.,



Figuur 1.4: Radiative forcing (RF) of climate change during the Industrial Era shown by emitted components from 1750 to 2011. The picture is taken from IPCC Climate Change 2013 The Physical Science Basis [Stocker et al., 2013].

1999]:

$$X = \frac{\pi D}{\lambda} \tag{1.1}$$

Depending on X, there are three major scattering regimes. When X is considerably larger than 1, which means the particle size D is much larger than the wavelength of incident radiation λ (unit: usually expressed in nm or μm), **geometric scattering** happens. This is usually the cases for the larger water droplets (e.g. clouds, drizzle, raindrops, $D > 10\mu m$) in the visible to mid-infrared bands (0.4 to 8 μm). Geometric scattering is non-selective, i.e. it scatters all wavelengths equally. That is why clouds look white. If clouds are optically thick enough, little light can penetrate through the clouds and the cloud bases appear dark (Figure 1.5a).

By contrast, when the particle size is far smaller than the wavelength of incident radiation ($X << 1$), **Rayleigh scattering** occurs. Rayleigh scattering is mainly contributed by small gas molecules ($D < 0.1\mu m$) and is highly wavelength-dependent (λ^{-4}). The scattering in ultra-violet to blue channels (0.2 to 0.5 μm) is stronger than that in longer wavelengths, thus the color of the sky is highly-saturated blue (if no particles exist, Figure 1.5b). At dawn and dusk, the incoming solar radiation travels a longer distance so that the shorter blue band is completely scattered, resulting in a reddish sunrise or sunset (Figure 1.5c).

When X is near one (also called resonant regions [Barber and Yeh, 1975]), i.e. the particle size and the wavelength are of similar magnitude, then **Mie scattering** applies. Mie scattering influences the spectrum ranging from near ultra-violet to the mid-infrared band (0.3 to 8 μm). Aerosols and small cloud particles ($0.1 < D < 10\mu m$) are the major contributor to Mie scattering. The spectral dependence of Mie scattering is much weaker than Rayleigh scattering. Thus, in the aerosol-abundant regions, the sky looks gray or white (Figure 1.5d). Mie scattering is also responsible for glares around the sun.

(a) White clouds with dark bases due to geometric scattering.

(b) Blue sky due to Rayleigh scattering.

(c) Red sunrise due to Rayleigh scattering.

(d) Low saturation of view and reduced visibility due to Mie scattering.

Figuur 1.5: Examples of scattering: (a) picture taken in Utrecht, the Netherlands on 11-Aug-2018 (geometric scattering); (b) picture taken in Utrecht, the Netherlands on 04-Apr-2020 (Rayleigh scattering); (c) picture taken in De Bilt, the Netherlands on 22-Jan-2019 (Rayleigh scattering); (d) picture taken in Tokyo, Japan on 14-Jan-2017 (Mie scattering).

1.3.2. CALCULATION OF AEROSOL OPTICAL PROPERTIES

Maxwell's equations are the basis of theoretical and computational methods describing light scattering. Given the aerosol refractive index, particle size distribution, morphology and relative humidity, the aerosol optical properties, i.e. the single scattering albedo (ω_0) and the phase function (or asymmetry parameter g) and the optical depth (τ), can be calculated via different methods that solves the Maxwell's equations.

For extreme small particles and large particles, Rayleigh scattering and geometric optics give approximation solutions of Maxwell's equations. In Mie scattering regimes, the German physicist Gustav Mie developed an analytical solution that is used for spherical particles, known as the Mie Theory [Mie, 1908]. Although the Mie Theory is satisfactory to model optical properties for the majority of aerosols, non-spherical dust aerosols shows significantly different properties compared with Mie calculations [Dubovik et al., 2002a, 2006, Hu et al., 2017]. Now with the improvement of computational efficiency, many exact computational solutions for non-spherical particles have been proposed and applied, e.g. the Finite-Difference Time-Domain (FDTD) [Kane Yee, 1966], the Discrete Dipole Approximation (DDA) [Purcell and Pennypacker, 1973] (used in CALIOP [Liu et al., 2005]),

T-matrix [Mishchenko et al., 1996, Waterman, 1965] (used in MODIS [Remer et al., 2006] and OMAERO [Torres et al., 2002b]), etc.

The following sections focus on how to calculate optical properties for spherical particles using the Mie Theory.

1.3.3. MICRO-PHYSICS OF A SINGLE AEROSOL PARTICLE (MIE THEORY)

Given a certain relative humidity (RH, unit: %), the chemical composition of a particle determines its **complex refractive index** $m_\lambda = n_\lambda + k_\lambda \boldsymbol{i}$, where n_λ is the real part of the refractive index at wavelength λ determined by the light travelling speed through a medium. The real part is the scattering component, and the imaginary part of the refractive index k_λ is responsible for absorption. A particle with $k_\lambda = 0$ means purely scattering.

For a spherical particle with diameter of D or radius r ($D = 2r$, unit: usually expressed in nm or μm) and complex refractive index $m_\lambda = n_\lambda + k_\lambda \boldsymbol{i}$, the scattering, absorption and extinction efficiency (Q_{sca}, Q_{abs} and Q_{ext}, dimensionless) can be calculated by the Mie Theory. Scattering redirects the incoming electromagnetic radiation, while absorption transfers the radiation into longer-wavelength internal energy heating the surroundings. The summed effect of scattering and absorption is called extinction (or attenuation).

The corresponding **extinction cross section** (σ_{ext}, unit: usually expressed in nm^2 or μm^2) is calculated (similar relationships for σ_{sca} and σ_{abs}):

$$\sigma_{ext,\lambda} = \frac{\pi D^2}{4} Q_{ext,\lambda} = \pi r^2 Q_{ext,\lambda} \tag{1.2}$$

The scattered intensity of radiation is angular-dependent, which is described as the scattering **phase function** ($P(\Theta, \lambda)$) [Seinfeld and Pandis, 1998]:

$$\frac{1}{2}\int_0^\pi P(\Theta,\lambda)\sin\Theta d\Theta = 1 \tag{1.3}$$

where Θ is the scattering angle between the incident and scattered radiation. The phase function can be concisely presented by the **asymmetry parameter** g ($-1 \le g \le 1$, dimensionless), defined as:

$$g_\lambda = \frac{1}{2}\int_0^\pi \sin(\theta)\cos(\theta)P(\theta,\lambda)d\theta \tag{1.4}$$

g describes the tendency of forward scattering of a particle. Forward scattering refers to $g > 0$ while backward scattering refers to $g < 0$. Aerosol and cloud particles tend to have a relatively large and positive g, i.e. strong forward scattering. That is why the scene of aerosols and clouds looks bright if an observer is looking at the forward scattered light.

1.3.4. MICRO-PHYSICS OF A GROUP OF PARTICLES (MIE THEORY)

So far only optical properties of a single particle have been discussed. In reality, the aerosol properties are measured for a larger number of particles with different types. Thus, for each aerosol type i, a **particle size distribution function** ($N^i(r)$) is introduced:

$$N^i(r) = \int_{r1}^{r2} n(r)dr \tag{1.5}$$

The size distribution function $N^i(r)$ can be mathematically characterized by different functions, such as a power-law, gamma and log-normal function, etc. The concept of **effective radius** (r_e, unit: usually expressed in nm or μm) is proposed to represent the particle size of the whole bulk of aerosols:

$$r_e = \frac{\int_{r1}^{r2}\pi r^3 n(r)dr}{\int_{r1}^{r2}\pi r^2 n(r)dr} \tag{1.6}$$

At a given relative humidity, the volume **extinction coefficients** ($\beta_{ext,\lambda}$, unit: usually expressed in nm^{-1} or μm^{-1}) for a group of mixed particles is the summation of the integrated optical properties over size distribution functions for all aerosol types (similar calculations for volume scattering $\beta_{sca,\lambda}$ and absorption coefficient $\beta_{abs,\lambda}$):

$$\beta_{ext,\lambda} = \sum_i \int_{r1}^{r2} \sigma^i_{ext,\lambda}(r) N^i(r) dr \tag{1.7}$$

, where σ^i_{ext} and $N^i(r)$ are the extinction cross section and the size distribution function for an aerosol type i. The **single scattering albedo** (SSA, ω_0, dimensionless), which describes the fraction of scattering of the total extinction, is defined as:

$$\omega_{0,\lambda} = \frac{\beta_{sca,\lambda}}{\beta_{ext,\lambda}} = \frac{\beta_{sca,\lambda}}{\beta_{sca,\lambda} + \beta_{abs,\lambda}} \tag{1.8}$$

The value of ω_0 is between 0 and 1. ω_0 of 1 indicates purely scattering and ω_0 of 0 indicates that all light is absorbed. The bulk phase function is weighed by scattering cross section σ^i_{sca} of each aerosol type:

$$P(\Theta, r, \lambda) = \sum_i \frac{\int_{r1}^{r2} P^i(\Theta, r, \lambda) \sigma^i_{sca,\lambda}(r) N^i(r) dr}{\int_{r1}^{r2} \sigma^i_{sca.\lambda}(r) N^i(r) dr} \tag{1.9}$$

The corresponding g_λ can be calculated by Equation 1.4.

The total light attenuation due to atmospheric aerosols is presented by the **aerosol optical depth** (AOD, τ, dimensionless), which is the calculated as the aerosol extinction coefficients β_{ext} from the surface to the top of the atmosphere (TOA):

$$\tau_\lambda = \int_0^{TOA} \beta_{ext,\lambda}(z) dz \tag{1.10}$$

Under single scattering approximation, the Beer-Lambert Law can be applied to calculate the τ_λ, thus:

$$\frac{I}{I_0} = \exp(-\beta_{ext,\lambda} z) = \exp(-\tau_\lambda) \tag{1.11}$$

where I and I_0 are the irradiance at any distance z and the incident irradiance, respectively. The aerosol extinction due to absorption is presented by the absorption aerosol optical depth (AAOD, $\tau_{a,\lambda}$, dimensionless):

$$\tau_{a,\lambda} = \tau_\lambda \times (1 - \omega_{0,\lambda}) \tag{1.12}$$

The wavelength dependence of τ_λ is described by the **Ångström Exponent** (ÅE, α, dimensionless):

$$\alpha = -\frac{\log(\tau_{\lambda_1} / \tau_{\lambda_2})}{\log(\lambda_1 / \lambda_2)} \tag{1.13}$$

ÅE is reversely related to the aerosol particle size. The smaller the particle, the larger ÅE, i.e. the stronger wavelength-dependence. Likewise, the spectral dependency of $\tau_{a,\lambda}$ (dimensionless), i.e. absorption Ångström Exponent (AÅE, α_a, dimensionless), can be calculated by:

$$\alpha_a = -\frac{\log(\tau_{a,\lambda_1} / \tau_{a,\lambda_2})}{\log(\lambda_1 / \lambda_2)} \tag{1.14}$$

1

1.3.5. AEROSOL MODELS IN APPLICATIONS

Aerosol models are used in chemical transport models and/or atmospheric retrieval algorithms. Aerosol models describe the aerosol chemical, physical and/or optical properties as a function of wavelength and relative humidity.

In reality, aerosol particles will not consist of a single substance, but are a mixture of different substances (externally mixed). That is why aerosol models in the majority of operational satellite retrieval algorithms are named by aerosol optical properties (scattering or absorbing), combined with information of sources (natural or anthropogenic, oceanic or continental, rural or urban), instead of using individual chemical component (e.g. [Liu et al., 2005, Remer et al., 2006, Torres and Chen, 2013, Torres et al., 2002b]). Note that even if the same naming used, the assumed chemical compositions and the calculated optical properties can be different among algorithms. For example, the biomass burning aerosols in OMAERO has a mean fine radius of 0.07 μm and a mean coarse radius of 0.87 μm [Torres et al., 2002b], while those in CALIOP aerosol retrieval algorithm are 0.14 μm and 3.73 μm, respectively [Liu et al., 2005].

Instead of aerosol models for mixtures, it is easier to understand optical properties of aerosols consisting of homogeneous chemical composition, which is usually used in chemical transport models. Below, aerosols of five representative types are introduced, i.e. sea salt, sulfates, mineral dust, black carbon and organic carbon based on the aerosol models in the chemical transport model (CTM) MERRA-2 [Randles et al., 2017] and CAMS [Bozzo et al., 2017]. In addition, although excluded in many global CTMs [Bauer et al., 2007a, Bian et al., 2017, Dall'Osto et al., 2009], the nitrate aerosols will also be introduced in this section as they have significant impacts on tropospheric chemistry and aerosol radiative forcing [Adams et al., 2001, Bellouin et al., 2011, Liao et al., 2003]. Besides, with the trend of decreasing sulfur dioxide and increasing ammonia emission, nitrate aerosols are expected to become more important in the future [Bian et al., 2017]. Below the aerosol optical properties are reported at (approximately) 550 nm by default, unless specifically indicated.

SEA SALT

Sea salt particles are generated by natural bubble bursting and sea spray. They are widely distributed and account for about 30% of the global column AOD [Holzer-Popp et al., 2013]. Sea salt particles are highly hygroscopic (tend to take up water and increase in size with increasing relative humidify), and they are considered as efficient cloud condensation nuclei [Boucher et al., 2013, Solomon, 2007]. The chemical composition of sea salt is complex and varies with location and time, but mainly includes chemical ion of Na^+, Cl^-, K^+, Mg^{2+}, Ca^{2+}, SO_4^{2-}, etc. [Tang et al., 1997].

Sea salt has a wide size range from fine to coarse particles. The dry effective radius (r_e) can have orders of difference in magnitude from 0.01 to 10 μm. The asymmetry factor g varies from 0.2 to 0.9, depending on particle size and relative humidity. Sea salt is assumed to be purely scattering, thus its single scattering albedo ω_0 is near unity and barely varies with wavelength (within the visible band, $\lambda \leq 1 \mu m$) and relative humidity [Bozzo et al., 2017, Randles et al., 2017].

In operational algorithms, sea salt is included in the category oceanic aerosols (or marine aerosols). The oceanic aerosols also contain the non-sea-salt sulfates (mainly from dimethyl sulfide, i.e. DMS) and nitrates (both marine and continental sources), organic matters, and mineral dust from continents [Fitzgerald, 1991, O'Dowd et al., 1997]. Despite of the presence of absorbers, the oceanic aerosols are still predominantly scattering [Liu et al., 2005, Remer et al., 2006, Torres et al., 2002b].

SULFATES

Sulfates are usually formed by gas-to-particle conversions of sulfur dioxide (SO_2) from geological (volcanic eruption) and anthropogenic activities (industrial pollution), as well as DMS from biogenic processes [Ghahremaninezhad et al., 2019]. Sulfate aerosols consist of sulfuric acid (SO_4^{-2}) or ammonium sulfate ($(NH_4)_2SO_4$), depending on the ammonia availability [Boucher et al., 2013]. Note that sea salt is also a source of sulfates, but usually it is not included in the sulfate aerosol model.

Sulfates are a part of industrial aerosol (urban aerosols) in some algorithms [Liu et al., 2005, Remer et al., 2006, Torres et al., 2002b] (industrial/urban aerosols may also consist of nitrates, black carbon and organic carbon, dust, etc. that produced by anthropogenic processes), but in some algorithms they are treated as an individual category (OMAERUV).

Sulfates are hygroscopic, thus their size varies with relative humidity. Nevertheless, the dry sulfate particles are relatively small with r_e at magnitude of 0.15 μm. Their g varies from 0.6 to 0.8 (slightly decreases with wavelength and increases with relative humidity). Sulfate particles are also efficient in scattering. The typical value of ω_0 is near unity and barely changes with wavelength (within the visible band) and relative humidity [Bozzo et al., 2017, Randles et al., 2017].

MINERAL DUST

Mineral dust consists of naturally-formed particles over desert, dry lake beds and semi-arid regions, as well as that produced by agricultural activities, and industrial activities like transportation, coal combustion, cement manufacturing, metallurgy and waste incineration [Boucher et al., 2013, Solomon, 2007]. Mineral dust has the largest mass emission rates, column mass burden and optical depth among all aerosol types [Satheesh and Krishna Moorthy, 2005]. In retrieval algorithms, dust aerosols usually indicate desert dust [Liu et al., 2005, Remer et al., 2006, Torres et al., 2002b]. Note that dust particles are usually non-spherical. Thus, instead of the Mie Theory, the optical properties of dust aerosol are calculated by DDA (used in CALIOP, [Liu et al., 2005]) or T-matrix (used in OMAERO, [Remer et al., 2006, Torres et al., 2002b]).

Compared with other aerosol types, dust particles are rather coarse, with r_e at magnitude from 0.1 to 100 μm, but only those with size smaller than 10 μm are subject to long-range transport [Moosmüller et al., 2009]. Dust aerosols are hydrophobic, which means their optical properties do not change with relative humidity. The g of dust particles is usually larger than 0.7, presenting a strong forward scattering tendency. The absorption of dust is driven by iron oxide content and increases towards the shorter wavelengths [Moosmüller et al., 2009]. The ω_0 of dust usually varies between 0.6 and 0.9 at 550 nm and increases with wavelength within the visible band [Bozzo et al., 2017, Randles et al., 2017].

BLACK CARBON

Black carbon, also referred to as elemental carbon, is mainly produced by incomplete combustion of biomass and fossil fuels of both natural and anthropogenic processes [Albrecht, 1989, Boucher et al., 2013, Solomon, 2007]. Unlike other aerosol types, black carbon is dark-colored, indicative of its strong capability of light absorption from the UV-visible band [Hyslop, 2009].

Freshly emitted black carbon particles are hydrophobic and mostly externally mixed with non-refractory compounds. However, these particles acquire hygroscopic materials from condensation of sulfates and organic matter, or coagulation with hygroscopic particles or cloud drops, which is called the coating process of black carbon [McMeeking and Coe, 2011]. The coated black carbon can serve as the cloud condensation nuclei. The absorption is significantly enhanced by the coating. According to Shiraiwa et al. [2010], the mass absorption cross section (at near 532 nm) can be doubled if the shell size is two times of the core size.

The dry black carbon is fine particles with r_e of 0.05 μm. Since they are small, the scattering direction is less forward (g of 0.3). Black carbon is a strong absorber with a constant imaginary refractive index of about 0.8 within the visible band [Bond and Bergstrom, 2006] (i.e. AÅE≈1). Its ω_0 is usually less than 0.4 at 550 nm and slightly decreases with wavelength (within the visible band) [Bozzo et al., 2017, Randles et al., 2017]. For hydrophilic particles, not only the absorption increases during the aging process, but also the particle size and scattering cross section [Cheng et al., 2014].

ORGANIC CARBON

Organic carbon is usually generated together with black carbon [Boucher et al., 2013, Solomon, 2007]. The directly produced organic carbon from combustion is also called primary organic aerosol. The secondary organic aerosol is formed as the oxidation products of certain volatile organic compounds

(VOCs) condense on existing aerosols [Chung and Seinfeld, 2002]. Organic carbon can also be separated into water-soluble and water-insoluble. The former accounts for 20%-70% of the total organic carbon and is mainly secondary organic aerosols. Similar to hydrophilic black carbon, water-soluble organic carbon also serves as the cloud condensation nuclei [Tang et al., 2016].

The dry organic carbon particles have a r_e of 0.1 μm. Its g is typically 0.6, showing a moderate forwarding scattering. Unlike black carbon, organic carbon has a ω_0 usually larger than 0.95 which slightly decreases with wavelength (within the visible band, $\lambda \leq 1\mu m$) [Bozzo et al., 2017, Randles et al., 2017]. Both g and ω_0 increase with relative humidity for hydrophilic organic carbon.

Although organic carbon is considered to be scattering-efficient, several studies have shown that the brown carbon, a sub-group of organic carbon, is also a light absorber [Kirchstetter et al., 2004, Mukai and Ambe, 1986, Patterson and McMahon, 1984]. It is produced by tar materials from smoldering fires or coal combustion, breakdown products from biomass burning, organic compounds emitted from soil, and VOCs emitted by vegetation [Yan et al., 2018]. Brown carbon is less absorbing than black carbon, but its absorption has a stronger spectral dependence (absorption decreases towards longer wavelength). The AÅE of brown carbon can larger than 3 while that of black carbon is near unity [Yang and Huebert, 2009]. The large variation of AÅE is due to the mixture of black carbon and brown carbon. It has been reported that combustion aerosols with a low SSA have an AÅE near unity, whereas combustion aerosols with a high SSA have an higher AÅE [Lewis et al., 2008]. In satellite retrieval algorithms, black carbon and organic carbon are the major components of biomass burning aerosols (also called smoke, soot, carbonaceous aerosols). Due to the presence of absorbing particles, biomass burning aerosols are also referred to as fine absorbing aerosols (corresponding to the mineral dust, the coarse absorbing aerosols).

NITRATES

Nitrates are usually chemically formed from ammonia (NH_3) and nitrogen acid (HNO_3), and photochemical products of the oxidation of nitrogen oxides (NO_x) [Hauglustaine et al., 2014]. The common sources of NH_3 and NO_x consist of both natural and anthropogenic processes, including excreta from domestic and wild animals, synthetic fertilizers, oceans, biomass burning, crops, human populations and pets, soils, lightening, industrial processes and fossil fuels [Bauer et al., 2007a, Bouwman et al., 1997]. Compared with other aerosols, less attention was given to ammonia (NH_3) and ammonium (NH_4^+) aerosols because of their complex chemical formation and reactions [Bian et al., 2017, Dall'Osto et al., 2009, Myhre et al., 2013]. However, with the decreasing SO_2 and increasing NH_3 emission in future, less NH_3 is needed to neutralize the the strong sulfuric acid (H_2SO_4), yielding more NO_3^- and NH_4^+. As a result, recently there is an increasing number of CTMs that include these two components [Bian et al., 2017].

Studies have shown that the light-scattering properties of nitrates do not significantly differ from that of sulfates as long as their dry particles have the same size distribution [Tang and Munkelwitz, 1994, Tang et al., 1997, Tang, 1996]. Some studies treat the micro-physical properties of nitrates in the same way as sulfates (e.g. [Drugé et al., 2019, Lowenthal et al., 2000]). Nitrates usually have a constant ω_0 of near unity and g between 0.6 and 0.8 with wavelength (within the visible band) and relative humidity. Nitrates particles in the nucleation mode are small with r of the order of 0.015 μm, while they can coat onto the existing particulate matter to form coarse particles with r of typically 0.15 μm [Bauer et al., 2007b, Zhang et al., 2012].

In operational satellite retrieval algorithms, nitrates are usually included together with sulfates in the industrial/urban aerosol group [Liu et al., 2005, Remer et al., 2006, Torres et al., 2002b].

1.4. AEROSOL MEASUREMENT PRINCIPLES

Aerosol properties can be measured in laboratory experiments, or by in situ and remote sensing techniques. In situ measurements require contacting with the target objects (e.g. an air parcel), whereas remote sensing techniques derive information from propagating signals (e.g. light) without

physical interactions with the targets. Satellite measurements, one of the remote sensing techniques, have the advantage of global coverage. Satellite remote sensing techniques developed in recent years play a key role in the global monitoring of aerosols.

Deriving physical parameters from remote sensing is an **inverse problem**: the properties that we want to know are not directly measured but have to be derived from the measured signals. Inverse problems are usually ill-conditioned, in other words, the amount of information in the measurements is not sufficient to derive a unique solution. To find a solution, the observations have to be combined with assumptions on one or more parameters, which is commonly referred as a priori information.

The following subsections will introduce the basic principles and techniques of tropospheric aerosol measurements with a particular emphasis on satellite remote sensing.

1.4.1. PASSIVE SENSORS (SHORTWAVE)

Passive sensors measure the Earth's reflectance at the top of the atmosphere (TOA), which is written as:

$$R(\lambda) = \frac{\pi I(\lambda)}{\mu_0 E_0(\lambda)} \tag{1.15}$$

where I and E_0 are the outgoing radiance and the solar irradiance as a function of wavelength λ at TOA; μ_0 is the cosine of the solar zenith angle. In the simplest case, by assuming a cloud-free horizontally homogeneous atmosphere over a Lambertian surface (isotropically reflecting), the TOA reflectance is affected by the interaction between photons and the Earth's atmosphere-surface system:

$$R(\lambda) = R_a(\lambda) + \frac{A_s(\lambda)}{1 - A_s(\lambda)s(\lambda)} T(\lambda) \tag{1.16}$$

where R_a is the atmospheric contribution (molecules and particles) to the total TOA reflectance, A_s is the surface albedo, s is the spherical albedo of the atmosphere, and T is the direct and diffuse transmittance of the atmosphere. The factor $1/(1 - A_s(\lambda)s(\lambda))$ represents the multiple interactions between the atmosphere and the surface. Note that all terms are dependent on the solar and viewing geometries. Except for A_s, all other terms also depend on the surface pressure. Note that the Equation 1.16 does not hold if the surface reflectance is significantly non-isotropic. In such cases, the surface reflectance has to be described by the bidirectional reflectance distribution function (BRDF). Due to the simplicity of Equation 1.16, the Lambertian surface assumption is often applied in satellite remote sensing retrievals.

In practice, satellite instruments only measure the total TOA reflectance. The molecular contribution (Rayleigh scattering) can be calculated with known pressure and temperature profiles [Torres et al., 2002b], but determining the surface contribution is challenging, particularly over land where the surface reflectivity in the visible and near-infrared is much higher than that over water [Kaufman et al., 1997a]. As as result, when a sensor observes a scene from only one direction, assumptions usually have to be made on the surface reflectance.

Due to the ill-posed inverse problem, not all aerosol properties mentioned in Section 1.3.4 can be retrieved simultaneously from the observed reflectance, and there is a hierarchy between various parameters [Kaufman et al., 1997a]. A single measurement is limited to retrieving a singular parameter, and the priority is usually given to AOD. Multiple measurements with multi-angle and/or polarization techniques allow to retrieve additional aerosol properties, but depending on what is to be retrieved and what measurements are made, assumptions have to be made on one or more of the following parameters: particle size distribution, refractive index, aerosol vertical distribution and surface reflectance.

1.4.2. ACTIVE SENSORS

Active Light Detection and Ranging (lidar) is widely used to determine vertical profiles of aerosol optical properties. A lidar emits laser pulse and measures the backscattered signals from which extinction

1

coefficient profiles and other aerosol optical properties can be estimated. Under cloud-free conditions, the lidar equation that relates the backscattered light power and the signal detected by the lidar receiver, can be written as (molecular absorption is ignored):

$$P(R) = \frac{E_0 \eta_L}{R^2} O(R) \beta(R) exp\left[-2\int_0^R \alpha(r)dr\right] \tag{1.17}$$

, where P (unit: Wsr^{-1}) is the signal due to molecules and particles from distance R (unit: km); E_0 is the transmitted laser energy (unit: J); η_L describes the efficiency of the optical and detection unit (unit: kms^{-1}); O is the overlap between the outgoing beam and the field of view of the receiver (unit: km^2), which is usually assumed to be complete (i.e. $O(R) = 1$); β (unit: $km^{-1}sr^{-1}$) and α (unit: km^{-1}) are the backscatter coefficient and extinction coefficient, respectively.

For elastic backscatter lidar (i.e. the emitted and the received pulse are at the same wavelength), the molecular contribution to the total signal can be determined from the temperature and pressure profiles or approximated from the standard atmosphere. The variables in the lidar equation α and β are the target to be derived from the lidar signals. One of the most common solutions, also called Klett method [Ansmann and Müller, 2005, Fernald, 1984, Klett, 1981], is introducing the concept of aerosol extinction to backscatter ratio, i.e. lidar ratio (unit: sr):

$$L_{aer}(r) = \frac{\alpha_{aer}(r)}{\beta_{aer}(r)} \tag{1.18}$$

L_{aer} depends on particle size, shape and chemical compositions (typical values at 532 nm of marine/oceanic aerosol: 20-30 sr; biomass burning/smoke aerosol: 40-70 sr; dust aerosol: 50-60 sr; urban/industrial aerosol: 40-60 sr) [Kim et al., 2018, Liu et al., 2005, Müller et al., 2007, Omar et al., 2009].

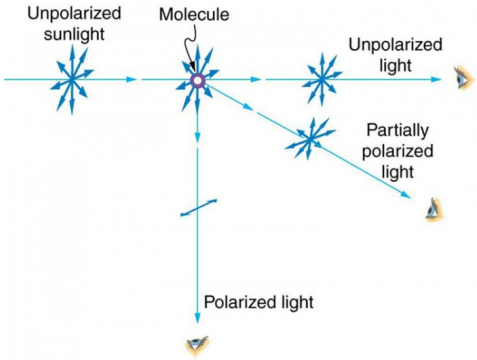
In practice, L_{aer} is always a first guess without knowing the truth, which becomes the major uncertainty in determining the aerosol extinction profiles. This unsatisfactory situation can be improved with the Raman lidar (the received signal has a wavelength shift relative to the emitted one) and the high spectral resolution lidar (HSRL). They can directly and unambiguously determine the vertical profiles of α_{aer}, β_{aer} and thus L_{aer}, without making assumptions on atmospheric scattering properties [Ansmann and Müller, 2005, Eloranta, 2005, Müller et al., 2007].

1.4.3. POLARIMETRY

Electromagnetic radiation is a transverse wave, whose electric field of light (E) oscillates perpendicularly to the direction of propagation (Z). Sunlight, lamps, and candle flames are usually unpolarized as the direction of their electric fields fluctuate randomly in time. On the contrary, a polarized wave (e.g. a laser) has a preferred orientation over time.

Transforming unpolarized light into polarized light is known as polarization, which can be done through transmission, reflection, refraction and scattering. In the context of atmospheric remote sensing, polarization is mainly caused by scattering. The unpolarized sunlight passes through the atmosphere, strikes the air molecules and particles and vibrates their electrons perpendicularly to the light travelling direction. The vibrating electrons then produce an electromagnetic wave that is radiated outward in all directions and strikes neighboring atoms, leading to their electrons into vibrations at the same frequency. Through scattering, the unpolarized sunlight eventually becomes partially polarized skylight (dominated by Rayleigh scattering) [Talmage and Curran, 1986, Zhou et al., 2013] (Figure 1.6a). The partial polarization of scattered light leads to photographs characterized by a washed-out sky, which can be corrected by a Polaroid filter to block the polarized light (Figure 1.6b).

Depending on the orientation of an electric field, polarization can be categorized into linear, circular and elliptical polarization (Figure 1.7). A linearly polarized light is when its electric field is confined to a plane along the propagation direction. If the electric field has a constant amplitude but its direction rotates at a constant rate along the propagation direction, then the wave is circularly polarized.



(a) (Source: https://courses.lumenlearning.com/physics/chapter/27-8-polarization/))

(b) (Source: https://www.goldenvisionclinic.com/optical-boutique/

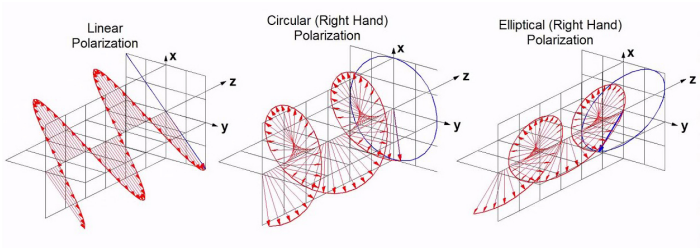
Figuur 1.6: Partially polarized skylight.

Circular polarization consists of two linear components with a same amplitude, and perpendicular directions with a phase difference of $\pi/2$. The elliptical polarization results from the combination of two linear components with differing amplitudes and/or a phase difference that is not $\pi/2$. Both the circular and linear polarized light are special cases of elliptically polarized light [Hansen and Travis, 1974, Sekera, 1957, Talmage and Curran, 1986].

The state of polarized light at a given wavelength can be described by the Stokes parameters [Stokes, 1851]:

$$\boldsymbol{I} = \begin{bmatrix} I \\ Q \\ U \\ V \end{bmatrix} = \begin{bmatrix} I_{0^\circ} + I_{90^\circ} \\ I_{0^\circ} - I_{90^\circ} \\ I_{45^\circ} + I_{135^\circ} \\ I_+ + I_- \end{bmatrix} \tag{1.19}$$

where I_α denotes the relative intensity at orientation α with respect to light traveling direction; + and − are the right and left-handed circular polarization components; I is the total intensity of the incident light beam; Q and U are the magnitude and orientation of the linear polarization; V describes the circularity of the polarization ellipse. The state of the scattered light can be written as the a linear



Figuur 1.7: Polarization types.

transformation of the incident light:

$$\begin{bmatrix} I_s \\ Q_s \\ U_s \\ V_s \end{bmatrix} \begin{bmatrix} F_{11}(\Theta) & F_{12}(\Theta) & 0 & 0 \\ F_{21}(\Theta) & F_{22}(\Theta) & 0 & 0 \\ 0 & 0 & F_{33}(\Theta) & F_{34}(\Theta) \\ 0 & 0 & F_{43}(\Theta) & F_{44}(\Theta) \end{bmatrix} \begin{bmatrix} I_0 \\ Q_0 \\ U_0 \\ V_0 \end{bmatrix} \tag{1.20}$$

where the matrix $\boldsymbol{F}$ consisting of elements $F_{ij}(\Theta)$, is the scattering matrix; Θ is the scattering angle; the subscript s and 0 indicate the scattered and incident light. For radiative transfer calculations in the atmosphere, $F_{11}(\Theta)$ is the scattering phase function as mentioned previously (Equation 1.3); $F_{12}(\Theta) = F_{21}(\Theta)$ and $F_{34}(\Theta) = -F_{43}(\Theta)$. The ratio $-F_{12}(\Theta)/F_{11}(\Theta)$ represents the degree of linear polarization if the incoming light is unpolarized [Boesche et al., 2006]. Polarization is dominated by small particles (radii$\leq 0.5\mu m$), thus a high degree of polarization implies that small particles are dominant [Tanré et al., 2011]. The circular polarization is negligible compared with linear counterparts [Hasekamp and Landgraf, 2007, Talmage and Curran, 1986, Waquet et al., 2009], as a result, $F_{34}(\Theta)$, $F_{43}(\Theta)$ and $F_{44}(\Theta)$ can be ignored.

The scattering matrix F highly depends on the aerosol micro-physical properties [Hansen and Travis, 1974]. Compared with intensity (I) that measured by many traditional sensors, the linear polarized terms (Q and U) are more sensitive to the particle size, shape and refractive index, allowing retrieval of detailed aerosol physical information in addition to AOD [Boesche et al., 2006, Chen et al., 2020a, Hasekamp and Landgraf, 2007, Mishchenko and Travis, 1997].

The passive sensors equipped with polaroids can measure the polarized radiance, from which the Stokes parameters (excluding the circular components) can be derived (Equation 1.20). The dominant molecular contribution to the polarized radiance can be parameterized and removed before the aerosol properties inversion [Herman et al., 2005]. Lidar can actively transmit linearly (nearly 100%) polarized beams and measure the degree of the linear polarization of the returned signal. The backscatter coefficients are separated by a beam-spilter into a parallel and a perpendicular component with respect to the plane of the outgoing pulse. The ratio between the two components (perpendicular to parallel) gives the volume linear depolarization ratio from which the particle shape-related aerosol depolarization ratio can be calculated [Nishizawa et al., 2007, Winker et al., 2007]. A large depolarization ratio usually indicates non-spherical particles.

Polarization can provide additional constraints on aerosol retrievals, resulting from lower sensitivity to the surface reflectance and higher measurement accuracy. The contribution of the surface to the total polarized light is uniform and constant, and less important compared to the atmospheric contribution [Nadal and Breon, 1999, Tanré et al., 2011]. Moreover, polarimetric sensors also benefit from instrumental design and calibration advantages. The linear polarization is just a ratio of two orthogonal intensity components (Equation 1.20) [Mishchenko and Travis, 1997]. A well-designed polarimetric instrument can reach a high accuracy of 0.2% on an operational basis [Chowdhary et al., 2002]. Polarimetric remote sensing is one of the most promising techniques to characterize the aerosol properties and has increasing applications in the modern remote sensing era [Dubovik et al., 2019].

1.5. OVERVIEW OF AEROSOL SATELLITE REMOTE SENSING

Satellite measurements of aerosol properties can be categorized into three major aspects: (1) columnar aerosol optical depth (AOD); (2) other columnar aerosol optical properties except for AOD; and (3) aerosol vertical distribution. This section will show the major missions that fall into the above three categories.

Kaufman et al. [1997a] summarized the fundamental aerosol remote sensing techniques, which are: (1) a single measurement, (2) multi-wavelength measurement, (3) multi-angle measurement, (4) polarization, (5) active lidar, (6) limb scanning and (7) occultation. Due to the ill-posed inverse problem, there is a hierarchy between the importance of various parameters in the observed reflectance. (1-2) techniques are commonly used to retrieve AOD; but now it is also possible to retrieve the aerosol

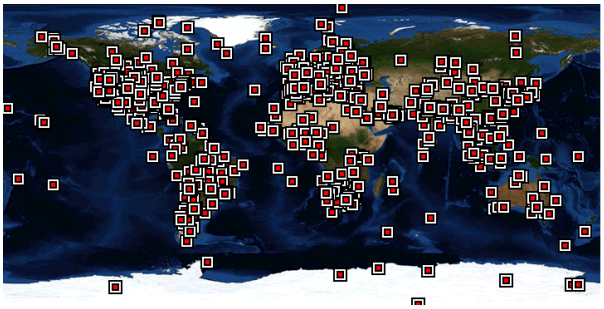
layer height (ALH) using some particular channels. (3-4) allow to retrieve additional information on particle size and refractive index. (5-7) are used to measure the aerosol vertical profiles. The last two are limited to stratospheric aerosols as the extinction along the light path is very large in the troposphere, which may lead to high uncertainties [Kaufman et al., 1997a]. Combinations of above basic techniques are also common in the modern remote sensing era (after 21st century), which will be introduced in following subsections.

1.5.1. MEASUREMENT OF AEROSOL OPTICAL DEPTH

The columnar AOD is the primary parameter derived from satellite observations, with assumptions made on other aerosol properties [Kaufman et al., 1997a]. Mono-spectral measurements can only retrieve AOD at a certain wavelength without guaranteeing high retrieval accuracy [Mishchenko and Travis, 1997], while multi-spectral measurements cannot only retrieve the wavelength-dependent AOD with improved accuracy, but also derive ÅE that contains the information on particle size.

The first global view of AOD was at a single wavelength (0.63 μm) over ocean provided by AVHRR/TIROS-N (launched in 1978) [Husar et al., 1997]. Later AVHRR instruments applied multi-channel measurements and retrieved additional information on particle size [Geogdzhayev et al., 2002, 2005]. Efforts were also made on AOD retrieval over land [Hauser et al., 2005a,b, Hsu et al., 2017, Li et al., 2013]. TOMS, another early sensor (the first one was launched on the Nimbus-7 in 1978), retrieved AOD over both land and ocean but is limited to the ultra-violet band [Torres et al., 2002a]. Modern satellite instruments, such as MODIS/Terra and Aqua [Levy et al., 2013, Remer et al., 2006, 2008], the family of ATSRs (ATSR/ERS-1, ATSR-2/ERS-2 and AATSR/ENVISAT) [Flowerdew and Haigh, 1995, Llewellyn-Jones et al., 2001, Veefkind et al., 1998], VIIRS/Suomi-NPP and NOAA-20 [Jackson et al., 2013], SeaWiFS/OrbView-2 [Hooker and McClain, 2000], MERIS/ENVISAT [Bezy et al., 2000], OMI/Aura [Torres et al., 2007], MISR/Terra [Kahn et al., 2009], POLDER/ADEOS I, ADEOS II and PARASOL [Deuzé et al., 2000, Dubovik et al., 2011, Herman et al., 1997b], ABI/GOES-R series [Schmit et al., 2017], CALIOP/CALIOPSO [Winker et al., 2009], SLSTR/Sentinel-3 [P.Coppo et al., 2010], etc., make great contributions to map the global wavelength-dependent AOD climatology.

An important source providing global AOD data is the AERONET direct sun product (ground-based) [Holben et al., 1998]. AERONET is a sun photometer network initialized by NASA and is greatly expanded by collaborators from national agencies, institutes, universities, individual scientists, and partners. Although it is a global network, the distribution of AERONET sites is uneven (Figure 1.8). Most sites measure aerosols over land, whereas the AOD over ocean is mainly provided by the Maritime Aerosol Network (MAN) component of AERONET by the ship-borne Microtops II sun photometers (since 2004, [Smirnov et al., 2009]). AERONET has provided a long-term (since 1993), continuous and readily accessible database of aerosol optical and radiative properties for aerosol research and characterization, validation of satellite retrievals and model simulations, and synergism with other databases.



Figuur 1.8: AERONET site distribution (will be changed later). Most sites are located in North America and West Europe, while there are less sites in the rest of the world.

The AERONET measures either direct solar irradiance or sky radiance (will be introduced in Section 1.5.2). The direct sun measurements are made in 8 spectral bands (340, 380, 440, 500, 670, 870, 940 and 1020 nm). The AOD is retrieved by the Beer-Bouguer Law, where attenuation due to Rayleigh scatter, and absorption by ozone, and gaseous pollutants are estimated and removed [Kinne, 1999]. The estimated near-real time AOD has a bias up to 0.02 with a standard deviation uncertainty of 0.02 [Giles et al., 2019].

1.5.2. MEASUREMENT OF MICRO-PHYSICS

The wavelength-dependent radiance measurements from a fixed viewing geometry cannot ensure reliable retrievals of aerosol micro-physical properties, as the absolute surface reflectance or its spectral variation at different wavelengths are usually unknown thus have to be pre-described [Chu et al., 2002, Hsu et al., 2006, Kaufman et al., 2002a]. Multi-angular measurements can put more strict constraints on aerosol characteristics and remove much ambiguity in decoupling the aerosol and the surface radiance [Kaufman et al., 1997a]: by viewing the same ground pixel from different angles within a relatively short period, the atmospheric conditions can be assumed to be constant. Multi-angle technique also provides information on the aerosol scattering phase function [Kahn and Gaitley, 2015, Kalashnikova et al., 2005, Kalashnikova and Kahn, 2006]. A representative multi-spectral and multi-angular instrument is MISR/Terra [Diner et al., 1991]. MISR is equipped with multiple cameras with one pointing to the nadir and others at forward and backward viewing angle of 26.1°, 45.6°, 60.0° and 70.5°) [Diner et al., 1998].

The aerosol size and type can be retrieved with higher accuracy by the multi-angle technique compared with single viewing measurements. However, as the multi-spectral measurements, multi-angular technique only measures the intensity of the radiance (i.e. the Stoke parameter I). As a result, it still leads to ambiguous solutions for the aerosol refractive index and effective radius [Chowdhary et al., 2002, Mishchenko and Travis, 1997]. To overcome this, in addition measurements of the polarization is necessary. The merit of polarization is the high measurement accuracy and the high sensitivity to particle size, shape and refractive index as a function of scattering angle and wavelength [Chowdhary et al., 2002, Mishchenko and Travis, 1997].

Experience with polarimetry remote sensing on planetary missions has demonstrated that the measurement of polarization and intensity together can resolve the non-uniqueness problem [Mishchenko and Travis, 1997]. Currently, there are several sensors applying polarimetry: POLDER/ADEOS I, ADEOS II and PARASOL [Deschamps et al., 1994, Goloub et al., 1999], APS/Glory (failed) [Mishchenko et al., 2007, Peralta et al., 2007], GOME/ERS-2, GOME-2/Metop-A, Metop-B and Metop-C [Munro et al., 2016, Veefkind et al., 2000], SCIAMACHY/ENVISAT [Burrows et al., 2011], CALIOP/CALIPSO [Winker et al., 2009], etc.

Future sensors will focus on combined uses of multi-spectrum, multi-angle and polarization, for example, MAIA developed by JPL that will be launched in 2022. It applies radiometric and polarimetric measurements to characterize the sizes, compositions and quantities of particulate matter in air pollution [Diner et al., 2018, Liu and Diner, 2017]; SPEXone Polarimeter/PACE that is scheduled for launch in 2023. It has 5 viewing angles (0°, ±20°, ±50°) and an improved the polarimetric accuracy (0.3%) [van Amerongen et al., 2019]; 3MI/EPS-SG is planned to operate in the time-frame 2020–2040 [Fougnie et al., 2018, Marbach et al., 2015], etc. For more planned missions can refer to Dubovik et al. [2019].

Another technique used in space-borne instruments to determine the aerosol absorption is the ultra-violet measurements [Moosmüller et al., 2009]. Due to the presence of UV light absorbing aerosols, it is possible to derive a qualitative parameter termed ultra-violet aerosol index (UVAI) or absorbing aerosol index (AAI). Sensors with at least two UV channels are capable to retrieve it. Note that UVAI is not a physical parameter like refractive index and effective radius. More detailed information will be introduced in Section 1.7.

The ground-based AERONET instruments also make use of the multi-spectral, multi-angular and polarization technique. They measure the sky radiance in four spectral bands (440, 670, 870 and 1020

nm) along the solar principal plane (at constant azimuth angle, with varied scattering angles) up to nine times a day, and along the solar almucantar (at constant elevation angle, with varied azimuth angles) up to six times a day. The sky radiance measurements observe the same aerosol profile through multiple scattering angles to retrieve aerosol absorption (SSA and AAOD), and other aerosol microphysical parameters (particle volume size distribution, complex refractive index, and scattering phase function) [Dubovik and King, 2000]. However, the potential value of polarization is not fully employed by AERONET [Xu and Wang, 2015]. The latest AERONET version 3 product is generated by a new polarized radiative transfer code (SORD) that provides the potential to add polarization measurements in inversion, but currently there are not many AERONET sites acquiring polarization measurements due to the long time interval of measurements. Additional research is required to determine the optimal number of the observational angles and spectral channels needed to reduce the acquisition time of these measurements [Sinyuk et al., 2020]. As a result, despite studies that have proven that polarization data can improve retrievals [Dubovik et al., 2006, Fedarenka et al., 2016, Li et al., 2009], the current version 3 algorithms still employs the same inversion as version 2 which is described in [Dubovik and King, 2000] and [Dubovik et al., 2006].

1.5.3. MEASUREMENT OF AEROSOL VERTICAL DISTRIBUTION

The vertical distribution of aerosols is strongly associated with aerosol radiative forcing assessments [Boucher et al., 2013, Mishra et al., 2015]. Moreover, (for passive sensors) it is a unknown parameter and has to be assumed, triggering uncertainties in retrieving other aerosol optical properties [Torres et al., 1998, Wu et al., 2017]. Wu et al. [2016c] demonstrated that for elevated smoke or dust layers (e.g., altitude = 7 km), the MODIS C6_DT algorithm can negatively bias the AOD retrievals by over 10%.

The measurements of aerosol vertical distribution mainly rely on the active instruments, i.e. lidar. There are many operational missions, such as the ground-based EARLINET [Pappalardo et al., 2014], MPLNET [Welton et al., 2001], the space-borne CALIOP/CALIPSO [Winker et al., 2009], GLAS/ICESat [Schutz et al., 2005], and the recently launched ADM-Aeolus [Flamant et al., 2008]. Although lidar can provide great details in the vertical direction, the measured profiles are subject to limited spatial and temporal coverage. Besides, the presence of clouds or optically dense aerosol layers may significantly attenuate lidar signals, resulting in large uncertainties or missing data in the measured profiles.

Nowadays, many algorithms have been developed to retrieve the columnar aerosol layer heights (ALH), a compact representation of aerosol profiles. Multiple angular measurements can determine elevated aerosol plume height by stereo photogrammetry, e.g. MISR/Terra [Nelson et al., 2013] and AATSR/ENVISAT [Virtanen et al., 2014]. ALH derived by spectrum fitting over oxygen (O_2) absorption spectroscopy is more commonly applied. Representative instruments include POLDER/PARASOL, MERIS/ENVISAT [Dubuisson et al., 2009, Duforêt et al., 2007], SCIAMACHY/ENVISAT [Sanghavi et al., 2012], GOME-2/ Metop-A, Metop-B and Metop-C [Sanders et al., 2015], TROPOMI/Sentinel-5P [Sanders and De Haan, 2016], and EPIC/DSCOVR [Xu et al., 2017]. Studies also have proposed improved solutions by adding polarization measurements [Boesche et al., 2009, Ding et al., 2016, Wang et al., 2014]. Chimot et al. [2017] also attempted to retrieve ALH from the $O_2 - O_2$ absorption band (477 *nm*) of OMI/Aura. Furthermore, the absorbing aerosol layer height (AAH) has been developed as an official product of the GOME-2 instruments [Tilstra and Stammes, 2019]. Instruments equipped with thermal infrared (thermal IR) band, such as AIRS/Aqua and IASI/Metop-A, can retrieve height information for dust because the thermal IR band is highly sensitive to dust aerosols [Pierangelo et al., 2004, Vandenbussche et al., 2013]. For a detailed review of ALH retrievals from observations, one can refer to Xu et al. [2018b]. Although extending the spatial coverage of lidar measurements, ALH retrievals from passive sensors are only applicable under certain conditions, e.g. elevated aerosol layers, over dark surfaces, in the absence of clouds, etc. Moreover, different retrieval algorithms may have different assumptions on aerosol optical properties, the aerosol profile shape and the ALH definition in the forward simulations, making it difficult to compare different ALH products. As a result, the validation of the retrieved ALH still relies on the extinction profiles retrieved from lidar measurements.

1

1.6. ADVANCES IN AEROSOL REMOTE SENSING RETRIEVALS

As introduced in Section 1.4, retrieving physical parameters from remote sensing measurements is an under-determined problem, i.e. the number of observations is less than the number of unknowns. A unique and stable solution requires additional constraints. Conventionally, a measured signal is calculated by a forward model based on physical processes (**forward simulation**). Then, by comparing the real measurements and the calculated ones, the target parameter is numerically retrieved using the least square methods or optimal estimation (**numerical inversion**). Nowadays, with the increasing amount of the geoscience data and the improved computational capability, many data-driven retrieval methods have been developed. By capturing the behaviors and characteristics from data and determining their underlying probability distributions, these methods can provide new knowledge regarding the variable of interest. The following subsections introduce the above two types of retrievals in detail.

1.6.1. PHYSICALLY-BASED RETRIEVALS

At the forward simulation phase of physically-based retrievals, a forward model is built based on physical processes (e.g. radiative transfer) with a priori assumptions on other physical parameters (except for the target) in order to constrain the under-determined system.

As the forward simulation could be computationally intensive, many operational algorithms use Look-Up Table (LUT) approach. A LUT is calculated beforehand for a set of nodal points of measurement geometries and physical parameters, thus in operation it is only needed to apply the numerical inversion. LUT has been the dominant method to retrieval aerosol properties for many years due to its high operational processing speed [Dubovik et al., 2011, Kokhanovsky et al., 2010]. It is widely employed in many mainstream satellite products, such as MODIS/Terra and Aqua [Remer et al., 2006], OMI/Aura [Torres et al., 1998], MISR/Terra [Kahn et al., 2007], POLDER/ADEOS I and ADEOS II [Lenoble et al., 2007] and CALIOP/CALIPSO [Winker et al., 2009].

But LUT approach in observations with multi-angular and/or polarization measurements reveals deficiencies. These observations have notably higher sensitivity to the details of aerosol and surface properties, and the retrieval of larger number of parameters is expected. This requires a comprehensive LUT with a larger dimension that is less suitable for operational use [Dubovik et al., 2011, Kokhanovsky et al., 2010]. One solution is creating a LUT relying only on the selected sub-set of the observations with highest sensitivity to the aerosol parameters and retrieve reduced set of characteristics, e.g. the POLDER operational retrieval over ocean [Herman et al., 2005] and land [Deuzé et al., 2001]. Another method is online radiative transfer simulations, owing to the improved computational capability. In this case, forward simulations are conducted during the retrieval process rather than computed in advance. Applications of online radiative transfer simulations require fast and accurate radiative transfer codes. This method has been either tentatively or operationally implemented in MERIS/ENVISAT [Katsev et al., 2009], AATSR/ENVISAT [Grey et al., 2006, Grey et al., 2006] and POLDER/PARASOL [Dubovik et al., 2014, Lenoble et al., 2007] retrievals. Such an inversion strategy allows for the use of a continuous space of solutions instead of a limited set of predetermined solutions as used in LUT-based algorithms [Dubovik et al., 2011].

1.6.2. STATISTICALLY-BASED RETRIEVALS

BIG DATA ERA OF GEOSCIENCE

Geoscience has witnessed the evolution from a data-poor field to a data-rich field, owing to the state-of-the-art measuring techniques (e.g.satellite remote sensing), improvements of computational capability for global-scale Earth system models, and the Internet-based democratization of data making most geoscience data publicly available [Karpatne et al., 2019b]. The storage volumes have been beyond the dozen of petabytes and the transmission rates exceed hundreds of terabytes per day [Reichstein et al., 2019]. Undoubtedly, like other scientific fields with big data, geoscience is welcoming the era of data-driven research and the **Machine Learning** comes into the eyes of geoscientists.

Machine Learning was first coined by Aurther Samuel in 1959 [Samuel, 1959] originally for the

quest of Artificial Intelligence (AI). A more widely quoted Machine Learning definition was proposed by Mitchell [Mitchell, 1997]: *A computer program is said to learn from experience E with respect to some class of tasks T and performance measure P if its performance at tasks in T, as measured by P, improves with experience E.*

Machine Learning is deeply rooted in statistics and probability theories. It extracts information and knowledge from data deluge by building numerical models based on behaviours and characteristics of empirical data without explicit rules [Reichstein et al., 2019]. Machine Learning has three major types (supervised learning, unsupervised learning and reinforcement learning), among which only the supervised learning requires the training data consisting of input features and corresponding output variables, i.e. supervisions. Depending on whether the output parameter is categorical or quantitative, supervised learning can be further grouped into classification and regression problems.

Lary et al. [2016a] states Machine Learning is ideal to address problems where our theoretical knowledge is incomplete but for which there exists a great number of data. Retrieving physical parameters from electromagnetic waves measured by satellite remote sensing is a typical problem of such kind. Another benefit using Machine Learning is its high computational efficiency. Although a preliminary training phase can be time-consuming for supervised learning, their applications are essentially instantaneous once the training phase is complete [Di Noia and Hasekamp, 2018]. Early (1990s) applications of Machine Learning in geoscience has succeeded in land and cloud classification (e.g.[Benediktsson et al., 1990, Lee et al., 1990]). It has further thrived in event detection (e.g. [Kim et al., 2019, Perol et al., 2018]), variable estimation (e.g. [Lary et al., 2014, Zhan et al., 2017]), data assimilation (e.g. [Gilbert et al., 2010]), bias correction (e.g. [Just et al., 2018, Lary et al., 2009]), interpolation (e.g. Li et al. [2011], Requia et al. [2019]), forecasting (e.g. [Rasouli et al., 2012, Sharma et al., 2011]), etc.

NEW POSSIBILITIES OF AEROSOL REMOTE SENSING RETRIEVALS

In the field of satellite remote sensing of aerosols, Machine Learning primarily carries out the following tasks [Di Noia and Hasekamp, 2018, Dramsch, 2020, Lary et al., 2016a]:

- An approximation of forward model (radiative transfer simulations);
- An surrogate of numerical inversion;
- Object detection and classification.

Applications of Machine Learning to aerosol remote sensing bring benefits of: (1) it is more computationally efficient than the forward simulation and/or optimal estimation; and (2) it can build numerical models based on behaviours and characteristics of empirical data without explicitly describe the physical processes. In other word, it is possible to retrieve aerosol optical properties without making assumptions on aerosol micro-physics (particle size, shape and composition) that are usually unknown.

There are many applications of Machine Learning thrive in aerosol remote sensing, including approximating radiative transfer simulations (e.g. [Berdnik and Gallyamova, 2012, Di Noia and Hasekamp, 2018, Krasnopolsky, 2007]), retrieving aerosol properties (micro-physics: e.g. [Berdnik and Loiko, 2009, 2016, Berdnik et al., 2004, Dong et al., 2020, Ishimaru et al., 1990, Taylor et al., 2014, Ulanowski et al., 1998, Wang et al., 1999]; aerosol optical depth: e.g. [Han et al., 2006, Radosavljevic et al., 2010, Ristovski et al., 2012, Taylor et al., 2014, Vucetic et al., 2008]; aerosol vertical distribution: e.g. [Chimot et al., 2017, Dong et al., 2020, Nanda et al., 2019, Yao et al., 2018]), classifying aerosol types (e.g. [Niang et al., 2006, 2003]), correcting/validating retrievals ([Lary et al., 2009, Levy et al., 2007]), detecting volcanic ash/wildfire plume (e.g. [Donida Labati et al., 2013, Gray and Bennartz, 2015, Larsen et al., 2020, Picchiani et al., 2011, Piscini et al., 2014]), etc.

CONCERNS ON MACHINE LEARNING APPLICATIONS IN GEOSCIENCE

Despite of the growing trend in geoscience, applying Machine Learning also encounters several challenges that should be paid attention to [Karpatne et al., 2019b].

The most challenging thing is that geoscience is governed by physical laws and principles, while Machine Learning can only build numerical relations between variables. This may lead to the predicted results that cannot be explained by physical processes. Thus, feature selection and engineering where geoscientists can bring their domain knowledge into the Machine Learning models, are particularly important in order to improve the physical interpretability [Guyon and Elisseeff, 2003, Weston et al., 2001].

Another issue is the inherit properties of geo-data, e.g. spatio-temporal auto-correlation, heterogeneity in space and time, etc. Recent studies have paid attention to the spatial and temporal structure of geo-data when applying Machine Learning algorithms, either by directly involving spatial and temporal information in the feature space (e.g. [Li et al., 2017, 2016]), or combining geo-statistical tools with Machine Learning (e.g. [Ishitsuka et al., 2018, Li et al., 2014, Michalak and Shlomi, 2013, Singh et al., 2017, Tapoglou et al., 2014, Tarasov et al., 2018]). There is also an increasing trend using the Convolutional Neural Network (CNN) and the Recurrent Neural Network (RNN) to present the temporal and spatial correlation (e.g. [Biancofiore et al., 2015, Chen et al., 2020b, Wen et al., 2019]), as these techniques can share weight parameters in space and time domain.

1.7. RESEARCH TOPIC

The research topic of my PhD project is **aerosol absorption in the UV band**. As introduced in Section 1.2, most aerosols can fully scatter sunlight, but there also exist aerosols transferring electromagnetic energy into thermal energy, termed absorbing aerosols. Absorbing aerosols are responsible for air quality degradation and visibility impairment. Moreover, the direct and semi-direct effect of absorbing aerosols are important mechanisms influencing the Earth's climate system. It is reported that the black carbon deposited on snow becomes the second largest positive radiative forcing terms after the greenhouse gases (Figure 1.4).

The single scattering albedo (SSA, Equation 3.2) is a common parameter describing the aerosol absorption. It is defined as the ratio between particle scattering and absorbing efficiency. Scattering aerosols usually have a SSA near unity ($\omega_0 \approx 1$) from the UV ($0.1\ \mu m$) to the short-wavelength infrared ($3\mu m$) band. On the other hand, SSA of absorbing aerosols varies significantly, depending on particle chemical compositions as well as shape, size and mixing state [Bergstrom et al., 2007]. Three major absorbing aerosol types are black carbon, brown carbon and mineral dust (see Section 3.1) [Moosmüller et al., 2009, Samset et al., 2018].

There are three types of techniques measuring aerosol light absorptive properties (represented by SSA, AAOD or refractive index): (1) filter methods, (2) in situ instruments and (3) remote sensing. The former two direct contact with the measuring targets, either collecting particles or measuring the particles in their natural suspended state (in situ instruments). Remote sensing retrieves aerosol properties without touching the targets, thus they are potential to provide aerosol absorption on a global scale. Although many advanced techniques have been applied to aerosol satellite remote sensing (multi-angle, polarimetry, Raman lidar, etc. See Section 1.5.2), so far there is no practical way to provide global, self-consistent measurements from a single source [Moosmüller et al., 2009, Samset et al., 2018]. AERONET routinely provides information on the aerosol micro-physics, but it is a unevenly-distributed-ground-based system that cannot produce a complete global view of aerosols. Satellite instruments such as OMI/Aura [Torres et al., 2002b], MISR/Terra [Diner et al., 1998, Kahn et al., 2009, Kahn et al., 2007], and POLDER/ADEOS I, ADEOS II and PARASOL [Deuzé et al., 2001, Dubovik et al., 2014, Herman et al., 2005, Tanré et al., 2011] provide aerosol SSA or AAOD products, however, the retrieved aerosol absorptive properties highly depend on the pre-assumed particle size, shape and complex refractive index that are usually in the vector of unknowns.

Another remote sensing technique to retrieve aerosol absorption is the Ultra-Violet (Absorbing) Aerosol Index (UVAI), also called the (Absorbing) Aerosol Index (AAI or AI) [Herman et al., 1997b]. Differing from SSA and AAOD, UVAI is a qualitative measure of aerosol absorption and it is calculated directly from radiance contrast without assumptions on aerosol micro-physics.

1.7.1. ULTRA-VIOLET ABSORBING AEROSOL INDEX

UV-absorbing aerosols were first paid attention to in research regarding the aerosol effects on ozone retrievals in UV band from space [Dave, 1978]. Absorbing aerosols reduce more radiance at shorter wavelength compared to a Rayleigh atmosphere, whereas scattering aerosols have an opposite effect. Based on this fact, the radiance difference between two near-UV wavelengths ($R_{340} - R_{380}$) was proposed to detect the biomass burning aerosols based on the released TOMS measurements in 1996 [Hsu et al., 1996]. Later, Herman et al. [1997a] devised the concept of UVAI which is well-known today: UVAI is the radiance difference between the logarithm of the ratio between a measured and a calculated radiance at two UV wavelengths λ_1 and λ_2 ($\lambda_1 < \lambda_2$) [Torres et al., 1998], defined as:

$$UVAI = -100\left[log_{10}\left(\frac{I_{\lambda_1}}{I_{\lambda_2}}\right)^{obs} - log_{10}\left(\frac{I_{\lambda_1}}{I_{\lambda_2}}\right)^{Ray}\right] \tag{1.21}$$

The radiance is calculated at a surface albedo A_{s,λ_2} found by requiring the calculated and the measured radiance at λ_2 are numerically identical ($I^{cal}_{\lambda_2}(A_{s,\lambda_2}) = I^{obs}_{\lambda_2}$). Then the above equation can be rewritten into:

$$UVAI = -100 log_{10}\frac{I^{obs}_{\lambda_1}}{I^{Ray}_{\lambda_1}(A_{s,\lambda_2})} \tag{1.22}$$

where the $I^{Ray}_{\lambda_1}$ is calculated by radiative transfer simulations assuming no aerosol present (i.e. Rayleigh atmosphere). A positive value of UVAI indicates the presence of absorbing aerosols, while non-absorbing aerosols and clouds yield near-zero or even negative values [Herman et al., 1997a]. UVAI is mainly determined by the interaction between an aerosol layer and the Rayleigh scattering beneath the layer.

In the calculation of the Rayleigh atmosphere radiance, the combined effect of molecules, clouds and surface are usually presented by a Lambert Equivalent Reflectivity (LER) model. Torres et al. [2018] indicated that such an approximation cannot capture the angular variability associated with the scattering phase function of clouds, which may cause the UVAI bias in across-scan direction of OMI. They thus proposed an improvement that separates the effects of surface and clouds. The surface contribution to the total radiance is calculated using wavelength dependent climatology of surface albedo, and that of clouds is calculated by the Mie Theory for an assumed water cloud model. The newly calculated UVAI shows a reduced across-scan bias, however, it is overall higher than the UVAI calculated by LER model, especially over oceans (shown in Chapter 5).

The first UVAI product was provided by TOMS/Nimbus-7 in 1978. Since then, many space-borne instruments have made contributions to the global UVAI record. Table 1.1 lists these sensors.

Compared with other aerosol optical parameters, UVAI can be calculated for all observations regardless of cloudiness conditions or over bright surfaces (e.g. snow and ice) [Torres et al., 2018], thus a global coverage is possible. Besides, the retrieval of UVAI does not require assumptions on aerosol micro-physics that is usually unknown. The main drawback of UVAI is that it is not a geophysical parameter like SSA that can quantitatively interpret aerosol absorption.

1.7.2. SENSITIVITIES OF UVAI

The most influential parameters to UVAI are aerosol loading, aerosol absorption and aerosol vertical distribution [Wang et al., 2012], but viewing geometries, clouds and surface reflectance also have (but less important) effects on UVAI calculation.

Sensitivity studies have investigated UVAI dependence on these parameters for either Rayleigh scattering or Mie Theory calculated aerosols based on radiative transfer simulations [Colarco et al., 2017, de Graaf et al., 2005, Herman et al., 1997a, Hsu et al., 1999, Penning de Vries et al., 2014, Sun et al., 2018, Torres et al., 1998]. Major conclusions are summarized here. Note that the following conclusions are not applicable to (nearly) purely scattering aerosols ($\omega_0 \approx 1$).

Tabel 1.1: Satellite missions with UVAI measurements.

Instrument	Platform	Period	$\lambda_1[nm]$	$\lambda_2[nm]$
TOMS	Nimbus-7	1978-1993	340	380
TOMS	METEOR-3	1991-1994	340	380
TOMS	ADEOS	1996-1997	340	380
TOMS	Earth Probe	1996-2005	331	360
GOME	ERS-2	1995-2003	335	380
SCIAMACHY	Envisat	2002-2012	340	380
OMI	Aura	2004-present	354	388
GOME-2	Metop-A	2007-present	340	380
GOME-2	Metop-B	2012-present	340	380
GOME-2	Metop-C	2019-present	340	380
OMPS	Suomi-NPP	2011-present	340	378.5
TROPOMI	Sentinel-5P	2017-present	340/354	380/388

AEROSOL ABSORPTION

Low SSA indicates strong absorption efficiency, thus UVAI decreases with SSA. This conclusion holds for both gray aerosols (spectrally-flat) and aerosols with a spectrally-dependent refractive index.

ASYMMETRY FACTOR (PARTICLE SIZE)

Most aerosols tend to scatter in the forward direction with the asymmetry factor g in the range from 0.4 to 0.85 [Dubovik et al., 2002b]. As g increases, more radiation is scattered forward to the surface rather than reflected backward to TOA. With less radiance measured by satellite, the UVAI becomes larger.

AEROSOL OPTICAL DEPTH

UVAI increases with aerosol loading. As AOD increases, more radiation is absorbed and less radiation returns back to TOA. It is found that the increase of UVAI with AOD is nearly linear with a slope proportional to SSA [de Graaf et al., 2005, Herman et al., 1997a, Hsu et al., 1999, Torres et al., 1998].

AEROSOL VERTICAL DISTRIBUTION

UVAI and aerosol layer height (relative to the topographic surface) are positively correlated and the dependence is nearly linear. The aerosol layer mainly interacts with radiation beneath it. The higher the aerosol layer, the more interaction occurs, yielding a higher UVAI. The altitude dependency increases with decreasing SSA and increasing AOD, while no significant dependence is found for non-absorbing aerosols.

This UVAI dependence on altitude is useful to detect elevated absorbing aerosols. But note that gray aerosols only show positive values above a certain altitude ($1-2km$), while non-gray aerosols can be detected from near surface [de Graaf et al., 2005].

SURFACE REFLECTANCE AND PRESSURE

Compared with the visible band, the UV channel is less sensitive to the brightness of the surface, but the surface effects on UVAI still exist. On the one hand, brighter surfaces directly increase reflected radiation at the TOA, which reduces UVAI. On the other hand, higher surface albedo reflects more radiation to the aerosol layer and thus facilitates the interaction between the aerosol layer and the reflected radiation, yielding a higher UVAI.

Surface pressure can simulate effects of topography. With a fixed aerosol layer height relative to the reference ellipsoid, decreasing the surface pressure is equivalent to raised topographic surface, which reduced the amount of air below the aerosol layer. The less shielding of multiple Rayleigh scattering by the aerosol layer leads to a smaller UVAI.

CLOUDS

The cloud effect on UVAI is similar to that of a bright surface when an aerosol layer overlies clouds. However, when clouds are overlying an aerosol layer, the impact on UVAI is determined by the optical properties of the clouds. If there is an optically thick cloud, it could intercept all the radiation coming from the upper atmosphere, yielding a near neutral UVAI. In this case, detection of the underlying aerosol layer becomes impossible.

GEOMETRIES

The scattering geometries effect on UVAI is complex as it depends on the relative locations of the sun and the instrument as well as the particle scattering phase function. Nevertheless, UVAI varies significantly when solar and viewing zenith angles are larger than 60 °, since the light path through the atmosphere becomes longer and less certain.

OTHER INFLUENCING FACTORS

There are many possible choices of the UVAI wavelength pair (λ_1 and λ_2) of UVAI (Table 1.1). Most sensors use (approximately) 340/380 nm, while others use 331/360 nm, 354/388 nm, or 335/380 nm. If the reference wavelength λ_2 is same but λ_1 at a shorter wavelength, the multiple Rayleigh scattering is more important. This is equivalent to adding more optical path through the absorbing aerosol layer, increasing the possibility for absorption, thus UVAI increases. The UVAI calculated with 335/380 nm pair is 10% higher than that with the 340/380 nm. Besides, the larger difference between the wavelength pair usually implies a larger contrast, i.e. larger UVAI [de Graaf et al., 2005].

The reduction of the TOA radiance due to ozone absorption has a positive effect on UVAI, but this can be easily corrected by linear fitting.

1.7.3. APPLICATIONS OF UVAI

UVAI is a qualitative measure that is widely used for absorbing aerosol detection (e.g. smoke: [Duncan et al., 2003a,b, Hsu et al., 1999, Israelevich et al., 2002, Spichtinger et al., 2001]; dust: [Alpert and Ganor, 2001, Chiapello et al., 1999, El-Askary et al., 2006, Ginoux and Torres, 2003, Moulin and Chiapello, 2004, Wang et al., 2008b]; volcanic ash: [Krotkov et al., 1999, Krueger et al., 2008, Seftor et al., 1997, Yuan et al., 2011]; mixed aerosols: [Duncan et al., 2003b]), aerosol classification (e.g. [Koukouli et al., 2006, Lee et al., 2007, Torres and Chen, 2013]), and CTM validation (e.g. [Buchard et al., 2015, 2017, Chiapello et al., 1999, Hammer et al., 2016]). Nevertheless, many studies also made efforts on quantitative uses of UVAI. For example, Hsu et al. [1999] presented a linear relationship between the TOMS UVAI and the AOD retrieved independently from the ground-based Sun-photometer over biomass burning regions and African dust regions. Ginoux and Torres [2003] built an empirical relation between UVAI with SSA, AOD, surface pressure and altitude of the plume for dust aerosols. The UVAI calculated by the empirical relation correctly reproduces the daily, seasonal and inter-annual variability of the TOMS observed UVAI over desert regions. Guan et al. [2010] found a linear relationship between UVAI and biomass burning plume height, which allows to determine the height of elevated plume and to validate that in CTMs.

Furthermore, UVAI is an important source holding information on long-term global aerosol absorption. In principle, it is possible to estimate SSA under a reasonable assumption on particle size, if UVAI, AOD and ALH are known [Jeong and Hsu, 2008]. For example, Hu et al. [2007] retrieved SSA from the TOMS UVAI based on radiative transfer simulations with constraints of the MODIS retrieved AOD and aerosol vertical profiles provided by CTM. Their retrieval uncertainty is 15% and shows good correlation with AERONET; Jeong and Hsu [2008] retrieved SSA together with ALH based on a pre-calculated LUT. Applications to biomass burning events in North American and South Asia show that the retrieved SSA is in good agreements with AERONET measurements.

1.8. SCOPE AND OUTLINE

Given the potential value of the long-term UVAI record and the promising results of previous studies, the core research question of this dissertation can be formulated as: **How to derive quantitative aerosol optical properties from satellite UVAI records?**

The research questions investigated in the following chapters are:

1. What are physical parameters that UVAI is most sensitive to?
2. What are the major challenges/uncertainties of deriving quantitative aerosol optical properties using traditional radiative transfer simulations?
3. How can we best obtain an aerosol vertical distribution data set for quantitative studies of UVAI?
4. How to apply Machine Learning techniques to derive quantitative aerosol optical properties from UVAI?

Correspondingly, this dissertation is organized in 6 chapters to answer the above questions. Chapter 2 presents a case study of biomass burning event whose SSA is retrieved in combination with ALH from the OMI UVAI, based on radiative transfer simulations. The assumptions made on aerosol micro-physics and the lack of aerosol vertical distribution information are the major uncertainties of this method (related to Question 2). The UVAI dependency on many physical parameters is also studied [Sun et al., 2018] (related to Question 1).

Chapter 3 is a transitional study where both the conventional radiative transfer simulations and Machine Learning algorithms are applied to retrieve aerosol absorption from UVAI. It further investigates the limitations of the physically-based methods and demonstrates the feasibility of data-driven method by comparing of SSA retrieved by two methods for a common case. Although Machine Learning does not have to make assumptions on aerosol micro-physics, it also faces the problem of the lack of aerosol vertical distribution data [Sun et al., 2019] (related to Question 2 and 4).

Chapter 4 attempts to solve one of the major uncertainties raised in Chapter 3: how to build a global database of the aerosol vertical distribution in order to reduce the uncertainties in aerosol absorption prediction. Such a database of ALH is derived from MERRA-2 aerosol reanalysis based on the investigation of the relation between UVAI and various definitions of ALH. The relation between UVAI and ALH is also validated by various satellite ALH products [Sun et al., 2020] (related to Question 3).

Chapter 5 gathers conclusions of previous chapters and creates an aerosol absorption database derived from the OMI UVAI between 2006 and 2019. Here, a Deep Neural Network is applied to build a numerical relation with UVAI and AAOD. The result is in good agreement with AERONET and might be useful for better understanding of the aerosol absorption distribution and validation of other satellite retrievals and CTMs (related to Question 4).

In the end, Chapter 6 summarizes the experience of the quantitative derivation of aerosol absorption from UVAI and provides suggestions and recommendations for future research.

2

QUANTIFYING THE SINGLE SCATTERING ALBEDO FROM THE OMI ABSORBING AEROSOL INDEX BY RADIATIVE TRANSFER SIMULATIONS

The content of this chapter has been published as: **Sun, J.**, Veefkind, J. P., van Velthoven, P., and Levelt, P. F.: Quantifying the single-scattering albedo for the January 2017 Chile wildfiress from simulations of the OMI absorbing aerosol index, Atmos. Meas. Tech., 11, 5261–5277, https://doi.org/10.5194/amt-11-5261-2018, 2018.

2.1. INTRODUCTION

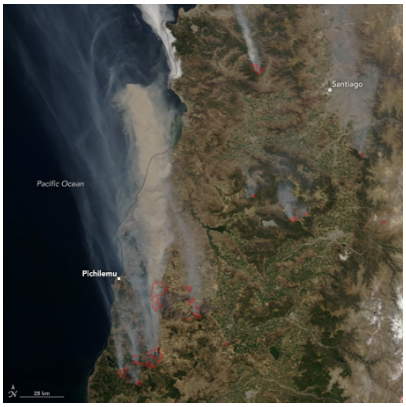
Biomass burning aerosols are generated from combustion of carbon-containing fuels by natural and anthropogenic processes [Bond et al., 2004, Pachauri et al., 2014]. They have great impacts on the Earth's climate [Kaufman et al., 2002b, Koch and Del Genio, 2010, Myhre et al., 2013, Pachauri et al., 2014, Solomon, 2007]. For example, the radiative forcing of black carbon (BC), a major component of biomass burning aerosols, is around 0.4 Wm^{-2} (0.05 – 0.80 Wm^{-2}). However, this value is highly uncertain [Bond et al., 2013, Myhre et al., 2013, Ramanathan and Carmichael, 2008]. Accurate measurements of the global aerosol single scattering albedo (SSA, symbol: ω_0) are therefore important to reduce the uncertainties in aerosol radiative forcing assessments [Hu et al., 2007]. SSA is defined as the ratio of aerosol scattering over aerosol extinction (Equation 3.2 on page 47). Routine operational SSA products are mainly contributed by ground-based instruments (e.g. [Corr et al., 2009, Dubovik et al., 1998, Eck et al., 2003, Kassianov et al., 2005, Petters et al., 2003, Yin et al., 2015]). Space-borne instruments usually derive SSA together with the aerosol optical depth (AOD, symbol: τ) retrieval based on a priori assumptions, e.g. the near-UV aerosol product of OMI/Aura [Torres et al., 2005, 2007]. Advanced instruments also apply multi-angular (e.g. MISR/Terra [Diner et al., 1998, Diner et al., 1991]) or polarimetric (e.g. POLDER/ADEOS-1/ADEOS-2/PARASOL [Deschamps et al., 1994, Goloub et al., 1999]) measurements to add additional constraints on the SSA retrieval.

Another parameter describing aerosol absorption is the near Ultra-Violet absorbing aerosol index (AAI) [Herman et al., 1997a], providing an alternative solution to retrieve SSA. AAI as qualitative measure of absorbing aerosols was first retrieved by TOMS/Nimbus-7 in 1978. Since then, various instruments have contributed to the AAI data record.

Compared with SSA retrieved from multi-angular and polarimetric measurements, AAI is 'cheaper' to obtain since it only requires measured UV-radiance at two wavelengths and a radiatively simulated Rayleigh atmosphere. The calculation of AAI is independent of assumptions on aerosol optical properties, which significantly reduces retrieval uncertainties. Ginoux et al. [2004] suggested that using AAI to compare model simulation and satellite retrievals to retrieve SSA allows a better control of discrepancies than using AOD, because the only error source is from models. The near-UV channel also brings advantages of the low surface reflectivity and the absence of significant molecular absorption. Moreover, the sensitivity of AOD to SSA in the visible band is lower over dark surfaces [Kaufman et al., 1997a], whereas AAI is by definition highly sensitive to aerosol absorption. In fact, it is possible to estimate SSA if AAI, AOD and the aerosol vertical distribution are known under a reasonable assumption on particle size [Jeong and Hsu, 2008]. Previous studies have proven the potential of AAI in retrieving aerosol properties. For example, Torres et al. [1998] applied an inversion method to derive AOD and SSA from the backscattered radiation. This method was validated by ground-based observations during the SAFARI 2000 measurement campaign. The agreement between their retrievals and AERONET measurements for AOD and SSA reached ±30% and ±0.03, respectively [Torres et al., 2005]. Hu et al. [2007] retrieved the global columnar SSA based on the TOMS AAI with an average uncertainty of 15%. Jeong and Hsu [2008] retrieved SSA together with the aerosol layer height (ALH, symbol: z_{aer}) for biomass burning aerosols, and the results showed a good agreement with both AERONET and CALIOP measurements.

Inspired by the above research, in this study, we use the AAI provided by OMI/Aura, the successor of TOMS, to derive the aerosol optical properties for the biomass burning aerosols generated by the wildfires in central Chile in January 2017. The fires were triggered by a combination of long-term drought and high temperature, and regarded as the worst wildfire season in the national history, leading to massive losses of the local forestry industry (pine and eucalyptus forests) [Guardian]. As shown in Figure 2.1, the smoke plume was transported away from the continent towards the tropical area of the Pacific Ocean by the north-westward winds.

In this study, we quantitatively retrieve the SSA together with ALH for this fire event by simulating the OMI retrieved AAI with a radiative transfer model (DISAMAR). In Section 2.2, we provide an introduction on AAI and its sensitivity study. The methodology and data sources are described in



Figuur 2.1: Wildfires in central Chile (Pichilemu 34.39°S, 72.00°W and Consititución 35.33°S, 72.42°W) detected by MODIS/Terra on 20 January 2017. (Source: NASA's Earth Observatory https://earthobservatory.nasa.gov/IOTD/view.php?id=89496).

Section 2.3. Section 2.4 presents the retrieved results with uncertainty analysis, followed by main conclusions in Section 2.5.

2.2. AAI SENSITIVITY STUDIES BASED ON DISAMAR

2.2.1. ULTRA-VIOLET ABSORBING AEROSOL INDEX

AAI was derived from the spectral contrast of satellite measurements, known as the residue method [Herman et al., 1997a]. The basic idea is that in a pure Rayleigh atmosphere, the radiance (I_λ)) decreases strongly with the wavelength in the UV channel. The presence of light-absorbing aerosols can reduce this spectral dependency. The change in the wavelength contrast between λ_1 and λ_2 ($\lambda_1 < \lambda_2$) is presented as:

$$AAI = -100\left[log_{10}\left(\frac{I_{\lambda_1}}{I_{\lambda_2}}\right)^{obs} - log_{10}\left(\frac{I_{\lambda_1}}{I_{\lambda_2}}\right)^{Ray}\right] \tag{2.1}$$

where obs and Ray denote the retrieved radiance and the calculated radiance, respectively. The longer wavelength λ_2 is treated as the reference wavelength where the surface albedo (a_{s,λ_2}) is determined by fitting the following relationship: $I^{Ray}_{\lambda_2}(a_{s,\lambda_2}) = I^{obs}_{\lambda_2}$. a_s is assumed to be spectrally flat, so that $I^{Ray}_{\lambda_1}$ can be calculated using the same a_s as $I^{Ray}_{\lambda_2}$. Then, the above equation can be reformed as:

$$AAI = -100log_{10}\frac{I^{obs}_{\lambda_1}}{I^{Ray}_{\lambda_1}(a_{s,\lambda_2})} \tag{2.2}$$

By defining $\Delta I_{\lambda_1} = I^{Ray}_{\lambda_1} - I^{obs}_{\lambda_1}$, the above equation can also be written as:

$$AAI = 100log_{10}(\frac{\Delta I_{\lambda_1}}{I^{obs}_{\lambda_1}} + 1) \tag{2.3}$$

The following sensitivity study will be based on Equation 2.3.

2.2.2. AAI SENSITIVITY STUDY

In this part, we present a sensitivity study of AAI based on a radiative transfer model developed by KNMI: DISAMAR. DISAMAR can perform simulations of the forward radiance spectrum in a wide spectral coverage (from 0.27 μm to 2.4 μm) and model scattering and absorption by gases, aerosol, clouds, and the reflection by the surface [de Haan, 2011]. It uses either the Doubling-Adding method or the Layer Based Orders of Scattering (LABOS) for the radiative transfer calculations. In this study the latter is employed as it is less computationally intensive [de Haan, 2011, Dehaan et al., 1987].

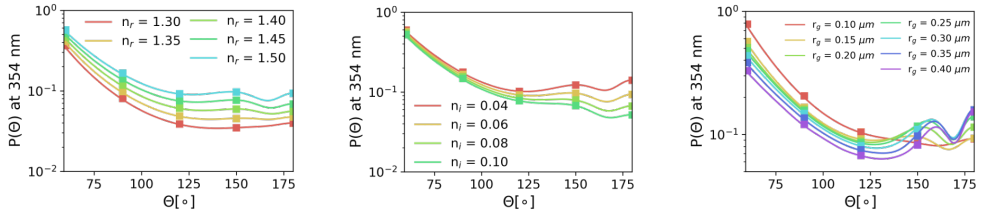
The sensitivity study is exclusively designed for biomass burning aerosols generated by the Chile wildfires. Table 2.1 lists the model parameters and corresponding values. The default values of micro-physical parameters are taken from the daily average of the AERONET station Santiago_Beauchef (33.46°S, 70.66°W) on 27 January 2017. We obtain the particle size distribution function and complex refractive index at 440, 675, 880 and 1018 nm from AERONET, and apply a linear interpolation/extrapolation to derive the complex refractive index over the spectrum from 340 to 675 nm (spectral resolution of 2 nm). Then we apply the Mie Theory to calculate the aerosol phase function $P(\Theta)$ and ω_0 (Θ is the scattering angle). The corresponding $P(\Theta)$ at 354 nm for different cases is presented in Figure 2.2. The aerosol profile is parameterized as a single homogeneous layer (box-shape), with its bottom at $z_{aer} - \Delta z/2$ and top at $z_{aer} + \Delta z/2$, where z_{aer} and Δz are the geometric central height and the geometric thickness of the aerosol layer. The whole sensitivity analysis is performed under cloud-free condition. The default wavelength pair to compute the AAI is the same as OMI (354/388 nm). To make the results of different situations comparable, the AAI calculated in this section is normalized by the maximum value in each case. Note that the sensitivity study always uses the default settings listed in Table 2.1 unless different values are explicitly mentioned.

Aerosol optical properties are determined by micro-physical properties, such as the complex refractive index (n_r and n_i) and the particle size (r_g). Figure 2.3 shows the impacts of these parameters on AAI, ΔI_{λ_1}, $I^{obs}_{\lambda_1}$, the corresponding ω_0 and the asymmetry factor (the averaged cosine of the scattering angle Θ, weighted by $P(\Theta)$, symbol: g). An increasing n_r directly enhances the $I^{obs}_{\lambda_1}$, whereas reduces ΔI_{λ_1} (Figure 2.3 (a)). This results in low values of AAI, which corresponds to a high ω_0 (Figure 2.3 (b)). Conversely, n_i has an opposite influence (Figure 2.3 (c, d)). The particle size distribution has more complicated influences on AAI. As shown in Figure 2.3 (e, f), AAI primarily follows the behaviour of ΔI_{λ_1}, whereas ω_0 is continuously decreasing and g is continuously increasing with the particle size.

Aerosol amount and vertical distribution also have influences on AAI. AAI is positively correlated with AOD, as $I^{obs}_{\lambda_1}$ measured by satellite decreases and ΔI_{λ_1} increases (Figure 2.4 (a)). Previous studies show that AAI is highly sensitive to the aerosol vertical distribution [de Graaf et al., 2005, Herman et al., 1997a, Torres et al., 1998]. Aerosol particles primarily interact with scattered photons beneath the aerosol layer. As the aerosol layer ascends (Figure 2.4 (b)), more underlying molecular scattering is

Tabel 2.1: Parameters and their values used in the sensitivity study.

Parameters	Default value	Sensitivity range	Unit
Geometric mean radius (r_g)	0.15	0.1, 0.15, 0.2, 0.25, 0.3, 0.35, 0.4	μm
Geometric standard deviation (σ_g)	1.5	-	μm
Real part of refractive index (n_r) at 354 nm	1.5	1.3, 1.35, 1.4, 1.45, 1.5	-
Imaginary part of refractive index (n_i) at 354 nm	0.06	0.04, 0.06, 0.08, 0.1	-
Aerosol layer geometric central height (z_{aer})	4.5	2.5, 4.5, 6.5, 8.5	km
Aerosol layer geometric thickness (Δz)	1	0.5, 1, 1.5, 2	km
Aerosol optical thickness (τ) at 550 nm	1	0.5, 1, 1.5, 2	-
Surface albedo (a_s) at 354 and 388 nm	0.05	0.05, 0.1, 0.5, 1.0	-
Surface pressure (P_s)	1013	1013, 963, 913, 863, 813	hPa
Solar zenith angle (θ_0)	30	0, 15, 30, 45, 60, 75	°
Viewing zenith angle (θ)	0	0, 15, 30, 45, 60, 75	°
Relative azimuth angle ($\Delta\phi = \phi - \phi_0 + 180°$)	0	0, 45, 90, 135, 180	°

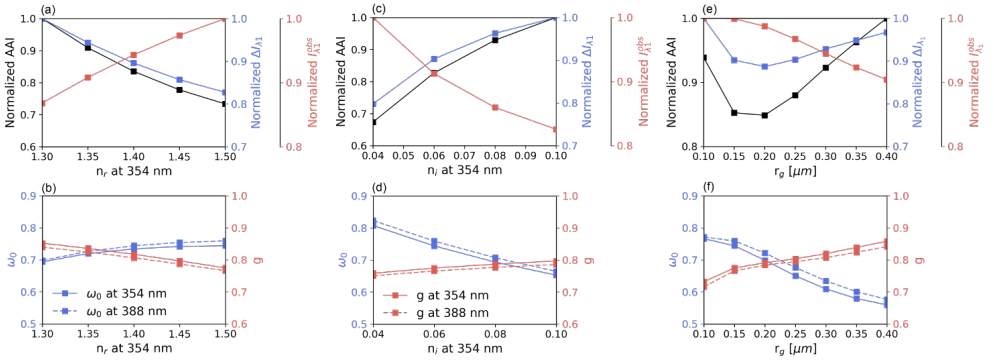


(a) $P_{354}(\Theta)$ as a function of n_r (with r_g = 0.15 μm and n_i = 0.06).

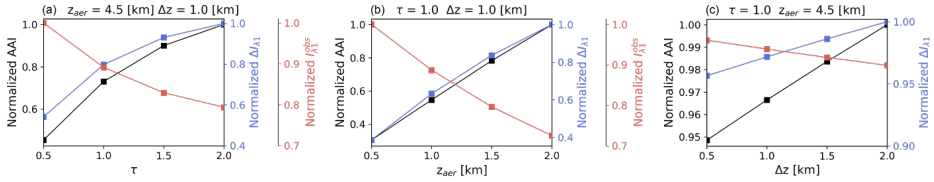
(b) $P_{354}(\Theta)$ as a function of n_i (with r_g = 0.15 μm and n_r = 1.5).

(c) $P_{354}(\Theta)$ as a function of r_g (with n_r = 1.5 and n_i = 0.06).

Figuur 2.2: $P(\Theta)$ at 354 nm of the parameterized Mie scattering aerosols in sensitivity study as a function of (a) n_r (with $r_g = 0.15\ \mu m$ and $n_i = 0.06$), (b) n_i (with $r_g = 0.15\ \mu m$ and $n_r = 1.5$), and (c) r_g (with $n_r = 1.5$ and $n_i = 0.06$). The markers in the plot correspond to the value when $\Theta = 60^\circ, 90^\circ, 120^\circ, 150^\circ, 180^\circ$.



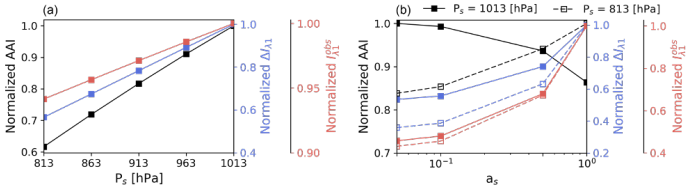
Figuur 2.3: AAI sensitivity to micro-physical parameters: (a, b) n_r , (c, d) n_i, and (e, f) r_g. The upper panels (a, c and e) show the sensitivity of the normalized AAI (black), the normalized ΔI_{λ_1} (blue) and the normalized $I^{obs}_{\lambda_1}$ (red). The lower panels (b, d and f) show SSA (ω_0, blue) and g (red) at wavelength 354 (solid line) and 388 (dashed line) nm, respectively.



Figuur 2.4: AAI sensitivity to macro-physical parameters: (a) AOD (τ) at 550 nm, (b) ALH (z_{aer}, unit: km), and (c) Δz (unit: km).

shielded, reducing $I^{obs}_{\lambda_1}$ whereas increasing ΔI_{λ_1}. The thickness of the aerosol layer, however, has a positive but limited effect (up to 5%) on AAI (Figure 2.4 (c)). The reason could be that a larger Δz allows the sunlight to have a higher possibility to be absorbed by aerosols, slightly enhancing the aerosol absorption.

AAI does not only depend on aerosols themselves, but also on the surface and clouds. Although AAI is capable to distinguish absorbing and non-absorbing agents [Herman et al., 1997a] and even to



Figuur 2.5: AAI sensitivity to surface parameters: (a) a_s and (b) P_s. The solid line and dashed line in (b) indicates a terrain height at sea level ($P_s = 1013hPa$) and an elevated terrain height ($P_s = 813hPa$), respectively.

Figuur 2.6: AAI sensitivity to θ and θ_0 at $\Delta\phi = 180°$. The black dashed contour in panel (a) indicates $\Theta = 60°$, 90°, 120°, 150°. The white dashed line in panel (a) indicates the cross section along viewing angles, with its corresponding normalized AAI, ΔI^{obs} and $I^{obs}_{\lambda_1}$ presented in panel (b).

retrieve aerosol information over clouds [Torres et al., 2012], the uncertainty triggered by clouds in the radiance is high, thus cloudy conditions are excluded from this study. Surface conditions are described by the surface pressure (P_s) and the surface albedo (a_s). P_s can represent the effect of topography. A decline in P_s, or equivalently an increase in terrain height, leads to less Rayleigh scattering shielded between the surface and the aerosol layer (Figure 2.5 (a)), which agrees with previous studies [Colarco et al., 2017, de Graaf et al., 2005]. Besides, increasing a_s has two counteracting effects [de Graaf et al., 2005]: it increases the amount of directly reflected radiation at the top of the atmosphere, i.e. a larger $I^{obs}_{\lambda_1}$; on the other hand, it enhances the role of absorption by the aerosol layer, i.e. a larger ΔI_{λ_1}. Which is the dominant effect depends on P_s (Figure 2.5 (b)). When the aerosol layer is relatively closer to the sea level ($P_s = 1013hPa$), the first effect dominates. However, a brighter surface compensates the loss of molecular scattering shielded by the aerosols when the terrain height rises ($P_s = 813hPa$), which makes the absorbing layer more detectable.

AAI also depends on the sun-satellite geometry. Here we provide AAI as a function of the measurement geometries for the default case with the relative azimuth angle $\Delta\phi = 180°$. As presented in Figure 2.6 (a), AAI becomes very sensitive to the geometries for the zenith angles larger than 60°, which confirms previous research [de Graaf et al., 2005, Herman et al., 1997a, Torres et al., 1998]. This is mainly due to the significant growth of $P(\Theta)$ when Θ becomes smaller (Figure 2.2). Thus, it is suggested that satellite measurements with the solar zenith angle (θ_0) or the viewing zenith angle (θ) larger than this value should be removed due to the large variations of AAI. To analyse the radiance behaviour as previously, we plot ΔI_{λ_1} and $I^{obs}_{\lambda_1}$ as a functions of Θ along the cross section, respectively (Figure 2.6 (b)). It is noted that $I^{obs}_{\lambda_1}$ increases when Θ is larger than 90°, whereas $P(\Theta)$ is generally low at this range (Figure 2.2). The reason could be that the Rayleigh scattering has an increasing contribution to the total radiance at those measurement angles.

2.3. METHODOLOGY AND DATA SETS

2.3.1. DATA SETS

OMI AND GOME-2 AAI

In this study, the OMI level 2 product OMAERO (`https://disc.gsfc.nasa.gov`) is used to provide AAI, together with the corresponding viewing geometry and the surface parameters. The samples are included in the radiative transfer simulation only if both θ_0 and θ is smaller than 60°, and satellite pixels are not contaminated by sun-glint, clouds, and row anomalies of the instrument. The simulation is only applied to pixels inside the biomass burning plume, which is defined as pixels with both OMI and GOME-2 AAI larger than 1.

MODIS AOD

MODIS/Aqua and Terra is a sensor specifically designed for atmosphere and climate research. The two satellites together ensure a daily global coverage. In this study, we use the MODIS/Aqua data as its overpass time (13:30 LT) is close to that of OMI. MODIS has a spatial resolution that ranges from 0.25 to 1 km and it has 36 spectral bands from 400 nm to 14.4 μm [Remer et al., 2005]. It employs different algorithms for aerosol retrieval over land and ocean [Hsu et al., 2004, Kaufman et al., 1997b, Remer et al., 2005, Tanré et al., 1997]. Currently the AOD provided by MODIS is one of the most reliable AOD sources [Lee et al., 2009, Sogacheva et al., 2020], with an estimated uncertainty of only $-0.04-10\%$ to $+0.02+10\%$ over ocean and $\pm(0.05+15\%)$ (Dark Target) / $\pm(0.05+20\%)$ (Deep Blue) over land [Levy et al., 2013, Remer et al., 2005]. Besides, the MODIS retrieved AOD is insensitive to the aerosol layer height [Satheesh et al., 2009]. This study uses cloud-filtered AOD at 550 nm from the Collection 6 level 2 product MYD04 as the input for radiative transfer calculations (`https://ladsweb.modaps.eosdis.nasa.gov`).

In addition, the AOD measured by OMI and MODIS are used to compare with AERONET. The OMAERO uses multi-spectral fitting techniques, and the retrieved AOD is reported in good accordance with AERONET and highly correlated with MODIS [Torres et al., 2007], with a correlation of 0.66 over land and 0.79 over ocean [Curier et al., 2008]. It suffers from cloud sub-pixel contamination due to the relatively coarse spatial resolution of OMI. Considering the wavelength difference, the AOD measured by OMI at 442 nm has to be transferred to 550 nm using the Ångström Exponent (ÅE) calculated between 440 and 675 nm of AERONET near the time when OMI flew over the selected site. The AERONET data used in this study is introduced in the next section.

AERONET MICRO-PHYSICAL PARAMETERS

AERONET is a ground-based sun photometer network. With standardized instruments, calibration, processing and distribution, AERONET provides a long-term global database for aerosol research, and air-borne and space-borne measurement validation [Holben et al., 1998]. The system takes two basic measurements. The AOD and ÅE are retrieved from the direct solar irradiance measurements [Eck et al., 1998]; the particle size, the phase function, the complex refractive index, and SSA are derived from multiple-angular measurements of sky radiance [Dubovik et al., 1998, Dubovik and King, 2000, Nakajima et al., 1983, 1996].

The AERONET site nearest to the Chile wildfires is the Santiago_Beauchef (33.46°S, 70.66°W, `https://aeronet.gsfc.nasa.gov`). The data used is version 2 level 1.5 product (cloud-screened but not quality-assured). To minimize the influence of time difference, the parameters of AERONET closest to the the OMI overpass time are used to simulate the aerosol optical properties. Note that the location of the site is in downtown of Santiago City and close to major roads, where the presence of scattering urban aerosols may bias the measurements as compared with lager scale satellite observations.

The AERONET retrieved AOD and SSA are used to evaluate the MODIS AOD and the retrieved SSA, respectively. The AERONET measured AOD is transferred to 550 nm using the ÅE calculated between 440 and 675 nm, while the SSA at 550 nm is linearly interpolated by the values at 440 and 675 nm.

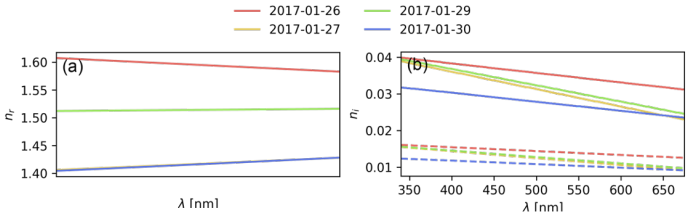
The complex refractive index and the particle size distribution provided by the AERONET need to be processed into the format required by DISAMAR. First, a conversion from the volume size distribution $V(r_v,\sigma_v)$ provided by AERONET to the number size distribution $N(r_g,\sigma_g)$ used in DISAMAR is required:

$$N(r_g,\sigma_g) = V(r_v,\sigma_v)\frac{3}{4\pi r_g^3}e^{-4.5\sigma_g^2} \tag{2.4}$$

The following relationships between the geometric and volumetric mean radii (r_g and r_v) and standard deviation (σ_g and σ_v) are assumed:

$$r_g = r_v e^{-3\sigma_g^2} \tag{2.5}$$

$$\sigma_g = \sigma_v \tag{2.6}$$



Figuur 2.7: Retrieved complex refractive index for each case: (a) n_r and (b) n_i. The dashed line in lower panel is the wavelength dependent n_i provided by AERONET.

The fine and coarse mode particle size are derived by finding the two peaks of the log-normal distribution function provided by AERONET. The complex refractive index is assumed to be the same for both modes. Since a bi-modal aerosol is not available in DISAMAR yet, we first calculate optical properties of two modes individually, then we externally combine the optical properties of two modes into a bi-modal aerosol with a fraction:

$$w_f = \frac{N_f(r_{g,f}, \sigma_{g,f})}{N_f(r_{g,f}, \sigma_{g,f}) + N_c(r_{g,c}, \sigma_{g,c})} \tag{2.7}$$

$$w_c = 1 - w_f \tag{2.8}$$

Then the weights for calculating the total ω_0 of the mixed aerosol are:

$$w_{\sigma,f} = \frac{w_f \sigma_f}{w_f \sigma_f + w_c \sigma_c} \tag{2.9}$$

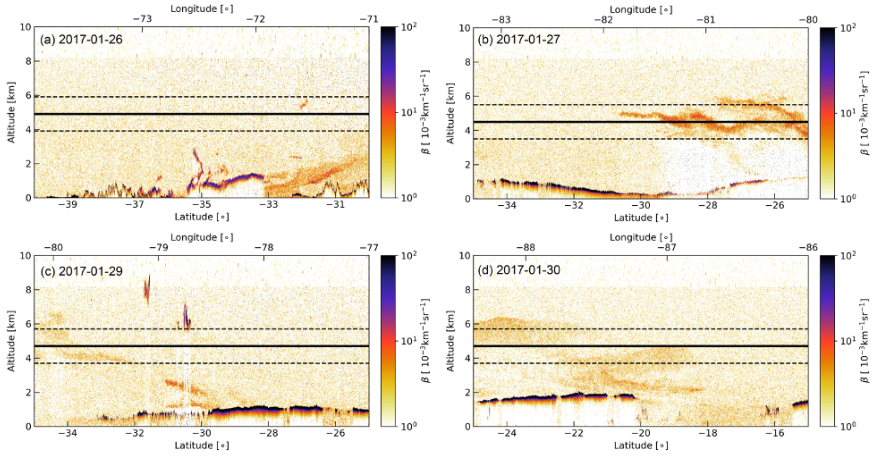
$$w_{\sigma,c} = 1 - w_{\sigma,f} \tag{2.10}$$

, where σ_f and σ_c are the extinction cross section of the fine and coarse aerosols. The expansion coefficients of the mixed aerosol are weighed by the ω_0 of the fine and coarse aerosols ($\omega_{0,f}$ and $\omega_{0,c}$), respectively:

$$w_{\omega_{0,f}} = \frac{w_f \sigma_f \omega_{0,f}}{w_f \sigma_f \omega_{0,f} + w_c \sigma_c \omega_{0,c}} \tag{2.11}$$

$$w_{\omega_{0,c}} = 1 - w_{\omega_{0,f}} \tag{2.12}$$

The AERONET instrument at this site only covers the visible and infrared band (from 440 nm to 1018 nm) for sky radiance measurements, i.e. no information in the UV band. Due to the absence of observations, assumptions have to be made on the spectral dependency of aerosol properties to obtain their values in the UV range. The properties of biomass burning aerosols depend on the type of fuel, the procedure producing the smoke, the age of the smoke, and also the atmospheric conditions [Reid et al., 2005]. Using measurements to constrain the input aerosol refractive index may reduce the uncertainties due to a priori knowledge. Our treatment of the complex refractive index is as follows: (1) take the complex refractive index in the visible band (from 440 to 675 nm) of the AERONET measurements; (2) linearly extrapolate the complex refractive index to the UV band. The real part n_r for radiative transfer calculation is obtained in this step. A slight wavelength dependency of n_r is found (Figure 2.7 (a)); (3) for the imaginary part n_i, we multiply it (for the entire wavelength from UV to visible) with a scaling factor as we set it as a free parameter. By varying the value of the scaling factor, both the magnitude and the wavelength dependency of n_i can change to meet the requirement for retrieval (Figure 2.7 (b)).



Figuur 2.8: CALIOP backscatter coefficient β at 532 nm. The solid and dashed line indicate the retrieved z_{aer} and Δz, respectively. The red to black dots indicate clouds and the orange dots indicate aerosol layers, respectively.

CALIOP BACKSCATTER COEFFICIENT

The CALIOP/CALIPSO launched in 2006 provides high-resolution profiles of aerosols and clouds. It has three channels with one measuring the backscatter intensity at 1064 nm and the rest measuring orthogonally polarized components at 532 nm [Winker et al., 2009]. In this study, we only use the total attenuated backscatter at 532 nm from level 1B Version 4.10 Standard data to evaluate the retrieved parameterized aerosol profiles (`https://eosweb.larc.nasa.gov/project/calipso/calipso_table`). But note that due to the limited spatial coverage, CALIOP did not capture the plume produced by the wildfires for all the cases where there were OMI observations.

2.3.2. METHODOLOGY

In this study, we employ the radiative transfer model DISAMAR to simulate the AAI from OMI and to derive the ω_0 for a biomass burning event, i.e. the Chile wildfires in January 2017. We select the period from 26 to 30 January 2017 (28 January is excluded due to a lack of data) when the OMI AAI value reached its peak during the fire event.

The forward simulation consists of two steps. First, DISAMAR calculates the Mie aerosol optical properties with aerosol micro-physical information taken from AERONET measurements (r_g, n_r and n_i). As mentioned in Section 5.2.1, we set the spectral-dependent imaginary refractive index n_i as a free parameter to vary SSA. Then, DISAMAR operates radiative transfer calculations with the input aerosol properties and environmental conditions of OMI.

It must be noted that the observed aerosol vertical distribution information is limited for this case. Previous research suggested that AAI cannot be quantitatively used without AOD or ALH information [Gassó and Torres, 2016]. Therefore, we implement the same parameterization on aerosol profiles as that used in the sensitivity study. Since AAI dependency on Δz is minor (Figure 2.4 (c)), and to reduce the computational cost, Δz is set to be a fixed value of 2 km based on the information from the CALIOP backscatter coefficient measurements (β) at 532 nm (Figure 2.8). The ALH, to which the AAI is highly sensitive, is treated as an unknown variable to be retrieved together with SSA.

With various combinations of ALH and SSA, a Look-Up Table (LUT) of the calculated AAI is constructed by radiative transfer calculations. It should be noted that for all pixels in the plume we assume the same aerosol micro-physical properties and the same aerosol layer height. Pixels outside the plume (even if they have AAI values larger than 1) may have significantly different properties which

will affect the retrievals. Consequently, we apply a data quality control procedure before retrieving SSA. First, we manually remove the pixels that are geographically isolated from the main plume. Then, we remove potential outliers based on statistical analysis. We filter the data set using an outlier detection criterion based on the inter-quartile range (IQR) of the AAI difference between the DISAMAR simulations and the OMI observations. According to Tukey's fences [Tukey, 1977], an AAI difference falling outside range between Q1-1.5 IQR and Q3+1.5 IQR may be regarded as an outlier and removed, where Q1 and Q3 are the first and third quartile of the AAI difference, and the IQR is the range between Q1 and Q3. Only pixels passing the outlier detection criterion are used to calculate the cost function:

$$RMSE = \sqrt{\frac{\sum_i^n \left(AAI_{sim,i}^q - AAI_{obs,i}\right)^2}{n}} \tag{2.13}$$

Here AAI_i indicates the AAI for i-th satellite pixel; q means the qualified pixels passing the outlier detection; sim and obs indicate the DISAMAR simulations and the OMI observations, respectively. The combination of ALH and SSA that leads to the minimum residue is used as retrieval results. The estimated ALH and SSA at 550 nm are evaluated by independent observations from CALIOP and AERONET, respectively.

2.4. RESULTS AND DISCUSSION

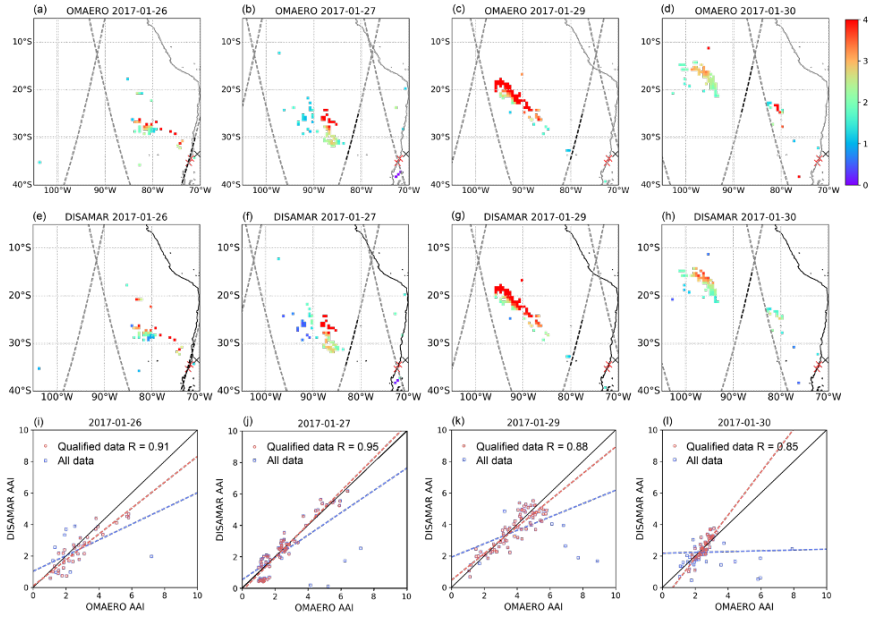
2.4.1. VALIDATION OF THE RETRIEVED SSA

The OMI measurements of the plume are displayed in Figure 2.9 (first row). The satellite pixels are presented for AAI values larger than 1, and are free of cloud contamination, sun-glint and row anomaly of the instrument. Fortunately, the remaining data are still able to capture the plume features. It can be clearly seen that from 26 to 30 January, the plume was transported by the south-easterly trade wind from the continent towards the lower latitude region of the Pacific Ocean. The plume travelled over 3000 km during the period.

The vertical movement of the plume is given by in Figure 2.8. The CALIOP overpasses closest to the plume are marked by a black dashed line in Figure 2.9. It is noted that due to the spatial coverage and the time difference, CALIOP is not able to represent the entire plume detected by OMI. The aerosol layer captured by CALIOP is located between 2 and 6 km, with an average height at approximately 4-5 km. The ascent of the plume was driven by the heat from the fire and aerosol sunlight absorption, as well as the atmospheric vertical motions.

Figure 2.9 (e-h) shows the simulated AAI 2.3.2. The spatial distribution of the simulated AAI shows similar patterns to the OMI observations. Some data points that are geographically isolated from the plume, e.g. in case 26 and 30 January, may differ from what are observed inside the plume. Including these outliers in the optimization could bias the retrieved aerosol properties. This can also be seen in Figure 2.9 (i-l), where the points passing the data quality control described in Section 2.3.2 are highlighted in red color. By removing outliers, the correlation coefficients (R) are generally no less than 0.85.

Table 2.2 lists the statistics of the qualified AAI data, including median, relative difference and RMSE. The median of measured AAI ranges from 2 to 4 during the research period. Except for 26 January, the median of the simulated AAI is in good agreement with the measurements, with relative differences within ±6%. The majority of the simulated AAI for 26 January is negatively biased, which is reflected in Figure 2.9 (a, e). A systematic bias in the inputs might exist. In terms of SSA, both the AERONET measurements and our retrievals show that aerosol absorption become weaker with time (Table 2.2). Although the simulated and observed AAI are generally in good agreement, the difference in SSA is significant. The mean of the retrieved SSA at 550 nm for the whole period is 0.84, whereas that of the AERONET measurements is 0.90, which is outside the expected AERONET uncertainty of ±0.03 [Dubovik et al., 2000, 2002b].



Figuur 2.9: The OMI AAI observations (a–d) and DISAMAR simulations (e–h) of the Chile wildfires on 26, 27, 29 and 30 January 2017. The black and red cross symbols are the AERONET station and the main fire sources (Pichilemu W34.39° S72.00° and Consititución S35.33°, W72.42°), respectively. The gray dashed line indicates the CALIOP overpasses, and the part that is used to validate the plume height are marked by black dashed line. The scatter plots (i–l) present the OMI observations against the DISAMAR simulations for only qualified data (red dot) and all data (blue dot), respectively.

Tabel 2.2: Summary of retrieved results (after applying IQR outlier detection).

Date		01/26	01/27	01/29	01/30
Number of pixels in the plume		44	70	82	75
AAI	AAI median (OMAERO)	2.52	2.38	4.05	2.61
	AAI median (DISAMAR)	2.17	2.48	3.81	2.49
	Relative difference	-13.88%	4.2%	-5.93%	-4.6%
	RMSE	0.67	0.51	0.6	0.41
Aerosol profile	z_{aer} $[km]$	4.9	4.5	4.7	4.7
	Δz $[km]$	2	2	2	2
n_i	n_i at 354 nm	0.0395	0.0382	0.0388	0.0314
	n_i at 388 nm	0.0386	0.0366	0.0373	0.0306
	Relative difference between 354 and 388 nm	2.33%	4.37%	4.02%	2.61%
ω_0 at 550 nm	ω_0 (AERONET)	0.89	0.89	0.92	0.91
	ω_0 (DISAMAR)	0.83	0.81	0.87	0.85
	Relative difference	-6.74%	-8.99%	-5.43%	-6.59%

2.4.2. ERROR ANALYSIS OF THE RETRIEVED SSA

There are many error sources contributing to the discrepancy in SSA. First, the nearest site Santiago_Beauchef did not directly measure the smoke plume as mentioned in Section 5.2.1. The site is located in downtown, where reflective urban or industrial aerosols may have been mixed with the

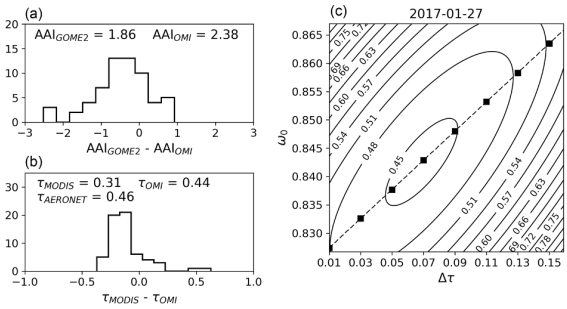
smoke and enhanced the SSA. This would also affect the spectral dependency of the complex refractive index used in the radiative transfer calculations. Table 2.2 reveals that the difference of n_i between 354 and 388 nm is less than 5%. This small spectral dependency of n_i is mainly determined by AERONET measurements in the visible band (dashed lines in Figure 2.7). We thus find a much weaker wavelength dependency than Jethva and Torres [2011a], where a 20% difference of n_i between the two UV wavelengths was applied to the OMAERUV algorithm and achieved a result 70% of the retrieved SSA falls within the ±0.03 AERONET SSA uncertainty. This 20% spectral dependency adopted in their work is associated with findings of [Kirchstetter et al., 2004]. According to them, the absorbing Ångström exponent (AÅE) of biomass burning aerosols is approximately 2 between 0.3 μm to 1 μm. In terms of n_i, a 20% increase at 354 nm with respect to the value at 388 nm is equivalent to an AÅE value between 2.5 and 3, depending on the aerosol models of OMAERUV [Jethva and Torres, 2011a]. Hoffer et al. [2006] also found a similar result. They conducted in situ measurements on humic-like substances (HULIS) of Amazonia biomass burning aerosols and found that around 35 to 50% light absorption occurred at 300 nm, whereas only around 15% at 400 nm. Bergstrom et al. [2007] also confirmed this conclusion from several field programs (SAFARI 2000, ACE Asia, PRIDE, TARFOX and INTEX-A). From the sensitivity study of [Jethva and Torres, 2011a], a stronger spectral dependency of n_i between 354 and 388 nm would allow to reach the same AAI with n_i at a relatively low level. In our study, this leads to retrieve a higher SSA at 550 nm. The presence of non-absorbing aerosols may reduce the measured n_i at visible band and the linear extension would weaken the spectral dependency particularly in the near-UV spectral range.

On the other hand, the AERONET inversion product is not error-free. The uncertainty of the size distribution retrieval may be minor for biomass burning aerosols [Dubovik et al., 2000], but under optically thick circumstances, even when retrievals are quality-assured (i.e. level 2 data), the reported accuracy of complex refractive index is 0.04 for n_r and 30-50% for n_i, respectively [Dubovik et al., 2002b]. It is also reported that AERONET tends to underestimate the absorption of biomass burning aerosols compared with in situ measurements [Dubovik et al., 2002b, Reid et al., 2005]. The uncertainty of SSA is 0.03 under high aerosol loading (AOD at 440 $nm > 0.5$) and 0.05-0.07 under low aerosol loading [Dubovik et al., 2002b, Holben et al., 2006].

Last but not least, the spatial representation of the AERONET site is also concerning. Santese et al. [2007] showed that the selected AERONET aerosol parameters can be representative of a 300×300 km^2 area. For the Chile wildfires with the most remote pixels over 3000 km away from the continent, the measurements at the selected site cannot fully represent the plume detected by the satellite.

Apart from AERONET, other parameters in the radiative transfer simulations could also bias the estimate of aerosol absorption, among which the parameterization of the one-layer-box-shape profile could be the largest error source. Although the influence of Δz on AAI is negligible (Figure 2.4 (c)), AAI calculation highly depends on ALH (z_{aer}, Figure 2.4 (b)). As shown in Table 2.2, the estimated plume altitude varies from 4.5 to 4.9 km. As the black solid lines indicated in Figure 2.8, the retrieved ALH can approximately capture the plumes. The ALH on 26 January is overestimated because of the large temporal and spatial difference. CALIOP sampled the plume near the fire source that is close to the surface, whereas the plume observed by OMI had been already elevated and transported to the open ocean. The lack of information on the real plume height makes it challenging to determine whether the plume height is responsible for the systematic bias in Figure 2.9 (i). Except for 26 January, ALH in other cases is generally consistent with CALIOP. Although the retrieved ALH are convincing to some extent, one should keep in mind that CALIOP and OMI observations are not exactly co-located. Besides, the parameterized aerosol profiles may fail to represent the spatial variation of the plume. Therefore, the uncertainty cannot be directly determined due to the lack of validation.

Compared with other cases, the retrieved SSA of 27 January is significantly lower than others. For this day, the agreement in terms of ALH with CALIOP is reasonable. We therefore explore the effect of observational biases of AAI and AOD on the retrieved SSA. We investigate the potential bias by plotting the histogram of the AAI measurement difference between GOME-2 and OMI (Figure 2.10 (a)) against the AOD measurement difference between MODIS and OMI (Figure 2.10 (b)), both are at 550 nm). It is



Figuur 2.10: Histogram of (a) the AAI difference between GOME-2 and OMI, against (b) the AOD difference at 550 nm between MODIS and OMI for 27 January. Contour of (c) shows the AAI RMSE as a function of variation in AOD and SSA for this date. The dashed line is the best estimation for each pair of $\Delta\tau$ and ω_0.

clear that on 27 January, the AAI from OMI seems to be overestimated compared to that of GOME-2. Despite the difference in instrumental design and wavelength pair for AAI retrieval, measurement conditions, etc., could contribute to the discrepancy between GOME-2 and OMI, exploring the difference between the two is beyond the scope of this study. In terms of input aerosol concentration, the AOD from MODIS is generally lower than that of OMI. Fitting higher AAI with lower AOD leads to an overestimation in aerosol absorption. Thus, we analytically quantify the impact of AOD for this specific case by systematically enhancing the MODIS AOD with a constant variation ($\Delta\tau$) added to all pixels, with AAI and ALH unchanged. Figure 2.10 (c) presents how the AAI RMSE and the estimated SSA respond to the enhanced AOD. It can be clearly seen that an increase in AOD by 0.07 raises SSA to 0.84 and optimizes the AAI simulation to a RMSE less than 0.45. If we apply this AOD adaption, the retrieved SSA of 27 January becomes more consistent with the other days as well as the AERONET measurements.

Apart from the observational errors in AERONET, OMI and MODIS, the assumption that the plume features are homogeneous could also result in the SSA discrepancy. In reality, the plume altitude, the optical properties and even the chemical compositions could vary from pixel to pixel, whereas our simulations cannot consider those effects.

2.5. CONCLUSION

Biomass burning aerosol is a major type of absorbing aerosols, posing a significant contribution on the climate warming. Quantitatively characterizing the absorption of biomass burning aerosols is therefore important to reduce the uncertainty of global radiative forcing assessments. Facing the lack of continuous and global SSA data, this study proposes as an alternative method to retrieve SSA based on satellite AAI. Although AAI is not a geophysical parameter, its calculation is independent of a priori aerosol properties and its high sensitivity to aerosol absorption, offer potentials for aerosol absorption research.

In this study, we retrieve SSA together with ALH by fitting the radiative transfer simulated AAI to the OMI retrieved AAI for the wildfires that occurred in Chile in January 2017. After filtering outliers, the high correlation coefficients (R>0.85) between the simulated and the observed AAI shows the potential of this method. The retrieved aerosol profiles indicate the smoke plume was elevated to 4.5-4.9 km. These results are in agreement with the independent CALIOP measurements. The retrieved SSA at 550 nm of the plume is approximately 0.84, which is 0.06 lower than that of AERONET. The retrieved SSA is acceptable, if one takes into account: (1) the typical uncertainty in the SSA retrieved from AERONET (±0.03); (2) the locations of the plume are outside the AERONET representative range; (3) the assumption of homogeneous and static plume properties, which ignores the plume evolution over

space and time; (4) the simplified parameterization of the aerosol profiles; and (5) the observational errors in AAI and AOD, as well as the retrieval errors in the aerosol micro-physics provided by AERONET.

This study proves the possibility of utilizing the satellite AAI to quantitatively characterize aerosol optical properties (i.e. SSA). However, apart from the observational uncertainties, the current method is probably limited by: (1) the lack information on the aerosol vertical distribution; and (2) the poorly-understood aerosol absorption spectral dependence in the near-UV band, which should be further studied in future.

3

QUANTIFYING THE SINGLE SCATTERING ALBEDO FROM THE ABSORBING AEROSOL INDEX: RADIATIVE TRANSFER VERSUS MACHINE LEARNING

The content of this chapter has been published as: **Sun, J.**, Veefkind, P., Nanda, S., van Velthoven, P., and Levelt, P.: The role of aerosol layer height in quantifying aerosol absorption from ultraviolet satellite observations, Atmos. Meas. Tech., 12, 6319–6340, https://doi.org/10.5194/amt-12-6319-2019, 2019.

3.1. INTRODUCTION

The concept of the near Ultra-Violet (near-UV) absorbing aerosol index (UVAI) initially appeared in the TOMS/Nimbus-7 ozone product. It is used to detect elevated UV-absorbing aerosol layers by measuring the spectral contrast difference between an observed radiance in reality and a simulated radiance in a Rayleigh atmosphere [Herman et al., 1997a]:

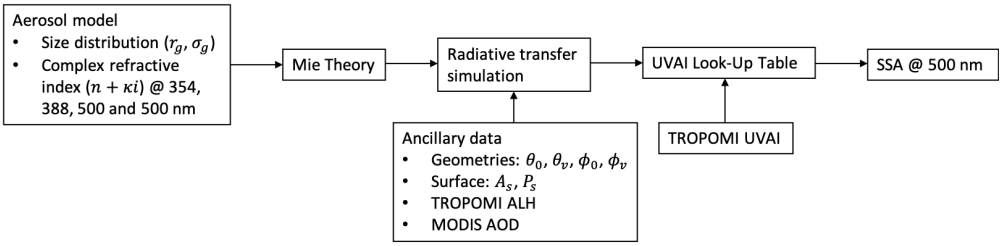
$$UVAI = -100\left[log_{10}\left(\frac{I_{\lambda_1}}{I_{\lambda_2}}\right)^{obs} - log_{10}\left(\frac{I_{\lambda_1}}{I_{\lambda_2}}\right)^{Ray}\right] \tag{3.1}$$

where the superscripts obs and Ray denote the radiance from observations and simulations; I_{λ_1} and I_{λ_2} are the radiance at wavelength λ_1 and λ_2 ($\lambda_1 < \lambda_2$); λ_2 is the longer wavelength at which a spectrally constant scene reflectivity is assumed for the calculation of I_{λ}^{Ray}. A positive UVAI value indicates the presence of absorbing aerosols, whereas a negative or near-zero value implies non-absorbing aerosols or clouds [Herman et al., 1997a]. The over 4 decades of UVAI observations (1978 to present) would be beneficial to derive aerosol absorption properties, e.g. the single scattering albedo (SSA). Aerosols are considered to be the largest error source of radiative forcing assessments [Pachauri et al., 2014], and SSA is one of the key parameters to reduce this uncertainty [Haywood and Shine, 1995].

The magnitude of UVAI depends on many factors [de Graaf et al., 2005, Herman et al., 1997a, Hsu et al., 1999, Sun et al., 2018, Torres et al., 1998], among which the most dominant are aerosol concentration, aerosol vertical distribution and aerosol optical properties [Buchard et al., 2017, Wang et al., 2012]. To derive SSA from UVAI, information on two other parameters are necessary. The aerosol concentration is usually represented by the aerosol optical depth (AOD). There are many AOD products available. By contrast, there is much less observations on the aerosol vertical distribution. The most well-known product is provided by the space-borne lidar CALIOP/CALIPSO, but the number of measurements is limited due to its narrow tracks [Winker et al., 2009]. Passive sensors also make efforts to retrieve the aerosol layer height (ALH) from radiance. For example, Chimot et al. [2017] presented the feasibility to retrieve ALH using the OMI oxygen band at 447 nm; Xu et al. [2019, 2017] attempted to retrieve ALH from the EPIC oxygen absorption band for dust and carbonaceous aerosols over land and ocean; Tilstra and Stammes [2019] developed an algorithm to derive the absorbing aerosol layer height from the GOME-2 FRESCO cloud layer height product. Recently a new ALH product has been run operationally, based on the measurements in the near-infrared (NIR) oxygen A-band of TROPOMI/S5-P [Sanders and De Haan, 2016]. TROPOMI has a wide swath of 2600 km, providing daily global coverage with a spatial resolution of $7 \times 3.5 km^2$ in nadir. The instrument is equipped with both UV–visible (270–500 nm) and NIR (675–775 nm) channel, which can simultaneously provide UVAI and ALH product [Veefkind et al., 2012].

Using the TROPOMI ALH, this work attempts to quantify aerosol absorption from UVAI for biomass burning aerosols. Two experiments are conducted. First, following previous studies [Colarco et al., 2002, Hu et al., 2007, Jeong and Hsu, 2008, Sun et al., 2018], we create a Look-Up Table (LUT) of simulated UVAI for various aerosol optical properties by a radiative transfer model (RTM). The SSA is derived by minimizing the difference between the pre-calculated UVAI and the satellite observed one. The major uncertainties in the retrieved SSA are caused by assumptions regarding the wavelength dependent complex refractive index and the lack of the aerosol vertical distribution information [Sun et al., 2018]. Now, with the TROPOMI ALH constraining forward simulations, it is expected to partly reduce the SSA retrieval uncertainty and to quantify the influence of the assumed aerosol refractive index on the retrieved SSA.

Although the availability of ALH can better constrain the radiative transfer calculations and improve the SSA retrieval, assumptions regarding aerosol micro-physics is still inevitable. Therefore, in the second experiment, we propose an numerical method to predict aerosol absorption based on the long-term record of co-located UVAI, ALH, AOD and absorbing aerosol optical depth (AAOD) using a



Figuur 3.1: Flowchart of the radiative transfer simulations of UVAI. The aerosol models are from the ESA Aerosol_cci project [Holzer-Popp et al., 2013] and the OMAERUV algorithm [Torres and Chen, 2013, Torres et al., 2007]. The satellite inputs are the TROPOMI geometries and ALH, the MODIS AOD and the OMI surface climatology. The aerosol profile is parameterized as a one-layer-box-shape, with the geometric central height set to the TROPOMI ALH and an assumed constant pressure thickness of 50 hPa.

Machine Learning algorithm. Machine Learning allows algorithms to learn the underlying behavior of a system from a given training data set. They are particularly useful to address ill-defined inversion problems of geoscience and remote sensing, where the theoretical understanding is incomplete but there is a large number of observations [Lary et al., 2016a]. We employ Machine Learning to quantify aerosol absorption from UVAI in order to avoid explicit assumptions on aerosol micro-physics as made in the first experiment. Currently, operational ALH products covering the globe on a daily basis are not abundantly available (See Chapter 4 for more information); therefore, we will use the ALH reported in the OMAERUV product to train the Machine Learning algorithm. Various Machine Learning algorithms have been developed to address regression problems. In this paper, we choose the Support Vector Regression (SVR), a regression variant form of the Support Vector Machines (SVM, [Drucker et al., 1997]). Compared with other algorithms (e.g., the Artificial Neural Network), SVR is less sensitive to the training data set size and can successfully work with a limited quantity of data [Mountrakis et al., 2011, Shin et al., 2005].

This work is organized as below: the first experiment is outlined in Section 3.2, including a description of the radiative transfer simulation settings and the analysis of the uncertainty trigger by the assumed aerosol aerosol absorption properties; Section 3.3 starts with introduction of SVR, followed by descriptions of the training data set, SVR model hyper-parameter tuning, error analysis and case applications. The major conclusions and implications for future research are summarized in Section 3.4.

3.2. EXPERIMENT 1: SSA RETRIEVAL USING RADIATIVE TRANSFER SIMULATIONS

In this section, we retrieve SSA using a radiative transfer model as done in our previous study [Sun et al., 2018]. Forward radiative transfer simulations are applied by DISAMAR [de Haan, 2011]. Figure 3.1 illustrates the model inputs and the procedure. For each pixel, the aerosol optical properties are computed by the Mie Theory for various predefined aerosol models. Then, DISAMAR calculates UVAI for each aerosol model with the corresponding information of AOD, ALH, measuring geometries (θ_0, θ_v, ϕ_0 and ϕ_v), surface albedo (A_s) and the surface pressure (P_s) of the target pixel. The output of the forward simulations is a LUT of UVAI as a function of the input SSA (determined by the predefined aerosol models) that is fit by a second-order polynomial function. Finally, by giving the corresponding satellite retrieved UVAI, the SSA of the target pixel is estimated from the pre-calculated UVAI–SSA relationship. The retrieved SSA is reported at 500 nm in order to compare it with the results of the SVR method in Section 3.3. Section 3.2.1 introduces the input parameters for the radiative transfer simulations, followed by retrieval results in Section 3.2.2.

Tabel 3.1: Aerosol models used in the radiative transfer simulations. $\Delta\kappa$ is the relative difference between κ_{354} and κ_{388}, defined as $\Delta\kappa = (\kappa_{354} - \kappa_{388})/\kappa_{388}$.

Geometric radius (r_g)	Effective radius (r_{eff})	Geometry standard deviation (σ_g)	Variance ($ln\sigma_{eff}$)	Refractive index real part (n)	Spectral dependence ($\Delta\kappa$)	Refractive index imaginary part at 354 nm (κ_{354})	Refractive index imaginary part of other wavelengths ($388nm$)
0.07 μm	0.14 μm	1.7	0.53	1.5	0%, 5%, 10%, 15%, 20%, 25%, 30%, 35%, 40%	$(1+\Delta\kappa)\times\kappa_{388}$	0.005 0.01 0.02 0.03 0.04 0.048 0.06

3

3.2.1. RADIATIVE TRANSFER SIMULATION CONFIGURATIONS

AEROSOL MODELS

The aerosol models are constructed from the ESA Aerosol_cci project [Holzer-Popp et al., 2013] and the OMAERUV algorithm [Torres and Chen, 2013, Torres et al., 2007]. We assume a fine-mode smoke aerosol type and further divide it into 7 sub-types as listed Table 3.1. We use the particle size distribution of the fine-mode strongly absorbing aerosol from the ESA Aerosol_cci project: with the geometric radius (r_g) of 0.07 μm (effective radius r_{eff} of 0.14 μm), and the geometric standard deviation (σ_g) of 1.7 (logarithmic variance $ln\sigma_g$ of 0.53). The real part of the refractive index (n) uses the same value as that in the OMAERUV algorithm, which is set to be 1.5 for all sub-types and spectrally flat. We adopt the imaginary part of the refractive index at 388 nm (κ_{388}) of the OMAERUV smoke sub-types (except for BIO-1 whose κ_{388} is 0) and add an extra sub-type with a κ_{388} of 0.06.

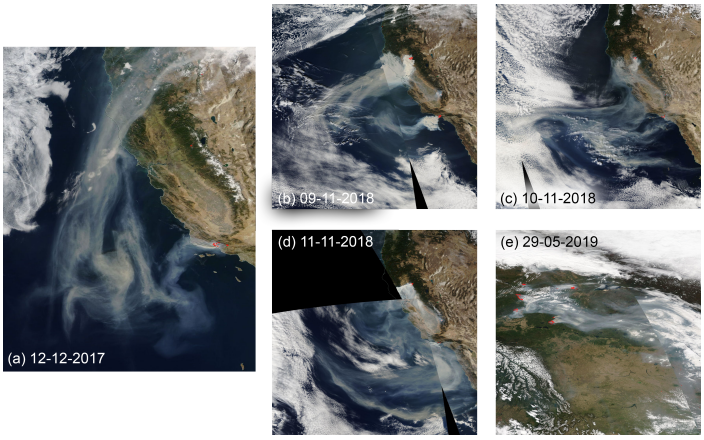
Many studies have shown evidence that the absorption by biomass burning aerosols in the near-UV band has a spectral dependence [Bergstrom et al., 2007, Kirchstetter et al., 2004, Russell et al., 2010]. Accordingly, a constant 20% $\Delta\kappa$ has been applied to all smoke sub-types in the recent OMAERUV algorithm [Jethva and Torres, 2011a], where $\Delta\kappa$ is defined as the relative difference between κ_{354} and κ_{388} (i.e. $\Delta\kappa = (\kappa_{354} - \kappa_{388})/\kappa_{388}$). In this experiment, we will investigate how the retrieved SSA responds to the assumed spectral dependence by considering nine different $\Delta\kappa$ values from 0% (i.e. "gray" aerosols) to 40% (very strong spectral dependence). This corresponds to an Absorbing Ångström Exponent (AÅE, α_{abs}) ranges from 1 to 3.4 and from 1.3 to 4.7, depending on aerosol sub-types. Note that the $\Delta\kappa$ is only applied between 354 and 388 nm. As we only investigate the influence due to aerosol absorption spectral dependence in the near-UV range. Aerosol absorption at wavelengths larger than 388 nm is set equal to that at 388 nm.

To summarize, the first experiment consists of 9 cases represented by different $\Delta\kappa$. Within each case, there are 7 predefined aerosol sub-types with varying κ_{388}. Thus, totally 63 forward simulations are performed for each individual pixel.

INPUTS FROM SATELLITES

Figure 3.1 presents the input parameters for the radiative transfer simulations of UVAI. Satellite measurement geometries (θ_0, θ_v, ϕ_0 and ϕ_v) and the surface pressure (P_s) of the TROPOMI UVAI reprocessed product (`https://scihub.copernicus.eu`) are used as input for forward simulations. The TROPOMI UVAI is calculated for two different wavelength pairs. One uses the conventional 340/380 nm wavelengths to continue the heritage of UVAI records from multiple sensors, and the other uses 354/388 nm in order to allow for comparisons with OMI measurements [Stein Zweers, 2016]. In this study we use the 354/388 nm pair.

The TROPOMI ALH is retrieved at the oxygen A-band (759–770 nm), where the strong absorption of oxygen causes the highly structured spectrum (`https://scihub.copernicus.eu`). This feature is



Figuur 3.2: Smoke plumes captured by MODIS/Aqua (source: https://gibs.earthdata.nasa.gov). The red regions indicate fires and thermal anomalies.

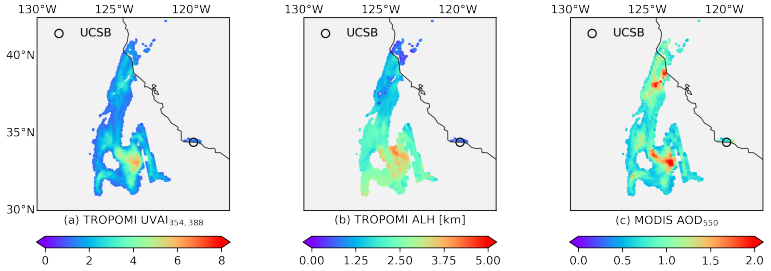
particularly suitable for elevated and optically dense aerosol layers, thus currently the TROPOMI ALH is only available for pixels with $UVAI_{354,388}$ larger than 1 [Sanders et al., 2015, Sanders and De Haan, 2016]. The ALH is reported in both altitude and pressure. For the radiative transfer calculations, the input aerosol profile is parameterized in the same way as that in the retrieval algorithm: a one-layer-box-shaped profile with a central layer height to be derived and an assumed constant pressure thickness of 50 hPa [Sanders and De Haan, 2016]. At the same band, the TROPOMI FRESCO cloud support product provides cloud fraction (CF) for detecting cloud effects, as will be explained in the following (https://scihub.copernicus.eu, [Wang et al., 2008a]).

The TROPOMI AOD product is not operational yet; thus, we use AOD from the level 2 MYD04 product (Collection 6) of MODIS/Aqua (https://doi.org/10.5067/MODIS/MYD04_L2.006). Aqua has an overpass time similar to S-5P (both are at 13:30 LT). The MODIS AOD at 550 nm used in the RTM-based method is a combination of *Deep Blue Aerosol Optical Depth 550 Land* and *Effective Optical Depth Op55um Ocean* [Levy et al., 2013]. A_s used to retrieve TROPOMI UVAI is not reported in the product. Instead, we use the OMI/Aura level 3 Lambertian Equivalent Reflectance (LER) monthly climatology calculated from the measurements between 2005 and 2009 (https://doi.org/10.5067/Aura/OMI/DATA3006, [Kleipool et al., 2008]). TROPOMI and OMI have similar overpass times (both are at 13:30 LT) and measurement geometries [Levelt and Noordhoek, 2002, Veefkind et al., 2012]. Due to the different spatial resolutions, the TROPOMI ALH, the OMI LER climatology and the MODIS AOD are re-sampled onto the TROPOMI UVAI grid and pre-processed to exclude pixels with large solar zenith angles ($\theta_0 > 70°$), weak aerosol absorption ($UVAI_{354,388} < 1$), insignificant aerosol amount ($AOD_{550} < 0.5$) or cloud contamination ($CF > 0.3$).

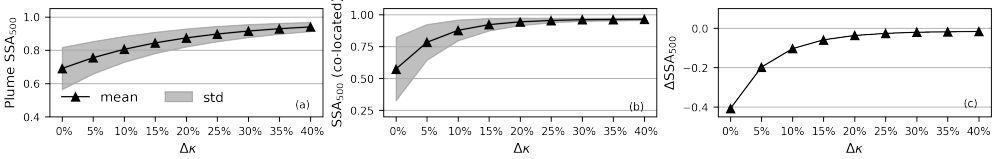
3.2.2. SSA RETRIEVED BY RADIATIVE TRANSFER SIMULATIONS

In the first experiment, we focus on one of the largest fire events that occurred in southern California in 2017, i.e., the Thomas Fire (http://www.fire.ca.gov/current_incidents/incidentdetails/Index/1922). Figure 3.2 (a) shows the RGB plume captured by MODIS. The major part of the plume was over the ocean and under cloud-free conditions, which is favorable for space-borne aerosol observations. There is a total of 5,217 pixels in this case. Figure 3.3 presents the corresponding UVAI, ALH and AOD data after pre-processing. The highest UVAI appeared at the southern part of the plume, where both the aerosol loading and aerosol layering were also high ($AOD_{550} > 2$ and $ALH > 2.5\ km$).

Figure 3.4 (a) displays the mean SSA of all plume pixels retrieved by the RTM-based method



Figuur 3.3: Satellite data from the California fire event on 12 December 2017: (a) TROPOMI UVAI calculated from the radiance at 354 and 388 nm; (b) TROPOMI ALH (unit: km); (c) MODIS AOD at 550 nm.



Figuur 3.4: The SSA retrieved by radiative transfer simulations as a function of $\Delta\kappa$ ($\Delta\kappa = (\kappa_{354} - \kappa_{388})/\kappa_{388}$): (a) SSA mean and standard deviation (filled region) of plume pixels; (b) SSA mean and standard deviation (filled region) of the 15 AERONET co-located pixels; (c) absolute difference between the mean SSA of the 15 co-located pixels and the AERONET retrieval.

as a function of $\Delta\kappa$. The retrieved aerosol absorption decreases with $\Delta\kappa$. This finding is in good agreement with [Jethva and Torres, 2011a]. Gray aerosols (spectrally flat) require stronger absorption to reach the same level of UVAI calculated by colored aerosols. This also explains the high SSA standard deviation (filled area) in the cases with little or no spectral dependence on aerosol absorption. The large difference in SSA (from 0.69±0.13 to 0.94±0.03) demonstrates that inappropriate assumptions of the spectral dependence of particle refractive index can significantly bias the result, which should be carefully handled in forward radiative transfer calculations.

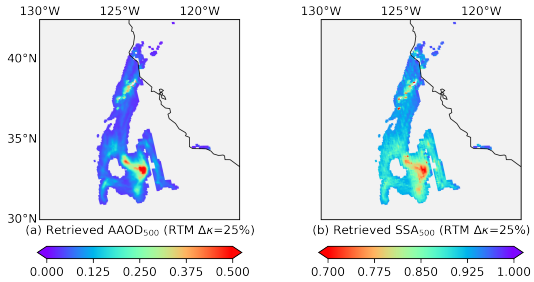
The retrieved SSA is compared with the version 3 level 1.5 AERONET inversion product (https://aeronet.gsfc.nasa.gov, [Holben et al., 1998]). Only one site is within 50 km from the TROPOMI plume pixels (UCSB, located at 119.845°W, 34.415°N) with only one record at 18:54:47 UTC, which is nearly 3 hr ahead of the TROPOMI overpass. The SSA at 500 nm for this case is 0.98 (sky radiance error 15.8 %). There are 15 TROPOMI pixels co-located with UCSB. As illustrated in Figure 3.4 (b), the mean SSA of the AERONET co-located pixels also increases with $\Delta\kappa$ and eventually levels off at 0.96.

The differences between the mean SSA of the co-located pixels and the AERONET are shown in Figure 3.4 (c). The retrieved SSA starts falling inside the uncertainty range of AERONET (±0.03, [Holben et al., 2006]) when $\Delta\kappa$ is 25%, where the plume SSA is 0.90±0.05 and the AERONET co-located SSA is 0.96±0.02. Table 3.2 also presents the SSA provided by OMAERUV version 3 product (https://doi.org/10.5067/Aura/OMI/DATA2004). OMI pixels are co-located with the AERONET site in the same way as TROPOMI. The SSA of OMAERUV is 0.06 lower than that of AERONET, which indicates that a 20% spectral dependence of the aerosol absorption in OMAERUV algorithm may be not sufficient for this case. Although our retrieved SSA seems closer to the AERONET than OMAERUV, one should keep in mind that there is only one record for this event, and the meteorological conditions, combustion phases and even the aerosol compositions may change during the 3 hr time difference.

Figure 3.5 presents the spatial distribution of retrieved AAOD (derived by AAOD = (1-SSA)×AOD) and SSA when $\Delta\kappa$ is 25 %, which shows a strong heterogeneity in the space. The plume center is

Tabel 3.2: The retrieved SSA by radiative transfer simulations for the California fire on 2017-12-12.

Retrieval methods	Number of plume pixels	Retrieved SSA (plume pixels)	$SSA_{max} - SSA_{min}$	Retrieved SSA (co-located pixels)	AERONET SSA	OMAERUV SSA
RTM with $\Delta\kappa$ = 25%	5217	0.90±0.05	0.38	0.95±0.02	0.98	0.92±0.01



Figuur 3.5: The AAOD and SSA retrieved by radiative transfer simulations for the California fire event on 12 December 2017 when $\Delta\kappa$ = 25%: (a) retrieved AAOD at 500 nm (derived by $AAOD = (1 - SSA) \times AOD$); (b) retrieved SSA at 500 nm.

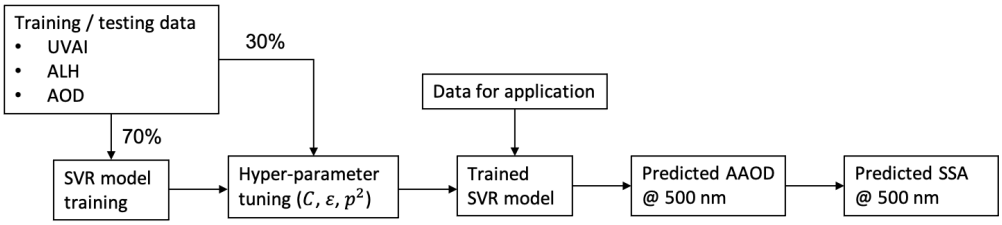
most absorbing where the SSA is even less than 0.70. The SSA gradually increases when the plume is transported northwards. SSA is expected to be low near source flaming regions [Eck et al., 2003, 2013, 1998], whereas it may become higher when aerosols are aging during transport [Lewis et al., 2009, Reid et al., 2005]. The strong spatial variability in SSA is mainly dominated by the heterogeneity of the UVAI (Figure 3.4 (a)) via the pixel-to-pixel numerical relationship. Depending on the combustion phase and meteorological conditions, heterogeneity of the aerosol properties is expected for a plume of this scale. Nevertheless, whether such a large SSA difference of 0.38 is reasonable requires further investigations (will be discussed in Section 3.3.6).

3.3. EXPERIMENT 2: SSA RETRIEVAL USING SUPPORT VECTOR REGRESSION

In this section, we propose an empirical method to derive SSA as an alternative to the RTM-based method in the first experiment. The motivation is that assumptions regarding aerosol micro-physics in forward simulations are inevitable, and our knowledge of them is inadequate (particularly the aerosol absorption spectral dependence). An inappropriate assumption may lead to significant bias in retrieved SSA (Figure 3.4). Conversely, Machine Learning is applicable to solve ill-posed inversion problems by learning the underlying behavior of a system from a given data set without explicitly assuming aerosol micro-physics. In this paper, we construct an SVR model with UVAI, AOD and ALH as input features and AAOD as the output variable, and then derive the SSA using the following relationship:

$$SSA = 1 - \frac{AAOD}{AOD} \tag{3.2}$$

The procedure of SVR retrieval is presented in Figure 3.6. This section starts with a brief introduction of SVR algorithm, followed by input feature selection (Section 3.3.2), training and testing data set preparation (Section 3.3.3), SVR model hyper-parameter tuning (Section 3.3.4), error analysis (Section 3.3.5), and applications to specific cases (Section 3.3.6).



Figuur 3.6: Flowchart of Support Vector Regression (SVR).

3.3.1. SUPPORT VECTOR REGRESSION

SVR is the regression variant of SVM [Drucker et al., 1997], a supervised non-parametric statistical algorithm initially devised by Cortes and Vapnik [1995]. SVM is suitable for solving problems of small training data sets with high-dimensional feature spaces and it can provide excellent generalization performance [Durbha et al., 2007, Yao et al., 2008]. It has been extensively applied to solve remote sensing problems [Di Noia and Hasekamp, 2018, Lary et al., 2009, Mountrakis et al., 2011]. The basic idea of SVM for solving classification problems is finding an optimal hyperplane in a high-dimensional feature space that maximizes the margin between the two classes to minimize mis-classification [Durbha et al., 2007]. The same principle is applied to regression problems, where SVR attempts to find an optimal hyper-plane that maximizes the margin of tolerance in order to minimize the prediction error. The error within the margin does not contribute to the total loss function. The samples on the margin are called support vectors. For the detailed mathematical formulation of SVR one can refer to [Smola and Schölkopf, 2004]. Briefly, given the training data with n observations $(x_1, y_1), (x_2, y_2), ..., (x_n, y_n)$, the statistical model is assumed to be as follows:

$$y = r(x) + \delta \tag{3.3}$$

where x is a multivariate input and y is a scalar output and both are with length n; δ is the independent zero mean random noise. The input x is first mapped onto a feature space with dimension of m by a nonlinear transformation, and then a linear model $f(x)$ is constructed:

$$f(x) = \sum_{j+1}^{n} \omega_j g_j(x) + b \tag{3.4}$$

where the $g_j(x)$ is the nonlinear transformation, ω_j is the model parameter vector and b is the bias. SVR tries to find an optimal model from a set of approximate functions $f(x)$. An approximate function is assessed by a loss function (also called ϵ-insensitive loss in SVR), defined as:

$$L\big(y, f(x)\big) = \begin{cases} 0, & |y - f(x)| \le \epsilon \\ |y - f(x)| - \epsilon, & otherwise \end{cases} \tag{3.5}$$

then the total empirical risk is:

$$R(\omega) = \frac{1}{n} \sum_{i+1}^{n} L\big(y_i, f(x_i)\big) \tag{3.6}$$

SVR performs linear regression in a high-dimension feature space using ϵ-insensitive loss, and reduces the model complexity by minimizing the norm $\|\omega\|^2$. By introducing non-negative slack variables (ξ_i and ξ_i^*) to measure the deviations of errors outside ϵ, SVR problems can be formulated as

below:

$$minimize \frac{1}{2}\|\omega\|^2 + C\sum_{i+1}^{n}(\xi_i + \xi_i^*)$$
$$s.t. \begin{cases} y_i - f(x_i) \leq & \epsilon + \xi_i^* \\ f(x_i) - y_i \leq & \epsilon + \xi_i \\ \xi_i^*, \xi_i \geq & 0 \end{cases} \tag{3.7}$$

where C is a positive regularization constant determining the trade-off between model complexity and the degree to which deviations larger than ϵ are penalized. The optimization problem can be transferred into a dual problem by introducing the Lagrange multipliers (α_i and α_i^*) and thus the solution becomes:

$$f(x) = \sum_{i=1}^{n}(\alpha_i - \alpha_i^*)K(x_i, x) + b$$
$$s.t. 0 \leq \alpha_i, \alpha_i^* \leq C \tag{3.8}$$

where $K(x_i, x)$ is the kernel function that is positively semi-definite in order to satisfy the Mercer's theorem. The kernel function enables the SVR to solve nonlinear problems.

According to the description above, we know that SVR generalization performance and estimation accuracy depend on the regularization constant C, the width of the tolerance margin ϵ and the kernel function $K(x_i, x)$. We will discuss how to determine the three hyper-parameters in Section 3.3.4.

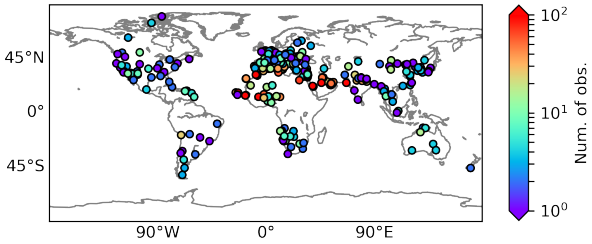
3.3.2. FEATURE SELECTION BASED ON OMI AND AERONET OBSERVATIONS

Although SVR is able to deal with high-dimensional input features, feature selection is still important for generalization performance, computational efficiency and interpretability issues [Weston et al., 2001]. Many sophisticated approaches have been devised for feature selection [Guyon and Elisseeff, 2003]. In this study, we choose features based on our empirical knowledge of UVAI and the Spearman's rank correlation coefficients (ρ).

CO-LOCATING OMI AND AERONET

The feature selection is based on the co-located OMAERUV version 3 product (https://doi.org/10.5067/Aura/OMI/DATA2004) and AERONET version 3 level 1.5 inversion product (https://aeronet.gsfc.nasa.gov). OMAERUV is currently the only satellite product containing a long-term record of UVAI, AOD and SSA [Torres and Chen, 2013, Torres et al., 2007]. Its AOD was validated by the multiyear AERONET record [Ahn et al., 2014], and its SSA was evaluated by AERONET retrievals [Jethva et al., 2014]. The OMAERUV also provided ALH information, which is a best-guess from either of pre-assumed values [Torres and Chen, 2013], for example, for retrieved carbonaceous particles (CB), if its UVAI smaller than 0.5, then its height is set to be 1.5 km, otherwise, CALIOP ALH climatology will be used (if available).

We collect the measurements of OMAERUV and AERONET from 1 January 2005 to 31 December 2017. OMI pixels with $\theta_0 > 70°$ or CF > 0.1 are excluded. An OMI pixel is co-located to an AERONET site if their spatial distance is within 50 km and their temporal difference is within 3 hr. To ensure consistency between the different measurement techniques (ground-based versus space-borne), we also exclude samples if the SSA difference between OMAERUV and AERONET is larger than 0.03, or the AOD difference between OMAERUV and AERONET is larger than 5%. The AERONET SSA and AAOD are linearly interpolated to 500 nm in order to compare with OMAERUV retrievals. In total, there are 5,679 samples obtained. Figure 3.7 shows the global distribution of the co-located OMAERUV–AERONET samples. Note that these samples are not limited to biomass burning areas, but also contain other aerosol types.



Figuur 3.7: Global distribution of the OMAERUV-AERONET joint data set. The color indicates the number of observations.

FEATURE SELECTION

The OMAERUV–AERONET joint data set consists of the following selected parameters: UVAI, geometries (θ_0, θ and $\Delta\phi = \phi - \phi_0 + 180°$), A_s, P_s and ALH from OMAERUV, and SSA, AOD and AAOD from AERONET. Note that the UVAI used here is the 'residue' term in OMAERUV product, where the simulated radiance (I_λ^{Ray} in Equation 3.1) is calculated by a simple Lambertian approximation that is also used for the TROPOMI UVAI [Torres et al., 2018].

Figure 3.8 presents the Spearman's rank correlation coefficients matrix (ρ) of those parameters. It is clear that except for AAOD, SSA is barely associated with other parameters. The correlation between UVAI and SSA is rather low (ρ = -0.25). Conversely, AAOD is highly correlated with UVAI (ρ = 0.66) as well as AOD (ρ = 0.75), as AAOD carries the information of both aerosol absorption and aerosol loading. Therefore, it is preferred to predict AAOD from given UVAI and to derive SSA via Equation 3.2 rather than directly predict SSA from UVAI. Furthermore, as mentioned previously, AOD and ALH are the major factors influencing UVAI, which is also reflected by the relatively stronger correlation (ρ = 0.66 and 0.4, respectively). Consequently, we construct an SVR model with UVAI, ALH and AOD as the input features, and AAOD as the target variable to be predicted. UVAI also shows a relatively high dependence on θ_0 (ρ = -0.5), but in this study, we only focus on the aerosol-related features. Figure 3.9 presents the basic statistics of the relevant parameters.

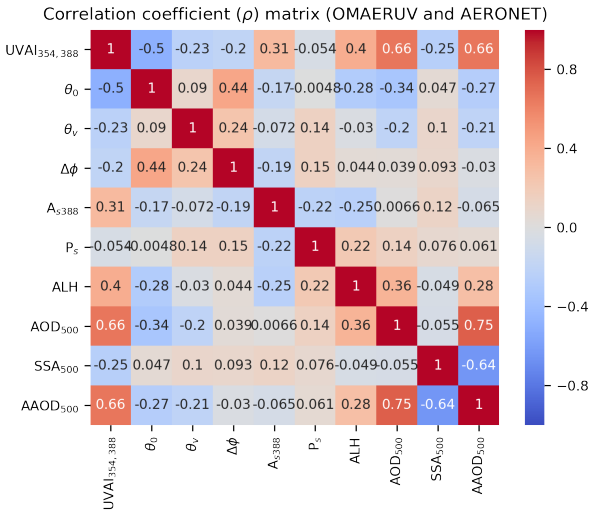
3.3.3. PREPARING TRAINING AND TESTING DATA SETS

The SVR model is trained and tested based on the OMAERUV–AERONET joint data set that contains 5,679 samples. We further partition it into a training data set and a testing data set. The testing data set is used to evaluate the generalization performance of the trained Machine Learning model in order to avoid high bias (under-fitting) or high variance (over-fitting) problems. For a small data set, the empirical ratio between the training data set and the testing data set is 70% versus 30%; thus, there are 3,975 samples in the training data set and 1,704 samples in the testing data set.

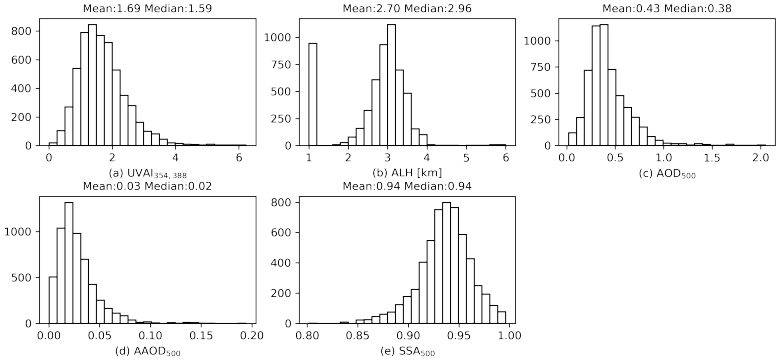
3.3.4. SVR HYPER-PARAMETERS TUNING

As described in Section 3.3.1, the generalization performance and model accuracy of SVR depends on the following hyper-parameters: (1) the width of the insensitive zone ϵ. The cost function does not consider errors in the training data as long as their deviation from the truth is smaller than ϵ; (2) the regularization constant C that determines the trade-off between model complexity and the degree to which deviations larger than ϵ are penalized; and (3) the choice of the kernel and the corresponding parameters. Here we adopt the methodology from [Cherkassky and Ma, 2004], where the SVR parameter C and ϵ can be directly determined from the statistics of the training data set:

$$C = max\left(|\bar{y} + 3\sigma_y|, |\bar{y} - 3\sigma_y|\right) \tag{3.9}$$



Figuur 3.8: Spearman's rank correlation coefficient matrix (ρ) of parameters from the OMAERUV-AERONET joint data set.



Figuur 3.9: Statistics of the selected features and target variables: (a) OMAERUV UVAI calculated by radiance at 354/388 nm; (b) OMAERUV ALH (unit: km); (c) AERONET AOD at 500 nm; (d) AERONET AAOD at 500 nm; (e) AERONET SSA at 500 nm.

$$\epsilon = 3\sigma\sqrt{\frac{ln(n)}{n}} \tag{3.10}$$

where $\bar{y}$ and σ_y are the mean and standard deviation of the output parameter in the training data set; σ is the input noise level (we set it to 0.001); and n is the number of training samples. The values determined for C and ϵ are shown in Table 3.3.

We employ the widely used radial basis function (RBF) kernel function to solve the non-linearity in the SVR model. Compared with other kernel functions, RBF is relatively less complex and more efficient. The RBF kernel is defined as:

$$K(x_i, x) = exp\left(-\frac{\|x_i - x\|^2}{2p^2}\right) \tag{3.11}$$

Tabel 3.3: Values for hyper-parameter: the regularization constant C, the width of the insensitive zone ϵ and the BRF kernel parameter p^2.

Parameters	C	ϵ	p^2
Values	0.09	0.0001	1.67

where p is the kernel width parameter that reflects the influencing area of the support vectors, which is determined by hyper-parameter tuning on the testing data set [Durbha et al., 2007].

The RMSE of the training process may overestimate the accuracy of a SVR model if the validation is based on the same data set. Instead, an independent testing data set is used to represent the accuracy of the SVR model. The difference of the model accuracy between training and testing process reflects the generalization performance of the SVR model. An ideal SVR model should output a low-level RMSE and a small discrepancy between the RMSE of the training and testing processes. If the RMSE of the testing process is much larger than that of the training process, the SVR may suffer from over-fitting problems. Figure 3.10 shows the hyper-parameter tuning process of the RBF kernel parameter p^2. It is clear that when $p^2 = 1.67$, the RMSE of the training process is relatively small, and so is the model accuracy difference between the training process and the testing process. The final values of C, ϵ and p that will be applied in the case studies are listed in Table 3.3. The corresponding RMSE of AAOD values predicted by the training process and the testing process are at a level of 0.01 (Figure 3.11 (a)).

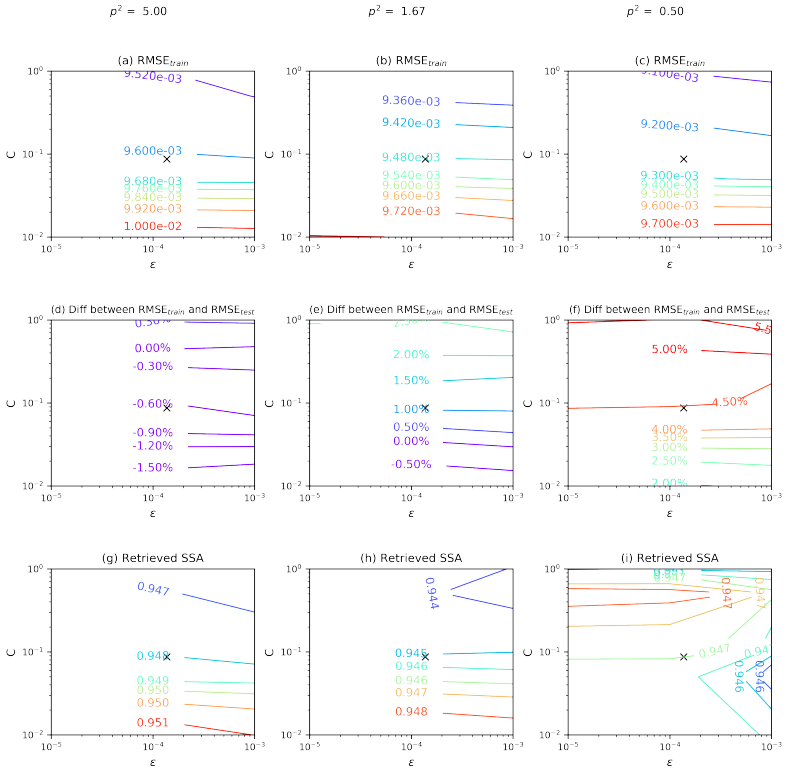
3.3.5. ERROR ANALYSIS

The quality of the SVR-predicted SSA depends on the model accuracy as well as the quality of input data. The model accuracy can be represented by the RMSE of the testing process (0.01). As shown in Figure 3.11 (a), the SVR model has difficult predicting samples with AAOD values larger than 0.05, reflected by significant biases at this range. The uncertainty in AAOD is passed to the SSA via Equation 3.2. Figure 3.11 (b) shows the retrieved SSA in the training and testing processes. It is noted that the predicted SSA is generally positively biased, particularly in strong absorption cases (SSA < 0.90). This is possibly due to the bias in the feature domain, where the UVAI is relatively strongly correlated with other factors (i.e., AOD and ALH) that may contain redundant information which adversely impacts model performance [Durbha et al., 2007, Weston et al., 2001]. In future, a more sophisticated feature selection scheme is suggested to reduce the redundancy, e.g. the Minimum Redundancy Maximum Relevance (mRMR, [Hanchuan Peng et al., 2005]). Moreover, the RBF kernel function may not capable enough to solve the non-linearity between the features. The accuracy of SSA predicted by the testing data set is ±0.02, with 82% of samples falling into the uncertainty range (±0.03) of the AERONET SSA.

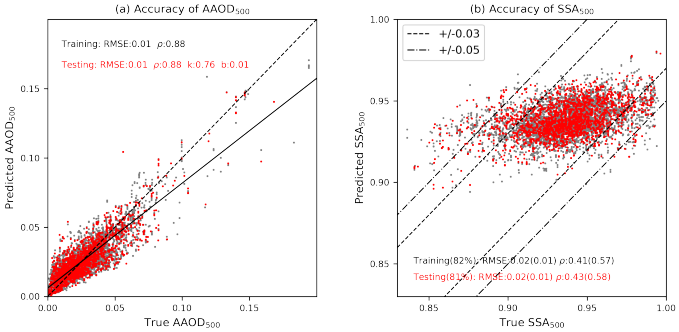
The error of the retrieved SSA due to the quality of the input data may also come from the observational or retrieval uncertainties in input features. The SSA sensitivity to the errors in input data is presented in Figure 3.12. We use the mean value of each parameter in the OMAERUV–AERONET data set as reference values (Figure 3.9, UVAI = 1.59, ALH = 2.96 km, AOD = 0.39), and the corresponding SSA value is 0.94. The positive bias of UVAI always leads to an underestimation of SSA, unless the aerosol layer is located at a relatively high altitude or there is little aerosols. Conversely, the small UVAI causes the overestimation of SSA, except for cases where the ALH is low or the AOD is high. The sensitivity of SSA to UVAI is weaker when the aerosol layer is close to the surface or at a very high altitude, and the sensitivity of SSA to UVAI always increases with AOD.

3.3.6. CASE APPLICATIONS

Once the hyper-parameters are determined (Section 3.3.4), the trained SVR model is ready to predict aerosol absorption. We apply it to the California fire event in December 2017, the same as that in the first experiment. To demonstrate the generalization capability of the SVR model, we also apply it to other fire events as long as there are co-located TROPOMI, MODIS and AERONET-retrieved AAOD/SSA

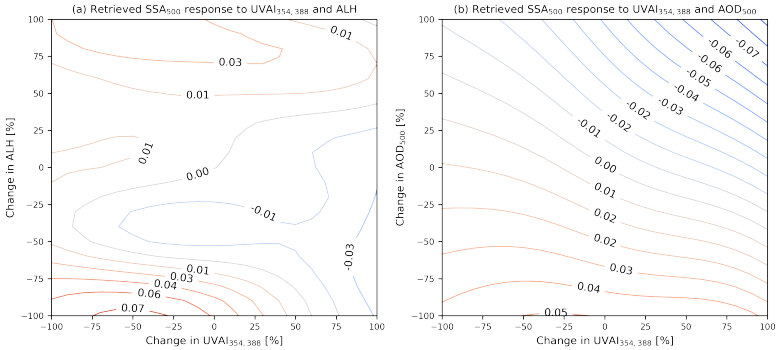


Figuur 3.10: The performance of the SVR model as a function of hyper-parameters (C, ϵ and p^2). The cross markers represent the values of C and ϵ according to [Cherkassky and Ma, 2004]. A p^2 value equal to 1.67 is sufficient to obtain a relatively high accuracy meanwhile preventing over-fitting problem.



Figuur 3.11: The accuracy of the trained SVR model: (a) the predicted AAOD at 500 nm against the truth. The dashed line is the 1:1 line, and the solid line is the linear fitting for the testing data set; (b) the predicted SSA at 500 nm against the truth. Gray and red indicate samples in training and testing data sets, respectively. The values in parentheses are the statistics for samples that fall within the AERONET SSA uncertainty of ±0.03.

available for comparisons.



Figuur 3.12: The sensitivity of the SVR retrieved SSA: (a) SSA at 500 nm as a function of changes in UVAI and ALH; (b) SSA at 500 nm as a function of changes in UVAI and AOD.

For all applications, the input parameters in the trained SVR model are the TROPOMI UVAI (calculated using the 354/388 nm), the TROPOMI ALH and the MODIS AOD, respectively. The MODIS AOD at 550 nm is converted to 500 nm using α provided by the co-located AERONET site. Note that the data include pixels with a CF>0.1 in order to ensure that there are satellite measurements co-located with the AERONET sites.

CALIFORNIA FIRE EVENT ON 12 DECEMBER 2017

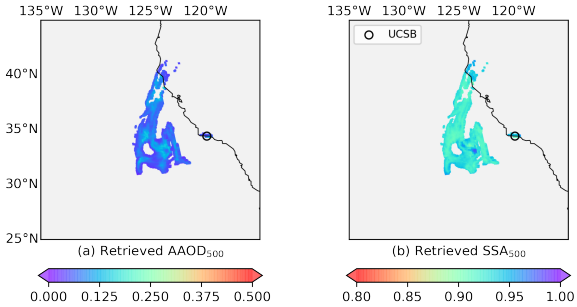
Figure 3.13 presents the retrieved AAOD and corresponding SSA. Recall that that UVAI and AOD are higher in the center of the plume, whereas ALH is relatively lower (Figure 3.3), the retrieved SSA should be smaller to compensate for the low altitude of the aerosol layer according to Figure 3.12. However, the SVR retrieved SSA is even higher than its surroundings. This is because the UVAI and AOD values at this region are outside the distribution of training data set, as shown in Figure 3.9. The OMAERUV–AERONET joint data cannot cover these extreme situations. Consequently, the SVR model fails to handle the input values outside of the range of the training data set. The reason for this may be that the joint data set is relatively small as a result of the limited data availability and the strict co-location criteria, or that the quality of the joint data suffers from observational or retrieval uncertainties.

The mean SSA of the plume is 0.94±0.01 (including the failed pixel predictions) and that for the AERONET co-located pixels is 0.97±0.01 (Table 3.4). The SVR-predicted SSA may be overestimated, whereas the standard deviation may be underestimated due to the SVR prediction failures of some samples. The difference between the mean SSA of validation pixels (co-located to the AERONET site) and the AERONET retrieval is only 0.01, which is within the typical uncertainty range of AERONET (±0.03).

OTHER CASE APPLICATIONS

To present the generalization performance of the SVR model, we apply it to other fire events as long as there is co-located data of TROPOMI, MODIS and AERONET. The same pre-processing is applied as in the previous case in order to exclude pixels with UVAI<1, AOD<0.5 or CF>0.3.

Figures 3.2 (b-d) present the California fire events during the period from 9 to 11 November 2018. The plumes were over ocean but were partly contaminated by the underlying clouds. Figures 3.2 (e) shows the Canadian fire event on 29 May 2019. This case was over land, which means that the brighter surface may cause a higher bias in the input AOD and ALH than cases over dark surfaces [Remer et al., 2005, Sanders et al., 2015].



Figuur 3.13: SVR prediction for California fire event on 2017-12-12: (a) the predicted AAOD at 500 nm; (b) the predicted SSA at 500 nm.

Tabel 3.4: SVR predicted SSA for different cases. If there is no standard deviation, then it indicates there is only one record.

Case	Number of Plume pixels	Retrieved SSA (plume pixels)	SSA_{max} – SSA_{min}	Co-located AERONET	SSA (co-located-pixels)	AERONET SSA	OMAERUV SSA
2017/12/12	5217	0.94±0.01	0.09	UCSB	0.97±0.01	0.98	0.92±0.01
2018/11/9	1944	0.94±0.01	0.1	Santa_Monica_Colg	0.93±0.01	0.89±0.06	0.89±0.06
2018/11/10	2184	0.94±0.02	0.1	CalTech	0.96±0.01	0.89±0.07	-
				Fresno_2	0.93±0.02	0.91±0.01	-
				Modesto	0.94±0.01	0.92±0.01	0.96±0.01
				USC_SEAPRISM_2	0.93±0.00	0.9	0
2018/11/11	2815	0.95±0.02	0.09	Modesto	0.98±0.00	0.96±0.01	0.95±0.00
2019/5/29	8013	0.97±0.02	0.10	Fort_McKay	0.97±0.02	0.95±0.00	0.93
			0.1	Fort_McMurray	0.98±0.01	0.93	1.00

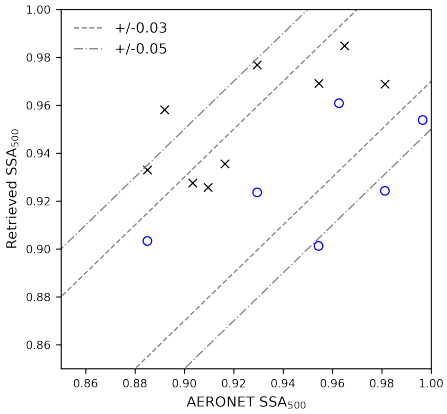
The predicted SSA for the above mentioned events is listed in Table 3.4. Similar to the California case on 12 December 2017, the SVR fails to predict reasonable SSA for pixels if input features fall outside the histogram in of the training data set (Figure 3.9), causing positive biases in the plume mean SSA. The plume SSA of two California fire events are similar, with values of around 0.94–0.95. The predicted SSA for the Canadian fire is relatively higher (0.97).

We further plot the predicted SSA against the co-located AERONET records (black crosses in Figure 3.14). There are 9 co-located records obtained. The difference between SVR and AERONET retrieved SSA are almost within ±0.05, among which over half (5 out of 9) fall within the AERONET SSA uncertainty range (±0.03). We also provide SSA from OMAERUV for these cases (Table 3.4 (blue circles in Figure 3.14). Compared with OMAERUV, the SSA predicted by SVR shows a better consistency with AERONET, although one should keep in mind that the accuracy of SVR model is ±0.02 and the model tends to overestimate the SSA for relatively stronger absorbing cases.

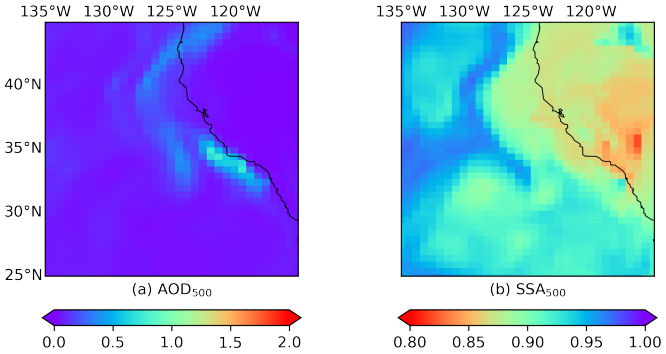
SPATIAL VARIABILITY OF RETRIEVED SSA

Compared with Figure 3.5 (b), the spatial variability of the SVR-predicted SSA is smaller: the difference between maximum and minimum SSA is at level of around 0.10 (Table 3.4). In the first experiment, SSA is determined by UVAI of each pixel individually. In the SVR model, the spatial variability of the intermediate output AAOD depends on 3 input features. Furthermore, SVR predicts AAOD for each pixel based on the common relationship between UVAI, AOD and ALH in the training data set.

Heterogeneity in aerosol properties is expected for a plume of this size. However, the degree of this heterogeneity requires further investigations. Here we assess the SSA spatial variability of



Figuur 3.14: SVR-retrieved SSA (black cross) and OMAERUV-retrieved SSA (blue circle) against AERONET SSA at 500 *nm* for all 5 cases in this study.



Figuur 3.15: MERRA-2 M2T1NXAER averaged between 12:00 and 15:00 local time for the California fire event on 2017-12-12: (a) AOD at 500 *nm*; (b) SSA at 500 *nm*

an independent data set. We employ the SSA from the MERRA-2 aerosol reanalysis hourly single-level product (https://disc.gsfc.nasa.gov/datacollection/M2T1NXAER_5.12.4.htm). The aerosol properties of MERRA-2 have proved to be in good agreement with independent measurements [Buchard et al., 2017, Randles et al., 2017]. MERRA-2 AOD and SSA of the California fire on 12 December 2017 are shown in Figure 3.15. The plume can be recognized by high AOD against the background value. Although the plume shape presented by satellite observations differs from that of model simulations, the SSA spatial difference within the plume is at an approximate magnitude of 0.1. We thus find that the spatial variability of SSA retrieved by the SVR model is in better agreement with that of MERRA-2. The same conclusion can be drawn for other cases (not shown).

3.4. CONCLUSION

The long-term record of global UVAI data is a treasure for deriving aerosol optical properties such as SSA. To quantify aerosol absorption from UVAI, information on AOD and ALH is necessary. Currently various AOD products are available, whereas operational ALH products are much less accessible.

In the first experiment, we derive the SSA by forward radiative transfer simulating UVAI for a

fire event in California on 12 December 2017. Using the TROPOMI ALH, we are able to quantify the influence of assumed spectral dependence of near-UV aerosol absorption (represented by the relative difference between κ_{354} and κ_{388}) on the retrieved SSA. A significant gap in plume mean SSA (0.25) between the spectrally flat aerosols ($\Delta\kappa$ = 0%) and aerosols with strong spectral dependence($\Delta\kappa$ = 40%) implies that inappropriate assumptions regarding spectral dependence may significantly bias the retrieved aerosol absorption. The SSA difference becomes smaller than the uncertainty of AERONET (±0.03) when $\Delta\kappa$ = 25%. The corresponding plume SSA is 0.90±0.05.

In the second part of this study, we propose a statistical method based on the long-term joint data of UVAI, AOD, ALH and AAOD using a SVR algorithm, in order to avoid making the assumptions of the aerosol absorption spectral dependence in the near-UV band. Considering all applications, the results are encouraging: the SSA discrepancy between retrievals and AERONET for almost all co-located samples is within ±0.05, and over half of them fall within the AERONET uncertainty range (±0.03). However, one should keep in mind that the SVR model tends to overestimate the SSA for strongly absorbing cases (e.g. SSA < 0.90), and fails to predict reasonable SSA when input values fall outside the distribution of the training data set.

In terms of spatial variability, the SSA derived by radiative transfer simulations significantly differs from that predicted by SVR. We employ the MERRA-2 SSA as an independent reference, and the spatial difference of this data within smoke plume is at a magnitude of approximately 0.1. The spatial pattern of the SVR-predicted SSA shows better agreement with this finding.

In this study, we present the potential to retrieve SSA using a Machine Learning method. The motivation is to avoid a priori assumptions on aerosol micro-physics such as those made in the radiative transfer simulations. In the current phase, we choose SVR because the size of the training data set is relatively small (less than 5,000). The input features are selected by the Spearman's rank correlation coefficients and a priori knowledge on the relationship between UVAI and aerosol-related features. The model hyper-parameters are analytically determined. The accuracy of SVR model predictions is ±0.02, with a tendency to overestimate the SSA for highly absorbing cases. The OMAERUV–AERONET data set cannot cover some extreme situations. As a result, prediction failures appear when the input values fall outside the distribution of the corresponding parameters in the training data set. In the future, more sophisticated feature selection techniques should be used to improve the prediction. Other non-aerosol features affecting UVAI could also be taken into consideration. Moreover, we use the OMAERUV ALH to train the SVR model, whereas it is not retrieved from measurements but a best-guess from model and a priori assumptions. The high-resolution TROPOMI level 2 ALH product brings the potential to increase the data availability and improve the quality of the training data set, which allows to use to more powerful algorithms (e.g. ANN) in this kind of applications.

4

TOWARDS A GLOBAL AEROSOL HEIGHT DATABASE FOR SATELLITE AEROSOL INDEX ANALYSIS

The content of this chapter has been adapted from **Sun, J.**, Veefkind, J. P., van Velthoven, P., Tilstra, L. G., Chimot, J., Nanda, S., and Levelt, P. F.: Defining aerosol layer height for UVAI interpretation using aerosol vertical distributions characterized by MERRA-2, Atmos. Chem. Phys. Discuss. [preprint], https://doi.org/10.5194/acp-2020-39, 2020.

4.1. INTRODUCTION

The satellite Ultra-Violet Aerosol Index (UVAI) is a qualitative measure of aerosol absorption which has a long-term global record since 1978 [Herman et al., 1997a]. It describes the change of the radiance contrast between two UV channels ($\lambda_1 < \lambda_2$) due to the presence of absorbing aerosols:

$$UVAI = -100 \left[log_{10} \left(\frac{I_{\lambda_1}}{I_{\lambda_2}} \right)^{obs} - log_{10} \left(\frac{I_{\lambda_1}}{I_{\lambda_2}} \right)^{Ray} \right] \tag{4.1}$$

where obs indicates the satellite measured radiance in a real atmosphere, and Ray indicates the simulated radiance in a Rayleigh atmosphere. A positive UVAI value implies the presence of absorbing aerosols, whereas a neutral or negative value implies the non-absorbing atmospheric components. UVAI has been widely used for absorbing aerosol detection (e.g. [Alpert and Ganor, 2001, Chiapello et al., 1999, Duncan et al., 2003a,b, El-Askary et al., 2006, Ginoux and Torres, 2003, Hsu et al., 1999, Israelevich et al., 2002, Krotkov et al., 1999, Krueger et al., 2008, Moulin and Chiapello, 2004, Seftor et al., 1997, Spichtinger et al., 2001, Wang et al., 2008b, Yuan et al., 2011]), aerosol classification (e.g. [Koukouli et al., 2006, Lee et al., 2007, Torres and Chen, 2013]), and chemistry transport model validation (e.g. [Buchard et al., 2015, 2017, Chiapello et al., 1999, Hammer et al., 2016]). It is also possible to derive quantitative aerosol absorptive properties (e.g. SSA and AAOD) from UVAI, if both the aerosol optical depth (AOD) and the aerosol vertical distribution are known (e.g. [Hu et al., 2007, Jeong and Hsu, 2008, Sun et al., 2018, 2019]). This kind of applications is meaningful to reduce uncertainties of aerosol radiative forcing assessments [Loeb and Su, 2010, McComiskey et al., 2008, Penner et al., 2001], as the aerosol absorption is a key parameter but poorly understood due to limited observations.

UVAI has a long-term global record since 1978, as listed in Table 1.1 (page 24). Correspondingly, AOD products are also contributed by various satellites covering a long time-span (e.g. ATSR families, MODIS/Terra and Aqua since 1999/2002, MISR/Terra since 2002, OMI/Aura since 2004, SLSTR/Sentinel 3 since 2016, etc.) [Sogacheva et al., 2020], and there are also ground-based systems for AOD validation and synergism (e.g. AERONET). However, observations of the aerosol vertical distribution do not have such a spatial and temporal coverage comparable to that of AOD and UVAI. Currently, the observational aerosol vertical distribution is mainly offered by two types of measurements [Xu et al., 2018b]. The first is aerosol profiles retrieved from the backscattered signal of either space-borne lidar (e.g. [Flamant et al., 2008, Schutz et al., 2005, Winker et al., 2009]) or ground-based lidar networks (e.g. [Pappalardo et al., 2014, Welton et al., 2001]). Another is aerosol layer heights (ALH) retrieved from columnar radiation measurements of passive sensors. There are various techniques developed, including stereo photogrammetry (e.g. [Nelson et al., 2013, Virtanen et al., 2014]), UV polarization (e.g. [Dubovik et al., 2011, Kokhanovsky et al., 2015, Wu et al., 2016a]), dust properties in the infrared-thermal band (e.g. [Pierangelo et al., 2004, Vandenbussche et al., 2013]), and oxygen absorption spectroscopy (e.g. [Boesche et al., 2009, Chimot et al., 2017, Ding et al., 2016, Dubuisson et al., 2009, Duforêt et al., 2007, Nanda et al., 2019, Sanders et al., 2015, Sanders and De Haan, 2016, Sanghavi et al., 2012, Tilstra and Stammes, 2019, Wang et al., 2014, Xu et al., 2017]).

Whereas lidar measurements provide a great amount of details on aerosol profiles, they are subjected to limited spatial and/or temporal coverage. Compared with aerosol profiles, aerosol heights are easier for quantitative analysis, but most aerosol height retrieval algorithms are only applicable under certain conditions, e.g. elevated aerosol layers, over dark surfaces, exclusively for dust aerosols, cloud-free, etc. As a result, in order to derive global aerosol absorptive properties from UVAI on a daily basis, a corresponding observational aerosol vertical distribution data is currently still absent.

The purpose of this study is to break the limitation posed by observations. We will construct a global daily aerosol height database with help of the aerosol profiles provided by a chemistry transport model. We choose the MERRA-2 aerosol reanalysis, whose aerosol fields are generally in good agreement with independent observations [Buchard et al., 2017, Randles et al., 2017]. In the next section, we first present the UVAI dependence on the aerosol height using radiative transfer simulations, and provide the evidence found in satellite observations (Section 4.2). The results of this section are used to the

evaluate the aerosol heights derived from the chemistry transport model. In Section 4.3, after a brief introduction of the MERRA-2 aerosol reanalysis, we propose several methods that derive the aerosol height from the MERRA-2 aerosol profiles, and inspect the relation between the derived aerosol heights and UVAI based on the findings in Section 4.2. Section 5.4 summaries the conclusions and outlooks for future applications.

4.2. UVAI DEPENDENCE ON AEROSOL HEIGHT

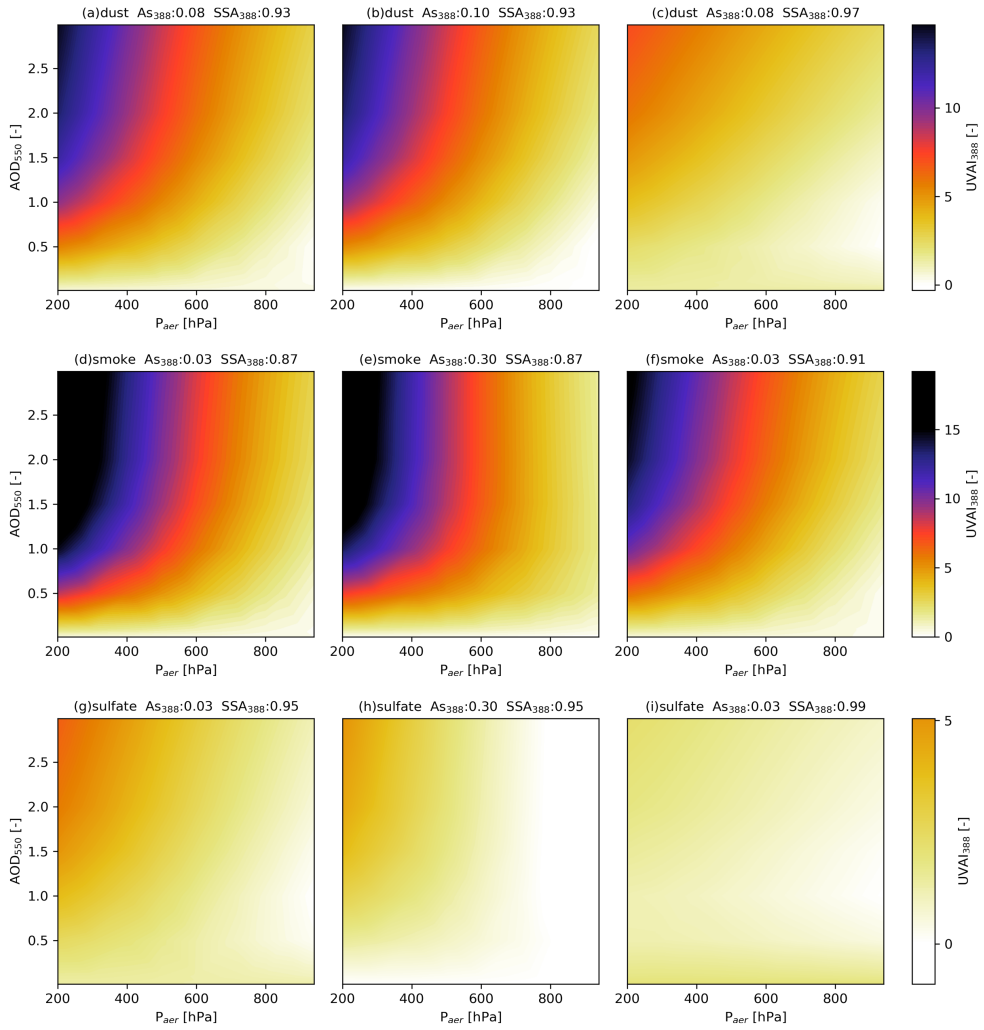
4.2.1. RADIATIVE TRANSFER SIMULATIONS

The UVAI dependence on the aerosol height is well-studied [de Graaf et al., 2005, Herman et al., 1997a, Hsu et al., 1999, Sun et al., 2018, Torres et al., 1998]. In general, UVAI is positively related to the aerosol height, but their relation also depends on AOD, aerosol absorption and surface brightness. Here, we set up a sensitivity study to recall how UVAI changes with the aerosol height under different situations by the radiative transfer model DISAMAR [de Haan, 2011]. In the forward simulations, the aerosol types are characterized by the Mie Theory with aerosol size distribution and refractive index taken from the OMAERUV algorithm [Jethva and Torres, 2011a]. The aerosol profile is parameterized as a homogeneous layer with a constant depth of 50 hPa. The input AOD at 550 nm is set to vary between 0.01 and 3, and the aerosol height is set to vary between 0.5 and 12 km. The wavelength dependency of surface albedo (A_s) is specified by the OMI level 3 Lambertian equivalent reflectance (LER) monthly climatology calculated from measurements between 2005 and 2009 (https://doi.org/10.5067/Aura/OMI/DATA3006, [Kleipool et al., 2008]). Other inputs for radiative transfer simulations are listed in Table 4.1. We design several scenarios by choosing either absorbing (lower SSA) or less-absorbing/scattering (higher SSA) aerosols over either a dark/bright surface.

Figure 4.1 shows the UVAI as a function of the aerosol height under different AOD and A_s for dust (first row), smoke (second row), and sulfate (third row) aerosols from the OMAERUV algorithm. When aerosols are located at higher altitude, more photons scattered by the air molecules below the layer are either absorbed or shielded by aerosols. With reduced radiation measured by a satellite, the change in the radiance contrast becomes larger, resulting in a larger UVAI. The higher the aerosol loading, the stronger UVAI dependence on ALH. However, this dependence becomes weaker over brighter surfaces, particularly when the aerosol loading is small (Figure 4.1 (b, e and h)). On the other hand, a weaker altitude dependence is found for less-absorbing (dust and smoke with higher SSA, Figure 4.1 (c, f)) and scattering aerosols (Figure 4.1 (g, h and i)). The above findings are in agreement with previous studies, although some of them using different aerosol profile parameterizations de Graaf et al. [2005], Herman et al. [1997a], Hsu et al. [1999], Sun et al. [2018], Torres et al. [1998]. As the influence of A_s is negligible, we will not include it in the further analysis.

Tabel 4.1: Configurations of UVAI radiative transfer simulations by DISAMAR.

Parameter	Value
Aerosol optical thickness at 550 nm (AOD_{550})	From 0.01 to 3
Aerosol layer pressure [hPa]	From 200 to 950 hPa (approximately 0.5 to 12 km)
Aerosol layer thickness [hPa]	50
Solar zenith angle (SZA) [°]	30
Viewing zenith angle (VZA) [°]	0
Relative azimuth angle (RAA) [°]	120
Surface pressure (P_s) [hPa]	1013
Surface albedo at 388 nm ($A_{s,388}$)	0.03, 0.3



Figuur 4.1: Sensitivity study of UVAI dependence on ALH for dust (a-c), smoke (d-f) and sulfate (g-i) aerosols with different A_s and SSA.

4.2.2. EVIDENCE FROM SATELLITE OBSERVATIONS

The above UVAI altitude dependence is also found in satellite observations. For example, Xu et al. [2019] showed the relation of UVAI, ALH and AOD retrieved from Earth Polychromatic Imaging Camera onboard DISCOVR for smoke aerosols, where an increasing correlation between UVAI and ALH with growing AOD was found. Here, we present the evidence retrieved from TROPOMI and GOME-2, as they can provide both UVAI and ALH products.

TROPOMI O_2 A-BAND AEROSOL LAYER HEIGHT

The TROPOMI O_2 A-band ALH is developed to detect vertically localized aerosol layers in the free troposphere, e.g. dust storms, biomass burning, and volcanic plumes [Sanders and De Haan, 2016]. The ALH is based on the O_2 absorption in the near-infrared (758-770 nm). A higher O_2 absorption implies

a longer absorption light path above an aerosol layer, i.e. a lower aerosol height [Sanders et al., 2015]. The operational TROPOMI ALH is produced by a Neural Network model for higher computational efficiency [Nanda et al., 2019]. The training data is prepared by radiative transfer simulations for a fixed Henyey-Greenstein aerosol model (SSA, the Ångström Exponent and the asymmetry factor at 550 nm are 0.95, 0, and 0.7, respectively) under different environmental and measurement conditions. The validation study shows that the CALIOP ALH is generally higher than TROPOMI by approximately 1 km over ocean, whereas a larger difference (2.4 km) is found over land [Nanda et al., 2020].

In this study, we collect the TROPOMI level 2 offline ALH product from 2018-11-03 to 2019-08-31 (https://s5phub.copernicus.eu/dhus/#/home). We only retain pixels with successful retrievals (Processing Quality Flags = 0) and full quality data (qa_value = 1). This automatically excludes pixels affected by sun-glint, clouds, bright surfaces (snow and ice) and UVAI (calculated by 354-388 nm wavelength pair) smaller than 1. Furthermore, to avoid the Neural Network extrapolation, samples are kept only if they fall within the distribution of the training data as described in Nanda et al. [2019]: AOD at 550 nm is between 0.05 and 5, SZA is between 8.2° and 70°, P_s is between 520 and 1048.5 hPa, A_s is (at 550 nm) smaller than 0.7, and the retrieved ALP is between 75 and 1000 hPa.

GOME-2 ABSORBING AEROSOL HEIGHT

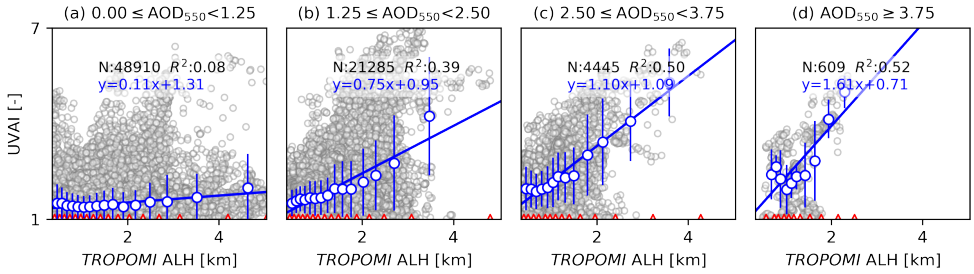
The GOME-2 absorbing aerosol height (AAH) is developed based on the FRESCO cloud product [Wang et al., 2008a]. FRESCO retrieves the effective cloud pressure and the cloud fraction using the reflectance of O_2 A-band at 760 nm. This wavelength is also suitable to retrieve ALH for cloud-free cases [Boesche et al., 2009, Dubuisson et al., 2009, Sanders et al., 2015], and since aerosols are treated in the same way as clouds in FRESCO, Wang et al. [2012] attempted to derive the aerosol layer pressure from the FRESCO cloud product. Their study led to the operational GOME-2 AAH product [Tilstra and Stammes, 2019], where heights of absorbing aerosols (UVAI calculated by 340-380 nm wavelength pair > 2) is determined by either cloud pressure or scene pressure according to the cloud fraction and the cloud albedo.

We collect AAH products provided by the GOME-2/Metop-A, B and C from 2018-08-01 to 2019-10-31. Pixels with SZA larger than 70°, or those affected by sun-glint or solar eclipse events are removed (AAH_Error_Flag = 0). Unconverged pixels due to computational failures with AAH set to be 15 km are also excluded. To assure the higher reliability, it is suggested to use samples with UVAI larger than 4 and the could fraction lower than 0.25 Tilstra and Stammes [2019].

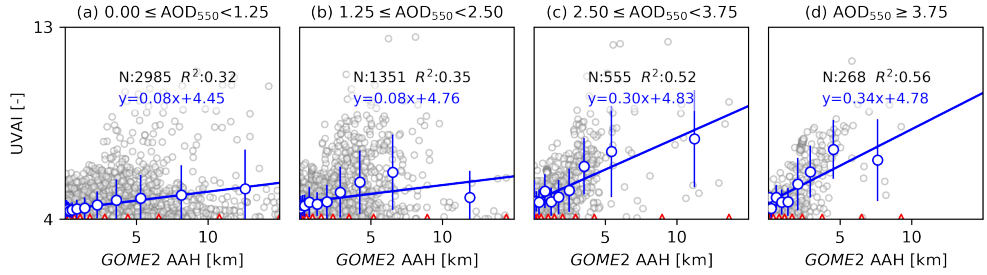
UVAI DEPENDENCE ON AEROSOL HEIGHTS RETRIEVED FROM SATELLITE OBSERVATIONS

Previous studies found that the relationship between UVAI and ALH is almost linear [de Graaf et al., 2005, Herman et al., 1997b, Torres et al., 1998], thus, similar to the analysis in Xu et al. [2019], we investigate the UVAI dependence on the TROPOMI and GOME-2 ALH products as a function of AOD, and use the correlation coefficient (R^2) and the linear regression fit to quantify the dependence. In order to avoid extreme outliers of the retrievals, an additional 1.5 times inter-quartile range ($IQR = Q3 - Q1$, where $Q1$ and $Q3$ are the first and the third quartile, respectively) is applied [Tukey, 1977]. There are total 75,249 TROPOMI samples and 5,159 GOME-2 samples left after the IQR filtering.

As both TROPOMI and GOME-2 do not have official AOD products yet (the TROPOMI ALH algorithm retrieves AOD, but it is only a diagnostic tool to indicate influence of bright surfaces and undetected clouds [Nanda et al., 2019]), we co-locate the Dark Target and Deep Blue combined AOD at 550 nm of the MODIS level 3 Daily Atmosphere Gridded Product [King et al., 2013] to the TROPOMI ALH (MYD08_D3, http://dx.doi.org/10.5067/MODIS/MYD08_D3.006) and the GOME-2 ALH (overpass time: 10:30 local time, MOD08_D3, http://dx.doi.org/10.5067/MODIS/MOD08_D3.006). Next, we manually divide each data into 4 groups according to AOD, and provide the R^2 between UVAI and ALH. In each AOD group, we further divided the data into $log_2(n)$ clusters based on the distribution of the ALH products using K-means, where n is the number of the samples in each AOD group. A linear fitting is applied to the mean UVAI and the mean ALH of the binned data.



Figuur 4.2: The TROPOMI UVAI (354-388 nm) against the TROPOMI O_2 A-band ALH as a function of the MODIS AOD at 550 nm. In each AOD group, the number of samples (gray dots) N and the correlation coefficient R^2 between ALH and UVAI are provided. Red triangles indicate the boundaries of ALH bins. The blue dot is the mean ALH and UVAI of each ALH bin, and the error bar is the standard deviation of UVAI of that bin. A linear regression is fit for the binned ALH and AOD values (blue dots).



Figuur 4.3: The GOME-2 UVAI (340-380 nm) against the GOME-2 AAH as a function of the MODIS AOD at 550 nm. In each AOD group, the number of samples (gray dots) N and the correlation coefficient R^2 between ALH and UVAI are provided. Red triangles indicate the boundaries of ALH bins. The blue dot is the mean ALH and UVAI of each ALH bin, and the error bar is the standard deviation of UVAI of that bin. A linear regression is fit for the binned ALH and AOD values (blue dots).

Figure 4.2 presents the relationship between the TROPOMI ALH and its corresponding UVAI product as a function of the MODIS AOD. The R^2 is only at level of 0.1 when the MODIS AOD is smaller than 1.25, but it increases gradually with AOD, and reaches 0.52 in the highest AOD group. The growing slope (k) of linear fitting on the binned data (blue dots) also presents an increasing sensitivity of UVAI to the changes in ALH. A similar relationship is found between the GOME-2 UVAI and AAH (Figure 4.3) but with higher R^2 than that of TROPOMI (ranges from 0.32 to 0.56). It is because that the TROPOMI ALH retains data with UVAI larger than 1, where weakly-absorbing or non-absorbing aerosols may weaken the dependence, whereas the GOME-2 AAH is only retrieved for strongly absorbing aerosols (UVAI larger than 4). The above is in good agreement with that presented by the radiative transfer simulations, and that found in the EPIC UVAI, ALH and AOD [Xu et al., 2019].

4.3. MERRA-2 AEROSOL HEIGHTS

As mentioned in Section 4.1, we will construct an aerosol height data set from the MERRA-2 aerosol profiles. This section starts with a brief introduction and validation of MERRA-2 aerosol reanalysis data. Then, we introduce various methods to derive an aerosol height from a given aerosol extinction profile, and investigate whether the UVAI dependence on the derived aerosol heights is in agreement with that found in the radiative simulations and satellite observations described above.

4.3.1. MERRA-2 AEROSOL REANALYSIS

MERRA-2 is the latest modern satellite era (1980 onward) atmospheric reanalysis from NASA GMAO [Buchard et al., 2017]. The spatial resolution is 0.5° × 0.625° latitude by longitude with 72 hybrid-eta layers from the surface up to 0.01 hPa. MERRA-2 assimilates multiple observational AOD data, including MODIS, AVHRR, AERONET and MISR [Randles et al., 2017]. Compared with the control run without AOD assimilation, an improved agreement between MERRA-2 and independent observations is found for aerosol optical properties and aerosol vertical distributions [Buchard et al., 2017].

In this work, we use the MERRA-2 3-hourly instantaneous aerosol mass mixing ratio profiles (MERRA-2 inst3_3d_aer_Nv, 10.5067/LTVB4GPCOTK2) to derive aerosol heights. The selected period is from 2006-01-01 to 2016-12-31. We use the mean value between 12:00 and 15:00 local time in order to be consistent with the satellite observations that are of most interests to us (i.e. OMI and TROPOMI, both have an overpass at 13:30 local time, [Levelt et al., 2006, Veefkind et al., 2012]).

CONVERSION FROM THE MASS MIXING RATIO TO THE EXTINCTION COEFFICIENTS

MERRA-2 provides mass mixing ratio profiles for 15 aerosol sub-species, including dust (5 size bins), sea salt (5 size bins), hydrophobic/hydrophilic black and organic carbon (BC and OC), and sulfate (SO_4^{-2}). A conversion from mass concentrations (c, unit: $kgkg^{-1}$) to extinction coefficients (β, unit: km^{-1}) is as follows:

$$\rho_x(z) = c_x(z) \times \rho_{air}(z) \tag{4.2}$$

$$\beta_x(z) = \rho_x(z) \times \beta_{m,x}(z) \tag{4.3}$$

where $c_x(z)$ is the mass mixing ratio (unit: $kgkg^{-1}$) of an aerosol type x at an altitude z (unit: m). $\rho_{air}(z)$ and $\rho_x(z)$ are the mass density (unit: kgm^{-3}) of the air and the aerosol species x. $\beta_{m,x}(z)$ and $\beta_x(z)$ are the mass extinction coefficients (unit: m^2kg^{-1}) and the extinction coefficients (unit: m^{-1}) for the aerosol species x. The aerosol mass extinction coefficient $\beta_{m,x}(z)$ is a function of relative humidity, whose values are provided in the supplementary document of [Randles et al., 2017]. The total extinction profile $\beta(z)$ is the summation of the contribution by X aerosol species:

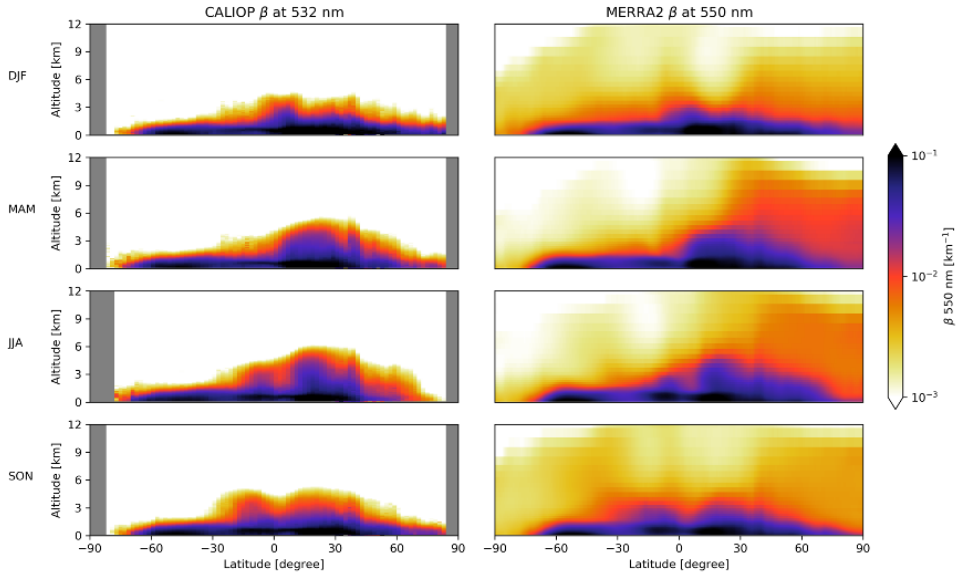
$$\beta(z) = \sum_{x=1}^{X} \beta_x(z) \tag{4.4}$$

Hereinafter, the term 'MERRA-2 aerosol profiles' indicates the 'aerosol extinction coefficient profiles' by default.

MERRA-2 VALIDATION WITH CALIOP

Aerosol extinction profiles simulated by chemistry transport models may have an order of magnitude difference [Kipling et al., 2016, Koffi et al., 2012]. Here we use the CALIOP level 3 all-sky aerosol extinction profiles from 2006 to 2016 and http://10.0.19.203/CALIOP/CALIPSO) to validate the MERRA-2 extinction profiles in the troposphere (below 12 km). The level 3 climatology is monthly available since June 2006. Its vertical and horizontal resolution is 60 m and 2°×5° (latitude by longitude). As the CALIOP climatology reports aerosol optical properties at 532 nm only, one should keep in mind that the extinction coefficient in the visible band has a lower measurement sensitivity in the lower part of the atmosphere due to strong attenuation by smoke [Kacenelenbogen et al., 2014, Kim et al., 2013, Liu et al., 2015, Torres and Chen, 2013].

Figure 4.4 shows the seasonal zonal aerosol extinction coefficient profiles ($\beta(z)$) as a function of latitude. The magnitude of the MERRA-2 extinction coefficients in the upper troposphere are generally higher than that of CALIOP. It may because that the MERRA-2 aerosol fields has higher background level of extinction coefficients. Besides, the CALIOP measurements suffer from problems of missing data or attenuated signal due to presence of clouds, heavy aerosol layers, underlying bright surfaces, etc., which may reduce the measurement sensitivity. Figure 4.5 shows the seasonal AOD maps calculated by the extinction coefficients below 12 km. The largest differences between CALIOP and MERRA-2



Figuur 4.4: Seasonal extinction coefficient ($\beta(z)$) profile as a function of latitude of CALIOP (left column) and MERRA-2 (right column) during period from 2006 to 2016.

occur at the Sahara region over all seasons, the biomass burning regions in the southern Africa during SON, and smoke plumes over the southern Atlantic Ocean during JJA and SON. Nevertheless, the spatial distribution and temporal variation of the MERRA-2 aerosol fields on a global scale generally correspond well with that of CALIOP.

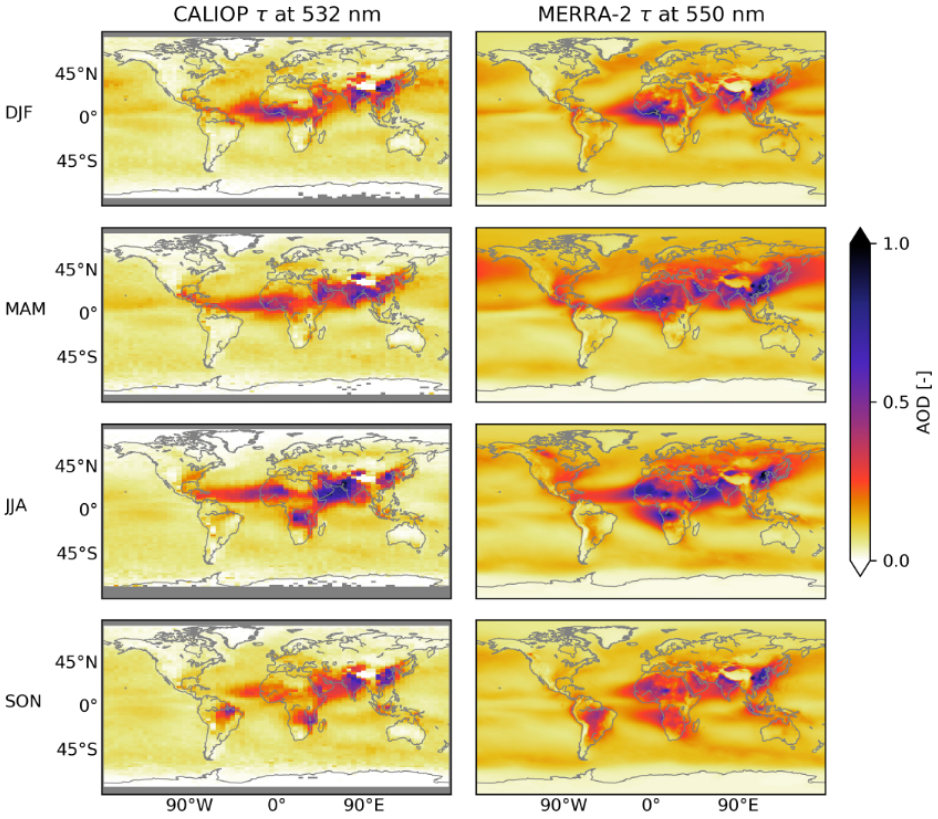
4.3.2. DERIVING AEROSOL HEIGHTS FROM THE MERRA-2 AEROSOL PROFILES

There are various methods to derive an aerosol height from a given aerosol extinction profile [Xu et al., 2018b]. It could be an averaged height weighed by the aerosol optical properties in vertical layers. An alternative is the scale height at which the aerosol extinction profile or the cumulative extinction profile passes a pre-determined threshold. An other common method is to find the geometrical boundary or center of a given profile. Below, we introduce the above methods (also summarized Table 4.2) in detail, and apply them to the MERRA-2 aerosol extinction profiles. Note that all aerosol heights are calculated for full profiles from the surface to the top of atmosphere (TOA), and they are relative to the terrain height by default unless it is mentioned otherwise.

AEROSOL OPTICAL PROPERTIES-WEIGHTED MEAN HEIGHT

Given an aerosol extinction coefficient profile ($\beta(z)$) with n layers, a common way to derive the aerosol height is calculating the mean height weighted by the extinction coefficient in each atmospheric height interval (Equation 4.5, [Chimot et al., 2018, Koffi et al., 2012, Kylling et al., 2018, Liu et al., 2019] or by the AOD in each atmospheric height interval (Equation 4.6, [Wu et al., 2016b]).

$$H^{\beta}_{aer} = \frac{\sum_{i=1}^{n} \beta(z_i) z_i}{\sum_{i=1}^{n} \beta(z_i)} \tag{4.5}$$



Figuur 4.5: Seasonal AOD climatology of CALIOP (left column) and MERRA-2 (right column) during period from 2006 to 2016.

Tabel 4.2: Definitions of the aerosol layer height.

Aerosol layer height	Symbols	Derivation method
Extinction-weighted mean aerosol layer height	H^{β}_{aer}	Equation 4.5
AOD-weighted mean aerosol layer height	H^{τ}_{aer}	Equation 4.6
Aerosol scale height with 63% AOD present	H^{63}_{aer}	Equation 4.7
Aerosol layer top height	H^{t}_{aer}	The first height searched from the surface where $\|\gamma ext(z)\| < 0.01$ for $z \geq H^{t}_{aer}$

$$H^{\tau}_{aer} = \frac{\sum_{i=1}^{n} \beta(z_i) dz_i z_i}{\sum_{i=1}^{n} \beta(z_i) dz_i} \tag{4.6}$$

where $\beta(z_i)$ and dz_i are the extinction coefficient and the geometric thickness of each atmospheric height interval i. The superscript β or τ indicates either the extinction coefficient or AOD of the layer is

used as averaging weight. If the thickness of atmospheric intervals is constant in vertical direction, then the H^{β}_{aer} and H^{τ}_{aer} will give the same result.

AEROSOL SCALE HEIGHT

Originally, the aerosol scale height is based on the assumption that the extinction profiles exponentially decay with altitude [Hayasaka et al., 2007, He et al., 2008, Tsai et al., 2011]. This restricts the application to the condition when the extinction coefficient peaks near the surface. Alternatively, a more generalized scale height definition is proposed [Hayasaka et al., 2007, Léon et al., 2009, Turner et al., 2001, Yu et al., 2010]:

$$\int_0^{H^{63}_{aer}} \beta(z)dz = \sum_{i=1}^{m} \beta(z_i)dz_i = \tau(1-e^{-1}) \approx 0.63\tau \tag{4.7}$$

where $\beta(z_i)$ and dz_i are the extinction coefficient and the geometric thickness of each atmospheric interval i, m is the number of layers up to the scale height, and τ is the total columnar AOD. The scale height H^{63}_{aer} indicates the altitude where 63% of total AOD is presented.

AEROSOL TOP HEIGHT

The above definitions are effective heights, where the aerosol loading should be placed to be representative of the columnar radiative properties. In contrast, the aerosol geometric height describes the 'actual' location of the aerosols [Kylling et al., 2018]. The geometric height is commonly presented by the upper and lower boundary or the center (the average of the upper and lower boundary) of an aerosol profile. For satellite measurements, the bottom boundary is more difficult to determine than the top boundary. Consequently, we only focus on how to define the top boundary of the highest aerosol layer (H^{t}_{aer}).

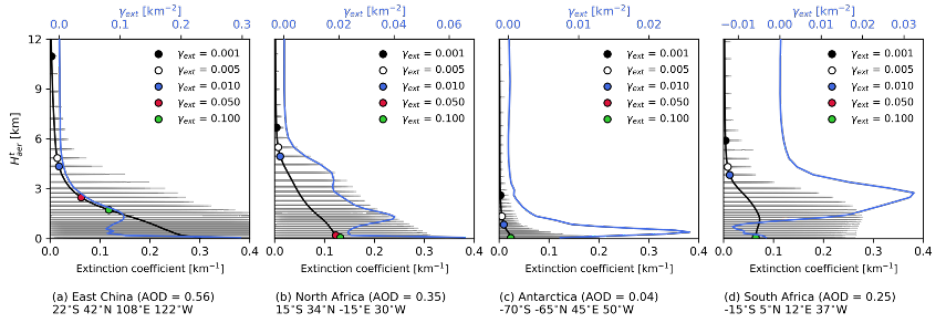
For a box-shape profile, H^{t}_{aer} is explicitly indicated by a clear sharp decrease of the extinction coefficient in a transition layer. For other profile types, there is no uniform method to determine H^{t}_{aer}. Welton et al. [2001] found the top boundary if the lidar signal strength is greater than the Rayleigh signal by a predetermined threshold setting. The mean signal over the next 500 m is also checked in order to avoid effects of noise. Léon et al. [2009] detected the top layer boundary by retaining the first altitude below TOA at which the signal is 3 times of the standard deviation larger than the average in the reference altitude (6.5 to 7 km). The CALIOP product employs a much more comprehensive layer detection algorithm (SIBYL, [Vaughan et al., 2009]), where the magnitude of the threshold is adapted according to the characteristics of the lidar signal [Winker et al., 2009].

For an extinction profile provided by MERRA-2, the extinction coefficient lapse rate (γ_{ext}, unit: km^{-2}) is introduced to help to determine the top boundary of an aerosol profile:

$$\gamma_{ext}(z) = -\frac{d\beta(z)}{dz} \tag{4.8}$$

where $d\beta(z)$ is the extinction coefficient difference between two continuous layers and dz is the atmospheric interval geometric thickness. The concept of γ_{ext} is proposed by Tian et al. [2017] to explore the relationship between atmospheric stability and aerosol vertical distributions. Stable meteorological conditions lead to a large positive γ_{ext}, whereas an elevated aerosol layers result in a negative value. Given an aerosol profile, we search upwards from the surface and retain the first height at which the magnitude of γ_{ext} above this height is always smaller than a certain value.

The choice of the threshold is empirical, which is based on a sensitivity study. As shown in Figure 4.6, we select 4 regions of interest and use their mean extinction profiles (averaged during the period from 2006-01-01 to 2016-12-31) as representative profile shapes: (a) a gradually decaying profile (North Africa), (b) a fast decaying profile (East China), (c) a profile with little variation (Antarctica) and (d) a profile with an inversion layer (South Africa). For each region, the mean profile shape (black lines) and their standard deviation (gray bars), together with γ_{ext} (blue lines) are shown. Next, we calculate



Figuur 4.6: Sensitivity study to determine the threshold of the extinction lapse rate (γ_{ext}) for the aerosol top boundary height (H^t_{aer}). The sensitivity study is conducted for 4 representative aerosol extinction coefficient profiles ($\beta(z)$) for (a) East China, (b) North Africa, (c) Antarctica and (d) South Africa. The black lines are the mean profiles and gray bars are the standard deviation during the period from 2006-01-01 to 2016-12-31. The blue lines are γ_{ext} of the mean profiles.

H^t_{aer} using different γ_{ext} ranging from 0.001 to 0.1 km^{-2}. It is clear that a threshold between 0.005 and 0.01 km^{-2} may be suitable for all profile shapes. Thus, in this paper, we use the γ_{ext} threshold of 0.01 km^{-2} to determine H^t_{aer}.

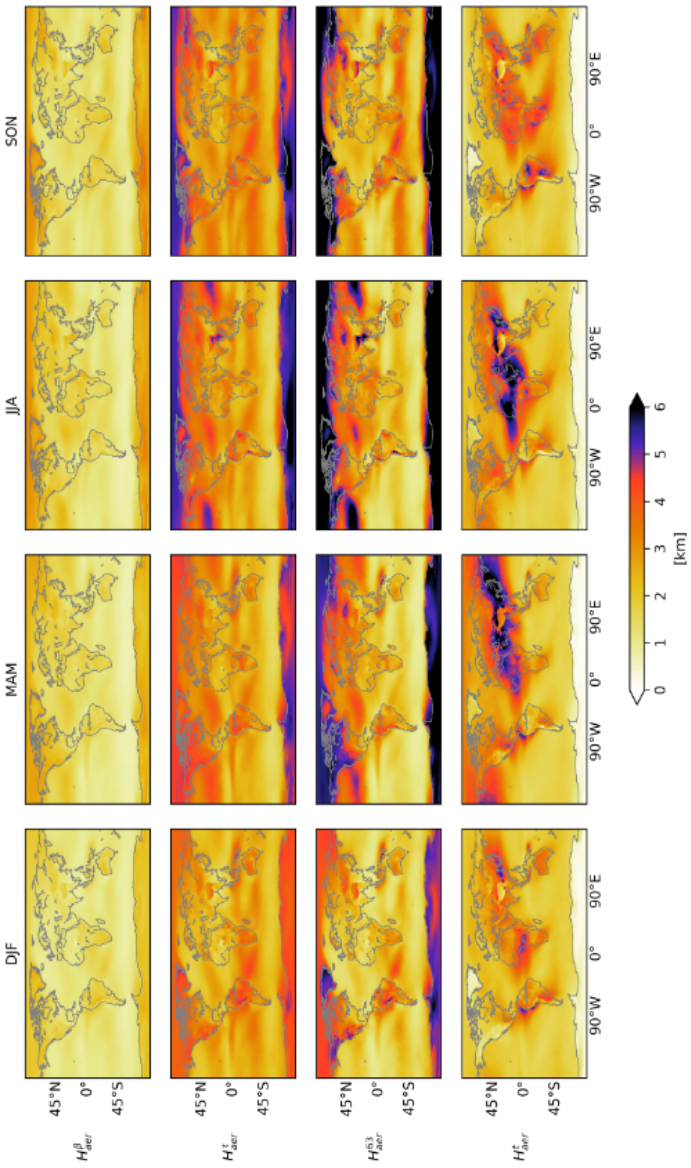
SPATIAL AND TEMPORAL DISTRIBUTION OF THE MERRA-2 AEROSOL LAYER HEIGHTS

The above aerosol height deriving methods are applied to the MERRA-2 aerosol extinction profiles collected from 2006-01-01 to 2016-01-01. Figure 4.7 shows the global seasonal climatology of these aerosol heights. The spatial distributions of the 3 effective heights (H^{β}_{aer}, H^{τ}_{aer} and H^{63}_{aer}) are similar. Compared with high aerosol loading regions (Figure 4.5), e.g. the dust belt and the biomass burning regions in Africa, the effective heights in the rest of the world are significantly higher, particularly in the high-latitude clean regions. The seasonal variations also show that the effective heights are more variable over these regions. It is noted that the magnitude of H^{β}_{aer} is lower compared with H^{τ}_{aer} and H^{63}_{aer}, because H^{β}_{aer} does not account for the varying thickness of each vertical interval. By contrast, the spatial-temporal variation of H^t_{aer} is associated with changes in AOD. One can recognize the seasonal aerosol sources from H^t_{aer} maps, e.g. the biomass burning events in the central Africa during DJF, the Sahara dust and its outflows over the northern Atlantic during JJA, etc.

Figure 4.8 shows the zonal average and standard deviation of the 4 aerosol heights against the corresponding zonal average of the MERRA-2 extinction coefficient profiles. The 3 effective heights behave in a similar way. They present significantly higher values over high-latitude regions where the extinction coefficients are low. Over the aerosol source regions in mid- and low-latitude, the effective heights are lower as more weights are given to the lower part of the atmosphere. H^t_{aer} generally follows the contours of the extinction coefficients. The magnitude and variability of H^t_{aer} is lower in the Southern Hemisphere and higher latitude regions (of both Hemispheres), because the major aerosol sources are located in the low- and mid-latitude of the Northern Hemisphere (Figure 4.5). The reason behind the behaviour of different aerosol heights will be explained in Section 4.3.3.

4.3.3. UVAI DEPENDENCE ON THE MERRA-2 AEROSOL HEIGHTS

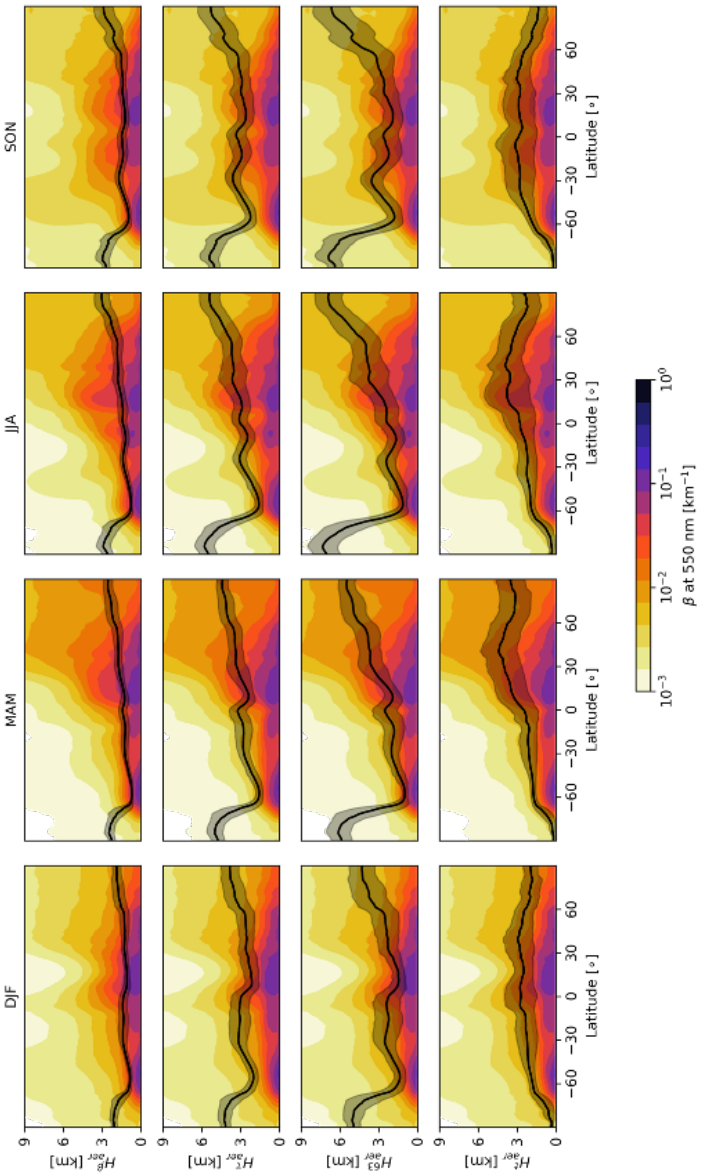
We further investigate the UVAI dependence on the derived MERRA-2 aerosol heights. The UVAI data is provided by the OMI/Aura level 2 OMAERUV product (http://dx.doi.org/10.5067/Aura/OMI/DATA2004) from 2006-01-01 to 2016-12-31. Pixels with SZA larger than 70°, or contaminated by clouds (cloud fraction > 0.3), sun-glint (glint angle > 20° over water) or the row anomaly (XTrackQualityFlags ≠ 0) are removed before analysis. Since the two data sets are independent, we use the AERONET



Figuur 4.7: Spatial distribution of MERRA-2 ALH seasonal climatology during period from 2006-01-01 to 2016-12-31. Rows represent different aerosol heights and columns represent different seasons.

version 3 level 1.5 products (https://aeronet.gsfc.nasa.gov) to ensure the consistency between them.

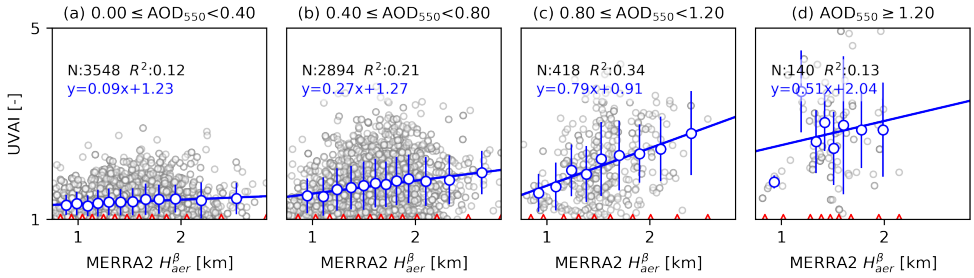
Both MERRA-2 and OMAERUV are co-located to an AERONET record if their time difference is smaller than ±3 hr and their spatial distance is within ≤50 km [Bréon et al., 2011, Jethva et al., 2014,



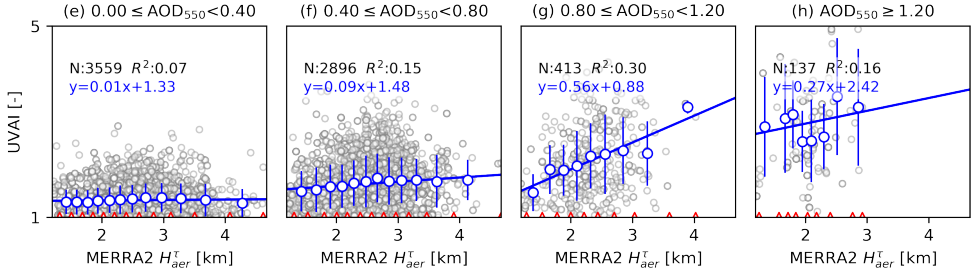
Figuur 4.8: The zonal average (black solid line) and the standard deviation (grey filling area) of the MERRA-2 ALH and the zonal average of the MERRA-2 extinction coefficients (contours) for the period from 2006-01-01 to 2016-12-31. Rows represent different ALH definitions and columns represent different seasons.

Lacagnina et al., 2015, Remer et al., 2002, Torres et al., 2002a]. Then, a AOD filter is applied as below:

- $|AOD_A - AOD_O| / AOD_O \leq 30\%$
- $|AOD_A - AOD_M| / AOD_M \leq 30\%$



Figuur 4.9: The OMAERUV UVAI (354-388 nm) against the MERRA-2 H^{β}_{aer} as a function of the MERRA-2 AOD at 550 nm. In each AOD regime, N is the number of samples (gray dots). Red triangles indicate the boundaries of ALH bins. The blue dot is the mean ALH and UVAI of each ALH bin, and the error bar is the standard deviation of UVAI of that bin. A linear fitting and correlation coefficient (R^2) are calculated for the binned ALH and AOD values (blue dots).

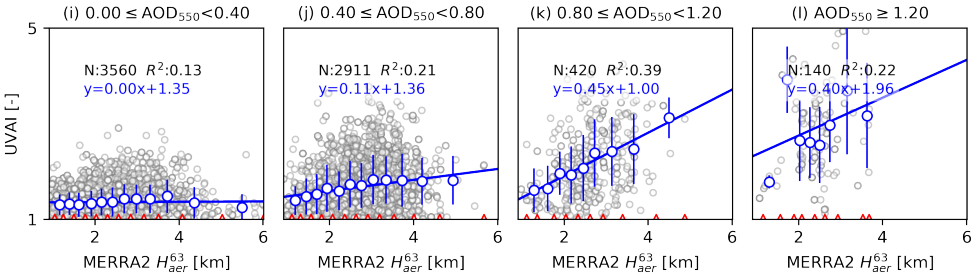


Figuur 4.10: Same as Figure 4.9 but for the MERRA-2 H^{τ}_{aer}.

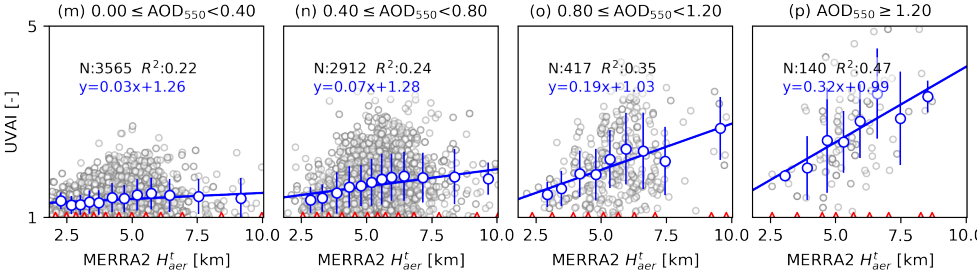
where A, O and M indicate AERONET, OMAERUV and MERRA-2. The threshold of 30% refers to the uncertainty of OMAERUV AOD [Ahn et al., 2014, Jethva and Torres, 2011b]. Since OMAERUV only reports aerosol properties up to 500 nm and MERRA-2 only reports at 550 nm, a wavelength conversion is necessary. Specifically, the OMAERUV AOD is extrapolated at 550 nm by the Ångström Exponent calculated from AOD at 388 and 500 nm. Besides, same as the TROPOMI and GOME-2 data, an 1.5 times IQR is applied to remove potential outliers.

Figure 4.9 to 4.12 present the OMAERUV UVAI against the derived MERRA-2 aerosol heights as a function of the AERONET AOD at 550 nm. The UVAI dependence on H^{β}_{aer}, H^{τ}_{aer} and H^{63}_{aer} (Figure 4.9, 4.10 and 4.11) is similar. For the first 3 AOD groups, both the R^2 between UVAI and the 3 effective heights, and the linear fitting slope k of the binned data increases with AOD, i.e. UVAI sensitivity to the changes in these aerosol heights increases with AOD. In the regime where AOD is highest, however, the UVAI dependence on the 3 effective heights becomes weaker. On the other hand, the UVAI becomes increasingly sensitive to H^{t}_{aer} when AOD is getting larger (Figure 4.12, k rises from 0 to 0.4). The corresponding R^2 is also relatively smaller when the aerosol loading is low, and it reaches 0.47 in the highest AOD cluster. In addition, compared with effective heights whose R^2 is no larger than 0.4, the correlation between H^{t}_{aer} and UVAI is generally higher. Based on above, we can conclude that among all aerosol heights derived from the MERRA-2 aerosol profiles, the UVAI dependence on H^{t}_{aer} matches best the evidence from sensitivity study and satellite observations.

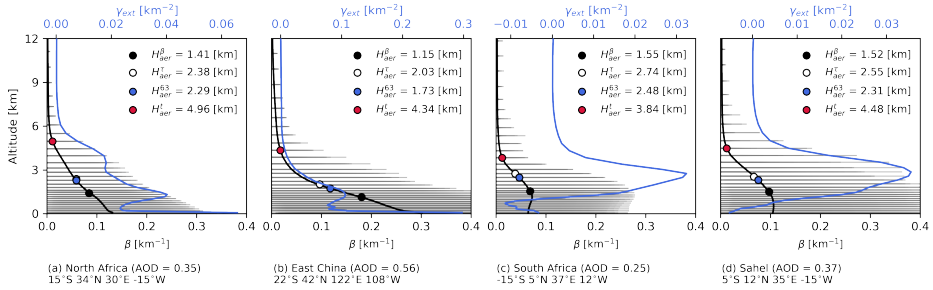
The 3 effective heights behaves similarly since they are all sensitive to the shape of a profile. More specifically, their calculations are sensitive to (1) the location of layers with the strongest aerosol



Figuur 4.11: Same as Figure 4.9 but for the MERRA-2 H^{63}_{aer}.

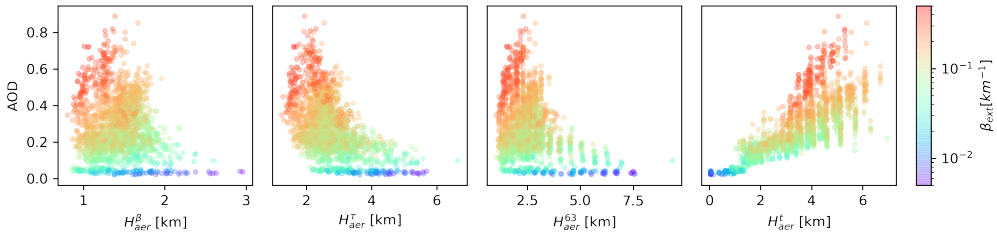


Figuur 4.12: Same as Figure 4.9 but for the MERRA-2 H^{t}_{aer}.



Figuur 4.13: Representative aerosol extinction coefficient profiles for (a) North Africa, (b) East China, (c) South Africa and (d) Sahel. The black lines are the mean profiles and gray bars are the standard deviation during the period from 2006-01-01 to 2016-12-31. The blue lines are the extinction lapse rate (γ_{ext}) of the mean profile.

extinctions (peak extinctions), and (2) the contrast of these strongest extinctions relative to extinctions of other layers. Here mean profiles derived from 4 regions of interest are shown to explain this (Figure 4.13). The profiles of (a) North Africa and (b) East China monotonically decrease with altitude, the profile of (c) South Africa and (d) Sahel represent elevated aerosol layers. Compared with profile (a) and (b) with their peak extinction (β_{max}) at the surface, the β_{max} of profile (c) and (d) are located at around 1.5 km. As a result, the derived effective heights are higher for profile (c) and (d). The magnitude of the peak extinction also matters. Compared with profile (a), profile (b) has a faster decay rate (larger Γ_{ext}), the contrast between its highest extinction layers and extinctions at upper layers is stronger than that of (a). Consequently, more weights are put on lower altitudes, leading to effective heights of profiles (b) smaller than (a). The same explanation is also applied to profile (c) and (d).



Figuur 4.14: Relationship between the aerosol heights derived from MERRA-2, AOD and the peak extinction (β_{mas}).

4

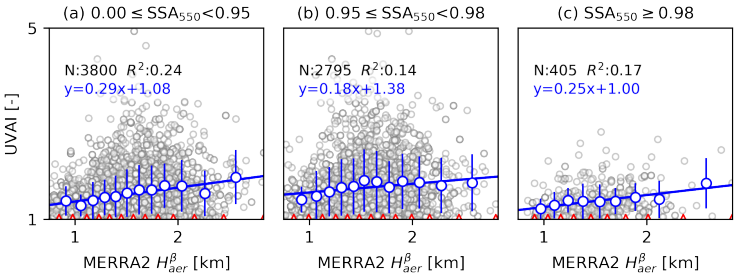
The weakest UVAI dependence on the effective heights for the highest AOD regime can be explained by the relation between AOD, β_{max} and the influence of profile shapes on the effective heights. Figure 4.14 presents the relationship between the AOD, β_{max}, and the derived MERRA-2 aerosol heights for regions of interest (Figure **??**). The effective heights become smaller and less variable with increasing AOD. Since AOD is positively associated with β_{max}, meanwhile from Figure 4.4 and 4.8, we also know that the location of the layers with strongest extinctions usually appears at the lower altitudes. In other word, as the aerosol loading becomes higher, the β_{max} tends to be larger but located at a lower altitude. Moreover, considering the influences of the profile shape discussed above, the effective heights become smaller and less variable with increasing AOD. As a result, the derived effective heights tend to be at a lower altitude with little variations, resulting in a lower correlation with the corresponding UVAI.

On the other hand, H^{t}_{aer} is better associated with AOD and become more variable when AOD is larger than 0.1. The top height captures the altitude above which aerosol makes little contribution to the total extinction, hence, it is more sensitive to the total extinction rather than the location or the magnitude of β_{max}. The stronger correlation between H^{t}_{aer} and AOD (Figure 4.5 and Figure 4.7) also enhance the correlation between H^{t}_{aer} and UVAI. One may note that the UVAI dependence on H^{t}_{aer} is not as strong as that found in satellite observations, because the satellite UVAI and ALH are both retrieved from the measured radiance of the same instrument (although at different wavelengths), whereas the OMAERUV UVAI is an independent data set to the MERRA-2 ALH and AOD.

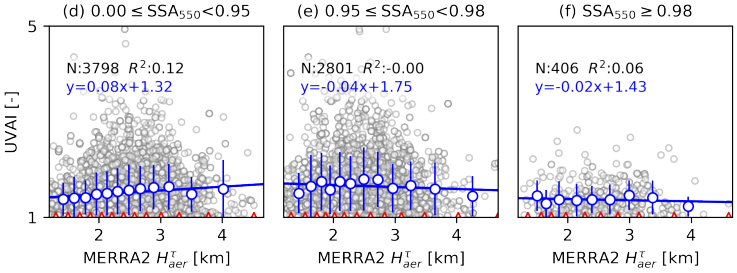
Figure 4.15 to 4.18 present the UVAI dependence on the derived MERRA-2 aerosol heights as a function of the SSA at 550 nm provided by AERONET. All the aerosol heights show the strongest correlation in the most absorbing regime (i.e. the lowest SSA). The sensitivity of UVAI on the 3 effective heights does not decreases along with increasing SSA. On the other hand, UVAI becomes less sensitive to H^{t}_{aer}, reflected by a reduced linear fitting slope k. Compared with the low correlation between UVAI and the effective heights (R^2 no more than 0.25), that between UVAI and H^{t}_{aer} is stronger (R^2 is between 0.28 and 0.40). The UVAI dependence on H^{t}_{aer} as a function of SSA is consistent with the findings of the sensitivity study by radiative transfer simulations (Figure 4.1).

4.4. SUMMARY

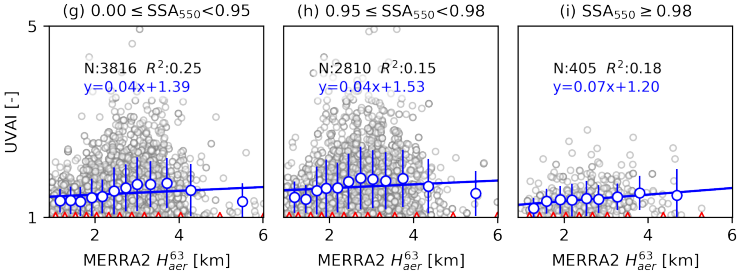
UVAI is a long-term global record with potentials to derive quantitative aerosol absorptive properties (e.g. AAOD and SSA), if both AOD and the aerosol height are known. Whereas AOD is widely accessible from various products, the availability of the aerosol height data is either restricted by the limited measurements of lidar instruments or the retrieval challenges (e.g. bright surfaces, clouds, low aerosol loading, etc.) of passive sensors. Currently, there is no independent observational data set that can provide the aerosol vertical distribution with a global coverage on a daily basis. To break the restrictions posed by the limited measurements, in this paper we create an aerosol height database from the aerosol profiles provided by the MERRA-2 for applications of quantifying aerosol absorption from satellite



Figuur 4.15: The OMAERUV UVAI (354-388 nm) against the MERRA-2 H^{β}_{aer} as a function of the MERRA-2 SSA at 550 nm. In each SSA regime, N is the number of samples (gray dots). Red triangles indicate the boundaries of ALH bins. The blue dot is the mean ALH and UVAI of each ALH bin, and the error bar is the standard deviation of UVAI of that bin. A linear fitting and correlation coefficient (R^2) are calculated for the binned ALH and AOD values (blue dots).



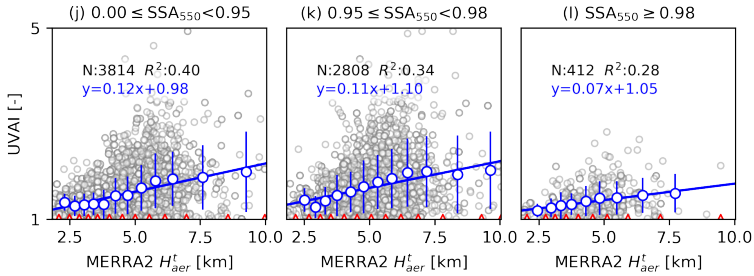
Figuur 4.16: Same as Figure 4.9 but for the MERRA-2 H^{τ}_{aer}.



Figuur 4.17: Same as Figure 4.9 but for the MERRA-2 H^{63}_{aer}.

UVAI.

We propose 4 candidate aerosol heights derived from the given MERRA-2 aerosol extinction profiles, including (1) the extinction-weighted mean height (H^{β}_{aer}), (2) the AOD-weighted mean height (H^{τ}_{aer}), (3) the scale height (H^{63}_{aer}) and (4) the top boundary height (H^{t}_{aer}). Through investigating the UVAI dependence on these aerosol heights, we find that the 3 effective heights (H^{β}_{aer}, H^{τ}_{aer} and H^{63}_{aer}) can perform as expected under majority circumstance except when AOD is large. Since effective heights are sensitive to the the altitude and the magnitude of layers with the strongest extinctions,



Figuur 4.18: Same as Figure 4.9 but for the MERRA-2 H^t_{aer}.

together with the fact that AOD is positively related to the peak extinction that is likely to be at a lower altitude, the UVAI dependence on the effective heights is lower for the highest AOD regimes. On the other hand, the UVAI dependence on H^t_{aer} matches best with the findings from sensitivity studies and satellite observations: when AOD increases, the sensitivity of UVAI to the changes of the H^t_{aer} becomes stronger, with a rising correlation R^2 from 0.22 to 0.47. Moreover, the UVAI dependence on H^t_{aer} also decreases with SSA, whereas UVAI shows little relation with the effective heights with the variation of SSA.

We thus conclude that the top boundary height (H^t_{aer}) derived from the MERRA-2 aerosol reanalysis can be an alternative aerosol height database, to avoid the restriction of the limited ALH products. We have succeeded using this aerosol height data set to derive global aerosol absorptive properties from the OMAERUV UVAI for the period from 2006 to 2019 and achieved encouraging outputs compared with the AERONET measurements.

5

AEROSOL ABSORPTION DERIVED FROM ULTRA-VIOLET AEROSOL INDEX BY DEEP LEARNING

The content of this chapter has been published as: **J. Sun**, P. Veefkind, P. van Velthoven and P. F. Levelt, Aerosol Absorption Over Land Derived From the Ultra-Violet Aerosol Index by Deep Learning, IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, vol. 14, pp. 9692-9710, https://doi.org/10.1109/JSTARS.2021.3108669, 2021.

5.1. INTRODUCTION

Atmospheric aerosols are solid or liquid particles suspended in the air. Most aerosols have a dominating cooling effect on the Earth's climate (-0.9 W/m^2 [-1.9 to 0.2 W/m^2]), however, the presence of light absorbing aerosols poses a partially offsetting warming effect [Pachauri et al., 2014]. Absorbing aerosols are mainly produced by combustion processes and dust uplifts [Moosmüller et al., 2009]. They affect the Earth's climate directly by absorbing the solar radiation or indirectly by altering the cloud and surface properties [Hansen et al., 1997, Hansen and Nazarenko, 2004, Nair et al., 2013, Ramanathan et al., 2001]. Quantifying the aerosol absorption is an important task to determine the aerosol radiative forcing and its uncertainties [Loeb and Su, 2010, McComiskey et al., 2008, Penner et al., 2001].

Quantitative aerosol absorptive properties, e.g. the aerosol single scattering albedo (SSA), which presents the ratio of scattering efficiency to total extinction efficiency, and the absorbing aerosol optical depth (AAOD), which describes the fraction of total columnar extinction due to particle absorption, are usually provided by the ground-based AERONET [Dubovik et al., 2002b, Holben et al., 1998], but the global distribution of AERONET stations is sparse and unbalanced. Observations from space have the advantage to have a better global coverage, but only few satellite sensors are capable to retrieve aerosol absorptive properties. Currently, satellite retrievals of AAOD and/or SSA use either multi-angle measurements, optionally combined with polarimetry [Kaufman et al., 1997a, Moosmüller et al., 2009], e.g. MISR/Terra (1999-present) [Diner et al., 1998, Garay et al., 2020], and POLDER-3/PARASOL (2005-2013) [Deschamps et al., 1994, Formenti et al., 2018, Wei et al., 2020].

Another parameter holding the information of the qualitative aerosol absorption is the Ultra-Violet Aerosol Index (UVAI) [Herman et al., 1997a]. Compared with the AAOD and SSA retrieved by MISR and POLDER-3, UVAI is simple to derive without multi-angular and polarimetric techniques. Moreover, its calculation does not rely on a priori assumptions of aerosol properties (e.g. the size distribution function and the complex refractive index). By contrast, the aerosol properties retrieved by MISR are based on Look-Up Tables (LUTs) produced for 74 pre-defined aerosol mixtures [Kahn and Gaitley, 2015]. The POLDER retrievals, either based on LUTs [Deuzé et al., 2000, Formenti et al., 2018, Tanré et al., 2011] or the recently developed GRASP algorithm that online calculates radiative transfer for multiple pixels simultaneously [Chen et al., 2018, Dubovik et al., 2014], also require pre-defined information on aerosol properties. On the other hand, the calculation of UVAI does not need to make such assumptions, because UVAI is defined as the change of the spectral contrast at two UV channels ($\lambda_1 < \lambda_2$) between a measured and a calculated radiance due to the presence of absorbing aerosols [Herman et al., 1997a, Torres et al., 1998]:

$$UVAI = -100\left[log_{10}\left(\frac{I_{\lambda_1}}{I_{\lambda_2}}\right)^{obs} - log_{10}\left(\frac{I_{\lambda_1}}{I_{\lambda_2}}\right)^{Ray}\right] \tag{5.1}$$

where Ray indicates the radiance calculated by radiative transfer simulations assuming no aerosol present. By assuming $I^{obs}_{\lambda_2} = I^{Ray}_{\lambda_2}(a_{s,\lambda_2})$ (a_{s,λ_2} is the surface albedo at the longer wavelength λ_2), the above can be rewritten as:

$$UVAI = -100 log_{10}\frac{I^{obs}_{\lambda_1}}{I^{Ray}_{\lambda_1}} \tag{5.2}$$

UVAI is a qualitative measure of aerosol absorption. A positive value indicates the presence of absorbing aerosols, whereas non-absorbing aerosols and clouds yield neutral or negative values [Herman et al., 1997a].

The UVAI has been continuously produced on a global scale by multiple satellites since 1978. It is of interest to quantify the aerosol absorption from such a long-term record. We have attempted to derive AAOD and SSA from UVAI based on radiative transfer simulations [Sun et al., 2018] and the Support Vector Regression (SVR) [Sun et al., 2019] for specific cases. Following previous studies, this work generates a database of quantitative aerosol absorptive properties over land from the UVAI between 2006 and 2019 provided by OMI/Aura using a Deep Neural Network (DNN) model. We create a

training data set using both observations and simulations, employ a filter as well as a wrapper method for feature selection, and apply the cross validation for hyper-parameter tuning. We provide an error analysis of the DNN predictions and validate the outputs with the AERONET observations. The result of this study leads to an aerosol climatology (2006-2019) of quantitative absorptive properties over land. Section 5.2 introduces the training data and the DNN algorithm. Section 5.3 presents the assessments of the derived aerosol absorption and the climatology for the period from 2006 to 2019. Section 5.4 summarizes the major findings and potential improvements in the future.

5.2. MATERIALS AND METHODS

5.2.1. DATA SETS AND METHODOLOGY

The data sets used in this study include the OMI/Aura level 2 version 3 aerosol product OMAERUV (http://dx.doi.org/10.5067/Aura/OMI/DATA2004), the MODIS/Aqua Collection 6 level 3 daily gridded aerosol product MYD08 (http://dx.doi.org/10.5067/MODIS/MYD08_M3.006, and the AERONET level 1.5 inversion almucantar product (https://aeronet.gsfc.nasa.gov). Besides, the MERRA-2 aerosol reanalysis is used to calculate the aerosol vertical distributions (MERRA-2 inst3_3d_aer_Nv, 10.5067/LTVB4GPCOTK2) and to provide AAOD and SSA (MERRA-2 tavg1_2d_aer_Nx, 10.5067/KLICLTZ8EM9D) for comparisons with the predicted results. All the data are collected globally from 2006-01-01 to 2019-12-31. All the observational data (i.e. satellite and ground-based data) are regarded to be cloud-free and all the satellite data are regarded to be not affected by large viewing angles and sun-glint after pre-processing.

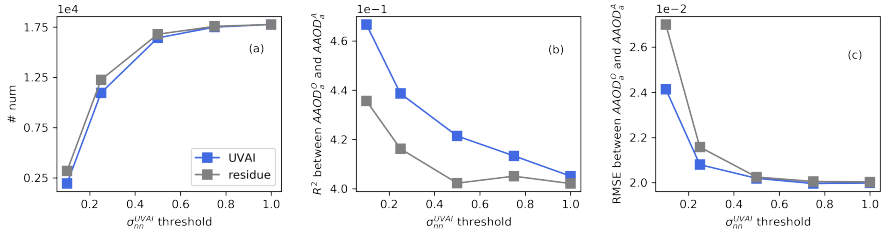
OMI OBSERVATIONS

OMI is a UV/Visible spectrometer onboard Aura (2004-present) [Levelt et al., 2006]. The instrument has a large swath of 2600 km, providing around 14 orbits per day to complete a global coverage (overpass on 13:45 local time). The spatial resolution at nadir is 13 $km\times$ 24 km (13 $km\times$ 48 km in the UV-1 band).

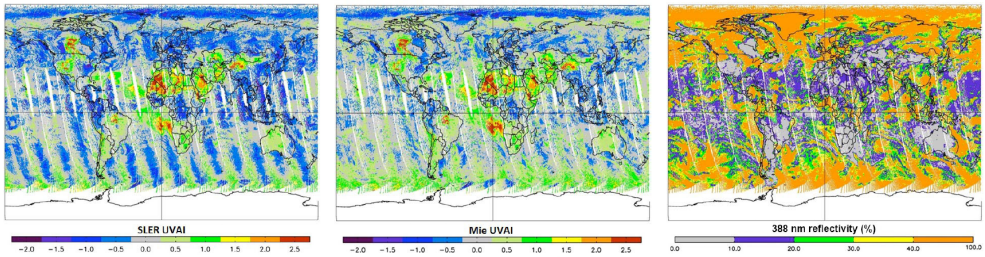
We take UVAI, the solar zenith angle (SZA), the satellite viewing zenith angle (VZA), the solar-satellite relative azimuth angle (RAA), the surface reflectance (a_s) and the surface pressure (P_s) from the level 2 OMAERUV version 3 product (http://dx.doi.org/10.5067/Aura/OMI/DATA2004). Pixels are excluded if they have SZA larger than 70°, or they are over bright surface (a_s at 388 nm higher than 0.3), contaminated by clouds (cloud fraction larger than 0.3) or sun-glint (glint angle larger than 20° over water).

OMI data suffers from the so-called row anomaly issue since 2008, which affects the quality of the level 1B radiance data at all wavelengths and consequently the level 2 products [Schenkeveld et al., 2017]. The row anomaly is caused by a partial blockage of the Earth telescope, affecting part of the across track swath. More information on row anomaly can refer to http://projects.knmi.nl/omi/research/product/rowanomaly-background.php, [Schenkeveld et al., 2017]. Even after filtering the data using the quality flag provided in OMAERUV product, the row anomaly can still be observed in some pixels. Therefore, we apply a moving standard deviation method to further eliminate the effects due to the row anomaly: for a given pixel, the standard deviation of UVAI (σ_{UVAI}) is calculated over a sliding window of its 8 neighboring elements. Pixels with σ_{UVAI} larger than a threshold are discarded. The threshold is determined by sensitivity studies based on co-located data of OMAERUV and AERONET.

According to Figure 5.1, the final determined σ_{UVAI} threshold is 0.5. This value on the one hand retains a large data size, meanwhile ensuring the consistency between the satellite and the ground-based AERONET measurements in terms of AAOD. We use AAOD as the criteria is because our study is based on the relationship between UVAI and AAOD. Both parameters contain information on the aerosol loading and the aerosol absorption. Moreover, AAOD is the common aerosol absorption parameter provided by both OMAERUV and AERONET.



Figuur 5.1: Sensitivity studies to determine the σ_{UVAI} threshold. (a) the number of OMAERUV-AERONET coincidences as a function of the σ_{UVAI} threshold; (b) the correlation coefficient (R^2) between the OMAERUV AAOD ($AAOD^O$) and the AERONET AAOD ($AAOD^A$) as a function of the σ_{UVAI} threshold; (c) the root mean square error (RMSE) between $AAOD^O$ and $AAOD^A$ as a function of the σ_{UVAI} threshold.



Figuur 5.2: Global depiction of SLER UVAI ('residue'), Mie UVAI ('UV Index') and scene reflectivity on 20 August 2007 (Source: Figure 12 in [Torres et al., 2018]).

The OMAERUV contains two UVAI products. One is called 'residue', where I_λ^{cal} is calculated by a Lambert Equivalent Reflectivity (LER) model as that done in UVAI products provided by other sensors [Herman et al., 1997a, Torres et al., 1998]. The other one is 'UV index', where I_λ^{cal} is reprocessed by considering the effects of clouds on scattering angular variability [Torres et al., 2018]. Although the 'UV index' shows a reduced across-scan bias and is better associated with the AERONET AAOD (Figure 5.1), it is overall higher than the 'residue', especially over ocean where neutral or negative should have been found (Figure 5.2). The reason behind is not explicitly explained. On the other hand, the method used to calculate 'residue' is widely used in other satellite products, whereas 'UV Index' is exclusively used for the OMAERUV product. Therefore, the term UVAI in this work refers to the 'residue'.

MODIS OBSERVATIONS

MODIS is a multi-spectral radiometer covering 36 wavelengths from 0.4 to 14.4 μm. A large swath of 2330 km allows it to complete a global coverage within 1 or 2 days. The spatial resolution ranges from 0.25 to 1 km. Currently, two MODIS instruments are operational: one is on EOS-Terra satellite launched in 1999 (descending node, overpass on 10:30 local time) and the other one is on EOS-Aqua launched in 2002 (ascending node, overpass on 13:30 local time) [Remer et al., 2005].

MODIS has two aerosol retrieval algorithms for (1) aerosols over ocean and vegetated dark surface ('Dark Target', DT) [Kaufman et al., 1997b, Levy et al., 2013, Tanré et al., 1997] and (2) aerosols over desert and arid regions ('Deep Blue', DB) [Hsu et al., 2004, 2006, Hsu et al., 2013]. The retrieved AOD uncertainty of DT is $\pm(0.05+15\%)$ over land and from $-0.04-10\%$ to $+0.02+10\%$ over ocean, and the expected uncertainty of DB is estimated better than $\pm(0.05+20\%)$. A hybrid AOD is also reported in a combination of DT and DB retrievals, where the method is based on the normalized difference vegetation index (NDVI) [Levy et al., 2013].

In this work, we use the DT and DB combined AOD at 550 nm from Collection 6.1 level 3 daily gridded data (MYD08_D3) of MODIS/Aqua (http://dx.doi.org/10.5067/MODIS/MYD08_M3.006) to provide information of aerosol loading. The resolution of the level 3 data is $1^\circ \times 1^\circ$. Pixels with geometric cloud fraction larger than 0.3 are excluded.

AERONET OBSERVATIONS

AERONET is a ground-based remote sensing network to retrieve aerosol optical, micro-physical and radiative properties for aerosol research and characterization, validation of satellite retrievals, and synergism with other databases [Holben et al., 1998].

The AERONET radiometer takes two types of measurements: (1) direct sun irradiance and (2) sky radiance. AOD is retrieved by the Beer-Bouguer Law from the direct sun measurements at a wide spectrum from 340 to 1020 nm [Kinne, 1999]. The bias of the near-real time AOD data is estimated as 0.02 with a standard deviation of 0.02 [Giles et al., 2019]. The sky radiance measurements scan an aerosol profile at multiple scattering angles to retrieve aerosol absorption (SSA and AAOD) and other aerosol micro-physical parameters (e.g. volume size distribution, complex refractive index, and the aerosol scattering phase function, etc.) at 440, 670, 870 and 1020 nm [Dubovik and King, 2000]. The AERONET level 2 (quality assured) SSA has a typical uncertainty of ±0.03 (for AOD at 440 nm larger than 0.2 [Dubovik et al., 2000, 2002b] or 0.3 [Sinyuk et al., 2020]).

In this work we use the AOD, AAOD and SSA from the AERONET version 3 level 1.5 inversion almucantar product (https://aeronet.gsfc.nasa.gov). Although the level 2 data is recommended (cloud screened and quality assured) [Smirnov et al., 2000], we use the level 1.5 product (only cloud screened) because the level 2 data availability is insufficient for our applications. Using the expected error of AOD (assuming σ_{AOD} = 0.02) and SSA (σ_{SSA} =0.03), and the definition of AAOD (Equation 5.3), we can calculate the expected error level of the AERONET AAOD using the error propagation equation (Equation 5.4):

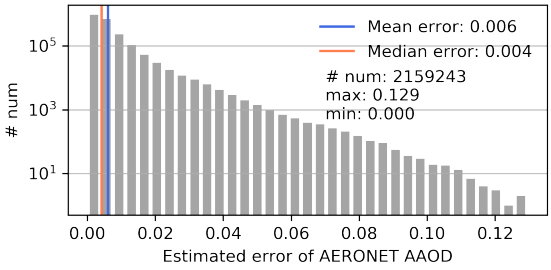
$$AAOD = (1 - SSA) \times AOD \tag{5.3}$$

$$\sigma_{AAOD} = \sqrt{\sigma_{SSA}^2 (\frac{\partial AAOD}{\partial SSA})^2 + \sigma_{AOD}^2 (\frac{\partial AAOD}{\partial AOD})^2} \tag{5.4}$$

Figure 5.3 presents the distribution of the AERONET AAOD at 550 nm uncertainty for all AERONET records collected between 2006 and 2019. For more than 80% of the samples, the AAOD uncertainty is less than 0.01. The mean and median of the estimated AAOD error are 0.006 and 0.004, respectively.

MERRA-2 AEROSOL REANALYSIS

MERRA-2 is the latest modern satellite era (1980 on-wards) atmospheric reanalysis [Buchard et al., 2017, Randles et al., 2017]. The model resolution is $0.5^\circ \times 0.625^\circ$ latitude by longitude with 72 hybrid-eta



Figuur 5.3: The estimated error of AERONET AAOD at 550 nm calculated by (5.3) with AERONET AOD uncertainty of ±0.02 and AERONET SSA uncertainty of ±0.03.

layers from the surface up to 0.01 hPa. MERRA-2 assimilates multiple observational AOD data sets, including MODIS, MISR, AVHRR and AERONET. MERRA-2 aerosol assimilation and the total column AOD evaluation are well-documented in [Randles et al., 2017]. They elaborated that the MERRA-2 AOD constrained by observations better matches independent measurements. Improved agreement is also found for aerosol optical properties and aerosol vertical distributions [Buchard et al., 2017].

In this work, we use the MERRA-2 columnar AOD and AAOD (reported at 550 nm) provided by the MERRA-2 1-hourly time-averaged aerosol (MERRA-2 tavg1_2d_aer_Nx, 10.5067/KLICLTZ8EM9D) to compare with our derived aerosol absorption results. Furthermore, we also derive the aerosol vertical distribution information from MERRA-2. This is important because UVAI is sensitive to the aerosol layer vertical location [de Graaf et al., 2005, Herman et al., 1997a, Sun et al., 2018, Torres et al., 1998]. Many efforts have been made on measuring aerosol vertical structures [Islam et al., 2017], including ground-based lidar systems [Pappalardo et al., 2014, Welton et al., 2001], space-borne lidar missions [Schutz et al., 2005, Winker et al., 2009], multi-angle measurements [Nelson et al., 2013, Virtanen et al., 2014], polarimetry [Dubovik et al., 2011], oxygen absorption at A-band [Dubuisson et al., 2009, Duforêt et al., 2007, Nanda et al., 2019, Sanders et al., 2015, Sanders and De Haan, 2016, Sanghavi et al., 2012, Wang et al., 2012, Xu et al., 2017], oxygen absorption in the visible band [Chimot et al., 2017] and thermal infrared [Pierangelo et al., 2004, Vandenbussche et al., 2013]. However, currently an aerosol vertical distribution product based on observations that has a daily global coverage as that of UVAI is still missing [Sun et al., 2020]. Consequently, we turn to derive the ALH from the aerosol vertical profiles provided by MERRA-2.

We derive the geometric aerosol layer top height as the ALH from the MERRA-2 3-hourly instantaneous aerosol mass mixing ratio profiles (MERRA-2 inst3_3d_aer_Nv, 10.5067/LTVB4GPCOTK2). The mass mixing ratio is converted to extinction coefficients first. Then for a MERRA-2 extinction profile, we attempt to find the top boundary of an aerosol layer with the extinction coefficient lapse rate (γ_{ext}, unit: km^{-2}, [Tian et al., 2017]), which is defined as:

$$\gamma_{ext}(z) = -\frac{d\beta(z)}{dz} \tag{5.5}$$

where $d\beta(z)$ is the extinction coefficient difference between two continuous layers, and dz is the atmospheric interval geometric thickness. Given an aerosol profile, we search upwards from the surface and retain the first height at which the magnitude of γ_{ext} above this height is always smaller than a certain value. The choice of the threshold is empirical, which we select 0.01 km^{-2} based on sensitivity studies (Figure 4.6 in Chapter 4). The relation between the derived ALH and the UVAI matches our knowledge of UVAI dependence on altitude [de Graaf et al., 2005, Herman et al., 1997a, Sun et al., 2018, Torres et al., 1998]. More details about the geometric aerosol layer top height derived from MERRA-2 can refer to Section 4.3.3 of Chapter 4.

5.2.2. CONSTRUCTION OF THE TRAINING DATA SET

The relevant parameters used in this work are listed in Table 5.1, including satellite-solar geometries (SZA, VZA, RAA), surface properties (a_s and P_s) and UVAI from OMI; AOD from MODIS; and AAOD and SSA from AERONET. Extinction Ångström Exponent (EÅE) and Absorption Ångström Exponent (AÅE) calculated between 440 and 870 nm from AERONET are used for AOD and AAOD spectral conversion according to the power law. Except for UVAI, all aerosol optical properties in this study is at 550 nm by default, unless other wavelengths are explicitly mentioned. The AOD, AAOD and SSA provided by OMI and MERRA-2 are also included for discussions and comparisons.

All the variables in Table 5.1 are merged into one hybrid data set. First, the daily OMI and MODIS satellite data are projected onto the MERRA-2 grid ($0.5° \times 0.625°$). Next, the satellite-model data is co-located to the ground-based AERONET observations. According to previous studies [Bréon et al., 2011, Ichoku et al., 2002, Jethva et al., 2014, Lacagnina et al., 2015, Remer et al., 2002, Torres et al., 2002a], the satellite-model joint data passing a time window of ±3 hours and a spatial distance ≤50 km is allocated to an AERONET record after averaging. To ensure the consistency between different

Tabel 5.1: Summary of data sets and relevant variables used in this study. All the data are collected globally for the period from 2006-01-01 to 2019-12-31. Parameters used as input features in DNN are bold.

OMI	**Solar zenith angle** (SZA), **viewing zenith angle** (VZA), **relative azimuth angle** (RAA), **surface albedo** (a_s, at 388 nm), **surface pressure** (P_s), **UVAI** (calculated by radiance at 354 and 388 nm), AOD (at 500 nm), AAOD (at 500 nm), SSA (at 500 nm)
MODIS	**AOD** (at 550 nm)
AERONET	AOD (at 550 nm), **AAOD** (at 550 nm), SSA (at 440 and 550 nm), Extinction Ångström Exponent (EÅE, between 440 and 870 nm), Absorption Ångström Exponent (AÅE, between 440 and 870 nm), **latitude** (°), **longitude** (°), **time** (day of year)
MERRA-2	AOD (at 550 nm), AAOD (at 550 nm), SSA (at 550 nm), **aerosol layer height** (ALH, derived from aerosol profiles)

measurement techniques (i.e. space-borne versus ground-based), we apply the following filtering criteria, based on the expected errors of the MODIS AOD ($\pm(0.05+15\%)$ over land and from $-0.04+10\%$ to $+0.02+10\%$) over ocean [Levy et al., 2013, Remer et al., 2005]), the OMAERUV AOD (0.1 or 30% [Ahn et al., 2014, Jethva and Torres, 2011b]) and the AERONET SSA (± 0.03 [Dubovik et al., 2002b]):

- $|AOD^M_{550} - AOD^A_{550}| \leq 0.05 + 15\% \times AOD^A$ (over land)
- $-0.04 - 10\% \times AOD^A \leq AOD^M_{550} - AOD^A_{550} \leq 0.02 + 10\% \times AOD^A$ (over ocean)
- $|AOD^O_{500} - AOD^A_{500}| \leq 0.1$ or $|AOD^O_{500} - AOD^A_{500}|/AOD^A_{500} \leq 30\%$)
- $|SSA^O_{500} - SSA^A_{500}| \leq 0.03$

where the upper-script M, A and O indicate MODIS, AERONET, and OMAERUV. Note that OMAERUV only provides aerosol properties on 354, 388 and 500 nm. Therefore, the corresponding AERONET AOD and AAOD are converted to 550 nm according to the power law:

$$\tau_{\lambda_1} = \tau_{\lambda_2} \left(\frac{\lambda_1}{\lambda_2} \right)^{-\alpha} \tag{5.6}$$

where τ_{λ_1} is the AOD at wavelength λ_1 to be estimated, and τ_{λ_2} is the AOD at wavelength λ_2 that is known; α is the EÅE listed in Table 5.1. The same conversion method is applied to AAOD using corresponding AÅE.

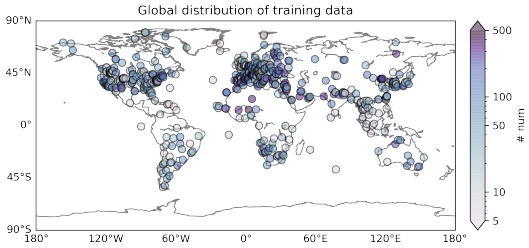
The final hybrid data set after quality filtering has 48,080 coincidences. Figure 5.4 shows its distribution. Most samples are located in North America and Western Europe. On the contrary, few sites measure aerosol absorptive properties over ocean. As a result, this study focus on deriving the aerosol absorptive properties over land.

5.2.3. FEATURE SELECTION

CANDIDATE FEATURES BASED ON DOMAIN KNOWLEDGE

In Machine Learning, features are treated as explanatory variables used to predict the target variable. In our algorithm, the target variable is the quantitative aerosol absorptive parameter AAOD and/or SSA. From our previous study [Sun et al., 2019], we have proven that compared with SSA, AAOD is better correlated with UVAI, because both AAOD and UVAI are sensitive to aerosol loading and aerosol absorption. Therefore, deriving AAOD from UVAI is expected to give better results than deriving SSA from UVAI.

UVAI is calculated from the satellite measured radiance. It does not only depend on the aerosol properties (e.g. AOD, ALH, aerosol absorption), but also on other parameters, including SZA, VZA, RAA, surface albedo (a_s) and surface pressure (P_s) [de Graaf et al., 2005, Herman et al., 1997a, Sun et al., 2018, Torres et al., 1998]. The above variables are thus listed as candidate features. Note that the



Figuur 5.4: Global distribution of the training data. The color indicates the number of observations. Note that the temporal coverage of each AERONET site varies.

influence of clouds has been eliminated by discarding pixels with cloud fraction larger than 0.3 (see Section 5.2.1) and will not be considered in this work.

Machine Learning algorithms solve the numerical relationship between given variables, but they neglect the fact that the environmental variables are correlated in space and time [Legendre, 1993]. Recently, there is an increasing trend involving the spatial and/or temporal auto-correlation in Machine Learning applications. This can be done in many ways, for example, Li et al. [2016] and Li et al. [2017] directly used the measurements in the past and/or the measurements from surrounding sites. In other studies, geo-statistical methods were used in combination with Machine Learning algorithms (e.g. Ishitsuka et al. [2018], Li et al. [2014], Michalak and Shlomi [2013], Singh et al. [2017], Tapoglou et al. [2014], Tarasov et al. [2018]). There is also an increasing trend using the Convolutional Neural Network (CNN) and/or the Recurrent Neural Network (RNN) to present the temporal and spatial correlation (e.g. [Biancofiore et al., 2015, Chen et al., 2020b, Wen et al., 2019]), as these techniques can share weight parameters in space and/or time domain. However, applying above techniques in this study is not feasible due to the limitation of training data sets (limited spatial coverage and different time spans, see Figure 5.4). Instead, we just directly add geo-coordinates (longitude and latitude) and time information (day of the year) into the feature space as that done in other studies (e.g. [Gupta and Christopher, 2009, Mauceri et al., 2019, Nanda et al., 2019, Olcese et al., 2015, Qin et al., 2018, Xiong et al., 2020, Zhu et al., 2019]). In total, we selected 11 variables as the candidate features based on our domain knowledge.

FEATURE SELECTION BY FILTER AND WRAPPER METHODS

We have chosen 11 features according to our experience and knowledge, but whether those features are favorable to derive AAOD from UVAI needs further investigations. Feature selection is the process of selecting a subset of features that is most relevant to the target variable. This process is important to enhance the model interpretability, computational efficiency, generalization performance, etc. [Guyon and Elisseeff, 2003].

In our previous study, we only selected 3 features that have relatively high Spearman's correlation coefficients (R_s^2) with the target variable (the AERONET AAOD) [Sun et al., 2019]. However, R_s^2 only measures the monotonic relationship between features and the target variable. In this work, we apply two independent feature selection methods: (1) the Maximum Information Coefficients (MIC) [Reshef et al., 2011] and (2) the Recursive Feature Elimination (RFE) [Guyon and Elisseeff, 2003].

MIC is a filter method that was first introduced by [Reshef et al., 2011]. It measures the dependence between two variable X and Y, no matter the relation is linear or non-linear. It uses a constrained adaptive bin method to apply mutual information (I) on continuous variables. The mutual information measures the mutual dependence between two random variables, which is defined as:

$$I(X;Y) = \int p(x,y) \log_2 \frac{p(x,y)}{p(x)p(y)} dxdy \tag{5.7}$$

where $p(x, y)$ is the joint probability density function of X and Y, and $p(x)$ and $p(y)$ are the marginal probability density function of X and Y. Then MIC uses binning to compute the normalized mutual information on continuous variables:

$$MIC(X,Y) = \max_{n_x * n_y < N^z} \frac{I(X;Y)}{\log_2 \min(n_x, n_y)} \tag{5.8}$$

where n_x and n_y are the bin numbers of X and Y. Their product is supposed to be smaller than a number N^z, where N is the size of the data and z usually takes 0.6 [Reshef et al., 2011]. Lastly, MIC is the maximum of the normalized mutual information values calculated by different combinations of n_x and n_y (MIC ranges between 0 and 1).

MIC is a filter method, which uses statistical measures to independently evaluate the correlation between the target variable and the input features, and to filter out the least relevant features. However, filter methods only look at one individual feature at a time, thus ignoring relationship between features. Wrapper methods like RFE can also detect the interactions between features [Phuong et al., 2005]. RFE recursively removes the least relevant feature by a certain metric and aims to find the feature combination that leads to the optimal model performance. The feature ranking metrics of RFE is given by an external estimator that assigns weights to each feature, such as the coefficients of the Linear Regression or the Support Vector Machines [Guyon and Elisseeff, 2003], or the feature importance of the Random Forests (RF) [Granitto et al., 2006]. In this work, we use the RF-based RFE, as RF can deal with non-linear problems and it does not have many hyper-parameters to tune. Since RF predictions will not always be the same, we set up 100 RFE experiments and use the average value to select the features of interest.

The MIC and REF evaluation on the 11 candidate features are presented in Figure 5.5. MIC (blue bars) indicates the UVAI, AOD, ALH are the most relevant features to AAOD, which is in agreement with our previous study [Sun et al., 2019]. The geo-coordinates (Lat and Lon) are more important compared to the rest of the parameters. The average of 100 RFE experiments (white bars) shows how the model performance changes (in terms of the averaged root mean square error, RMSE) with an eliminated feature. It is clear that the most important feature is UVAI, followed by AOD and ALH. The remaining features do neither significantly improve nor hurt the model performance. As there is no solid reason to exclude them, we decide to keep all 11 features.

The basic statistics of the selected features are shown in Figure 5.6. But to show the impact of the spatial and temporal information (Lat, Lon, DOY) on predictions, we also build a feature space without this information to allow a comparison. Consequently, we have the following two feature spaces:

- Feature space with 11 features (F11): UVAI, AOD, ALH, Lat, Lon, DOY, SZA, VZA, RAA, a_s, P_s
- Feature space with 8 features (F8): UVAI, AOD, ALH, SZA, VZA, RAA, a_s, P_s

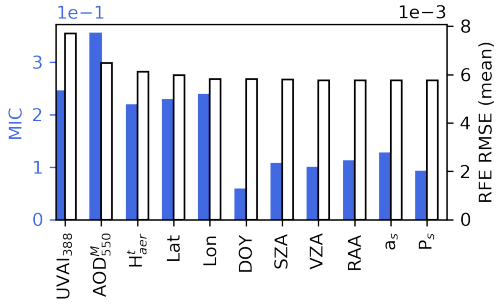
5.2.4. DEEP NEURAL NETWORK (DNN)

PRINCIPLE OF DEEP NEURAL NETWORK

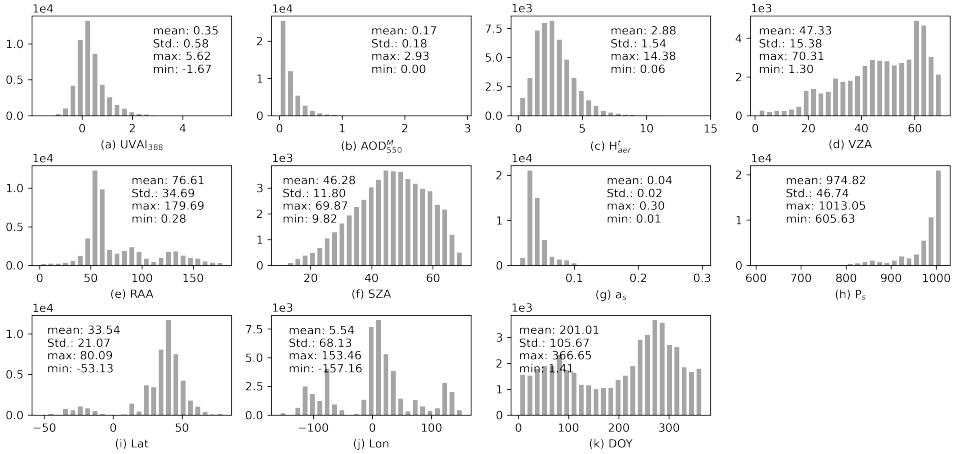
Machine Learning algorithms learn the underlying behavior of a system from a set of training data [Lary et al., 2016b]. They have been widely applied in geosciences and remote sensing [Karpatne et al., 2019a, Lary et al., 2016b]. Recently, driven by the increasing size of geo-data, algorithms such as Deep Learning becomes more and more popular [Reichstein et al., 2019, Zhang et al., 2016]. Deep Learning is characterized by Neural Networks with no less than 2 hidden layers [Zhu et al., 2017].

We use a feed-forward Deep Neural Network (DNN) with multiple hidden layers (Figure 5.7). A DNN consists of an input layer, n intermediate hidden layers and an output layer. The input layer contains input features, and the output layer gives the predicted variable(s). A hidden layer consists of m neurons. The j-th neuron in l-th hidden layer (a_l^j) is calculated by neurons in the previous layer:

$$z_l^j = \sum_{i=1}^{n} w_l^{ji} a_{l-1}^i + b_l^j \tag{5.9}$$



Figuur 5.5: Feature selection: the Maximum Information Coefficient (MIC, blue bars) and the Recursive Feature Elimination (RFE, white bars). The higher MIC values, the stronger correlation between a feature and the target variable (AAOD). RFE presents how the model performance (in format of averaged RMSE from 100 random forest experiments) varies with an eliminated feature. The higher the RMSE (the worse the model performance), the more important the corresponding feature.



Figuur 5.6: Histogram and basic statistics of the selected features.

where w_l^{ji} is the weight of the i-th neuron in $(l-1)$-th layer (a_{l-1}^j) given to the a_l^j; b_l^j is a bias term. Then, the computed z_l^j is fed into an activation function σ:

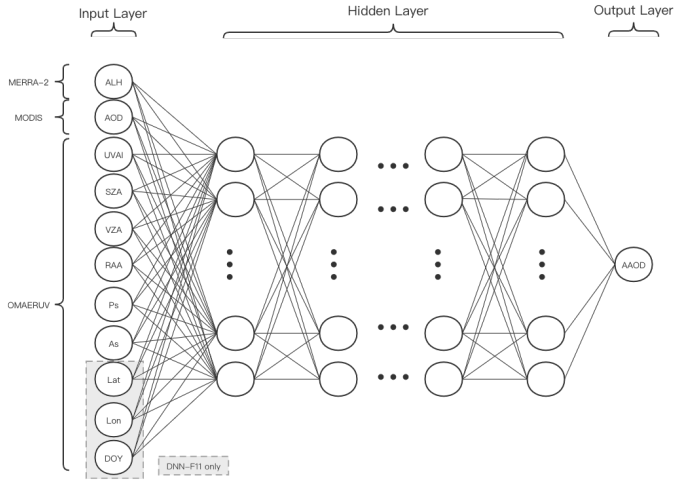
$$a_l^j = \sigma(z_l^j) \tag{5.10}$$

The activation function is used to add the non-linear properties to a neural network. In this work, the Rectified Linear Unit (ReLU) is used as the activation function:

$$\sigma(x) = max(0, x) \tag{5.11}$$

The DNN is to find the optimal weight matrix W containing all w_l^{ji} that can minimized the loss function (L). Here we use the RMSE as the error metric:

$$L = \left[\frac{1}{k} \sum_{i=1}^{k} (\hat{y}_i - y_i) \right]^{1/2} \tag{5.12}$$



Figuur 5.7: A conceptual scheme of the Deep Neural Network (DNN) in this study.

where k is the number of samples; $\hat{y}_i$ and y_i are the DNN-predicted and the true values. In this work, the optimization of loss function is realized by the Adam solver [Kingma and Ba, 2017] with a fixed learning rate of 10^{-4}. The Adam solver is more computationally efficient than classical optimization algorithms and performs well for large data sets [Alfadda et al., 2018].

HYPER-PARAMETER TUNING

The structure of DNN (the number of hidden layers n and the number of neurons in each layer m) highly varies with applications. Therefore, we applied an exhaustive grid search over specified hyper-parameter values. The process is based on 10-fold cross validations: the data set is randomly split into 10 sets. For each fold, 9 out of 10 sets are used to train the DNN model and the remaining set is used for validation.

Table 5.2 and 5.3 present the training results and the validation of the model performance for different layer and neuron configurations. Compared to the model performance of DNN-F8, the DNN-F11 predictions show a better consistency with the true values (R^2 =0.9). The RMSE and the mean absolute error (MAE) of the DNN-11 predictions are also smaller than that of DNN-8. The comparison reveals that the additional spatial and temporal information indeed improves the DNN model performance. However, the improvement is not significant as that in other studies (e.g. [Li et al., 2017]), because we only provide coordinate and time information in a straightforward way, instead of giving the explicit spatial and temporal auto-correlation calculated by geo-statistics methods (e.g. deterministic interpolation, Kriging, or Gaussian process regression), or sharing weights in space and time domain by CNN or RNN.

The optimal model of DNN-F11 is constructed by 3 hidden layers with 64 neurons in each layer. The final trained model has a prediction accuracy of 0.0045 (Figure 5.8 (a)). This model performance is significantly better than the SVR model used in the previous work ([Sun et al., 2019], RMSE = 0.01). Considering the representative uncertainty of AERONET AAOD is at the level between 0.004 and 0.006 (calculated by the expected errors of AERONET AOD and SSA via the error propagation equation, see Figure 5.3), the model performance is encouraging. The optimal model of DNN-F8 is constructed by 2 hidden layers with 128 neurons in each layer. Without the information on coordinate and time, the final trained model has a prediction accuracy of 0.0056, slightly higher than that of DNN-F11. Consequently, we only use the DNN-F11 in the following applications.

Tabel 5.2: Model performance of DNN with 11 features under difference configurations in terms of the linear fitting slope (k), the intercept (b), the correlation coefficient (R^2), the root mean square error (RMSE) and the mean absolute error (MAE).

		Training					Validation				
Layer	Neuron	k	b	R^2	RMSE	MAE	k	b	R^2	RMSE	MAE
1	64	0.78	0	0.88	5.34E-03	3.42E-03	0.78	0	0.88	5.47E-03	3.48E-03
	128	0.79	0	0.89	5.19E-03	3.34E-03	0.79	0	0.88	5.46E-03	3.44E-03
	256	0.79	0	0.9	5.16E-03	3.29E-03	0.77	0	0.88	5.34E-03	3.40E-03
2	64	0.85	0	0.9	4.92E-03	3.19E-03	0.82	0	0.89	5.21E-03	3.36E-03
	128	0.84	0	0.91	4.71E-03	3.05E-03	0.82	0	0.89	5.18E-03	3.31E-03
	256	0.85	0	0.92	4.51E-03	2.92E-03	0.83	0	0.89	5.06E-03	3.23E-03
3	64	0.83	0	0.91	4.71E-03	3.06E-03	0.82	0	0.89	5.25E-03	3.33E-03
	128	0.86	0	0.93	4.19E-03	2.69E-03	0.82	0	0.9	4.96E-03	3.18E-03
	256	0.85	0	0.93	4.10E-03	2.61E-03	0.82	0	0.9	4.96E-03	3.13E-03
4	64	0.83	0	0.92	4.53E-03	2.94E-03	0.81	0	0.89	5.10E-03	3.28E-03
	128	0.84	0	0.94	3.92E-03	2.51E-03	0.81	0	0.9	4.86E-03	3.10E-03
	256	0.84	0	0.93	4.11E-03	2.56E-03	0.8	0	0.9	4.94E-03	3.10E-03
5	64	0.84	0	0.92	4.39E-03	2.84E-03	0.81	0	0.89	5.03E-03	3.23E-03
	128	0.86	0	0.94	3.99E-03	2.56E-03	0.84	0	0.9	4.81E-03	3.09E-03
	256	0.86	0	0.94	3.79E-03	2.36E-03	0.82	0	0.91	4.76E-03	3.03E-03
6	64	0.82	0	0.92	4.52E-03	2.94E-03	0.79	0	0.9	4.98E-03	3.23E-03
	128	0.86	0	0.94	3.78E-03	2.37E-03	0.82	0	0.9	4.82E-03	3.05E-03
	256	0.83	0	0.94	3.87E-03	2.35E-03	0.79	0	0.9	4.88E-03	3.03E-03

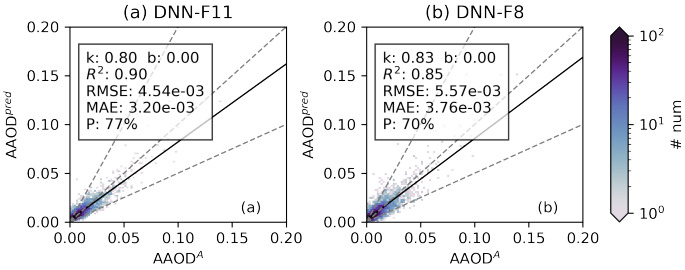
Tabel 5.3: Model performance of DNN with 8 features under different configurations in terms of the linear fitting slope (k), the intercept (b), the correlation coefficient (R^2), the root mean square error (RMSE) and the mean absolute error (MAE).

1	64	0.76	0	0.86	5.84E-03	3.65E-03	0.75	0	0.85	5.88E-03	3.68E-03
	128	0.76	0	0.86	5.70E-03	3.60E-03	0.75	0	0.86	5.84E-03	3.65E-03
	256	0.76	0	0.87	5.54E-03	3.55E-03	0.75	0	0.86	5.73E-03	3.62E-03
2	64	0.75	0	0.88	5.51E-03	3.52E-03	0.74	0	0.86	5.76E-03	3.63E-03
	128	0.8	0	0.88	5.35E-03	3.42E-03	0.79	0	0.86	5.77E-03	3.59E-03
	256	0.79	0	0.89	5.15E-03	3.28E-03	0.77	0	0.87	5.63E-03	3.52E-03
3	64	0.78	0	0.89	5.17E-03	3.32E-03	0.76	0	0.87	5.63E-03	3.56E-03
	128	0.83	0	0.91	4.82E-03	3.01E-03	0.8	0	0.87	5.61E-03	3.44E-03
	256	0.84	0	0.91	4.64E-03	2.94E-03	0.81	0	0.87	5.45E-03	3.40E-03
4	64	0.77	0	0.9	4.94E-03	3.16E-03	0.74	0	0.87	5.51E-03	3.48E-03
	128	0.82	0	0.92	4.49E-03	2.84E-03	0.77	0	0.87	5.41E-03	3.37E-03
	256	0.8	0	0.91	4.71E-03	2.91E-03	0.76	0	0.87	5.48E-03	3.38E-03
5	64	0.82	0	0.9	4.92E-03	3.13E-03	0.79	0	0.87	5.53E-03	3.45E-03
	128	0.82	0	0.92	4.49E-03	2.85E-03	0.78	0	0.88	5.33E-03	3.36E-03
	256	0.81	0	0.92	4.43E-03	2.72E-03	0.77	0	0.88	5.30E-03	3.28E-03
6	64	0.77	0	0.9	4.92E-03	3.14E-03	0.73	0	0.87	5.58E-03	3.47E-03
	128	0.81	0	0.92	4.39E-03	2.76E-03	0.77	0	0.88	5.34E-03	3.33E-03
	256	0.81	0	0.92	4.42E-03	2.67E-03	0.78	0	0.88	5.31E-03	3.26E-03

5.3. RESULTS AND DISCUSSION

5.3.1. ASSESSMENTS OF THE DNN-PREDICTED AEROSOL ABSORPTION

The DNN-F11 model is applied to predict the aerosol absorption from 2006 to 2019. The AAOD estimated by the DNN and further derived SSA (SSA = 1 - AAOD/AOD) are validated by the co-located AERONET records and compared with the aerosol absorption provided by OMAERUV and MERRA-2. To ensure the consistency between different products, we apply the same quality filtering described in Section 5.2.2 before analysis. This filtering is also applied to MERRA-2 AOD and SSA. There are in total 39,504 coincidences for validation.



Figuur 5.8: Performance of the optimal model selected by grid search. The linear fitting slope (k), intercept (b), the correlation coefficient (R^2), the root mean square error (RMSE) and the mean absolute error (MAE) between the observed and predicted values, and the percentage of AAOD difference within AERONET AAOD estimated error in the total samples (P) are provided: (a) DNN with 11 features (DNN-F11); (b) DNN with 8 features (DNN-F8). The dashed lines are 1:1 and 50% difference lines.

5

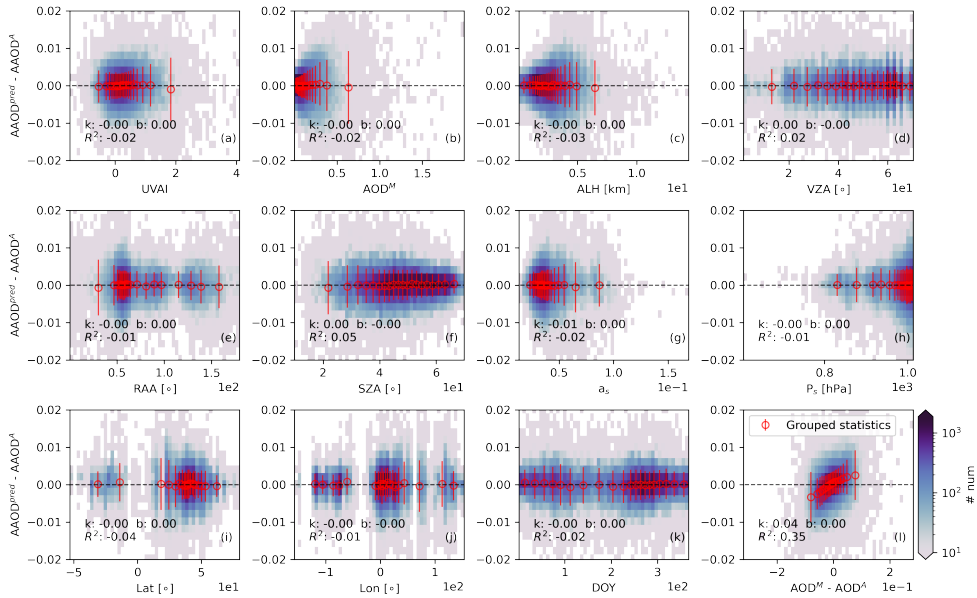
THE INFLUENCE OF INPUT FEATURES ON THE DNN PREDICTIONS

First, we investigate the relationship between the errors of the DNN predictions and input features. According to Figure 5.9, the difference between the DNN-predicted and the AERONET-retrieved AAOD shows little dependence on the input features. However, it is slightly related to the AOD difference between MODIS and AERONET (Figure 5.9 (l)). The positive bias of MODIS AOD leads to overestimation of the DNN-predicted AAOD. On the other hand, Figure 5.10 (b) reflects that low AOD may cause large biases in the DNN-derived SSA, because MODIS observations become less sensitive to low aerosol loading [Levy et al., 2013]. This indicates that a quality filtering on AOD is necessary to ensure the quality of the DNN predictions. As a result, based on the sensitivity study of the DNN-derived SSA accuracy (Figure 5.11), a threshold of 0.1 on the MODIS AOD at 550 nm is applied to the DNN predictions. The same threshold is also applied to the OMAERUV AOD (converted to 550 nm) and the MERRA-2 AOD for direct comparisons. There are total 21,600 samples after AOD-screening. The following discussions are based on the filtered data.

VALIDATING THE DNN PREDICTIONS WITH AERONET

According to Figure 5.12 (a), the filtered DNN-predicted AAOD show a high correlation with the AERONET retrievals (R^2=0.89). The linear fitting shows that the DNN predictions have a slight negative bias, which is mainly caused by the underestimation for cases with high aerosol absorption. Both RMSE and MAE are at level around 0.005, and about 83% samples falling within the AERONET AAOD estimated error. The corresponding SSA is also positively biased (Figure 5.13 (a)). Over 80% of the samples are within the AERONET SSA typical uncertainty of ±0.03. Compared with the SVR predictions in our previous study with only 66% data is inside the ±0.03 confidence interval [Sun et al., 2019], the DNN-predictions show a significant improvement. Despite of the larger bias for high absorption cases, the averaged error of DNN-derived SSA is at level around 0.02.

Aerosol absorptive properties of OMAERUV (at 500 nm) and MERRA-2 (at 550 nm) are also compared with the AERONET retrievals. Although both the OMAERUV AOD and SSA are quality assured by AERONET, its AAOD is still considerably lower than the corresponding AERONET retrievals, resulting in a RMSE larger than the expected magnitude (Figure 5.12 (b)). Similar as the DNN predictions, the gap is mainly due to the underestimation of the high aerosol absorption cases. The MERRA-2 AAOD tends to be overestimated for lower values and underestimated for higher values, leading to the highest error level (both RMSE and MAE) among the three data sets (Figure 5.12 (c)). In most cases, the MERRA-2 SSA is also higher than that of AERONET, with only 66% of samples within the ±0.03 confidence range (Figure 5.13 (c)).



Figuur 5.9: The AAOD difference between DNN and AERONET against input features. The gray dashed line indicates the neutral. The color of density plots indicates the number of samples. The data is further grouped into 20 sets, and the red circles and error bars are the mean and standard deviation of each group.

Tabel 5.4: Aerosol classification based on the Extinction Ångström Exponent (EÅE), Absorption Ångström Exponent (AÅE) and SSA at 440 nm adapted from [Cazorla et al., 2013, Giles et al., 2012].

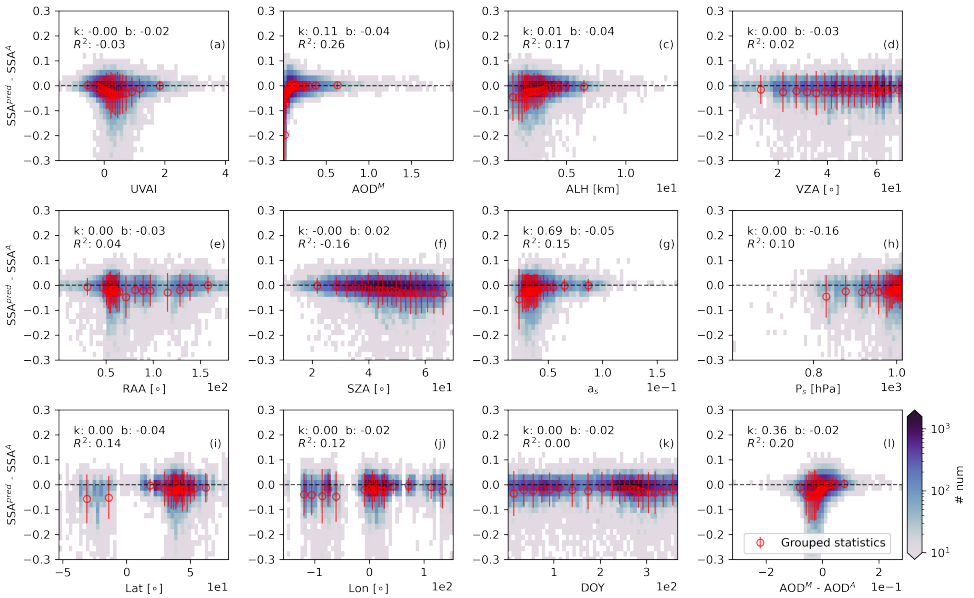
Aerosol types	$EÅE_{440-870}$	$AÅE_{440-870}$	SSA_{440}	Comment
Smoke	$EÅE \geq 1.5$	$AÅE \geq 0$	$SSA \leq 0.95$	Small absorbing aerosols
Dust	$EÅE \leq 0.5$	$AÅE \geq 1$	$SSA \leq 0.95$	Large absorbing aerosols
Mixed	$0.5 \leq EÅE \leq 1.5$	$AÅE \geq 1$	$SSA \leq 0.95$	Mixture of small and large absorbing aerosols
Other		All other conditions		Non-absorbing aerosols

VALIDATING THE DNN PREDICTIONS BY AEROSOL TYPES

We further validate the DNN predictions according to aerosol types. We categorize the aerosols into four types (i.e. smoke, dust, smoke-and-dust-mixed and other non-absorbing aerosols) by the EÅE between 440 and 870 nm, the AÅE between 440 and 870 nm, and the SSA at 440 nm reported in the AERONET inversion product (Table 5.4). This classification method is adapted from [Cazorla et al., 2013, Giles et al., 2012].

The comparisons of the DNN, OMAERUV and MERRA-2 against the AERONET AAOD and SSA for the four aerosol types are presented in Figure 5.14 and Table 5.5. All three data sets show their best consistency with the AERONET retrievals for smoke aerosols (R^2 >0.90). R^2 slightly decreases for larger absorbing aerosols (mixed and dust aerosols), but most of the samples is still within the ±50% difference range (dashed lines). On the other hand, all three data sets are not well correlated with AERONET for non-absorbing aerosols (R^2 <0.8), and tend to overestimate the aerosol absorption.

Compared to OMAERUV and MERRA-2, the DNN predictions have higher R^2 and lower error



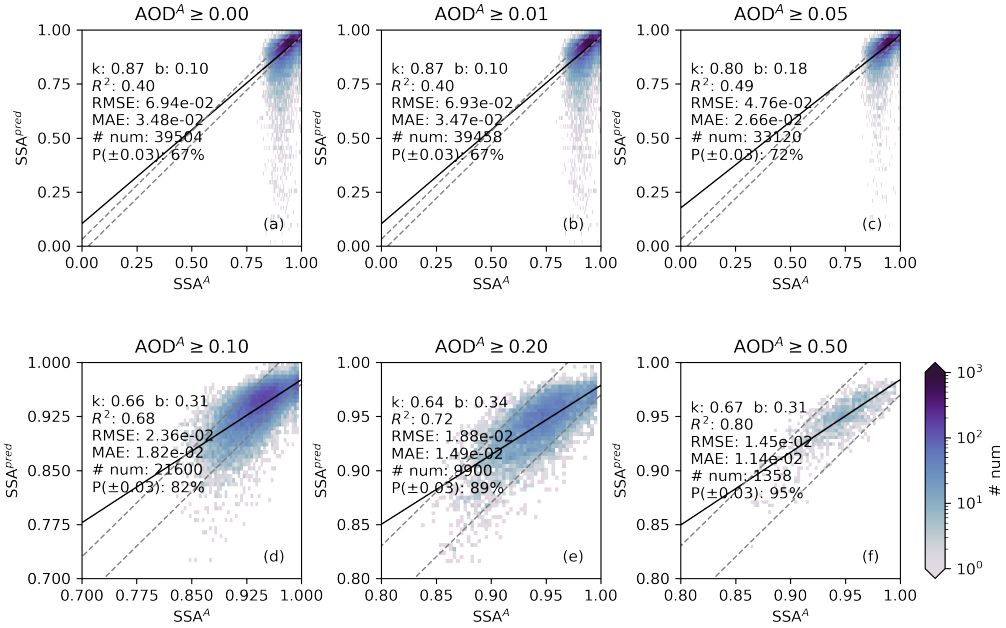
Figuur 5.10: The SSA difference between DNN and AERONET against input features. The gray dashed line indicates the neutral. The color in density plots indicates the number of samples. The data is further grouped into 20 sets, and the red circles and error bars are the mean and standard deviation of each group.

in general. The agreement between the DNN predictions and the AERONET retrievals varies with aerosol types. The DNN is best at predicting aerosol absorption for smaller absorbing aerosols. This is because the DNN-predicted AAOD is most sensitive to the input MODIS AOD and OMAERUV UVAI. The uncertainty of the MODIS AOD in the visible band is higher for dust aerosols due to the stronger surface reflectance [Hsu et al., 2013, Levy et al., 2013, Remer et al., 2005, Wei et al., 2019]. Consequently, the DNN predictions for dust aerosols are less consistent with the AERONET retrievals than smoke aerosols. The gap between DNN-predictions and AERONET is largest for non-absorbing aerosols. It is because that the OMAERUV UVAI is more sensitive to absorbing aerosols than scattering aerosols by definition [Herman et al., 1997a, Torres et al., 1998].

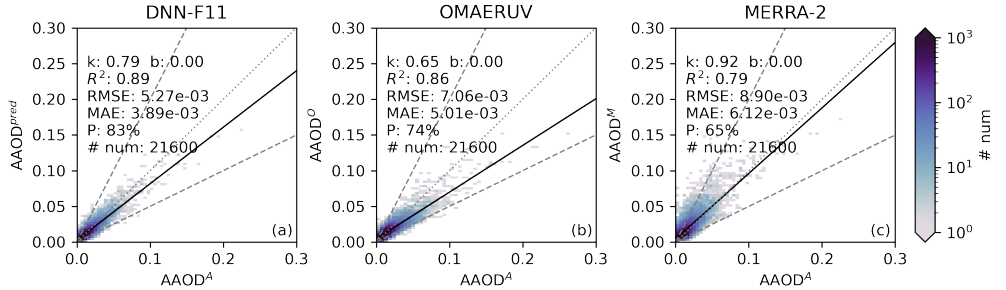
The OMAERUV AAOD is overall lower than the AERONET retrievals, with k <0.8 for all aerosol types. This is caused by the higher OMAERUV SSA compared to AERONET for the majority of the data. This agrees with findings in [Jethva et al., 2014]. According to [Jethva and Torres, 2011a], the potential reasons behind the SSA difference to AERONET could come from the clouds contamination, the assumptions on ALH, the high surface albedo (in desert and arid areas), and the assumed aerosol micro-physical properties. On the contrary, the MERRA-2 AAOD shows slightly positive biases (except for smoke), particularly for non-absorbing aerosols, reflecting that the MERRA-2 aerosol models may overestimate the aerosol absorption.

ASSESSMENTS OF THE DNN-PREDICTED AEROSOL ABSORPTION BY CASES

Here we selected 4 cases to further investigate the performance of the DNN predictions, covering dust storms and biomass burning events in Africa. Figure 5.15 and 5.16 present the MODIS/Aqua true color images and the OMAERUV UVAI for the selected cases. Case 2019-02-14 and 2019-02-18 show the biomass burning events in the central Africa. Case 2019-07-10 and 2019-08-07 mainly present the dust



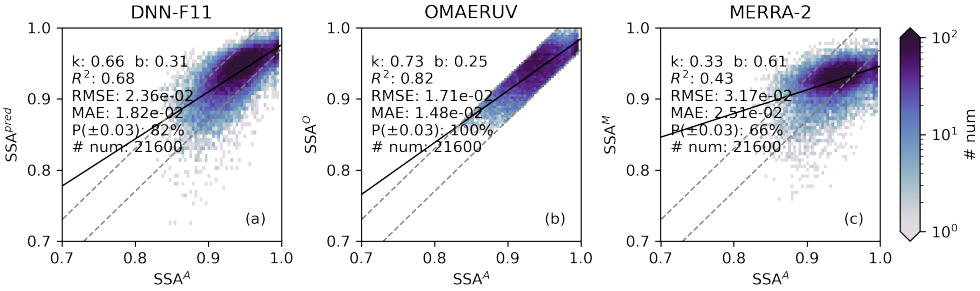
Figuur 5.11: Sensitivity of DNN-derived SSA accuracy to AOD. The gray dashed lines are the ±0.03 uncertainty of AERONET SSA.



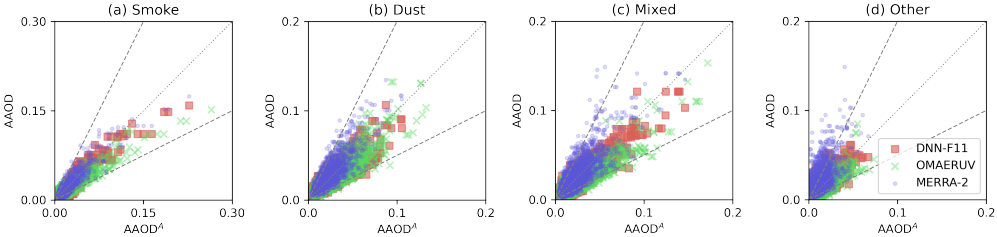
Figuur 5.12: Comparison between the AERONET AAOD and the AAOD of (a) DNN-F11 predictions, (b) OMAERUV and (c) MERRA-2 after the AOD-screening. The black lines are the linear fittings. The dotted and dashed lines are 1:1 and 50% difference lines, respectively. *P* is the percentage of the difference between the DNN-predicted AAOD and AERONET AAOD smaller than the expected error of the AERONET AAOD calculated by (Equation 5.4) in the total samples.

storms in the northern Africa with some biomass burning events in the southern Africa.

Figure 5.17-5.20 show the DNN-predicted AAOD, SSA and the input MODIS AOD (left column), and their difference with the corresponding parameters of OMAERUV (middle column) and MERRA-2 (right column). In general, the DNN-derived AAOD can reflect the distribution of absorbing aerosols (cloud-free parts) as that shown in Figure 5.15 and 5.16, whereas it is difficult to tell the dust or smoke plumes from the derived-SSA alone (same for SSA of OMAERUV and MERRA-2). The distribution of the DNN-derived AAOD is dominant by the input MODIS AOD, whereas the effect of UVAI is weaker.



Figuur 5.13: AERONET SSA against the SSA of (a) DNN-F11 predictions, (b) OMAERUV and (c) MERRA-2 after the AOD-screening. The black lines are linear fittings. The gray dashshed lines are the ±0.03 uncertainty of AERONET SSA. *P* is the percentage of the difference between the DNN-predicted SSA and AERONET SSA smaller than the expected error of the AERONET SSA (±0.03) in the total samples.



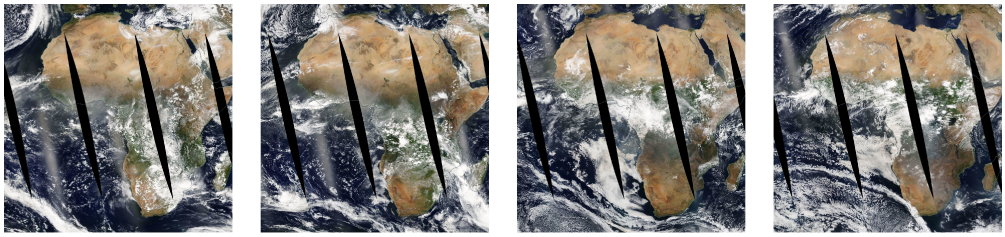
Figuur 5.14: AERONET AAOD against that of DNN-F11 (red square), OMAERUV (green cross) and MERRA-2 (blue circle) for four aerosol types based on Table 5.4: (a) smoke, (b) dust, (c) mixed and (d) other aerosols. The dashed lines are 1:1 and 50% difference lines.

Tabel 5.5: AAOD validation by aerosol types. The columns are statistics between AAOD of each data set and AERONET: the root mean square error (RMSE), the mean absolute error (MAE), the linear fitting slope (k), intercept (b) and correlation coefficient (R^2), the percentage of difference within the expected AERONET error level in the total samples (P), the number of samples (N), and the mean error of AAOD, AOD and SSA (ME_{AAOD}, ME_{AOD} and ME_{SSA}) of each aerosol type.

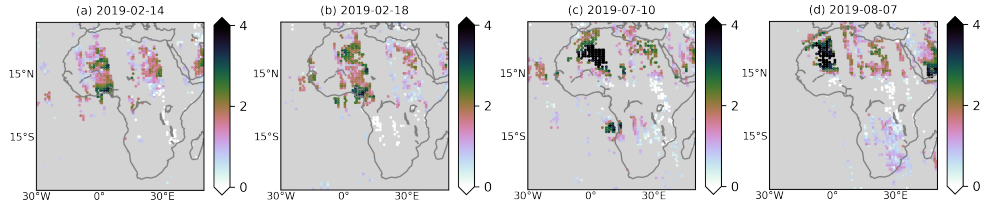
Type	Data	RMSE	MAE	k	b	R^2	P[%]	N	ME_{AAOD}	ME_{AOD}	ME_{SSA}
Smoke	DNN-F11	5.46E-03	3.94E-03	0.9	0	0.93	79	3533	-0.003	-0.02	0.00
	MERRA-2	6.09E-03	4.17E-03	0.96	0	0.91	77	3533	-0.001	-0.01	0.00
	OMAERUV	7.51E-03	4.74E-03	0.65	0	0.91	72	3533	-0.002	0.00	0.01
Dust	DNN-F11	6.50E-03	4.75E-03	0.73	0.01	0.86	89	2410	0.00	0.00	0.00
	MERRA-2	1.01E-02	7.42E-03	1.02	0	0.83	74	2410	0.005	0.01	-0.01
	OMAERUV	8.92E-03	6.66E-03	0.73	0	0.85	81	2410	-0.003	-0.01	0.00
Mixed	DNN-F11	5.21E-03	3.86E-03	0.82	0	0.9	82	7322	-0.002	-0.01	0.00
	MERRA-2	7.92E-03	5.21E-03	1.02	0	0.82	73	7322	0.001	-0.01	0.00
	OMAERUV	7.77E-03	5.47E-03	0.6	0	0.83	69	7322	-0.003	-0.01	0.01
Other	DNN-F11	4.83E-03	3.64E-03	0.73	0.00	0.73	84	8335	0.002	0.00	-0.01
	MERRA-2	1.02E-02	7.37E-03	1.04	0.01	0.63	50	8335	0.007	0.00	-0.03
	OMAERUV	5.41E-03	4.24E-03	0.67	0.00	0.7	76	8335	0.002	0.01	-0.01

For example, in case 2019-07-10 and 2019-08-07, both MODIS AOD and DNN-derived AAOD show high values in biomass burning regions, whereas the corresponding UVAI is very low.

The difference plots of AAOD and SSA show that the DNN-predicted aerosol absorption is generally higher than that of OMAERUV, whereas it is lower than that of MERRA-2. This finding is consistent



Figuur 5.15: MODIS/Aqua true color maps of the selected cases: (a) 2019-02-14; (b) 2019-02-18; (c) 2019-07-10; (d) 2019-08-07. Source: https://worldview.earthdata.nasa.gov/.



Figuur 5.16: OMAERUV UVAI of the selected cases: (a) 2019-02-14; (b) 2019-02-18; (c) 2019-07-10; (d) 2019-08-07.

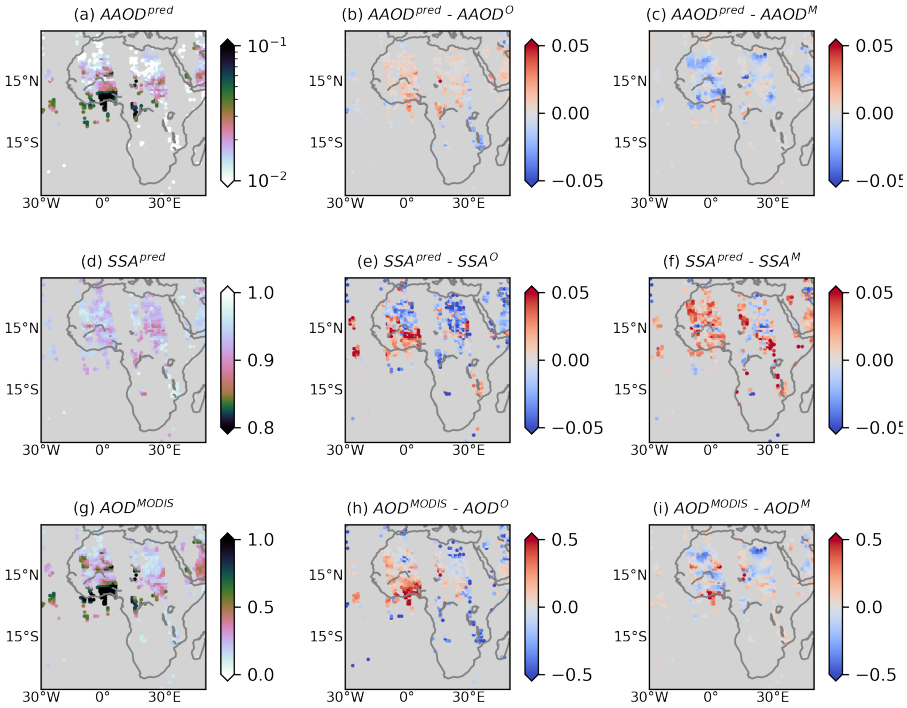
with that found in previous sections. The difference in aerosol absorption is mainly associated with the difference in AOD. For example, in case 2019-02-18, the MODIS AOD of the smoke plume is significantly higher than that of OMAERUV (Figure 5.18 (h)), and the corresponding AAOD and SSA difference shows the same pattern (Figure 5.18 (b, e)).

5.3.2. AEROSOL ABSORPTION CLIMATOLOGY

Here we discuss the aerosol absorption climatology over land derived by DNN for the period from 2006 to 2019. Although it is applicable for aerosols over ocean, the DNN model is mainly trained for aerosols over land, as the most of AERONET sites measure aerosol absorptive properties of continental aerosols (Figure 5.4). The aerosol properties over ocean may have a different distribution from that of continental aerosols, which could bias the predictions. Therefore, we only discuss the aerosol absorption over land in this section, and provide a tentative analysis on the predicted aerosol absorption over ocean.

GLOBAL AEROSOL ABSORPTION CLIMATOLOGY OVER LAND

Figure 5.21 and 5.22 present the seasonal aerosol absorption of DNN, OMAERUV and MERRA-2 over land averaged between 2006 and 2019. All three data sets show the major absorbing aerosol sources in each season, for instance, the biomass burning events in Central Africa in DJF and MAM, and dust storms from Sahara in MAM and JJA, etc. Compared with MERRA-2, both the DNN and OMAERUV AAOD have lower magnitudes, especially over desert regions. It is because both MODIS and OMI have difficulty in retrieving AOD over bright surfaces in the visible band. Besides, in OMAERUV, dust particles are assumed to be spherical which is not a proper assumption to calculated the dust optical properties [Torres et al., 2002b]. On the contrary, MERRA-2 assimilates the MISR and AERONET AOD observations over bright surfaces [Buchard et al., 2017, Randles et al., 2017]. Owing to MISR's multi-angle measurement technique providing more constraints on the inversion, the MISR AOD is better associated with AERONET than MODIS over desert regions [Kalashnikova and Kahn, 2008, Mishchenko et al., 2010]. Both MERRA-2 and DNN show that the northern India is a considerable source in all seasons, although the latter has a lower magnitude. However, this is not observed in OMAERUV. By



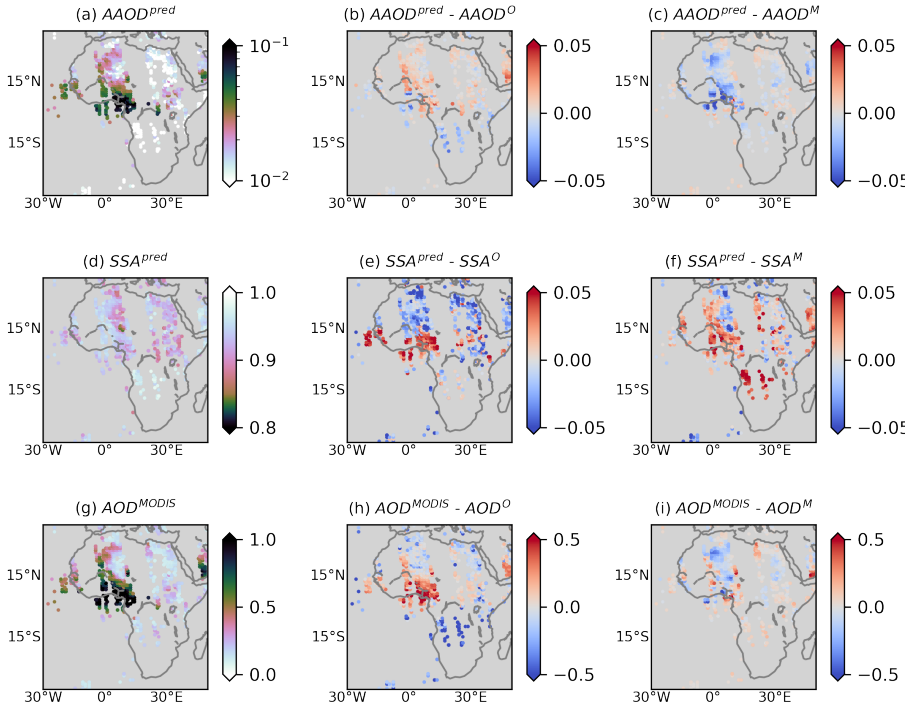
Figuur 5.17: Case study of 2019-02-14. The upper-script *O* and *M* indicate OMAERUV and MERRA-2 respectively. First row: DNN-predicted AAOD and its difference with the OMAERUV and MERRA-2 AAOD; second row: DNN-derived SSA and its difference with the OMAERUV and MERRA-2 SSA; third row: input UVAI, the MODIS AOD difference with OMAERUV, and MODIS AOD difference with MERRA-2.

contrast, the OMAERUV AAOD is higher in South America during all seasons, whereas the other two data only show significant AAOD during the fire season (SON). These disagreements are associated with the difference in AOD climatology (not shown).

In terms of SSA, the correlation between three data sets is not explicit. Because the OMAERUV and MERRA-2 SSA are calculated from independent a priori aerosol models, whereas the DNN derives SSA from observational data. The OMAERUV SSA shows the highest magnitude and the MERRA-2 SSA is lower than others, which agrees with findings in validation section. The spatial patterns of the DNN and MERRA-2 SSA are more similar to each other than that of OMAERUV over the Southern Africa and South America biomass burning regions, although the DNN predictions are slightly higher than that of MERRA-2.

GLOBAL AEROSOL ABSORPTION CLIMATOLOGY OVER OCEAN

In this section we provide a tentative analysis for the DNN-derived aerosol absorption climatology over ocean. Figure 5.23 and 5.24 present the AAOD and SSA, respectively. All AAOD data sets show the absorbing aerosol outflows from major continental sources, e.g. the dust storms from the Sahara Desert and smoke plumes from Central and Southern Africa. The OMAERUV climatology map is more noisy compared with others due to observational and retrieval errors, missing data, etc. There are high values of the OMAERUV AAOD over remote oceans which are far away from absorbing aerosol sources, especially in the Northern Pacific during MAM and in the Southern Pacific during SON 5.23 (e, k)).

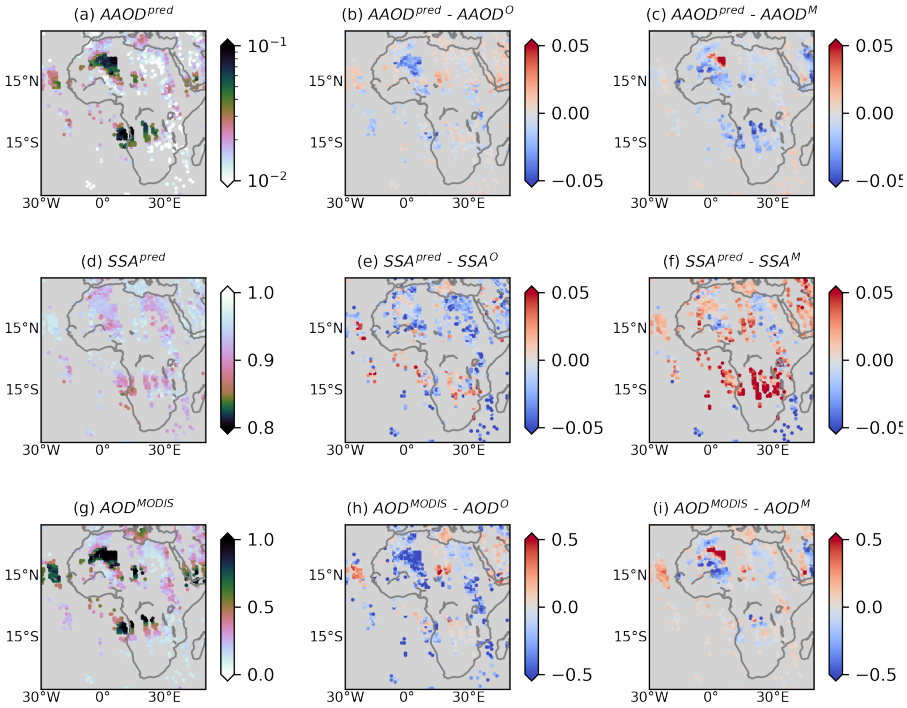


Figuur 5.18: The same as Figure 5.17 but for 2019-02-18.

The outflow of the Sahara dust and the Africa biomass burning plumes can be observed in the OMAERUV SSA climatology. The MERRA-2 SSA is also associated with the corresponding AAOD, particularly for smoke plumes. But the dust outflows over the northern Atlantic Ocean and the Arabian Sea can hardly be observed in the SSA maps. This may be caused by the aerosol models used in MERRA-2. The SSA of dust aerosols is between 0.77 to 0.96 (depending on particle size), whereas the absorption of black carbon is much stronger (SSA varies from 0.21 to 0.38, depending on relative humidity) [Randles et al., 2017]. On the other hand, the outflows of Africa dust and smoke are observable in the DNN-derived SSA map, but the magnitude is much less than other two data sets. The DNN predictions over ocean is overall not as good as that over land due to the limited spatial distribution of the training data.

REGIONAL AEROSOL ABSORPTION CLIMATOLOGY

We further select the major absorbing aerosol sources for regional analysis, as shown in Figure 5.25. We analyze the monthly averaged aerosol absorption time series of the selected regions of interest with the mean of the corresponding AERONET measurements as a reference. According to Figure 5.26a, the DNN-predicted AAOD and MERRA-2 AAOD show better agreeemnt with AERONET, whereas in terms of SSA (Figure 5.26b), the fluctuation of the AERONET SSA is larger than other data sets, and generally shows a stronger aerosol absorption in almost all regions. But note that the selected AERONET sites can hardly represent the whole features of the selected regions, thus we only provide statistical analysis between DNN predictions, OMAERUV and MERRA-2. All the three data sets show similar seasonal cycles of aerosol absorption. Overall, the DNN-predicted AAOD are better associated with the other two data sets than the DNN-predicted SSA.



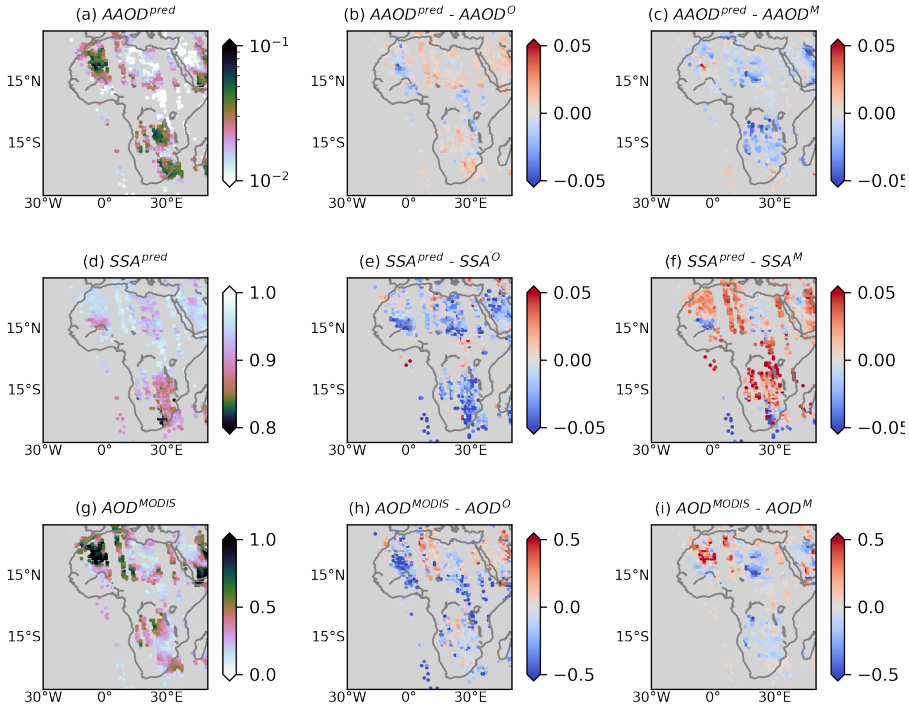
Figuur 5.19: The same as Figure 5.17 but for 2019-07-10.

In biomass burning regions, e.g. Southeast Asia, Central Africa and Southern Africa, the DNN-predicted AAOD is highly correlated with OMAERUV and MERRA-2 (R^2 >0.8), but the magnitude of the DNN predictions is more similar to the latter. An exception is South America, where the OMAERUV AAOD has the highest magnitude but with little seasonal variations, showing very low correlations with other two data sets. This is also reflected in the climatology maps (Figure 5.21). However, the corresponding SSA of DNN and MERRA-2 are in good agreement over this region (R^2 =0.93). The DNN-derived SSA is also well correlated with that of MERRA-2 in Africa biomass burning regions (R^2 =0.9), whereas lower correlations are found in Southeast Asia and the Amazonian regions (R^2 <0.8). The reason could be that the satellite observations are less sensitive to these regions as the aerosol loading is lower compared to that in Africa.

Compared with biomass burning regions, the AAOD and SSA of all data sets show less correlations in the desert regions of Arabia and Northern Africa. MERRA-2 presents the strongest absorption, whereas OMAERUV presents the lowest absorption, which is consistent with the findings in validation section. The DNN predictions are between the two data sets, as the low measurement sensitivity of AOD in the visible band due to high surface albedo is partially compensated by the advantage of UVAI that can detect the presence of absorbing aerosols even over bright surfaces [Herman et al., 1997a, Torres et al., 1998].

5.4. CONCLUSION

A global aerosol absorption database is important to reduce uncertainties of aerosol radiative forcing assessments. In this study, we introduced a DNN model to retrieve quantitative aerosol absorptive

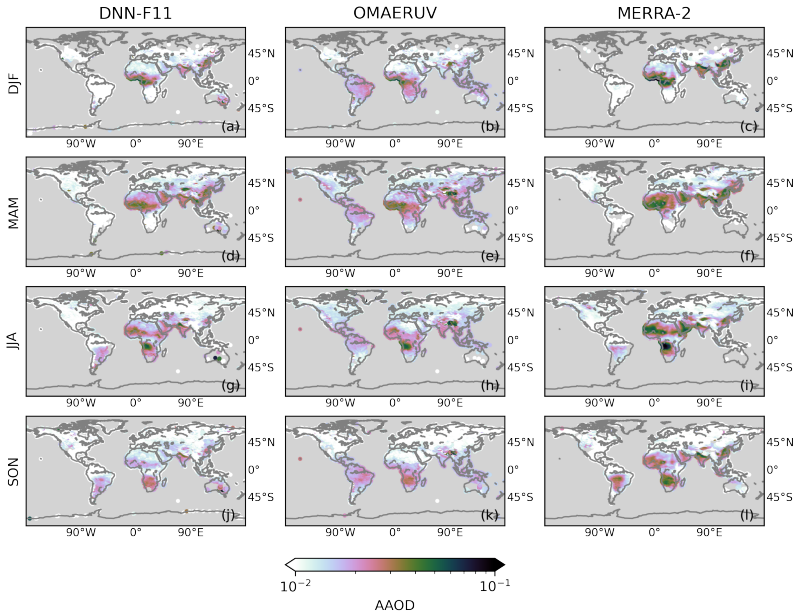


Figuur 5.20: The same as Figure 5.17 but for 2019-08-07.

properties (AAOD and SSA) from the long-term OMI UVAI product. The input features are selected by both filter and wrapper methods. The hyper-parameters of the DNN model are determined by grid search with 10-fold cross validation. The final trained model has an accuracy of 0.0045 for AAOD prediction, which is within the expected AAOD error (0.004 to 0.006).

The trained DNN model has been applied to predict aerosol absorption over land for the period between 2006 and 2019. It is recommended to use DNN-predicted AAOD and SSA with AOD at 550 nm above 0.1, as the low aerosol loading will lead to significant biases in the DNN predictions. The point-to-point validation shows that the DNN AAOD is negatively biased, but still highly associated with the AERONET retrievals (R^2=0.89 and RMSE=0.0050). There are 82% samples fall in the excepted uncertainty of the AERONET SSA (±0.03). The analysis based on aerosol types shows that our DNN model is better at predicting fine absorbing aerosols (e.g. smoke) rather than coarse ones (e.g. mineral dust). This is mainly caused by the property of the MODIS AOD in the visible band. The consistency is lowest for non-absorbing aerosols, since UVAI is only sensitive to absorbing aerosols according to its definition.

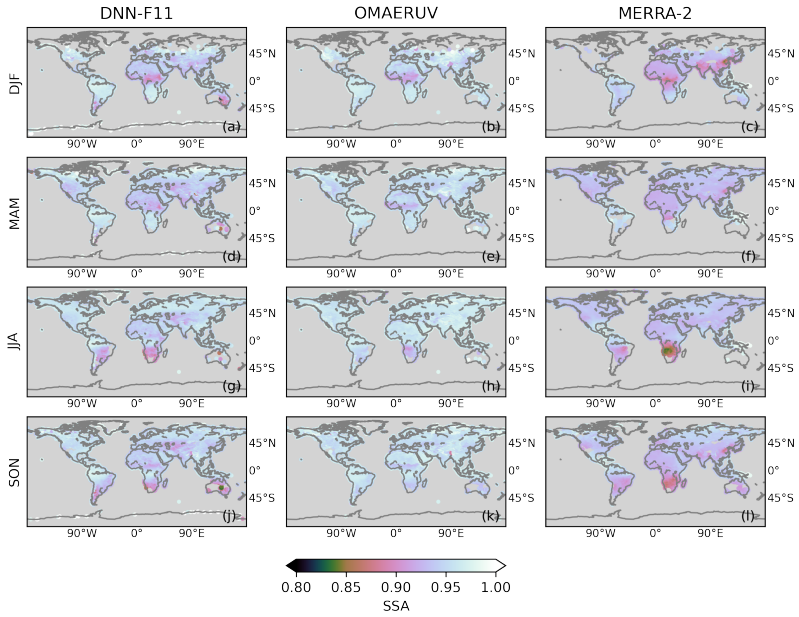
The DNN predictions are also compared with the OMAERUV retrievals and the MERRA-2 aerosol reanalysis. It should be noted that for both OMAERUV and MERRA-2, the SSA is calculated from a priori assumptions on aerosol micro-physical properties, whereas for no such constraints in our DNN model. Using AERONET as the reference, the DNN AAOD outperforms that of OMAERUV and MERRA-2 for all aerosol types. Specifically, OMAERUV underestimates the aerosol absorption particularly for absorbing aerosols, whereas MERRA-2 has the tendency to overestimate aerosol absorption in general. The AAOD spatial distribution and temporal variation of the DNN predictions and MERRA-2 are highly similar, particularly in the biomass burning regions. The spatial pattern of SSA is less comparable.



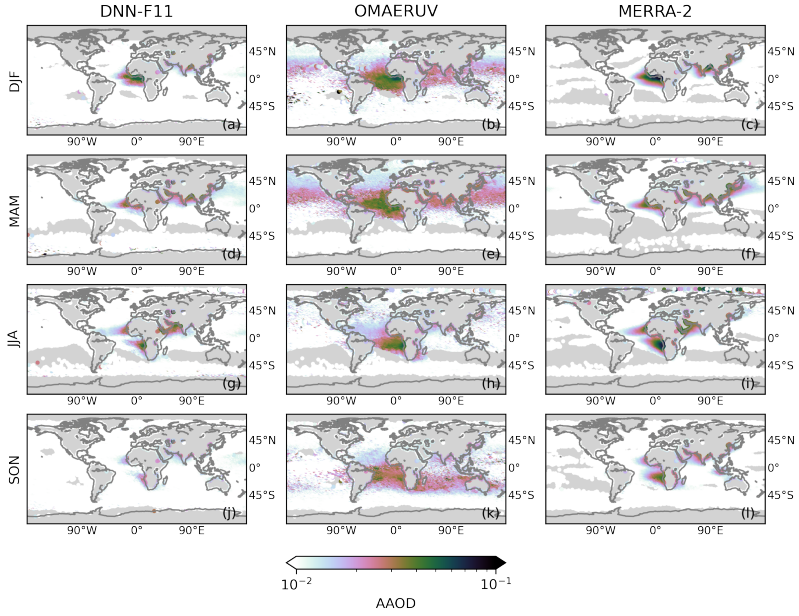
Figuur 5.21: Seasonal AAOD of DNN (left column), OMAERUV (middle column) and MERRA-2 (right column) over land.

Our DNN model shows encouraging results in aerosol absorptive properties retrieval over land. However, the result is only satisfying for AOD above a certain level, since the low measurement sensitivity to low aerosol loading is a general problem of aerosol remote sensing. Besides, the satellite retrieved AOD in the visible band is still challenging for dust particles over bright surfaces. The application to aerosols over ocean is still restricted by the availability of reliable aerosol absorption sources used for training Machine Learning models. Future applications should pay attention to the above aspects.

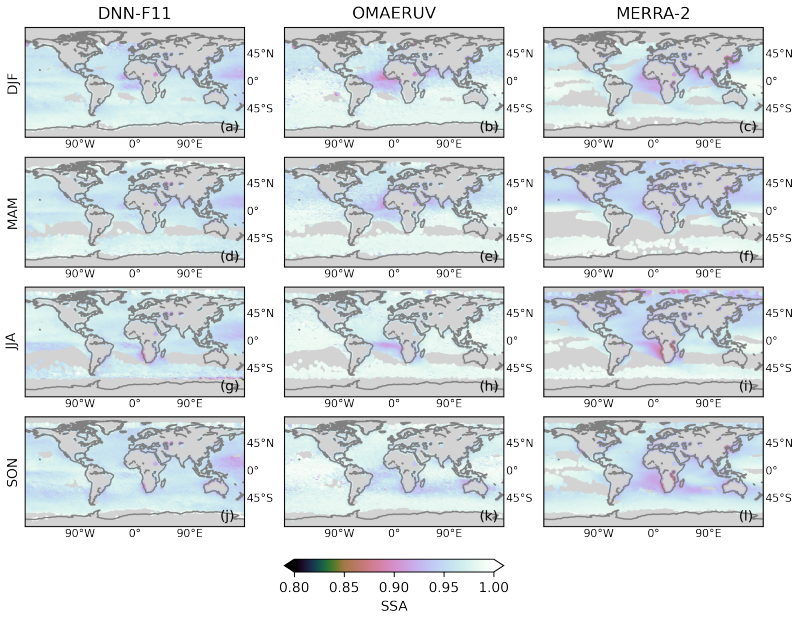
Currently, due to the limitation of the training data set (a hybrid data sets of satellite observations, model simulations and ground-based network measurements), our study did not properly account for spatial and temporal auto-correlation of geo-parameters. With more efforts put on the global monitoring of aerosol properties and vertical distribution, and better resolution of satellite instruments, it is expected to solve this problem either using geo-statistics that explicitly provide the spatial and temporal auto-correlation, or using advanced Deep Learning techniques such as CNN and/or RNN that can share weight parameters in space and/or time domain.



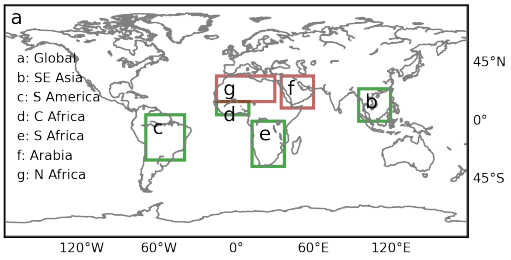
Figuur 5.22: Seasonal SSA of DNN (left column), OMAERUV (middle column) and MERRA-2 (right column) over land.



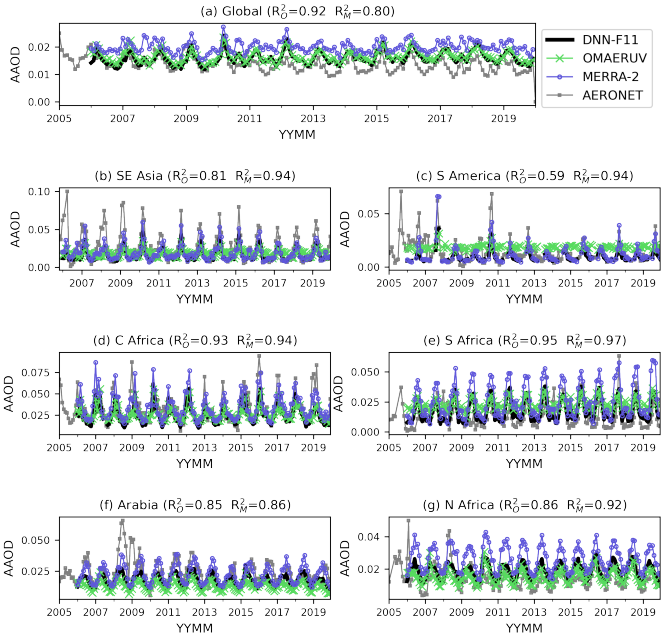
Figuur 5.23: Seasonal AAOD of DNN (left column), OMAERUV (middle column) and MERRA-2 (right column) over ocean.



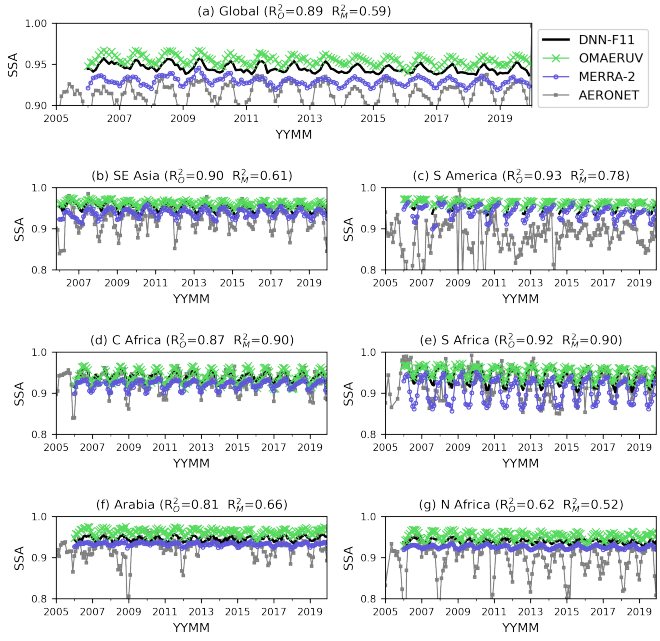
Figuur 5.24: Seasonal SSA of DNN (left column), OMAERUV (middle column) and MERRA-2 (right column) over ocean.



Figuur 5.25: Regions of interest. Green and red indicates regions dominated by smoke aerosols and dust aerosols, respectively.



(a) AAOD time series.



(b) SSA time series.

Figuur 5.26: Time series of DNN (black), OMAERUV (green), MERRA-2 (blue) and AERONET (gray) for regions of interest.

6

CONCLUSION AND OUTLOOK

Aerosol absorptive properties, e.g. the absorbing aerosol optical depth (AAOD) and the single scattering albedo (SSA), are one of the major uncertainties in aerosol radiative forcing assessments. Monitoring the global aerosol absorptive properties from space is therefore essential but challenging. Currently most satellite sensors are not equipped with multi-angular and/or polarimetric techniques, which can add constraints to the inversion process for retrieving aerosol optical properties in addition to the aerosol optical depth (AOD). Consequently, we turn to the radiance measurements in the Ultra-Violet (UV) band, from which a qualitative aerosol absorptive parameter, i.e. the Ultra-Violet Aerosol Index (UVAI), can be calculated.

UVAI is directly calculated from the measured radiance and the radiance simulated for a Rayleigh atmosphere. As such, the UVAI is free from the uncertainties due to a priori assumed aerosol properties. Many satellite sensors provide UVAI products, together contributing to a global database holding the information of aerosol absorption for over four decades (Table 1.1 in Chapter 1).

However, UVAI cannot be directly used in quantitative calculations, e.g. the assessment of aerosol radiative forcing. Hence, the objective of this thesis is to derive quantitative aerosol absorptive properties, i.e. AAOD and/or SSA from UVAI. We conducted a sensitivity study of UVAI, based on which we investigated the feasibility to derive AAOD and/or SSA using either physically-based or data-driven methods (Chapter 2 and 3). The lack of observations of the aerosol vertical distribution restricts the application of UVAI, therefore we created an aerosol layer height (ALH) data set from a chemical transport model (Chapter 4). Synthesizing the previous results, we derived an aerosol absorption database from the long-term OMI UVAI record (Chapter 5). The potential of the method is demonstrated by evaluating the results against independent observations.

6.1. CONCLUSIONS

The research questions addressed in this dissertation are listed in Section 1.8 of Chapter 1. Below, the main findings related to each of these research questions are presented.

1. WHAT ARE PHYSICAL PARAMETERS THAT UVAI IS MOST SENSITIVE TO?

Section 1.7.2 in Chapter 1, section 2.2 in Chapter 2, and section 4.1 in Chapter 4 present sensitivity studies of UVAI. UVAI is calculated from the radiance at the top of the atmosphere measured by satellite. Theoretically, in addition to aerosol related parameters, UVAI also depends on the satellite-solar geometries, the relative location to clouds, the surface albedo, the surface pressure, the choice of the wavelength pair, etc. Nevertheless, influences of these non-aerosol parameters are relatively limited. The feature selection in Chapter 3 section 3.3.2 and Chapter 5 section 5.2.3, where we applied Machine Learning algorithms to derive aerosol absorbing properties from UVAI, also confirms this conclusion.

UVAI is only sensitive to elevated absorbing aerosol layers with relatively high aerosol loading. In other words, the most important parameters to UVAI are:

- **Aerosol absorption**: UVAI was proposed for detecting the presence of absorbing aerosols. The stronger the absorption, the larger the detected UVAI. Aerosol absorption is usually described by SSA, which is usually determined by the complex refractive index and the size distribution function. However, as found in Chapter 3 and 5, compared with SSA, UVAI is better correlated with another quantitative aerosol absorptive parameter, i.e. AAOD, since both UVAI and AAOD contain information on aerosol absorption and aerosol amount.
- **Aerosol loading**: in the context of remote sensing, the aerosol amount is represented by the AOD. For elevated absorbing aerosols, the higher the AOD, the higher the UVAI value, and the AOD effect becomes stronger with altitude. But the AOD effect is not significant if aerosol layers are close to the ground. The above conclusions are not applicable for non-absorbing aerosols such as sea salt or sulfates.
- **Aerosol vertical distribution**: the UVAI signal is caused by absorbing the up-welling Rayleigh scattering. The higher an absorbing aerosol layer, the more scattered radiance beneath the aerosol layer is absorbed by the layer, resulting a higher UVAI. This UVAI altitude dependence also increases with aerosol absorption and aerosol loading. On the other hand, this vertical distribution effect on scattering aerosols is limited.

2. WHAT ARE THE MAJOR CHALLENGES/UNCERTAINTIES OF DERIVING QUANTITATIVE AEROSOL OPTICAL PROPERTIES USING TRADITIONAL RADIATIVE TRANSFER SIMULATIONS?

In Chapter 2 and 3, we derived the SSA from UVAI using radiative transfer simulations for several fire events. Specifically, a Look-Up Table (LUT) of UVAI is calculated as a function of SSA, and the SSA is calculated by the Mie Theory with given aerosol models. Then the SSA is estimated by minimizing the difference between the satellite measured and the simulated UVAI.

- One difficulty of this method is that assumptions have to be made on aerosol properties. The uncertainty is caused by insufficient knowledge of aerosol micro-physics, especially the spectral dependence of the imaginary refractive index. Moreover, the assumed aerosol models used in the LUT cannot represent the real aerosol properties that vary considerably in space and time.
- Another challenge is the lack of information on the aerosol vertical distribution. In radiative transfer simulations, it is possible to describe the complete aerosol profiles, whereas the calculation would be very time consuming. Alternatively, parameterized aerosol profiles are more often used. The most commonly used profiles are the exponential-decay profile, the Gaussian or log-normal profile, and the homogeneous box-shape profile. But no matter whether one uses real or parameterized profiles, there is little information of aerosol vertical distribution from measurements. This will be explained further in the next Question.

3. HOW TO OBTAIN AN AEROSOL VERTICAL DISTRIBUTION DATA SET FOR QUANTITATIVELY USING UVAI?

UVAI is sensitive to the aerosol vertical distribution, and its altitude dependence increases with aerosol absorption and aerosol loading. Without doubt, the aerosol vertical distribution is important for the quantitative use of UVAI.

Currently, the observational aerosol vertical distribution is mainly offered by two types of measurements: (1) aerosol extinction profiles from active lidar, and (2) aerosol heights from passive sensors. Whereas lidar measurements provide a great amount of details on aerosol profiles, they are subjected to limited spatial and/or temporal coverage. Compared with aerosol profiles, aerosol heights are easier for quantitative analysis, but most aerosol height retrieval algorithms are only applicable under certain conditions, e.g. elevated aerosol layers, over dark surfaces, cloud-free, etc.

Facing the lack of observations, we attempt to build a global aerosol height database based on the aerosol profiles of a well-validated chemistry transport model (i.e. the MERRA-2 aerosol reanalysis) for future applications of quantifying aerosol absorption from satellite UVAI. Since aerosol profiles are too complicated to analyze, it is more common using ALH to study the UVAI dependence on

the aerosol vertical distribution. But this comes to another issue that a widely accepted standard metric to convert aerosol profiles into a single height parameter (either an effective height or an actual height) is presently missing. Thus in Chapter 4, based on the relationship between UVAI and aerosol heights found in radiative transfer simulations and current satellite observations, we investigated 4 methods that convert aerosol profiles into aerosol heights, they are: (1) the mean height weighed by the extinction coefficient of each layer; (2) the mean height weighed by the aerosol optical depth of each layer; (3) the scale height where 63% of the columnar AOD is retained; and (4) the geometric height indicating the top boundary of aerosol layers.

We found that the UVAI dependence on 3 effective heights (derived by first 3 methods) is consistent with findings in the simulations and observations under majority circumstance, except when the aerosol loading is high. Since effective heights are sensitive to the the altitude and the magnitude of layers with the strongest extinctions, together with the fact that AOD is positively related to the peak extinction that is likely to appear at a lower altitude, the UVAI dependence on the effective heights is lower for the highest AOD regimes. On the other hand, the UVAI dependence on the top height matches best with the findings from sensitivity studies and satellite observations: when AOD increases, the sensitivity of UVAI to the changes of the top height becomes stronger. Moreover, the UVAI dependence on the top height also decreases with SSA, whereas UVAI shows little relation with the effective heights as SSA varies.

Thus, the answer to this research question is: currently there is no observational aerosol vertical distribution data set can be used to quantify aerosol absorptive properties from the long-term UVAI record. However, it is possible to create such a data set from a well-validated chemistry transport model, e.g. the top boundary height derived from the MERRA-2 aerosol reanalysis, as an alternative of limited aerosol height measurements. We have succeeded using this aerosol height data set to deriving global aerosol absorptive properties from the OMAERUV UVAI for the period from 2006 to 2019 with encouraging outputs compared with the AERONET measurements.

4. HOW TO APPLY MACHINE LEARNING TECHNIQUES TO DERIVE QUANTITATIVE AEROSOL OPTICAL PROPERTIES FROM UVAI?

Existing retrieval methods of aerosol properties (either using LUT or online radiative transfer simulations) cannot avoid assumptions on aerosol micro-physical properties in the forward radiative transfer simulations. This is one of the major error sources in the retrieved aerosol optical properties, since our knowledge on aerosol micro-physics is too limited to well present aerosols in reality. Fortunately, Machine Learning techniques provide an alternative solution for inversion problems of aerosol remote sensing by building a numerical relationship between variables learned from the given data without explicitly describing aerosol micro-physical properties.

In Chapter 3, we found that the AERONET AAOD and the OMI UVAI are well correlated. We conducted a tentative experiment using the Support Vector Regression (SVR) to derive AAOD and SSA from UVAI for several individual fire events. Compared with radiative transfer simulations (as that in Chapter 2 and 3), the SVR-predicted aerosol absorption shows a better agreement with the AERONET retrievals. But the SVR model suffers from a high bias problem as only three features (UVAI, AOD and ALH) are selected to train the model. Furthermore, the size of the training data set (less than 5,000) is significantly limited by the availability of ALH data.

Based on the previous experience, in Chapter 5, we applied a more comprehensive Machine Learning algorithm to retrieve the aerosol absorptive properties from the OMI UVAI for the period from 2006 to 2019. With the ALH derived from the MERRA-2 aerosol profiles (i.e. the geometric top boundary of an aerosol layer as defined in Chapter 4), the training data set extends to near 50,000. With training data of this size, it is more efficient to use Deep Learning algorithms, e.g. the Deep Neural Network (DNN). The selected features do not only include all the parameters that UVAI is sensitive to, but also take spatial and temporal information into account. The comparison with AERONET shows that the DNN-derived aerosol absorption is only slightly underestimated with an AAOD error level of 0.004 and over 80% samples fall inside AERONET SSA uncertainty range. The validation results show

that the DNN-predicted aerosol absorption outperforms that provided by OMI and MERRA-2, and the DNN is better at predicting aerosol properties for fine absorbing aerosols than for coarser ones.

The DNN-derived aerosol absorption shows encouraging results and the potential to produce a global daily aerosol absorption database from the long-term OMI UVAI record. However, our DNN model is currently restricted by the low satellite measurement sensitivity of low aerosol loading cases, the large uncertainty of AOD over bright surfaces in the visible band, the limited measurements of aerosol absorptive properties over ocean to train the model. These factors indicate directions for future improvements. Furthermore, due to the limitation of the training data set (a hybrid data sets of satellite observations, model simulations and ground-based network measurements), our study did not properly account for spatial and temporal auto-correlation of geo-parameters. With more efforts put on the global monitoring of aerosol properties and vertical distribution, and better resolution of satellite instruments, it is expected to solve this problem either using geo-statistics that explicitly provide the spatial and temporal auto-correlation, or using advanced Deep Learning techniques such as CNN and/or RNN that can share weight parameters in space and/or time domain.

6.2. OUTLOOK FOR FUTURE RESEARCH

UVAI is a treasure that holds information on global aerosol absorption over several decades. This dissertation presents encouraging results of deriving quantitative aerosol absorptive properties from the satellite UVAI record. It shows the potential of UVAI as an alternative source to complete global aerosol absorption monitoring. However, we are still facing challenges. Here we provide some suggestions and recommendations for future research on this topic.

HIGH UNCERTAINTIES OF AEROSOL OPTICAL DEPTH IN VISIBLE BAND OVER BRIGHT SURFACES

AOD is one of the most important parameters that determines the value of UVAI. UVAI is calculated by radiance in the UV band which is insensitive to surface albedo. However, the MODIS AOD used in this dissertation is retrieved in the visible band, under which it is difficult to separate aerosol contribution to the total reflected radiance from that of surface, particularly over bright surfaces, e.g. arid regions. This results in a lower accuracy of retrieved aerosol absorption for continental dust aerosols. AOD retrieved at shorter (e.g. Ultra-Violet) or longer wavelengths (e.g. Infrared channel) is less sensitive to the surface albedo. Joint use of AOD in different bands could be a potential solution.

AVAILABILITY OF THE AEROSOL VERTICAL DISTRIBUTION

The aerosol vertical distribution is another important parameter that UVAI depends on. Currently, despite the fact that many sensors try to retrieve aerosol profiles or heights, there hardly exists an individual product that can provide a consistent global vision of the aerosol vertical distribution. Combined use of current aerosol profile or height products is an option. Nevertheless, the consistency between different products should be investigated first. Another solution is deriving aerosol profiles from chemistry transport models. This can ensure the consistency of the data set, but these models should be carefully validated before use.

LIMITATIONS OF AERONET TO MACHINE LEARNING APPLICATIONS

Our Machine Learning method to derive quantitative aerosol absorptive properties from UVAI is also restricted by the availability of AERONET. AERONET is currently the most widely distributed ground-based remote sensing network monitoring aerosols. Although it is a global system, most AERONET sites providing inversion products are located at continents. In other word, the AERONET network is basically blind for aerosol micro-physics over ocean. Oceanic aerosols are generally assumed to be scattering, but continental pollutants can be transported over the ocean and lead to different optical properties. In this case, an alternative aerosol absorption data is needed, or a careful data transfer to the AERONET observations should be considered.

Moreover, the majority of AERONET sites are located in developed regions (America and Europe). There are only a few stations located at absorbing aerosol source regions, e.g. the Sahara desert and the Amazonian rain-forest. The uneven distribution also biases the training data. In other words, there are much more cases of urban and/or mixed aerosols to train a Machine Learning model, than cases of absorbing aerosols which are more important to aerosol radiative forcing. Data from satellites or models can compensate the insufficient observations, but it is also necessary to increase the number of ground-based sites in these regions.

CONCERNS ON MACHINE LEARNING APPLICATIONS IN AEROSOL REMOTE SENSING

Machine Learning brings new possibilities in solving remote sensing inversion problems. However, there are some concerns when using this kind of method. Here the most important ones are listed (but they are not limited to these):

First, Machine Learning algorithms only build numerical relations between variables, thereby ignoring the underlying physical processes. In order to pursue better physical interpretability of Machine Learning outputs, as a data scientist with a background in atmospheric science and remote sensing, one should bring his/her domain knowledge to constrain the numerical model. The interface where a data scientist teaches the model how to work is the feature selection and engineering.

Second, as our target variables are geo-parameters, the spatial and temporal auto-correlation should be taken into consideration. Providing geo-coordinate and date-time information, or information of past data and/or nearby pixels are simple and straightforward methods. There is also an increasing trend applying hybrid approaches in both Machine Learning and geo-statistics. The latter can explicitly describe the auto-correlation of geo-parameters in space and time. The advanced Deep Learning techniques such as CNN and/or RNN are also an alternative option, as they can share weight parameters in space and/or time domain.

The last is a general problem of Machine Learning when the training data set and test data set have different distributions. In other words, the model may execute 'extrapolation' if the test data falls outside the training data domain. The 'extrapolated' outputs can be easily detected for a model with less than 4 features by visualization, but it is not the case for models with a higher-dimensional feature space. Fortunately, some algorithms have been developed to detect the drift of feature distributions (e.g. Concept Drift Detection), which are worth adapting to future applications.

AEROSOL PROFILES VERSUS AEROSOL HEIGHTS

Previous research (also included in this dissertation) focused on the relationship between UVAI and the aerosol layer height. The advantage of the aerosol layer height is it is concise for quantitative analysis. Also, for radiative transfer simulations, using full aerosol profiles may slow down the computational speed. However, Machine Learning brings the possibility to use aerosol profiles directly: the aerosol optical property in each layer can be treated as an individual feature. In this way, one does not have to make choices between different aerosol height definitions or check the consistency between different aerosol height products.

TOWARDS A LONG-TERM SYNTHETIC AEROSOL ABSORPTION DATABASE

UVAI has a long-term record since 1978, contributed by platforms listed in Table 1.1 in Chapter 1. It would be of great value to build a synthesized aerosol absorption database from the long-term UVAI. Since the wavelength pairs, the radiative transfer simulations, and the instrumental characteristics vary with different UVAI products, and UVAI cannot be directly used for quantitative calculations, direct assimilation of different UVAI products into one data set can lead to errors. Instead, it may be more reasonable to derive AAOD and SSA from each UVAI product first using separate Machine Learning models, then synthesize the quantitative absorptive properties into a complete data set. To ensure the consistency between different products, the AOD and aerosol vertical distribution used for quantifying different UVAI should be as consistent as possible.

BIBLIOGRAFIE

Peter J. Adams, John H. Seinfeld, Dorothy Koch, Loretta Mickley, and Daniel Jacob. General circulation model assessment of direct radiative forcing by the sulfate-nitrate-ammonium-water inorganic aerosol system. *Journal of Geophysical Research: Atmospheres*, 106(D1):1097–1111, 2001. doi: 10.1029/2000JD900512. URL https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2000JD900512.

Changwoo Ahn, Omar Torres, and Hiren Jethva. Assessment of omi near-uv aerosol optical depth over land. *Journal of Geophysical Research: Atmospheres*, 119(5):2457–2473, 2014. doi: 10.1002/2013JD020188. URL https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1002/2013JD020188.

Bruce A Albrecht. Aerosols, cloud microphysics, and fractional cloudiness. *Science*, 245(4923):1227–1230, 1989.

Abdullah Alfadda, Saifur Rahman, and Manisa Pipattanasomporn. Solar irradiance forecast using aerosols measurements: A data driven approach. *Solar Energy*, 170:924 – 939, 2018. ISSN 0038-092X. doi: https://doi.org/10.1016/j.solener.2018.05.089. URL http://www.sciencedirect.com/science/article/pii/S0038092X18305309.

P. Alpert and E. Ganor. Sahara mineral dust measurements from toms: Comparison to surface observations over the middle east for the extreme dust storm, march 14–17, 1998. *Journal of Geophysical Research: Atmospheres*, 106(D16):18275–18286, 2001. doi: 10.1029/2000JD900366. URL https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2000JD900366.

Steven Q Andrews. Inconsistencies in air quality metrics: 'blue sky' days and PM10 concentrations in beijing. *Environmental Research Letters*, 3(3):034009, jul 2008. doi: 10.1088/1748-9326/3/3/034009. URL https://doi.org/10.1088%2F1748-9326%2F3%2F3%2F034009.

Albert Ansmann and Detlef Müller. *Lidar and Atmospheric Aerosol Particles*, pages 105–141. Springer New York, New York, NY, 2005. ISBN 978-0-387-25101-1. doi: 10.1007/0-387-25101-4_4. URL https://doi.org/10.1007/0-387-25101-4_4.

P. Barber and C. Yeh. Scattering of electromagnetic waves by arbitrarily shaped dielectric bodies. *Appl. Opt.*, 14(12):2864–2872, Dec 1975. doi: 10.1364/AO.14.002864. URL http://ao.osa.org/abstract.cfm?URI=ao-14-12-2864.

S. E. Bauer, D. Koch, N. Unger, S. M. Metzger, D. T. Shindell, and D. G. Streets. Nitrate aerosols today and in 2030: a global simulation including aerosols and tropospheric ozone. *Atmospheric Chemistry and Physics*, 7(19):5043–5059, October 2007a. URL https://hal.archives-ouvertes.fr/hal-00296344.

S. E. Bauer, M. I. Mishchenko, A. A. Lacis, S. Zhang, J. Perlwitz, and S. M. Metzger. Do sulfate and nitrate coatings on mineral dust have important effects on radiative properties and climate modeling? *Journal of Geophysical Research: Atmospheres*, 112(D6), 2007b. doi: 10.1029/2005JD006977. URL https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2005JD006977.

Nicolas Bellouin, Jamie Rae, Andy Jones, Colin Johnson, Jim Haywood, and Olivier Boucher. Aerosol forcing in the climate model intercomparison project (cmip5) simulations by hadgem2-es and the role of ammonium nitrate. *Journal of Geophysical Research: Atmospheres*, 116(D20), 2011. doi: 10.1029/2011JD016074. URL https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2011JD016074.

J. A. Benediktsson, P. H. Swain, and O. K. Ersoy. Neural network approaches versus statistical methods in classification of multisource remote sensing data. *IEEE Transactions on Geoscience and Remote Sensing*, 28(4):540–552, July 1990. ISSN 1558-0644. doi: 10.1109/TGRS.1990.572944.

V. V. Berdnik and G. I. Gallyamova. Calculation of the reflectance and transmittance of a disperse layer by the neural network method. *Optics and Spectroscopy*, 112(4):618–623, 2012. doi: 10.1134/S0030400X12040078. URL https://doi.org/10.1134/S0030400X12040078.

Vladimir V. Berdnik and Valery A. Loiko. Retrieval of size and refractive index of spherical particles by multiangle light scattering: neural network method application. *Appl. Opt.*, 48(32):6178–6187, Nov 2009. doi: 10.1364/AO.48.006178. URL http://ao.osa.org/abstract.cfm?URI=ao-48-32-6178.

Vladimir V. Berdnik and Valery A. Loiko. *Neural networks for particle parameter retrieval by multi-angle light scattering*, pages 291–340. Springer Berlin Heidelberg, Berlin, Heidelberg, 2016. ISBN 978-3-662-46762-6. doi: 10.1007/978-3-662-46762-6_7. URL https://doi.org/10.1007/978-3-662-46762-6_7.

Vladimir V. Berdnik, Robert D. Mukhamedyarov, and Valery A. Loiko. Sizing of soft spheroidal particles by multiangle scattered light intensity data: application of neural networks. *Journal of Quantitative Spectroscopy and Radiative Transfer*, 89(1):279 – 289, 2004. ISSN 0022-4073. doi: https://doi.org/10.1016/j.jqsrt.2004.05.028. URL http://www.sciencedirect.com/science/article/pii/S0022407304001864. VII Electromagnetic and Light Scattering by Non-Spherical Particles: Theory, Measurement, and Applications.

R. W. Bergstrom, P. Pilewskie, P. B. Russell, J. Redemann, T. C. Bond, P. K. Quinn, and B. Sierau. Spectral absorption properties of atmospheric aerosols. *Atmospheric Chemistry and Physics*, 7(23):5937–5943, 2007. doi: 10.5194/acp-7-5937-2007. URL https://acp.copernicus.org/articles/7/5937/2007/.

JL Bezy, S Delwart, and M Rast. Meris-a new generation of ocean-colour sensor onboard envisat. *ESA bulletin*, 103:48–56, 2000.

H. Bian, M. Chin, D. A. Hauglustaine, M. Schulz, G. Myhre, S. E. Bauer, M. T. Lund, V. A. Karydis, T. L. Kucsera, X. Pan, A. Pozzer, R. B. Skeie, S. D. Steenrod, K. Sudo, K. Tsigaridis, A. P. Tsimpidi, and S. G. Tsyro. Investigation of global particulate nitrate from the aerocom phase iii experiment. *Atmospheric Chemistry and Physics*, 17(21):12911–12940, 2017. doi: 10.5194/acp-17-12911-2017. URL https://acp.copernicus.org/articles/17/12911/2017/.

Fabio Biancofiore, Marco Verdecchia, Piero Di Carlo, Barbara Tomassetti, Eleonora Aruffo, Marcella Busilacchio, Sebastiano Bianco, Sinibaldo Di Tommaso, and Carlo Colangeli. Analysis of surface ozone using a recurrent neural network. *Science of The Total Environment*, 514:379–387, 2015. doi: https://doi.org/10.1016/j.scitotenv.2015.01.106. URL https://www.sciencedirect.com/science/article/pii/S004896971500128X.

Eyk Boesche, Piet Stammes, Thomas Ruhtz, Réne Preusker, and Juergen Fischer. Effect of aerosol microphysical properties on polarization of skylight: sensitivity study and measurements. *Appl. Opt.*, 45(34):8790–8805, Dec 2006. doi: 10.1364/AO.45.008790. URL http://ao.osa.org/abstract.cfm?URI=ao-45-34-8790.

Eyk Boesche, Piet Stammes, and Ralf Bennartz. Aerosol influence on polarization and intensity in near-infrared o2 and co2 absorption bands observed from space. *Journal of Quantitative Spectroscopy and Radiative Transfer*, 110(3):223 – 239, 2009. ISSN 0022-4073. doi: https://doi.org/10.1016/j.jqsrt.2008.09.019. URL http://www.sciencedirect.com/science/article/pii/S0022407308002008.

T. C. Bond, S. J. Doherty, D. W. Fahey, P. M. Forster, T. Berntsen, B. J. DeAngelo, M. G. Flanner, S. Ghan, B. Kärcher, D. Koch, S. Kinne, Y. Kondo, P. K. Quinn, M. C. Sarofim, M. G. Schultz, M. Schulz, C. Venkataraman, H. Zhang, S. Zhang, N. Bellouin, S. K. Guttikunda, P. K. Hopke, M. Z. Jacobson, J. W. Kaiser, Z. Klimont, U. Lohmann, J. P. Schwarz, D. Shindell, T. Storelvmo, S. G. Warren, and C. S. Zender. Bounding the role of black carbon in the climate system: A scientific assessment. *Journal of Geophysical Research: Atmospheres*, 118(11):5380–5552, 2013. doi: 10.1002/jgrd.50171. URL https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1002/jgrd.50171.

Tami C. Bond and Robert W. Bergstrom. Light absorption by carbonaceous particles: An investigative review. *Aerosol Science and Technology*, 40(1):27–67, 2006. doi: 10.1080/02786820500421521. URL https://doi.org/10.1080/02786820500421521.

Tami C. Bond, David G. Streets, Kristen F. Yarber, Sibyl M. Nelson, Jung-Hun Woo, and Zbigniew Klimont. A technology-based global inventory of black and organic carbon emissions from combustion. *Journal of Geophysical Research: Atmospheres*, 109(D14), 2004. doi: 10.1029/2003JD003697. URL https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2003JD003697.

Olivier Boucher, David Randall, Paulo Artaxo, Christopher Bretherton, Gragam Feingold, Piers Forster, V-M Kerminen, Yutaka Kondo, Hong Liao, Ulrike Lohmann, et al. Clouds and aerosols. In *Climate change 2013: the physical science basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*, pages 571–657. Cambridge University Press, 2013.

A. F. Bouwman, D. S. Lee, W. A. H. Asman, F. J. Dentener, K. W. Van Der Hoek, and J. G. J. Olivier. A global high-resolution emission inventory for ammonia. *Global Biogeochemical Cycles*, 11(4):561–587, 1997. doi: 10.1029/97GB02266. URL https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/97GB02266.

Alessio Bozzo, S. Remy, Angela Benedetti, Johannes Flemming, Peter Bechtold, Mark Rodwell, and J.-J. Morcrette. Implementation of a cams-based aerosol climatology in the ifs. (801), 2017. doi: 10.21957/84ya94mls. URL https://www.ecmwf.int/node/17219.

Charlie S. Bristow, Karen A. Hudson-Edwards, and Adrian Chappell. Fertilizing the amazon and equatorial atlantic with west african dust. *Geophysical Research Letters*, 37(14), 2010. doi: 10.1029/2010GL043486. URL https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2010GL043486.

Francois-Marie Bréon, Anne Vermeulen, and Jacques Descloitres. An evaluation of satellite aerosol products against sunphotometer measurements. *Remote Sensing of Environment*, 115(12):3102 – 3111, 2011. ISSN 0034-4257. doi: https://doi.org/10.1016/j.rse.2011.06.017. URL http://www.sciencedirect.com/science/article/pii/S0034425711002410.

V. Buchard, A. M. da Silva, P. R. Colarco, A. Darmenov, C. A. Randles, R. Govindaraju, O. Torres, J. Campbell, and R. Spurr. Using the omi aerosol index and absorption aerosol optical depth to evaluate the nasa merra aerosol reanalysis. *Atmospheric Chemistry and Physics*, 15(10):5743–5760, 2015. doi: 10.5194/acp-15-5743-2015. URL https://acp.copernicus.org/articles/15/5743/2015/.

V. Buchard, C. A. Randles, A. M. da Silva, A. Darmenov, P. R. Colarco, R. Govindaraju, R. Ferrare, J. Hair, A. J. Beyersdorf, L. D. Ziemba, and H. Yu. The MERRA-2 Aerosol Reanalysis, 1980 Onward. Part II: Evaluation and Case Studies. *Journal of Climate*, 30(17):6851–6872, 07 2017. ISSN 0894-8755. doi: 10.1175/JCLI-D-16-0613.1. URL https://doi.org/10.1175/JCLI-D-16-0613.1.

J Burrows, A. Goede, Christian Muller, and H Bovensmann. Sciamachy - exploring the changing earth's atmosphere. 01 2011. doi: 10.1007/978-90-481-9896-2_1.

KS Carslaw, LA Lee, CL Reddington, KJ Pringle, A Rap, PM Forster, GW Mann, DV Spracklen, MT Woodhouse, LA Regayre, et al. Large contribution of natural aerosols to uncertainty in indirect forcing. *Nature*, 503(7474):67–71, 2013. doi: 10.1038/nature12674.

A. Cazorla, R. Bahadur, K. J. Suski, J. F. Cahill, D. Chand, B. Schmid, V. Ramanathan, and K. A. Prather. Relating aerosol absorption due to soot, organic carbon, and dust to emission sources determined from in-situ chemical measurements. *Atmospheric Chemistry and Physics*, 13(18):9337–9350, 2013. doi: 10.5194/acp-13-9337-2013. URL https://acp.copernicus.org/articles/13/9337/2013/.

C. Chen, O. Dubovik, D. K. Henze, T. Lapyonak, M. Chin, F. Ducos, P. Litvinov, X. Huang, and L. Li. Retrieval of desert dust and carbonaceous aerosol emissions over africa from polder/parasol products generated by the grasp algorithm. *Atmospheric Chemistry and Physics*, 18(16):12551–12580, 2018. doi: 10.5194/acp-18-12551-2018. URL https://acp.copernicus.org/articles/18/12551/2018/.

Wei Chen, Shuang Bai, Dexuan Wang, Haimeng Zhao, Hao Sun, Lina Yi, Hengqian Zhao, Donghui Xie, Jouni Peltoniemi, and Zhanqing Li. Aerosol-induced changes in sky polarization pattern: potential hint on applications in polarimetric remote sensing. *International Journal of Remote Sensing*, 41(13):4963–4980, 2020a. doi: 10.1080/01431161.2019.1685724. URL https://doi.org/10.1080/01431161.2019.1685724.

Xingfeng Chen, Gerrit de Leeuw, Antti Arola, Shumin Liu, Yang Liu, Zhengqiang Li, and Kainan Zhang. Joint retrieval of the aerosol fine mode fraction and optical depth using modis spectral reflectance over northern and eastern china: Artificial neural network method. *Remote Sensing of Environment*, 249:112006, 2020b. doi: https://doi.org/10.1016/j.rse.2020.112006. URL https://www.sciencedirect.com/science/article/pii/S003442572030376X.

Tianhai Cheng, Yu Wu, and Hao Chen. Effects of morphology on the radiative properties of internally mixed light absorbing carbon aerosols with different aging status. *Opt. Express*, 22(13):15904–15917, Jun 2014. doi: 10.1364/OE.22.015904. URL http://www.opticsexpress.org/abstract.cfm?URI=oe-22-13-15904.

Vladimir Cherkassky and Yunqian Ma. Practical selection of svm parameters and noise estimation for svm regression. *Neural Networks*, 17(1):113 – 126, 2004. ISSN 0893-6080. doi: https://doi.org/10.1016/S0893-6080(03)00169-2. URL http://www.sciencedirect.com/science/article/pii/S0893608003001692.

I. Chiapello, J. M. Prospero, J. R. Herman, and N. C. Hsu. Detection of mineral dust over the north atlantic ocean and africa with the nimbus 7 toms. *Journal of Geophysical Research: Atmospheres*, 104(D8):9277–9291, 1999. doi: 10.1029/1998JD200083. URL https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/1998JD200083.

J. Chimot, J. P. Veefkind, T. Vlemmix, J. F. de Haan, V. Amiridis, E. Proestakis, E. Marinou, and P. F. Levelt. An exploratory study on the aerosol height retrieval from omi measurements of the 477 nm $o_2 - o_2$ spectral band using a neural network approach. *Atmospheric Measurement Techniques*, 10(3):783–809, 2017. doi: 10.5194/amt-10-783-2017. URL https://amt.copernicus.org/articles/10/783/2017/.

J. Chimot, J. P. Veefkind, T. Vlemmix, and P. F. Levelt. Spatial distribution analysis of the omi aerosol layer height: a pixel-by-pixel comparison to caliop observations. *Atmospheric Measurement Techniques*, 11(4):2257–2277, 2018. doi: 10.5194/amt-11-2257-2018. URL https://amt.copernicus.org/articles/11/2257/2018/.

Jacek Chowdhary, Brian Cairns, and Larry D. Travis. Case Studies of Aerosol Retrievals over the Ocean from Multiangle, Multispectral Photopolarimetric Remote Sensing Data. *Journal of the Atmospheric Sciences*, 59(3):383–397, 02 2002.

D. A. Chu, Y. J. Kaufman, C. Ichoku, L. A. Remer, D. Tanré, and B. N Holben. Validation of modis aerosol optical depth retrieval over land. *Geophysical Research Letters*, 29(12):MOD2–1–MOD2–4, 2002. doi: 10.1029/2001GL013205. URL https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2001GL013205.

Serena H. Chung and John H. Seinfeld. Global distribution and climate forcing of carbonaceous aerosols. *Journal of Geophysical Research: Atmospheres*, 107(D19):AAC 14–1–AAC 14–33, 2002. doi: 10.1029/2001JD001397. URL https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2001JD001397.

P. R. Colarco, S. Gassó, C. Ahn, V. Buchard, A. M. da Silva, and O. Torres. Simulation of the ozone monitoring instrument aerosol index using the nasa goddard earth observing system aerosol reanalysis products. *Atmospheric Measurement Techniques*, 10(11):4121–4134, 2017. doi: 10.5194/amt-10-4121-2017. URL https://amt.copernicus.org/articles/10/4121/2017/.

Peter R. Colarco, Owen B. Toon, Omar Torres, and Philip J. Rasch. Determining the uv imaginary index of refraction of saharan dust particles from total ozone mapping spectrometer data using a three-dimensional model of dust transport. *Journal of Geophysical Research: Atmospheres*, 107(D16):AAC 4–1–AAC 4–18, 2002. doi: 10.1029/2001JD000903. URL https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2001JD000903.

C. A. Corr, N. Krotkov, S. Madronich, J. R. Slusser, B. Holben, W. Gao, J. Flynn, B. Lefer, and S. M. Kreidenweis. Retrieval of aerosol single scattering albedo at ultraviolet wavelengths at the t1 site during milagro. *Atmospheric Chemistry and Physics*, 9(15):5813–5827, 2009. doi: 10.5194/acp-9-5813-2009. URL https://acp.copernicus.org/articles/9/5813/2009/.

Corinna Cortes and Vladimir Vapnik. Support-vector networks. *Machine Learning*, 20(3):273–297, 1995. doi: 10.1007/BF00994018. URL https://doi.org/10.1007/BF00994018.

R. L. Curier, J. P. Veefkind, R. Braak, B. Veihelmann, O. Torres, and G. de Leeuw. Retrieval of aerosol optical properties from omi radiances using a multiwavelength algorithm: Application to western europe. *Journal of Geophysical Research: Atmospheres*, 113(D17), 2008. doi: 10.1029/2007JD008738. URL https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2007JD008738.

M. Dall'Osto, R. M. Harrison, H. Coe, P. I. Williams, and J. D. Allan. Real time chemical characterization of local and regional nitrate aerosols. *Atmospheric Chemistry and Physics*, 9(11):3709–3720, 2009. doi: 10.5194/acp-9-3709-2009. URL https://acp.copernicus.org/articles/9/3709/2009/.

J. V. Dave. Effect of Aerosols on the Estimation of Total Ozone in an Atmospheric Column from the Measurements of Its Ultraviolet Radiance. *Journal of the Atmospheric Sciences*, 35(5):899–911, 05 1978. ISSN 0022-4928. doi: 10.1175/1520-0469(1978)035<0899:EOAOTE>2.0.CO;2. URL https://doi.org/10.1175/1520-0469(1978)035<0899:EOAOTE>2.0.CO;2.

Cliff I. Davidson, Robert F. Phalen, and Paul A. Solomon. Airborne particulate matter and human health: A review. *Aerosol Science and Technology*, 39(8):737–749, 2005. doi: 10.1080/02786820500191348. URL https://doi.org/10.1080/02786820500191348.

M. de Graaf, P. Stammes, O. Torres, and R. B. A. Koelemeijer. Absorbing aerosol index: Sensitivity analysis, application to gome and comparison with toms. *Journal of Geophysical Research: Atmospheres*, 110(D1), 2005. doi: 10.1029/2004JD005178. URL https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2004JD005178.

J. F. de Haan. Disamar algorithm description and background information, 2011.

Johan Dehaan, P. Bosma, and J. Hovenier. The adding method for multiple scattering of polarized light. *Astronomy and Astrophysics*, 183:371–391, 08 1987.

P-Y Deschamps, F-M Bréon, Marc Leroy, Alain Podaire, Annick Bricaud, J-C Buriez, and Genevieve Seze. The polder mission: Instrument characteristics and scientific objectives. *IEEE Transactions on geoscience and remote sensing*, 32(3):598–615, 1994.

J. L. Deuzé, P. Goloub, M. Herman, A. Marchand, G. Perry, S. Susana, and D. Tanré. Estimate of the aerosol properties over the ocean with polder. *Journal of Geophysical Research: Atmospheres*, 105(D12):15329–15346, 2000. doi: 10.1029/2000JD900148. URL https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2000JD900148.

J. L. Deuzé, F. M. Bréon, C. Devaux, P. Goloub, M. Herman, B. Lafrance, F. Maignan, A. Marchand, F. Nadal, G. Perry, and D. Tanré. Remote sensing of aerosols over land surfaces from polder-adeos-1 polarized measurements. *Journal of Geophysical Research: Atmospheres*, 106(D5):4913–4926, 2001. doi: 10.1029/2000JD900364. URL https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2000JD900364.

Antonio Di Noia and Otto P. Hasekamp. *Neural Networks and Support Vector Machines and Their Application to Aerosol and Cloud Remote Sensing: A Review*, pages 279–329. Springer International Publishing, Cham, 2018. ISBN 978-3-319-70796-9. doi: 10.1007/978-3-319-70796-9_4. URL https://doi.org/10.1007/978-3-319-70796-9_4.

D. J. Diner, J. C. Beckert, T. H. Reilly, C. J. Bruegge, J. E. Conel, R. A. Kahn, J. V. Martonchik, T. P. Ackerman, R. Davies, S. A. W. Gerstl, H. R. Gordon, J. . Muller, R. B. Myneni, P. J. Sellers, B. Pinty, and M. M. Verstraete. Multi-angle imaging spectroradiometer (misr) instrument description and experiment overview. *IEEE Transactions on Geoscience and Remote Sensing*, 36(4):1072–1087, July 1998. ISSN 1558-0644. doi: 10.1109/36.700992.

David J. Diner, Carol J. Bruegge, John V. Martonchik, Graham W. Bothwell, Eric D. Danielson, Elmer L. Floyd, Virginia G. Ford, Larry E. Hovland, Kenneth L. Jones, and Mary L. White. A multiangle imaging spectroradiometer for terrestrial remote sensing from the earth observing system. *International Journal of Imaging Systems and Technology*, 3(2):92–107, 1991. doi: 10.1002/ima.1850030206. URL https://onlinelibrary.wiley.com/doi/abs/10.1002/ima.1850030206.

David J. Diner, Stacey W. Boland, Michael Brauer, Carol Bruegge, Kevin A. Burke, Russell Chipman, Larry Di Girolamo, Michael J. Garay, Sina Hasheminassab, Edward Hyer, Michael Jerrett, Veljko Jovanovic, Olga V. Kalashnikova, Yang Liu, Alexei I. Lyapustin, Randall V. Martin, Abigail Nastan, Bart D. Ostro, Beate Ritz, Joel Schwartz, Jun Wang, and Feng Xu. Advances in multiangle satellite remote sensing of speciated airborne particulate matter and association with adverse health effects: from MISR to MAIA. *Journal of Applied Remote Sensing*, 12(4):1 – 22, 2018. doi: 10.1117/1.JRS.12.042603. URL https://doi.org/10.1117/1.JRS.12.042603.

S. Ding, J. Wang, and X. Xu. 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, 2016. doi: 10.5194/amt-9-2077-2016. URL https://amt.copernicus.org/articles/9/2077/2016/.

Y. Dong, E. Spinei, and A. Karpatne. A feasibility study to use machine learning as an inversion algorithm for aerosol profile and property retrieval from multi-axis differential absorption spectroscopy measurements. *Atmospheric Measurement Techniques*, 13(10):5537–5550, 2020. doi: 10.5194/amt-13-5537-2020. URL https://amt.copernicus.org/articles/13/5537/2020/.

R. Donida Labati, A. Genovese, V. Piuri, and F. Scotti. Wildfire smoke detection using computational intelligence techniques enhanced with synthetic smoke plume generation. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 43(4):1003–1012, July 2013. ISSN 2168-2232. doi: 10.1109/TSMCA.2012.2224335.

Jesper Sören Dramsch. 70 years of machine learning in geoscience in review. *Advances in Geophysics*, 61:1–55, 2020. doi: 10.1016/bs.agph.2020.08.002. URL https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7500415/.

Harris Drucker, Christopher JC Burges, Linda Kaufman, Alex J Smola, and Vladimir Vapnik. Support vector regression machines. In *Advances in neural information processing systems*, pages 155–161, 1997.

T. Drugé, P. Nabat, M. Mallet, and S. Somot. Model simulation of ammonium and nitrate aerosols distribution in the euro-mediterranean region and their radiative and climatic effects over 1979–2016. *Atmospheric Chemistry and Physics*, 19(6):3707–3731, 2019. doi: 10.5194/acp-19-3707-2019. URL https://acp.copernicus.org/articles/19/3707/2019/.

O. Dubovik, B. N. Holben, Y. J. Kaufman, M. Yamasoe, A. Smirnov, D. Tanré, and I. Slutsker. Single-scattering albedo of smoke retrieved from the sky radiance and solar transmittance measured from ground. *Journal of Geophysical Research: Atmospheres*, 103(D24):31903–31923, 1998. doi: 10.1029/98JD02276. URL https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/98JD02276.

O. Dubovik, A. Smirnov, B. N. Holben, M. D. King, Y. J. Kaufman, T. F. Eck, and I. Slutsker. Accuracy assessments of aerosol optical properties retrieved from aerosol robotic network (aeronet) sun and sky radiance measurements. *Journal of Geophysical Research: Atmospheres*, 105(D8):9791–9806, 2000. doi: 10.1029/2000JD900040. URL https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2000JD900040.

O. Dubovik, B. N. Holben, T. Lapyonok, A. Sinyuk, M. I. Mishchenko, P. Yang, and I. Slutsker. Non-spherical aerosol retrieval method employing light scattering by spheroids. *Geophysical Research Letters*, 29(10):54–1–54–4, 2002a. doi: 10.1029/2001GL014506. URL https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2001GL014506.

O. Dubovik, M. Herman, A. Holdak, T. Lapyonok, D. Tanré, J. L. Deuzé, F. Ducos, A. Sinyuk, and A. Lopatin. Statistically optimized inversion algorithm for enhanced retrieval of aerosol properties from spectral multi-angle polarimetric satellite observations. *Atmospheric Measurement Techniques*, 4(5):975–1018, 2011. doi: 10.5194/amt-4-975-2011. URL https://amt.copernicus.org/articles/4/975/2011/.

Oleg Dubovik and Michael D. King. A flexible inversion algorithm for retrieval of aerosol optical properties from sun and sky radiance measurements. *Journal of Geophysical Research: Atmospheres*, 105(D16):20673–20696, 2000. doi: 10.1029/2000JD900282. URL https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2000JD900282.

Oleg Dubovik, Brent Holben, Thomas F. Eck, Alexander Smirnov, Yoram J. Kaufman, Michael D. King, Didier Tanré, and Ilya Slutsker. Variability of Absorption and Optical Properties of Key Aerosol Types Observed in Worldwide Locations. *Journal of the Atmospheric Sciences*, 59(3):590–608, 02 2002b. ISSN 0022-4928. doi: 10.1175/1520-0469(2002)059<0590:VOAAOP>2.0.CO;2. URL https://doi.org/10.1175/1520-0469(2002)059<0590:VOAAOP>2.0.CO;2.

Oleg Dubovik, Alexander Sinyuk, Tatyana Lapyonok, Brent N. Holben, Michael Mishchenko, Ping Yang, Tom F. Eck, Hester Volten, Olga Muñoz, Ben Veihelmann, Wim J. van der Zande, Jean-Francois Leon, Michael Sorokin, and Ilya Slutsker. Application of spheroid models to account for aerosol particle nonsphericity in remote sensing of desert dust. *Journal of Geophysical Research: Atmospheres*, 111(D11), 2006. doi: 10.1029/2005JD006619. URL https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2005JD006619.

Oleg Dubovik, Tatyana Lapyonok, Pavel Litvinov, Maurice Herman, David Fuertes, Fabrice Ducos, Anton Lopatin, Anatoli Chaikovsky, Benjamin Torres, Yevgeny Derimian, et al. Grasp: a versatile algorithm for characterizing the atmosphere. *SPIE Newsroom*, 25, 2014. doi: https://doi.org/10.1117/2.1201408.005558.

Oleg Dubovik, Zhengqiang Li, Michael I. Mishchenko, Didier Tanré, Yana Karol, Bojan Bojkov, Brian Cairns, David J. Diner, W. Reed Espinosa, Philippe Goloub, Xingfa Gu, Otto Hasekamp, Jin Hong, Weizhen Hou, Kirk D. Knobelspiesse, Jochen Landgraf, Li Li, Pavel Litvinov, Yi Liu, Anton Lopatin, Thierry Marbach, Hal Maring, Vanderlei Martins, Yasjka Meijer, Gennadi Milinevsky, Sonoyo Mukai, Frederic Parol, Yanli Qiao, Lorraine Remer, Jeroen Rietjens, Itaru Sano, Piet Stammes, Snorre Stamnes, Xiaobing Sun, Pierre Tabary, Larry D. Travis, Fabien Waquet, Feng Xu, Changxiang Yan, and Dekui Yin. Polarimetric remote sensing of atmospheric aerosols: Instruments, methodologies, results, and perspectives. *Journal of Quantitative Spectroscopy and Radiative Transfer*, 224:474 – 511, 2019. ISSN 0022-4073. doi: https://doi.org/10.1016/j.jqsrt.2018.11.024. URL http://www.sciencedirect.com/science/article/pii/S0022407318308409.

Philippe Dubuisson, Robert Frouin, David Dessailly, Lucile Duforêt, Jean-François Léon, Kenneth Voss, and David Antoine. Estimating the altitude of aerosol plumes over the ocean from reflectance ratio measurements in the o2 a-band. *Remote Sensing of Environment*, 113(9):1899 – 1911, 2009. ISSN 0034-4257. doi: https://doi.org/10.1016/j.rse.2009.04.018. URL http://www.sciencedirect.com/science/article/pii/S0034425709001333.

Lucile Duforêt, Robert Frouin, and Philippe Dubuisson. Importance and estimation of aerosol vertical structure in satellite ocean-color remote sensing. *Appl. Opt.*, 46(7):1107–1119, Mar 2007. doi: 10.1364/AO.46.001107. URL http://ao.osa.org/abstract.cfm?URI=ao-46-7-1107.

B. N. Duncan, I. Bey, M. Chin, L. J. Mickley, T. D. Fairlie, R. V. Martin, and H. Matsueda. Indonesian wildfires of 1997: Impact on tropospheric chemistry. *Journal of Geophysical Research: Atmospheres*, 108(D15), 2003a. doi: 10.1029/2002JD003195. URL https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2002JD003195.

Bryan N. Duncan, Randall V. Martin, Amanda C. Staudt, Rosemarie Yevich, and Jennifer A. Logan. Interannual and seasonal variability of biomass burning emissions constrained by satellite observations. *Journal of Geophysical Research: Atmospheres*, 108(D2): ACH 1–1–ACH 1–22, 2003b. doi: 10.1029/2002JD002378. URL https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2002JD002378.

Surya S. Durbha, Roger L. King, and Nicolas H. Younan. Support vector machines regression for retrieval of leaf area index from multiangle imaging spectroradiometer. *Remote Sensing of Environment*, 107(1):348 – 361, 2007. ISSN 0034-4257. doi: https://doi.org/10.1016/j.rse.2006.09.031. URL http://www.sciencedirect.com/science/article/pii/S0034425706004159. Multi-angle Imaging SpectroRadiometer (MISR) Special Issue.

T. F. Eck, B. N. Holben, D. E. Ward, M. M. Mukelabai, O. Dubovik, A. Smirnov, J. S. Schafer, N. C. Hsu, S. J. Piketh, A. Queface, J. Le Roux, R. J. Swap, and I. Slutsker. Variability of biomass burning aerosol optical characteristics in southern africa during the safari 2000 dry season campaign and a comparison of single scattering albedo estimates from radiometric measurements. *Journal of Geophysical Research: Atmospheres*, 108(D13), 2003. doi: 10.1029/2002JD002321. URL https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2002JD002321.

T. F. Eck, B. N. Holben, J. S. Reid, M. M. Mukelabai, S. J. Piketh, O. Torres, H. T. Jethva, E. J. Hyer, D. E. Ward, O. Dubovik, A. Sinyuk, J. S. Schafer, D. M. Giles, M. Sorokin, A. Smirnov, and I. Slutsker. A seasonal trend of single scattering albedo in southern african biomass-burning particles: Implications for satellite products and estimates of emissions for the world's largest biomass-burning source. *Journal of Geophysical Research: Atmospheres*, 118(12):6414–6432, 2013. doi: 10.1002/jgrd.50500. URL https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1002/jgrd.50500.

Thomas F. Eck, Brent N. Holben, Ilya Slutsker, and Alberto Setzer. Measurements of irradiance attenuation and estimation of aerosol single scattering albedo for biomass burning aerosols in amazonia. *Journal of Geophysical Research: Atmospheres*, 103(D24):31865–31878, 1998. doi: 10.1029/98JD00399. URL https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/98JD00399.

H. El-Askary, R. Gautam, R.P. Singh, and M. Kafatos. Dust storms detection over the indo-gangetic basin using multi sensor data. *Advances in Space Research*, 37(4):728 – 733, 2006. ISSN 0273-1177. doi: https://doi.org/10.1016/j.asr.2005.03.134. URL http://www.sciencedirect.com/science/article/pii/S0273117705003972. Natural Hazards and Oceanographic Processes from Satellite Data.

Edwin E. Eloranta. *High Spectral Resolution Lidar*, pages 143–163. Springer New York, New York, NY, 2005. ISBN 978-0-387-25101-1. doi: 10.1007/0-387-25101-4_5. URL https://doi.org/10.1007/0-387-25101-4_5.

A. Roderick Escombe, Clarissa Oeser, Robert H. Gilman, Marcos Navincopa, Eduardo Ticona, Carlos Martínez, Luz Caviedes, Patricia Sheen, Armando Gonzalez, Catherine Noakes, David A. J. Moore, Jon S. Friedland, and Carlton A. Evans. The Detection of Airborne Transmission of Tuberculosis from HIV-Infected Patients, Using an In Vivo Air Sampling Model. *Clinical Infectious Diseases*, 44(10):1349–1357, 05 2007. ISSN 1058-4838. doi: 10.1086/515397. URL https://doi.org/10.1086/515397.

Anton Fedarenka, Oleg Dubovik, Philippe Goloub, Zhengqiang Li, Tatyana Lapyonok, Pavel Litvinov, Luc Barel, Louis Gonzalez, Thierry Podvin, and Didier Crozel. Utilization of aeronet polarimetric measurements for improving retrieval of aerosol microphysics: Gsfc, beijing and dakar data analysis. *Journal of Quantitative Spectroscopy and Radiative Transfer*, 179:72 – 97, 2016. ISSN 0022-4073. doi: https://doi.org/10.1016/j.jqsrt.2016.03.021. URL http://www.sciencedirect.com/science/article/pii/S0022407316301534.

Graham Feingold, William R. Cotton, Sonia M. Kreidenweis, and Janel T. Davis. The Impact of Giant Cloud Condensation Nuclei on Drizzle Formation in Stratocumulus: Implications for Cloud Radiative Properties. *Journal of the Atmospheric Sciences*, 56(24): 4100–4117, 12 1999. ISSN 0022-4928. doi: 10.1175/1520-0469(1999)056<4100:TIOGCC>2.0.CO;2. URL https://doi.org/10.1175/1520-0469(1999)056<4100:TIOGCC>2.0.CO;2.

Frederick G. Fernald. Analysis of atmospheric lidar observations: some comments. *Appl. Opt.*, 23(5):652–653, Mar 1984. doi: 10.1364/AO.23.000652. URL http://ao.osa.org/abstract.cfm?URI=ao-23-5-652.

James W. Fitzgerald. Marine aerosols: A review. *Atmospheric Environment. Part A. General Topics*, 25(3):533 – 545, 1991. ISSN 0960-1686. doi: https://doi.org/10.1016/0960-1686(91)90050-H. URL http://www.sciencedirect.com/science/article/pii/096016869190050H.

Pierre Flamant, Juan Cuesta, Marie-LAURE Denneulin, Alain Dabas, and Dorit Huber. Adm-aeolus retrieval algorithms for aerosol and cloud products. *Tellus A: Dynamic Meteorology and Oceanography*, 60(2):273–286, 2008. doi: 10.1111/j.1600-0870.2007.00287.x. URL https://doi.org/10.1111/j.1600-0870.2007.00287.x.

Roland J. Flowerdew and Joanna D. Haigh. An approximation to improve accuracy in the derivation of surface reflectances from multi-look satellite radiometers. *Geophysical Research Letters*, 22(13):1693–1696, 1995. doi: https://doi.org/10.1029/95GL01662. URL https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/95GL01662.

P. Formenti, L. Mbemba Kabuiku, I. Chiapello, F. Ducos, F. Dulac, and D. Tanré. Aerosol optical properties derived from polder-3/parasol (2005–2013) over the western mediterranean sea – part 1: Quality assessment with aeronet and in situ airborne observations. *Atmospheric Measurement Techniques*, 11(12):6761–6784, 2018. doi: 10.5194/amt-11-6761-2018. URL https://www.atmos-meas-tech.net/11/6761/2018/.

Bertrand Fougnie, Thierry Marbach, Antoine Lacan, Ruediger Lang, Peter Schlüssel, Gabriele Poli, Rosemary Munro, and André B. Couto. The multi-viewing multi-channel multi-polarisation imager – overview of the 3mi polarimetric mission for aerosol and cloud characterization. *Journal of Quantitative Spectroscopy and Radiative Transfer*, 219:23 – 32, 2018. ISSN 0022-4073. doi: https://doi.org/10.1016/j.jqsrt.2018.07.008. URL http://www.sciencedirect.com/science/article/pii/S002240731830373X.

M. J. Garay, M. L. Witek, R. A. Kahn, F. C. Seidel, J. A. Limbacher, M. A. Bull, D. J. Diner, E. G. Hansen, O. V. Kalashnikova, H. Lee, A. M. Nastan, and Y. Yu. Introducing the 4.4 km spatial resolution multi-angle imaging spectroradiometer (misr) aerosol product. *Atmospheric Measurement Techniques*, 13(2):593–628, 2020. doi: 10.5194/amt-13-593-2020. URL https://amt.copernicus.org/articles/13/593/2020/.

S. Gassó and O. Torres. The role of cloud contamination, aerosol layer height and aerosol model in the assessment of the omi near-uv
retrievals over the ocean. *Atmospheric Measurement Techniques*, 9(7):3031–3052, 2016. doi: 10.5194/amt-9-3031-2016. URL https://amt.copernicus.org/articles/9/3031/2016/.

Igor V. Geogdzhayev, Michael I. Mishchenko, William B. Rossow, Brian Cairns, and Andrew A. Lacis. Global Two-Channel AVHRR Retrievals of Aerosol Properties over the Ocean for the Period of NOAA-9 Observations and Preliminary Retrievals Using NOAA-7 and NOAA-11 Data. *Journal of the Atmospheric Sciences*, 59(3):262–278, 02 2002. ISSN 0022-4928. doi: 10.1175/1520-0469(2002)059<0262:GTCARO>2.0.CO;2. URL https://doi.org/10.1175/1520-0469(2002)059<0262:GTCARO>2.0.CO;2.

Igor V. Geogdzhayev, Michael I. Mishchenko, Edward I. Terez, Galina A. Terez, and Genady K. Gushchin. Regional advanced very high resolution radiometer–derived climatology of aerosol optical thickness and size. *Journal of Geophysical Research: Atmospheres*, 110(D23), 2005. doi: 10.1029/2005JD006170. URL https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2005JD006170.

R. Ghahremaninezhad, W. Gong, M. Galí, A.-L. Norman, S. R. Beagley, A. Akingunola, Q. Zheng, A. Lupu, M. Lizotte, M. Levasseur, and W. R. Leaitch. Dimethyl sulfide and its role in aerosol formation and growth in the arctic summer – a modelling study. *Atmospheric Chemistry and Physics*, 19(23):14455–14476, 2019. doi: 10.5194/acp-19-14455-2019. URL https://acp.copernicus.org/articles/19/14455/2019/.

Robin C. Gilbert, Michael B. Richman, Theodore B. Trafalis, and Lance M. Leslie. Machine Learning Methods for Data Assimilation. In *Intelligent Engineering Systems through Artificial Neural Networks, Volume 20*. ASME Press, 01 2010. ISBN 9780791859599. doi: 10.1115/1.859599.paper14. URL https://doi.org/10.1115/1.859599.paper14.

D. M. Giles, B. N. Holben, T. F. Eck, A. Sinyuk, A. Smirnov, I. Slutsker, R. R. Dickerson, A. M. Thompson, and J. S. Schafer. An analysis of aeronet aerosol absorption properties and classifications representative of aerosol source regions. *Journal of Geophysical Research: Atmospheres*, 117(D17), 2012. doi: 10.1029/2012JD018127. URL https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2012JD018127.

D. M. Giles, A. Sinyuk, M. G. Sorokin, J. S. Schafer, A. Smirnov, I. Slutsker, T. F. Eck, B. N. Holben, J. R. Lewis, J. R. Campbell, E. J. Welton, S. V. Korkin, and A. I. Lyapustin. Advancements in the aerosol robotic network (aeronet) version 3 database – automated near-real-time quality control algorithm with improved cloud screening for sun photometer aerosol optical depth (aod) measurements. *Atmospheric Measurement Techniques*, 12(1):169–209, 2019. doi: 10.5194/amt-12-169-2019. URL https://www.atmos-meas-tech.net/12/169/2019/.

P. Ginoux, J.M. Prospero, O. Torres, and M. Chin. Long-term simulation of global dust distribution with the gocart model: correlation with north atlantic oscillation. *Environmental Modelling Software*, 19(2):113 – 128, 2004. ISSN 1364-8152. doi: https://doi.org/10.1016/S1364-8152(03)00114-2. URL http://www.sciencedirect.com/science/article/pii/S1364815203001142. Modelling of Wind Erosion and Aeolian Processes.

Paul Ginoux and Omar Torres. Empirical toms index for dust aerosol: Applications to model validation and source characterization. *Journal of Geophysical Research: Atmospheres*, 108(D17), 2003. doi: 10.1029/2003JD003470. URL https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2003JD003470.

P. Goloub, D. Tanre, J. L. Deuze, M. Herman, A. Marchand, and F. . Breon. Validation of the first algorithm applied for deriving the aerosol properties over the ocean using the polder/adeos measurements. *IEEE Transactions on Geoscience and Remote Sensing*, 37(3):1586–1596, 1999.

Pablo M. Granitto, Cesare Furlanello, Franco Biasioli, and Flavia Gasperi. Recursive feature elimination with random forest for ptr-ms analysis of agroindustrial products. *Chemometrics and Intelligent Laboratory Systems*, 83(2):83 – 90, 2006. ISSN 0169-7439. doi: https://doi.org/10.1016/j.chemolab.2006.01.007. URL http://www.sciencedirect.com/science/article/pii/S0169743906000232.

T. M. Gray and R. Bennartz. Automatic volcanic ash detection from modis observations using a back-propagation neural network. *Atmospheric Measurement Techniques*, 8(12):5089–5097, 2015. doi: 10.5194/amt-8-5089-2015. URL https://amt.copernicus.org/articles/8/5089/2015/.

W. M. F. Grey, P. R. J. North, S. O. Los, and R. M. Mitchell. Aerosol optical depth and land surface reflectance from multiangle aatsr measurements: global validation and intersensor comparisons. *IEEE Transactions on Geoscience and Remote Sensing*, 44(8):2184–2197, Aug 2006. ISSN 1558-0644. doi: 10.1109/TGRS.2006.872079.

William M. F. Grey, Peter R. J. North, and Sietse O. Los. Computationally efficient method for retrieving aerosol optical depth from atsr-2 and aatsr data. *Appl. Opt.*, 45(12):2786–2795, Apr 2006. doi: 10.1364/AO.45.002786. URL http://ao.osa.org/abstract.cfm?URI=ao-45-12-2786.

H. Guan, R. Esswein, J. Lopez, R. Bergstrom, A. Warnock, M. Follette-Cook, M. Fromm, and L. T. Iraci. A multi-decadal history of biomass burning plume heights identified using aerosol index measurements. *Atmospheric Chemistry and Physics*, 10(14): 6461–6469, 2010. doi: 10.5194/acp-10-6461-2010. URL https://acp.copernicus.org/articles/10/6461/2010/.

The Guardian. Chile battles devastating wildfires: 'we have never seen anything on this scale'. *The Guardian*. URL https://www.theguardian.com/world/2017/jan/25/chile-fire-firefighting-international-help?utm_source=Climate+News+Network&utm_campaign=afdf3cf10c-EMAIL_CAMPAIGN_2017_02_03&utm_medium=email&utm_term=0_1198ea8936-afdf3cf10c-38798061.

Pawan Gupta and Sundar A. Christopher. Particulate matter air quality assessment using integrated surface, satellite, and meteorological products: 2. a neural network approach. *Journal of Geophysical Research: Atmospheres*, 114(D20), 2009. doi: https://doi.org/10.1029/2008JD011497. URL https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2008JD011497.

Isabelle Guyon and André Elisseeff. An introduction to variable and feature selection. *J. Mach. Learn. Res.*, 3(null):1157–1182, March 2003. ISSN 1532-4435.

M. S. Hammer, R. V. Martin, A. van Donkelaar, V. Buchard, O. Torres, D. A. Ridley, and R. J. D. Spurr. Interpreting the ultraviolet aerosol index observed with the omi satellite instrument to understand absorption by organic aerosols: implications for atmospheric oxidation and direct radiative effects. *Atmospheric Chemistry and Physics*, 16(4):2507–2523, 2016. doi: 10.5194/acp-16-2507-2016. URL https://acp.copernicus.org/articles/16/2507/2016/.

GW Hammond, RL Raddatz, and DE Gelskey. Impact of atmospheric dispersion and transport of viral aerosols on the epidemiology of influenza. *Reviews of infectious diseases*, 11(3):494–497, 1989. doi: 10.1093/clinids/11.3.494.

Bo Han, Slobodan Vucetic, Amy Braverman, and Zoran Obradovic. A statistical complement to deterministic algorithms for the retrieval of aerosol optical thickness from radiance data. *Engineering Applications of Artificial Intelligence*, 19(7):787 – 795, 2006. ISSN 0952-1976. doi: https://doi.org/10.1016/j.engappai.2006.05.009. URL http://www.sciencedirect.com/science/article/pii/S0952197606000947. Special issue on Engineering Applications of Neural Networks - Novel Applications of Neural Networks in Engineering.

Hanchuan Peng, Fuhui Long, and C. Ding. Feature selection based on mutual information criteria of max-dependency, max-relevance, and min-redundancy. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 27(8):1226–1238, Aug 2005. ISSN 1939-3539. doi: 10.1109/TPAMI.2005.159.

J. Hansen, M. Sato, and R. Ruedy. Radiative forcing and climate response. *Journal of Geophysical Research: Atmospheres*, 102 (D6):6831–6864, 1997. doi: 10.1029/96JD03436. URL https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/96JD03436.

J. E. Hansen and L. D. Travis. Light scattering in planetary atmospheres. *Space Sci. Rev.*, 16:527–610, 1974. doi: 10.1007/BF00168069.

James Hansen and Larissa Nazarenko. Soot climate forcing via snow and ice albedos. *Proceedings of the National Academy of Sciences*, 101(2):423–428, 2004. ISSN 0027-8424. doi: 10.1073/pnas.2237157100. URL https://www.pnas.org/content/101/2/423.

Otto P. Hasekamp and Jochen Landgraf. Retrieval of aerosol properties over land surfaces: capabilities of multiple-viewing-angle intensity and polarization measurements. *Appl. Opt.*, 46(16):3332–3344, Jun 2007. doi: 10.1364/AO.46.003332. URL http://ao.osa.org/abstract.cfm?URI=ao-46-16-3332.

D. A. Hauglustaine, Y. Balkanski, and M. Schulz. A global model simulation of present and future nitrate aerosols and their direct radiative forcing of climate. *Atmospheric Chemistry and Physics*, 14(20):11031–11063, 2014. doi: 10.5194/acp-14-11031-2014. URL https://acp.copernicus.org/articles/14/11031/2014/.

A. Hauser, D. Oesch, and N. Foppa. Aerosol optical depth over land: Comparing aeronet, avhrr and modis. *Geophysical Research Letters*, 32(17), 2005a. doi: 10.1029/2005GL023579. URL https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2005GL023579.

A. Hauser, D. Oesch, N. Foppa, and S. Wunderle. Noaa avhrr derived aerosol optical depth over land. *Journal of Geophysical Research: Atmospheres*, 110(D8), 2005b. doi: 10.1029/2004JD005439. URL https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2004JD005439.

Tadahiro Hayasaka, Shinsuke Satake, Atsushi Shimizu, Nobuo Sugimoto, Ichiro Matsui, Kazuma Aoki, and Yoshitaka Muraji. Vertical distribution and optical properties of aerosols observed over japan during the atmospheric brown clouds–east asia regional experiment 2005. *Journal of Geophysical Research: Atmospheres*, 112(D22), 2007. doi: 10.1029/2006JD008086. URL https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2006JD008086.

J. M. Haywood and K. P. Shine. The effect of anthropogenic sulfate and soot aerosol on the clear sky planetary radiation budget. *Geophysical Research Letters*, 22(5):603–606, 1995. doi: 10.1029/95GL00075. URL https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/95GL00075.

James Haywood and Olivier Boucher. Estimates of the direct and indirect radiative forcing due to tropospheric aerosols: A review. *Reviews of Geophysics*, 38(4):513–543, 2000. doi: 10.1029/1999RG000078. URL https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/1999RG000078.

Qianshan He, Chengcai Li, Jietai Mao, Alexis Kai-Hon Lau, and D. A. Chu. Analysis of aerosol vertical distribution and variability in hong kong. *Journal of Geophysical Research: Atmospheres*, 113(D14), 2008. doi: 10.1029/2008JD009778. URL https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2008JD009778.

Dean A. Hegg, Ronald J. Ferek, and Peter V. Hobbs. Light scattering and cloud condensation nucleus activity of sulfate aerosol measured over the northeast atlantic ocean. *Journal of Geophysical Research: Atmospheres*, 98(D8):14887–14894, 1993. doi: 10.1029/93JD01615. URL https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/93JD01615.

J. R. Herman, P. K. Bhartia, O. Torres, C. Hsu, C. Seftor, and E. Celarier. Global distribution of uv-absorbing aerosols from nimbus 7/toms data. *Journal of Geophysical Research: Atmospheres*, 102(D14):16911–16922, 1997a. doi: 10.1029/96JD03680. URL https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/96JD03680.

M. Herman, J. L. Deuzé, C. Devaux, P. Goloub, F. M. Bréon, and D. Tanré. Remote sensing of aerosols over land surfaces including polarization measurements and application to polder measurements. *Journal of Geophysical Research: Atmospheres*, 102(D14): 17039–17049, 1997b. doi: 10.1029/96JD02109. URL https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/96JD02109.

M. Herman, J.-L. Deuzé, A. Marchand, B. Roger, and P. Lallart. Aerosol remote sensing from polder/adeos over the ocean: Improved retrieval using a nonspherical particle model. *Journal of Geophysical Research: Atmospheres*, 110(D10), 2005. doi: 10.1029/2004JD004798. URL https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2004JD004798.

G.M. Hidy. Aerosols. In Robert A. Meyers, editor, *Encyclopedia of Physical Science and Technology (Third Edition)*, pages 273 – 299. Academic Pres, New York, third edition edition, 2003. ISBN 978-0-12-227410-7. doi: https://doi.org/10.1016/B0-12-227410-5/00014-4. URL http://www.sciencedirect.com/science/article/pii/B0122274105000144.

Eleanor J. Highwood, James M. Haywood, Michael D. Silverstone, Stuart M. Newman, and Jonathan P. Taylor. Radiative properties and direct effect of saharan dust measured by the c-130 aircraft during saharan dust experiment (shade): 2. terrestrial spectrum. *Journal of Geophysical Research: Atmospheres*, 108(D18), 2003. doi: 10.1029/2002JD002552. URL https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2002JD002552.

A. Hoffer, A. Gelencsér, P. Guyon, G. Kiss, O. Schmid, G. P. Frank, P. Artaxo, and M. O. Andreae. Optical properties of humic-like substances (hulis) in biomass-burning aerosols. *Atmospheric Chemistry and Physics*, 6(11):3563–3570, 2006. doi: 10.5194/acp-6-3563-2006. URL https://acp.copernicus.org/articles/6/3563/2006/.

B. N. Holben, T. F. Eck, I. Slutsker, A. Smirnov, A. Sinyuk, J. Schafer, D. Giles, and O. Dubovik. Aeronet's Version 2.0 quality assurance criteria. In Si-Chee Tsay, Teruyuki Nakajima, Ramesh P. Singh, and R. Sridharan, editors, *Remote Sensing of the Atmosphere and Clouds*, volume 6408, pages 134 – 147. International Society for Optics and Photonics, SPIE, 2006. doi: 10.1117/12.706524. URL https://doi.org/10.1117/12.706524.

B.N. Holben, T.F. Eck, I. Slutsker, D. Tanré, J.P. Buis, A. Setzer, E. Vermote, J.A. Reagan, Y.J. Kaufman, T. Nakajima, F. Lavenu, I. Jankowiak, and A. Smirnov. Aeronet—a federated instrument network and data archive for aerosol characterization. *Remote Sensing of Environment*, 66(1):1 – 16, 1998. ISSN 0034-4257. doi: https://doi.org/10.1016/S0034-4257(98)00031-5. URL http://www.sciencedirect.com/science/article/pii/S0034425798000315.

T. Holzer-Popp, G. de Leeuw, J. Griesfeller, D. Martynenko, L. Klüser, S. Bevan, W. Davies, F. Ducos, J. L. Deuzé, R. G. Graigner, A. Heckel, W. von Hoyningen-Hüne, P. Kolmonen, P. Litvinov, P. North, C. A. Poulsen, D. Ramon, R. Siddans, L. Sogacheva, D. Tanre, G. E. Thomas, M. Vountas, J. Descloitres, J. Griesfeller, S. Kinne, M. Schulz, and S. Pinnock. Aerosol retrieval experiments in the esa aerosol_cci project. *Atmospheric Measurement Techniques*, 6(8):1919–1957, 2013. doi: 10.5194/amt-6-1919-2013. URL https://amt.copernicus.org/articles/6/1919/2013/.

S.B. Hooker and C.R. McClain. The calibration and validation of seawifs data. *Progress in Oceanography*, 45(3):427 – 465, 2000. ISSN 0079-6611. doi: https://doi.org/10.1016/S0079-6611(00)00012-4. URL http://www.sciencedirect.com/science/article/pii/S0079661100000124.

H. Horvath. Atmospheric visibility. *Atmospheric Environment (1967)*, 15(10):1785 – 1796, 1981. ISSN 0004-6981. doi: https://doi.org/10.1016/0004-6981(81)90214-6. URL http://www.sciencedirect.com/science/article/pii/0004698181902146. Plumes and Visibility Measurements and Model Components.

N. C. Hsu, J. R. Herman, P. K. Bhartia, C. J. Seftor, O. Torres, A. M. Thompson, J. F. Gleason, T. F. Eck, and B. N. Holben. Detection of biomass burning smoke from toms measurements. *Geophysical Research Letters*, 23(7):745–748, 1996. doi: 10.1029/96GL00455. URL https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/96GL00455.

N. C. Hsu, J. R. Herman, J. F. Gleason, O. Torres, and C. J. Seftor. Satellite detection of smoke aerosols over a snow/ice surface by toms. *Geophysical Research Letters*, 26(8):1165–1168, 1999. doi: 10.1029/1999GL900155. URL https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/1999GL900155.

N. C. Hsu, Si-Chee Tsay, M. D. King, and J. R. Herman. Aerosol properties over bright-reflecting source regions. *IEEE Transactions on Geoscience and Remote Sensing*, 42(3):557–569, March 2004. ISSN 1558-0644. doi: 10.1109/TGRS.2004.824067.

N. C. Hsu, S. . Tsay, M. D. King, and J. R. Herman. Deep blue retrievals of asian aerosol properties during ace-asia. *IEEE Transactions on Geoscience and Remote Sensing*, 44(11):3180–3195, Nov 2006. ISSN 1558-0644. doi: 10.1109/TGRS.2006.879540.

N. C. Hsu, M.-J. Jeong, C. Bettenhausen, A. M. Sayer, R. Hansell, C. S. Seftor, J. Huang, and S.-C. Tsay. Enhanced deep blue aerosol retrieval algorithm: The second generation. *Journal of Geophysical Research: Atmospheres*, 118(16):9296–9315, 2013. doi: 10.1002/jgrd.50712. URL https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1002/jgrd.50712.

N. C. Hsu, J. Lee, A. M. Sayer, N. Carletta, S.-H. Chen, C. J. Tucker, B. N. Holben, and S.-C. Tsay. Retrieving near-global aerosol loading over land and ocean from avhrr. *Journal of Geophysical Research: Atmospheres*, 122(18):9968–9989, 2017. doi: 10.1002/2017JD026932. URL https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1002/2017JD026932.

R.-M. Hu, R. V. Martin, and T. D. Fairlie. Global retrieval of columnar aerosol single scattering albedo from space-based observations. *Journal of Geophysical Research: Atmospheres*, 112(D2), 2007. doi: 10.1029/2005JD006832. URL https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2005JD006832.

Shuai Hu, Taichang Gao, Hao Li, Bo Yang, Feng Zhang, Ming Chen, and Lei Liu. Light scattering computation model for non-spherical aerosol particles based on multi-resolution time-domain scheme: model development and validation. *Opt. Express*, 25(2):1463–1486, Jan 2017. doi: 10.1364/OE.25.001463. URL http://www.opticsexpress.org/abstract.cfm?URI=oe-25-2-1463.

Rudolf B. Husar, Joseph M. Prospero, and Larry L. Stowe. Characterization of tropospheric aerosols over the oceans with the noaa advanced very high resolution radiometer optical thickness operational product. *Journal of Geophysical Research: Atmospheres*, 102(D14):16889–16909, 1997. doi: 10.1029/96JD04009. URL https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/96JD04009.

Nicole Pauly Hyslop. Impaired visibility: the air pollution people see. *Atmospheric Environment*, 43(1):182 – 195, 2009. ISSN 1352-2310. doi: https://doi.org/10.1016/j.atmosenv.2008.09.067. URL http://www.sciencedirect.com/science/article/pii/S1352231008009217. Atmospheric Environment - Fifty Years of Endeavour.

Charles Ichoku, D. Allen Chu, Shana Mattoo, Yoram J. Kaufman, Lorraine A. Remer, Didier Tanré, Ilya Slutsker, and Brent N. Holben. A spatio-temporal approach for global validation and analysis of modis aerosol products. *Geophysical Research Letters*, 29(12):MOD1–1–MOD1–4, 2002. doi: https://doi.org/10.1029/2001GL013206. URL https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2001GL013206.

Akira Ishimaru, Robert J. Marks, Leung Tsang, Chi M. Lam, Dong C. Park, and Shinzo Kitamura. Particle-size distribution determination using optical sensing and neural networks. *Opt. Lett.*, 15(21):1221–1223, Nov 1990. doi: 10.1364/OL.15.001221. URL http://ol.osa.org/abstract.cfm?URI=ol-15-21-1221.

K. Ishitsuka, T. Mogi, K. Sugano, Y. Yamaya, T. Uchida, and T. Kajiwara. Resistivity-based temperature estimation of the kakkonda geothermal field, japan, using a neural network and neural kriging. *IEEE Geoscience and Remote Sensing Letters*, 15(8):1154–1158, Aug 2018. ISSN 1558-0571. doi: 10.1109/LGRS.2018.2832647.

T. Islam, Y. Hu, A.A. Kokhanovsky, and J. Wang. *Remote Sensing of Aerosols, Clouds, and Precipitation*. Elsevier Science, 2017. ISBN 9780128104385. URL https://books.google.nl/books?id=ooyZDgAAQBAJ.

P. L. Israelevich, Z. Levin, J. H. Joseph, and E. Ganor. Desert aerosol transport in the mediterranean region as inferred from the toms aerosol index. *Journal of Geophysical Research: Atmospheres*, 107(D21):AAC 13–1–AAC 13–13, 2002. doi: 10.1029/2001JD002011. URL https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2001JD002011.

John M. Jackson, Hongqing Liu, Istvan Laszlo, Shobha Kondragunta, Lorraine A. Remer, Jingfeng Huang, and Ho-Chun Huang. Suomi-npp viirs aerosol algorithms and data products. *Journal of Geophysical Research: Atmospheres*, 118(22):12,673–12,689, 2013. doi: 10.1002/2013JD020449. URL https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1002/2013JD020449.

Myeong-Jae Jeong and N. Christina Hsu. Retrievals of aerosol single-scattering albedo and effective aerosol layer height for biomass-burning smoke: Synergy derived from "a-train" sensors. *Geophysical Research Letters*, 35(24), 2008. doi: 10.1029/2008GL036279. URL https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2008GL036279.

H. Jethva and O. Torres. Satellite-based evidence of wavelength-dependent aerosol absorption in biomass burning smoke inferred from ozone monitoring instrument. *Atmospheric Chemistry and Physics*, 11(20):10541–10551, 2011a. doi: 10.5194/acp-11-10541-2011. URL https://acp.copernicus.org/articles/11/10541/2011/.

H. Jethva and O. Torres. Satellite-based evidence of wavelength-dependent aerosol absorption in biomass burning smoke inferred from ozone monitoring instrument. *Atmospheric Chemistry and Physics*, 11(20):10541–10551, 2011b. doi: 10.5194/acp-11-10541-2011. URL https://acp.copernicus.org/articles/11/10541/2011/.

Hiren Jethva, Omar Torres, and Changwoo Ahn. Global assessment of omi aerosol single-scattering albedo using ground-based aeronet inversion. *Journal of Geophysical Research: Atmospheres*, 119(14):9020–9040, 2014. doi: 10.1002/2014JD021672. URL https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1002/2014JD021672.

Allan Just, Margherita De Carli, Alexandra Shtein, Michael Dorman, Alexei Lyapustin, and Itai Kloog. Correcting measurement error in satellite aerosol optical depth with machine learning for modeling pm2.5 in the northeastern usa. *Remote Sensing*, 10 (5):803, May 2018. ISSN 2072-4292. doi: 10.3390/rs10050803. URL http://dx.doi.org/10.3390/rs10050803.

M. Kacenelenbogen, J. Redemann, M. A. Vaughan, A. H. Omar, P. B. Russell, S. Burton, R. R. Rogers, R. A. Ferrare, and C. A. Hostetler. An evaluation of caliop/calipso's aerosol-above-cloud detection and retrieval capability over north america. *Journal of Geophysical Research: Atmospheres*, 119(1):230–244, 2014. doi: 10.1002/2013JD020178. URL https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1002/2013JD020178.

R. A. Kahn, D. L. Nelson, M. J. Garay, R. C. Levy, M. A. Bull, D. J. Diner, J. V. Martonchik, S. R. Paradise, E. G. Hansen, and L. A. Remer. Misr aerosol product attributes and statistical comparisons with modis. *IEEE Transactions on Geoscience and Remote Sensing*, 47(12):4095–4114, 2009.

Ralph A. Kahn and Barbara J. Gaitley. An analysis of global aerosol type as retrieved by misr. *Journal of Geophysical Research: Atmospheres*, 120(9):4248–4281, 2015. doi: 10.1002/2015JD023322. URL https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1002/2015JD023322.

Ralph A. Kahn, Michael J. Garay, David L. Nelson, Kevin K. Yau, Michael A. Bull, Barbara J. Gaitley, John V. Martonchik, and Robert C. Levy. Satellite-derived aerosol optical depth over dark water from misr and modis: Comparisons with aeronet and implications for climatological studies. *Journal of Geophysical Research: Atmospheres*, 112(D18), 2007. doi: 10.1029/2006JD008175. URL https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2006JD008175.

O. V. Kalashnikova, R. Kahn, I. N. Sokolik, and Wen-Hao Li. Ability of multiangle remote sensing observations to identify and distinguish mineral dust types: Optical models and retrievals of optically thick plumes. *Journal of Geophysical Research: Atmospheres*, 110(D18), 2005. doi: 10.1029/2004JD004550. URL https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2004JD004550.

Olga V. Kalashnikova and Ralph Kahn. Ability of multiangle remote sensing observations to identify and distinguish mineral dust types: 2. sensitivity over dark water. *Journal of Geophysical Research: Atmospheres*, 111(D11), 2006. doi: 10.1029/2005JD006756. URL https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2005JD006756.

Olga V. Kalashnikova and Ralph A. Kahn. Mineral dust plume evolution over the atlantic from misr and modis aerosol retrievals. *Journal of Geophysical Research: Atmospheres*, 113(D24), 2008. doi: 10.1029/2008JD010083. URL https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2008JD010083.

Kane Yee. Numerical solution of initial boundary value problems involving maxwell's equations in isotropic media. *IEEE Transactions on Antennas and Propagation*, 14(3):302–307, 1966.

A. Karpatne, I. Ebert-Uphoff, S. Ravela, H. A. Babaie, and V. Kumar. Machine learning for the geosciences: Challenges and opportunities. *IEEE Transactions on Knowledge and Data Engineering*, 31(8):1544–1554, Aug 2019a. ISSN 1558-2191. doi: 10.1109/TKDE.2018.2861006.

A. Karpatne, I. Ebert-Uphoff, S. Ravela, H. A. Babaie, and V. Kumar. Machine learning for the geosciences: Challenges and opportunities. *IEEE Transactions on Knowledge and Data Engineering*, 31(8):1544–1554, Aug 2019b. ISSN 1558-2191. doi: 10.1109/TKDE.2018.2861006.

Evgueni I. Kassianov, James C. Barnard, and Thomas P. Ackerman. Retrieval of aerosol microphysical properties using surface multifilter rotating shadowband radiometer (mfrsr) data: Modeling and observations. *Journal of Geophysical Research: Atmospheres*, 110(D9), 2005. doi: 10.1029/2004JD005337. URL https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2004JD005337.

Iosif L. Katsev, Alexander S. Prikhach, Eleonora P. Zege, Arkadii P. Ivanov, and Alexander A. Kokhanovsky. *Iterative procedure for retrieval of spectral aerosol optical thickness and surface reflectance from satellite data using fast radiative transfer code and its application to MERIS measurements*, pages 101–133. Springer Berlin Heidelberg, Berlin, Heidelberg, 2009. ISBN 978-3-540-69397-0. doi: 10.1007/978-3-540-69397-0_4. URL https://doi.org/10.1007/978-3-540-69397-0_4.

Y. J. Kaufman, D. Tanré, H. R. Gordon, T. Nakajima, J. Lenoble, R. Frouin, H. Grassl, B. M. Herman, M. D. King, and P. M. Teillet. Passive remote sensing of tropospheric aerosol and atmospheric correction for the aerosol effect. *Journal of Geophysical Research: Atmospheres*, 102(D14):16815–16830, 1997a. doi: 10.1029/97JD01496. URL https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/97JD01496.

Y. J. Kaufman, D. Tanré, L. A. Remer, E. F. Vermote, A. Chu, and B. N. Holben. Operational remote sensing of tropospheric aerosol over land from eos moderate resolution imaging spectroradiometer. *Journal of Geophysical Research: Atmospheres*, 102(D14): 17051–17067, 1997b. doi: 10.1029/96JD03988. URL https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/96JD03988.

Yoram J. Kaufman, Nadine Gobron, Bernard Pinty, Jean-Luc Widlowski, and Michel M. Verstraete. Relationship between surface reflectance in the visible and mid-ir used in modis aerosol algorithm - theory. *Geophysical Research Letters*, 29(23):31–1–31–4, 2002a. doi: 10.1029/2001GL014492. URL https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2001GL014492.

Yoram J. Kaufman, Didier Tanré, and Olivier Boucher. A satellite view of aerosols in the climate system. *Nature*, 419(6903):215–223, 2002b. doi: 10.1038/nature01091. URL https://doi.org/10.1038/nature01091.

M.-H. Kim, A. H. Omar, J. L. Tackett, M. A. Vaughan, D. M. Winker, C. R. Trepte, Y. Hu, Z. Liu, L. R. Poole, M. C. Pitts, J. Kar, and B. E. Magill. The calipso version 4 automated aerosol classification and lidar ratio selection algorithm. *Atmospheric Measurement Techniques*, 11(11):6107–6135, 2018. doi: 10.5194/amt-11-6107-2018. URL https://amt.copernicus.org/articles/11/6107/2018/.

Man-Hae Kim, Sang-Woo Kim, Soon-Chang Yoon, and Ali H. Omar. Comparison of aerosol optical depth between caliop and modis-aqua for caliop aerosol subtypes over the ocean. *Journal of Geophysical Research: Atmospheres*, 118(23):13,241–13,252, 2013. doi: 10.1002/2013JD019527. URL https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1002/2013JD019527.

Minsang Kim, Myung-Sook Park, Jungho Im, Seonyoung Park, and Myong-In Lee. Machine learning approaches for detecting tropical cyclone formation using satellite data. *Remote Sensing*, 11(10):1195, May 2019. ISSN 2072-4292. doi: 10.3390/rs11101195. URL http://dx.doi.org/10.3390/rs11101195.

M. D. King, S. Platnick, W. P. Menzel, S. A. Ackerman, and P. A. Hubanks. Spatial and temporal distribution of clouds observed by modis onboard the terra and aqua satellites. *IEEE Transactions on Geoscience and Remote Sensing*, 51(7):3826–3852, July 2013. ISSN 1558-0644. doi: 10.1109/TGRS.2012.2227333.

Diederik P. Kingma and Jimmy Ba. Adam: A method for stochastic optimization, 2017.

T. F. Eck B. N. Holben J. S. Reid O. Dubovik A. Smirnov N. T. O'Neill I. Slutsker S. Kinne. Wavelength dependence of the optical depth of biomass burning, urban, and desert dust aerosols. *Journal of Geophysical Research: Atmospheres*, 104(D24): 31333–31349, 1999. doi: 10.1029/1999JD900923. URL https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/1999JD900923.

Z. Kipling, P. Stier, C. E. Johnson, G. W. Mann, N. Bellouin, S. E. Bauer, T. Bergman, M. Chin, T. Diehl, S. J. Ghan, T. Iversen, A. Kirkevåg, H. Kokkola, X. Liu, G. Luo, T. van Noije, K. J. Pringle, K. von Salzen, M. Schulz, Ø. Seland, R. B. Skeie, T. Takemura, K. Tsigaridis, and K. Zhang. 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, 2016. doi: 10.5194/acp-16-2221-2016. URL https://acp.copernicus.org/articles/16/2221/2016/.

Thomas W. Kirchstetter, T. Novakov, and Peter V. Hobbs. Evidence that the spectral dependence of light absorption by aerosols is affected by organic carbon. *Journal of Geophysical Research: Atmospheres*, 109(D21), 2004. doi: 10.1029/2004JD004999. URL https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2004JD004999.

Q. L. Kleipool, M. R. Dobber, J. F. de Haan, and P. F. Levelt. Earth surface reflectance climatology from 3 years of omi data. *Journal of Geophysical Research: Atmospheres*, 113(D18), 2008. doi: 10.1029/2008JD010290. URL https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2008JD010290.

James D. Klett. Stable analytical inversion solution for processing lidar returns. *Appl. Opt.*, 20(2):211–220, Jan 1981. doi: 10.1364/AO.20.000211. URL http://ao.osa.org/abstract.cfm?URI=ao-20-2-211.

D. Koch and A. D. Del Genio. Black carbon semi-direct effects on cloud cover: review and synthesis. *Atmospheric Chemistry and Physics*, 10(16):7685–7696, 2010. doi: 10.5194/acp-10-7685-2010. URL https://acp.copernicus.org/articles/10/7685/2010/.

Brigitte Koffi, Michael Schulz, Francois-Marie Bréon, Jan Griesfeller, David Winker, Yves Balkanski, Susanne Bauer, Terje Berntsen, Mian Chin, William D. Collins, Frank Dentener, Thomas Diehl, Richard Easter, Steven Ghan, Paul Ginoux, Sunling Gong, Larry W. Horowitz, Trond Iversen, Alf Kirkevåg, Dorothy Koch, Maarten Krol, Gunnar Myhre, Philip Stier, and Toshihiko Takemura. 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), 2012. doi: 10.1029/2011JD016858. URL https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2011JD016858.

A. A. Kokhanovsky, J. L. Deuzé, D. J. Diner, O. Dubovik, F. Ducos, C. Emde, M. J. Garay, R. G. Grainger, A. Heckel, M. Herman, I. L. Katsev, J. Keller, R. Levy, P. R. J. North, A. S. Prikhach, V. V. Rozanov, A. M. Sayer, Y. Ota, D. Tanré, G. E. Thomas, and E. P. Zege. The inter-comparison of major satellite aerosol retrieval algorithms using simulated intensity and polarization characteristics of reflected light. *Atmospheric Measurement Techniques*, 3(4):909–932, 2010. doi: 10.5194/amt-3-909-2010. URL https://amt.copernicus.org/articles/3/909/2010/.

A. A. Kokhanovsky, A. B. Davis, B. Cairns, O. Dubovik, O. P. Hasekamp, I. Sano, S. Mukai, V. V. Rozanov, P. Litvinov, T. Lapyonok, I. S. Kolomiets, Y. A. Oberemok, S. Savenkov, W. Martin, A. Wasilewski, A. Di Noia, F. A. Stap, J. Rietjens, F. Xu, V. Natraj, M. Duan, T. Cheng, and R. Munro. Space-based remote sensing of atmospheric aerosols: The multi-angle spectropolarimetric frontier. *Earth-Science Reviews*, 145:85–116, 2015. doi: https://doi.org/10.1016/j.earscirev.2015.01.012. URL http://www.sciencedirect.com/science/article/pii/S0012825215000215.

M.E. Koukouli, D.S. Balis, V. Amiridis, S. Kazadzis, A. Bais, S. Nickovic, and O. Torres. Aerosol variability over thessaloniki using ground based remote sensing observations and the toms aerosol index. *Atmospheric Environment*, 40(28):5367 – 5378, 2006. ISSN 1352-2310. doi: https://doi.org/10.1016/j.atmosenv.2006.04.046. URL http://www.sciencedirect.com/science/article/pii/S1352231006004341.

Vladimir M. Krasnopolsky. Reducing uncertainties in neural network jacobians and improving accuracy of neural network emulations with nn ensemble approaches. *Neural Networks*, 20(4):454 – 461, 2007. ISSN 0893-6080. doi: https://doi.org/10.1016/j.neunet.2007.04.008. URL http://www.sciencedirect.com/science/article/pii/S0893608007000421. Computational Intelligence in Earth and Environmental Sciences.

N. A. Krotkov, O. Torres, C. Seftor, A. J. Krueger, A. Kostinski, W. I. Rose, G. J. S. Bluth, D. Schneider, and S. J. Schaefer. Comparison of toms and avhrr volcanic ash retrievals from the august 1992 eruption of mt. spurr. *Geophysical Research Letters*, 26 (4):455–458, 1999. doi: 10.1029/1998GL900278. URL https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/1998GL900278.

Arlin Krueger, Nickolay Krotkov, and Simon Carn. El chichon: The genesis of volcanic sulfur dioxide monitoring from space. *Journal of Volcanology and Geothermal Research*, 175(4):408 – 414, 2008. ISSN 0377-0273. doi: https://doi.org/10.1016/j.jvolgeores.2008.02.026. URL http://www.sciencedirect.com/science/article/pii/S0377027308001625. The 25th Anniversary of the El Chichón Eruption.

A. Kylling, S. Vandenbussche, V. Capelle, J. Cuesta, L. Klüser, L. Lelli, T. Popp, K. Stebel, and P. Veefkind. Comparison of dust-layer heights from active and passive satellite sensors. *Atmospheric Measurement Techniques*, 11(5):2911–2936, 2018. doi: 10.5194/amt-11-2911-2018. URL https://amt.copernicus.org/articles/11/2911/2018/.

Carlo Lacagnina, Otto P. Hasekamp, Huisheng Bian, Gabriele Curci, Gunnar Myhre, Twan van Noije, Michael Schulz, Ragnhild B. Skeie, Toshihiko Takemura, and Kai Zhang. Aerosol single-scattering albedo over the global oceans: Comparing parasol retrievals with aeronet, omi, and aerocom models estimates. *Journal of Geophysical Research: Atmospheres*, 120(18):9814–9836, 2015. doi: 10.1002/2015JD023501. URL https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1002/2015JD023501.

Alexandra Larsen, Ivan Hanigan, Brian J. Reich, Yi Qin, Martin Cope, Geoffrey Morgan, and Ana G. Rappold. A deep learning approach to identify smoke plumes in satellite imagery in near-real time for health risk communication. *Journal of Exposure Science & Environmental Epidemiology*, 2020. doi: 10.1038/s41370-020-0246-y. URL https://doi.org/10.1038/s41370-020-0246-y.

D. J. Lary, L. A. Remer, D. MacNeill, B. Roscoe, and S. Paradise. Machine learning and bias correction of modis aerosol optical depth. *IEEE Geoscience and Remote Sensing Letters*, 6(4):694–698, Oct 2009. ISSN 1558-0571. doi: 10.1109/LGRS.2009.2023605.

David J. Lary, Fazlay S. Faruque, Nabin Malakar, Alex Moore, Bryan Roscoe, Zachary L. Adams, and York Eggelston. Estimating the global abundance of ground level presence of particulate matter (pm2.5). *Geospatial Health*, 8(3):S611–S630, Dec. 2014. doi: 10.4081/gh.2014.292. URL https://www.geospatialhealth.net/index.php/gh/article/view/292.

David J. Lary, Amir H. Alavi, Amir H. Gandomi, and Annette L. Walker. Machine learning in geosciences and remote sensing. *Geoscience Frontiers*, 7(1):3 – 10, 2016a. ISSN 1674-9871. doi: https://doi.org/10.1016/j.gsf.2015.07.003. URL http://www.sciencedirect.com/science/article/pii/S1674987115000821. Special Issue: Progress of Machine Learning in Geosciences.

David J. Lary, Amir H. Alavi, Amir H. Gandomi, and Annette L. Walker. Machine learning in geosciences and remote sensing. *Geoscience Frontiers*, 7(1):3 – 10, 2016b. ISSN 1674-9871. doi: https://doi.org/10.1016/j.gsf.2015.07.003. URL http://www.sciencedirect.com/science/article/pii/S1674987115000821. Special Issue: Progress of Machine Learning in Geosciences.

Jeanne M. Leclair, John A. Zaia, Myron J. Levin, Richard G. Congdon, and Donald A. Goldmann. Airborne transmission of chickenpox in a hospital. *New England Journal of Medicine*, 302(8):450–453, 1980. doi: 10.1056/NEJM198002213020807. URL https://doi.org/10.1056/NEJM198002213020807. PMID: 7351951.

J. Lee, R. C. Weger, S. K. Sengupta, and R. M. Welch. A neural network approach to cloud classification. *IEEE Transactions on Geoscience and Remote Sensing*, 28(5):846–855, Sep. 1990. ISSN 1558-0644. doi: 10.1109/36.58972.

Jaehwa Lee, Jhoon Kim, Hee Choon Lee, and Toshihiko Takemura. Classification of aerosol type from modis and omi over east asia. *Asia-Pacific Journal of Atmospheric Sciences*, 43(4):343–357, 2007.

Kwon H. Lee, Zhanqing Li, Young J. Kim, and Alexander Kokhanovsky. *Atmospheric Aerosol Monitoring from Satellite Observations: A History of Three Decades*, pages 13–38. Springer Netherlands, Dordrecht, 2009. ISBN 978-1-4020-9674-7. doi: 10.1007/978-1-4020-9674-7_2. URL https://doi.org/10.1007/978-1-4020-9674-7_2.

Pierre Legendre. Spatial autocorrelation: Trouble or new paradigm? *Ecology*, 74(6):1659–1673, 1993. doi: 10.2307/1939924. URL https://esajournals.onlinelibrary.wiley.com/doi/abs/10.2307/1939924.

J. Lenoble, M. Herman, J.L. Deuzé, B. Lafrance, R. Santer, and D. Tanré. A successive order of scattering code for solving the vector equation of transfer in the earth's atmosphere with aerosols. *Journal of Quantitative Spectroscopy and Radiative Transfer*, 107 (3):479 – 507, 2007. ISSN 0022-4073. doi: https://doi.org/10.1016/j.jqsrt.2007.03.010. URL http://www.sciencedirect.com/science/article/pii/S0022407307001306.

J.-F. Léon, Y. Derimian, I. Chiapello, D. Tanré, T. Podvin, B. Chatenet, A. Diallo, and C. Deroo. Aerosol vertical distribution and optical properties over m'bour (16.96deg; w; 14.39deg; n), senegal from 2006 to 2008. *Atmospheric Chemistry and Physics*, 9(23): 9249–9261, 2009. doi: 10.5194/acp-9-9249-2009. URL https://acp.copernicus.org/articles/9/9249/2009/.

P. F. Levelt, G. H. J. van den Oord, M. R. Dobber, A. Malkki, Huib Visser, Johan de Vries, P. Stammes, J. O. V. Lundell, and H. Saari. The ozone monitoring instrument. *IEEE Transactions on Geoscience and Remote Sensing*, 44(5):1093–1101, May 2006. ISSN 1558-0644. doi: 10.1109/TGRS.2006.872333.

PF Levelt and R Noordhoek. Omi algorithm theoretical basis document volume i. *ATBD. OMI-Ol, Version*, 1(1), 2002.

R. C. Levy, S. Mattoo, L. A. Munchak, L. A. Remer, A. M. Sayer, F. Patadia, and N. C. Hsu. The collection 6 modis aerosol products over land and ocean. *Atmospheric Measurement Techniques*, 6(11):2989–3034, 2013. doi: 10.5194/amt-6-2989-2013. URL https://amt.copernicus.org/articles/6/2989/2013/.

Robert C. Levy, Lorraine A. Remer, Shana Mattoo, Eric F. Vermote, and Yoram J. Kaufman. Second-generation operational algorithm: Retrieval of aerosol properties over land from inversion of moderate resolution imaging spectroradiometer spectral reflectance. *Journal of Geophysical Research: Atmospheres*, 112(D13), 2007. doi: 10.1029/2006JD007811. URL https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2006JD007811.

K. A. Lewis, W. P. Arnott, H. Moosmüller, R. K. Chakrabarty, C. M. Carrico, S. M. Kreidenweis, D. E. Day, W. C. Malm, A. Laskin, J. L. Jimenez, I. M. Ulbrich, J. A. Huffman, T. B. Onasch, A. Trimborn, L. Liu, and M. I. Mishchenko. Reduction in biomass burning aerosol light absorption upon humidification: roles of inorganically-induced hygroscopicity, particle collapse, and photoacoustic heat and mass transfer. *Atmospheric Chemistry and Physics*, 9(22):8949–8966, 2009. doi: 10.5194/acp-9-8949-2009. URL https://acp.copernicus.org/articles/9/8949/2009/.

Kristin Lewis, William P. Arnott, Hans Moosmüller, and Cyle E. Wold. Strong spectral variation of biomass smoke light absorption and single scattering albedo observed with a novel dual-wavelength photoacoustic instrument. *Journal of Geophysical Research: Atmospheres*, 113(D16), 2008. doi: 10.1029/2007JD009699. URL https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2007JD009699.

Jin Li, Andrew D. Heap, Anna Potter, and James J. Daniell. Application of machine learning methods to spatial interpolation of environmental variables. *Environmental Modelling Software*, 26(12):1647 – 1659, 2011. ISSN 1364-8152. doi: https://doi.org/10.1016/j.envsoft.2011.07.004. URL http://www.sciencedirect.com/science/article/pii/S1364815211001654.

Long Li, Runhe Shi, Lu Zhang, Jie Zhang, and Wei Gao. The data fusion of aerosol optical thickness using universal kriging and stepwise regression in East China. In Wei Gao, Ni-Bin Chang, and Jinnian Wang, editors, *Remote Sensing and Modeling of Ecosystems for Sustainability XI*, volume 9221, pages 219 – 229. International Society for Optics and Photonics, SPIE, 2014. doi: 10.1117/12.2061764. URL https://doi.org/10.1117/12.2061764.

Tongwen Li, Huanfeng Shen, Qiangqiang Yuan, Xuechen Zhang, and Liangpei Zhang. Estimating ground-level pm2.5 by fusing satellite and station observations: A geo-intelligent deep learning approach. *Geophysical Research Letters*, 44(23):11,985–11,993, 2017. doi: 10.1002/2017GL075710. URL https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1002/2017GL075710.

Xiang Li, Ling Peng, Yuan Hu, Jing Shao, and Tianhe Chi. Deep learning architecture for air quality predictions. *Environmental Science and Pollution Research*, 23(22):22408–22417, 2016. doi: 10.1007/s11356-016-7812-9.

Yingjie Li, Yong Xue, Gerrit de Leeuw, Chi Li, Leiku Yang, Tingting Hou, and Farhi Marir. Retrieval of aerosol optical depth and surface reflectance over land from noaa avhrr data. *Remote Sensing of Environment*, 133:1 – 20, 2013. ISSN 0034-4257. doi: https://doi.org/10.1016/j.rse.2013.01.020. URL http://www.sciencedirect.com/science/article/pii/S0034425713000382.

Zhengqiang Li, Philippe Goloub, Oleg Dubovik, Luc Blarel, Wenxing Zhang, Thierry Podvin, Alexander Sinyuk, Mikhail Sorokin, Hongbin Chen, Brent Holben, Didier Tanré, Marius Canini, and Jean-Pierre Buis. Improvements for ground-based remote sensing of atmospheric aerosol properties by additional polarimetric measurements. *Journal of Quantitative Spectroscopy and Radiative Transfer*, 110(17):1954 – 1961, 2009. ISSN 0022-4073. doi: https://doi.org/10.1016/j.jqsrt.2009.04.009. URL http://www.sciencedirect.com/science/article/pii/S002240730900154X.

Chih-Hsiang Liao, Shyh-Fang Kang, and Yu-Wei Hsu. Zero-valent iron reduction of nitrate in the presence of ultraviolet light, organic matter and hydrogen peroxide. *Water Research*, 37(17):4109 – 4118, 2003. ISSN 0043-1354. doi: https://doi.org/10.1016/S0043-1354(03)00248-3. URL http://www.sciencedirect.com/science/article/pii/S0043135403002483.

J. Liu, E. Scheuer, J. Dibb, G. S. Diskin, L. D. Ziemba, K. L. Thornhill, B. E. Anderson, A. Wisthaler, T. Mikoviny, J. J. Devi, M. Bergin, A. E. Perring, M. Z. Markovic, J. P. Schwarz, P. Campuzano-Jost, D. A. Day, J. L. Jimenez, and R. J. Weber. Brown carbon aerosol in the north american continental troposphere: sources, abundance, and radiative forcing. *Atmospheric Chemistry and Physics*, 15(14):7841–7858, 2015. doi: 10.5194/acp-15-7841-2015. URL https://acp.copernicus.org/articles/15/7841/2015/.

M. Liu, J. Lin, K. F. Boersma, G. Pinardi, Y. Wang, J. Chimot, T. Wagner, P. Xie, H. Eskes, M. Van Roozendael, F. Hendrick, P. Wang, T. Wang, Y. Yan, L. Chen, and R. Ni. Improved aerosol correction for omi tropospheric NO_2 retrieval over east asia: constraint from caliop aerosol vertical profile. *Atmospheric Measurement Techniques*, 12(1):1–21, 2019. doi: 10.5194/amt-12-1-2019. URL https://amt.copernicus.org/articles/12/1/2019/.

Yang Liu and David J. Diner. Multi-angle imager for aerosols: A satellite investigation to benefit public health. *Public Health Reports*, 132(1):14–17, 2017. doi: 10.1177/0033354916679983. URL https://doi.org/10.1177/0033354916679983. PMID: 28005476.

Yuan Liu, Zhi Ning, Yu Chen, Ming Guo, Yingle Liu, Nirmal Kumar Gali, Li Sun, Yusen Duan, Jing Cai, Dane Westerdahl, et al. Aerodynamic analysis of sars-cov-2 in two wuhan hospitals. *Nature*, pages 1–4, 2020. doi: https://doi.org/10.1038/s41586-020-2271-3.

Zhaoyan Liu, Ali H. Omar, Yongxiang Hu, and Mark A. Vaughan. Caliop algorithm theoretical basis document part 3 : Scene classification algorithms. 2005.

D Llewellyn-Jones, MC Edwards, CT Mutlow, AR Birks, IJ Barton, and H Tait. Aatsr: Global-change and surface-temperature measurements from envisat. *ESA bulletin*, 105(10-21):25, 2001.

Norman G. Loeb and Wenying Su. Direct Aerosol Radiative Forcing Uncertainty Based on a Radiative Perturbation Analysis. *Journal of Climate*, 23(19):5288–5293, 10 2010. ISSN 0894-8755. doi: 10.1175/2010JCLI3543.1. URL https://doi.org/10.1175/2010JCLI3543.1.

Douglas H Lowenthal, John G Watson, and Pradeep Saxena. Contributions to light extinction during project mohave. *Atmospheric Environment*, 34(15):2351 – 2359, 2000. ISSN 1352-2310. doi: https://doi.org/10.1016/S1352-2310(99)00449-5. URL http://www.sciencedirect.com/science/article/pii/S1352231099004495.

T. Marbach, J. Riedi, A. Lacan, and P. Schlüssel. The 3MI mission: multi-viewing-channel-polarisation imager of the EUMETSAT polar system: second generation (EPS-SG) dedicated to aerosol and cloud monitoring. In Joseph A. Shaw and Daniel A. LeMaster, editors, *Polarization Science and Remote Sensing VII*, volume 9613, pages 271 – 278. International Society for Optics and Photonics, SPIE, 2015. doi: 10.1117/12.2186978. URL https://doi.org/10.1117/12.2186978.

S. Mauceri, B. Kindel, S. Massie, and P. Pilewskie. Neural network for aerosol retrieval from hyperspectral imagery. *Atmospheric Measurement Techniques*, 12(11):6017–6036, 2019. doi: 10.5194/amt-12-6017-2019. URL https://amt.copernicus.org/articles/12/6017/2019/.

Allison McComiskey and Graham Feingold. Quantifying error in the radiative forcing of the first aerosol indirect effect. *Geophysical Research Letters*, 35(2), 2008. doi: 10.1029/2007GL032667. URL https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2007GL032667.

Allison McComiskey, Stephen E. Schwartz, Beat Schmid, Hong Guan, Ernie R. Lewis, Paul Ricchiazzi, and John A. Ogren. Direct aerosol forcing: Calculation from observables and sensitivities to inputs. *Journal of Geophysical Research: Atmospheres*, 113(D9), 2008. doi: 10.1029/2007JD009170. URL https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2007JD009170.

Good N. Petters M. D. McFiggans G. McMeeking, G. R. and H. Coe. Influences on the fraction of hydrophobic and hydrophilic black carbon in the atmosphere. *Atmos. Chem. Phys.*, 11:5099–5112, 2011. doi: 10.5194/acp-11-5099-2011.

Anna M. Michalak and Shahar Shlomi. *A Geostatistical Data Assimilation Approach for Estimating Groundwater Plume Distributions from Multiple Monitoring Events*, pages 73–88. American Geophysical Union (AGU), 2013. ISBN 9781118666463. doi: 10.1029/171GM08. URL https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/171GM08.

Gustav Mie. Beiträge zur optik trüber medien, speziell kolloidaler metallösungen. *Annalen der Physik*, 330(3):377–445, 1908. doi: 10.1002/andp.19083300302. URL https://onlinelibrary.wiley.com/doi/abs/10.1002/andp.19083300302.

Michael I. Mishchenko and Larry D. Travis. Satellite retrieval of aerosol properties over the ocean using measurements of reflected sunlight: Effect of instrumental errors and aerosol absorption. *Journal of Geophysical Research: Atmospheres*, 102(D12): 13543–13553, 1997. doi: 10.1029/97JD01124. URL https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/97JD01124.

Michael I. Mishchenko, Larry D. Travis, and Daniel W. Mackowski. T-matrix computations of light scattering by nonspherical particles: A review. *Journal of Quantitative Spectroscopy and Radiative Transfer*, 55(5):535 – 575, 1996. ISSN 0022-4073. doi: https://doi.org/10.1016/0022-4073(96)00002-7. URL http://www.sciencedirect.com/science/article/pii/0022407396000027. Light Scattering by Non-Spherical Particles.

Michael I. Mishchenko, Brian Cairns, Greg Kopp, Carl F. Schueler, Bryan A. Fafaul, James E. Hansen, Ronald J. Hooker, Tom Itchkawich, Hal B. Maring, and Larry D. Travis. Accurate Monitoring of Terrestrial Aerosols and Total Solar Irradiance: Introducing the Glory Mission. *Bulletin of the American Meteorological Society*, 88(5):677–692, 05 2007. ISSN 0003-0007. doi: 10.1175/BAMS-88-5-677. URL https://doi.org/10.1175/BAMS-88-5-677.

Michael I. Mishchenko, Li Liu, Igor V. Geogdzhayev, Larry D. Travis, Brian Cairns, and Andrew A. Lacis. Toward unified satellite climatology of aerosol properties.: 3. modis versus misr versus aeronet. *Journal of Quantitative Spectroscopy and Radiative Transfer*, 111(4):540 – 552, 2010. ISSN 0022-4073. doi: https://doi.org/10.1016/j.jqsrt.2009.11.003. URL http://www.sciencedirect.com/science/article/pii/S0022407309003343.

Amit Kumar Mishra, Ilan Koren, and Yinon Rudich. Effect of aerosol vertical distribution on aerosol-radiation interaction: A theoretical prospect. *Heliyon*, 1(2):e00036, 2015. ISSN 2405-8440. doi: https://doi.org/10.1016/j.heliyon.2015.e00036. URL http://www.sciencedirect.com/science/article/pii/S2405844015303169.

Thomas M. Mitchell. *Machine Learning*. McGraw-Hill, Inc., USA, 1 edition, 1997. ISBN 0070428077.

H. Moosmüller, R.K. Chakrabarty, and W.P. Arnott. Aerosol light absorption and its measurement: A review. *Journal of Quantitative Spectroscopy and Radiative Transfer*, 110(11):844 – 878, 2009. ISSN 0022-4073. doi: https://doi.org/10.1016/j.jqsrt.2009.02.035. URL http://www.sciencedirect.com/science/article/pii/S0022407309000879. Light Scattering: Mie and More Commemorating 100 years of Mie's 1908 publication.

C. Moulin and I. Chiapello. Evidence of the control of summer atmospheric transport of african dust over the atlantic by sahel sources from toms satellites (1979–2000). *Geophysical Research Letters*, 31(2), 2004. doi: 10.1029/2003GL018931. URL https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2003GL018931.

Giorgos Mountrakis, Jungho Im, and Caesar Ogole. Support vector machines in remote sensing: A review. *ISPRS Journal of Photogrammetry and Remote Sensing*, 66(3):247 – 259, 2011. ISSN 0924-2716. doi: https://doi.org/10.1016/j.isprsjprs.2010.11.001. URL http://www.sciencedirect.com/science/article/pii/S0924271610001140.

Hitoshi Mukai and Yoshinari Ambe. Characterization of a humic acid-like brown substance in airborne particulate matter and tentative identification of its origin. *Atmospheric Environment (1967)*, 20(5):813 – 819, 1986. ISSN 0004-6981. doi: https://doi.org/10.1016/0004-6981(86)90265-9. URL http://www.sciencedirect.com/science/article/pii/0004698186902659.

R. Munro, R. Lang, D. Klaes, G. Poli, C. Retscher, R. Lindstrot, R. Huckle, A. Lacan, M. Grzegorski, A. Holdak, A. Kokhanovsky, J. Livschitz, and M. Eisinger. The gome-2 instrument on the metop series of satellites: instrument design, calibration, and level 1 data processing
– an overview. *Atmospheric Measurement Techniques*, 9(3):1279–1301, 2016. doi: 10.5194/amt-9-1279-2016. URL https://amt.copernicus.org/articles/9/1279/2016/.

G. Myhre, B. H. Samset, M. Schulz, Y. Balkanski, S. Bauer, T. K. Berntsen, H. Bian, N. Bellouin, M. Chin, T. Diehl, R. C. Easter, J. Feichter, S. J. Ghan, D. Hauglustaine, T. Iversen, S. Kinne, A. Kirkevåg, J.-F. Lamarque, G. Lin, X. Liu, M. T. Lund, G. Luo, X. Ma, T. van Noije, J. E. Penner, P. J. Rasch, A. Ruiz, Ø. Seland, R. B. Skeie, P. Stier, T. Takemura, K. Tsigaridis, P. Wang, Z. Wang, L. Xu, H. Yu, F. Yu, J.-H. Yoon, K. Zhang, H. Zhang, and C. Zhou. Radiative forcing of the direct aerosol effect from aerocom phase ii simulations. *Atmospheric Chemistry and Physics*, 13(4):1853–1877, 2013. doi: 10.5194/acp-13-1853-2013. URL https://acp.copernicus.org/articles/13/1853/2013/.

D. Müller, A. Ansmann, I. Mattis, M. Tesche, U. Wandinger, D. Althausen, and G. Pisani. Aerosol-type-dependent lidar ratios observed with raman lidar. *Journal of Geophysical Research: Atmospheres*, 112(D16), 2007. doi: 10.1029/2006JD008292. URL https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2006JD008292.

F. Nadal and F. . Breon. Parameterization of surface polarized reflectance derived from polder spaceborne measurements. *IEEE Transactions on Geoscience and Remote Sensing*, 37(3):1709–1718, May 1999. ISSN 1558-0644. doi: 10.1109/36.763292.

Vijayakumar S. Nair, S. Suresh Babu, K. Krishna Moorthy, Arun Kumar Sharma, Angela Marinoni, and Ajai. Black carbon aerosols over the himalayas: direct and surface albedo forcing. *Tellus B: Chemical and Physical Meteorology*, 65(1):19738, 2013. doi: 10.3402/tellusb.v65i0.19738. URL https://doi.org/10.3402/tellusb.v65i0.19738.

Teruyuki Nakajima, Masayuki Tanaka, and Toyotato Yamauchi. Retrieval of the optical properties of aerosols from aureole and extinction data. *Appl. Opt.*, 22(19):2951–2959, Oct 1983. doi: 10.1364/AO.22.002951. URL http://ao.osa.org/abstract.cfm?URI=ao-22-19-2951.

Teruyuki Nakajima, Glauco Tonna, Ruizhong Rao, Paolo Boi, Yoram Kaufman, and Brent Holben. Use of sky brightness measurements from ground for remote sensing of particulate polydispersions. *Appl. Opt.*, 35(15):2672–2686, May 1996. doi: 10.1364/AO.35.002672. URL http://ao.osa.org/abstract.cfm?URI=ao-35-15-2672.

S. Nanda, M. de Graaf, J. P. Veefkind, M. ter Linden, M. Sneep, J. de Haan, and P. F. Levelt. A neural network radiative transfer model approach applied to the tropospheric monitoring instrument aerosol height algorithm. *Atmospheric Measurement Techniques*, 12(12):6619–6634, 2019. doi: 10.5194/amt-12-6619-2019. URL https://amt.copernicus.org/articles/12/6619/2019/.

6

S. Nanda, M. de Graaf, J. P. Veefkind, M. Sneep, M. ter Linden, J. Sun, and P. F. Levelt. A first comparison of tropomi aerosol layer height (alh) to caliop data. *Atmospheric Measurement Techniques*, 13(6):3043–3059, 2020. doi: 10.5194/amt-13-3043-2020. URL https://amt.copernicus.org/articles/13/3043/2020/.

David L. Nelson, Michael J. Garay, Ralph A. Kahn, and Ben A. Dunst. Stereoscopic height and wind retrievals for aerosol plumes with the misr interactive explorer (minx). *Remote Sensing*, 5(9):4593–4628, 2013. ISSN 2072-4292. doi: 10.3390/rs5094593. URL https://www.mdpi.com/2072-4292/5/9/4593.

A. Niang, F. Badran, C. Moulin, M. Crépon, and S. Thiria. Retrieval of aerosol type and optical thickness over the mediterranean from seawifs images using an automatic neural classification method. *Remote Sensing of Environment*, 100(1):82 – 94, 2006. ISSN 0034-4257. doi: https://doi.org/10.1016/j.rse.2005.10.005. URL http://www.sciencedirect.com/science/article/pii/S0034425705003421.

Awa Niang, Lidwine Gross, Sylvie Thiria, Fouad Badran, and Cyril Moulin. Automatic neural classification of ocean colour reflectance spectra at the top of the atmosphere with introduction of expert knowledge. *Remote Sensing of Environment*, 86(2):257 – 271, 2003. ISSN 0034-4257. doi: https://doi.org/10.1016/S0034-4257(03)00113-5. URL http://www.sciencedirect.com/science/article/pii/S0034425703001135.

Tomoaki Nishizawa, Hajime Okamoto, Nobuo Sugimoto, Ichiro Matsui, Atsushi Shimizu, and Kazuma Aoki. An algorithm that retrieves aerosol properties from dual-wavelength polarized lidar measurements. *Journal of Geophysical Research: Atmospheres*, 112(D6), 2007. doi: 10.1029/2006JD007435. URL https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2006JD007435.

Colin D. O'Dowd, Michael H. Smith, Ian E. Consterdine, and Jason A. Lowe. Marine aerosol, sea-salt, and the marine sulphur cycle: a short review. *Atmospheric Environment*, 31(1):73 – 80, 1997. ISSN 1352-2310. doi: https://doi.org/10.1016/S1352-2310(96)00106-9. URL http://www.sciencedirect.com/science/article/pii/S1352231096001069.

Luis E. Olcese, Gustavo G. Palancar, and Beatriz M. Toselli. A method to estimate missing aeronet aod values based on artificial neural networks. *Atmospheric Environment*, 113:140–150, 2015. ISSN 1352-2310. doi: https://doi.org/10.1016/j.atmosenv.2015.05.009. URL https://www.sciencedirect.com/science/article/pii/S1352231015300832.

Ali H. Omar, David M. Winker, Mark A. Vaughan, Yongxiang Hu, Charles R. Trepte, Richard A. Ferrare, Kam-Pui Lee, Chris A. Hostetler, Chieko Kittaka, Raymond R. Rogers, Ralph E. Kuehn, and Zhaoyan Liu. The CALIPSO Automated Aerosol Classification and Lidar Ratio Selection Algorithm. *Journal of Atmospheric and Oceanic Technology*, 26(10):1994–2014, 10 2009. ISSN 0739-0572. doi: 10.1175/2009JTECHA1231.1. URL https://doi.org/10.1175/2009JTECHA1231.1.

Rajendra K Pachauri, Myles R Allen, Vicente R Barros, John Broome, Wolfgang Cramer, Renate Christ, John A Church, Leon Clarke, Qin Dahe, Purnamita Dasgupta, et al. *Climate change 2014: synthesis report. Contribution of Working Groups I, II and III to the fifth assessment report of the Intergovernmental Panel on Climate Change*. Ipcc, 2014.

Gelsomina Pappalardo, Aldo Amodeo, Arnoud Apituley, Adolfo Comeron, Volker Freudenthaler, H. Linné, A. Ansmann, Jens Bösenberg, G. D'Amico, I. Mattis, L. Mona, Ulla Wandinger, V. Amiridis, Lucas Alados-Arboledas, D. Nicolae, and M. Wiegner. Earlinet. *Atmospheric Measurement Techniques*, (8):2389–2409, 2014. URL http://nbn-resolving.de/urn/resolver.pl?urn=nbn:de:bvb:19-epub-33841-0.

E.M. Patterson and C.K. McMahon. Absorption characteristics of forest fire particulate matter. *Atmospheric Environment (1967)*, 18(11):2541 – 2551, 1984. ISSN 0004-6981. doi: https://doi.org/10.1016/0004-6981(84)90027-1. URL http://www.sciencedirect.com/science/article/pii/0004698184900271.

P.Coppo, B.Ricciarelli, F.Brandani, J.Delderfield, M.Ferlet, C.Mutlow, G.Munro, T.Nightingale, D.Smith, S.Bianchi, P.Nicol, S.Kirschstein, T.Hennig, W.Engel, J.Frerick, and J.Nieke. Slstr: a high accuracy dual scan temperature radiometer for sea and land surface monitoring from space. *Journal of Modern Optics*, 57(18):1815–1830, 2010. doi: 10.1080/09500340.2010.503010. URL https://doi.org/10.1080/09500340.2010.503010.

Joyce E Penner, MO Andreae, H Annegarn, L Barrie, J Feichter, D Hegg, A Jayaraman, R Leaitch, D Murphy, J Nganga, et al. Aerosols, their direct and indirect effects. In *Climate Change 2001: The Scientific Basis. Contribution of Working Group I to the Third Assessment Report of the Intergovernmental Panel on Climate Change*, pages 289–348. Cambridge University Press, 2001.

M. J. M. Penning de Vries, S. Dörner, J. Puķīte, C. Hörmann, M. D. Fromm, and T. Wagner. Characterisation of a stratospheric sulfate plume from the nabro volcano using a combination of passive satellite measurements in nadir and limb geometry. *Atmospheric Chemistry and Physics*, 14(15):8149–8163, 2014. doi: 10.5194/acp-14-8149-2014. URL https://acp.copernicus.org/articles/14/8149/2014/.

Richard J. Peralta, Carl Nardell, Brian Cairns, Edgar E. Russell, Larry D. Travis, Michael I. Mishchenko, Bryan A. Fafaul, and Ronald J. Hooker. Aerosol polarimetry sensor for the Glory Mission. In Tianxu Zhang, Carl Anthony Nardell, Duane D. Smith, Hangqing Lu, Tianxu Zhang, Carl Anthony Nardell, and Hanqing Lu, editors, *MIPPR 2007: Automatic Target Recognition and Image Analysis; and Multispectral Image Acquisition*, volume 6786, pages 1460 – 1476. International Society for Optics and Photonics, SPIE, 2007. doi: 10.1117/12.783307. URL https://doi.org/10.1117/12.783307.

Thibaut Perol, Michaël Gharbi, and Marine Denolle. Convolutional neural network for earthquake detection and location. *Science Advances*, 4(2), 2018. doi: 10.1126/sciadv.1700578. URL https://advances.sciencemag.org/content/4/2/e1700578.

J. L. Petters, V. K. Saxena, J. R. Slusser, B. N. Wenny, and Sasha Madronich. Aerosol single scattering albedo retrieved from measurements of surface uv irradiance and a radiative transfer model. *Journal of Geophysical Research: Atmospheres*, 108(D9), 2003. doi: 10.1029/2002JD002360. URL https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2002JD002360.

T. M. Phuong, Z. Lin, and R. B. Altman. Choosing snps using feature selection. In *2005 IEEE Computational Systems Bioinformatics Conference (CSB'05)*, pages 301–309, Aug 2005. doi: 10.1109/CSB.2005.22.

M. Picchiani, M. Chini, S. Corradini, L. Merucci, P. Sellitto, F. Del Frate, and S. Stramondo. Volcanic ash detection and retrievals using modis data by means of neural networks. *Atmospheric Measurement Techniques*, 4(12):2619–2631, 2011. doi: 10.5194/amt-4-2619-2011. URL https://amt.copernicus.org/articles/4/2619/2011/.

C. Pierangelo, A. Chédin, S. Heilliette, N. Jacquinet-Husson, and R. Armante. Dust altitude and infrared optical depth from airs. *Atmospheric Chemistry and Physics*, 4(7):1813–1822, 2004. doi: 10.5194/acp-4-1813-2004. URL https://acp.copernicus.org/articles/4/1813/2004/.

Robert Pincus and Marcia B Baker. Effect of precipitation on the albedo susceptibility of clouds in the marine boundary layer. *Nature*, 372(6503):250–252, 1994. doi: 10.1038/372250a0.

A. Piscini, M. Picchiani, M. Chini, S. Corradini, L. Merucci, F. Del Frate, and S. Stramondo. A neural network approach for the simultaneous retrieval of volcanic ash parameters and so_2 using modis data. *Atmospheric Measurement Techniques*, 7(12):4023–4047, 2014. doi: 10.5194/amt-7-4023-2014. URL https://amt.copernicus.org/articles/7/4023/2014/.

Edward M Purcell and Carlton R Pennypacker. Scattering and absorption of light by nonspherical dielectric grains. *The Astrophysical Journal*, 186:705–714, 1973.

Wenmin Qin, Lunche Wang, Aiwen Lin, Ming Zhang, and Muhammad Bilal. Improving the estimation of daily aerosol optical depth and aerosol radiative effect using an optimized artificial neural network. *Remote Sensing*, 10(7), 2018. ISSN 2072-4292. doi: 10.3390/rs10071022. URL https://www.mdpi.com/2072-4292/10/7/1022.

V. Radosavljevic, S. Vucetic, and Z. Obradovic. A data-mining technique for aerosol retrieval across multiple accuracy measures. *IEEE Geoscience and Remote Sensing Letters*, 7(2):411–415, April 2010. ISSN 1558-0571. doi: 10.1109/LGRS.2009.2037720.

V. Ramanathan and G. Carmichael. Global and regional climate changes due to black carbon. *Nature Geoscience*, 1(4):221–227, 2008. doi: 10.1038/ngeo156. URL https://doi.org/10.1038/ngeo156.

V. Ramanathan, P. J. Crutzen, J. T. Kiehl, and D. Rosenfeld. Aerosols, climate, and the hydrological cycle. *Science*, 294(5549):2119–2124, 2001. ISSN 0036-8075. doi: 10.1126/science.1064034. URL https://science.sciencemag.org/content/294/5549/2119.

C. A. Randles, A. M. da Silva, V. Buchard, P. R. Colarco, A. Darmenov, R. Govindaraju, A. Smirnov, B. Holben, R. Ferrare, J. Hair, Y. Shinozuka, and C. J. Flynn. The MERRA-2 Aerosol Reanalysis, 1980 Onward. Part I: System Description and Data Assimilation Evaluation. *Journal of Climate*, 30(17):6823–6850, 07 2017. ISSN 0894-8755. doi: 10.1175/JCLI-D-16-0609.1. URL https://doi.org/10.1175/JCLI-D-16-0609.1.

Kabir Rasouli, William W. Hsieh, and Alex J. Cannon. Daily streamflow forecasting by machine learning methods with weather and climate inputs. *Journal of Hydrology*, 414-415:284 – 293, 2012. ISSN 0022-1694. doi: https://doi.org/10.1016/j.jhydrol.2011.10.039. URL http://www.sciencedirect.com/science/article/pii/S0022169411007633.

Markus Reichstein, Gustau Camps-Valls, Bjorn Stevens, Martin Jung, Joachim Denzler, Nuno Carvalhais, and Prabhat. Deep learning and process understanding for data-driven earth system science. *Nature*, 566(7743):195–204, 2019. doi: 10.1038/s41586-019-0912-1. URL https://doi.org/10.1038/s41586-019-0912-1.

J. S. Reid, T. F. Eck, S. A. Christopher, R. Koppmann, O. Dubovik, D. P. Eleuterio, B. N. Holben, E. A. Reid, and J. Zhang. A review of biomass burning emissions part iii: intensive optical properties of biomass burning particles. *Atmospheric Chemistry and Physics*, 5(3):827–849, 2005. doi: 10.5194/acp-5-827-2005. URL https://acp.copernicus.org/articles/5/827/2005/.

L. A. Remer, D. Tanré, Y. J. Kaufman, C. Ichoku, S. Mattoo, R. Levy, D. A. Chu, B. Holben, O. Dubovik, A. Smirnov, J. V. Martins, R.-R. Li, and Z. Ahmad. Validation of modis aerosol retrieval over ocean. *Geophysical Research Letters*, 29(12):MOD3–1–MOD3–4, 2002. doi: 10.1029/2001GL013204. URL https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2001GL013204.

L. A. Remer, Y. J. Kaufman, D. Tanré, S. Mattoo, D. A. Chu, J. V. Martins, R.-R. Li, C. Ichoku, R. C. Levy, R. G. Kleidman, T. F. Eck, E. Vermote, and B. N. Holben. The MODIS Aerosol Algorithm, Products, and Validation. *Journal of the Atmospheric Sciences*, 62(4):947–973, 04 2005. ISSN 0022-4928. doi: 10.1175/JAS3385.1. URL https://doi.org/10.1175/JAS3385.1.

Lorraine A Remer, Didier Tanré, Yoram J Kaufman, R Levy, S Mattoo, et al. Algorithm for remote sensing of tropospheric aerosol from modis: Collection 005. *National Aeronautics and Space Administration*, 1490, 2006.

Lorraine A. Remer, Richard G. Kleidman, Robert C. Levy, Yoram J. Kaufman, Didier Tanré, Shana Mattoo, J. Vanderlei Martins, Charles Ichoku, Ilan Koren, Hongbin Yu, and Brent N. Holben. Global aerosol climatology from the modis satellite sensors. *Journal of Geophysical Research: Atmospheres*, 113(D14), 2008. doi: 10.1029/2007JD009661. URL https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2007JD009661.

Weeberb J. Requia, Brent A. Coull, and Petros Koutrakis. Evaluation of predictive capabilities of ordinary geostatistical interpolation, hybrid interpolation, and machine learning methods for estimating pm2.5 constituents over space. *Environmental Research*, 175:421 – 433, 2019. ISSN 0013-9351. doi: https://doi.org/10.1016/j.envres.2019.05.025. URL http://www.sciencedirect.com/science/article/pii/S001393511930283X.

David N. Reshef, Yakir A. Reshef, Hilary K. Finucane, Sharon R. Grossman, Gilean McVean, Peter J. Turnbaugh, Eric S. Lander, Michael Mitzenmacher, and Pardis C. Sabeti. Detecting novel associations in large data sets. *Science*, 334(6062):1518–1524, 2011. doi: 10.1126/science.1205438.

K. Ristovski, S. Vucetic, and Z. Obradovic. Uncertainty analysis of neural-network-based aerosol retrieval. *IEEE Transactions on Geoscience and Remote Sensing*, 50(2):409–414, Feb 2012. ISSN 1558-0644. doi: 10.1109/TGRS.2011.2166120.

C. A. Rivera-Carpio, C. E. Corrigan, T. Novakov, J. E. Penner, C. F. Rogers, and J. C. Chow. Derivation of contributions of sulfate and carbonaceous aerosols to cloud condensation nuclei from mass size distributions. *Journal of Geophysical Research: Atmospheres*, 101(D14):19483–19493, 1996. doi: 10.1029/95JD01077. URL https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/95JD01077.

P. B. Russell, R. W. Bergstrom, Y. Shinozuka, A. D. Clarke, P. F. DeCarlo, J. L. Jimenez, J. M. Livingston, J. Redemann, O. Dubovik, and A. Strawa. Absorption angstrom exponent in aeronet and related data as an indicator of aerosol composition. *Atmospheric Chemistry and Physics*, 10(3):1155–1169, 2010. doi: 10.5194/acp-10-1155-2010. URL https://acp.copernicus.org/articles/10/1155/2010/.

Bjørn H. Samset, Camilla W. Stjern, Elisabeth Andrews, Ralph A. Kahn, Gunnar Myhre, Michael Schulz, and Gregory L. Schuster. Aerosol absorption: Progress towards global and regional constraints. *Current Climate Change Reports*, 4(2):65 – 83, 2018. ISSN 2198-6061. doi: https://doi.org/10.1007/s40641-018-0091-4.

A. L. Samuel. Some studies in machine learning using the game of checkers. *IBM Journal of Research and Development*, 3(3): 210–229, July 1959. ISSN 0018-8646. doi: 10.1147/rd.33.0210.

A. F. J. Sanders, J. F. de Haan, M. Sneep, A. Apituley, P. Stammes, M. O. Vieitez, L. G. Tilstra, O. N. E. Tuinder, C. E. Koning, and J. P. Veefkind. Evaluation of the operational aerosol layer height retrieval algorithm for sentinel-5 precursor: application to o_2 a band observations from gome-2a. *Atmospheric Measurement Techniques*, 8(11):4947–4977, 2015. doi: 10.5194/amt-8-4947-2015. URL https://amt.copernicus.org/articles/8/4947/2015/.

AFJ Sanders and JF De Haan. Tropomi atbd of the aerosol layer height product, 2016.

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 o_2 a- and b-band sciamachy observations over kanpur: a case study. *Atmospheric Measurement Techniques*, 5(5):1099–1119, 2012. doi: 10.5194/amt-5-1099-2012. URL https://amt.copernicus.org/articles/5/1099/2012/.

M. Santese, F. De Tomasi, and M. R. Perrone. Aeronet versus modis aerosol parameters at different spatial resolutions over southeast italy. *Journal of Geophysical Research: Atmospheres*, 112(D10), 2007. doi: 10.1029/2006JD007742. URL https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2006JD007742.

S. K. Satheesh, O. Torres, L. A. Remer, S. Suresh Babu, V. Vinoj, T. F. Eck, R. G. Kleidman, and B. N. Holben. Improved assessment of aerosol absorption using omi-modis joint retrieval. *Journal of Geophysical Research: Atmospheres*, 114(D5), 2009. doi: 10.1029/2008JD011024. URL https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2008JD011024.

S.K. Satheesh and K. Krishna Moorthy. Radiative effects of natural aerosols: A review. *Atmospheric Environment*, 39(11):2089 – 2110, 2005. ISSN 1352-2310. doi: https://doi.org/10.1016/j.atmosenv.2004.12.029. URL http://www.sciencedirect.com/science/article/pii/S135223100500018X.

V. M. E. Schenkeveld, G. Jaross, S. Marchenko, D. Haffner, Q. L. Kleipool, N. C. Rozemeijer, J. P. Veefkind, and P. F. Levelt. In-flight performance of the ozone monitoring instrument. *Atmospheric Measurement Techniques*, 10(5):1957–1986, 2017. doi: 10.5194/amt-10-1957-2017. URL https://amt.copernicus.org/articles/10/1957/2017/.

Timothy J. Schmit, Paul Griffith, Mathew M. Gunshor, Jaime M. Daniels, Steven J. Goodman, and William J. Lebair. A closer look at the abi on the goes-r series. *Bulletin of the American Meteorological Society*, 98(4):681 – 698, 2017. doi: 10.1175/BAMS-D-15-00230.1. URL https://journals.ametsoc.org/view/journals/bams/98/4/bams-d-15-00230.1.xml.

B. E. Schutz, H. J. Zwally, C. A. Shuman, D. Hancock, and J. P. DiMarzio. Overview of the icesat mission. *Geophysical Research Letters*, 32(21), 2005. doi: 10.1029/2005GL024009. URL https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2005GL024009.

C. J. Seftor, N. C. Hsu, J. R. Herman, P. K. Bhartia, O. Torres, W. I. Rose, D. J. Schneider, and N. Krotkov. Detection of volcanic ash clouds from nimbus 7/total ozone mapping spectrometer. *Journal of Geophysical Research: Atmospheres*, 102(D14): 16749–16759, 1997. doi: 10.1029/97JD00925. URL https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/97JD00925.

J Seinfeld and S Pandis. *Atmospheric Chemistry and Physics: From Air Pollution to Climate Change*. 1998.

Zdeněk Sekera. *Polarization of Skylight*, pages 288–328. Springer Berlin Heidelberg, Berlin, Heidelberg, 1957. ISBN 978-3-642-45881-1. doi: 10.1007/978-3-642-45881-1_4. URL https://doi.org/10.1007/978-3-642-45881-1_4.

N. Sharma, P. Sharma, D. Irwin, and P. Shenoy. Predicting solar generation from weather forecasts using machine learning. In *2011 IEEE International Conference on Smart Grid Communications (SmartGridComm)*, pages 528–533, Oct 2011. doi: 10.1109/SmartGridComm.2011.6102379.

Kyung-Shik Shin, Taik Soo Lee, and Hyun jung Kim. An application of support vector machines in bankruptcy prediction model. *Expert Systems with Applications*, 28(1):127 – 135, 2005. ISSN 0957-4174. doi: https://doi.org/10.1016/j.eswa.2004.08.009. URL http://www.sciencedirect.com/science/article/pii/S095741740400096X.

M. Shiraiwa, Y. Kondo, T. Iwamoto, and K. Kita. Amplification of light absorption of black carbon by organic coating. *Aerosol Science and Technology*, 44(1):46–54, 2010. doi: 10.1080/02786820903357686. URL https://doi.org/10.1080/02786820903357686.

M. K. Singh, R. Gautam, and P. Venkatachalam. Bayesian merging of misr and modis aerosol optical depth products using error distributions from aeronet. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 10(12):5186–5200, Dec 2017. ISSN 2151-1535. doi: 10.1109/JSTARS.2017.2734331.

A. Sinyuk, B. N. Holben, T. F. Eck, D. M. Giles, I. Slutsker, S. Korkin, J. S. Schafer, A. Smirnov, M. Sorokin, and A. Lyapustin. The aeronet version 3 aerosol retrieval algorithm, associated uncertainties and comparisons to version 2. *Atmospheric Measurement Techniques*, 13(6):3375–3411, 2020. doi: 10.5194/amt-13-3375-2020. URL https://amt.copernicus.org/articles/13/3375/2020/.

A. Smirnov, B.N. Holben, T.F. Eck, O. Dubovik, and I. Slutsker. Cloud-screening and quality control algorithms for the aeronet database. *Remote Sensing of Environment*, 73(3):337 – 349, 2000. ISSN 0034-4257. doi: https://doi.org/10.1016/S0034-4257(00)00109-7. URL http://www.sciencedirect.com/science/article/pii/S0034425700001097.

A. Smirnov, B. N. Holben, I. Slutsker, D. M. Giles, C. R. McClain, T. F. Eck, S. M. Sakerin, A. Macke, P. Croot, G. Zibordi, P. K. Quinn, J. Sciare, S. Kinne, M. Harvey, T. J. Smyth, S. Piketh, T. Zielinski, A. Proshutinsky, J. I. Goes, N. B. Nelson, P. Larouche, V. F. Radionov, P. Goloub, K. Krishna Moorthy, R. Matarrese, E. J. Robertson, and F. Jourdin. Maritime aerosol network as a component of aerosol robotic network. *Journal of Geophysical Research: Atmospheres*, 114(D6), 2009. doi: https://doi.org/10.1029/2008JD011257. URL https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2008JD011257.

Alex J. Smola and Bernhard Schölkopf. A tutorial on support vector regression. *Statistics and Computing*, 14(3):199–222, 2004. doi: 10.1023/B:STCO.0000035301.49549.88. URL https://doi.org/10.1023/B:STCO.0000035301.49549.88.

L. Sogacheva, T. Popp, A. M. Sayer, O. Dubovik, M. J. Garay, A. Heckel, N. C. Hsu, H. Jethva, R. A. Kahn, P. Kolmonen, M. Kosmale, G. de Leeuw, R. C. Levy, P. Litvinov, A. Lyapustin, P. North, O. Torres, and A. Arola. Merging regional and global aerosol optical depth records from major available satellite products. *Atmospheric Chemistry and Physics*, 20(4):2031–2056, 2020. doi: 10.5194/acp-20-2031-2020. URL https://acp.copernicus.org/articles/20/2031/2020/.

Susan Solomon. Ipcc (2007): Climate change the physical science basis. *AGUFM*, 2007:U43D–01, 2007.

N. Spichtinger, M. Wenig, P. James, T. Wagner, U. Platt, and A. Stohl. Satellite detection of a continental-scale plume of nitrogen oxides from boreal forest fires. *Geophysical Research Letters*, 28(24):4579–4582, 2001. doi: 10.1029/2001GL013484. URL https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2001GL013484.

DC Stein Zweers. Tropomi atbd of the uv aerosol index, 2016.

T.F. Stocker, D. Qin, G.-K. Plattner, L.V. Alexander, S.K. Allen, N.L. Bindoff, F.-M. Bréon, J.A. Church, U. Cubasch, S. Emori, P. Forster, P. Friedlingstein, N. Gillett, J.M. Gregory, D.L. Hartmann, E. Jansen, B. Kirtman, R. Knutti, K. Krishna Kumar, P. Lemke, J. Marotzke, V. Masson-Delmotte, G.A. Meehl, I.I. Mokhov, S. Piao, V. Ramaswamy, D. Randall, M. Rhein, M. Rojas, C. Sabine, D. Shindell, L.D. Talley, D.G. Vaughan, and S.-P. Xie. *Technical Summary*, book section TS, page 33–115. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 2013. ISBN ISBN 978-1-107-66182-0. doi: 10.1017/CBO9781107415324.005. URL www.climatechange2013.org.

G. G. Stokes. On the Composition and Resolution of Streams of Polarized Light from different Sources. *Transactions of the Cambridge Philosophical Society*, 9:399, January 1851.

J. Sun, J. P. Veefkind, P. van Velthoven, and P. F. Levelt. Quantifying the single-scattering albedo for the january 2017 chile wildfires from simulations of the omi absorbing aerosol index. *Atmospheric Measurement Techniques*, 11(9):5261–5277, 2018. doi: 10.5194/amt-11-5261-2018. URL https://amt.copernicus.org/articles/11/5261/2018/.

J. Sun, P. Veefkind, S. Nanda, P. van Velthoven, and P. Levelt. The role of aerosol layer height in quantifying aerosol absorption from ultraviolet satellite observations. *Atmospheric Measurement Techniques*, 12(12):6319–6340, 2019. doi: 10.5194/amt-12-6319-2019. URL https://amt.copernicus.org/articles/12/6319/2019/.

J. Sun, J. P. Veefkind, P. van Velthoven, L. G. Tilstra, J. Chimot, S. Nanda, and P. F. Levelt. Defining aerosol layer height for uvai interpretation using aerosol vertical distributions characterized by merra-2. *Atmospheric Chemistry and Physics Discussions*, 2020:1–36, 2020. doi: 10.5194/acp-2020-39. URL https://acp.copernicus.org/preprints/acp-2020-39/.

Yele Sun, Zifa Wang, Pingqing Fu, Qi Jiang, Ting Yang, Jie Li, and Xinlei Ge. The impact of relative humidity on aerosol composition and evolution processes during wintertime in beijing, china. *Atmospheric Environment*, 77:927 – 934, 2013. ISSN 1352-2310. doi: https://doi.org/10.1016/j.atmosenv.2013.06.019. URL http://www.sciencedirect.com/science/article/pii/S1352231013004743.

Toshihiko Takemura, Toru Nozawa, Seita Emori, Takashi Y. Nakajima, and Teruyuki Nakajima. Simulation of climate response to aerosol direct and indirect effects with aerosol transport-radiation model. *Journal of Geophysical Research: Atmospheres*, 110(D2), 2005. doi: 10.1029/2004JD005029. URL https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2004JD005029.

D. A. Talmage and P. J. Curran. Remote sensing using partially polarized light. *International Journal of Remote Sensing*, 7(1):47–64, 1986. doi: 10.1080/01431168608954660. URL https://doi.org/10.1080/01431168608954660.

I. N. Tang and H. R. Munkelwitz. Water activities, densities, and refractive indices of aqueous sulfates and sodium nitrate droplets of atmospheric importance. *Journal of Geophysical Research: Atmospheres*, 99(D9):18801–18808, 1994. doi: 10.1029/94JD01345. URL https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/94JD01345.

I. N. Tang, A. C. Tridico, and K. H. Fung. Thermodynamic and optical properties of sea salt aerosols. *Journal of Geophysical Research: Atmospheres*, 102(D19):23269–23275, 1997. doi: 10.1029/97JD01806. URL https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/97JD01806.

Ignatius N. Tang. Chemical and size effects of hygroscopic aerosols on light scattering coefficients. *Journal of Geophysical Research: Atmospheres*, 101(D14):19245–19250, 1996. doi: 10.1029/96JD03003. URL https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/96JD03003.

Xiong Tang, Xiaoshan Zhang, Zhangwei Wang, and Zhijia Ci. Watersoluble organic carbon (wsoc) and its temperaturere-solved carbon fractions in atmospheric aerosols in beijing. *Atmospheric Research*, 181:200 – 210, 2016. ISSN 0169-8095. doi: https://doi.org/10.1016/j.atmosres.2016.06.019. URL http://www.sciencedirect.com/science/article/pii/S0169809516301727.

D. Tanré, F. M. Bréon, J. L. Deuzé, O. Dubovik, F. Ducos, P. François, P. Goloub, M. Herman, A. Lifermann, and F. Waquet. Remote sensing of aerosols by using polarized, directional and spectral measurements within the a-train: the parasol mission. *Atmospheric Measurement Techniques*, 4(7):1383–1395, 2011. doi: 10.5194/amt-4-1383-2011. URL https://amt.copernicus.org/articles/4/1383/2011/.

D. Tanré, Y. J. Kaufman, M. Herman, and S. Mattoo. Remote sensing of aerosol properties over oceans using the modis/eos spectral radiances. *Journal of Geophysical Research: Atmospheres*, 102(D14):16971–16988, 1997. doi: 10.1029/96JD03437. URL https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/96JD03437.

Evdokia Tapoglou, George P. Karatzas, Ioannis C. Trichakis, and Emmanouil A. Varouchakis. A spatio-temporal hybrid neural network-kriging model for groundwater level simulation. *Journal of Hydrology*, 519:3193 – 3203, 2014. ISSN 0022-1694. doi: https://doi.org/10.1016/j.jhydrol.2014.10.040. URL http://www.sciencedirect.com/science/article/pii/S002216941400835X.

D.A. Tarasov, A.G. Buevich, A.P. Sergeev, and A.V. Shichkin. High variation topsoil pollution forecasting in the russian subarctic: Using artificial neural networks combined with residual kriging. *Applied Geochemistry*, 88:188 – 197, 2018. ISSN 0883-2927. doi: https://doi.org/10.1016/j.apgeochem.2017.07.007. URL http://www.sciencedirect.com/science/article/pii/S0883292717302421. SI: ISEG 2016.

M. Taylor, S. Kazadzis, A. Tsekeri, A. Gkikas, and V. Amiridis. Satellite retrieval of aerosol microphysical and optical parameters using neural networks: a new methodology applied to the sahara desert dust peak. *Atmospheric Measurement Techniques*, 7(9):3151–3175, 2014. doi: 10.5194/amt-7-3151-2014. URL https://amt.copernicus.org/articles/7/3151/2014/.

Raymond Tellier. Review of aerosol transmission of influenza a virus. *Emerging infectious diseases*, 12(11):1657, 2006. doi: 10.3201/eid1211.060426.

Raymond Tellier. Aerosol transmission of influenza a virus: a review of new studies. *Journal of the Royal Society Interface*, 6 (suppl_6):S783–S790, 2009.

P. Tian, X. Cao, L. Zhang, N. Sun, L. Sun, T. Logan, J. Shi, Y. Wang, Y. Ji, Y. Lin, Z. Huang, T. Zhou, Y. Shi, and R. Zhang. Aerosol vertical distribution and optical properties over china from long-term satellite and ground-based remote sensing. *Atmospheric Chemistry and Physics*, 17(4):2509–2523, 2017. doi: 10.5194/acp-17-2509-2017. URL https://acp.copernicus.org/articles/17/2509/2017/.

Wang P. Tilstra, L. G. and P. Stammes. Algorithm theoretical gome-2 absorbing aerosol height. Technical report, Royal Netherlands Meteorological Institute, 2019. URL https://amt.copernicus.org/articles/6/2577/2013/.

Ahn C. Torres, O. and Z. Chen. Improvements to the omi near-uv aerosol algorithm using a-train caliop and airs observations. *Atmos. Meas. Tech.*, 6:3257–3270, 2013. doi: 10.5194/amt-6-3257-2013.

O. Torres, P. K. Bhartia, J. R. Herman, Z. Ahmad, and J. Gleason. Derivation of aerosol properties from satellite measurements of backscattered ultraviolet radiation: Theoretical basis. *Journal of Geophysical Research: Atmospheres*, 103(D14):17099–17110, 1998. doi: 10.1029/98JD00900. URL https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/98JD00900.

O. Torres, P. K. Bhartia, J. R. Herman, A. Sinyuk, Paul Ginoux, and Brent Holben. A Long-Term Record of Aerosol Optical Depth from TOMS Observations and Comparison to AERONET Measurements. *Journal of the Atmospheric Sciences*, 59(3):398–413, 02 2002a. ISSN 0022-4928. doi: 10.1175/1520-0469(2002)059<0398:ALTROA>2.0.CO;2. URL https://doi.org/10.1175/1520-0469(2002)059<0398:ALTROA>2.0.CO;2.

O Torres, R Decae, JP Veefkind, and G De Leeuw. Omi aerosol retrieval algorithm, omi algorithm theoretical basis document: clouds, aerosols, and surface uv irradiance, vol. 3, version 2. *OMI-ATBD-03. Available at http://eospso. gsfc. nasa. gov/eoshomepage/forscientists/atbd/docs/OMI/ATBD-OMI-03. pdf*, 2002b.

O. Torres, P. K. Bhartia, A. Sinyuk, E. J. Welton, and B. Holben. Total ozone mapping spectrometer measurements of aerosol absorption from space: Comparison to safari 2000 ground-based observations. *Journal of Geophysical Research: Atmospheres*, 110(D10), 2005. doi: 10.1029/2004JD004611. URL https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2004JD004611.

O. Torres, P. K. Bhartia, H. Jethva, and C. Ahn. Impact of the ozone monitoring instrument row anomaly on the long-term record of aerosol products. *Atmospheric Measurement Techniques*, 11(5):2701–2715, 2018. doi: 10.5194/amt-11-2701-2018. URL https://amt.copernicus.org/articles/11/2701/2018/.

Omar Torres, Aapo Tanskanen, Ben Veihelmann, Changwoo Ahn, Remco Braak, Pawan K. Bhartia, Pepijn Veefkind, and Pieternel Levelt. Aerosols and surface uv products from ozone monitoring instrument observations: An overview. *Journal of Geophysical Research: Atmospheres*, 112(D24), 2007. doi: 10.1029/2007JD008809. URL https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2007JD008809.

Omar Torres, Hiren Jethva, and P. K. Bhartia. Retrieval of Aerosol Optical Depth above Clouds from OMI Observations: Sensitivity Analysis and Case Studies. *Journal of the Atmospheric Sciences*, 69(3):1037–1053, 03 2012. ISSN 0022-4928. doi: 10.1175/JAS-D-11-0130.1. URL https://doi.org/10.1175/JAS-D-11-0130.1.

Tzu-Chin Tsai, Yung-Jyh Jeng, D. Allen Chu, Jen-Ping Chen, and Shuenn-Chin Chang. Analysis of the relationship between modis aerosol optical depth and particulate matter from 2006 to 2008. *Atmospheric Environment*, 45(27):4777 – 4788, 2011. ISSN 1352-2310. doi: https://doi.org/10.1016/j.atmosenv.2009.10.006. URL http://www.sciencedirect.com/science/article/pii/S1352231009008541.

John W. Tukey. *Exploratory Data Analysis*. Addison-Wesley, 1977.

D. D. Turner, R. A. Ferrare, and L. A. Brasseur. Average aerosol extinction and water vapor profiles over the southern great plains. *Geophysical Research Letters*, 28(23):4441–4444, 2001. doi: 10.1029/2001GL013691. URL https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2001GL013691.

S Twomey et al. Pollution and the planetary albedo. *Atmos. Environ*, 8(12):1251–1256, 1974.

Zbigniew Ulanowski, Zhenni Wang, Paul H. Kaye, and Ian K. Ludlow. Application of neural networks to the inverse light scattering problem for spheres. *Appl. Opt.*, 37(18):4027–4033, Jun 1998. doi: 10.1364/AO.37.004027. URL http://ao.osa.org/abstract.cfm?URI=ao-37-18-4027.

Aaldert van Amerongen, Jeroen Rietjens, Jochen Campo, Ersin Dogan, Jos Dingjan, Raj Nalla, Jerome Caron, and Otto Hasekamp. SPEXone: a compact multi-angle polarimeter. In Zoran Sodnik, Nikos Karafolas, and Bruno Cugny, editors, *International Conference on Space Optics — ICSO 2018*, volume 11180, pages 223 – 236. International Society for Optics and Photonics, SPIE, 2019. URL https://doi.org/10.1117/12.2535940.

Neeltje van Doremalen, Trenton Bushmaker, Dylan H. Morris, Myndi G. Holbrook, Amandine Gamble, Brandi N. Williamson, Azaibi Tamin, Jennifer L. Harcourt, Natalie J. Thornburg, Susan I. Gerber, James O. Lloyd-Smith, Emmie de Wit, and Vincent J. Munster. Aerosol and surface stability of sars-cov-2 as compared with sars-cov-1. *New England Journal of Medicine*, 382(16):1564–1567, 2020. doi: 10.1056/NEJMc2004973. URL https://doi.org/10.1056/NEJMc2004973.

S. Vandenbussche, S. Kochenova, A. C. Vandaele, N. Kumps, and M. De Mazière. Retrieval of desert dust aerosol vertical profiles from iasi measurements in the tir atmospheric window. *Atmospheric Measurement Techniques*, 6(10):2577–2591, 2013. doi: 10.5194/amt-6-2577-2013. URL https://amt.copernicus.org/articles/6/2577/2013/.

Mark A. Vaughan, Kathleen A. Powell, David M. Winker, Chris A. Hostetler, Ralph E. Kuehn, William H. Hunt, Brian J. Getzewich, Stuart A. Young, Zhaoyan Liu, and Matthew J. McGill. Fully Automated Detection of Cloud and Aerosol Layers in the CALIPSO Lidar Measurements. *Journal of Atmospheric and Oceanic Technology*, 26(10):2034–2050, 10 2009. ISSN 0739-0572. doi: 10.1175/2009JTECHA1228.1. URL https://doi.org/10.1175/2009JTECHA1228.1.

J. Pepijn Veefkind, Gerrit de Leeuw, and Phillip A. Durkee. Retrieval of aerosol optical depth over land using two-angle view satellite radiometry during tarfox. *Geophysical Research Letters*, 25(16):3135–3138, 1998. doi: https://doi.org/10.1029/98GL02264. URL https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/98GL02264.

J.P. Veefkind, I. Aben, K. McMullan, H. Förster, J. de Vries, G. Otter, J. Claas, H.J. Eskes, J.F. de Haan, Q. Kleipool, M. van Weele, O. Hasekamp, R. Hoogeveen, J. Landgraf, R. Snel, P. Tol, P. Ingmann, R. Voors, B. Kruizinga, R. Vink, H. Visser, and P.F. Levelt. 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 of Environment*, 120:70 – 83, 2012. ISSN 0034-4257. doi: https://doi.org/10.1016/j.rse.2011.09.027. URL http://www.sciencedirect.com/science/article/pii/S0034425712000661. The Sentinel Missions - New Opportunities for Science.

J.Pepijn Veefkind, Gerrit de Leeuw, Piet Stammes, and Robert B.A Koelemeijer. Regional distribution of aerosol over land, derived from atsr-2 and gome. *Remote Sensing of Environment*, 74(3):377 – 386, 2000. ISSN 0034-4257. doi: https://doi.org/10.1016/S0034-4257(00)00106-1. URL http://www.sciencedirect.com/science/article/pii/S0034425700001061.

Jean-Paul Vernier, T. Duncan Fairlie, Terry Deshler, Murali Natarajan, Travis Knepp, Katie Foster, Frank G. Wienhold, Kristopher M. Bedka, Larry Thomason, and Charles Trepte. In situ and space-based observations of the kelud volcanic plume: The persistence of ash in the lower stratosphere. *Journal of Geophysical Research: Atmospheres*, 121(18):11,104–11,118, 2016. doi: 10.1002/2016JD025344. URL https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1002/2016JD025344.

T. H. Virtanen, P. Kolmonen, E. Rodríguez, L. Sogacheva, A.-M. Sundström, and G. de Leeuw. Ash plume top height estimation using aatsr. *Atmospheric Measurement Techniques*, 7(8):2437–2456, 2014. doi: 10.5194/amt-7-2437-2014. URL https://www.atmos-meas-tech.net/7/2437/2014/.

S. Vucetic, B. Han, W. Mi, Z. Li, and Z. Obradovic. A data-mining approach for the validation of aerosol retrievals. *IEEE Geoscience and Remote Sensing Letters*, 5(1):113–117, Jan 2008. ISSN 1558-0571. doi: 10.1109/LGRS.2007.912725.

Jun Wang, Xiaoguang Xu, Shouguo Ding, Jing Zeng, Robert Spurr, Xiong Liu, Kelly Chance, and Michael Mishchenko. 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, 2014. doi: https://doi.org/10.1016/j.jqsrt.2014.03.020. URL http://www.sciencedirect.com/science/article/pii/S0022407314001368.

Kaicun Wang, Robert E. Dickinson, and Shunlin Liang. Clear sky visibility has decreased over land globally from 1973 to 2007. *Science*, 323(5920):1468–1470, 2009. ISSN 0036-8075. doi: 10.1126/science.1167549. URL https://science.sciencemag.org/content/323/5920/1468.

P. Wang, P. Stammes, R. van der A, G. Pinardi, and M. van Roozendael. Fresco+: an improved o_2 a-band cloud retrieval algorithm for tropospheric trace gas retrievals. *Atmospheric Chemistry and Physics*, 8(21):6565–6576, 2008a. doi: 10.5194/acp-8-6565-2008. URL https://acp.copernicus.org/articles/8/6565/2008/.

P. Wang, O. N. E. Tuinder, L. G. Tilstra, M. de Graaf, and P. Stammes. Interpretation of fresco cloud retrievals in case of absorbing aerosol events. *Atmospheric Chemistry and Physics*, 12(19):9057–9077, 2012. doi: 10.5194/acp-12-9057-2012. URL https://acp.copernicus.org/articles/12/9057/2012/.

Xin Wang, Jianping Huang, Mingxia Ji, and Kaz Higuchi. Variability of east asia dust events and their long-term trend. *Atmospheric Environment*, 42(13):3156 – 3165, 2008b. ISSN 1352-2310. doi: https://doi.org/10.1016/j.atmosenv.2007.07.046. URL http://www.sciencedirect.com/science/article/pii/S1352231007006218.

Z. Wang, Z. Ulanowski, and P. H. Kaye. On solving the inverse scattering problem with rbf neural networks: Noise-free case. *Neural Computing & Applications*, 8(2):177–186, 1999. doi: 10.1007/s005210050019. URL https://doi.org/10.1007/s005210050019.

F. Waquet, B. Cairns, K. Knobelspiesse, J. Chowdhary, L. D. Travis, B. Schmid, and M. I. Mishchenko. Polarimetric remote sensing of aerosols over land. *Journal of Geophysical Research: Atmospheres*, 114(D1), 2009. doi: 10.1029/2008JD010619. URL https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2008JD010619.

P. C. Waterman. Matrix formulation of electromagnetic scattering. *Proceedings of the IEEE*, 53(8):805–812, 1965.

D. Watson-Parris, N. Bellouin, L. T. Deaconu, N. A. J. Schutgens, M. Yoshioka, L. A. Regayre, K. J. Pringle, J. S. Johnson, C. J. Smith, K. S. Carslaw, and P. Stier. Constraining uncertainty in aerosol direct forcing. *Geophysical Research Letters*, 47(9):e2020GL087141, 2020. doi: 10.1029/2020GL087141. URL https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2020GL087141. e2020GL087141 2020GL087141.

Jing Wei, Zhanqing Li, Yiran Peng, and Lin Sun. Modis collection 6.1 aerosol optical depth products over land and ocean: validation and comparison. *Atmospheric Environment*, 201:428 – 440, 2019. ISSN 1352-2310. doi: https://doi.org/10.1016/j.atmosenv.2018.12.004. URL http://www.sciencedirect.com/science/article/pii/S1352231018308513.

Yuanyuan Wei, Zhengqiang Li, Ying Zhang, Cheng Chen, Oleg Dubovik, Yang Zhang, Hua Xu, Kaitao Li, Jie Chen, Haofei Wang, Bangyu Ge, and Cheng Fan. Validation of polder grasp aerosol optical retrieval over china using sonet observations. *Journal of Quantitative Spectroscopy and Radiative Transfer*, 246:106931, 2020. ISSN 0022-4073. doi: https://doi.org/10.1016/j.jqsrt.2020.106931. URL http://www.sciencedirect.com/science/article/pii/S0022407320300650.

Ellsworth J. Welton, James R. Campbell, James D. Spinhirne, and V. Stanley Scott III. Global monitoring of clouds and aerosols using a network of micropulse lidar systems. In Upendra N. Singh, Kazuhiro Asai, Toshihiro Ogawa, Upendra N. Singh, Toshikasu Itabe, and Nobuo Sugimoto, editors, *Lidar Remote Sensing for Industry and Environment Monitoring*, volume 4153, pages 151 – 158. International Society for Optics and Photonics, SPIE, 2001. doi: 10.1117/12.417040. URL https://doi.org/10.1117/12.417040.

Congcong Wen, Shufu Liu, Xiaojing Yao, Ling Peng, Xiang Li, Yuan Hu, and Tianhe Chi. A novel spatiotemporal convolutional long short-term neural network for air pollution prediction. *Science of The Total Environment*, 654:1091–1099, 2019. ISSN 0048-9697. doi: https://doi.org/10.1016/j.scitotenv.2018.11.086. URL https://www.sciencedirect.com/science/article/pii/S0048969718344413.

Jason Weston, Sayan Mukherjee, Olivier Chapelle, Massimiliano Pontil, Tomaso Poggio, and Vladimir Vapnik. Feature selection for svms. In T. K. Leen, T. G. Dietterich, and V. Tresp, editors, *Advances in Neural Information Processing Systems 13*, pages 668–674. MIT Press, 2001. URL http://papers.nips.cc/paper/1850-feature-selection-for-svms.pdf.

WHO. Health aspects of air pollution with particulate matter, ozone and nitrogen dioxide : report on a who working group, bonn, germany 13-15 january 2003. page EUR/03/5042688, 2003.

WHO. Air quality guidelines: global update 2005: particulate matter, ozone, nitrogen dioxide, and sulfur dioxide. page EUR/05/5046029, 2006.

David M. Winker, William H. Hunt, and Matthew J. McGill. Initial performance assessment of caliop. *Geophysical Research Letters*, 34(19), 2007. doi: 10.1029/2007GL030135. URL https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2007GL030135.

David M. Winker, Mark A. Vaughan, Ali Omar, Yongxiang Hu, Kathleen A. Powell, Zhaoyan Liu, William H. Hunt, and Stuart A. Young. Overview of the CALIPSO Mission and CALIOP Data Processing Algorithms. *Journal of Atmospheric and Oceanic Technology*, 26(11):2310–2323, 11 2009. ISSN 0739-0572. doi: 10.1175/2009JTECHA1281.1. URL https://doi.org/10.1175/2009JTECHA1281.1.

Lianghai Wu, Otto Hasekamp, Bastiaan van Diedenhoven, Brian Cairns, John E. Yorks, and Jacek Chowdhary. Passive remote sensing of aerosol layer height using near-uv multiangle polarization measurements. *Geophysical Research Letters*, 43(16):8783–8790, 2016a. doi: https://doi.org/10.1002/2016GL069848. URL https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1002/2016GL069848.

Lianghai Wu, Otto Hasekamp, Bastiaan van Diedenhoven, Brian Cairns, John E. Yorks, and Jacek Chowdhary. Passive remote sensing of aerosol layer height using near-uv multiangle polarization measurements. *Geophysical Research Letters*, 43(16):8783–8790, 2016b. doi: 10.1002/2016GL069848. URL https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1002/2016GL069848.

Yerong Wu, Martin de Graaf, and Massimo Menenti. The sensitivity of aod retrieval to aerosol type and vertical distribution over land with modis data. *Remote Sensing*, 8(9):765, Sep 2016c. ISSN 2072-4292. doi: 10.3390/rs8090765. URL http://dx.doi.org/10.3390/rs8090765.

Yerong Wu, Martin de Graaf, and Massimo Menenti. The impact of aerosol vertical distribution on aerosol optical depth retrieval using calipso and modis data: Case study over dust and smoke regions. *Journal of Geophysical Research: Atmospheres*, 122(16):8801–8815, 2017. doi: 10.1002/2016JD026355. URL https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1002/2016JD026355.

Jie Xiong, Runming Yao, Wenbo Wang, Wei Yu, and Baizhan Li. A spatial-and-temporal-based method for rapid particle concentration estimations in an urban environment. *Journal of Cleaner Production*, 256:120331, 2020. ISSN 0959-6526. doi: https://doi.org/10.1016/j.jclepro.2020.120331. URL https://www.sciencedirect.com/science/article/pii/S0959652620303784.

X. Xu, J. Wang, Y. Wang, J. Zeng, O. Torres, J. S. Reid, S. D. Miller, J. V. Martins, and L. A. Remer. Detecting layer height of smoke aerosols over vegetated land and water surfaces via oxygen absorption bands: hourly results from epic/dscovr in deep space. *Atmospheric Measurement Techniques*, 12(6):3269–3288, 2019. doi: 10.5194/amt-12-3269-2019. URL https://www.atmos-meas-tech.net/12/3269/2019/.

Xiaoguang Xu and Jun Wang. Retrieval of aerosol microphysical properties from aeronet photopolarimetric measurements: 1. information content analysis. *Journal of Geophysical Research: Atmospheres*, 120(14):7059–7078, 2015. doi: 10.1002/2015JD023108. URL https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1002/2015JD023108.

Xiaoguang Xu, Jun Wang, Yi Wang, Jing Zeng, Omar Torres, Yuekui Yang, Alexander Marshak, Jeffrey Reid, and Steve Miller. 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. *Geophysical Research Letters*, 44(14):7544–7554, 2017. doi: 10.1002/2017GL073939. URL https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1002/2017GL073939.

Xiaoguang Xu, Jun Wang, Yi Wang, and Alexander Kokhanovsky. Chapter 1 - passive remote sensing of aerosol height. In Tanvir Islam, Yongxiang Hu, Alexander Kokhanovsky, and Jun Wang, editors, *Remote Sensing of Aerosols, Clouds, and Precipitation*, pages 1 – 22. Elsevier, 2018a. ISBN 978-0-12-810437-8. doi: https://doi.org/10.1016/B978-0-12-810437-8.00001-3. URL http://www.sciencedirect.com/science/article/pii/B9780128104378000013.

Xiaoguang Xu, Jun Wang, Yi Wang, Alexander Kokhanovsky, Tanvir Islam, and Yongxiang Hu. *Chapter 1 - Passive Remote Sensing of Aerosol Height*, pages 1–22. Elsevier, 2018b. ISBN 978-0-12-810437-8. doi: https://doi.org/10.1016/B978-0-12-810437-8.00001-3. URL http://www.sciencedirect.com/science/article/pii/B9780128104378000013.

Juping Yan, Xiaoping Wang, Ping Gong, Chuanfei Wang, and Zhiyuan Cong. Review of brown carbon aerosols: Recent progress and perspectives. *Science of The Total Environment*, 634:1475 – 1485, 2018. ISSN 0048-9697. doi: https://doi.org/10.1016/j.scitotenv.2018.04.083. URL http://www.sciencedirect.com/science/article/pii/S0048969718312476.

Howell S. G. Zhuang J. Yang, M. and B. J. Huebert. Attribution of aerosol light absorption to black carbon, brown carbon, and dust in china – interpretations of atmospheric measurements during east-aire. *Atmos. Chem. Phys.*, 9:2035–2050, 2009. doi: 10.5194/acp-9-2035-2009.

Jiayun Yao, Sean M. Raffuse, Michael Brauer, Grant J. Williamson, David M.J.S. Bowman, Fay H. Johnston, and Sarah B. Henderson. Predicting the minimum height of forest fire smoke within the atmosphere using machine learning and data from the calipso satellite. *Remote Sensing of Environment*, 206:98 – 106, 2018. ISSN 0034-4257. doi: https://doi.org/10.1016/j.rse.2017.12.027. URL http://www.sciencedirect.com/science/article/pii/S003442571730603X.

X. Yao, L.G. Tham, and F.C. Dai. Landslide susceptibility mapping based on support vector machine: A case study on natural slopes of hong kong, china. *Geomorphology*, 101(4):572 – 582, 2008. ISSN 0169-555X. doi: https://doi.org/10.1016/j.geomorph.2008.02.011. URL http://www.sciencedirect.com/science/article/pii/S0169555X08000585.

Bangsheng Yin, Qilong Min, and Everette Joseph. Retrievals and uncertainty analysis of aerosol single scattering albedo from mfrsr measurements. *Journal of Quantitative Spectroscopy and Radiative Transfer*, 150:95 – 106, 2015. ISSN 0022-4073. doi: https://doi.org/10.1016/j.jqsrt.2014.08.012. URL http://www.sciencedirect.com/science/article/pii/S0022407314003562. Topical issue on optical particle characterization and remote sensing of the atmosphere: Part I.

Hongbin Yu, Mian Chin, David M. Winker, Ali H. Omar, Zhaoyan Liu, Chieko Kittaka, and Thomas Diehl. Global view of aerosol vertical distributions from calipso lidar measurements and gocart simulations: Regional and seasonal variations. *Journal of Geophysical Research: Atmospheres*, 115(D4), 2010. doi: 10.1029/2009JD013364. URL https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2009JD013364.

T. Yuan, L. A. Remer, and H. Yu. Microphysical, macrophysical and radiative signatures of volcanic aerosols in trade wind cumulus observed by the a-train. *Atmospheric Chemistry and Physics*, 11(14):7119–7132, 2011. doi: 10.5194/acp-11-7119-2011. URL https://acp.copernicus.org/articles/11/7119/2011/.

Yu Zhan, Yuzhou Luo, Xunfei Deng, Huajin Chen, Michael L. Grieneisen, Xueyou Shen, Lizhong Zhu, and Minghua Zhang. Spatiotemporal prediction of continuous daily pm2.5 concentrations across china using a spatially explicit machine learning algorithm. *Atmospheric Environment*, 155:129 – 139, 2017. ISSN 1352-2310. doi: https://doi.org/10.1016/j.atmosenv.2017.02.023. URL http://www.sciencedirect.com/science/article/pii/S1352231017300936.

H. Zhang, Z. Shen, X. Wei, M. Zhang, and Z. Li. Comparison of optical properties of nitrate and sulfate aerosol and the direct radiative forcing due to nitrate in china. *Atmospheric Research*, 113:113 – 125, 2012. ISSN 0169-8095. doi: https://doi.org/10.1016/j.atmosres.2012.04.020. URL http://www.sciencedirect.com/science/article/pii/S0169809512001330.

L. Zhang, L. Zhang, and B. Du. Deep learning for remote sensing data: A technical tutorial on the state of the art. *IEEE Geoscience and Remote Sensing Magazine*, 4(2):22–40, June 2016. ISSN 2168-6831. doi: 10.1109/MGRS.2016.2540798.

Jingxia Zhao, Richard P. Turco, and Owen B. Toon. A model simulation of pinatubo volcanic aerosols in the stratosphere. *Journal of Geophysical Research: Atmospheres*, 100(D4):7315–7328, 1995. doi: 10.1029/94JD03325. URL https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/94JD03325.

Guanhua Zhou, Wujian Xu, Chunyue Niu, and Huijie Zhao. The polarization patterns of skylight reflected off wave water surface. *Opt. Express*, 21(26):32549–32565, Dec 2013. doi: 10.1364/OE.21.032549. URL http://www.opticsexpress.org/abstract.cfm?URI=oe-21-26-32549.

X. X. Zhu, D. Tuia, L. Mou, G. Xia, L. Zhang, F. Xu, and F. Fraundorfer. Deep learning in remote sensing: A comprehensive review and list of resources. *IEEE Geoscience and Remote Sensing Magazine*, 5(4):8–36, Dec 2017. ISSN 2168-6831. doi: 10.1109/MGRS.2017.2762307.

Xiudi Zhu, Qiang Zhang, Chong-Yu Xu, Peng Sun, and Pan Hu. Reconstruction of high spatial resolution surface air temperature data across china: A new geo-intelligent multisource data-based machine learning technique. *Science of The Total Environment*, 665:300–313, 2019. ISSN 0048-9697. doi: https://doi.org/10.1016/j.scitotenv.2019.02.077. URL https://www.sciencedirect.com/science/article/pii/S0048969719305571.

A

ABBREVIATIONS

Tabel A.1: List of abbreviations.

3MI	Multi-View Multi-Channel Multi-Polarization Imaging
AATSR	Advanced Along-Track Scanning Radiometer
ACE Asia	Aerosol Characterization Experiments-Asia
ADEOS	Advanced Earth Observing Satellite
ADM-Aeolus	Atmospheric Dynamics Mission Aeolus
AERONET	AErosol RObotic NETwork
AIRS	Atmospheric InfraRed Sounder
APS	Aerosol Polarimetry Sensor
ATSR	Along Track Scanning Radiometer
AVHRR	Advanced Very High Resolution Radiometer
CALIOP	Cloud-Aerosol Lidar with Orthogonal Polarization
CALIPSO	Cloud-Aerosol Lidar and Infrared Pathfinder Satellite Observation
CAMS	Copernicus Atmosphere Monitoring Service
DISAMAR	Determining Instrument Specifications and Analysing Methods for Atmospheric Retrieval
EARLINET	European Aerosol Research Lidar Network
EPIC	Earth Polychromatic Imaging Camera
EPS-SG	EUMETSAT Polar System – Second Generation
ESA	European Space Agency
FRESCO	Fast Retrieval Scheme for Clouds from the Oxygen A band
GEOS-5	Goddard Earth Observing System, version 5
GLAS	Geoscience Laser Altimeter System
GMAO	Global Modeling and Assimilation Office
GOCART	Goddard Chemistry Aerosol Radiation and Transport
GOME	Global Ozone Monitoring Experiment
GRASP	Generalized Retrieval of Aerosol and Surface Properties
GSI	Gridpoint Statistical Interpolation
IASI	Infrared Atmospheric Sounding Interferometer
ICESat	Ice, Cloud and Elevation Satellite
INTEX-A	Intercontinental Chemical Transport Experiment–North America
IPCC	Intergovernmental Panel on Climate Change
JPL	Jet Propulsion Laboratory
KNMI	Royal Netherlands Meteorological Institute
LIVAS	LIdar climatology of Vertical Aerosol Structure for space-based lidar simulation studies
MAIA	Multi-Angle Imager for Aerosols
MERIS	MEdium Resolution Imaging Spectrometer
MERRA-2	Modern-Era Retrospective Analysis for Research and Applications, version 2
Metop	Meteorological Operational
MISR	The multi-angle imaging spectroradiometer
MODIS	MODerate-resolution Imaging Spectroradiometer
MPLNET	Micro-Pulse Lidar Network
NASA	National Aeronautics and Space Administration
OMI	Ozone Monitoring Instrument
OMPS	Ozone Mapping and Profiler Suite
PACE	Plankton, Aerosol, Cloud and ocean Ecosystem
PARASOL	Polarization & Anisotropy of Reflectances for Atmospheric Sciences coupled with Observations from a Lidar
POLDER	POLarization and Directionality of the Earth's Reflectances
PRIDE	Puerto RIco Dust Experiment
S-5P	Sentinel-5 Precursor

SAFARI	Southern African Regional Science Initiative
SCIAMACHY	SCanning Imaging Absorption spectroMeter for Atmospheric CartograpHY
SeaWiFS	Sea-viewing Wide Field-of-view Sensor
SLSTR	Sea and Land Surface Temperature Radiometer
TARFOX	Tropospheric Aerosol Radiative Forcing Observational eXperiment
TOMS	Total Ozone Mapping Spectrometer
TROPOMI	TROPOspheric Monitoring Instrument
VIIRS	Visible Infrared Imaging Radiometer Suite
WHO	World Health Organization

ACKNOWLEDGEMENTS

Throughout my Ph.D career I have received a great deal of support and assistance.

First of all, I would give appreciation to my promoter Prof. dr. Pieternel Levelt, co-promoter Dr. ir. Pepijn Veefkind and supervisor Dr. Peter van Velthoven for providing me the opportunity to working at Royal Netherlands Meteorological Institute. I would like to thank to my supervisory team Pepijn Veefkind and Peter van Velthoven for their patient guidance and professional suggestions. The thesis has been greatly improved with the comments by the committee members. Thank you for all your contributions. I would thank specially to Prof. dr. Gerrit de Leeuw and Dr. Joanna Joiner that provided me very detailed and helpful suggestion on my dissertation.

I appreciate my colleagues at Royal Netherlands Meteorological Institute and Delft University of Technology who provided professional and technical supports in my Ph.D career, Deborah Stein Zweers, Gijs Tilstra, Johan de Haan, Julien Chimot, Maarten Sneep, Mark ter Linden, Martin de Graaf, Swadhin Nanda, Tim Vlemmix and Wang Ping. Special thanks to Piet Stammes who is always ready to help international junior scientists like us in terms of work.

Many thanks to Jacqueline Baas and other colleagues at Royal Netherlands Meteorological Institute, and the secretaries of Geoscience and Remote Sensing, Suzanne Dollée and Debbie Rietdijk, who helped me with administrative issues.

I would thank to my Chinese colleagues and friends in the Netherlands who accompanies with me during my six-year life aboard, Chuxiao Yu, Fei Liu, Feixiong Huang, Jianye Chen, Jieying Ding, Mengyao Liu, Mingqian Wang, Taiping Yang, Weicheng Ni, Xing'ou Xu, Yipu Wei, Yue Chen, Yunhe Yu and Zhen Li (with her cat Guaiguai). Look forward the next hot pot party with you!

I want to thank to Richard Bintanja who introduced me to badminton groups from which I knew a lot of badminton teammates, Aart Vermetten, Calvin, Cedric, Cisco de Bruijn, Erik Tuenter, Frank Selten, Jan Fokke Meirink, Ludo, Marie, Michiel Baatsen, Li Fang, Minglong Liu, Xiaoli Zhao, Xiaoyi Chen, Weichao Wang, Wentao Li, Yongliang Zhang and Yulong Zhao. I would specially thank to Noud, who taught me a lot of techniques and practiced with me on badminton and Xiaobin Xie, who invited me to join the competitions in the Netherlands. Playing badminton is one of my most delightful components of my life in the Netherlands.

I would also thank to my friends from primary and middle school who accompanied with my growth, Hanyu Tang, Kang Yuan, Leman Feng, Tianwen Lan, Wenqian Bai, Xiaoye Ma, Yaqi Chen, Yining Yao, Yun He, Yunhao Sun, Zhaowen Luo and Zhenyue Yu.

Finally, I would also thank to my families and my pet dog Civil for their spiritual supports.

List of Publications

Liu, M., van der A, R., van Weele, M., Eskes, H., Lu, X., Veefkind, P., de Laat, J., Kong, H., Wang, J., **Sun, J.**, Ding, J., Zhao, Y. and Weng, H.: A new divergence method to quantify methane emissions using observations of Sentinel-5P TROPOMI. Geophysical Research Letters, 48, e2021GL094151. https://doi.org/10.1029/2021GL094151, 2021.

Sun, J., Veefkind, J. P., van Velthoven, P., and Levelt, P. F.: Aerosol Absorption Over Land Derived From the Ultra-Violet Aerosol Index by Deep Learning, IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, vol. 14, pp. 9692-9710, https://doi.org/10.1109/JSTARS.2021.3108669, 2021.

Sun, J., Veefkind, J. P., van Velthoven, P., Tilstra, L. G., Chimot, J., Nanda, S., and Levelt, P. F.: Defining aerosol layer height for UVAI interpretation using aerosol vertical distributions characterized by MERRA-2, Atmos. Chem. Phys. Discuss., https://doi.org/10.5194/acp-2020-39, preprint, 2020.

Nanda, S., de Graaf, M., Veefkind, J. P., Sneep, M., ter Linden, M., **Sun, J.**, and Levelt, P. F.: A first comparison of TROPOMI aerosol layer height (ALH) to CALIOP data, Atmos. Meas. Tech., 13, 3043–3059, https://doi.org/10.5194/amt-13-3043-2020, 2020.

Sun, J., Veefkind, P., Nanda, S., van Velthoven, P., and Levelt, P.: The role of aerosol layer height in quantifying aerosol absorption from ultraviolet satellite observations, Atmos. Meas. Tech., 12, 6319–6340, https://doi.org/10.5194/amt-12-6319-2019, 2019.

Sun, J., Veefkind, J. P., van Velthoven, P., and Levelt, P. F.: Quantifying the single-scattering albedo for the January 2017 Chile wildfires from simulations of the OMI absorbing aerosol index, Atmos. Meas. Tech., 11, 5261–5277, https://doi.org/10.5194/amt-11-5261-2018, 2018.

CONFERENCES, SEMINARS AND WORKSHOPS

2022 Jun	Oral presentation in the International Radiation Symposium (IRS), online
2020 Dec	Poster presentation in the American Geophysical Union (AGU), online
2020 Oct	Poster presentation in the ECMWF-ESA Workshop on Machine Learning for Earth System Observation and Prediction, online
2019 Mar	Oral presentation in the Nederlands Aardwetenschappelijk Congres (NAC), Utrecht, the Netherlands
2018 Sept	Oral presentation in the OMI Science Team Meeting, De Bilt, the Netherlands
2018 May	Poster presentation in the European Geosciences Union (EGU), Vienna, Austria
2017 Apr	Poster presentation in the European Geosciences Union General Assembly, Vienna, Austria

CURRICULUM VITÆ

Jiyunting SUN

1992 Jun Born in Nanjing, Jiangsu, China.

EDUCATION

2010–2014 **Bachelor of Science**
Major in Management of Resources, Environment and Urban-Rural Planning
Nanjing Agriculture University, China

2014–2016 **Master of Science**
Major in Geoscience and Remote Sensing
Delft University of Technology, the Netherlands

2016–2022 **Ph.D Atmospheric Remote Sensing**
Faculty of Civil Engineering & Geosciences
Delft University of Technology, the Netherlands

Thesis:	Aerosol Absorption from Global Satellite Measurements in the Ultra-Violet
Promotor:	Prof. dr. P. F. Levelt
Co-promotor:	Dr. ir. P. J. Veefkind

TRAINING

2018 Dec Machine Learning provided by Coursera
2019 Dec Deep Learning Specialization provided by Coursera
2019 Dec Geo-statistics provided by the Graduate Schools for Production Ecology and Resource Conservation and for Wage-ningen Institute for Environment and Climate Research

WORK EXPERIENCE

2012 Feb 2013 Dec	**Intern** Jiangsu Institute of Urban Planning and Design, Nanjing, China
2015 Aug 2015 Oct	**Research intern** Royal Netherlands Meteorological Institute, De Bilt, the Netherlands
2015 Oct 2015 Dec	**Research intern** Netherlands National Institute for Public Health and the Environment, Bilthoven, the Netherlands
2022 Feb	**Data Analyst** Nanjing Pulan Atmospheric Environment Research Institute Co. LTD., Nanjing, China

dissertation

Propositions

accompanying the dissertation

AEROSOL ABSORPTION FROM GLOBAL SATELLITE MEASUREMENTS IN THE ULTRA-VIOLET

FROM QUALITATIVE AEROSOL INDEX TO QUANTITATIVE AEROSOL ABSORPTIVE PROPERTIES

by

Jiyunting SUN

1. The Ultra-Violet Aerosol Index (UVAI) cannot be quantitatively analyzed without information on aerosol loading and aerosol vertical distribution. (This proposition pertains to this dissertation.)
2. The Ultra-Violet Aerosol Index (UVAI) should receive more appreciations than it currently has obtained. (This proposition pertains to this dissertation.)
3. In the context of Machine Learning, the Ultra-Violet Aerosol Index (UVAI) itself is a result of feature engineering. (This proposition pertains to this dissertation.)
4. Data deluge saves time meanwhile wasting time. Remote sensing enlarges the data size, but the question is how much data are actually valid.
5. The First Law of Geography: *Everything is related to everything else, but nearby things are more related than distant things*, should also be taken into account in remote sensing retrievals.
6. Online radiative transfer calculation is a future trend in atmospheric remote sensing inversion problem solving; Machine Learning is another.
7. In Machine Learning, features are more important than algorithms, whereas data are more important than features.
8. Remote sensing retrievals are actually data science with specific domain knowledge.
9. Do not attach too much value on what you value.
10. Sometimes, research is just providing scientific evidence for what one already knows from common sense.

These propositions are regarded as opposable and defendable, and have been approved as such by the promotor Prof. dr. P.F. Levelt.