UC Berkeley
UC Berkeley Electronic Theses and Dissertations

Title
Aerosol Retrieval Using Remote-sensed Observations

Permalink
https://escholarship.org/uc/item/7zx6z7wh

Author
Wang, Yueqing

Publication Date
2012

Peer reviewed|Thesis/dissertation

eScholarship.org
Powered by the California Digital Library
University of California

Aerosol Retrieval Using Remote-sensed Observations

by

Yueqing Wang

A dissertation submitted in partial satisfaction of the
requirements for the degree of
Doctor of Philosophy

in

Statistics

and the Designated Emphasis

in

Communication, Computation, and Statistics

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor Bin Yu, Chair
Professor Cari Kaufman
Professor Michael Jerrett

Fall 2012

Aerosol Retrieval Using Remote-sensed Observations

Copyright 2012
by
Yueqing Wang

Abstract

Aerosol Retrieval Using Remote-sensed Observations

by

Yueqing Wang

Doctor of Philosophy in Statistics

with the Designated Emphasis in

Communication, Computation, and Statistics

University of California, Berkeley

Professor Bin Yu, Chair

Atmospheric aerosols are solid particles and liquid droplets that are usually smaller than the diameter of a human hair. They can be found drifting in the air in every ecosystem on Earth, leaving significant impacts on human health and our climate. Understanding the spatial and temporal distribution of different atmospheric aerosols, therefore, is an important first step to decode the complex system of aerosols and further, their effects on public health and climate.

The development of remote-sensing radiometers provides a powerful tool to monitor the amount of atmospheric aerosols, as well as their compositions. Radiometers aboard satellites measure the amount of electromagnetic solar radiation. The amount of atmospheric aerosols is further quantified by aerosol optical depth (AOD), defined as the amount of solar radiation that aerosols scatter and absorb in the atmosphere and generally prevent from reaching the Earth surface. Despite efforts to improve remote-sensing instruments and a great demand for a detailed profile of aerosol spatial distribution, methods needed to provide AOD estimation at a reasonably fine resolution, are lacking. The quantitative uncertainties in the amount of aerosols, and especially aerosol compositions, limit the utility of traditional methods for aerosol retrieval at a fine resolution.

In Chapter 2 and 3 of this thesis, we exploit the use of statistical methods to estimate aerosol optical depth using remote-sensed radiation. A Bayesian hierarchy proves to be useful for modeling the complicated interactions among aerosols of different amount and compositions over a large spatial area. Based on the hierarchical model, Chapter 2 estimates and validates aerosol optical depth using Markov chain Monte Carlo methods, while chapter 3 resorts to an optimization-based approach for faster computation. We extend our study focus from the aerosol amount to the aerosol compositions in Chapter 4.

Chapter 1 briefly reviews the characteristics of atmospheric aerosols, including the different types of aerosols and their major impacts on human health. We also introduce a major

remote-sensing instrument, NASA's Multi-angle Imaging SpectroRadiometer (MISR), which collects the observations our studies base on. Currently, the MISR operational aerosol retrieval algorithm provides estimates of aerosol optical depth at the spatial resolution of 17.6 km.

In Chapter 2, we embed MISR's operational weighted least squares criterion and its forward calculations for aerosol optical depth retrievals in a likelihood framework. We further expand it into a hierarchical Bayesian model to adapt to finer spatial resolution of 4.4 km. To take advantage of the spatial smoothness of aerosol optical depth, our method borrows strength from data at neighboring areas by postulating a Gaussian Markov Random Field prior for aerosol optical depth. Our model considers aerosol optical depth and mixing vectors of different types of aerosols as continuous variables. The inference is then carried out using Metropolis-within-Gibbs sampling methods. Retrieval uncertainties are quantified by posterior variabilities. We also develop a parallel Markov chain Monte Carlo algorithm to improve computational efficiency. We assess our retrieval performance using ground-based measurements from the AErosol RObotic NETwork (AERONET) and satellite images from Google Earth. Based on case studies in the greater Beijing area, China, we show that 4.4 km resolution can improve both the accuracy and coverage of remote-sensed aerosol retrievals, as well as our understanding of the spatial and seasonal behaviors of aerosols. This is particularly important during high-AOD events, which often indicate severe air pollution.

Chapter 3 of this thesis continues to improve our statistical aerosol retrievals for better accuracy and more efficient computation by switching to an optimization-based approach. We first establish objective functions for aerosol optical depth and aerosol compositions, based upon MISR operational weighted least squares criterion and its forward calculations. Our method also borrows strength from aerosol spatial smoothness by constructing penalty terms in the objective functions. The penalties correspond to a Gaussian Markov Random Field prior for aerosol optical depth and a Dirichlet prior for aerosol mixing vectors under our hierarchical Bayesian scheme; the optimization-based approach corresponds to Bayesian Maximum a Posteriori (MAP) estimation. Our MAP retrieval algorithm provides computational efficiency almost 60 times that of our Bayesian retrieval algorithm presented in Chapter 2. To represent the increasing heterogeneity of urban aerosol sources, our model continues to expand the pre-fixed aerosol mixtures used in the MISR operational algorithm by considering aerosol mixing vectors as continuous variables. Our retrievals are again validated using ground-based AERONET measurements. Case studies in the greater Beijing and Zhengzhou areas of China reassure that 4.4 km resolution can improve the accuracy and spatial coverage of remotely-sensed retrievals and enhance our understanding of the spatial behaviors of aerosols.

When comparing our aerosol retrievals to the extensive ground-based measurements collected in Baltimore, Maryland, we encountered greater uncertainties of aerosol compositions. It is a result from both the complex terrain structures of Baltimore and its various aerosol emission sources. Chapter 4, as result, extends the flexibility of our previous aerosol retrievals by incorporating a complete set of the eight commonly observed types of aerosols.

The consequential rise in model complexity is met by a warm-start Markov chain Monte Carlo sampling scheme. We first design two Markov sub-chains, each representing an aerosol mixture containing only four types of the commonly observed aerosols. Combining the samples generated by these two sub-chains, we propose an initialization for the Markov chain that contains all eight types of commonly observed aerosols. Partial information on the interactions of different types of aerosols from the samples generated by the sub-chains proves to be useful in choosing a more efficient initial point for the complete Markov chain. Faster computation is achieved without compromising the retrieval accuracy nor the spatial resolution of the estimated aerosol optical depth. In the end, through case studies of aerosol retrievals for the Baltimore area, we explore the potentials of remote-sensed retrievals in improving our understanding of aerosol compositions.

Contents

List of Figures

List of Tables

Acknowledgments

First, I would like to express my appreciation and gratitude to my advisor, Professor Bin Yu, for her endless support and guidance. Her course on Applied Statistics, Theory and Practice, enlightened me tremendously and finally set the research focus of my PhD, as well as that of my future career path. It is truly a privilege to have been given the opportunity to learn from her vast knowledge and priceless experiences in using statistics to solve real-world problems, and to grow as an applied statistician under her close influence. I particularly admire Bin's ability to build connections among various research areas as well as researchers. Her guidance and advice went far beyond statistical research. She truly believes in each one of her students and helping us grow both as a researcher and a person has always been top of her priorities. Her trust and influence has guided me through difficult times in my PhD. I am ever grateful to her for inspiring me to thrive into a better self, to shoulder greater responsibilities, and to find one's passions in life and true appreciations of the people, through her constant encouragement, motivation and personal example.

Further, I would like to thank Professor Cari Kaufman for providing me with encouragement and support throughout my PhD. She always has an open door for me, be that research questions or career advice. Her course on Spatial Statistics is one of the most helpful and enjoyable classes during my PhD time at Berkeley. Her provision of teaching, time, and patience have always been an invaluable support.

I would also like to thank Professor Michael Jerrett, my outside member, for his great advice and support. I very much enjoyed the discussions with him, which have provided me with great exposure and insights to many different areas of research at the interface of statistics and public health.

Since the first semester at Berkeley, I have been engaged in various discussions and reading studies with Bin's research group, and in particular with Taesup Moon in the last year of my PhD. I would like to thank the whole group for countless great discussions and constant support and encouragement, as well as for teaching me so much on various topics at the frontier of statistical research, from causal inferences to geometric function analysis. Taesup has been a superb collaborator, his attentiveness to details and grasps of different segments of a project have always inspired me. I would also like to thank Yuval Benjamini and Derek Bean for their great help with my research and academic writing.

I have also received valuable guidance and support for my research from members outside Berkeley. I am especially grateful to Professor Ming Jiang and Xin Jiang at the Peking University, who helped introduce me to and get me started on the problem of remote-sensed aerosol retrievals, which led to Chapter 2 of my thesis; the work and influence continue on.

Further, I would like to thank Professor Yang Liu at the Emory University for his valuable perspectives and advice on public health related aerosol studies. I would also like to thank Doctor Amy Braverman, Doctor Ralph Kahm and the MISR team from the Jet Propulsion Laboratory at the California Institute of Technology, for many insightful discussions and great encouragements.

I am very grateful to the Department of Statistics at the University of California, Berkeley, for supporting me financially during the full five years of my PhD. Besides giving me the opportunity to attend one of the best graduate programs in statistics, the department provided me with valuable advice, a network of great mentors and peers, and very useful seminars and discussions.

Finally, my PhD would not have been as enlightening and enjoyable without the incredible interactions with and support from my family and friends. I would like to thank my parents for always being there for me with support and advice and for reminding me that there are greater responsibilities and values beyond my studies. Professor Sharon Murrel has always been an invaluable mentor and friend ever since my undergraduate studies. Her influence has guided me throughout my work and life and I am truly grateful. I am really fortunate to have met Caroline Uhler, Meromit Singer, Anna Sapfo and Garvesh Raskutti during the first semester here at Berkeley. I would like to thank them for growing with me throughout the past five years, for being wonderful and supportive friends and for all the fun time we spent together.

Chapter 1

Introduction

The goal of this first chapter is to provide a brief background on atmospheric aerosols and their impacts on human health.

1.1 Aerosols and Their Compositions

Atmospheric aerosols float between anywhere from the Earth's stratosphere to the surface and vary widely in size, shape and chemical compositions. Aerosols can further clump together to form even more complex mixtures of a larger variety. Commonly observed aerosols can be grouped into major categories according to their chemical compositions, including sulfates, organic carbon, black carbon, nitrates, mineral dust and sea salt [3]. Among them, sulfates from volcanic ash, sea salt and dust are three most common types.

Approximately 90 percent of atmospheric aerosols are generated by natural processes. Examples of such natural origins include volcanoes ejecting ash, sulfur dioxide and other gases, forest fires releasing burned organic carbon, certain microalgae in the ocean producing a sulfurous gas, as well as sandstorms carrying mineral dust from deserts into the atmosphere.

The remaining 10 percent of aerosols are generated by anthropogenic activities. Despite the small overall proportion of man-made aerosols, they can dominate the atmosphere above urban and industrial regions, as we will see in several case studies throughout this thesis. Examples of anthropogenic origins include fossil fuel combustion generating large amount of sulfate aerosols, biomass burning producing smoke of organic carbon and black carbon, automobiles and power plants creating sulfates, nitrates, black carbon and other particles.

1.2 Aerosols Impacts on Human Health

Atmospheric aerosols can significantly affect human health and life expectancy [29]. When inhaled, aerosols can penetrate cell membranes, then migrate and seriously damage human respiratory, cardiovascular systems [28] and the brain [24]. Short-term impacts include: irritation to eyes, nose and throat; upper respiratory infections including pneumonia and bronchitis; and stroke or death from cardiovascular causes. Continual exposure to hazardous aerosols can aggravate or complicate medical conditions in the elderly[4]; aerosols from silica and diesel can lead to diseases including silicosis and black lung. Aerosols with an aerodynamic diameter less than 2.5 μm, such as black carbon, can severely reduce ground-level visibility. Understanding public health concerns, therefore, can largely benefit from an informative profile of aerosols' optical properties, especially in urban areas with dense population and complex anthropogenic aerosol sources. In fact, efforts to broaden spatial coverage and measures of aerosol characteristics in order to assist public health studies are emphasized by NOAA, NASA and EPA [34].

Profiling spatial distribution of aerosols at fine resolution is thus critical for air quality and public health studies, especially in urban areas with complex anthropogenic aerosol sources discussed in the previous section.

China, one of the most aerosol-laden countries, harbors fast-developing cities such as Beijing, which produce heavy layers of complex mixtures of aerosols. Anthropogenic deforestation in northern China leads to more frequent and severe sandstorms, which carry dust from the deserts in the north to other parts of China and even Japan. Given these concerns and the author's Chinese background, it is only natural to start with cities in China as locations in our case studies of Chapter 2 and 3.

1.3 Focus of Our Studies

The main goal of our studies is to provide a profile of spatial distribution of aerosol optical depth at resolution fine enough to facilitate research on aerosols and their impact on urban public health. Such research is especially important in highly-populated urban areas with increasingly heterogeneous aerosol sources. Using the top-of-atmosphere radiances observed by NASA's Multi-angle Imaging SpectroRadiometer (MISR) aboard Terra satellite, the current MISR operational algorithm retrieves aerosol optical depth at 17.6 km resolution. Our work refines the retrieval resolution to 4.4 km to better support urban public health studies while also improving retrieval accuracy and spatial coverage.

In general, there are two types of remote-sensed aerosol retrieval algorithms: over-land and over-ocean. In this thesis, we restrict to aerosol retrievals over land.

Chapter 2

A Hierarchical Bayesian Approach for Aerosol Retrieval Using MISR Data

2.1 Motivation

There are two approaches to measure the spatial distribution of aerosols: through ground-based measurements or remote-sensed radiance imageries. Both quantify the amount of aerosols by spectral Aerosol Optical Depth (AOD), defined as the negative logarithm of the fraction of radiation (sunlight) not scattered or absorbed by aerosols on a path in the Earth's atmosphere[1]. Another major property of aerosols, aerosol single scattering albedo (SSA)[2], can also be estimated based on reflected solar radiation field measurements obtained from downward looking remote-sensing radiometers [20].

AOD at different spectral bands can be viewed as known functions of AOD at the green band using the Angstr$\ddot{o}$m power law[23]. For notational simplicity, this thesis refers to AOD at the green band. With either ground or remote-sensing approach, the spatial and temporal variabilities of aerosols require continual observations and computationally efficient analyses.

The AErosol RObotic NETwork (AERONET)[14] provides a data archive of local AOD values using a network of automatic sun photometers (Figure 2.1, left panel) located at more than 400 stations on the Earth's surface. It measures AOD from every half hour to every two hours, with uncertainties $< \pm 0.01$ at wavelengths > 440 nm[15]. AERONET measurements are widely accepted as a gold standard to validate AOD estimates based on other data sources. The sparse and heterogeneous locations of AERONET stations, however, make it difficult to directly use their measurements to study the spatial behaviors of aerosols.

---

[1] For example, an AOD value of 2.5 corresponds to 92% of radiation scattered or absorbed.

[2] SSA is defined as the ratio of scattered radiation to total extinct radiation (scattered and absorbed).

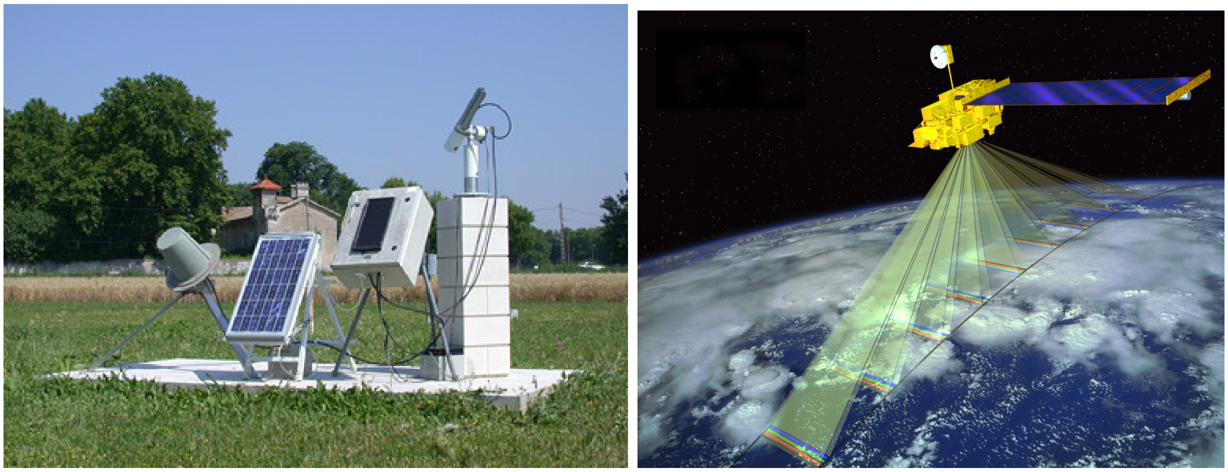


Figure 2.1: AERONET sun photometer at Avignon, France (left) and MISR cameras (right).

Remote-sensing radiometers offer a better spatial coverage by retrieving AOD from radiance imageries over the Earth's entire surface, such as the Multi-angle Imaging Spectro-Radiometer (MISR) aboard the NASA Earth Observing System Terra satellite (Figure 2.1, right panel). MISR views the day-lit Earth atmosphere almost simultaneously at nine angles along its track. On the observed imagery, the spacing between centers of each pixel is 275 m in the red band, and 1.1 km for the blue, green and near-infrared bands [7]. This unique design of multiple viewing angles renders an enhanced sensitivity to aerosol scattering, as well as an advantage in identifying near-source aerosol plume height [25][19]. MISR outputs four-spectral imageries at 1.1 km resolution for the blue, green, near-infrared bands, and at 275 m for the red band. Based on these imageries, MISR then produces AOD retrievals at 17.6 km resolution.

To quantitatively represent aerosol types, the MISR operational retrieval algorithm categorizes aerosol particles according to their properties such as characteristic radius and SSA. Each category, that is, each individual aerosol component, has a single type given by the refractive indices and a single particle size distribution. The resulting 21 individual aerosol components are used to define aerosol mixtures. To simplify remote-sensing retrieval, MISR considers only 74 aerosol mixtures, each defined by a set of two or three aerosol components and their mixing vector, indicating the relevant abundance of each component. For example, a dusty continental aerosol mixture can be represented by 80% of component "Sulfate mode 1", 10% of component "Mineral dust accumulation 2" and 10% of component "Mineral dust coarse".

In general, an aerosol mixture can now be identified by a notion of composition: a collection of M component aerosols and their mixing vector relative to these M components. Elements of the M-dimensional mixing vector sum up to 1, indicating mixing percentages of the M components. Based on the 74 MISR-prefixed mixtures and a discrete grid of

AOD values, forward radiative transfer (RT) calculations simulate radiances in the 36 MISR channels (9 viewing angles × 4 spectral bands). The results are stored in the Simulated MISR Ancillary Radiative Transfer (SMART) dataset. The MISR operational retrieval algorithm then compares the simulated radiances to the MISR observed radiances. A weighted least squares criterion is used to determine whether the simulated radiances provide good fits to the data, i.e., whether "successful" retrievals have been achieved. The averages of AOD values associated with successful retrievals are MISR retrieval outputs. They have been validated by ground measurements such as the AErosol RObotic NETwork (AERONET)[18][26], as well as by MISR field instruments and airplane campaigns[8]. Moreover, MISR retrieved AOD values have been shown to be informative in studies on wildfire smoke[17], mineral dust[21], and climate changing aerosols[32]; remote-sensed AOD estimates are used in conjunction with ground measurements for operational $\mathrm{PM}_{2.5}$ air quality forecasting [9][34].

MISR's ability to capture aerosol-related information makes it well suited to assist studies on aerosols' impact on public health. However, the heterogeneity of urban aerosols within an area of 17.6×17.6 km^2, the spatial resolution of MISR AOD retrievals makes finer resolution desirable. For example, San Francisco is represented by less than half of a MISR pixel. Yet the residents of San Francisco are exposed to varying levels of air pollution. Case studies in Delhi show that 5-km AOD has a significantly higher association with health-related particulate matters than AOD of rougher resolution [22]. As a result, we use 4.4 km as our retrieval resolution, also to be compatible with the MISR observations at 1.1 km. Also, observational studies indicate that tropospheric aerosol burden has increased at mid-latitudes and in the Arctic, probably due to anthropogenic activities[27][31]. Retrieval algorithms using pre-fixed aerosol mixtures are therefore limited. That is, when none of the 74 MISR aerosol mixtures prove to be good fits to the observations, no AOD retrievals are achieved. For example, 15.4% of the MISR missing retrievals for the greater Beijing area on March 15, 2009, was due to lack of fitting aerosol mixtures among the 74 MISR mixtures. This suggests that more varieties beyond the 74 pre-fixed aerosol compositions are to be considered in order to capture aerosols' growing heterogeneity.

Finer-resolution retrievals with greater varieties of aerosol compositions lead to a larger number of parameters to estimate. This is possible if we take advantage of AOD's spatial smoothness and reduce the 21 component aerosols to a smaller subset, say four, chosen according to current knowledge of the study region's aerosol conditions. This choice of incorporating four component aerosols is adequate for areas like Beijing which have relatively homogeneous spatial characteristics. Later, however, it proves to be insufficient for areas of a more complex terrain structure and we accordingly work to expand the model in Chapter 4.

In particular, a hierarchical Bayesian model is proposed to retrieve AOD values and mixing vectors based on MISR observations at 4.4 km resolution. We adopt a likelihood framework based on MISR's weighted least squares and construct the Bayesian hierarchy to incorporate AOD's spatial smoothness using a Gaussian Markov Random Field (GMRF) prior. The movement and dispersion of air particles in the atmosphere justify the spatial smoothness of AOD from a physical viewpoint. To flexibly describe various aerosol conditions, our model regards AOD values and mixing vectors as continuous parameters. This expands the set of possible compositions beyond the 74 pre-fixed choices of MISR. We show how this enriched variety is necessary to retrieve heterogeneous urban aerosols.

Throughout this thesis, our study takes a MISR Block[3] as a data unit to balance the coverage of a greater metropolitan area and computational cost.

The posterior inference of AOD and mixing vectors is carried out using Markov Chain Monte Carlo (MCMC) sampling methods, particularly Metropolis-within-Gibbs. Such sampling methods allow us to quantify the retrieval uncertainties by posterior variabilities. The algorithm, however, is computationally intense. We develop a parallel MCMC algorithm by partitioning a MISR Block into smaller patches, to enable parallel samplings while maintaining the overall smoothness level using summary statistics. We show that retrievals from the two algorithms are consistent, with an increase in computational speed for the parallel MCMC algorithm. To assess the performance of our methods, we apply them to retrieve AOD values for the greater Beijing area in China. Our retrievals are tested against ground-based measurements of AOD from two AERONET stations in the area. Results show improvement on retrieval accuracy and coverage, especially during high-AOD events. We also include geographical conditions and levels of anthropogenic activities from Google Earth to qualitatively validate our results.

The rest of this chapter is organized as follows: Section 2.2 provides the rationale and details of our Bayesian model for retrieving AOD values and mixing vectors, while Section 2.3 details our MCMC algorithms. Section 2.4 contains case studies for model validation and interpretation, comparing our results with MISR's retrievals and AERONET measurements. Section 2.4.3 illustrates the necessity to include a richer variety of aerosol compositions. Section 3.5 summarizes the results and suggests directions for future research.

2.2 Hierarchical Bayesian Model

Our objective is to establish a more detailed data-driven description of the relationship among radiances, AOD, and aerosol compositions to assist aerosol-related health studies. The MISR operational retrieval algorithm provides this information by comparing the observed and the radiative transfer calculated radiances, but it is limited within the 74 pre-fixed aerosol compositions and a discrete grid of AOD values. We propose to allow a greater variety of aerosol optical behaviors by considering AOD values and mixing vectors as continuous variables, given a fixed set of four component aerosols. For the greater Beijing area, this set includes spherical non-absorbing aerosols without sulfate, spherical non-absorbing aerosols with sulfate, spherical absorbing aerosols, and grains (dust).

Each MISR Block contains 256 pixels (8 rows × 32 columns) at 17.6 km resolution in the MISR retrievals. The number of pixels in a MISR Block rises to 4,096 (32 rows ×

---

[3] MISR observes the Earth's surface in 233 swaths; each swath contains 180 560×140 km^2 MISR Blocks.

128 columns) at 4.4 km resolution, presenting a more complex problem with approximately 16,384 parameters to estimate. On the other hand, air particles interact in the atmosphere within a certain range; they affect aerosol conditions in near neighborhoods[1][36]. This suggests a stronger spatial dependence among adjacent pixels at a finer scale. When modeling at fine resolution, therefore, it is necessary and beneficial to borrow strength from AOD's spatial smoothness to reduce model complexity. In particular, we construct a hierarchical Bayesian model with a built-in spatial dependence using a Gaussian Markov Random Field prior for AOD.

2.2.1 Defining the Likelihood Function

Let $p = 1, \ldots, P$ index the $P = 4,096$ pixels on a two-dimensional lattice in a MISR Block at 4.4 km resolution, and $\boldsymbol{L} = (\boldsymbol{L}_1, \ldots, \boldsymbol{L}_P)$ denote the MISR-observed top-of-atmosphere radiances. For each pixel p, $\boldsymbol{L}_p = (L_{1p}, \ldots, L_{Cp}) \in \mathbb{R}^C$ corresponds to MISR's $C = 36$ channels. For every channel $c = 1, \ldots, C$, the MISR retrieval algorithm sets a measurement error of size σ_c as 5% of the smaller value between 0.04 and $\bar{L}_c = (\sum_{p=1}^{P} L_{cp})/P$. For pixel p, our goal is to estimate its AOD value $\tau_p \in \mathbb{R}$ and mixing vector $\boldsymbol{\theta}_p = (\theta_{p1}, \ldots, \theta_{pM}) \in \mathbb{R}^M$, relative to the M component aerosols involved ($\boldsymbol{\theta}_p \geq 0$ and $\sum_{m=1}^{M} \theta_{pm} = 1$). Each of MISR's 74 pre-fixed aerosol mixtures contain two or three component aerosols. We expand to allow mixtures of four component aerosols by setting $M = 4$; case studies confirm the sufficiency of this choice.

Given the geolocation of pixel p, its AOD value τ_p, a set of component aerosols and their mixing vector $\boldsymbol{\theta}_p$, Radiative Transfer (RT) equations are used to simulate radiances $\boldsymbol{L}^{RT} = (L_1^{RT}, \ldots, L_C^{RT})$ [6]; their pre-computed values at discrete points are stored in MISR's SMART Dataset[4]. Thus, $\boldsymbol{L}^{RT}$ can be viewed as functions of $(\tau_p, \boldsymbol{\theta}_p)$, relative to the M component aerosols involved. For each pixel p independently, the MISR operational retrieval algorithm uses a weighted least squares criterion to measure the closeness of an observed radiance vector to a particular RT simulated radiance vector. The weighted least squares take the following form [5]:

$$\chi_p^2 = \sum_{c=1}^{C} \frac{(L_{cp} - L_c^{RT}(\tau_p, \boldsymbol{\theta}_p))^2}{2\sigma_c^2}. \tag{2.1}$$

The MISR retrieval algorithm exhaustively searches over all combinations of pre-fixed AOD values and 74 aerosol compositions to match $\boldsymbol{L}^{RT}$ to the observed $\boldsymbol{L}$. The combinations of AOD and compositions satisfying a pre-established threshold of χ_p^2 in (2.1) are considered good fits to the observations; the average of all such AODs is the MISR retrieval at pixel p.

---

[4] The other parameters, such as the ambient pressure, take the default values unless otherwise specified. The MISR team has kindly given us access to the SMART dataset.

Inspired by MISR's weighted least squares criterion, we propose to use the weighted differences between observed $\boldsymbol{L}$ and radiative transfer simulated $\boldsymbol{L}^{RT}$ in (2.1) to form the following operational likelihood function:

$$p(\boldsymbol{L}|\boldsymbol{\tau},\boldsymbol{\theta}) \propto \exp\left\{-\sum_{c=1}^{C}\sum_{p=1}^{P}\frac{(L_{cp}-L_c^{RT}(\tau_p,\boldsymbol{\theta}_p))^2}{2\sigma_c^2}\right\}. \tag{2.2}$$

If we carry out a Maximum Likelihood estimation, the above Gaussian likelihood function coincides with MISR's weighted least squares criterion, assessing how relatively probable are the unobserved parameters $\boldsymbol{\tau} = (\tau_1, \ldots, \tau_P)$ and $\boldsymbol{\theta} = (\boldsymbol{\theta}_1, \ldots, \boldsymbol{\theta}_P)$, given the MISR observations $\boldsymbol{L}$. More importantly, this operational likelihood provides a formal device for us to construct a spatial smoothness structure for the AOD values $\boldsymbol{\tau}$ into the Bayesian hierarchy.

Even though the exact distribution of the weighted differences in (2.2) is difficult to determine due to the complex origins for these differences[5], histograms of retrieval residuals based on (2.2) display a single modal distribution; this supports our choice for a Gaussian-shaped operational likelihood. Another assumption in both (2.1) and (2.2) is that the differences between $\boldsymbol{L}^{RT}$ and $\boldsymbol{L}$ are independent of the channel c[6], if the correct values of ($\boldsymbol{\tau}$, $\boldsymbol{\theta}$) have been selected.

Now we are ready to describe our hierarchical model through building conditional relationships within the Bayesian hierarchy and assigning reasonable priors to the unobserved variables.

2.2.2 Construction of Priors and Conditional Probabilities

For fixed atmospheric pressures, humidity, wind levels, and a set of component aerosols involved, the top-of-atmosphere radiances $\boldsymbol{L}$ are mainly determined by AOD $\boldsymbol{\tau}$ and aerosol mixing vectors $\boldsymbol{\theta}$. Our Bayesian hierarchy's first level depicts this dependence of $\boldsymbol{L}$ on $\boldsymbol{\tau}$ and $\boldsymbol{\theta}$. Prior distribution for $\boldsymbol{\tau}$ is postulated to capture the spatial smoothness, calibrated by hyperparameter κ. We further assume independence between priors for $\boldsymbol{\tau}$ and $\boldsymbol{\theta}$ to simplify computation, i.e., $p(\boldsymbol{\tau},\boldsymbol{\theta}) = p(\boldsymbol{\tau})p(\boldsymbol{\theta})$. The inference of parameters and hyperparameters using MCMC sampling methods, is discussed in Section 2.3.

---

[5] Such origins include MISR camera measurement errors, radiative transfer calculation noises, differences between the proposed and true values for AOD and mixing vectors, choices of component aerosols, and errors in estimating surface-leaving radiances.

[6] We found close-to-0 correlations (-0.0445) between our retrievals' residuals at different viewing angles, but nontrivial correlations (0.5714) between residuals at different spectral bands. In current work, we are building this dependence structure among different bands in our model.

Prior Beliefs about AOD's Spatial Dependence

We characterize the spatial dependence of AOD values $\boldsymbol{\tau}$ using an intrinsic Gaussian Markov Random Field (GMRF) prior of first order[30]. Define κ as the homogenous scaler precision and use $\sim$ to indicate spatial adjacency. The following prior is invariant to perturbation by the same constant to $\boldsymbol{\tau}$ of all pixels[2],

$$p(\boldsymbol{\tau}|\kappa) \propto \kappa^{\frac{P-1}{2}} \exp\left\{-\frac{\kappa}{2}\sum_{p'\sim p}(\tau_{p'}-\tau_p)^2\right\}. \tag{2.3}$$

This allows us to model AOD's spatial smoothness by penalizing sharp changes of $\boldsymbol{\tau}$ among adjacent pixels, regardless of their unknown overall level. The prior in (2.3) is calibrated by κ as AOD's precision. The larger κ is, the smoother the region's AOD values are. For some regions, however, a more complicated GMRF prior might be necessary. For example, a constant wind pattern might require distinguishing an upwind pixel from a downwind pixel. This chapter works with a homogenous precision κ and thus has its limitations.

To estimate κ, we assign it a hyperprior. Due to AOD's large variability within a day and the lack of pre-existing records to specify a prior belief of $\boldsymbol{\tau}$'s behaviors, we consider a noninformative prior: $p(\kappa) \propto 1/\kappa$. The posterior is a proper Gamma distribution,

$$p(\kappa|\boldsymbol{\tau}) \propto \kappa^{\frac{P-1}{2}-1} \exp\left\{-\frac{\kappa}{2}\sum_{p':p'\sim p}(\tau_{p'}-\tau_p)^2\right\}. \tag{2.4}$$

For the above prior to work well, the number of groups, namely P, is to be larger than 5 [11]. In our case, P is commonly larger than 1000 at 4.4 km resolution. The simulation to be described in Section 2.3.2 shows good agreement between the true and retrieved values of κ using our MCMC algorithm. For example, we observed 100 (true) and 92.08 (retrieved) in one simulation, 500 (true) and 485.76 (retrieved) in another.

Prior Specification for Aerosol Compositions

Prior information on aerosol compositions is incorporated in the model through choices of the $M = 4$ component aerosols involved, based on geophysical knowledge of the study region. To model the mixing vectors $\boldsymbol{\theta}$ of the M component aerosols, we use an M-dimensional Dirichlet prior with Dirichlet parameter $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_M)$. Conditioning on $\boldsymbol{\alpha}$, the mixing vectors $\{\boldsymbol{\theta}_p\}_{p=1}^P$ are considered to be independent of each other,

$$p(\boldsymbol{\theta}|\boldsymbol{\alpha}) = \prod_{p=1}^{P} p(\boldsymbol{\theta}_p|\boldsymbol{\alpha}) = \prod_{p=1}^{P} \frac{\Gamma(\sum_{m=1}^{M}\alpha_m)}{\prod_{m=1}^{M}\Gamma(\alpha_m)}\theta_{p1}^{\alpha_1-1}\cdots\theta_{pM}^{\alpha_M-1}. \tag{2.5}$$

Even though the mixing vectors' spatial smoothness is not explicitly formulated, it is still captured and implicitly enforced by the spatial structure of AOD $\boldsymbol{\tau}$ through their dependence on the observed radiances $\boldsymbol{L}$. In fact, our estimates of mixing vectors $\boldsymbol{\theta}$ indeed display spatial smoothness. The model and algorithms remain relatively simple and computationally efficient.

We can further control the overall sparsity of the mixings of component aerosols by adjusting the magnitude of $\boldsymbol{\alpha}$. In general, we obtain no prior information on the mixing's sparsity; we assign $\boldsymbol{\alpha}$ a hyperprior to estimate it. Since (2.5) belongs to an exponential family, we adopt its conjugate: $p(\boldsymbol{\alpha}) \propto \exp(\sum_{m=1}^{M}(1-\alpha_m))$. This prior of $\boldsymbol{\alpha}$ gives larger probability to a smaller sum of α_m's, which suggests a sparse mixing of component aerosols, i.e. mixtures with one or two dominant components. This is supported by results from observational studies on aerosol mixings[6]. The posterior has the following form,

$$p(\boldsymbol{\alpha}|\boldsymbol{\theta}) \propto \exp\left\{\sum_{m=1}^{M}(\alpha_m - 1)(\sum_{p=1}^{P}\log\theta_{pm} + 1) - P \times (\sum_{m=1}^{M}\log\Gamma(\alpha_m) - \log\Gamma(\sum_{m=1}^{M}\alpha_m))\right\}.$$

Hyperprior for $\boldsymbol{\sigma}^2$

In our approach, we regard $\{\sigma_c^2\}_{c=1}^{C}$ as unknown and they are estimated together with $(\boldsymbol{\tau}, \boldsymbol{\theta})$. The likelihood function for $\boldsymbol{\sigma}^2 = (\sigma_1^2, \ldots, \sigma_C^2)$, $p(\boldsymbol{L}|\boldsymbol{\sigma}^2, \boldsymbol{\tau}, \boldsymbol{\theta})$, follows a normal distribution with known mean and unknown variance. We adopt a noninformative scaled inverse-χ^2 hyperprior for $\boldsymbol{\sigma}^2$ to model the channel weights $\{\frac{1}{2\sigma_c^2}\}_{c=1}^{C}$: $p(\sigma_c^2) \propto \sigma_c^{-2}$. This hyperprior suggests that values for the unknown weights become less likely in inverse proportion to their values; it is also a choice of computational convenience. The conditional posterior also follows the scaled inverse-χ^2 distribution,

$$p(\sigma_c^2|\boldsymbol{\tau}, \boldsymbol{\theta}, \boldsymbol{L}) \propto (\sigma_c^2)^{-(\frac{P}{2}+1)} \exp\left\{-\frac{\sum_{p=1}^{P}(L_{cp} - L_c^{RT}(\tau_p, \boldsymbol{\theta_p}))^2}{2\sigma_c^2}\right\}.$$

2.3 MCMC Retrieval Algorithms

Based on the hierarchical Bayesian model previously developed, this section first derives marginal posterior distributions of AOD values $\boldsymbol{\tau}$ and mixing vectors $\boldsymbol{\theta}$. We then devise two MCMC algorithms to sample from the posteriors. Using MISR observed radiances as input, we take the sampled posterior means as outputs.

2.3.1 Posterior Distributions of AOD Values and Mixing Vectors

The full Bayesian model discussed above can be summarized as follows:

$$\begin{aligned}
\boldsymbol{L}_p|\tau_p, \boldsymbol{\theta_p} &\sim \mathcal{N}(\boldsymbol{L}^{RT}(\tau_p, \boldsymbol{\theta}_p), \boldsymbol{\sigma^2}),\ p = 1, \ldots, P,\\
\boldsymbol{\tau}|\kappa &\sim GMRF(\kappa),\\
\boldsymbol{\theta}|\boldsymbol{\alpha} &\sim Dirichlet(\boldsymbol{\alpha}),\\
\boldsymbol{\sigma}^2 &\sim scaled\ inverse - \chi^2(\nu_0),\\
\kappa &\sim Gamma(\alpha_0, \beta_0),\\
p(\boldsymbol{\alpha}) &\sim Exp(\sum_{m=1}^{M}(1-\alpha_m)).
\end{aligned}$$

With no additional information on the hyperparameters, ν_0, α_0, and β_0 are chosen to be 0 for convenience and later shown to be robust. The marginal posterior of AOD values $\boldsymbol{\tau}$ is,

$$p(\boldsymbol{\tau}|\boldsymbol{\theta}, \kappa, \boldsymbol{\sigma}^2, \boldsymbol{L}) \propto \exp\left\{-\frac{1}{2}\kappa \sum_{p':p'\sim p}(\tau_{p'} - \tau_p)^2 - \sum_{c=1}^{C}\sum_{p=1}^{P}\frac{(L_{cp} - L_c^{RT}(\tau_p, \boldsymbol{\theta}_p))^2}{2\sigma_c^2}\right\}. \quad (2.6)$$

The marginal posterior distribution of the mixing vectors $\boldsymbol{\theta}$ can be expressed as,

$$p(\boldsymbol{\theta}|\boldsymbol{\tau}, \boldsymbol{\alpha}, \boldsymbol{\sigma}^2, \boldsymbol{L}) \propto \exp\left\{\sum_{p=1}^{P}\sum_{m=1}^{M}(\alpha_m - 1)\log\theta_{pm} - \sum_{c=1}^{C}\sum_{p=1}^{P}\frac{(L_{cp} - L_c^{RT}(\tau_p, \boldsymbol{\theta}_p))^2}{2\sigma_c^2}\right\}. \quad (2.7)$$

Both posteriors contain radiative transfer simulated $\boldsymbol{L}^{RT}$, which can be obtained at necessary values through interpolations from the MISR SMART Dataset, using τ_p and $\boldsymbol{\theta}_p$ as inputs [6]. The resulted non-closed-form posteriors, however, are difficult to directly sample from. A Metropolis-within-Gibbs sampler is thus used.

2.3.2 Metropolis-within-Gibbs Sampling from the Posterior Distributions

The Gibbs sampler[13] is a numerical technique to sample from a joint distribution, $p(\boldsymbol{\tau}, \boldsymbol{\theta}, \boldsymbol{\sigma}^2, \kappa, \boldsymbol{\alpha}|\boldsymbol{L})$ in our case. We sample for τ_p and $\boldsymbol{\theta}_p$ using a Metropolis-Hastings (M-H) sampler, for each pixel p on the MISR Block column by column and pixel by pixel. The following proposal distribution is used in M-H sampler for τ_p,

$$p(\tau_p|\boldsymbol{\tau}_{-p}) \propto \exp\left(-\frac{n_p\kappa}{2}(\tau_p - \frac{1}{n_p}\sum_{p':p'\sim p}\tau_{p'})^2\right),$$

where n_p is the number of adjacent pixels to pixel p, and κ the scalar precision of the Markov Random Field. A Dirichlet proposal distribution with parameter $\boldsymbol{\alpha}$ is used in M-H for $\boldsymbol{\theta}_p$. Denote vector $(\tau_p, \ldots, \tau_{p'})$ by $\boldsymbol{\tau}_{p:p'}$ and similarly for $\boldsymbol{\theta}$ and their Dirichlet parameter $\boldsymbol{\alpha}$. Given initializations $(\boldsymbol{\tau}^{(0)}, \boldsymbol{\theta}^{(0)}, (\boldsymbol{\sigma}^2)^{(0)}, \kappa^{(0)}, \boldsymbol{\alpha}^{(0)})$, the sampler proceeds as described in the following Metropolis-within-Gibbs Algorithm.

Algorithm 1 *

Metropolis-within-Gibbs Algorithm (M-w-G) At step t, iterate the following process:

```
for $p = 1$ to $P$ do
    Use M-H to sample $\tau_p^{(t)} \sim p(\tau_p | \boldsymbol{\tau}_{1:(p-1)}^{(t)}, \boldsymbol{\tau}_{(p+1):P}^{(t-1)}, \boldsymbol{\theta}^{(t-1)}, (\boldsymbol{\sigma}^2)^{(t-1)}, \kappa^{(t-1)}, \boldsymbol{L})$.
for $p = 1$ to $P$ do
    Use M-H to sample $\boldsymbol{\theta}_p^{(t)} \sim p(\boldsymbol{\theta}_p | \boldsymbol{\tau}^{(t)}, \boldsymbol{\theta}_{1:(p-1)}^{(t)}, \boldsymbol{\theta}_{(p+1):P}^{(t-1)}, (\boldsymbol{\sigma}^2)^{(t-1)}, \boldsymbol{\alpha}^{(t-1)}, \boldsymbol{L})$.
for $c = 1$ to $C$ do
    Use M-H to sample $(\sigma_c^2)^{(t)} \sim p(\sigma_c^2 | \boldsymbol{\tau}^{(t)}, \boldsymbol{\theta}^{(t)}, (\boldsymbol{\sigma}_{1:(c-1)}^2)^{(t)}, (\boldsymbol{\sigma}_{(c+1):C}^2)^{(t-1)}, \boldsymbol{L})$.
Sample $\kappa^{(t)} \sim p(\kappa | \boldsymbol{\tau}^{(t)})$.
for $m = 1$ to $M$ do
    Use M-H to sample $\alpha_m^{(t)} \sim p(\alpha_m | \boldsymbol{\theta}^{(t)}, \boldsymbol{\alpha}_{1:(m-1)}^{(t)}, \boldsymbol{\alpha}_{(m+1):M}^{(t-1)})$.
```

Each cycle of the algorithm generates a realization of a Markov chain, which gives approximate samples from the marginal posteriors after a successful burn-in process[13]. We check that the acceptance rate of the Metropolis-Hastings sampler is roughly between 25% and 50% for adequate mixing of posterior samples [10]. The potential scale reduction $\hat{R}$ [12] is also used to check convergence of the Markov chains. We run the chains until $\hat{R}$ is less than 1.1 or 1.2, using $\hat{R}$ of the logarithm of the posterior distribution as a benchmark. A geometric decay of the autocorrelation as a function of the lag also suggests well mixing of our chains.

We also conduct a simulation study to verify the M-w-G's ability to converge to the target distribution:

The trace plots of the MCMC samples of AOD values (Figure 2.10[8]) show good convergence after approximately 400 iterations, whether the initialization is close to the true value or not. We observe similar convergence rates for mixing vectors. Assigning different values to the hyperparameters, the correlation between the true AOD and the MCMC-retrieved AOD ranges between 0.78 and 0.90, and the coefficient of variation of the rooted-mean-square error ranges between 4.24% and 9.26%.

Finally, we use the sampled posterior mean to estimate AOD values and mixing vectors. While the MCMC algorithm enables us to handle a hierarchy whose complexity precludes

---

[7] The noise's standard deviation $\boldsymbol{\sigma}$ is set as 10% of the averaged radiance, while the MISR operational algorithm estimates $\boldsymbol{\sigma}$ as 5% of the same average.

[8] We attach in appendix two trace plots showing one example of each type, up to the first 1000 iterations.

Algorithm 2 *

Algorithm Example Simulation to Verify Convergence of M-w-G

```
Select the same four component aerosols as in the Beijing case studies (Section 2.4).
$\kappa \leftarrow 100$ (or 500 for different runs).
$\boldsymbol{\alpha} \leftarrow (0.8, 0.4, 0.2, 0.2)$ (or (2, 4, 0.1, 0.1) for different runs).
Sample $\boldsymbol{\tau}^{(0)} \sim$ (2.3).
Sample $\boldsymbol{\theta}^{(0)} \sim$ (2.5).
Considering ($\boldsymbol{\tau}^{(0)}$, $\boldsymbol{\theta}^{(0)}$) as the true values, simulate radiances $\boldsymbol{L}^{sim}$ using the SMART lookup table and an additive Gaussian noise[7].
Input $\boldsymbol{L}^{sim}$ into M-w-G retrieval algorithm to estimate AOD and mixing vectors.
```

fitting by analytical methods, its computational intensity limits its operational use. Next, we propose a parallel MCMC algorithm to reduce computational cost.

2.3.3 A Parallel MCMC Algorithm

Many MCMC sampling algorithms for spatial data suffer from high computational cost caused by the large dimensionality of data. At 4.4 km resolution, our MCMC algorithm simulates samples for more than 16,000 variables[9] for one MISR Block. The large computational cost is exacerbated by the non-closed form of the posterior distributions. It is possible, however, to develop a faster algorithm to sample from a distribution which approximates the target posterior of the original MCMC algorithm.

By this token, we devise a parallel MCMC algorithm to improve the computational efficiency: each MISR block is divided into 2 × 8 patches of equal size with at least four overlapping columns and rows for adjacent patches; the M-w-G sampler is applied to each patch independently to generate samples for ($\boldsymbol{\tau}$, $\boldsymbol{\theta}$). This independent sampling on different patches can therefore benefit from parallel computing.

Information on AOD's spatial dependence structure is to be communicated across the entire MISR Block to estimate AOD's spatial smoothness level. On that account, we let the patches periodically exchange spatial smoothness information across the entire MISR Block. Given κ's conditional posterior,

$$p(\kappa|\boldsymbol{\tau}) \propto \kappa^{(P-1)/2-1} \exp\{-\frac{1}{2}\kappa \sum_{p\sim p'} (\tau_p - \tau_{p'})^2\},$$

and summary statistic, $T_\kappa = \sum_{p\sim p'} (\tau_p - \tau_{p'})^2$, it follows that $p(\kappa|\boldsymbol{\tau}) = p(\kappa|T_\kappa)$. Hence T_κ summarizes the information on calibration κ for AOD's spatial smoothness across the entire MISR block. Given that the hyperparameters control the spatial smoothness level of model

---

[9] Excluding cloudy pixels can sometimes reduce the total dimensions to around 5,000 for one MISR Block.

parameters in all patches, the parallel MCMC algorithm provides an approximation to the posterior of AOD values $\boldsymbol{\tau}$, while the patch-samplings in parallel improve the computational efficiency. We now describe the parallel MCMC algorithm in detail:

Algorithm 3 *

Parallel MCMC Algorithm Obtain a MISR Block of 32×128 pixels at 4.4 km resolution and divide the Block into 2×8 patches, each of 20×20 pixels, with at least 4 overlapping columns/rows between adjacent patches. At step t, iterate the following process:

```
Use M-w-G algorithm to sample $\boldsymbol{\tau} \sim p(\boldsymbol{\tau}|\boldsymbol{\theta}, \kappa, \boldsymbol{\sigma}^2, \boldsymbol{L})$, $\boldsymbol{\theta} \sim p(\boldsymbol{\theta}|\boldsymbol{\tau}, \boldsymbol{\alpha}, \boldsymbol{\sigma}^2, \boldsymbol{L})$, $\boldsymbol{\sigma}^2 \sim p(\boldsymbol{\sigma^2}|T_{\boldsymbol{\sigma}}^{(t)})$, $\kappa \sim p(\kappa|T_\kappa^{(t)})$, $\boldsymbol{\alpha} \sim p(\boldsymbol{\alpha}|T_{\boldsymbol{\alpha}}^{(t)})$ within each patch in parallel for 50 iterations.
Average the samples of the overlapping pixels between any two adjacent patches.
Calculate summary statistics using current samples,
    $T_{\sigma_c}^{(t+1)} = \sum_{p=1}^{P}(L_{cp} - L_c^{RT}(\tau_p, \boldsymbol{\theta}_p))^2, c = 1, \ldots, C,$
    $T_\kappa^{(t+1)} = \sum_{p \sim p'}(\tau_p - \tau_{p'})^2,$
    $T_{\alpha_m}^{(t+1)} = \sum_{p=1}^{P} \log \theta_{pm}, m = 1, \ldots, M.$
```

The above process can be automated using the Perl programming language. For a MISR block at 4.4 km resolution, the computational time of the parallel MCMC algorithm is less than one-fifth of that of the global MCMC sampling algorithm, accounting for overhead time of communication among different patches.

This parallel MCMC sampling scheme can be generalized to improve the computational efficiency of MCMC sampling based on spatial data of a large scale. By conditioning on a summary statistic which preserves the global spatial dependence level, we can partition the original sampling problem into many sub-samplings and distribute them to different processing units concurrently. Samples generated from each processing unit can be periodically collected to renew the summary statistic, which is then returned to each processing unit to update the sub-samplings. Though this scheme samples from an approximation to the target distribution, it can largely speed up the computation.

The global and the parallel MCMC algorithms produce reasonably consistent results. The outputs generally agree, except for a small group of pixels that mostly lie on the patch edges. The spatial smoothness is interrupted between patches; the benefits of a stabilizing factor from neighboring pixels are lost. This confirms that maintaining an appropriate spatial structure is important, and that our parallel MCMC algorithm's outputs are only an approximation to the target distribution. Increasing the number of iterations and communications of summary statistics, and smoothing the patch edges, reduce the disagreement. The next section evaluates the performance of our retrievals using case studies. For non-operational model validations, we apply the global MCMC algorithm to avoid inconsistency in number of iterations for different retrievals.

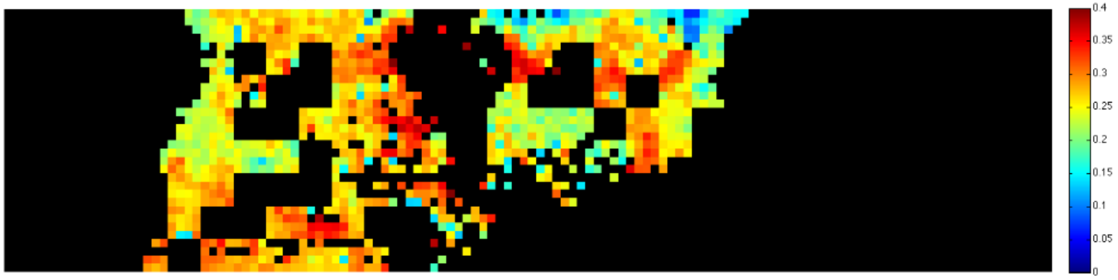
2.4 Validation and Results: Case Studies on Aerosol Retrievals for the Greater Beijing Area, China

In this section, we compare our retrievals to MISR outputs for the greater Beijing area (latitude: 38.95N∼40.15N; longitude: 115.57E∼119.50E) and discuss their differences. We validate our results using AERONET measurements and Google Earth satellite images. Through case studies, we demonstrate the importance of fine-resolution retrievals and a greater variety of compositions to improve retrieval accuracy and coverage.

2.4.1 Comparison with MISR Retrievals

Figure 2.2 displays the MISR AOD retrievals at 17.6 km resolution in panel (a) and our Bayesian AOD retrievals at 4.4 km resolution in panel (b). Shared information in MISR and our retrievals is observed, such as the coastline on the right, the overall AOD level, and its spatial patterns. This consistency is confirmed by the scatterplots of MISR outputs and Bayesian AOD retrievals aggregated to 17.6 km resolution (Figure 2.3, left panel). The black pixels in Figure 2.2 represent missing retrievals, mostly due to two common reasons. Firstly, aerosol retrievals are not attempted when clouds are detected. MISR averages the 1.1 km observations into a pixel at 17.6 km resolution and ignores clouds, when the cloudless areas are more than $\frac{1}{16}$ of the pixel. Clouds negligible at 17.6 km resolution, however, might

(a) MISR AOD retrievals at 17.6 km resolution.



(b) Bayesian AOD retrievals using MCMC at 4.4 km resolution.

Figure 2.2: AOD estimates from MISR and our Bayesian retrievals.

be significant at 4.4 km resolution; we tend to have more missing retrievals in some areas, intrinsically determined by the observations. Secondly, when none of the 74 MISR-designated compositions satisfy MISR's weighted least squares criterion, MISR operational algorithm marks the retrieval as missing. Our Bayesian retrievals, allowing for a richer variety of compositions, eliminate such unnecessarily missing retrievals (Section 2.4.3).

On the other hand, Figure 2.2 also demonstrates increased diversity in our Bayesian-retrieved AOD across the MISR Block, as the retrieval resolution improves. This is expected, since a finer resolution leads to more information observed and piped into the model. The reliability of such diversity needs to be further validated by other independent sources such as ground-based measurements, as discussed in the next section.

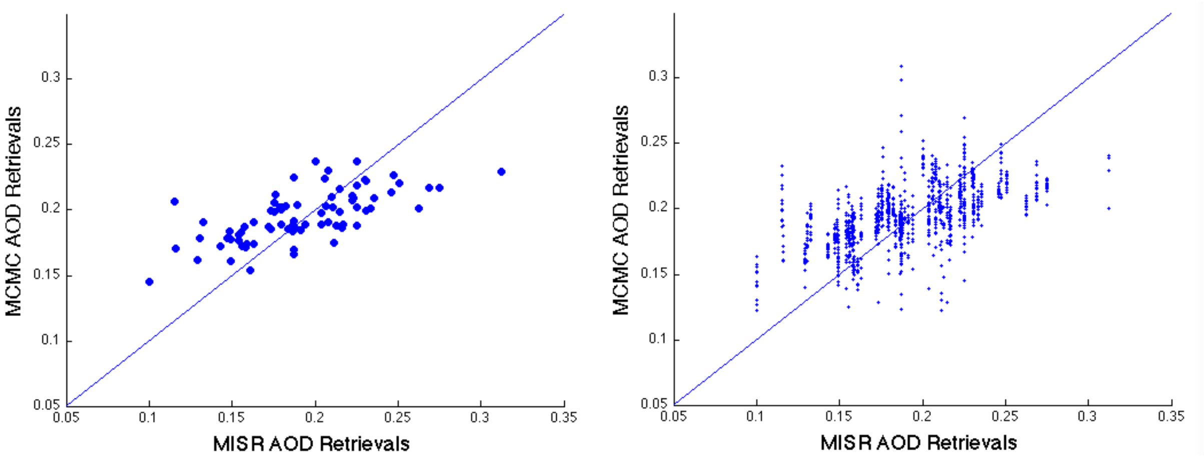


Figure 2.3: Scatterplots of MISR against MCMC retrievals at an aggregated 17.6 km resolution (left, r.m.s. = 0.0295) and a 4.4 km resolution (right, r.m.s. = 0.0309).

2.4.2 Model Validation for Bayesian Retrievals by Ground-based Measurements and Google Earth

Ground-based measurements are collected at AERONET Beijing and AERONET Xianghe stations, as well as via a hand-held MICROTOPS II Sunphotometer at several locations in urban Beijing area. The fixed locations of the AERONET stations and the limited travel range of the Sunphotometer's human operator make it impossible to validate retrievals of all pixels on the MISR Block under study. Instead, we focus on the pixels that contain the AERONET stations or our Sunphotometer-visited locations. To match the AOD values at the same wavelength, we first convert AERONET measurements to those at 550 nm using AERONET estimates of Angström exponent. We then average the measurements within a one-hour window when Terra carrying MISR passes over the AERONET stations, for Jiang, *et al.* [16] show that a narrower time window better captures the correlation between

AERONET measurements and MISR retrievals. The area's frequent cloudy weather and its latitude[10] contribute to the scarcity of the remote-sensed versus ground-based data pairs for validation.

As a result of this scarcity of ground-based validation, we also carry out qualitative validation using satellite images from Google Earth and discuss the findings in Section 2.4.2.

Retrieval Validation at AERONET Beijing Station

Figure 2.4 shows a boxplot of our Bayesian AOD retrievals for the pixel which contains the AERONET Beijing Station[11], with estimated uncertainties indicated by the box edges for inter-quartile ranges of posteriors and the whiskers drawn to the 5th and 95th percentiles. The three retrievals on March 15, April 30, and May 16, 2009 are plotted separately in the right panel to keep an appropriate scale for the left panel.

As long as a pixel is cloudless, our MCMC algorithms provide an AOD retrieval. However, the MISR operational retrieval algorithm shows missing values for 24% of the 21 cases in Figure 2.4. This results from the increasingly heterogeneous aerosol conditions in Beijing and the limited choices of aerosol compositions in MISR retrievals. In the coarse-resolution retrievals, high AOD values are averaged down by its neighbors and low AOD values averaged up, resulting in a loss of useful information. Our Bayesian retrievals show improvement in accuracy; detailed information on aerosols are revealed by the fine-resolution retrievals. The three high AOD values in the right panel of Figure 2.4 indicate Beijing's extreme air

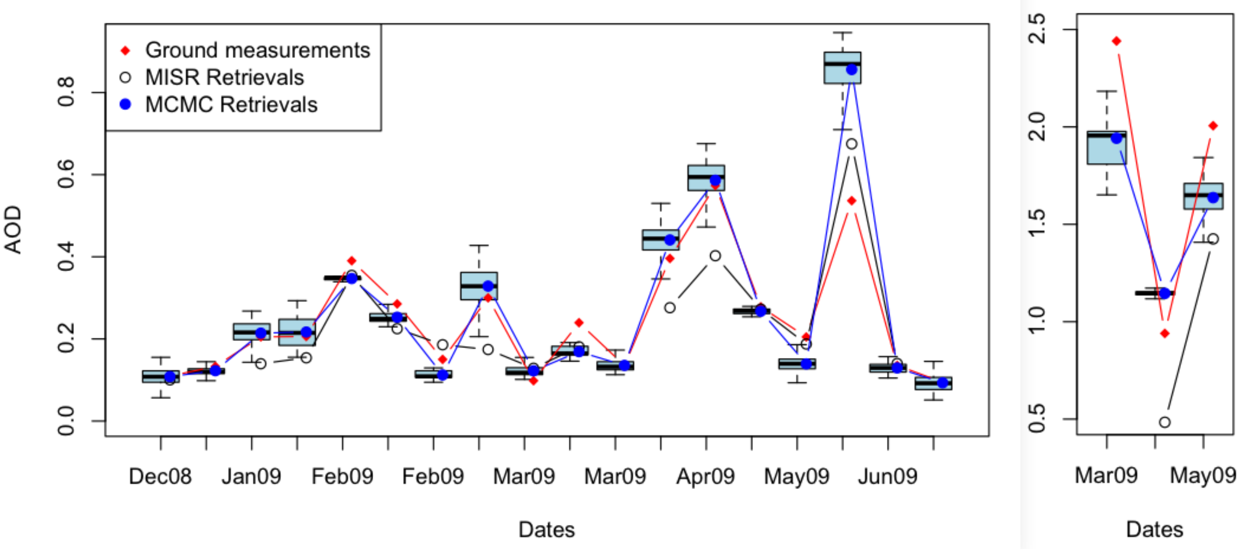


Figure 2.4: Validation of our AOD retrievals by measurements at AERONET Beijing Station.

---

[10] The Beijing city is visited by the Terra satellite every five to nine days.

[11] Latitude: 39.97689° North; longitude: 116.38137° East.

conditions, corresponding to 86%, 71%, and 81% reduced radiation by aerosols. For example, records of news from Xinhua Headlines show that on March 15, 2009, the city was trapped in a sandstorm originated in Inner Mongolia.

We would like to discuss one particular case when the Bayesian retrieval (0.8560) is much worse than MISR output (0.6750), compared to the AERONET measurement (0.5455): the third to last case in Figure 2.4 (left panel), May 25, 2009. AERONET reports no measurement when Terra carrying MISR passed above Beijing. Instead, we use the measurement of 0.5455, which is the closest in time but three hours earlier. This record reached the lowest of that day, with others between 0.6521 and 1.5366. It suggests that the particular AERONET record we used might not be ideal to validate the remote-sensed retrieval, but our best option.

Retrieval Validation at AERONET Xianghe Station

Figure 2.5 compares the remote-sensed retrievals to AERONET measurements at the pixel that contains the AERONET Xianghe station[12]. From December to February, AERONET measurements are mostly higher than remote-sensed retrievals, but no distinctive pattern afterwards.

AERONET Xianghe station has the Jingshen Expressway to its north, which is a major path connecting two hub cities: Beijing and Shenyang[13]. The northwest wind in winter carries car exhaust to the AERONET Xianghe station, possibly leading to high AOD measurements. Yet for remote-sensing retrievals, the green fields in a larger neighborhood balance this factor, which possibly results in a washed-out signal. However, the fine-resolution retrievals seem to suffer less from the balancing factors and display a better accuracy.

We would like to discuss one of the cases where our AOD retrieval is much higher than AERONET measurement: the first data point in Figure 2.5, December 25, 2008. MISR produced no output for this day. To the east of Xianghe station in Hebei Province lie several major malls for furniture exhibition and manufacture. On December 25, 2008, the furniture companies started renovating their exhibition halls. The construction could have caused localized aerosol loadings not observed by the Xianghe AERONET site 2 km away upwind within one day, but detected by the MISR instrument and captured by our retrievals.

Qualitative Validation using Google Earth in Absence of Ground Measurements

We observe other disagreements in our Bayesian-retrieved AOD values and those of MISR, in addition to those at the two pixels that contain the two AERONET stations in the MISR Block. Since they are retrievals at different spatial scales, they could as well be different, that is, they could both be valid. An indirect way to validate our retrieved AOD values is to see whether they reasonably reflect the region's geographical and anthropogenic conditions,

---

[12] Latitude: 39.75360° North; longitude: 116.96150° East.

[13] The capital and largest city of Liaoning Province in Northeast China.

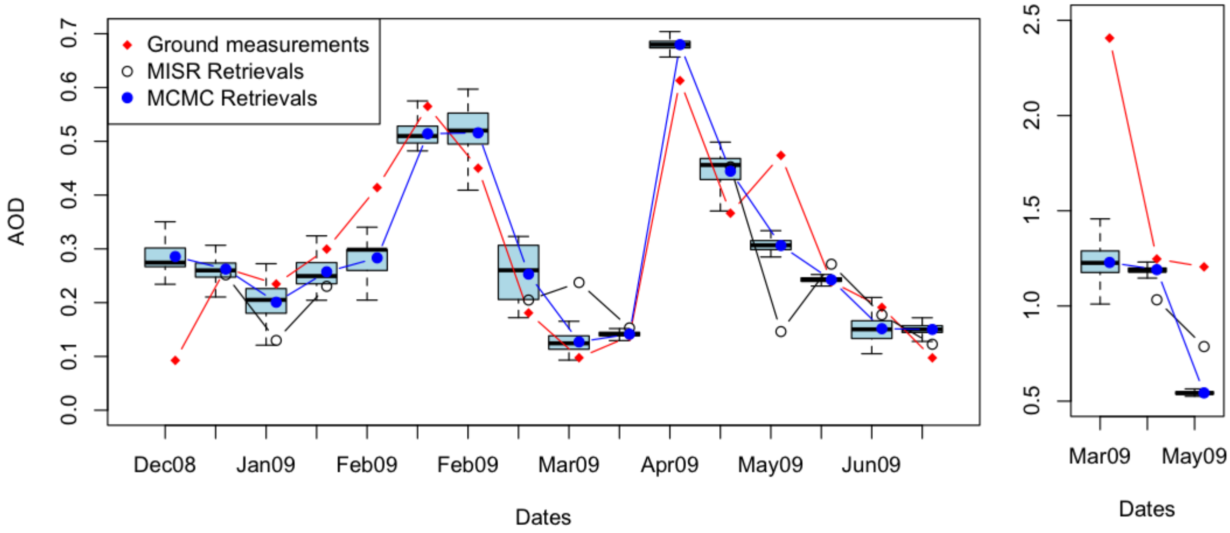


Figure 2.5: Validation of our AOD retrieval results by AERONET measurements in Xianghe.

such as existence of heavy industries and transportation patterns. These conditions can be easily assessed using the satellite images from Google Earth, making them indirect validation for our retrievals as a reasonable and detailed profiling of AOD spatial distribution. Here we focus on pixels with Bayesian AOD retrievals largely disagreeing with those of their adjacent pixels or pixels with locally highly variable Bayesian AOD retrievals. Since our retrieval pixels are only $\frac{1}{16}$ of the size of a MISR retrieval pixel, these AOD locations cannot be identified in the corresponding MISR retrievals.

In particular, we project our Bayesian AOD values onto Google Earth (Figure 2.6) and examine the pixels with locally highly variable AOD values. We thus identify a hub of the Jingshen and Jingtang Highways (pin A in Figure 2.6) and construction sites producing pollution (pin C), supporting the high AOD values indicated by only our Bayesian retrievals. The Olympic Park (pin D) and Beidaihe (pin F), a famous beach resort, also confirm the reasonability of the low AOD values captured by only the Bayesian retrievals at a finer resolution.

2.4.3 Case Study for Including a Richer Variety of Aerosol Compositions

This section emphasizes the necessity to expand MISR's 74 aerosol compositions. This expansion improves retrieval coverage and detects more features of aerosol behaviors, such as seasonality of component aerosols.

For an example of the improvement on retrieval coverage, we examine March 15, 2009. MISR failed to retrieve AOD for the majority of the Block (Figure 2.7, upper panel). Our

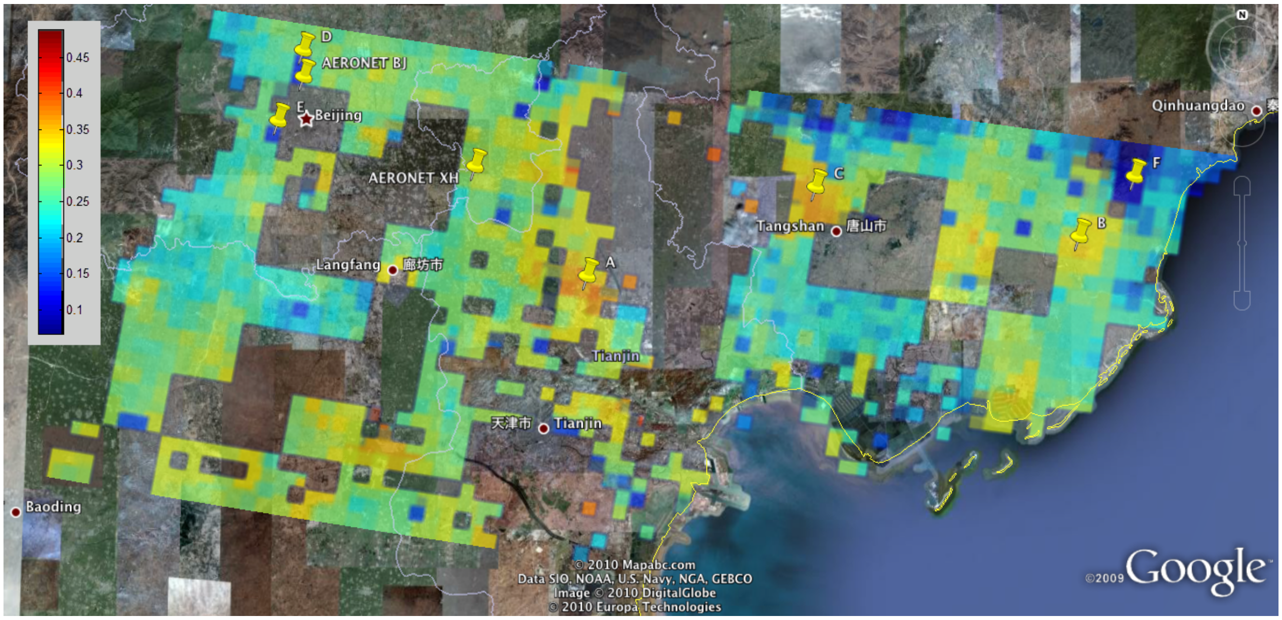


Figure 2.6: Retrieval results projected on Google Earth.

Bayesian retrievals provide better coverage and the retrievals give distinctly high AOD values with a clear path of aerosols migrating from west to east and into the ocean. This unusual discrepancy leads us to run through the weather records: on that day, the area suffered from a sandstorm originated in Inner Mongolia, which later passed into eastern China.

Figure 2.8 shows the corresponding mixing percentages of the four component aerosols involved. The fourth component aerosol with mixing percentage θ_4 represents the dust-related aerosols, listed as No. 19 in MISR documentation. In MISR retrieved mixing vectors (left panel, Figure 2.8), the fourth component showed barely any existence. In the MAP retrieved mixing vectors, the dust component dominates, which is consistent with the reported sandstorm.

Moreover, information on retrieved mixing vectors can be used to study aerosols' seasonal behaviors, and to further identify pollution types and sources. For example, results show that component No.6 which contains sulfate tends to dominate the mixture in winter due to burning coal for heating, while component No.19, grains (dust), dominates in spring due to sandstorms.

In general, in terms of retrieval coverage and accuracy, it is necessary and beneficial to expand the MISR-designated 74 mixtures. By correctly identifying the major pollutants in each season, we can better understand the movement and dispersion of aerosols. Measures can then be taken to improve air quality in a more specific and effective manner. More of such information in the retrieved aerosol mixing vectors is discussed in Chapter 4.

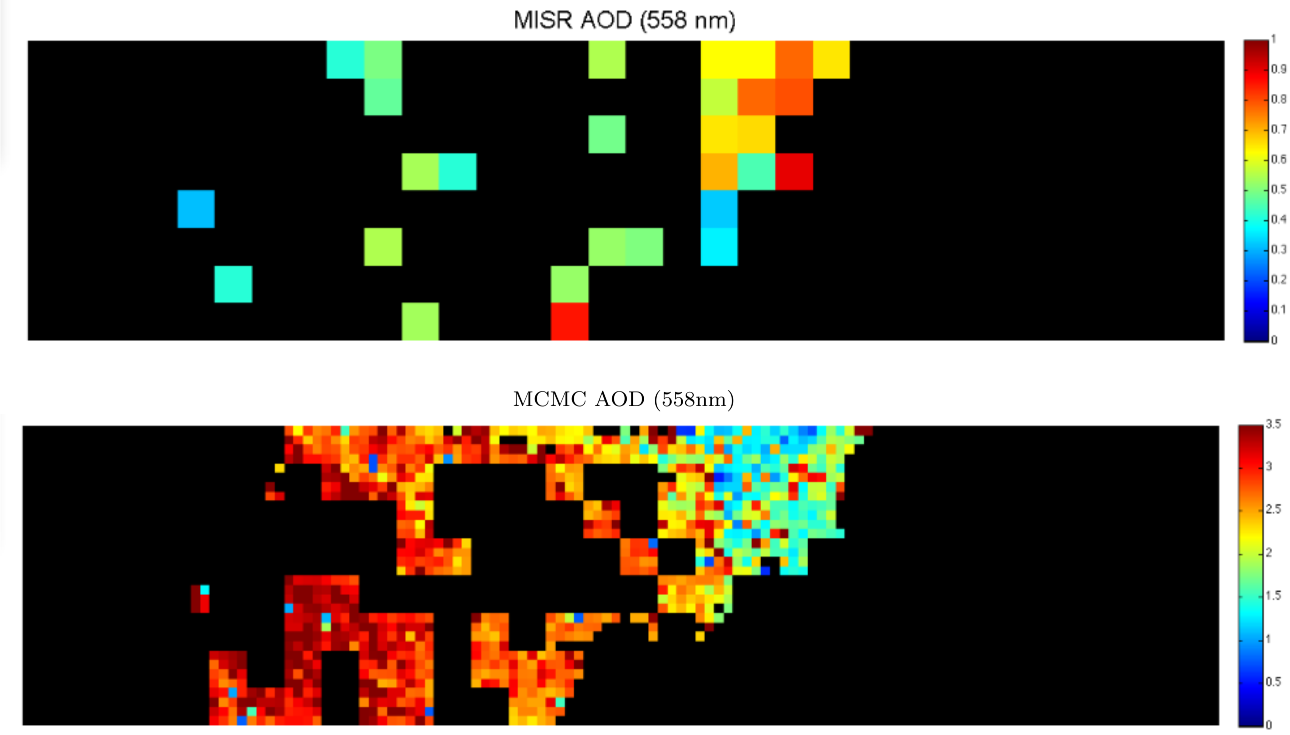


Figure 2.7: Case study of AOD retrievals on March 15, 2009.

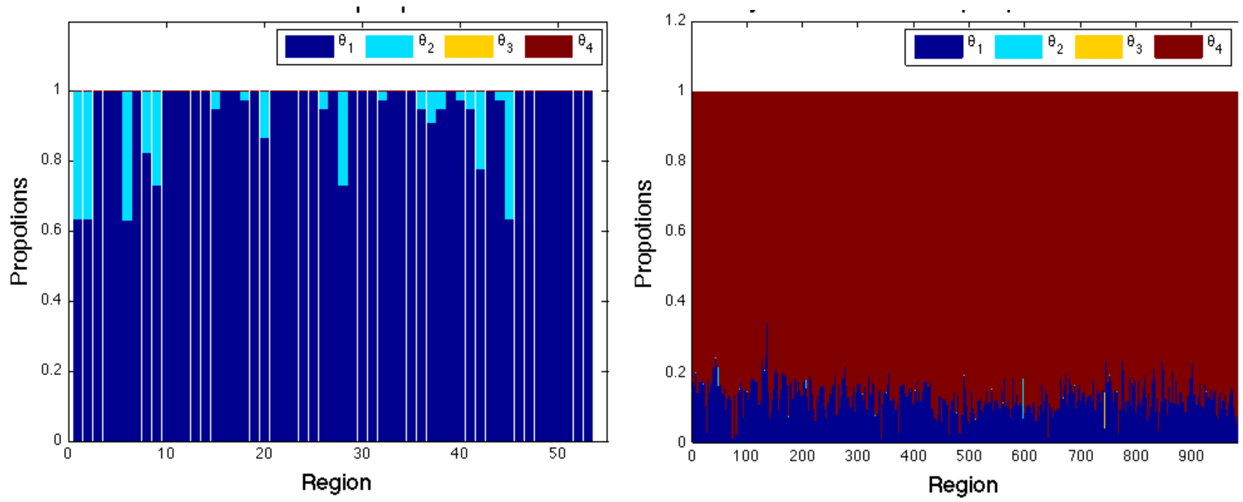


Figure 2.8: Mixing vectors of component aerosols for the greater Beijing area on March 15, 2009.

In general, for areas like Beijing, which experience occasional sandstorms, the limited compositions containing grains(dust) among MISR's 74 choices could easily result in a low coverage of MISR retrievals. Similar situations might exist for other locations with usual aerosol conditions. The retrieved mixing vectors also contains information on the regional aerosol composition and can be used to identify pollution type and source. For example, results show that component No.6 with sulfate tends to dominate the composition in winter due to coal burning for heating, with No.19, grains (dust) dominating in spring due to sandstorms.

Figure 2.9 shows the mixing percentages of component No.19 over December 2008 to June 2009, at four different locations: the AERONET Beijing station, the AERONET Xianghe station, location (A) and (F) marked in Figure 2.6. For AERONET Beijing station, the percentage of grains(dust) only rose in the spring, due to the sandstorms, while the constructions around AERONET Xianghe might have raised the percentage earlier in the year. Location (A), where major highways intersect, showed a high amount of dust in its aerosol compositions through the warm seasons when traffic typically increases. The mixing percentage of No.19 at Location (F), the Beidaihe Resort, moved relatively in consistence with AERONET Beijing station. We hope to explore this trend in future research. In general, by correctly identifying the major pollutants for each season, we can better understand the transitions of aerosols and, therefore, take efforts to improve air quality in a more specific and to the point manner. For accuracy and coverage, it is necessary to expand the MISR-designated 74 aerosol compositions to a richer variety.

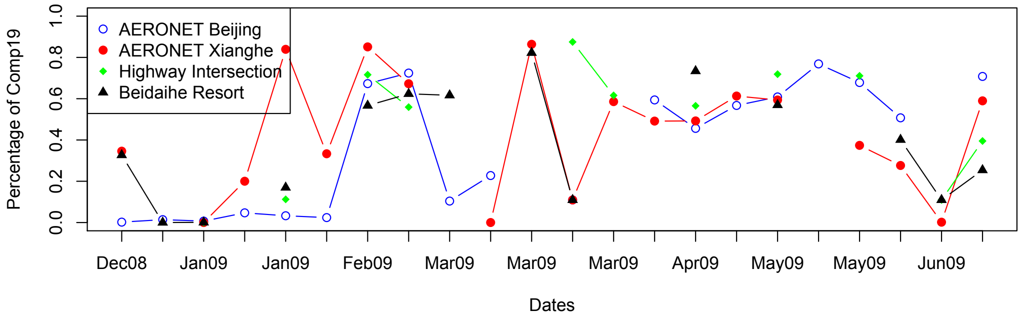


Figure 2.9: Mixing percentages of component No.19 from winter to spring.

2.5 Discussion

Aerosols serve as an important factor in air quality and public health. A profile of AOD's spatial distribution can eventually expand the potential of remote-sensed observations in facilitating urban air quality monitoring and public health studies [34][33]. The heterogeneity of urban aerosols due to anthropogenic activities calls for a profile of aerosols at a fine resolution and a larger variety of aerosol compositions.

In this chapter, we have presented a hierarchical Bayesian model to retrieve AOD values and mixing vectors relative to a collection of four component aerosols at an improved resolution of 4.4 km using MISR observations. The model incorporates a spatial dependence structure to gain strength from AOD's spatial smoothness; it also allows for a richer variety of aerosol mixing vectors to better capture the growing heterogeneity of urban aerosols and the increasingly severe weather conditions, such as dust storms. A more detailed AOD spatial profile is provided and further validated by AERONET and Google Earth; an improved accuracy and a better retrieval coverage is obtained due to the improved resolution and flexible choices of aerosol compositions. This improvement is particularly important during high-AOD events, which often indicate severe air pollution. We further develop a parallel MCMC algorithm to improve the computational efficiency, which can be generalized to speed up other MCMC sampling algorithms based on spatial data.

From the case studies, we become more aware of the complexity in aerosol conditions and thus hope to use our results to study the aerosols' impact on public health in urban areas at the enhanced resolution. We also hope to explore the possibility of improving the retrieval accuracy by incorporating more prior knowledge in the model, such as wind measurements and dependence among the four spectral bands.

2.6 Appendix: Example trace plots of MCMC samples for AOD in simulation study (Section 3.2)

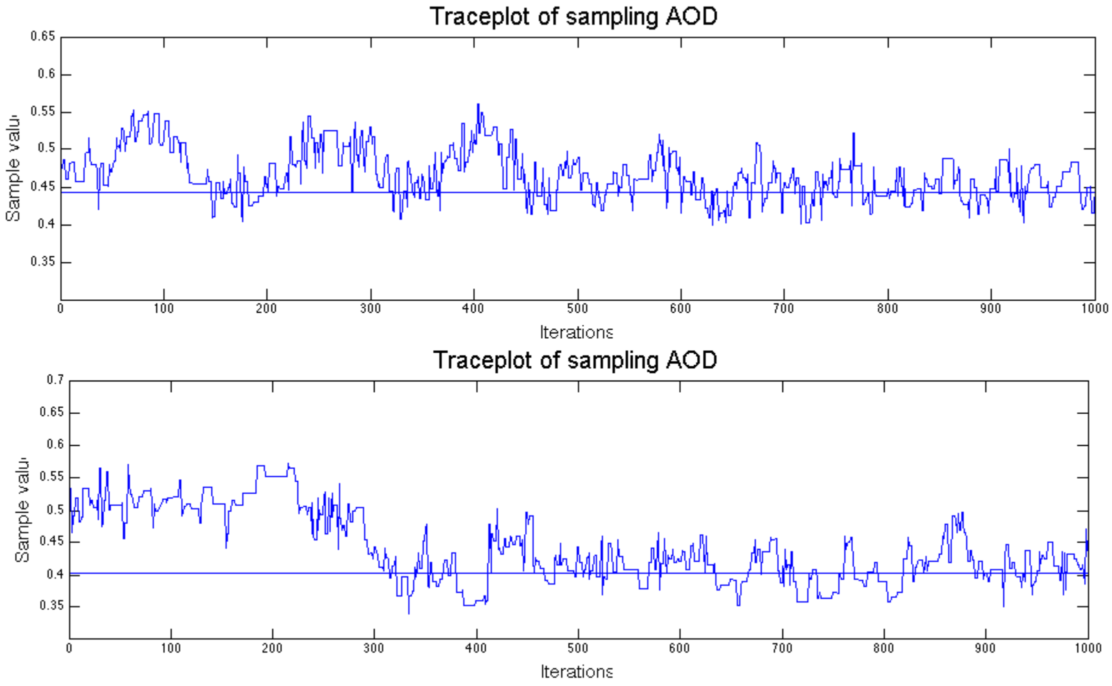


Figure 2.10: Example of sampling trace plots of AOD retrievals.

Chapter 3

An MAP Aerosol Retrieval for MISR Data at 4.4 km Resolution

3.1 Motivation

Our previous studies in Chapter 2 showed that 4.4 km resolution, together with flexibly modeled aerosol mixtures, can improve both the accuracy and spatial coverage of aerosol retrievals using MISR observed radiances. Our retrieval algorithm was based on Markov Chain Monte Carlo (MCMC) sampling methods. Like most MCMC algorithms, it was computationally intense. Continued development produced a parallel MCMC retrieval algorithm, which was approximately five times faster than the original global MCMC algorithm. It is still not fast enough to be effective for operational use: satellite Terra revolves once around the planet in 98.88 minutes; any operational algorithm needs to retrieve at a speed compatible with the observation speed.

This chapter proposes an optimization-based approach and a flexible model for aerosol mixtures at 4.4 km spatial resolution. While MISR uses an exhaustive search over a discrete grid of AOD values and the 74 MISR-prefixed mixtures, we treat both AOD and mixing vectors as continuous variables. We build upon MISR weighted least squares criterion to construct objective functions. The number of parameters to optimize, however, is large as is commonly encountered in problems involving spatial data. To reduce the model complexity, our objective functions incorporate penalty terms to borrow strength from spatial smoothness of AOD values and a global parameter for aerosol mixing vectors. The penalties further correspond to a Gaussian Markov Random Field prior for AOD values and a Dirichlet prior for aerosol mixing vectors under a hierarchical Bayesian scheme; the optimization-based approach corresponds to the Maximum a Posteriori (MAP) estimation in the Bayesian framework. This correspondence further serves us to better understand the relationship of these two frameworks: the optimization-based approach helps us see within the Bayesian hierarchy how elements interact with each other; the hierarchical Bayesian model, from another standpoint, provides a high-level picture of all variables in different objective functions. We then use gradient-based methods to find the optimal values of AOD and mixing vectors as retrieval outputs. Our MAP retrieval algorithm demonstrates a significant gain in computational efficiency, as well as an improved accuracy for high-AOD events. To assess the performance of the MAP retrieval method, we apply our retrieval algorithm to estimate AOD and mixing vectors for the greater Beijing area, using MISR observations. Our results are validated by ground measurements from the two AERONET stations in the area. They show improvement in both the accuracy and the retrieval coverage, especially in high-AOD events, enhancing our understanding of spatial and seasonal behaviors of aerosols.

The rest of the chapter is organized as follows: Section 3.2 defines the objective functions and further modifies them to allow for finer retrieval resolution and flexibility in aerosol mixing vectors. Section 3.3 provides the rationale and details of our MAP retrieval algorithm. Section 3.4 contains case studies for model validation and interpretation, comparing our results with AERONET ground measurements and MISR retrievals. In particular, a case study with severe air pollution conditions is discussed in Section 3.4.3. Section 3.5 summarizes the results and suggests directions for future research.

3.2 The Objective Functions

To assist aerosol-related public health studies, we first need to establish a data-driven description of the relationships among the remote-sensed radiances, AOD, and aerosol types. The goal of this study is to refine this description through finer retrieval resolution, specifically 4.4 km, as well as flexible models for aerosol mixtures.

The above goal, however, leads to inevitable challenges caused by the high dimensionality of problems involving spatial data at fine resolution. In our case, MISR observations during one revolution around the Earth are divided into 180 MISR Blocks, each measuring an area of 563.2 km (cross-track) × 140.8 km (along-track). In MISR retrievals at 17.6 km resolution, each MISR Block contains 256 pixels (8 rows × 32 columns). As we retrieve at 4.4 km resolution, the number of pixels rises to 4,096 (32 rows × 128 columns), presenting parameters to estimate at least 15 times more than those at the coarser resolution. Such retrievals are possible if we borrow strength from aerosols' spatial smoothness; movement and dispersion of air particles in the atmosphere justify from a physical viewpoint the inclusion of AOD spatial smoothness into our model [1] [36]. Section 3.2.3 discusses how we add a penalty term to the objective function to represent the spatial smoothness structure of AOD values.

Furthermore, the retrieval complexity continues to rise as we let component aerosols mix beyond the 74 MISR-prefixed aerosol mixtures. This freedom poses more challenges for modeling and computation. As we see through case studies in Section 3.4, however, it is necessary for retrieving heterogeneous urban aerosols. Of the 21 component aerosols established in the MISR retrieval algorithm, only eight are used in the MISR routine retrievals because they are commonly observed in the atmosphere. In fact, each of the 74 MISR mixtures contains only two or three component aerosols from the eight commonly observed ones. To control the number of parameters that must be estimated, we further choose a subset of four from the eight commonly observed components based on the study region's aerosol conditions. Then, by considering the mixing vectors as continuous variables, we allow for full flexibility in aerosol mixings relative to the four components involved. Section 3.2.2 explains the construction of the objective functions after incorporating such flexibility. In this chapter, we have not introduced an explicit spatial smoothness structure for the mixing vectors. Instead, the spatial smoothness of mixing vectors is implicitly enforced by the smoothness of AOD values; translating the relationship between variables into a hierarchical Bayesian viewpoint clarifies such implicit interactions.

We now review the notations. Let $p = 1, \ldots, P$ denote the $P = 4,096$ pixels in an MISR Block at 4.4 km resolution. For each pixel p, we obtain its top-of-atmosphere radiances $\boldsymbol{L}_p = (L_{p1}, \ldots, L_{pC})$ at $C = 36$ MISR channels (9 viewing angles × 4 spectral bands) by averaging MISR 1.1 km observations. Every channel c is weighted by $\frac{1}{2\sigma_c^2}$, $c = 1, \ldots, C$, to account for different errors that occur in measurement, radiative transfer calculation, etc. For pixel p, let $\tau_p \in \mathbb{R}$ denote its AOD value, and $\boldsymbol{\theta}_p = (\theta_{p1}, \ldots, \theta_{pM}) \in \mathbb{R}^M$ the mixing vector of the M component aerosols involved. The value of M is two or three for MISR operational retrievals, and $M = 4$ in our retrievals. We have $\sum_{m=1}^{M} \theta_{pm} = 1$. We further denote $\boldsymbol{L} = (\boldsymbol{L}_1, \ldots, \boldsymbol{L}_P)$, $\boldsymbol{\tau} = (\tau_1, \ldots, \tau_P)$, and $\boldsymbol{\theta} = (\boldsymbol{\theta}_1, \ldots, \boldsymbol{\theta}_P)$.

For each MISR output pixel at 17.6 km resolution, MISR first establishes a lookup table of simulated radiances $\boldsymbol{L}^{RT} = (L_1^{RT}, \ldots, L_C^{RT})$ using radiative transfer calculations based on a number of fixed aerosol settings. Each of these settings are characterized by an AOD level and a component aerosol[1]. Simulated radiances for an aerosol mixture are computed as a linear combination of simulated radiances for the mixing component aerosols. All simulated radiances are stores in the MISR SMART dataset. In routine retrievals, the values of $\boldsymbol{L}^{RT}$ can be obtained by interpolating at discrete points from SMART, and thus viewed as functions of AOD values and mixing vectors, relative to the component aerosols involved. For pixel p at 4.4 km resolution, we use the simulated radiances for the 17.6 km MISR pixel in which pixel p resides. Our goal is to estimate AOD values $\boldsymbol{\tau}$, and mixing vectors $\boldsymbol{\theta}$ relative to the M component aerosols involved.

The following section discusses the weighted least squares criterion used by the MISR operational retrieval algorithm, which is the basis upon which we build the objective functions in our MAP retrievals.

---

[1] Simulated radiances are computed using all combinations of AOD levels in the MISR pre-fixed grid, [0, 0.05, 0.1, 0.2, 0.4, 0.6, 0.8, 1, 1.5, 2, 3, 4, 6], and the 21 component aerosols.

3.2.1 MISR Weighted Least Squares Criterion

Use $\boldsymbol{ac}_1, \ldots, \boldsymbol{ac}_{74}$ to represent the 74 MISR-prefixed aerosol mixtures. For notational simplicity, we use $\boldsymbol{ac}_i$ for both the set of two or three component aerosols involved and their mixing vector.

For each pixel p, $p = 1, \ldots, P$, MISR observes top-of-atmosphere radiances $\boldsymbol{L}_p$ at $C = 36$ channels. Using all combinations of AOD values from the MISR-prefixed AOD grid, τ_p, and the 74 MISR aerosol mixtures, $\boldsymbol{ac}_p \in \{\boldsymbol{ac}_1, \ldots, \boldsymbol{ac}_{74}\}$, the MISR operational retrieval algorithm simulates radiative transfer radiances $\{\boldsymbol{L}^{RT}(\tau_p, \boldsymbol{ac}_p)\}$, interpolating from the MISR SMART dataset. The algorithm then exhaustively searches over all simulated radiances in order to match the observed radiances $\boldsymbol{L}_p$. In particular, the MISR operational retrieval algorithm uses the following weighted least squares criterion to measure good fits between $\boldsymbol{L}^{RT}$ and $\boldsymbol{L}_p$ [5]:

$$\chi_p^2 = \sum_{c=1}^{C} \frac{(L_{pc} - L_c^{RT}(\tau_p, \boldsymbol{ac}_p))^2}{2\sigma_c^2}. \tag{3.1}$$

The values of $\{\sigma_c^2\}_{c=1}^{C}$ are explicitly given as the minimum of 0.04 and $\bar{L}_c = (\sum_{p=1}^{P} L_{pc})/P$. The combinations of AOD and mixtures whose χ_p^2 satisfy a pre-established threshold are considered “successful” retrievals. The above search and match are carried out independently for each pixel p in an MISR Block, $p = 1, \ldots, P$; the MISR retrievals of AOD value for pixel p is the average of all successfully retrieved $\{\tau_p\}$ for that pixel.

3.2.2 Improving Retrieval Resolution and Allowing Flexibility for Mixings of Components

The observational resolution of MISR is at coarsest 1.1 km. To prepare the observations for retrieval, MISR retrieval algorithm averages the radiances to 17.6 km resolution. Instead, we aggregate MISR observed radiances from 1.1 km (or 275 m for red band and nadir camera) to 4.4 km. The reduced area over which the observed data are aggregated increases the noise level. It is thus necessary to take advantage of the spatial smoothness structure of AOD values. Details for establishing objective functions to construct this structure are introduced in Section 3.2.3.

Moreover, the MISR operational retrieval algorithm allows for only 74 aerosol mixtures. An early investigation of the missing retrievals in MISR outputs indicates that, besides the existence of clouds, many retrievals are missing because none of the 74 MISR mixtures provide good fits to the observed radiances. Models with all 21 or even the 8 common component aerosols while allowing flexible mixing vectors, however, have a large number of parameters to estimate. Such models are computationally expensive. They are also unnecessary, for current knowledge suggests that aerosol mixtures mostly consist of only two or three component aerosols[6]. We reduce the number of component aerosols involved from 21 to $M = 4$, and treat their mixing vectors as continuous variables. For example, the choices of component aerosols for the greater Beijing area include spherical non-absorbing aerosols without sulfate, spherical non-absorbing aerosols with sulfate, spherical absorbing aerosols, and grains (dust). Case studies in Section 3.4 indicate the sufficiency of this choice[2]. For each pixel p, $p = 1, \dots, P$, we use notation $\boldsymbol{\theta}_p = (\theta_{p1}, \dots, \theta_{pM}) \in \mathbb{R}^M$ for the mixing vector for the four selected components.

Now we are ready to turn the weighted least squares in (3.1) into an objective function to seek for minimal differences between the MISR observations $\boldsymbol{L}$ and the radiative transfer calculations $\boldsymbol{L}^{RT}$. The optimal values, $\boldsymbol{\tau}^*$ and $\boldsymbol{\theta}^*$, are used as point estimates for AOD values and aerosol mixing vectors. In particular, we quantify this objective through the Euclidean distance between $\boldsymbol{L}_p$ and $\boldsymbol{L}^{RT}$ and formalize the aerosol retrieval constraint in (3.1) as optimizing functions of $(\tau_p, \boldsymbol{\theta}_p)$. For each pixel p, $p = 1, \dots, P$, let

$$\tau_p^*, \boldsymbol{\theta}_p^* = \underset{\substack{\tau_p \geq 0 \\ \boldsymbol{\theta}_p \in \mathbb{R}^M, \boldsymbol{\theta}_p \cdot \mathbf{1} = 1}}{\arg\min} \quad F(\tau_p, \boldsymbol{\theta}_p), \tag{3.2}$$

where,

$$F(\tau_p, \boldsymbol{\theta}_p) = \sum_{c=1}^{C} \frac{(L_{pc} - L_c^{RT}(\tau_p, \boldsymbol{\theta}_p))^2}{2\sigma_c^2}. \tag{3.3}$$

For fixed atmospheric pressures, humidity and wind levels, the top-of-atmosphere radiances $\boldsymbol{L}$ are mainly determined by $\boldsymbol{\tau}$ and $\boldsymbol{\theta}$. Hence, if we view the objective (3.3) as a likelihood function to assess how probable are the observations $\boldsymbol{L}_p$ given τ_p and $\boldsymbol{\theta}_p$, the problem can also be formalized as inference of $(\tau_p, \boldsymbol{\theta}_p)$ based on the following Gaussian density,

$$p(\boldsymbol{L}_p | \tau_p, \boldsymbol{\theta}_p) \propto \exp\left\{ -\sum_{c=1}^{C} \frac{(L_{pc} - L_c^{RT}(\tau_p, \boldsymbol{\theta}_p))^2}{2\sigma_c^2} \right\}. \tag{3.4}$$

3.2.3 Establish Objective Functions with Spatial Smoothness

In order to build a spatial smoothness structure for AOD values of all pixels, we first combine the objectives in (3.3) for all pixels into one function:

$$F(\boldsymbol{\tau}, \boldsymbol{\theta}) = \sum_{p=1}^{P} \sum_{c=1}^{C} \frac{(L_{pc} - L_c^{RT}(\tau_p, \boldsymbol{\theta}_p))^2}{2\sigma_c^2}. \tag{3.5}$$

Due to the large number of parameters to estimate and the non-closed form of the radiative transfer calculations $\boldsymbol{L}^{RT}$ in (3.5), we simplify the multivariate optimization into

---

[2] For areas with more varying terrain types or aerosol sources, however, it can be necessary to include more component aerosols in the model. We are currently working on efficient retrievals using all of the 8 commonly observed component aerosols.

iterations of univariate optimization. Namely, for each pixel $p = 1, \ldots, P$, we first fix an aerosol mixing vector $\boldsymbol{\theta}_p$ and optimize (3.5) over its AOD value τ_p based on the observed radiance $\boldsymbol{L}_p$. Then we update its AOD value by τ_p^* and optimize (3.5) over the mixing percentage of the first component aerosol, θ_{p1}, and so forth. After updating all parameters related to the current pixel p, we move on to the next pixel. After updating all pixels, we repeat the above optimization steps in a reverse order for the MISR Block.

Now that the retrieval problem is formalized for all pixels, we continue to introduce penalty terms to the objective function in (3.5) in order to impose a spatial smoothness structure on the estimates of parameters.

Updating AOD Values $\boldsymbol{\tau}$ with Spatial Smoothness

We propose the following objective to optimize over AOD value τ_p of pixel p, a function of τ_p only:

$$F(\tau_p) = \sum_{c=1}^{C} \frac{(L_{pc} - L_c^{RT}(\tau_p, \boldsymbol{\theta}_p))^2}{2\sigma_c^2} + \frac{1}{2}\kappa \sum_{p':p'\sim p} (\tau_p - \tau_{p'})^2, \tag{3.6}$$

where $\sim$ indicates spatial adjacency[3]. The second term penalizes steep changes in AOD values among neighboring pixels. Model turning parameter κ, which calibrates the level of spatial smoothness, is also periodically updated according to (3.11). We can also obtain estimates of κ based on outputs from MISR coarse-resolution operational algorithm for the same region, or other estimates of the spatial smoothness precision, if such information is available.

The above penalty term in (3.6) for AOD's spatial smoothness can also be viewed under the hierarchical Bayesian scheme as introducing a Gaussian Markov Random Field (GMRF) prior for AOD values. Firstly, in correspondence with (3.5), the likelihood function (3.4) for a single pixel p can be expanded into one for all pixels together:

$$p(\boldsymbol{L}|\tau, \boldsymbol{\theta}) \propto \exp\left\{ -\sum_{p=1}^{P}\sum_{c=1}^{C} \frac{(L_{pc} - L_c^{RT}(\tau_p, \boldsymbol{\theta}_p))^2}{2\sigma_c^2} \right\}. \tag{3.7}$$

Then the spatial smoothness in $\boldsymbol{\tau}$ can be incorporated through prior distributions. We model the spatial smoothness structure of AOD values using an intrinsic GMRF prior of first order[30]. If we consider κ as the homogenous scalar precision, the following GMRF prior is invariant to perturbation by the same constant to $\boldsymbol{\tau}$ of all pixels[2],

$$p(\boldsymbol{\tau}|\kappa) \propto \kappa^{\frac{P-1}{2}} \exp\left\{ -\frac{\kappa}{2} \sum_{p'\sim p} (\tau_{p'} - \tau_p)^2 \right\}. \tag{3.8}$$

---

[3] In this study, we consider that each central pixel has four adjacent pixels.

The larger κ, the smoother the region's AOD values are, with $\frac{1}{\sqrt{\kappa}}$ as AOD's precision. The marginal posterior of $\boldsymbol{\tau}$ is then,

$$p(\boldsymbol{\tau}|\boldsymbol{\theta}, \kappa, \boldsymbol{\sigma}, \boldsymbol{L}) \propto \exp\left\{-\sum_{p=1}^{P}\sum_{c=1}^{C}\frac{(L_{pc} - L_c^{RT}(\tau_p, \boldsymbol{\theta}_p))^2}{2\sigma_c^2} - \frac{1}{2}\kappa\sum_{p'\sim p}(\tau_{p'} - \tau_p)^2\right\}. \tag{3.9}$$

In fact, we see that the cost function (3.6) is the negative logarithm of (3.9), and such optimization over posteriors is referred to as an Maximum a Posteriori (MAP) algorithm. The minimum of the objective functions in turn corresponds to the mode of the posterior distributions.

Updating AOD Spatial Precision κ

Given the above Bayesian interpretation, the posterior distribution of the homogenous precision $\kappa \in \mathbb{R}$ conditioning on AOD values $\boldsymbol{\tau}$ is a proper Gamma distribution. Namely,

$$p(\kappa|\boldsymbol{\tau}) \propto \kappa^{\frac{P-1}{2}-1}\exp\{-\frac{\kappa}{2}\sum_{p'\sim p}(\tau_{p'} - \tau_p)^2\}. \tag{3.10}$$

After updating the AOD values $\boldsymbol{\tau}$ according to the objective function in (3.6), we can use the updated $\boldsymbol{\tau}$ to renew the value of κ. In particular, we use the posterior mode of (3.10),

$$\kappa^* = \frac{P-3}{\sum_{p'\sim p}(\tau_{p'} - \tau_p)^2}. \tag{3.11}$$

Updating Aerosol Mixing Vectors $\boldsymbol{\theta}$ with a Global Dirichlet Parameter

Similarly, we assign the mixing vectors of component aerosols $\boldsymbol{\theta}$ a global prior distribution to impose an overall sparsity level of the mixings within an MISR Block. Mode of the posterior distribution is then used to update the value of $\boldsymbol{\theta}$ at each iteration of the MAP retrieval algorithm. Prior distributions for $\boldsymbol{\tau}$ and $\boldsymbol{\theta}$ are further assumed to be independent, i.e., $p(\boldsymbol{\tau}, \boldsymbol{\theta}) = p(\boldsymbol{\tau})p(\boldsymbol{\theta})$ to simplify the resulted optimization. There is no explicit spatial structure for $\boldsymbol{\theta}$, the interactions between $\boldsymbol{\tau}$, $\boldsymbol{\theta}$ and observed radiances $\boldsymbol{L}$ through the hierarchy implicitly enforce a spatial dependence structure for $\boldsymbol{\theta}$. Results of case studies display such spatial smoothness patterns as we anticipated (see Section 3.4).

To model the mixing vectors $\boldsymbol{\theta}$ for the M component aerosols involved, we adopt an M-dimensional Dirichlet prior with parameter $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_M)$. Define $\alpha_0 = \sum_{m=1}^{M}\alpha_m$. By adjusting the magnitude of $\boldsymbol{\alpha}$, we can control the sparsity of the mixings of the M components. Conditioning on $\boldsymbol{\alpha}$, the mixing vectors $\boldsymbol{\theta}_p$ for each pixel p are considered to be independent,

$$p(\boldsymbol{\theta}|\boldsymbol{\alpha}) = \prod_{p=1}^{P} p(\boldsymbol{\theta}_p|\boldsymbol{\alpha}) = \prod_{p=1}^{P} \frac{\Gamma(\sum_{m=1}^{M}\alpha_m)}{\prod_{m=1}^{M}\Gamma(\alpha_m)} \theta_{p1}^{\alpha_1-1}\cdots\theta_{pM}^{\alpha_M-1}. \tag{3.12}$$

The marginal posterior distribution of $\boldsymbol{\theta}$ can then be expressed as:

$$p(\boldsymbol{\theta}|\boldsymbol{\tau},\boldsymbol{\alpha},\boldsymbol{\sigma},\boldsymbol{L}) \propto \exp\left\{\sum_{p=1}^{P}\sum_{m=1}^{M}(\alpha_m-1)\log\theta_{pm} - \sum_{c=1}^{C}\sum_{p=1}^{P}\frac{(L_{pc}-L_c^{RT}(\tau_p,\boldsymbol{\theta}_p))^2}{2\sigma_c^2}\right\}. \tag{3.13}$$

Based on the above result, we establish the following objective function for aerosol mixing vectors,

$$\begin{aligned} F(\boldsymbol{\theta}) &= F(\boldsymbol{\theta}_1,\ldots,\boldsymbol{\theta}_p) && (3.14)\\ &= \sum_{p=1}^{P}\sum_{c=1}^{C}\frac{(L_{pc}-L_c^{RT}(\tau_p,\boldsymbol{\theta}_p))^2}{2\sigma_c^2} - \sum_{p=1}^{P}\sum_{m=1}^{M}(\alpha_m-1)\log\theta_{pm} && (3.15)\\ &= \sum_{p=1}^{P}\sum_{c=1}^{C}\frac{(L_{pc}-L_c^{RT}(\tau_p,\boldsymbol{\theta}_p))^2}{2\sigma_c^2} - \sum_{p=1}^{P}\sum_{m=1}^{M}\log\theta_{pm}^{(\alpha_m-1)}. && (3.16)\end{aligned}$$

The second term in the above objective serves as penalties for aerosol mixings $(\theta_{p1},\ldots,\theta_{pM})$ in large disagreement with the Dirichlet ratio $(\frac{\alpha_1}{\alpha_0},\ldots,\frac{\alpha_M}{\alpha_0})$. The Dirichlet parameter α thus implicitly serves to calibrate the general level of spatial smoothness for the mixing vectors over the MISR Block.

During optimization retrievals, the objective in (3.15) can be viewed as a univariate function of θ_{pm}, having fixed the values of the other variables.

Updating Aerosol Mixing Dirichlet Parameter $\boldsymbol{\alpha}$

In order to determine the values for the Dirichlet parameter $\boldsymbol{\alpha}$, we assign $\boldsymbol{\alpha}$ a hyperprior and add another layer of optimization to search for their appropriate values. Notice that (3.12) belongs to an exponential family. For convenience, we take the conjugate distribution as a prior for $\boldsymbol{\alpha}$. The posterior becomes,

$$p(\boldsymbol{\alpha}|\boldsymbol{\theta}) \propto \exp\{\sum_{m=1}^{M}(\alpha_m-1)(\sum_{p=1}^{P}\log\theta_{pm}+1) - P\times(\sum_{m=1}^{M}\log\Gamma(\alpha_m) - \log\Gamma(\sum_{m=1}^{M}\alpha_m))\}. \tag{3.17}$$

The above posterior includes Gamma functions; calculations of both derivatives and inverse functions of Gamma functions usually resort to numerical approximations. Due to computational concerns, instead of using gradient-based methods, we update the values of $\boldsymbol{\alpha}$ using estimates of moments based on the Dirichlet prior in (3.12). Recall that $\alpha_0 = \sum_{m=1}^{M}\alpha_m$.

We have the following:

$$E(\theta_m) = \frac{\alpha_m}{\alpha_0}, \tag{3.18}$$

$$Var(\theta_m) = \frac{\alpha_m(\alpha_0 - \alpha_m)}{\alpha_0^2(\alpha_0 + 1)}. \tag{3.19}$$

For $m = 1, \ldots, M$, we use the sample mean $\bar{\boldsymbol{\theta}}_m$ and sample variance s_P^2 as estimates for the mean and variance of the prior for $\boldsymbol{\theta_m}$,

$$\bar{\boldsymbol{\theta}}_m = \frac{1}{P}\sum_{p=1}^{P} \theta_{pm}, \tag{3.20}$$

$$s_m^2 = \frac{1}{P}\sum_{p=1}^{P} (\theta_{pm} - \bar{\boldsymbol{\theta}}_m)^2. \tag{3.21}$$

Relating (3.20) to (3.18) and (3.21) to (3.19), we can solve the following estimates for $\boldsymbol{\alpha}$:

$$\hat{\alpha}_m = \frac{\bar{\boldsymbol{\theta}}_m^2(1 - \bar{\boldsymbol{\theta}}_m)}{s_m^2} - \bar{\boldsymbol{\theta}}_m, \quad \text{for } \ m = 1, \ldots, M. \tag{3.22}$$

Updating MISR Channel Variance $\boldsymbol{\sigma}^2$

The likelihood function for $\boldsymbol{\sigma^2} = (\sigma_1^2, \ldots, \sigma_C^2)$, $p(\boldsymbol{L}|\boldsymbol{\sigma}^2, \boldsymbol{\tau}, \boldsymbol{\theta})$ in (3.7), follows a Gaussian distribution with known mean and unknown variance. We adopt a noninformative inverse-χ^2 hyperprior for $\boldsymbol{\sigma}^2$: $p(\sigma_c^2) \propto \sigma_c^{-2}$, as a choice of computational convenience. The conditional posterior follows a proper scaled inverse-χ^2 distribution,

$$p(\sigma_c^2|\boldsymbol{\tau}, \boldsymbol{\theta}, \boldsymbol{L}) \propto (\sigma_c^2)^{-\frac{P}{2}+1} \exp\left\{-\frac{\sum_{p=1}^{P}(L_{pc} - L_c^{RT}(\tau_p, \boldsymbol{\theta}_p))^2}{2\sigma_c^2}\right\}. \tag{3.23}$$

We can further update $\boldsymbol{\sigma}^2$ using the corresponding posterior mode:

$$(\sigma_c^2)^* = \frac{\sum_{p=1}^{P}(L_{pc} - L_c^{RT}(\tau_p, \boldsymbol{\theta}_p))^2}{P+2}. \tag{3.24}$$

Now that we have established univariate objective functions given fixed values of all other variables, the next section describes our optimization algorithm based on these objective functions. The algorithm takes MISR observed radiances $\boldsymbol{L}$ as input, and outputs the optimal AOD values and mixing vectors, namely the posterior modes, as our retrievals.

3.3 The Maximum a Posteriori Retrieval Algorithm (MAP)

For all P pixels in an MISR Block, our goal is to estimate their AOD values $\boldsymbol{\tau}$ and mixing vectors $\boldsymbol{\theta}$ relative to the M component aerosols, with the exception of unretrievable pixels[4]. Thus, we have more than $[1 + (M - 1)] \times P$ parameters to estimate, since the degree of freedom in mixing vectors is $M - 1$[5]. We break this multi-dimensional optimization problem into iterations of univariate optimizations.

In particular, Section 3.2.3 calculated the updating rules for AOD spatial precision κ, the Dirichlet parameter $\boldsymbol{\alpha}$ and MISR channel variance $\boldsymbol{\sigma}^2$. We also established univariate objective functions for AOD $\boldsymbol{\tau}$ and mixing vectors $\boldsymbol{\theta}$, having the values of all other variables fixed. This section describes how to further obtain the optimum of $\boldsymbol{\tau}$ and $\boldsymbol{\theta}$ based on the above objective functions, using gradient-descent methods.

Convexity, or local convexity, is highly desirable for efficiently using gradient-based optimization methods to acquire the correct optimum. In our case, it is difficult to analytically prove local convexity of the objective functions, due to the non-closed-form segment of radiative transfer simulated radiances $\boldsymbol{L}^{RT}$. Empirical results show that, in fact, the objective functions are not strictly convex, but the global minima exists. The examples in Figure 3.1 indicate that the objective functions, i.e. negative log-posterior[6], are locally convex to τ around the minima, and globally convex to the first element of aerosol mixing vector. Similar patterns are observed for other elements of the mixing vector.

Case studies suggest that the cost functions are close enough to being convex with respect to the choice of initial points. Hence, with an appropriately chosen initial step size δ, we can adopt the steepest direction optimization algorithm to find the global minima. The Armijo rule is used to choose the step size after the first step. The results also prove to be reliable and stable with different starting values.

Now we are going to discuss in detail our MAP retrieval algorithm.

---

[4] MISR-observed radiances are flagged as unusable for aerosol retrieval, if cloud occurs, or the observations fail to pass MISR Radiometric Data Quality test, etc.

[5] Based on cases in the greater Beijing area over the period of December 2008 to June 2009, the number of retrievable pixels in an MISR Block ranges between 100 and 1,600 at 4.4 km resolution (Section 3.4).

[6] The figures adopt the logarithm scale in order to clearly illustrate the shape of equivalent cost functions in range (0,2) for τ and (0, 0.7) for θ_1.

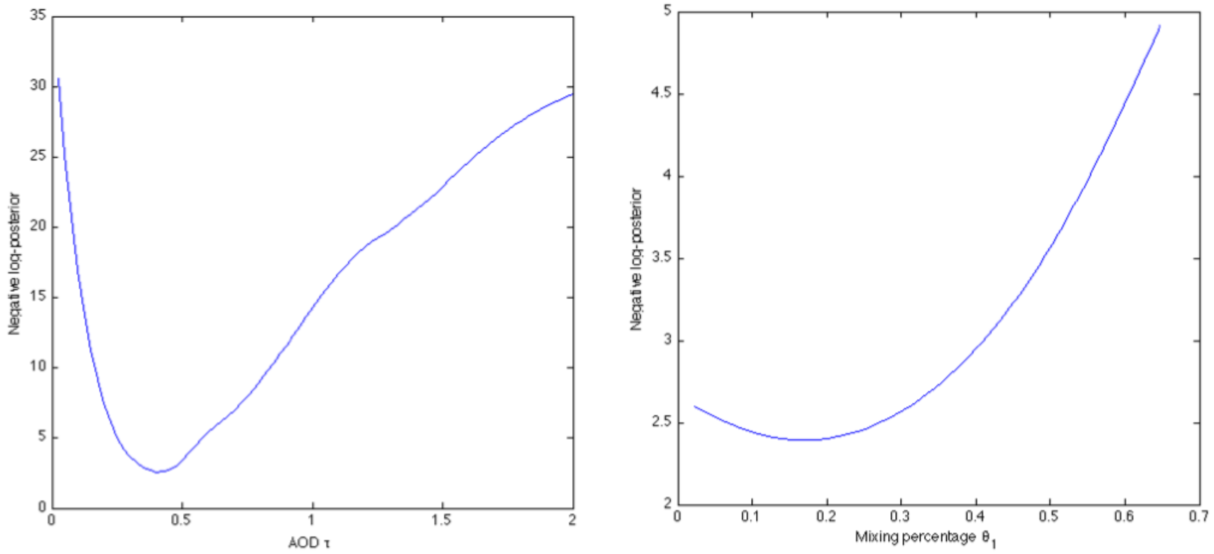


Figure 3.1: Convexity of the negative log-posterior to AOD τ (left panel, with $\theta_1 = 0.4782$) and mixing percentage θ_1 for component aerosol 1 (right panel, with $\tau = 0.5000$).

Optimizing Objectives over AOD $\boldsymbol{\tau}$

For each pixel p, $p = 1, \dots, P$, to minimize the objective function in (3.6) over AOD τ_p, we consider a small perturbation τ^δ of size δ to a current approximation of τ_p:

$$\tau_p^\delta = \tau_p - \delta \frac{\mathrm{d}F}{\mathrm{d}\tau_p}. \tag{3.25}$$

The derivative of the objective function can be approximated using numerical differences as follows:

$$\frac{\mathrm{d}F}{\mathrm{d}\tau_p} \approx \sum_{c=1}^{C} \frac{L_c^{RT}(\tau_p, \boldsymbol{\theta}_p) - L_{pc}}{\sigma_c^2} \left(\frac{L_c^{RT}(\tau_p + \triangle\tau, \boldsymbol{\theta}_p) - L_c^{RT}(\tau_p, \boldsymbol{\theta}_p)}{\triangle\tau} \right) + \kappa \sum_{p':p\sim p'} (\tau_p - \tau_{p'}), \tag{3.26}$$

where $\triangle\tau$ is taken as 0.025, the AOD grid-search gap size in MISR AOD retrieval algorithm.

From the first order Taylor expansion around τ_p we have

$$\begin{aligned} F(\tau_p^\delta) &= F(\tau_p) + \frac{\mathrm{d}F}{\mathrm{d}\tau_p}(\tau_p^\delta - \tau_p) + o(|\tau_p^\delta - \tau_p|) \\ &= F(\tau_p) - \delta(\frac{\mathrm{d}F}{\mathrm{d}\tau_p})^2 + o(\delta). \end{aligned}$$

Thus, for positive but sufficiently small δ, $F(\tau_p^\delta)$ is smaller than $F(\tau_p)$. Hence, by updating τ_p to τ_p^δ, the algorithm is approaching the minima of the objective function $F(\tau_p)$ in (3.6).

Optimizing Objectives over Mixing Vectors $\boldsymbol{\theta}$

We then optimize the objective functions in (3.15) over the mixing vector $\boldsymbol{\theta}_p$ in a similar manner, using the steepest direction optimization method.

One of the major difference in optimizing over $\boldsymbol{\theta}_p$ from optimizing over $\boldsymbol{\tau}_p$ is the constraint that the elements of the mixing vector sum up to 1, namely $\sum_{m=1}^M \theta_{pm} = 1$, for any pixel p. For $m = 1, \dots, M$, we maintain the constraint by perturbing θ_{pm} by $\triangle\theta$ and re-normalizing:

$$\theta_{pm}^\delta = \frac{\theta_{pm} - \delta\triangle\theta}{1 - \delta\triangle\theta}, \tag{3.27}$$

$$\theta_{pm'}^\delta = \frac{\theta_{pm'}}{1 - \delta\triangle\theta}, \qquad m' = 1, \dots, M,\, m' \neq m. \tag{3.28}$$

The perturbation size δ to optimize over $\boldsymbol{\theta}$ can take a different value than that for optimizing over τ. The perturbation unit $\triangle\theta$ is chosen to be along the steepest direction of the objective function, with an absolute value no greater than 0.01, while maintaining the non-negativity of $\boldsymbol{\theta}$. We also notice that a smaller θ_{pm} is more sensitive to such perturbations than a larger θ_{pm}:

$$|\theta_{pm}^\delta - \theta_{pm}| = \frac{\delta|\triangle\theta|}{1 - \delta\triangle\theta} \cdot (1 - \theta_{pm}). \tag{3.29}$$

In other words, a small value of θ_{pm} leads to a proportionally large difference between the mixing percentage before and after the perturbation. This property can be beneficial in two ways. Firstly, we would like to explore all possible mixings of component aerosols. With (3.29), even if the initial mixture approximation assigns a close-to-0 percentage to a certain component, the algorithm will not be forced to ignore this component as the algorithm iterates to optimize. Secondly, if a certain component aerosol currently fills up most of the aerosol mixture, the algorithm is reluctant to further increase its mixing percentage.

Description of MAP

The complete Maximum a Posteriori Retrieval Algorithm (MAP) can be summarized as follows. The number of iterations (line 2 of Algorithm 1) is generally a trade-off between computation resources and estimation accuracy. Case studies show that MAP typically converges within six iterations so that we can get the number of iterations $I = 6$.

Like many algorithms with alternations of univariate optimizations, one concern of MAP is whether optimizing over the variables in different orderS will significantly change the algorithm outputs, and if so, how to choose the order in which to optimize. For each pixel p, $p = 1, \dots, P$,the overall negative log-posterior is a function of five variables: AOD τ_p, aerosol

Algorithm 4 The Maximum a Posteriori Retrieval Algorithm (MAP)

```
For an MISR Block,
1: Initiate model parameters with random values drawn from their prior distributions.
2: for $i = 1$ to $I$ do
3:     for $p = 1$ to $P$ do
4:         Use (3.25) to update AOD value $\tau_p$ according to (3.6), based on MISR observations
           $\boldsymbol{L}$, the updated values of $\{\tau_q\}_{q \neq p}$, current values of $\boldsymbol{\theta}$, $\boldsymbol{\alpha}$ and $\boldsymbol{\sigma}^2$.
5:         Use (3.27) to update the mixing vector $\boldsymbol{\theta}_p$ according to (3.15), based on the updated
           value of $\{\tau_q\}_{q \leq p}$.
6:     Update AOD's spatial precision parameter $\kappa$ according to (3.11), using the updated
       values of $\boldsymbol{\tau}$.
7:     Update the tuning parameter $\boldsymbol{\alpha}$ according to (3.22), using the updated values of $\boldsymbol{\theta}$.
8:     Sweep the MISR Block in a reverse order and apply the same optimization steps as
       lines 3 to 7, in order to incorporate the information from all adjacent pixels regardless
       of the position.
```

mixing vector $\boldsymbol{\theta}_p = (\theta_{p1}, \ldots, \theta_{pM})$, where $M = 4$. To identify the best order, it is important to understand how these variables influence the behavior of the log-posterior, and how they interact with each other. Figure 3.2 shows that the relationship among the negative log-posterior, τ and θ_1 for a typical pixel on an MISR Block. The pattern we observe in Figure 3.2 is similar for other pixels and MISR Blocks: the negative log-posterior is more sensitive to the change of AOD $\boldsymbol{\tau}$, than to that of the mixing vectors $\boldsymbol{\theta}$. Therefore, our algorithm optimizes the objective function over $\boldsymbol{\tau}$ before it seeks the optimal values of $\boldsymbol{\theta}$. The updating order within $(\theta_{p1}, \ldots, \theta_{pm})$ is assigned randomly and independently during each iteration.

The next section discusses some results of aerosol retrieval using the MAP and further evaluates the algorithm's performance based on case studies for the greater Beijing area, China. It also discusses the informational and computational gain of our optimization approach.

3.4 Results

First of all, we compare our MAP AOD retrievals at 4.4 km resolution to MISR AOD retrievals at 17.6 km resolution, using the latter as a benchmark. Results show overall agreement between the two sets of retrievals, while the finer resolution is characterized by more varying levels of AOD values. To further validate the reliability of this increased variability of MAP AOD retrievals, we resort to ground measurements provided by the AErosol RObotic NETwork (AERONET)[14]. We will then demonstrate the necessity of flexibility in modeling aerosol mixtures through case studies. Finally, we showcase two retrieval instances with high-AOD conditions captured by our MAP AOD retrievals but not

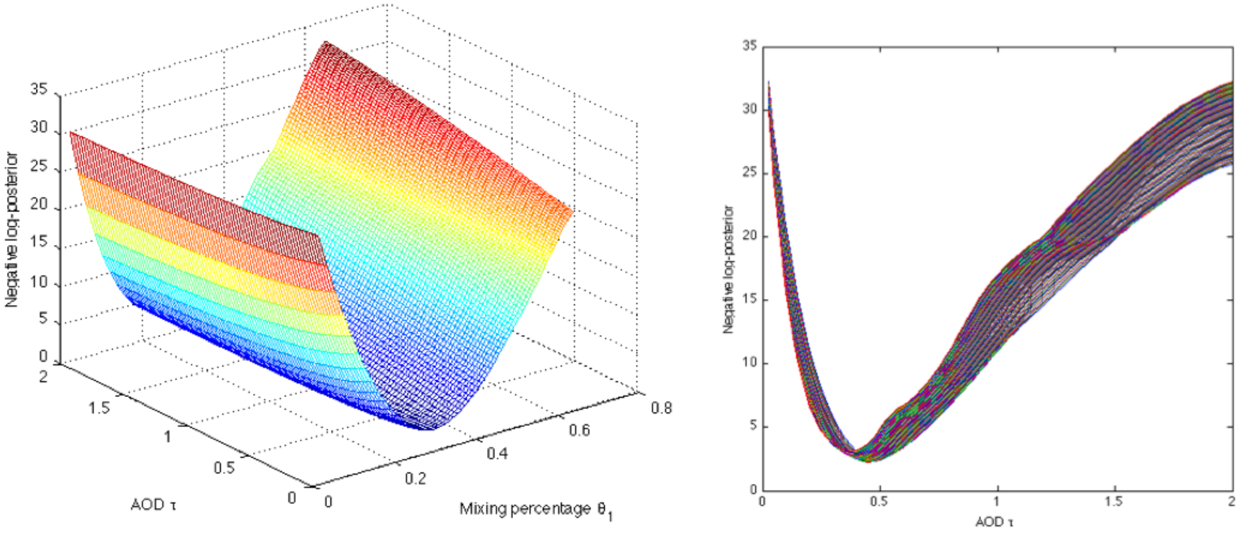


Figure 3.2: Left panel: convexity of negative log-posterior as a function of AOD τ and mixing percentage θ_1 of aerosol component 1 (spherical non-absorbing aerosols). Right panel: side view of the three dimensional surface.

the MISR AOD retrievals. Such events are especially important, since they indicate heavy air pollution sufficient to have a long term impact on human health.

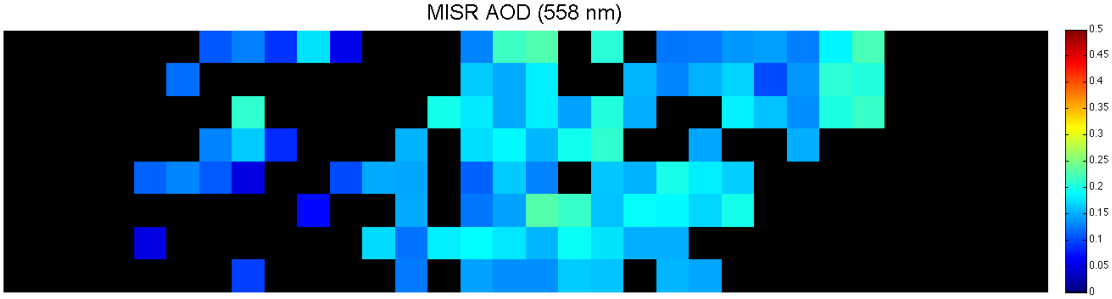
The retrievals studied in this section are based on the MISR observations over the greater Beijing area in China (latitude 38.95N∼40.15N; longitude: 115.57E∼119.50E).

3.4.1 Comparison with MISR Retrievals

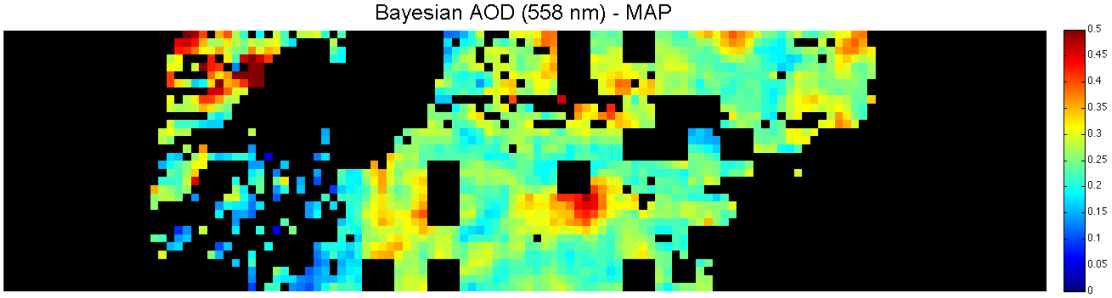
Figure 3.3 panel (a) shows the observed MISR imagery captured by one of the forward cameras (CF) on Mar 06, 2009. MISR AOD retrievals are plotted at 17.6 km resolution in panel (b). Our MAP AOD retrievals are plotted in panel (c) at 4.4 km. Figure 3.3 suggests overall agreement between MISR AOD and our MAP AOD retrievals, such as the coastline on the right and patterns of AOD spatial distribution. For example, both panel (b) and (c) display lower AOD values in the lower lefthand corner and higher values in the top righthand corner, as well as in the lower middle portion of this MISR Block.

The black pixels represent missing retrievals, which are primarily caused by two situations. Firstly, aerosol retrievals are not attempted in cloudy situations. Particularly, cloud detections are implemented at 1.1 km resolution on MISR observations. Each MISR output pixel at 17.6 km resolution corresponds to 256 “child” pixels at 1.1 km resolution, each of which is flagged cloudy or clear. This high level of information aggregation from 1.1 km to 17.6 km resolution allows MISR to retrieve for a 17.6 km pixel as long as no more than $\frac{1}{16}$ of its “child” pixels are cloudy. To maintain the same cloudy/clear ratio at 4.4 km resolution, however, is a much stricter requirement for a pixel to be treated as clear. This requirement is much less often met, since the common size of clouds is beyond 1.1 × 1.1 km^2. Hence, we sometimes observe more missing retrievals at 4.4 km resolution that are caused to cloud screening.

(a) MISR radiance imagery by the nadir camera on Mar 06, 2009.


MISR AOD (558 nm)

(b) MISR AOD retrievals at 17.6 km resolution.


Bayesian AOD (558 nm) - MAP

(c) Bayesian AOD retrievals using MAP at 4.4 km resolution.

Figure 3.3: AOD estimates from MISR and MAP retrievals.

The second situation where missing values occur is due to the limitation of the 74 MISR-prefixed aerosol mixtures. When none of these 74 mixtures produce good fits to the MISR observed radiances, the MISR operational retrieval algorithm marks the pixel with missing retrieval. Our MAP retrieval algorithm allows for a richer variety of aerosol mixtures, and thus is able to eliminate these unnecessary missing retrievals. More details and an example is given in Section 3.2.2.

Another advantage of our MAP retrieval algorithm lies in its computational efficiency. Table 3.1 displays one typical example of the computational expenses of three methods: the MISR operational retrieval algorithm, our MAP retrieval algorithm and an aerosol retrieval algorithm using MCMC sampling methods, developed in a previous study [35]. The MCMC-based algorithm is computationally the most intense.

Even though MAP allows for a much larger variety of aerosol mixtures, its steepest gradient direction method proves to be more efficient: the MAP at 4.4 km resolution is almost three times faster than the MISR operational algorithm at the much coarser resolution of 17.6 km. This is largely due to the nature of optimization. Namely, once MAP tests the agreement of a combination of AOD and mixing vector to the observed radiances, it does not have to test any combinations which perform worse than one already evaluated. Thus, computation is largely saved without the exhaustive search as in the MISR operational retrieval algorithm.

Table 3.1: Computational Expenses of Retrieval Algorithms.

Algorithm (resolution)	Computation Time for Retrieval
MISR (17.6 km)	$\sim$ 16.7 minutes (1004.03 sec)
MCMC sampling (4.4 km)	$\sim$ 6.2 hours (22452.26 sec)
MAP (4.4 km)	$\sim$ 6 minutes (378.91 sec)

On the other hand, though our MAP AOD retrievals appear to be consistent with the MISR AOD retrievals on the whole, our finer-resolution retrievals in panel (c) of Figure 3.3 reveal more widely varying levels of AOD values. We also observe high AOD values in the top lefthand corner, which are missing in MISR AOD retrieval outputs shown in panel (b) of Figure 3.3. To objectively evaluate the performance of our MAP and understand the above differences in the retrieval outputs, the next section conducts a validation study using ground measurements from AERONET.

3.4.2 Validation for MAP Retrievals by AERONET Measurements

The AERONET provides a data archive of local AOD values using a network of automatic sun photometers located at more than 400 stations on the Earth's surface. At each station, local AOD level is measured from every half hour to every two hours, with AOD uncertainties $< \pm 0.01$ at wavelengths > 440 nm[15]. AERONET measurements are widely accepted as the gold standard for validating AOD estimates based on remote-sensed imageries, and hence our choice of source for validation.

To validate aerosol retrievals over the greater Beijing area, we use ground measurements of AOD collected at AERONET Beijing and AERONET Xianghe stations. The fixed locations of the AERONET stations make it impossible to validate retrievals of all pixels in the MISR Block. Instead, we focus on the pixels that contain the AERONET stations. To match the MAP retrieval outputs at the wavelength of 550 nm, we first convert AERONET measurements to those at 550 nm using AERONET estimates of Angström exponent and a linear transformation. Jiang, *et al.* [16] showed that a narrower time window better captures the correlations between AERONET measurements and MISR retrievals. Thus, we then average the measurements within a one-hour window when satellite Terra carrying MISR passes over the AERONET stations. The frequent cloudy weather in the greater Beijing area and its latitude[7] contribute to the difficulty of matching remote-sensed versus ground-based data pairs for model validation.

Figure 3.4 presents our MAP retrievals plotted in blue on the y-axis against AERONET measurements on the x-axis, with MISR retrievals in black for comparison. All matched data pairs of remote-sensed AOD and ground measurements are included in the left panel; the right panel plots only the pairs with AERONET measurements ≤ 0.5, so as to obtain an appropriate scale.

In low-AOD situations, as is shown in the right panel of Figure 3.4, our MAP retrieval algorithm mostly overestimated the AOD values, compared to AERONET measurements. In high-AOD situations, on the other hand, the performance of our MAP AOD retrievals is much better compared to MISR retrievals, using AERONET measurements as a standard. For example, the three high AOD values in the top righthand corner of the left panel, Figure 3.4, indicate that Beijing is characterized by heavy air pollution relatively frequently. These

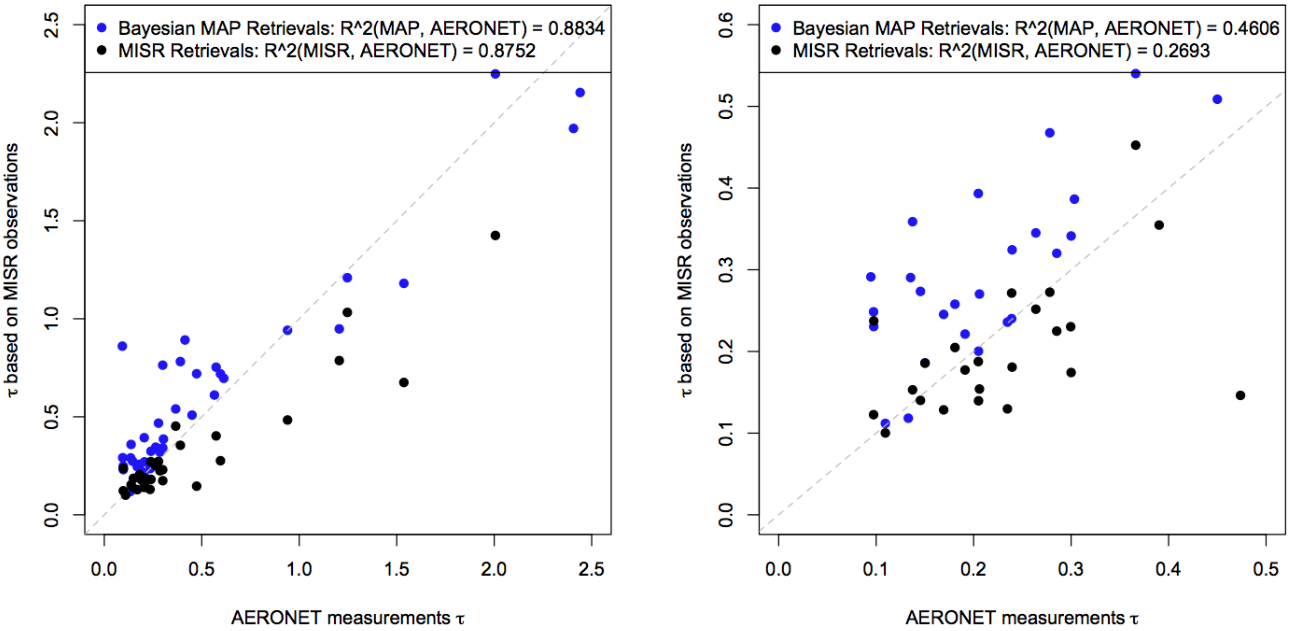


Figure 3.4: Validation of our MAP AOD retrievals.

---

[7] The Beijing city is visited by the Terra satellite every five to nine days due to the area's latitude.

three cases have AOD ground measurements larger than 1.9, corresponding to more than 85% reduced radiation by aerosols. Out of the three cases, the MISR operational algorithm reported two missing retrievals and underestimated the AOD level in the third case, when compared to AERONET measurements. The MAP AOD retrievals manage to agree well with the AERONET measurements in these three high-AOD situations. This suggests that for high aerosol loadings, the "optimized" combination of AOD and mixtures, as obtained in the MAP retrievals, is able to better describe the observed conditions; the average of all "successful" retrieved AOD, as produced by the MISR retrieval algorithm, is more suitable for low-AOD situations.

Meanwhile, a flexible model of aerosol mixtures improves the retrieval outputs' spatial coverage as well as accuracy, especially in high-AOD situations. In particular, as long as a pixel is cloudless, our MAP provided an AOD retrieval, as in all of the 37 cases in Figure 3.4. The MISR operational retrieval algorithm, however, produced missing retrievals for 19% of these 37 cases. These missing retrievals result from the limited choices of 74 aerosol mixtures in MISR retrievals and the increasingly heterogeneous aerosol conditions in Beijing. It further emphasizes the necessity in allowing for flexible aerosol mixings in retrieving urban aerosols.

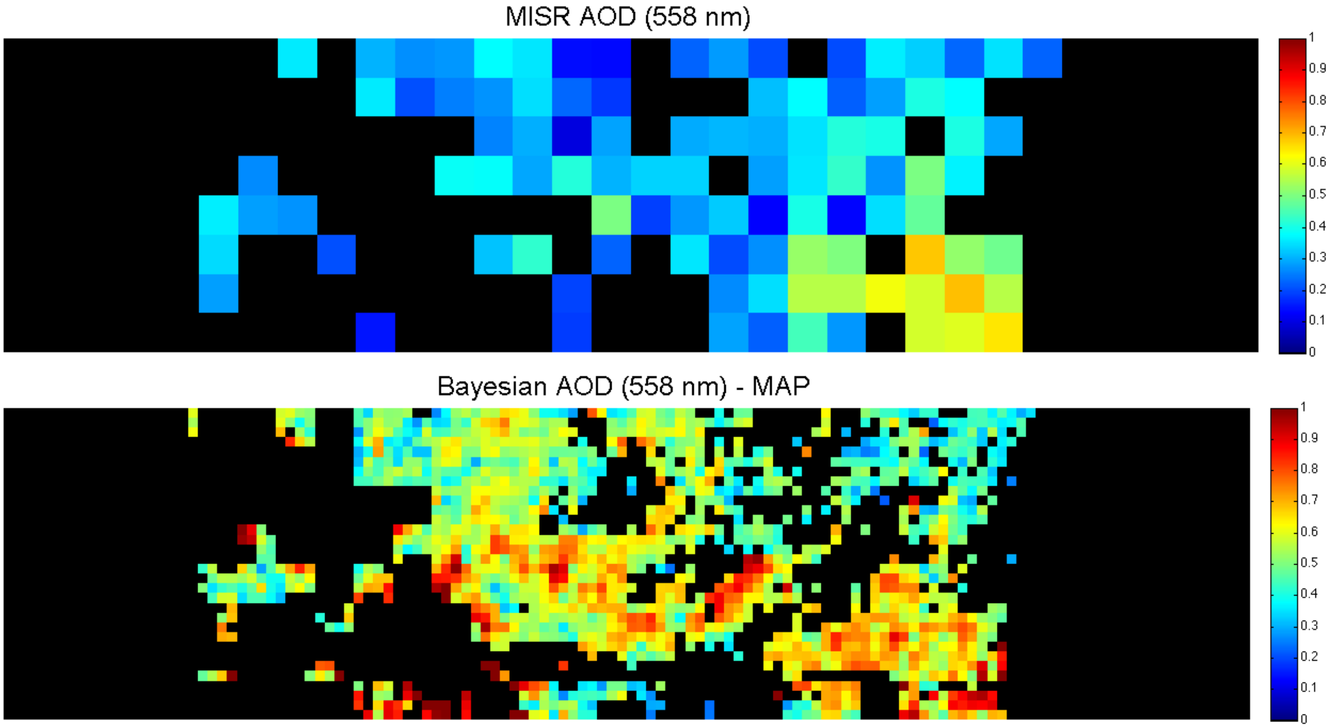
3.4.3 High/Medium-AOD Situations

Finally, we showcase one more high-AOD situation and a medium-AOD situation, where MAP managed to capture the weather conditions in mainland China. We first look at the aerosol conditions in Beijing on April 14, 2009.

Using MAP Retrievals to Depict a Dust Storm in Beijing and Zhengzhou on April 14, 2009

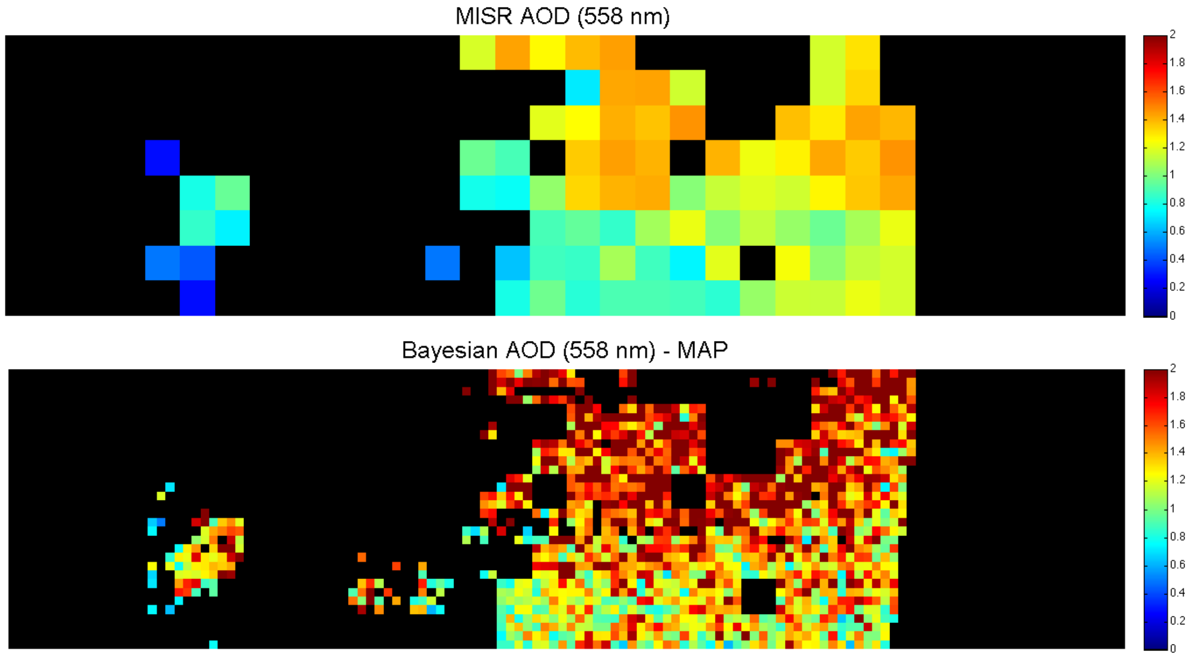
Figure 3.5 shows the MISR and MAP retrieval results over the greater Beijing area on April 14, 2009, at spatial resolutions of 17.6 km and 4.4 km respectively. The MAP AOD retrievals are overall higher than those of MISR AOD retrievals across the MISR Block. We also observe more high MAP-retrieved AOD values in the lower half of the MISR Block, which are absent in MISR retrievals.

Another search on the weather archive for April 14 provides us another dust storm alert for central China. This explains the overall high AOD values in MAP retrievals. However, we are currently unable to find concrete evidence for or against the high values in the middle of the Block, because of the lack of ground measurements in the great Beijing area. Instead, we turn our analysis to the center of this particular dust storm, Zhengzhou, Henan Province. AOD retrievals for the greater Zhengzhou area are plotted in Figure 3.6, with MISR retrievals in the upper panel and MAP retrievals in the lower panel. As is shown, the MAP retrievals display an apparent pattern of high AOD levels spreading, where the AOD values are consistent with the AERONET measurements of similar conditions. The


MISR AOD (558 nm)

Bayesian AOD (558 nm) - MAP

Figure 3.5: Case study of AOD retrievals on April 14, 2009 over the greater Beijing area.


MISR AOD (558 nm)

Bayesian AOD (558 nm) - MAP

Figure 3.6: Case study of AOD retrievals on April 14, 2009 at Zhengzhou, Henan.

MISR retrievals in the upper panel, however, seem to underestimate the AOD values in this situation.

From the two cases we discussed in this and the previous sections, we see that MAP retrievals are able to depict dust storms with an better coverage and accuracy than MISR. For regions like Beijing and Zhengzhou which experience occasional severe weather conditions, the 74 MISR-prefixed aerosol mixtures contain a very limited variety of how grains (dust) aerosols exist in the atmosphere. Failing to provide an aerosol mixture that reflects the observed situations could easily result in a low retrieval coverage and underestimation of AOD levels. In the aerosol-related public health studies, however, it is more crucial to accurately estimate the AOD values for high-AOD events. They have a larger impact on human health than the fine weather.

Using MAP Retrievals to Capture Air Pollution in the Beijing Area on January 24, 2009

Here we present an event with medium-level AOD values for the greater Beijing area. Figure 3.7 shows the MISR retrievals at 17.6 km resolution (upper panel) and our MAP retrievals at 4.4 km resolution (lower panel). The general level of AOD values are not as high and unusual as the previous two cases, but the disagreement in the two sets of AOD estimates is distinctive.

For validation, we resort to the ground measurements. The AERONET measurement at the Beijing station reached 0.2319 on that day; AERONET Xianghe station reported AOD as high as 0.3519. Such medium-AOD situations are relatively common in urban areas with constant air pollution, such as Beijing, due to its heavy industry development and highly concentrated population. It is important to be able to detect such events and further study their long-term impact on public health. Our MAP proves to be more accurate and relevant when it comes to high/medium-AOD situations.

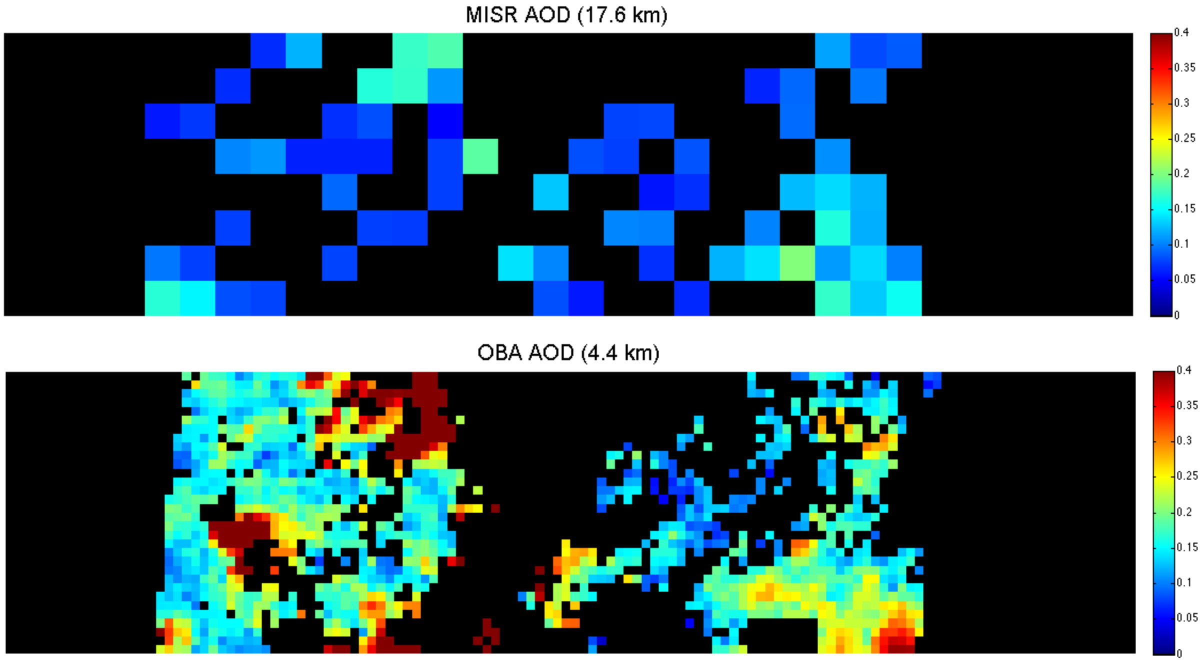


Figure 3.7: Case study of AOD retrievals on January 24, 2009 over the greater Beijing area.

3.5 Discussion

Urban public health and air quality suffer from hazardous aerosols due to heavy industrialization and dense population. Anthropogenic activities, global climate and regional wind patterns continue to shape aerosols into an increasingly heterogeneous system. This growing complexity calls for a flexible model and fine retrieval resolution in order to better understand urban aerosols and their impacts on human health.

In this chapter, we have presented an optimization-based approach to retrieve AOD values and mixing vectors relative to a selection of four component aerosols at 4.4 km resolution using MISR observations. Our model borrows strength from a spatial smoothness structure for AOD values to reduce model complexity, while treating mixing vectors as continuous variables to flexibly represent the mixing of the component aerosols. The free mixing of the selected component aerosols, instead of prefixed mixtures in the MISR retrieval algorithm, proves necessary in case of severely polluted situations, such as dust storms. Spatial profiles of AOD values at finer resolution are provided using our retrieval algorithm, and then validated using ground measurements collected by AERONET. Compared to the MISR operational retrieval algorithm, our algorithm achieves a better retrieval coverage due to flexible mixing of component aerosols. In the case studies, we also observed significantly improved accuracy in high-AOD and medium-AOD situations.

Our optimization approach can also be viewed as a Maximum a Posteriori (MAP) estimation problem in a hierarchical Bayesian scheme. The optimization-based retrieval algorithm (MAP algorithm) is almost 60 times faster than the Markov Chain Monte Carlo sampling algorithm developed based on the same hierarchical Bayesian model. The enhanced computation efficiency further gives room for developing retrievals at higher temporal resolution. Ideally, technology enhancement will improve both spatial and temporal resolution that facilitates in-depth assessments of air pollution and human health. For example, one possible solution exploits data fusion with observed radiances from other instruments, such as the Moderate Resolution Imaging Spectroradiometer (MODIS) aboard the NASA Terra and Aqua satellites.

Another possible continuation of this work is to develop a systematic approach to select four component aerosols involved in the retrievals. Also, a larger number of component aerosols might be necessary for regions with a more varying terrain structure or aerosol emission sources than those in the greater Beijing area. We are currently working on a full expansion of the 74 MISR mixtures by allowing free mixings of all eight commonly observed component aerosols. This will capture a wider range of aerosol mixtures and could potentially improve retrieval accuracy.

We hope that an aerosol profile obtained at fine resolution by an efficient algorithm will eventually expand the potential of MISR observations to assist urban air quality monitoring and public health studies.

Chapter 4

Aerosol Retrievals for Baltimore Area and A Full Inclusion of Component Aerosols

4.1 Motivation

The previous two chapters focused on Aerosol Optical Depth (AOD), quantification of the level of atmospheric aerosol concentration. The current chapter, instead, looks into another critical factor of aerosols' optical properties in determining the reflected solar radiations: the types of aerosols and their mixing percentages. In fact, information of aerosol types and mixing percentages, together with AOD levels, enter the forward radiative transfer simulations as a major input. The simulated results construct the SMART dataset. Both MISR operational and our Bayesian retrieval algorithms rely on comparisons between these simulations and MISR observations.

4.1.1 More on Component Aerosols

To represent and understand the mixing of different types of aerosols, MISR operational retrieval algorithm categorizes aerosol particles into 21 groups based on their properties such as particle size distribution, refractive index and their tendency to absorb water. Each group is referred to as a component aerosol. Eight of these 21 component aerosols represent commonly observed aerosol particles, upon which the radiative transfer simulations and the MISR routine retrievals are based. These eight component aerosols are *Black carbon* (soot), *Carbonaceous* (biomass burning particles), *Mineral dust accumulation 1*, *Mineral dust accumulation 2*, *Mineral dust coarse*, *Sea salt accumulation*, *Sea salt coarse* and *Sulfate mode 1*.

Using these eight commonly observed component aerosols, MISR constructs 74 aerosol mixtures to represent the atmospheric aerosol compositions. Each of the 74 MISR aerosol compositions consists of two or three of the eight component aerosols, accompanied by a fixed vector to indicate the relevant abundance of the components. In fact, the 74 MISR aerosol compositions include only six distinct combinations of the eight component aerosols, mixed according to different settings of their relevant abundance. These six combinations can be represented by the indices of component aerosols involved: $\{1, 6, 19\}$, $\{2, 6, 19\}$, $\{3, 6, 19\}$, $\{6, 8, 19\}$, $\{6, 14, 19\}$ and $\{2, 19, 21\}$. MISR operational retrieval algorithm exhaustively searches over radiative transfer simulations generated according to these 74 MISR compositions, in order to provide good fits to the observations. As the complexity of urban aerosols grows with metropolitan population and industrialization, the 74 MISR compositions are no longer sufficient to represent the aerosol mixing conditions.

In fact, previous case studies based on aerosol conditions in the greater Beijing area of China support the necessity to expand the 74 MISR compositions. MISR reported many missing retrievals for the Beijing area due to the failure to match any of the 74 MISR compositions to the observations. In Chapter 2, we included into our model four component aerosols that are commonly observed for the greater Beijing area. The four components are selected based on historical retrieval results from MISR; we did not include all eight components in order to achieve computational efficiency. Our model then considers the mixing vectors for the four selected component aerosols as continuous variables and thus allows for the maximum freedom for mixing of these components. For example, we discovered that aerosol compositions with a high abundance of coarse mineral dust are common in the greater Beijing area in spring due to regular sandstorms. However, such compositions are not represented by any of the 74 MISR compositions, but is considered by our model because of the free mixing of component aerosols. For such cases, MISR fails to match any of the 74 prefixed compositions to the observations and thus reports no retrievals of AOD nor aerosol composition. Our retrieval algorithm, on the other hand, is still able to provide retrievals of AOD and mixing vectors. Validation of our retrievals by ground-based measurements of AOD levels suggests the four selected component aerosols well represent the aerosol conditions in the greater Beijing area. For regions with more complex types of aerosol emissions, however, more than four component aerosols might be necessary to capture the region's aerosol mixing conditions.

In search of a second study region to validate our retrievals and to further understand aerosol mixings, we turn to Baltimore, Maryland.

4.1.2 Baltimore and the DRAGON Stations

Baltimore is chosen for its complex terrain conditions and various aerosol emission types, as well as the availability of ground-base AOD measurements from a dense network of AErosol RObotic NETwork (AERONET) stations, known as the Distributed Regional Aerosol Gridded Observation Networks (DRAGON).

The DRAGON stations use ground-based sun photometers to collect ground measurements of aerosol properties to validate remote-sensed retrievals at resolution from 50 m to 10 km. It is implemented under the NASA Earth Venture airborne mission, DISCOVER-AQ (Deriving Information on Surface Conditions from Column and Vertically Resolved Observations Relevant to Air Quality). Measurements of AOD levels, as well as size distribution and single scattering albedo among other optical properties, are collected for the Baltimore-Washington DC metropolitan area over the period of June to August, 2011. The entire region covers different types of terrain, including urban, agricultural and mountain landscapes. Figure 4.1 marks the locations of the DRAGON ground stations over the Baltimore-Washington region. The dense and rather gridded distribution of the DRAGON stations provides us an opportunity to extensively validate remote-sensed AOD retrievals. Such validation is especially important in justifying the gains in AOD accuracy by improving the retrieval resolution. Moreover, measurements of particle size distributions can be used as a preliminary validation for the retrieved mixing vectors, as discussed in Section 4.3.

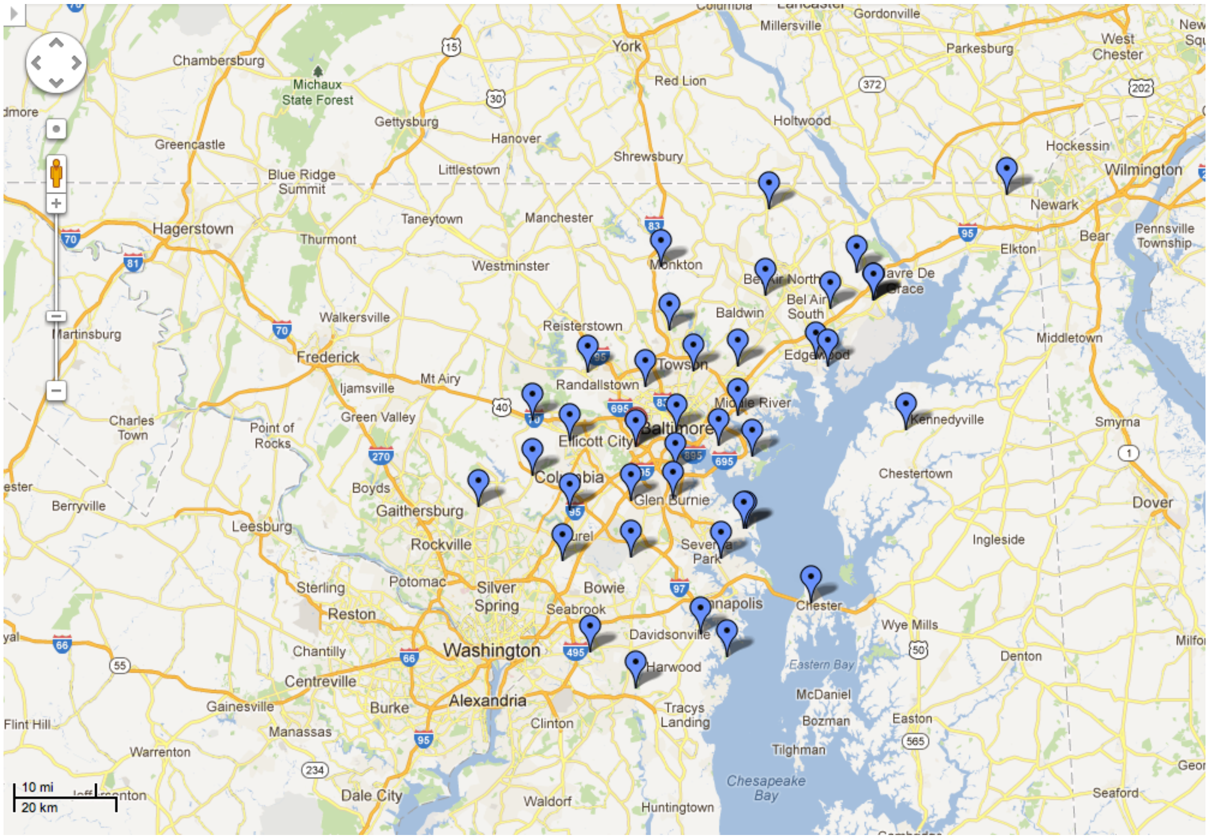


Figure 4.1: Locations of the DRAGON ground stations.

4.1.3 Terrain Structure around Baltimore

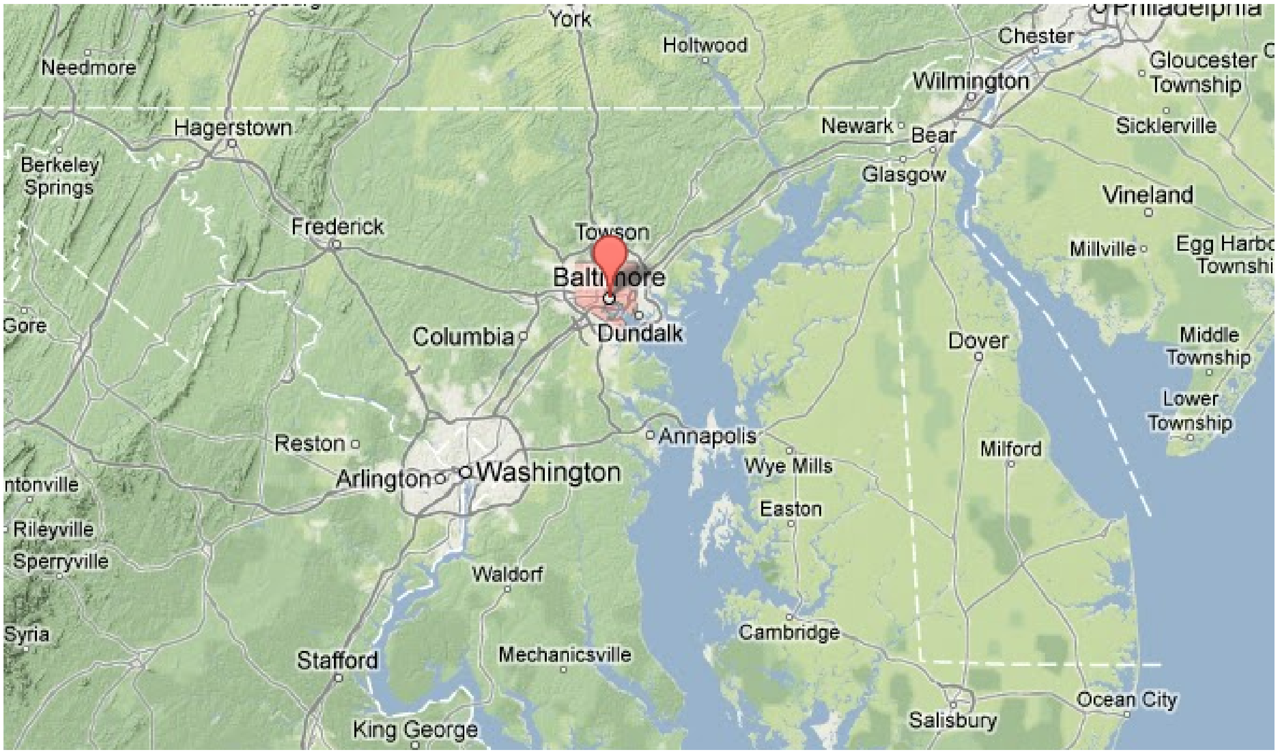


Figure 4.2: Display of terrain in Baltimore area.

The Baltimore-Washington metropolitan area has always been of interest to meteorologists and climatologists due to its reliably overall high levels of aerosol concentration for a large region. Besides the access to densely distributed ground measurements of aerosol optical properties, the complex terrain of Baltimore makes it a unique study region with various landscapes. Figure 4.2 displays the area's terrain conditions, with Baltimore in the center of map. With our final research goal to study the relationship between aerosol concentration and public health, as well as the available DRAGON ground-based measurements, Baltimore naturally become a unique site of great interest to us. It is the case especially in summer when the aerosol loading reaches the highest of the year. The generous precipitation makes it more important to develop timely and accurate aerosol retrievals, as rainfalls reduce the lifespan of atmospheric aerosols.

Another major reason that we set the study region to Baltimore-Washington is that the area has a complex terrain, which has the benefit of a gradient meteorology: mountains, forest, plain, port, river and city. The complexity of the terrain adds to the challenge and significance of the retrievals.

Moreover, the Baltimore metropolitan area has a wide range of different aerosol emissions and thus different types of aerosol particles. This calls into question the assumptions of spatial homogeneity of the selection of component aerosols and makes it interesting to study the mixing of component aerosols in the region. In fact, as we will see in Section 4.2, four component aerosols are not sufficient to represent the complex aerosol mixings in the Baltimore area. Different types of aerosols from this region include fine-mode urban or industrial and biomass burning, coarse-mode dust and sea salt, as well as in mixtures of the various types.

In particular, the following different aerosol emission sources contribute to the complexity of anthropogenic aerosols in the Baltimore-Washington region:

- Thirty-six miles southwest to the city of Baltimore stands the Duke Energy power plant in the Ohio River Valley. Its coal-burning process produces aerosols that contain sulfate.
- Baltimore is surrounded by several interstate highways, including I-70, I-83 (the Jones Falls Expressway), I-95 (the John F. Kennedy Memorial Highway), I-395, I-695 (the Baltimore Beltway), I-795 (the Northwest Expressway), I-895 (the Harbor Tunnel Thruway), and I-97. Baltimore also serves as a top destination for Amtrak trains along the Northeast traffic corridor. Baltimore's Penn Station is one of the busiest in the country. The exhaust fumes from heavy traffic congestion and general tranportaion generates more organic aerosols and cause air pollution.
- Diesel-burning ships and loading trucks at the Port of Baltimore generates more air-polluting aerosols, due to the large scale of the engines. In fact, the Port of Baltimore is one of two seaports on the U.S. east coast with a 15 m dredge to accommodate the largest shipping vessels.
- Aviation is another major source of gaseous and particulate aerosol emissions in the region, especially take-off processes from the Baltimore Washington International (BWI) airport.
- At the same time, the bay breeze effect from the Chesapeake Bay brings inland sea salt particles and adds to the complexity of aerosols mixings, from the southeast in Baltimore, and from the southwest on the Eastern Shore.

Due to the above reasons, AERONET established a mesoscale gridded network of sun photometers in Baltimore-Washington metropolitan area in 2011, namely the DRAGON campaign. DRAGON provide not only ground-based measurements for AOD levels, but also comprehensive measurements of other aerosol optical properties, such as size distribution, single scattering albedo and refractive index (imagery). We use this additional information to validate our retrieved mixing vectors for the eight commonly observed component aerosols, as to be discussed in Section 4.3.

4.1.4 Overview of Chapter 4

In this chapter, we provide aerosol retrievals for the Baltimore-Washington region with a fully flexible model to represent the mixing of component aerosols, based on MISR observed radiances. In our retrieval algorithm, we adopt a hierarchical Bayesian model that constructs a spatial smoothness structure for AOD values over a spatial lattice. We further develop our retrieval algorithm at 4.4 km resolution, using Markov Chain Monte Carlo (MCMC) sampling methods or Maximum a Posteriori (MAP) estimation methods. Aerosol Optical Depth (AOD) retrievals at fine resolution are provided for the Baltimore-Washington region for further studies on the relationship between aerosol concentration and public health. Section 4.2 summarizes the limitation of our previous work caused by large variety of landscapes in the region, as well as the necessity to fully expand the selection of component aerosols to all eight commonly observed components in an effective retrieval model.

The inclusion of all eight commonly observed component aerosols, however, adds to the complexity of the computation. To improve the computation efficiency, we propose a warm start for the Markov sampling algorithm. We first develop two separate Markov chains based on two sub-models, each of which incorporates only four component aerosols. Combining the samples generated by these two separate Markov chains renders a better initialization for the Markov chain incorporating all eight commonly observed component aerosols. Details on subsetting the eight component aerosols and combining the results from the two separate Markov chains are discussed in Section 4.3.2.

The retrieved results are displayed, validated by ground measurements and analyzed in Section 4.3, with future research direction discussed in Section 4.4.

4.2 Limitations of Previous Work

This section explains the limitation of our previously developed retrieval algorithm and the circumstances made limitation impair our retrieval performance. We allow for a wider range of possible values for certain model parameters to expand the selection of component aerosols in our previous algorithm to the complete set of the eight commonly observed components.

4.2.1 Retrieval Results Using the Previous Bayesian Algorithm with Four Components

We first retrieved AOD and aerosol compositions using our previous Bayesian aerosol retrieval algorithm, taking as input the MISR observed radiances for the Baltimore-Washington region. The previous algorithm included only four component aerosols, which are selected based on current knowledge of the study region's aerosol conditions. In particular, we tested the retrieval performance with several different subsets of four component aerosols for the Baltimore-Washington region:

(M1) Components 2, 3, 6 and 19: the collection of component aerosols based on domain knowledge suggested by Prof. Yang Liu at the Rollins School of Public Health at the Emory University;

(M2) Components 2, 3, 8 and 19: the collection of the most commonly observed components for Baltimore-Washington region, based on MISR historical retrievals for the study region;

(M3) Components 2, 6, 8 and 19: the collection of the second most commonly observed components for the study region, based on MISR historical retrievals.

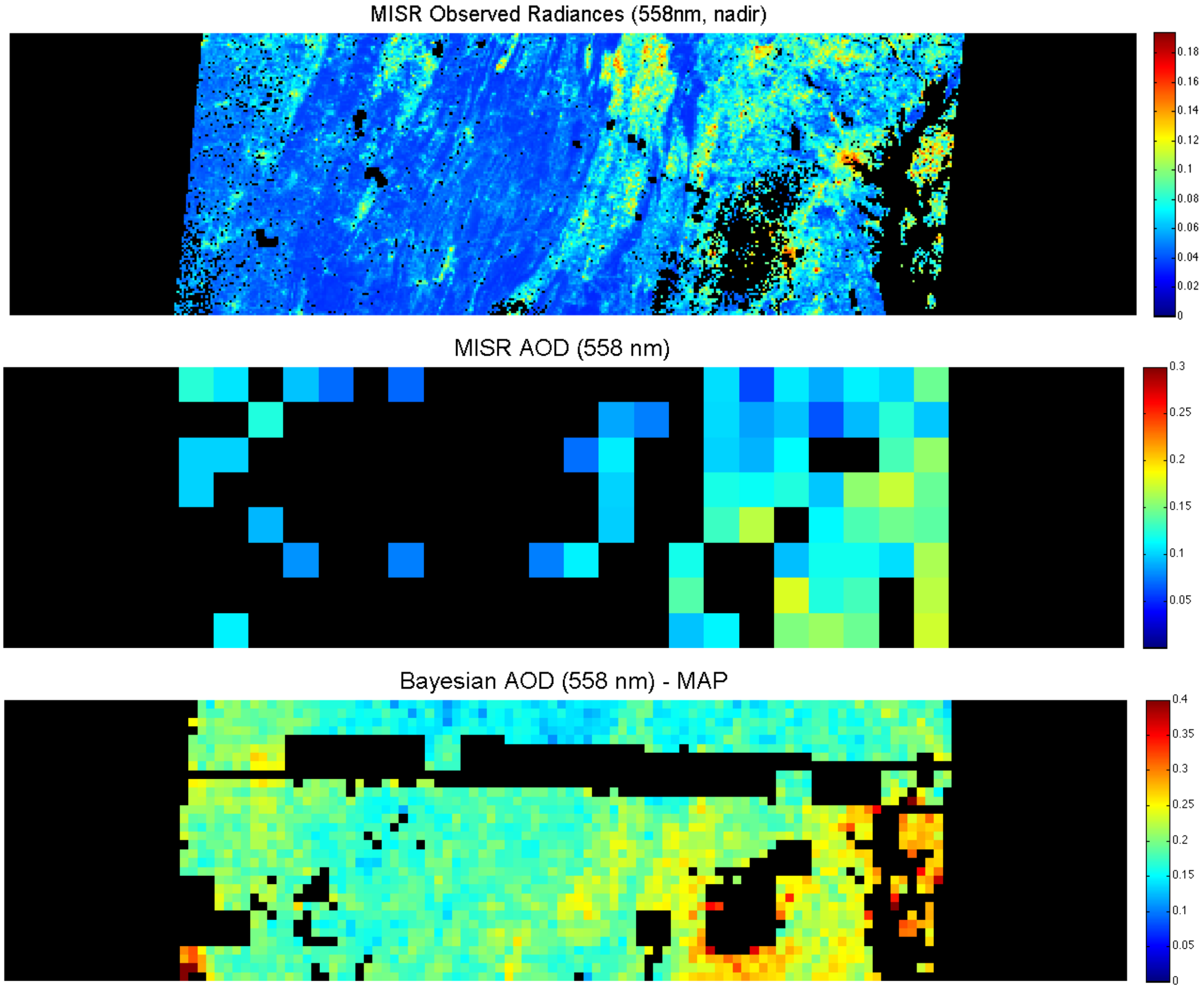
As an example, Figure 4.3 displays the MISR observed radiances (upper panel), MISR AOD retrievals at 17.6 km resolution (middle panel) and ours at 4.4 km resolution (lower panel) for the Baltimore-Washington region on June 2, 2011. MISR AOD estimates were obtained by fitting the 74 MISR prefixed compositions to MISR observations; our retrievals allowed for free mixings of components 2, 3, 8 and 19.

4.2.2 Comparison of MISR Retrievals, Bayesian Retrievals and DRAGON Measurements

Our Bayesian retrievals are reasonable when benchmarked against the MISR retrievals. Our Bayesian and MISR retrievals share a general pattern in AOD spatial distribution: relatively low AOD levels in the north and gradually increasing AOD levels in the southeast corner. Figure 4.3, however, also shows two major differences between MISR and our AOD retrievals, besides the retrieval resolution.

Firstly, our Bayesian retrievals have a better spatial coverage, compared to the MISR retrievals. There are primarily two situations where MISR reports missing retrievals: if clouds are detected or if none of the 74 MISR compositions provide good fits to the observed radiances. In the Baltimore-Washington region on June 2, the latter explanation caused the MISR missing retrievals. In fact, 186 (72.66% of the total 256 pixels and 87.74% of the total missing retrievals) of the 212 missing retrievals (82.82% of the total 256 pixels) in MISR output (Figure 4.3, upper panel) reported that none of the 74 MISR prefixed aerosol compositions could produce simulated radiances that matched the observed radiances. In summary, more than 87% of the missing retrievals were caused by the fact that none of the 74 MISR compositions could well represent the observed aerosol mixing conditions. Our Bayesian retrieval algorithm, on the other hand, is able to significantly improve the retrieval spatial coverage by letting the selected component aerosols vary according to any choice of relative abundance. As a result, we were able to provide estimates for AOD levels as long as the pixel was mostly cloudless.

Secondly, we also observed generally higher AOD levels in our Bayesian retrievals throughout the region, compared to the MISR retrievals. In our previous work, the Maximum a


MISR Observed Radiances (558nm, nadir)

MISR AOD (558 nm)

Bayesian AOD (558 nm) - MAP

Figure 4.3: MISR observed radiances at nadir camera (upper), MISR AOD retrievals at 17.6 km resolution (middle) and our Bayesian retrievals at 4.4 km resolution (lower) over the greater Baltimore area on June 2, 2011.

Posteriori (MAP) AOD retrieval algorithm tends to overestimate AOD during low-AOD situations as is the case for Baltimore-Washington region on June 2.

In order to objectively validate our AOD retrievals for June 2, we compared the results to ground-based measurements collected by DRAGON in Figure 4.4. It compares remote-sensed retrievals by MISR operational algorithm, as well as our MAP and MCMC AOD retrieval algorithms with three different sets of component aerosols. All remote-sensed AOD retrievals are plots on the y-axis against the ground measurements on the x-axis, with the index of involved component aerosols listed in the legend.

Figure 4.4 shows that the MAP retrieval algorithm tends to overestimate AOD levels in low-AOD cases. MISR operational algorithm, on the other hand, tends to underestimate AOD values in medium-AOD situations. We also observe that the selection of component aerosols in the Bayesian retrieval algorithms plays an important role in determining the overall levels of AOD estimates, and hence retrieval accuracy.

4.2.3 Choice of Relevant Component Aerosols and the Necessity to Consider the Complete Set of Eight Commonly-observed Component Aerosols

Our previous MAP and MCMC retrieval algorithms are limited in terms of incorporating all of the eight commonly observed component aerosols. The decision of including only four component aerosols in our previous algorithms results from rather homogenous aerosol conditions in the previous study region, as well as the balance to achieve computational efficiency. In our previous studies in Beijing area, AOD retrievals obtained by our MAP retrieval algorithm agreed with the ground measurements. That is because the aerosol compositions remained relatively homogenous over the area. The region around the city of Baltimore, however, has a very different terrain from the city, leading to largely varying aerosol compositions for this region, as is to be discussed in Section 4.1.3. The Baltimore-Washington cases, therefore, demonstrate necessity to expand the choice of components to the full set of all eight commonly observed components. Our previous algorithms were flexible in letting the mixing vector vary freely, once the subset of four component aerosols is hand-picked. We hence continue to allow the fully flexible mixing of component aerosols in this work.

To confirm the above arguments, we take a closer look at the distribution of each component aerosols in the retrievals over the Baltimore-Washington region. The upper panel in Figure 4.5 uses a range of colors to display the index of MISR-prefixed compositions chosen by MISR retrieval algorithm; the index ranges from 1 to 74. The choice of compositions seem to be indeed largely varying and not quite clustered. It suggests that even though most aerosol mixtures consist of two or three component aerosols, one MISR Block might be too big to maintain the same subset of three or even four component aerosols.

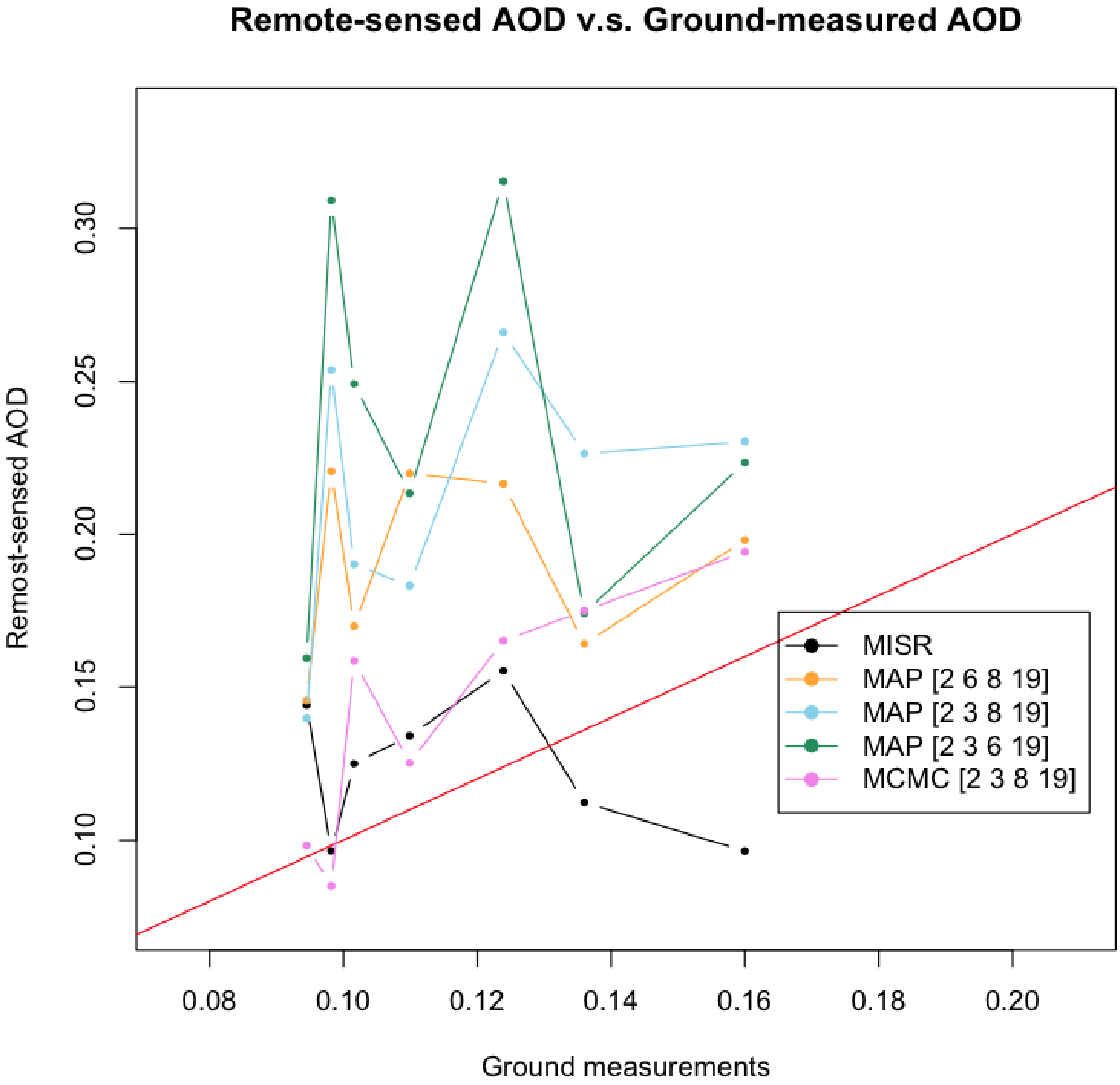


Figure 4.4: Comparisons of AOD retrievals against ground measurements by MISR, MAP and MCMC algorithms with different sets of component aerosols.

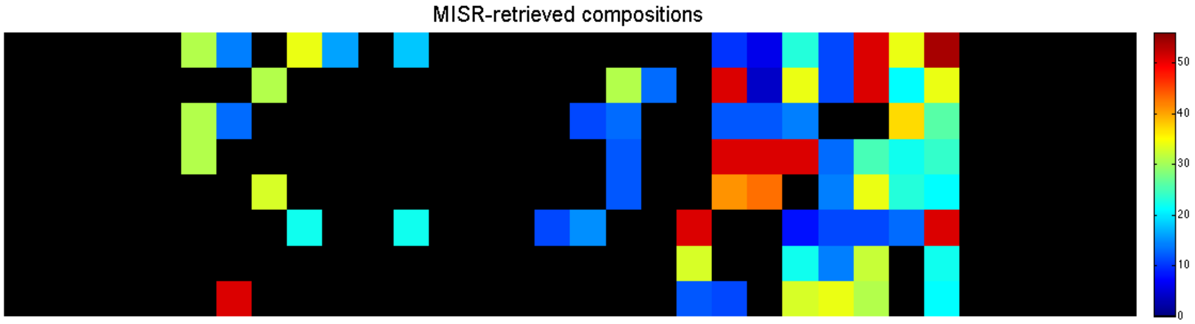


Figure 4.5: Plot of composition index of MISR retrievals (upper) and plot of MISR observed radiances at nadir camera (lower) for Baltimore-Washington region on June 2.

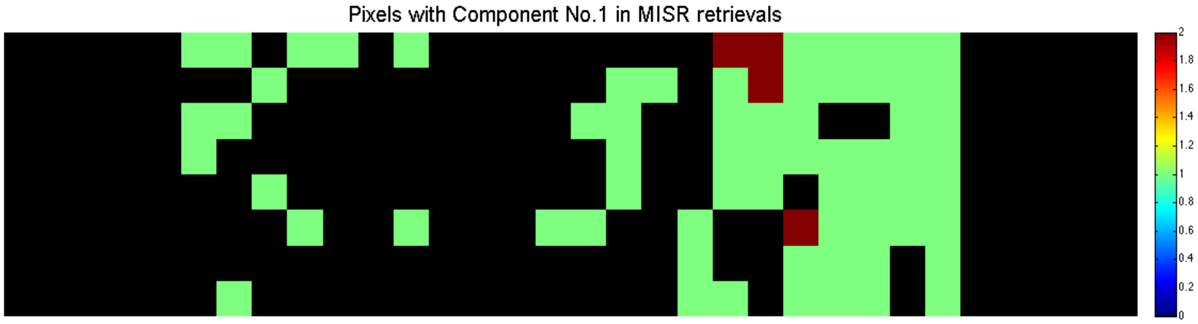
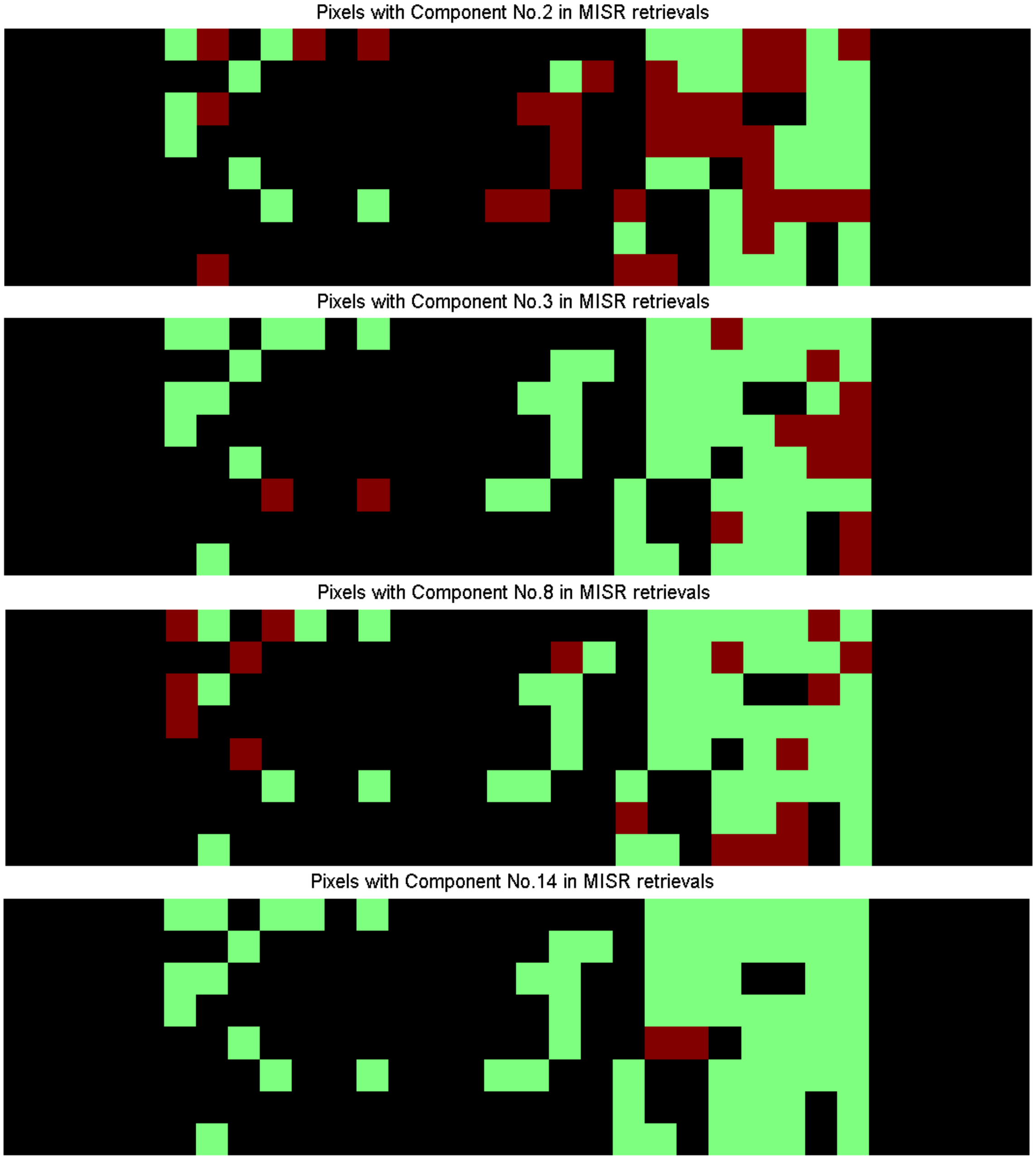


Figure 4.6: Existence of individual component aerosol 1 for Baltimore-Washington region on June 2.



Pixels with Component No.2 in MISR retrievals

Pixels with Component No.3 in MISR retrievals

Pixels with Component No.8 in MISR retrievals

Pixels with Component No.14 in MISR retrievals

Figure 4.7: Existence of individual component aerosols 2, 3, 8 and 14 for Baltimore-Washington region on June 2.

Specifically, Figure 4.6 and Figure 4.7 indicate the existence of each individual component aerosols in the MISR retrievals for the Baltimore-Washington region on June 2. A red pixel indicates appearance of the component aerosol in question in MISR retrieval for that pixel; a green pixel indicates that MISR obtains retrievals for this pixel but reports no appearance of the component in question; black pixels indicate missing retrievals. Component 6 and 19 appeared in all MISR retrieved pixels and component 21 made no appearance in MISR retrievals.

Figure 4.6 and Figure 4.7 confirmed the necessity to incorporate more than four component aerosols for the Baltimore-Washington region. In this case, our original MAP retrieval algorithm was limited and our performance was impaired, due to our fixed subset of component aerosols for the entire MISR Block. The exact existence of different types of component aerosols might result from the different atmospheric environments or ground-level geology of the spatial locations.

The following two sections will briefly summarize background information before we expand the subset of four component aerosols to all of the eight commonly observed components. In particular, Section 4.1.3 reviews the Baltimore-Washington terrain, while Section 4.3.4 introduces some crucial optical properties of the eight commonly observed component aerosols.

4.3 Retrieval Results and Validations using DRAGON Measurements

This section summarizes our retrieval results for the Baltimore-Washington region during June to August, 2011, using our Bayesian MCMC retrieval algorithm and incorporating all eight commonly observed component aerosols.

4.3.1 Retrieving AOD Values

Recall that Section 4.2 discussed the limitations of our previous algorithms and the motivation to remove the model restrictions for the aerosol mixings. All results in this section are generated using Bayesian MCMC aerosol retrieval algorithm which allows for the full free mixing of all eight commonly observed component aerosols.

We match the ground-based DRAGON measurements and the remote-sensed AOD retrievals for the duration of the DRAGON campaign, June to August, 2011. That is, for days when Satellite Terra carrying MISR flies over the Baltimore-Washington region, we locate the pixels with valid AOD retrievals and a DRAGON station inside.

Figure 4.8 displays all such pairs of remote-sensed AOD retrievals and DRAGON measurements. As is shown in the figure, our MCMC retrievals using all of the eight commonly observed component aerosols have a correlation of 0.9512 with the DRAGON measurements, while MISR retrievals have a correlation of 0.7751 with the DRAGON measurement. Out of

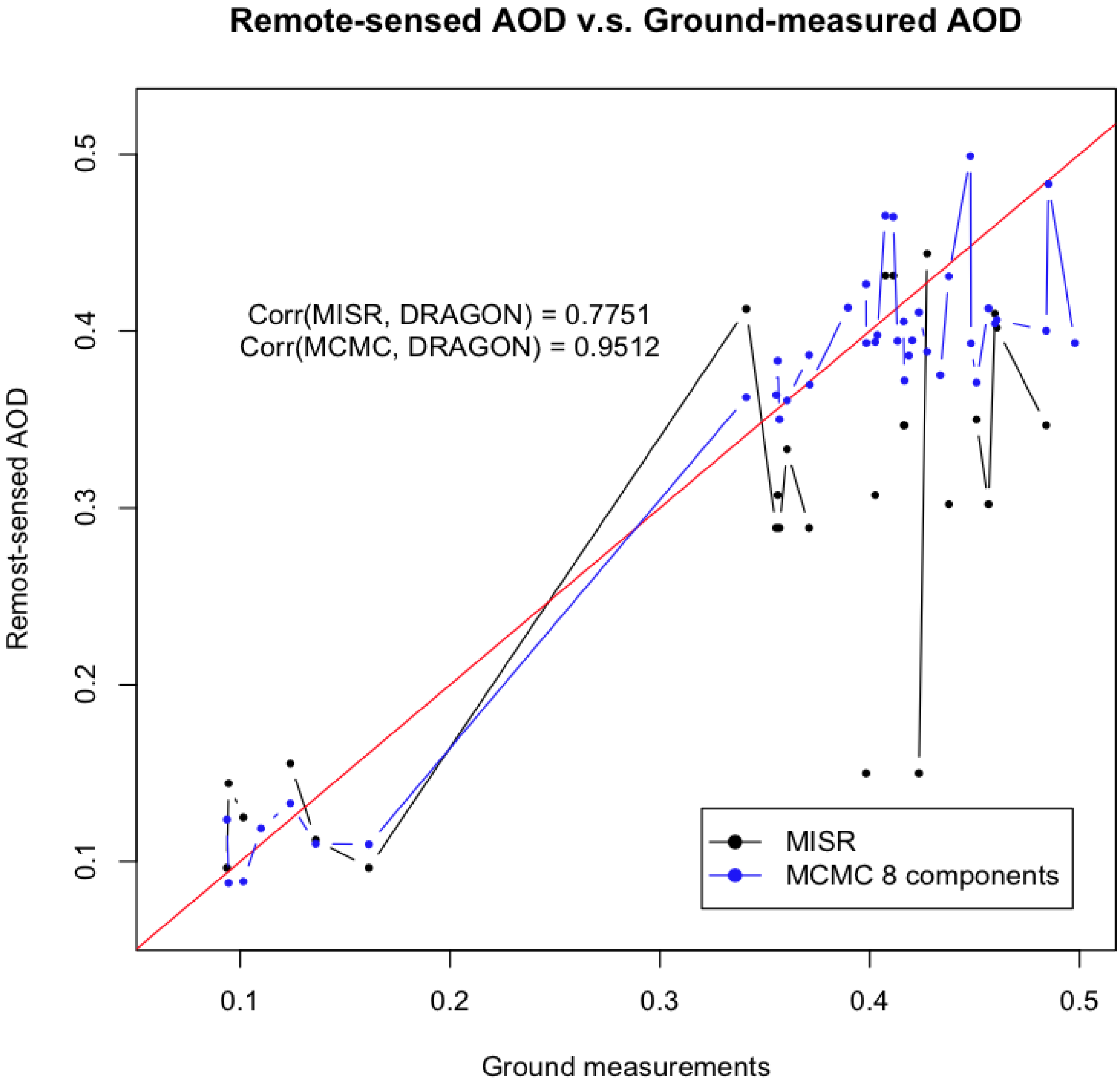


Figure 4.8: Comparisons of AOD retrievals by MISR, our Bayesian MCMC and MAP retrieval algorithms for Baltimore-Washington on June 02, July 20 and July 22 2011.

the 39 validation data pairs, MISR reports 13 cases of missing retrievals, while our MCMC algorithm is able to provide AOD retrievals as long as the pixel is cloudless. We further notice that MISR retrievals tend to underestimate AOD levels in high-AOD events, which is problematic if we want to use these results to provide guidance for policies on public health.

Results above show that expanding the set of component aerosols in consideration to all of the eight commonly observed component aerosols is beneficial for aerosol retrieval accuracy. The increased computation expenses, however, make this expansion less desirable. After confirming the gain in retrieval accuracy from incorporating all eight components, we now work to develop a faster approach to retrieve information on aerosol optical properties. In fact, the Markov chains take more than twice as many steps to converge and the sampling process at each step also consumes more computation. To speed up the mixing of the chains, the next section proposes an approach to search for a warmer initialization for the chains.

4.3.2 Warm-start for MCMC Sampling Using Samples from Subchains

This section explains our motivation and design of searching for a warm start of the Markov chains, as well as demonstrates the consequential computational gain.

We notice that the total run time of retrievals with only four component aerosols is less than a third of that for retrievals with eight component aerosols. It is possible to take advantage of the faster mixing of Markov chains that involve a smaller set of component aerosols and thus a much smaller number of parameters to sample at once. In particular, we split the set of the eight most commonly observed component aerosols into two subsets, each with four components. We first split the set into components $\{1, 3, 8, 19\}$ and components $\{2, 6, 14, 21\}$ to balance the different properties of the components, such as size distribution etc. For example, component 8 and component 14 are not distinguishable in many aspects. We tried to put them into different subsets. Also component 1, 2 and 3 are similar in their optical behaviors with their effective radii increasing from component 1 to component 3. Thus, we place them into two different chains. Later we also tried to split the eight components into two subsets as $\{1, 2, 3, 6\}$ and $\{8, 14, 19, 21\}$. Results are quite similar to those using the first split of the eight components.

We first establish some notations. Denote the first chain with superscript a and the second chain with superscript b. We run the MCMC algorithm on the chain a and obtained samples for: AOD values, τ^a; mixing vectors for components 1, 3, 8 and 19, denoted as $(\theta_1^a, \theta_3^a, \theta_8^a, \theta_{19}^a)$; AOD spatial precision parameter κ^a; the Dirichlet parameter for aerosol mixings, α^a; and variance terms for the 36 MISR channels, $(\sigma^a)^2$. Similarly, we have samples for the second chain: τ^b, κ^b, α^b, $(\sigma^b)^2$; and for the mixing vectors: $(\theta_2^b, \theta_6^b, \theta_{14}^b, \theta_{21}^b)$.

After we have split the eight component aerosols into two subsets, we run the Bayesian MCMC retrieval algorithm based on MISR observations and each of the two subset of components independently. We then combine the samples from both of the Markov chains to generate an initialization point for the chain that incorporates all eight components. We use the average of the two sets of samples as the new starting point of the model parameters and hyperparameters, for the full chain with eight component aerosols:

$$\boldsymbol{\tau}_0 = \frac{\boldsymbol{\tau}^a + \boldsymbol{\tau}^b}{2} \tag{4.1}$$

$$\boldsymbol{\theta}_{k0} = \boldsymbol{\theta}_k^a, \text{if } k = 1, 3, 8 \text{ or } 19 \tag{4.2}$$

$$\boldsymbol{\theta}_{k0} = \boldsymbol{\theta}_k^b, \text{if } k = 2, 6, 14 \text{ or } 21 \tag{4.3}$$

$$\kappa_0 = \frac{\kappa^a + \kappa^b}{2} \tag{4.4}$$

$$\boldsymbol{\alpha}_0 = \frac{\boldsymbol{\alpha}^{\boldsymbol{a}} + \boldsymbol{\alpha}^b}{2} \tag{4.5}$$

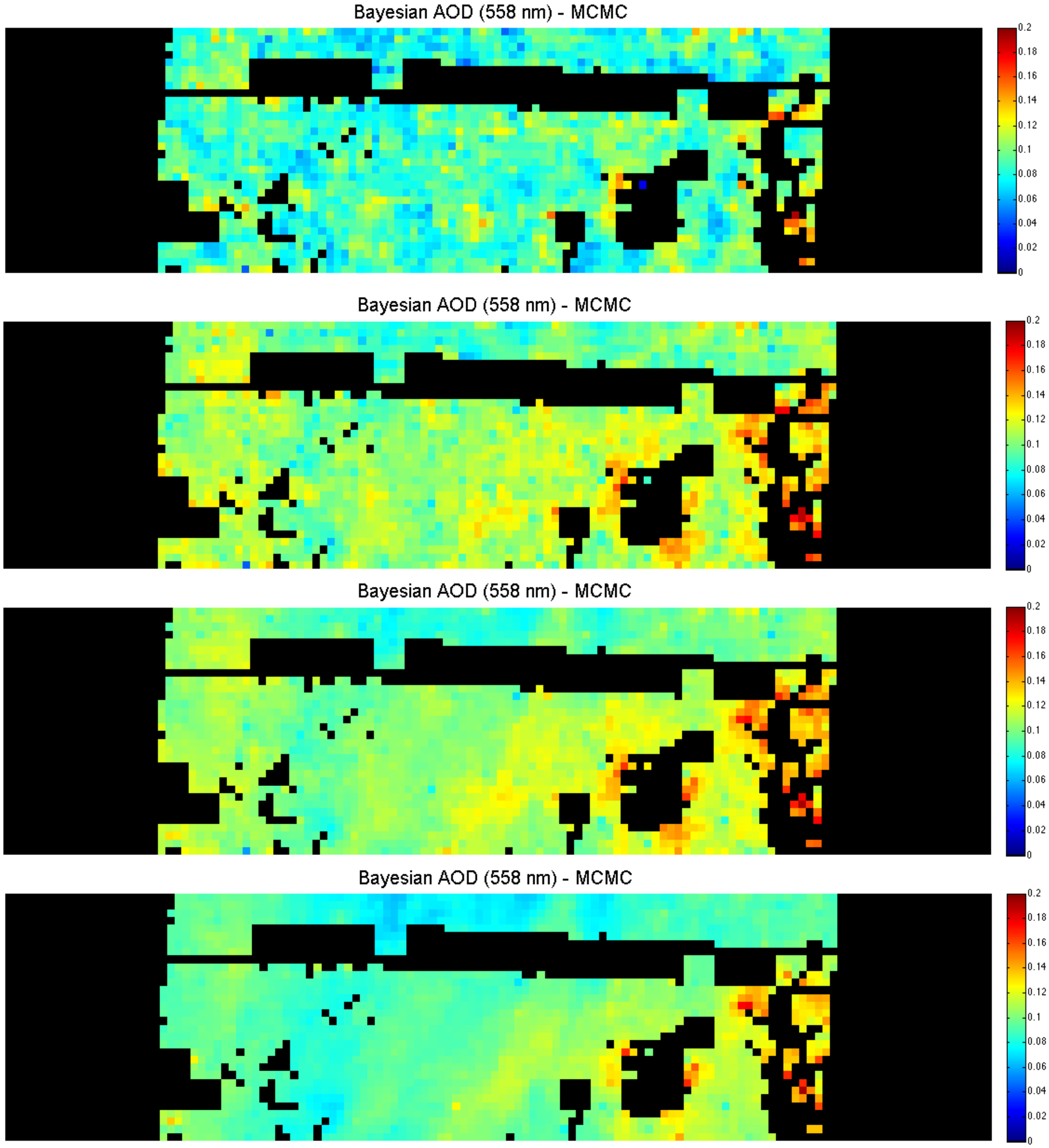
$$(\boldsymbol{\sigma}_0)^2 = (\frac{\sqrt{(\boldsymbol{\sigma}^a)^2} + \sqrt{(\boldsymbol{\sigma}^b)^2}}{2})^2. \tag{4.6}$$

To capture the progress of the Markov chains, we took snapshots of the AOD samples. Each snapshot is obtained by averaging over 100 steps along the chain to reduce unnecessary sampling noise. For example, Figure 4.9 displays four snapshots of the MCMC samples of AOD values over the Baltimore-Washington region for June 2. The snapshots are averaged over MCMC sampling steps 1:100, 301:400, 901:1000 and 1901:2000, generated by our MCMC retrieval algorithm with random initialization.

The four snapshots respectively display the progress of a random start, mixing, developing spatial shapes of the final retrievals and finally smoothing. One distinct shape in the final retrievals is the stripe of relatively high AOD values in the middle of the MISR Block, stretching from lower west to upper east. This stripe is not visible in the first two panels, but already falls into place in the third panel and is quite obvious in the last panel. The second panel already display an overall level similar to the final retrievals, while the first panel is still quite far away from the final results.

To explore the possibility of a warmer initialization than a random start, we ran two subchains as discussed above. Figure 4.10 displays estimates of AOD values based on samples from the two sub-chains. As we can see, both share the general pattern of the last panel in Figure 4.9. Their overall levels of AOD values, however, are very different from that of the last panel, Figure 4.9. In particular, the upper panel of Figure 4.9 allows for component aerosols 1, 3, 8 and 19; while the lower panel allows for component aerosols 2, 6, 14 and 21. In general, component aerosols with a large effective radius tend to be associated with higher level of aerosol concentrations, as is seen in the lower panel of Figure 4.10.

When we combine the samples from the two sub-chains according to (4.3.2), however, the MCMC sampling process seems to have skipped the random start and initial mixing. Figure 4.11 displays three snapshots of MCMC samplings over steps 1:10, 1:100 and 301:400.


Bayesian AOD (558 nm) - MCMC

Bayesian AOD (558 nm) - MCMC

Bayesian AOD (558 nm) - MCMC

Bayesian AOD (558 nm) - MCMC

Figure 4.9: Converging progress of AOD retrievals using Bayesian MCMC algorithm with random start for Baltimore-Washington on June 2, 2011.

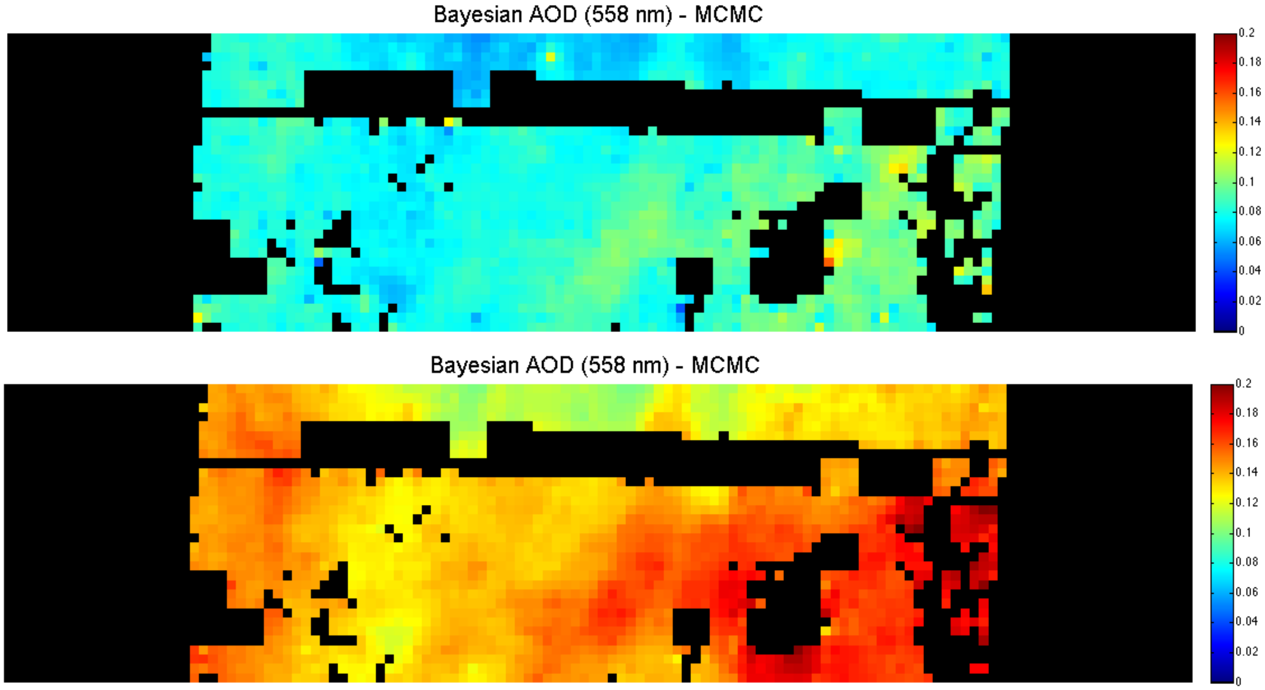


Figure 4.10: Retrieved AOD values by two separate Markov chains each including four components for Baltimore-Washington on June 2.

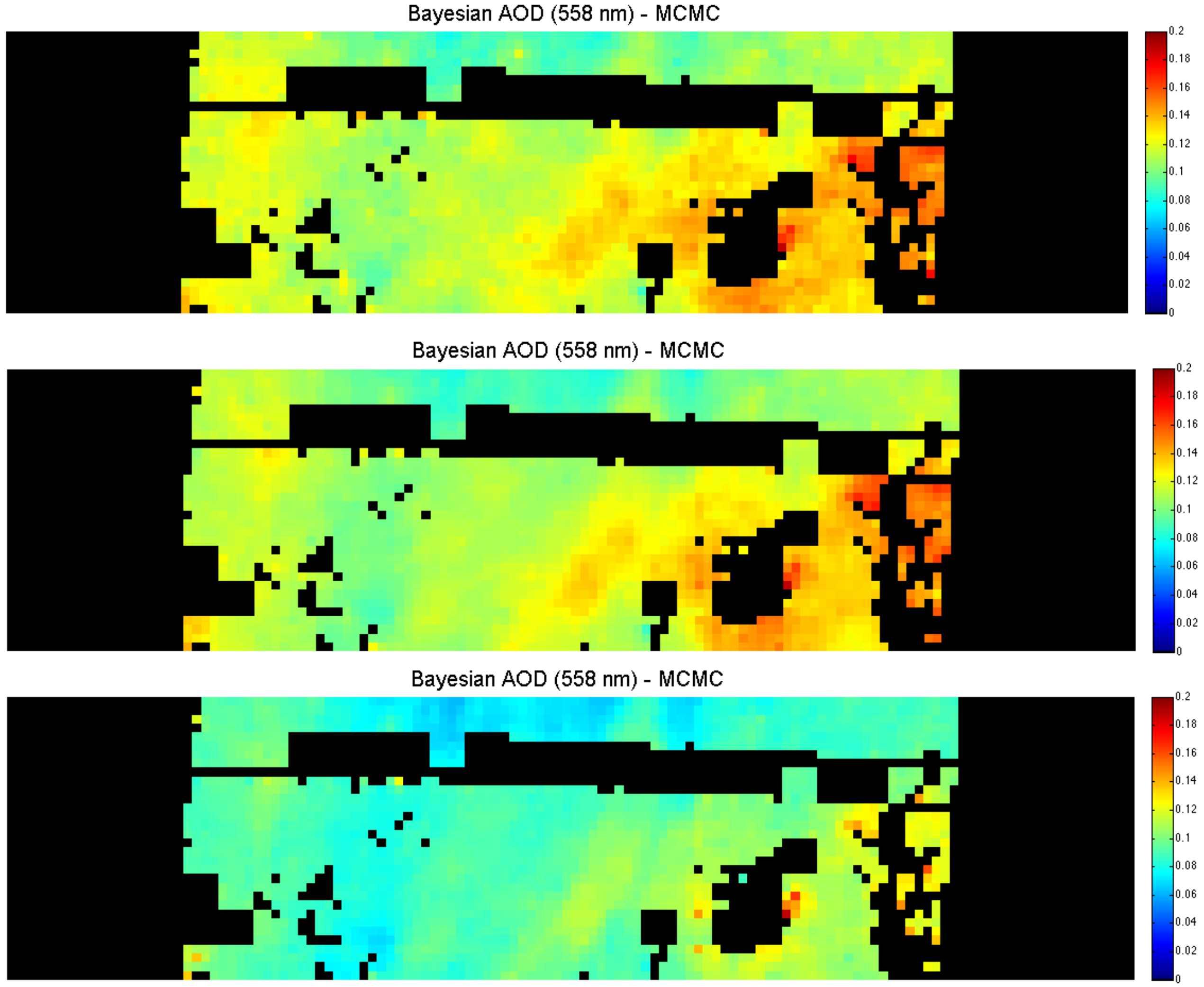


Figure 4.11: Converging progress of AOD retrievals using Bayesian MCMC algorithm with a warm start for Baltimore-Washington on June 2.

To reach the final smooth stage of the AOD estimates, the original MCMC retrieval algorithm with random start took 26251.22 seconds. Sampling along two sub-chains, which can be run independently in parallel, took 7699.35 seconds. After combining the samples from the two sub-chains to reach a warmer start, the new chain took up to 4195.16 seconds to reach the smoothed AOD estimates, as shown in Figure 4.11. Therefore, by running two sub-chains to construct a warm start, we are able to save more than 50% of the computation time. This is partially due to sampling in parallel for the 4-component sub-chains and the reduced number of steps for the 8-component chain to converge.

We also tested an approach to combine samples from the two sub-chains to construct a warm start for the Markov chain with eight component aerosols. Instead of taking the average of the two chains, we calculated MISR's χ^2 criterion and use them as inverse weight when combining the samples from the sub-chains. That is, we assign a heavier weight to components that result in a smaller χ^2 score and thus a better fit to the observations. Let,

$$\begin{aligned} d^a &= \sqrt{(\chi^a)^2} \\ d^b &= \sqrt{(\chi^b)^2}, \end{aligned}$$

and we further have,

$$\begin{aligned} \boldsymbol{\tau}_0 &= \frac{d^b}{d^a + d^b}\boldsymbol{\tau}^a + \frac{d^a}{d^a + d^b}\boldsymbol{\tau}^b \\ \boldsymbol{\theta}_{k0} &= \frac{d^b}{d^a + d^b}\boldsymbol{\theta}^a_k, \text{if } k = 1, 3, 8 \text{ or } 19 \\ \boldsymbol{\theta}_{k0} &= \frac{d^a}{d^a + d^b}\boldsymbol{\theta}^b_k, \text{if } k = 2, 6, 14 \text{ or } 21 \\ \kappa_0 &= \frac{d^b}{d^a + d^b}\kappa^a + \frac{d^a}{d^a + d^b}\kappa^b \\ \boldsymbol{\alpha}_0 &= \frac{\boldsymbol{\alpha}^{\boldsymbol{a}} + \boldsymbol{\alpha}^b}{2} \\ (\boldsymbol{\sigma}_0)^2 &= (\frac{\sqrt{(\boldsymbol{\sigma}^a)^2} + \sqrt{(\boldsymbol{\sigma}^b)^2}}{2})^2. \end{aligned}$$

But the mixing behaviors of the resulted Markov chain are similar to the one started with average values of the two sub-chains. For example, Figure 4.12 displays the MCMC samples over the first 10 steps after the warm start. Both the overall AOD level and the general pattern appear to be quite similar to the first panel in Figure 4.11.

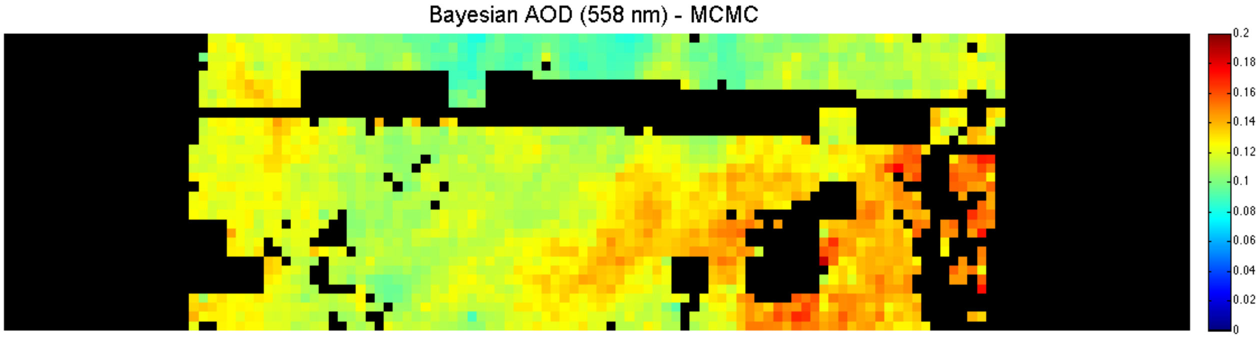


Figure 4.12: AOD retrievals by Bayesian MCMC retrieval algorithm using a warm start for Baltimore-Washington on June 2.

4.3.3 Retrieving Aerosol Mixing Vectors

In this section, we look at the retrieved aerosol mixing vectors. Incorporating all of the eight commonly observed components and letting their mixing vectors vary freely to completely expand the 74 MISR prefixed aerosol compositions.

We first review the basic optical properties of aerosol particles, which are categorized into the 21 component aerosols according to MISR operational retrieval algorithm.

4.3.4 Important Optical Properties of Component Aerosols

Ninety percent of the aerosols are generated by dust storms, desert and soil erosion, biogenic emissions, forest and grassland fires, and sea spray, while ten percent are due to anthropogenic activities. The terrain of Baltimore as shown in Section 4.1.3 suggests a complex mix of aerosols sources in the Baltimore area.

For relevance to our retrieval algorithms, here we only focus on the eight commonly observed component aerosols, namely the component aerosols 1, 2, 3, 6, 8, 14, 19 and 21.

Size Distribution

The Particle Size Distribution (PSD) of aerosol particles is a function indicating the relative abundance of particles present, sorted according to their radius. In particular, the MISR component aerosols are modeled using a log-normal particle size distribution, parameterized by the characteristic radius r_c and characteristic width σ [6]. Then the density function of the PSD can be written as:

$$f(r; r_c, \sigma) = \frac{1}{r\sigma\sqrt{2\pi}} \exp\left(-\frac{(\ln r - r_c)^2}{2\sigma^2}\right). \tag{4.7}$$

Sieve analysis is one of the most common approaches to determine a PSD. The particles are first separated by sieves of different sizes. Then their PSD are approximated by the

Component	1	2	3	6	8	14	19	21
Minimum Radius ($r_{min}, \mu m$)	0.0010	0.0010	0.0100	0.1000	0.0010	0.0010	0.1000	0.1000
Maximum Radius ($r_{max}, \mu m$)	0.4000	0.7500	1.5000	50	0.7500	0.7500	1	6
Effective Radius ($r_c, \mu m$)	0.0562	0.1212	0.2620	2.8009	0.1212	0.1212	0.7542	2.3997
Effective Size Variance (σ^2)	0.0808	0.1034	0.1281	0.2596	0.1034	0.1034	0.0319	0.0510

Table 4.1: Size distribution parameters for the eight commonly observed component aerosols.

relative abundance of particles that "pass" through the sieves with size of a list of discrete ranges.

For a mixture of MISR component aerosols, the density function of its PSD can be calculated as linear combinations of the density functions of each of the component aerosols involved. The characteristic radius and characteristic width for each component aerosol are stored in MISR Aerosol Climatology Product, aerosol physical and optical properties file.

In particular, for the eight commonly observed components, Table 1 displays their parameters for their size distributions.

Single Scattering Albedo

Single scattering albedo (SSA) is the ratio of scattering efficiency of aerosols to their total extinction efficiency, namely the sum of scattering and absorption efficiency.

For a mixture of n component aerosols, denote ω_n the single scattering albedo of the nth component and θ_n its relative abundance in the mixture. The single scattering albedo of the mixture is calculated as:

$$\omega_{mix} = \sum_n \theta_n \omega_n. \tag{4.8}$$

The single scattering albedo of each component aerosol is stored in MISR Aerosol Climatology Product, aerosol physical and optical properties file. The eight commonly observed component aerosols share a same SSA values as 1.

4.3.5 Comparison to MISR Retrieved Mixing Vectors

We first compare the Bayesian MCMC retrieved mixing vectors to those of MISR. Since MISR retrievals are based on 74 prefixed compositions, we first project the 74 aerosol mixtures into an eight-dimensional space. If one of the eight commonly observed component aerosols is absent in the MISR retrieved mixtures, we consider the corresponding mixing percentage as zero.

Figure 4.13 to 4.16 display the mixing percentage of each of the eight component aerosols. As we can see from the results, MISR retrieved mixing vectors disagree largely with our retrieved mixing vectors. The major difference lies in the relative abundance of component 1 and 21. Both components displayed almost zero appearance in MISR retrievals but are quite dominant in our MCMC retrievals, especially component 21. Specifically, 91.58% of the MCMC retrieved pixels have more than 10% component No.21, 70.66% of them have more then 50% component No.21 and 1% of them have more than 70% component No.21.

To validate these retrieved component aerosol mixing percentages, we turn to particle size distribution, collected by the DRAGON and available on AERONET website. For example, we downloaded data for the particle size distribution measured at Station Aldino in the Baltimore-Washington region on June 2, 2011. The upper panel of Figure 4.3.5 plots the corresponding density.

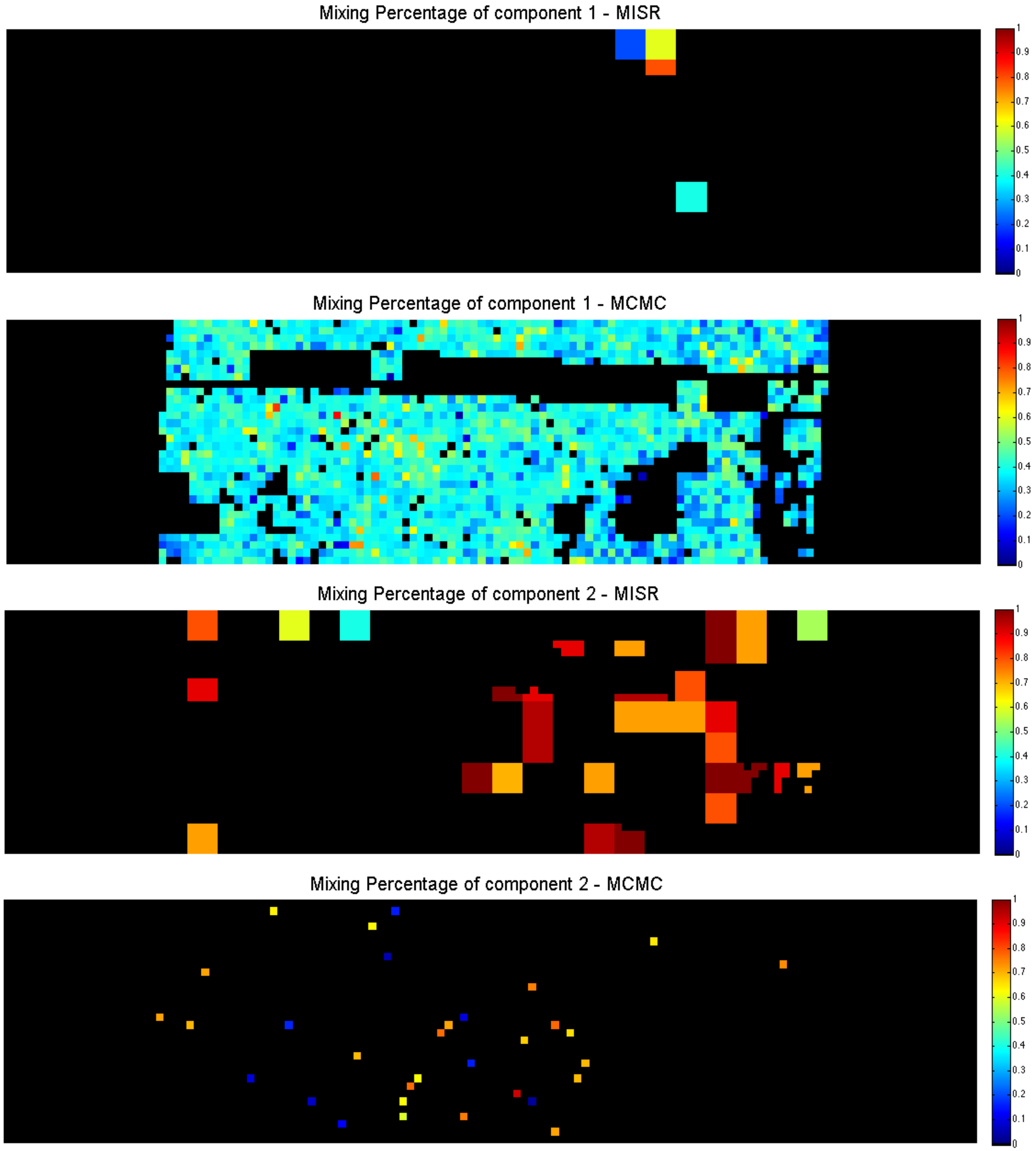
For remote-sensed aerosol compositions, we use mixture density models to combine the particle size distribution of each individual component aerosol and obtain the size distribution for the aerosol mixture. The size distribution of each component aerosol is obtained using the distribution parameters provided in Table 1 (Section 4.3.4). The middle panel of Figure 4.3.5 displays the estimated size distribution retrieved by MISR operational algorithm. It rather accurately captures the first high mode of the size distribution, indicating a high concentration of small-radius aerosols. The DRAGON measured Particle Size Distribution, however, shows another mode for particles with radius around 5 μm. Our Bayesian MCMC retrieved mixing vectors lead to a particle size distribution with a mode near zero and another mode with radius around 6 μm. This further confirms that our retrieved mixing percentage for component 21 is reasonably high, though this particular component aerosol is not commonly observed in urban areas. One possible explanation is that there could be fugitive dust particles or small dusts coated with nitrogen, forming aerosols with a large effective radius like coarse aerosols.

Though this comparison between retrieved particle size distribution and ground-measured particle size distribution does not agree to every detail as seen in Figure 4.3.5, it can already provide some guidance on validating the remote-sensed aerosol compositions.

4.4 Discussion

To obtain a warm start for the MCMC sampling algorithm, instead of using two exclusive subsets of the eight commonly observed components (Section 4.3.2), we could also consider using overlapping subsets of component aerosols. It could potentially take advantage of the relative abundance of components within each subset represented by the samples from each sub-chain. Moreover, splitting the eight commonly observed components into more than two subsets may further reduce the computation time due to efficient parallel sampling for the four separate Markov chains. For instance, four subsets of two components each.

Another direction for future research is to further refine the retrieval resolution to 1.1 km. We implemented our Bayesian retrieval algorithm at even finer resolution to test the potential reliability of retrievals at 1.1 km resolution.

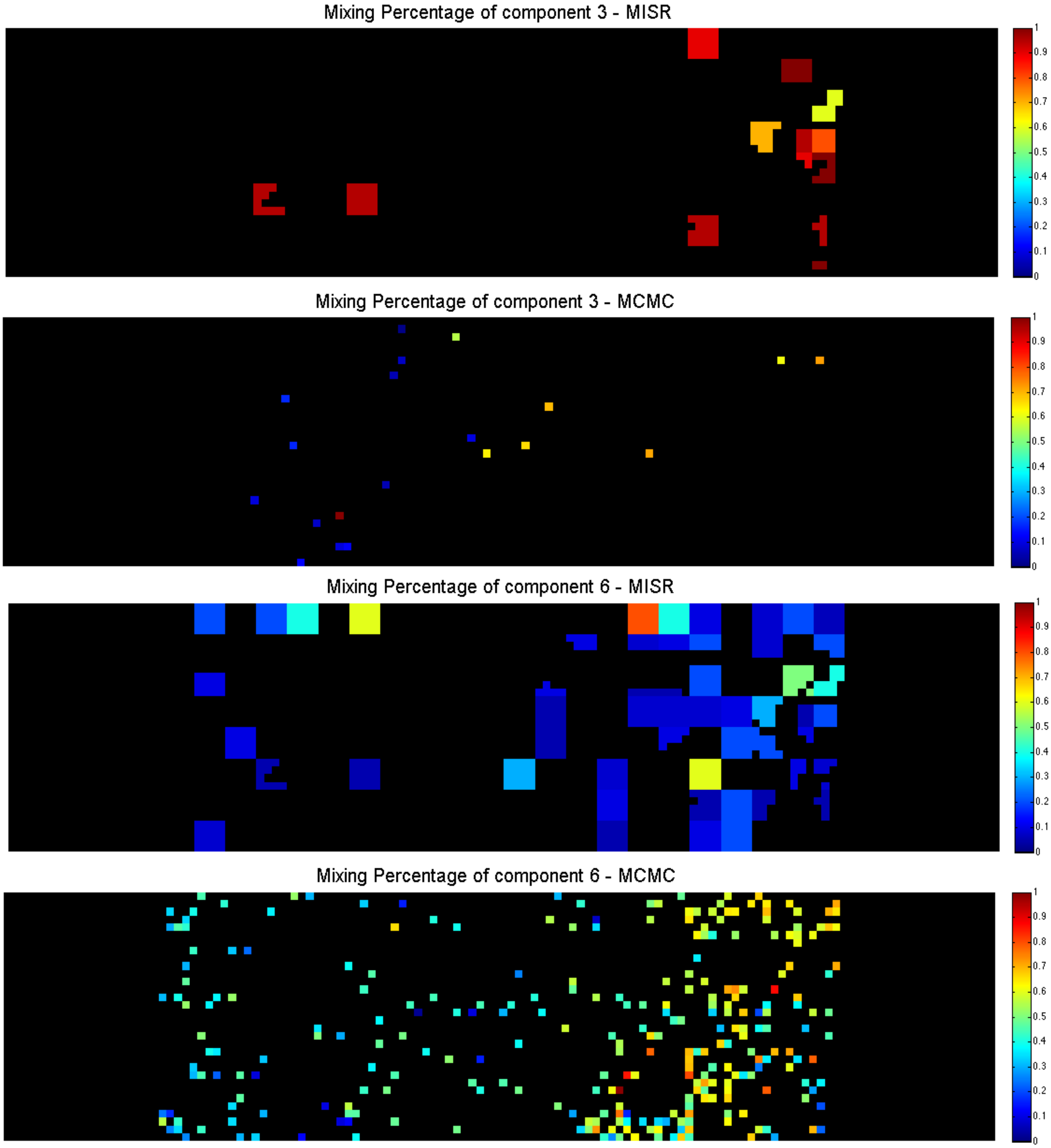

Mixing Percentage of component 1 - MISR

Mixing Percentage of component 1 - MCMC

Mixing Percentage of component 2 - MISR

Mixing Percentage of component 2 - MCMC

Figure 4.13: Comparisons of retrieved mixing percentage for component aerosols 1 and 2 on June 9, 2011.



Mixing Percentage of component 3 - MISR

Mixing Percentage of component 3 - MCMC

Mixing Percentage of component 6 - MISR

Mixing Percentage of component 6 - MCMC

Figure 4.14: Comparisons of retrieved mixing percentage for component aerosols 3 and 6 on June 9, 2011.

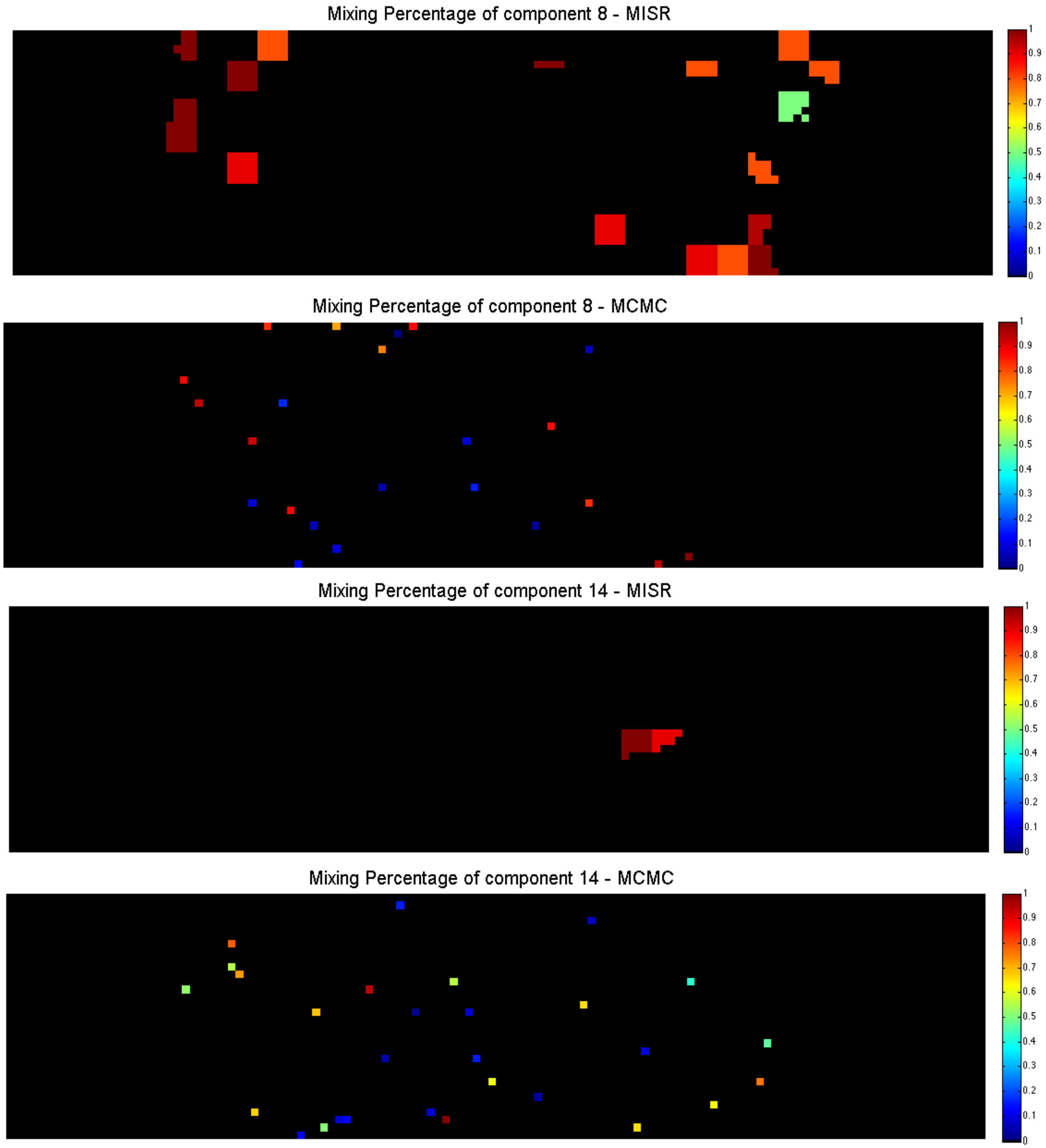


Figure 4.15: Comparisons of retrieved mixing percentage for component aerosols 8 and 14 on June 9, 2011.

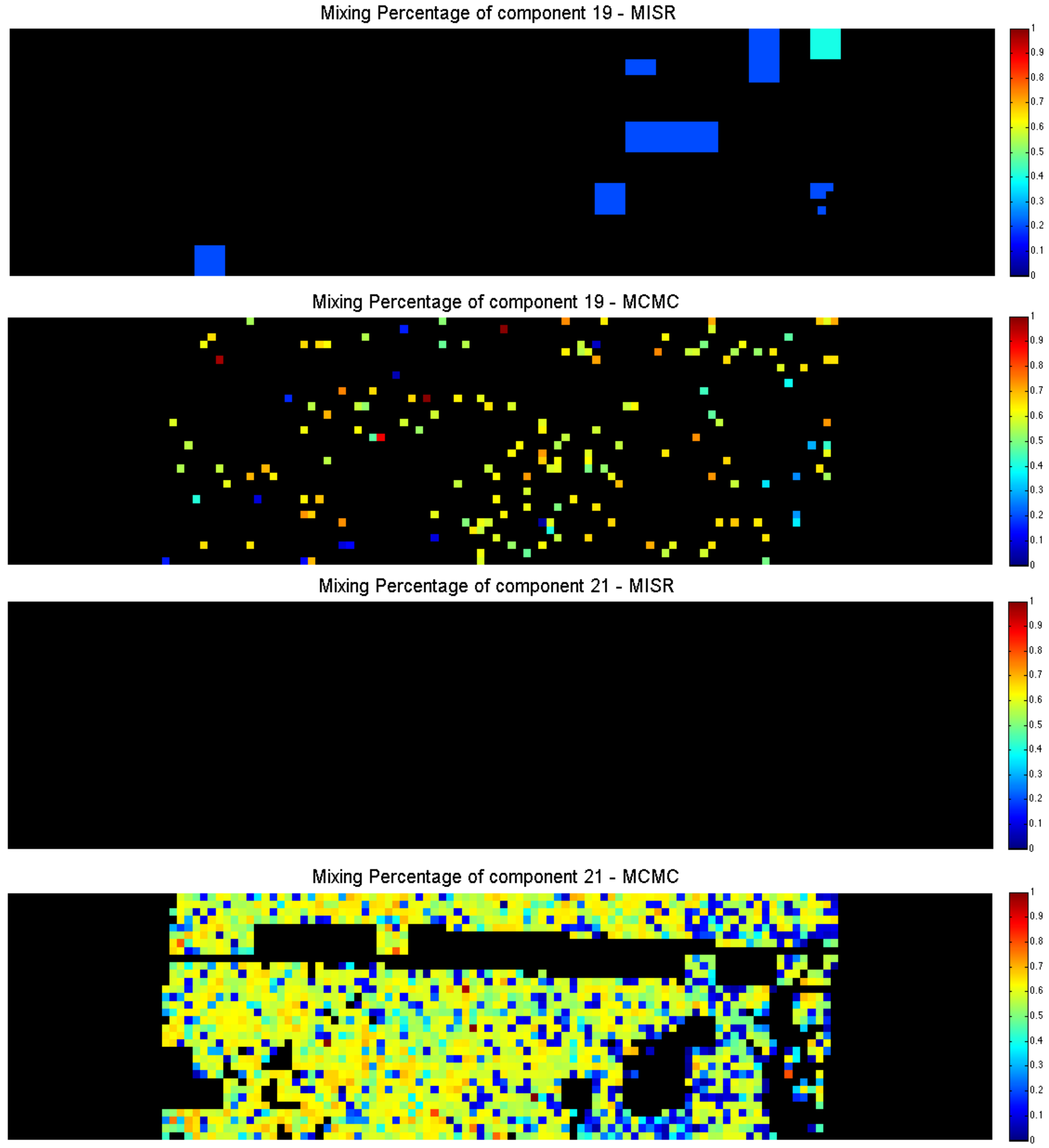
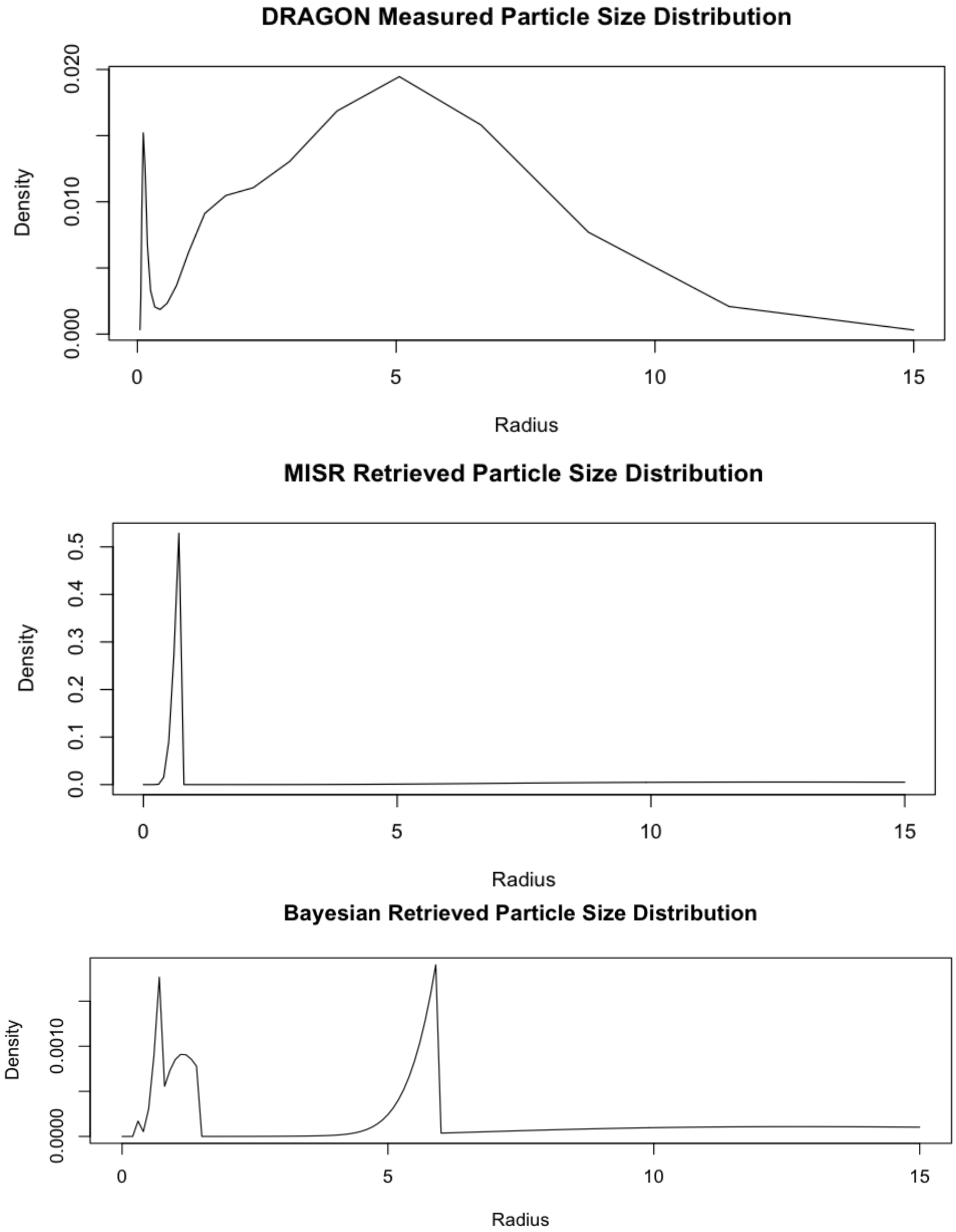


Figure 4.16: Comparisons of retrieved mixing percentage for component aerosols 19 and 21 on June 9, 2011.



DRAGON Measured Particle Size Distribution

MISR Retrieved Particle Size Distribution

Bayesian Retrieved Particle Size Distribution

Figure 4.17: Comparisons of DRAGON-measured and remote-sensed retrievals of size distributions for Baltimore-Washington on June 2, 2011.

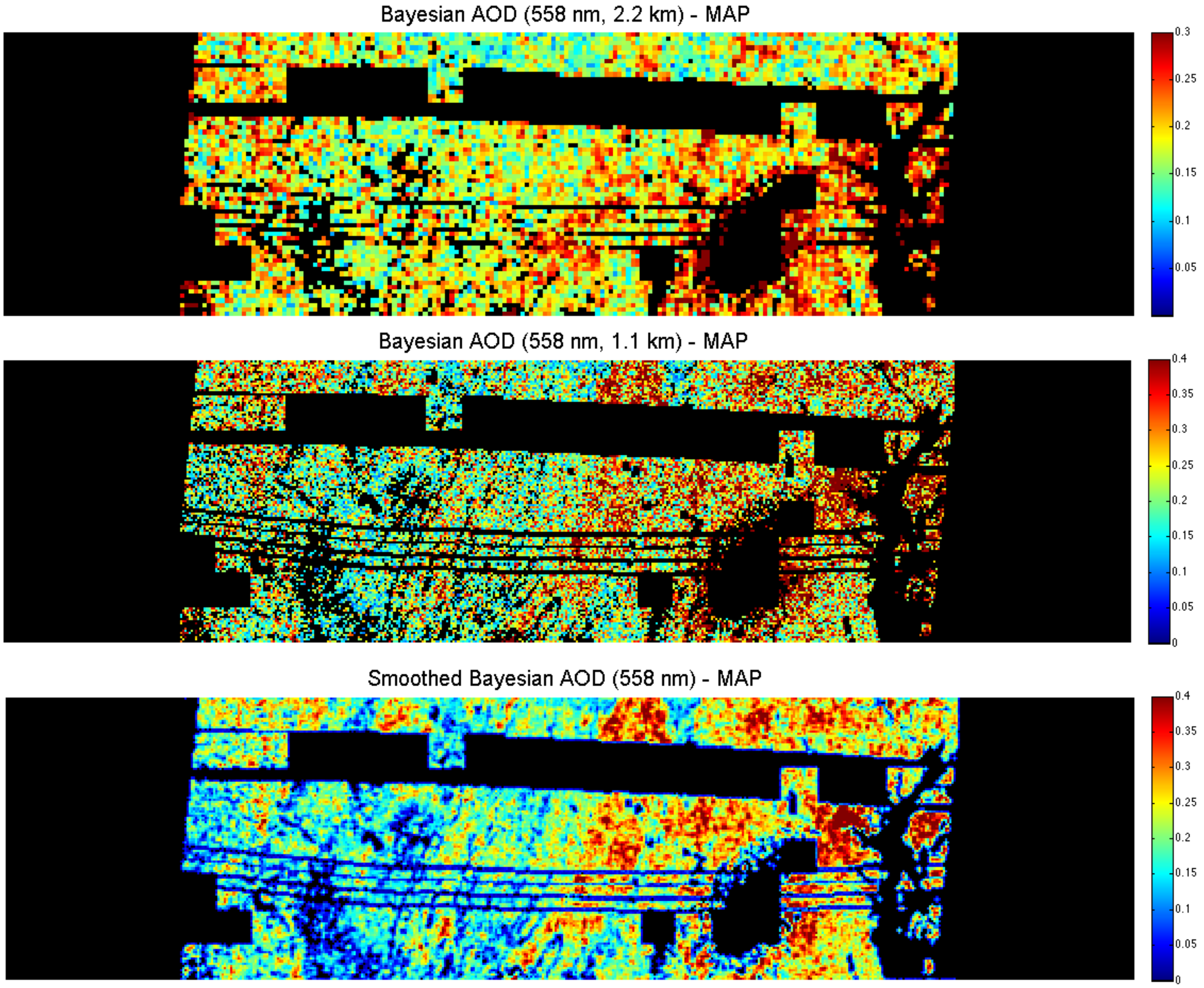


Figure 4.18: AOD retrievals over the greater Baltimore area on June 2, 2011.

Bibliography

[1] D. Bergametti, A. L. Dutot, P. BUAT-MENARD, R. LOSNO, and E. REMOUDAKI. Seasonal variability of the elemental composition of atmospheric aerosol particles over the northwestern mediterranean. *Tellus B*, 41B(3):353–361, 1989.

[2] Julian Besag. Spatial interaction and the statistical analysis of lattice systems. *Journal of the Royal Statistical Society*, B:192–236, 1974.

[3] Mian Chin, Ralph A. Kahn, and Stephen E. Schwartz. Atmospheric aerosol properties and climate impacts: A report by the U.S. climate change science program and the subcommittee on global change research. *National Aeronautics and Space Administration*, 2009.

[4] J. A. De Gouw, A. M. Middlebrook, C. Warneke, R. Ahmadov, R. Atlas, E.L.and Bahreini, D. R. Blake, C. A. Brock, J. Brioude, D. W. Fahey, F. C. Fensenfeld, J. S. Holloway, M. Le Henaff, R. A. Lueb, Mckeen S. A., J. F. Meagher, D. M. Murphy, C. Paris, D. D. Parrish, A. E. Perring, I. B. Pollack, A. R. Ravishankara, A. L. Robinson, T. B. Ryerson, J. P. Schwarz, Spackman J. R., A. Srinivasan, and L. A. Watts. Organic aerosol formation downwind from the deepwater horizon oil spill. *Science*, 331(6022):1295–1299, 2011.

[5] David J. Diner, Wedad A. Abdou, Thomas P. Ackerman, Kathleen Crean, Howard R. Gordon, et al. Level 2 aerosol retrieval algorithm theoretical basis. *JPL*, D(11400), 2008.

[6] David J. Diner, Wedad A. Abdou, Howard R. Gordon, Ralph A. Kahn, Yuri Knyazikhin, John V. Martonchik, Duncan McDonald, Stuart McMuldroch, Ranga Myneni, and Robert A. West. Level 2 ancillary products and datasets algorithm theoretical basis. *JPL*, D(13402), 1999.

[7] David J. Diner, Jewel C. Beckert, Terrence H. Reilly, Carol J. Bruegge, James E. Conel, Ralph A. Kahn, and John V. Martonchik. Multi-angle imaging spectroradiometer (MISR) - instrument description and experiment overview. *IEEE Transactions on Geoscience and Remote Sensing*, 36(4):1072–1087, 1998.

[8] J. C. Diner, D. J. Beckert, G. W. Bothwell, and J. I. Rodriguez. Performance of the MISR instrument during its first 20 months in earth orbit. *IEEE Transactions on Geoscience and Remote Sensing*, 40(7):1449–1466, 2002.

[9] Aaron van Donkelaar, Randall V. Martin, Michael Brauer, Ralph Kahn, Robert Levy, Carolyn Verduzco, and Paul J. Villeneuve. Global estimates of ambient fine particulate matter concentrations from satellite-based aerosol optical depth: development and application. *Environmental Health Perspectives*, 118(6):847–855, 2010.

[10] A. Gelman, G. O. Roberts, and W. R. Gilks. Efficient metropolis jumping rules. *Bayesian Statistics*, 5:599–607, 1996.

[11] Andrew Gelman. Prior distributions for variance parameters in hierarchical models. *Bayesian Analysis*, 1(3):515–533, 2006.

[12] Andrew Gelman and Donald B. Rubin. Inference from iterative simulation using multiple sequences. *Statistical Science*, 7(4):457–472, 1992.

[13] S. Geman and D. Geman. Stochastic relaxation, Gibbs distributions and the bayesian restoration of images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 6(6):721–741, 1984.

[14] David M. Giles. http://aeronet.gsfc.nasa.gov/, 2011.

[15] B. N. Holben, T. F. Eck, Slutsker I., D. Tanré, J. P. Buis, Setzer A., Vermote E., et al. Aeronet-a federated instrument network and data archive for aerosol characterization. *Remote Sens. Environ.*, 66(1):1–16, 1998.

[16] Xin Jiang, Yang Liu, Bin Yu, and Ming Jiang. Comparison of MISR aerosol optical thickness with aeronet measurements in Beijing metropolitan area. *Remote Sensing of Environment*, 107:45–53, 2007.

[17] Ralph A. Kahn, Yang Chen, David L. Nelson, Fok-Yan Leung, Qinbin Li, David J. Diner, and Jennifer A. Logan. Wild smoke injection heights: Two perspectives from space. *Geophysical Research Letters*, 35:L04809, 2008.

[18] Ralph A. Kahn, Barbara J. Gaitley, Michael J. Garay, David J. Diner, Thomas F. Eck, Alexander Smirnov, and Brent N. Holben. Multiangle Imaging SpectroRadiometer global aerosol product assessment by comparison with the Aerosol Robotic Network. *Journal of Geophysical Research*, 115:D23209, 2010.

[19] Ralph A. Kahn, W.-H. Li, Catherine Moroney, David J. Diner, John V. Martonchik, and Evan Fishbein. Aerosol source plume physical characteristics from space-based multiangle imaging. *Journal of Geophysical Research*, 112:D11205, 2007.

[20] Michael D. King, Yoram J. Kaufman, Didier Tanré, and Teruyuki Nakajima. Remote sensing of tropospheric aerosols from space: Past, present, and future. *Bulletin of the American Meteorological Society*, 80(11):2229–2259, 1999.

[21] C. D. Koven and I. Fung. Identifying global dust source areas using high-resolution land surface form. *Journal of Geophysical Research Atmospheres*, 113:D22204, 2008.

[22] Naresh Kumar, Allen Chu, and Andrew Foster. An empirical relationship between $PM_{2.5}$ and aerosol optical depth in Delhi Metropolitan. *Atmospheric Environment*, 41:4492–4503, 2007.

[23] Kuo-nan Liou. An introduction to atmospheric radiation. (84), 2002.

[24] Marjorie Monleau, Cyrill Bussy, Philippe Lestaevel, Pascale Houpert, François Paquet, and Valérie Chazel. Bioaccumulation and behavioral effects of depleted uranium in rats exposed to repeated inhalations. *Neuroscience Letters*, 390:31–36, 2005.

[25] Jan-Peter Muller, Athula Mandanayake, Catherine Moroney, Roger Davies, and David J. Diner. MISR Stereoscopic Image Matchers: Techniques and Results. *IEEE Transactions on Geoscience and Remote Sensing*, 40(7):1547–1559, 2002.

[26] The National Aeronautics and Space Administration (NASA). Aerosol Robotic Network (AERONET) Homepage, 2012. Online; accessed 10-August-2012. Available at: http://aeronet.gsfc.nasa.gov/.

[27] J. T. Peterson, E. C. Flowers, G. J. Beni, C. L. Reynolds, and J. H. Rudisill. Atmospheric turbidity over central North Caroline. *Journal of Applied Meteorology*, 20:229, 1982.

[28] C. Arden Pope III, Richard T. Burnett, Michael J. Thun, Eugenia E. Calle, Daniel Krewski, Kazuhiko Ito, and George D. Thurston. Lung cancer, cardiopulmonary mortality, and long-term exposure to fine particulate air pollution. *Journal of the American Medical Association*, 287(9):1132–1141, 2002.

[29] Ulrich Pöschl. Atmospheric aerosols: Composition, transformation, climate and health effects. *Atmospheric Chemistry*, 44:7520–7540, 2005.

[30] Havard Rue and Leonhard Held. Gaussian Markov random fields: Theory and applications. 2005.

[31] G. E. Shaw. Atmospheric turbidity in the polar regions. *Journal of Applied Meteorology*, 21:1080, 1982.

[32] F. Solmon, F. Giorgi, and C. Liousse. Aerosol modelling for regional climate studies: application to anthropogenic particles and evaluation over a European/African domain. *Tellus*, 58B:51–72, 2006.

[33] Andrew J. Tatem, Scott J. Goetz, and Simon I. Hay. Terra and Aqua: new data for epidemiology and public health. *International Journal of Applied Earth Observation and Geoinformation*, 6:33–46, 2004.

[34] Sally Tinkle, Mary Gant, Michael Humble, Gary Foley, Valerie Garcia, and Andy Bond. Integrated Earth Observations: Application to air quality and human health. *United States Environmental Protection Agency Research and Development*, 2007.

[35] Yueqing Wang, Xin Jiang, Ming Jiang, and Bin Yu. A Hierarchical Bayesian Approach for Aerosol Retrieval Using MISR Data. 2012.

[36] John W. Winchester and Gordon D. Nifong. Water pollution in lake Michigan by trace elements from pollution aerosol fallout. *Water, Air, and Soil Pollution*, 1(1):50–64, 1971.