UNIVERSITY OF OKLAHOMA

GRADUATE COLLEGE

DEVELOPMENT OF A CORRELATION-BASED INVERSION METHOD FOR AEROSOL REMOTE SENSING

A DISSERTATION

SUBMITTED TO THE GRADUATE FACULTY

in partial fulfillment of the requirements for the

Degree of

DOCTOR OF PHILOSOPHY

By

TAOZHONG HUANG
Norman, Oklahoma
2026

DEVELOPMENT OF A CORRELATION-BASED INVERSION METHOD FOR AEROSOL REMOTE SENSING

A DISSERTATION APPROVED FOR THE
SCHOOL OF METEOROLOGY

BY THE COMMITTEE CONSISTING OF

Dr. Feng Xu, Committee Chair

Dr. Lan Gao

Dr. Jens Redemann

Dr. Chenghao Wang

Dr. Binbin Weng

ACKNOWLEDGMENTS

I would like to appreciate my advisor, Dr. Feng Xu, for his guidance and help throughout the five-year period of my PhD study. His mentorship shaped not only the direction of this research but also my growth as a scientist, and I am grateful for the countless hours he dedicated to discussing ideas, reviewing drafts, and pushing me to think more rigorously about every aspect of this work.

I am also thankful to the members of my dissertation committee, Dr. Lan Gao, Dr. Jens Redemann, Dr. Chenghao Wang, and Dr. Binbin Wen, for their valuable feedback, constructive criticism, and encouragement at each stage of this project. Their patient guidance not only strengthened this work considerably but also gave me the encouragement to overcome difficulties.

Special appreciation to Dr. Yixin Wen, who guided me in my undergraduate research and brought me to the field.

To my colleagues in the ARRS group, Benting Chen, Wenzhi Zhang, and Congcong Qiao, and my friends in Norman, as well as my colleagues, senior scientists in the School of Meteorology, thank you for the discussions, for sharing your experience, and for your suggestions on my research. Our conversations made this journey enjoyable, and I am especially grateful for the friendships formed during this time.

I would like to thank all of my family and friends who have helped and encouraged me throughout this journey. This dissertation is as much a reflection of your support as it is of my own effort. Your support gave me the courage to face every challenge along the way. Special thanks to my father and mother, who never gave up and have been wonderful role models for how I approach life and interact with others, which has strengthened my spirit. Many thanks to my family in

Norman, who provided help whenever I needed it and made me feel at home while studying abroad. Heartfelt thanks to my fiancée, who, as a fellow PhD candidate, has shown an exemplary attitude toward research and has guided me when I didn't know what to do. I hope everything goes well for her, and I wish her all the best.

ABSTRACT

Accurate characterization of aerosol microphysical properties from diverse remote sensing platforms is essential for quantifying aerosol-radiation and aerosol-cloud interactions. However, aerosol retrieval remains a highly ill-posed inversion problem, particularly when operating under limited information content. To address this challenge, this dissertation introduces two distinct applications that improve state-of-the-art retrieval techniques by incorporating well-characterized a priori constraints.

The first application develops a surface model parameter conversion algorithm that bridges a priori information from semi-empirical, Ross–Li-based surface model parameters to physical, Rahman–Pinty–Verstraete (RPV)-based surface parameters. Validated through Directional Hemispherical Reflectance (DHR) and Bidirectional Reflectance Factor (BRF) comparisons, the algorithm demonstrates a highly successful conversion capability. Further retrieval tests indicate that the conversion algorithm introduces minimal impact to the aerosol inversion, providing a robust mechanism for surface a priori initialization.

The second application presents the Correlation based Inversion Method for Aerosol Properties (CIMAP) framework, an aerosol inversion strategy that utilizes Principal Component Analysis (PCA) to implement intrinsic correlation constraints among aerosol properties. By retaining a targeted number of Principal Components (PCs), CIMAP substantially reduces the state vector dimensionality, enhancing both numerical stability and computational efficiency. The framework is evaluated under two configurations: CIMAP-A: Utilizes ground-based AERONET observations across three distinct regions (southeast coast of the US, northern China, and southern Africa). Using only 7–8 PCs out of 30, CIMAP-A achieves a processing speed acceleration of over

80% while maintaining high accuracy relative to official AERONET inversion products, yielding Mean Absolute Difference (MAD) of 0.005, 0.019, 0.039, 0.003, 0.013 μm, 0.02 μm and for Aerosol Optical Depth (AOD), Single Scattering Albedo (SSA), Real and Imaginary parts of Refractive Index (RIR and RII), effective radius of fine and coarse modes of particle size distribution (PSD), respectively. CIMAP-P: Applies the framework to multi-angle, multi-spectral, space-borne polarimetric observations using POLDER data, integrating the surface conversion algorithm. Validation against collocated AERONET references shows strong agreement, with AOD MAD ranging from 0.015 to 0.042 and SSA MAD from 0.048 to 0.070 across reference wavelengths and 0.032 μm and 0.402 μm for the effective radius of fine- and coarse-mode PSD, exhibiting performance competitive with multiple Generalized Retrieval of Aerosol and Surface Properties (GRASP) algorithm configurations.

Limitations of the CIMAP configurations are discussed, alongside potential expansions toward globally covering observations and possible improvements to constraint construction. Ultimately, the flexible and computationally efficient architecture of CIMAP, with its surface initialization capability, positions it as a highly promising inversion framework for maximizing the rich information content offered by next-generation polarimetric missions.

Table of Contents

List of Figures

List of Tables

List of Abbreviations

3MI	Multi-Viewing Multi-Channel Multi-Polarization Imaging
AERONET	AErosol RObotic NETwork
AirMSPI	Airborne Multiangle SpectroPolarimetric Imager
AOD	Aerosol Optical Depth
BRF	Bidirectional Reflectance Factor
CIMAP	Correlation based Inversion Method for Aerosol Properties
CIMAP-A	CIMAP for AERONET
CIMAP-P	CIMAP for POLDER
CIMAP-W	CIMAP-P with no surface a priori
DB	MODIS Deep Blue algorithm
DHR	Directional-hemispherical reflectance
d-mRPV	degenerate form of mRPV model
d-RPV	degenerate form of RPV model
DT	MODIS Dark Target algorithm
GRASP	Generalized Retrieval of Aerosol and Surface Properties
GRASP-C	Component Configuration of GRASP
GRASP-HP	High-Precision Configuration of GRASP
GRASP-M	Models Configuration of GRASP
HARP-2	Hyper-Angular Rainbow Polarimeter-2
LUT	Look-Up-Table
MAD	Mean Absolute Difference
MAIA	Multi-Angle Imager for Aerosols
MAIAC	MODIS Multi-Angle Implementation of Atmospheric Correction algorithm
MD	Mean Differences
MISR	Multi-angle Imaging SpectroRadiometer

MODIS	Moderate Resolution Imaging Spectroradiometer
mRPV	modified RPV model
OCI	Ocean Color Instrument
PACE	Plankton, Aerosol, Cloud, and ocean Ecosystem
PCA	Principal Component Analysis
POLDER	POLarization and Directionality of the Earth's Reflectance
PSD	Particle Size Distribution
R^2	coefficient of determination
RII	Imaginary part of Refractive Index
RIR	Real part of Refractive Index
RMSE	Root Mean Square Error
RPV	Rahman-Pinty-Verstraete model
RT	Radiative Transfer
RTLS	Ross-Thick Li-Sparse model
SPEXone	Spectro-polarimeter for Planetary Exploration one
sRTLS	scale Ross-Thick Li-Sparse model
SSA	Single Scattering Albedo
STD	STandard Deviation
SZA	Solar Zenith Angle
TOA	Top-Of-Atmosphere

1. Introduction

1.1 Aerosol and Their Impacts

Aerosols are defined as solid or liquid particles suspended in the atmosphere. They originate from both natural sources, such as sandstorms, volcanic eruptions, and wildfires, and anthropogenic sources, including fuel combustion, industrial emissions, and agricultural activities. Aerosols play an important role in Earth's energy budget and climate system. Through their direct effects, aerosols can either warm or cool the atmosphere by absorbing and scattering incoming solar radiation [Chýlek and Coakley, 1974; Meyer et al., 2015]. In addition, aerosols can act as cloud condensation nuclei (CCN), altering cloud microphysical properties and influencing cloud formation processes through indirect effects [Twomey, 1974; Albrecht, 1989]. Aerosols also exert semi-direct effects by heating the surrounding atmosphere, thereby modifying cloud development and lifetime [Ackerman et al., 2000; Koch and Del Genio, 2010].

Consequently, aerosol-radiation and aerosol–cloud interactions remain among the largest sources of uncertainty in estimating radiative forcing and Earth's global energy budget [Pachauri et al., 2014]. The radiative and cloud-forming effects of aerosols strongly depend on their microphysical and chemical properties, including particle size, morphology, refractive index, and chemical composition. Accurate acquisition and characterization of aerosol properties is therefore essential for improving climate modeling, quantifying aerosol radiative effects, understanding aerosol–cloud interactions, and supporting environmental, agricultural, and public health applications.

Obtaining globally distributed measurements of aerosol properties with long-term temporal coverage remains challenging [Dubovik et al., 2021]. As a result, remote sensing techniques have become an essential approach for aerosol observation and data acquisition. Numerous passive remote sensing instruments have been developed to monitor global and regional aerosol distributions and to retrieve aerosol properties [Dubovik et al., 2019, 2021].

1.2 Remote Sensing Measurements

Setting ground-based instruments is a classic method to measure solar radiance attenuated by aerosols. Take AErosol RObotic NETwork (AERONET; Holben et al., 1998) as an example. AERONET is a ground-based network of passive CIMEL sun photometers that measures direct sun and sky radiances at visible and near-infrared wavelengths (440–1020 nm). Direct sun measurements are obtained using filtered detectors that quantify the spectral extinction of direct-beam solar radiation through the total atmospheric column. Based on the Beer–Lambert law, the Aerosol Optical Depth (AOD) can be derived from the measurements. Sky radiance data are acquired by scanning the sky at specific angular geometries, including almucantar and principal-plane scans. These angular measurements of scattered light are further processed through inversion algorithms to retrieve complex aerosol microphysical properties, such as the single scattering albedo (SSA), volume particle size distribution (PSD), and real and imaginary refractive index (RIR & RII) [Dubovik and King, 2000]. Because AERONET products are derived from direct or near-direct measurements, they possess significantly lower uncertainties than remote sensing products from air-borne or satellite-borne instruments. Consequently, AERONET is one of the primary references for validating retrieved aerosol properties from other instruments.

While ground-based networks like AERONET provide high-accuracy point measurements, spaceborne instruments offer the broad spatial coverage and high temporal repeatability necessary for aerosol monitoring at regional to global scales, especially over remote regions inaccessible to surface stations. Among these, the Moderate Resolution Imaging Spectroradiometer (MODIS;

Salomonson et al., 1989) has served as a cornerstone for global observations. Currently operating on NASA's Terra (launched in 1999) and Aqua (launched in 2002) satellites, MODIS has provided over 25 years of continuous climate data. As a 36-band spectrometer, it views the entire Earth's surface every one to two days, measuring reflected solar and emitted thermal radiation across a broad spectral range from 0.4 μm to 14.4 μm. This extensive spectral diversity enables the simultaneous retrieval of aerosol, cloud, and surface properties, providing a comprehensive view of the Earth's atmospheric and terrestrial systems [Remer et al., 2005].

However, a single-view instrument like MODIS faces limitations in providing enough information for solving ill-possdness in retrieval process. To address this, multi-angle sensors were developed to provide additional constraints on the retrieval process. The Multi-angle Imaging SpectroRadiometer (MISR; Diner et al., 1998) stands as one of the first satellite-borne instruments designed to monitor aerosols, clouds, and surfaces using multiple viewing geometries. MISR utilizes nine cameras to capture data at various angles (nadir, ± 26.1°, ± 45.6°, ± 60.0° and ± 70.5°) across four spectral bands (446, 558, 672, and 867 nm). By observing the same atmospheric column from multiple perspectives, MISR significantly increases the information content available for inversions, thereby enabling more robust constraints on aerosol microphysical properties [Martonchik et al., 1998; Diner et al., 1999, 2008; Limbacher et al., 2022].

Alongside the advancement of passive remote sensing over the past few decades, polarimetry, which measures both the intensity and polarimetric components of radiation, has emerged as a central trend in satellite-borne instrument design [Dubovik et al., 2019]. By capturing the unique scattering and absorption signatures of aerosols within polarimetric signals, these instruments provide information content beyond what conventional radiometers alone can offer. This additional data is crucial for mitigating ill-posedness during the inversion process, ultimately enabling more accurate retrievals [Xu et al., 2017; Yang et al., 2026]. Indeed, the polarimeter has been proven both theoretically and practically to be a powerful tool for improving the characterization of aerosol abundance and microphysical properties, as well as cloud and surface reflection properties [Mishchenko et al., 2011; Hasekamp et al., 2011; Knobelspiesse et al., 2012; Dubovik et al., 2019].

Among the pioneering instruments in this field, the POLarization and Directionality of the Earth's Reflectance (POLDER; Deschamps et al., 1994) stands out as the first spaceborne polarimetric instrument, accumulating nearly a decade of observations and establishing the longest time series of polarimetric measurements from space. The POLDER-3 instrument aboard PARASOL extended the maximum number of view angles to 16 and acquired observations at 443, 490 (polarized), 565, 670 (polarized), 763, 765, 865 (polarized), 910, and 1020 nm, with a spatial resolution of 5.3 × 6.2 km at nadir and a swath of 2100 × 1600 km enabling near-global coverage in approximately two days. As the first of its kind, POLDER fundamentally shaped the design of subsequent generations of polarimetric instruments.

Building directly on POLDER's legacy, the Multi-Viewing Multi-Channel Multi-Polarization Imaging (3MI; Fougnie et al., 2018) instrument, carried aboard the EUMETSAT Polar System Second Generation platform and launched in 2025, represents a significant evolution of these principles. Where POLDER operated with 3 polarization bands, 3MI expands this capability to 9 polarization bands across the visible and near-infrared, spanning 12 wavelengths from 410 to 2120 nm, providing substantially tighter constraints on aerosol and cloud property retrievals.

A parallel trajectory of innovation is evident in NASA's Plankton, Aerosol, Cloud, and ocean Ecosystem (PACE; Petro et al., 2020) mission, launched in 2024. PACE carries two dedicated polarimeters, the Hyper-Angular Rainbow Polarimeter-2 (HARP-2; McBride et al., 2024), which measures at 441, 549, 669, and 873 nm, and the Spectro-polarimeter for Planetary Exploration one (SPEXone; Hoogeveen et al., 2023), which provides continuous spectral coverage from 385 to 770 nm. Together, these polarimetric measurements are designed to constrain aerosol properties and support the mission's third instrument, the Ocean Color Instrument (OCI; Meister et al., 2024), by enabling improved correction of water-leaving radiances, illustrating how polarimetry increasingly serves as an integrating element within multi-instrument observing systems.

Looking ahead, the Multi-Angle Imager for Aerosols (MAIA; Diner et al., 2018), inspired by the MISR mission, extends polarimetric remote sensing toward a new application domain. MAIA will deliver multi-angle observations across 14 spectral bands spanning the ultraviolet to shortwave infrared, with three polarimetric aerosol bands centered at 445, 645, and 1039 nm. Distinctively, MAIA is primarily focused on characterizing speciated aerosol composition and its effects on human health, broadening the societal relevance of polarimetric Earth observation. The mission is currently anticipated to launch in 2027.

1.3 Aerosol Retrieval Algorithm

Alongside the advancement of polarimetric and multi-angle instruments, sophisticated inversion algorithms are equally essential for comprehensive aerosol research, particularly in the context of radiative forcing estimation. These algorithms exploit the rich spectral, angular, and polarimetric information provided by remote sensing instruments to generate optimal aerosol parameters by leveraging the physical relationships embedded in forward radiative transfer (RT) modeling. As instruments have grown more capable and diverse, inversion algorithms have similarly evolved to accommodate data from different observing systems, continuously improving the accuracy and computational efficiency of aerosol remote sensing.

One of the most established approaches is the look-up-table (LUT) method, which operates by comparing measured radiances against a precomputed, multi-dimensional table of simulated radiances generated from a range of prescribed aerosol combinations. The retrieval identifies the best match by minimizing fitting residuals between the simulated and observed signals. The LUT strategy has been widely adopted in operational retrieval systems precisely because of its computational tractability. By performing the radiative transfer calculations offline, retrieval speed is greatly accelerated. This efficiency makes LUT-based methods well suited for global, near-real-time processing. For instance, the three operational MODIS aerosol retrieval algorithms [Remer et al., 2005] all adopt this strategy [Dark Target (DT): Levy et al., 2013; Deep Blue (DB): Hsu et al., 2019; Multi-Angle Implementation of Atmospheric Correction (MAIAC): Lyapustin et al., 2018], as does the operational MISR aerosol algorithm, which searches for solutions within a LUT of radiances representing 74 distinct aerosol mixtures [Diner et al., 2008]. Despite these advantages, however, LUT methods are inherently constrained by the discrete parameter space. Because retrievals are limited to the aerosol combinations preloaded into the table, realistic variations in aerosol properties that fall between sampled states may not be adequately represented, potentially introducing systematic biases in retrieval outcomes.

To address this fundamental limitation, a variety of optimization-based retrieval algorithms have been developed over the past two decades for joint aerosol and surface property retrieval [Kokhanovsky et al., 2015]. Rather than searching a fixed discrete table, optimization methods solve for unknown parameters by iteratively minimizing the difference between observed and forward-modeled signals within a continuous parameter space. This continuous representation allows the retrieval to capture a far broader range of aerosol states, improving fidelity in complex scenarios where aerosol composition, loading, or surface reflectance deviates from simplified assumptions. A widely used example is the Levenberg–Marquardt (LM; Levenberg et al., 1944; Marquardt et al., 1963) algorithm, which blends gradient descent and the Gauss-Newton method to balance convergence efficiency with numerical robustness across varying retrieval conditions.

However, the increased flexibility of optimization-based approaches also introduces a critical challenge: ill-posedness. Since multiple combinations of aerosol properties can produce nearly identical remotely sensed signals, and the information content of any finite set of measurements is inherently limited, the inverse problem is often underdetermined. Without additional constraints, optimization retrievals risk converging to physically unrealistic solutions. Thus, the stable optimization retrieval of aerosol properties often relies on the imposition of additional constraints [Tikhonov and Arsenin, 1977]. A foundational approach in this regard is the smoothness constraint [Dubovik and King, 2000], which encodes the physical expectation that aerosol properties, such as the real and imaginary parts of the refractive index as a function of wavelength, or the size distribution as a function of particle radius, should vary smoothly and continuously rather than fluctuate erratically. By penalizing abrupt variations among retrieved parameters, this constraint effectively narrows the solution space to physically plausible states. The AERONET inversion algorithm adopted this strategy as a core component of its operational retrieval framework [Dubovik and King, 2000], demonstrating its practical value at the global scale.

Extending this concept further, Dubovik et al. (2011) introduced the multi-pixel inversion approach, which simultaneously exploits the spatial smoothness of aerosol properties across neighboring pixels and the temporal smoothness of surface reflectance over successive observations. By jointly inverting multiple pixels and time steps, this approach introduces additional degrees of constraint that a single-pixel retrieval cannot access, substantially improving both retrieval accuracy and computational efficiency. This multi-pixel framework is embedded at the Generalized Retrieval of Aerosol and Surface Properties (GRASP; Dubovik et al., 2014), which has since become one of the most comprehensive and widely applied aerosol inversion algorithms in the community. GRASP serves as a unifying framework which is capable of ingesting observations from diverse instruments and retrieval configurations [Chen et al., 2020, 2024; Lopatin et al., 2013, 2021; Litvinov et al., 2024].

1.4 A Priori Constraints

While the constraints discussed previously have demonstrated considerable success in stabilizing the retrieval, minimizing retrieval uncertainties remains an ongoing challenge in algorithm development. A straightforward strategy is to incorporate known or relatively reliable information about gas composition, surface reflectance, and aerosol climatologies into the retrieval configuration. Such information can guide the retrieval away from ill-posed solutions and thereby enhance accuracy.

One classic application of climatological database is the use of gas climatology to build LUTs for gaseous absorption correction [Patadia et al., 2018]. Atmospheric gases attenuate the signal measured at aerosol-sensitive spectral bands, and this attenuation must be corrected before the aerosol signal can be reliably isolated. To achieve this, climatological gas profiles are used to compute the Rayleigh optical depth and generate a spectrally resolved total gas transmittance correction factor, effectively restoring a "gas-free" signal at aerosol retrieval bands [Bodhaine et al., 1999]. Many operational algorithms, including MODIS DT, DB, and MAIAC, as well as the MISR algorithm [Levy et al., 2013; Hsu et al., 2004; Lyapustin et al., 2018; Diner et al., 2008], apply this approach to account for absorption by atmospheric gases including water vapor (H_2O), ozone (O_3), carbon dioxide (CO_2), methane (CH_4), oxygen (O_2), nitrous oxide (N_2O), and nitrogen dioxide (NO_2), among other trace gases.

Beyond gas correction, assumptions about aerosol microphysical and optical properties are widely used to reduce the degrees of freedom in the retrieval. Different algorithms adopt different levels of aerosol model complexity, ranging from regionally prescribed static models to flexible compositional mixtures.

The MAIAC algorithm employs eight regional aerosol models with corresponding LUTs over land, including a dedicated dust model, allowing the global domain to be discretized with sufficient regional specificity to capture aerosol variability [Lyapustin et al., 2018]. The models can be either static, with fixed parameters representative of arid environments, or dynamic [Remer and Kaufman, 1998], with parameters that vary as a function of AOD to reflect hygroscopic growth in moderate-to-high humidity regions. Model parameters including size distribution, volumetric concentration ratios, and refractive index are generally representative of AERONET regional climatology.

The MISR retrieval algorithm takes a discrete mixture approach, relying on a LUT generated for a predefined set of aerosol mixtures with known optical properties. The Version 22 (V22) operational algorithm defines 74 aerosol mixtures, each consisting of up to three unique particle types with prescribed optical and microphysical properties [Kahn et al., 2010; Kahn and Gaitley, 2015; Witek et al., 2018]. These 74 mixtures are built from combinations of eight primary particle types designed to represent compositional categories typically found in the atmosphere, including sea spray, sulfate, nitrate, mineral dust, carbonaceous, and urban soot aerosols.

In contrast to the discrete LUT approach, the GRASP algorithm adopts two distinct but related aerosol model frameworks that offer greater flexibility. GRASP-Models represents aerosol as an external mixture of climatological components, each with prescribed size distribution and optical properties with only the relative concentrations of those components to be retrieved [Chen et al., 2020; Lopatin et al., 2021; Dubovik et al., 2021]. GRASP-Component takes this further by assuming an internal mixture of aerosol compositions, directly retrieving the volume fractions of six chemical components: black carbon (BC), brown carbon (BrC), fine-mode non-absorbing insoluble (FNAI), coarse-mode non-absorbing insoluble (CNAI), coarse-mode absorbing insoluble (CAI, representing primarily iron oxides in mineral dust), size distribution and relative humidity of the host medium [Li et al., 2019; Zhang et al., 2022]. This application fixes the refractive index of possible mixtures, reducing retrieval degrees of freedom while preserving sensitivity to aerosol composition.

Inspired by the well-developed a priori constraint framework established in prior work, this dissertation introduces new tools and applications that systematically incorporate climatological information into the aerosol retrieval process, thereby improving both the robustness and overall performance of the retrievals.

The first application addresses the challenge of surface model compatibility for retrieval initialization [Huang et al., 2026a]. Specifically, a conversion algorithm is developed that transforms Ross-Thick Li-Sparse-based (RTLS; Ross, 1981; Li and Strahler, 1992) surface model parameters derived from climatology into the equivalent Rahman-Pinty-Verstraete-based (RPV; Rahman et al., 1993, Martonchik et al., 1998; Engelsen et al., 1996) surface model representation. This conversion broadens the range of climatological datasets available for retrieval initialization, enabling more effective constraint of surface reflectance properties during the decoupling of surface and aerosol signals which is a critical step in ensuring retrieval accuracy and convergence efficiency.

The second application presents a complete inversion framework: Correlation based Inversion Method for Aerosol Properties (CIMAP) that leverages well-tested a priori constraints, specifically correlation-based constraints [Xu et al., 2019; Huang et al., 2021, 2026b], to retrieve aerosol optical and microphysical properties. This framework extends the methodology previously demonstrated for the Airborne Multiangle SpectroPolarimetric Imager [AirMSPI; Diner et al., 2013] to two additional and complementary observation platforms: the ground-based AERONET and the satellite-borne POLDER instrument.

The extension to these platforms represents a meaningful step toward multi-platform consistency and broader applicability of the correlation-based retrieval approach. The AERONET application establishes the baseline performance of CIMAP by exploiting the rich spectral and temporal information available from ground-based sun-sky radiometer observations, offering a well-characterized testbed for evaluating retrieval accuracy. The POLDER application, in turn, demonstrates CIMAP's capacity to operate on multi-angle, multi-spectral Top-Of-Atmosphere (TOA) observations by jointly exploiting both the radiometric and polarimetric components of the measured signal, demonstrating capability that is particularly valuable for constraining aerosol microphysical properties and that establishes a methodological foundation for application to future satellite missions.

Together, these two retrieval tests aim to derive aerosol properties with improved computational efficiency and robustness across a diverse range of observational contexts. By incorporating well-characterized a priori constraints, CIMAP directly addresses the inherent ill-posedness of the aerosol retrieval problem which is a fundamental challenge arising from the limited information content of the measured signal relative to the large number of unknowns to be simultaneously retrieved.

Notably, the CIMAP POLDER retrieval further incorporates the surface model conversion algorithm developed in the first application, directly linking the two contributions of this dissertation and demonstrating the enhanced retrieval accuracy after surface parameter initialization. This connection illustrates a broader, systematic retrieval framework that is capable of accommodating multiple a priori constraints alongside fundamental inversion constraints, highlighting the flexibility and adaptability of CIMAP across different retrieval configurations and observational settings.

1.5 Sections of the Dissertation

This dissertation is organized as follows. Section 2 introduces the foundational retrieval structure of CIMAP, encompassing the universal configuration settings, the assumptions and formulations underlying the retrieved aerosol properties, observation from instruments, radiative transfer modeling, light scattering calculations, and surface reflectance calculations. Section 3 presents the surface a priori conversion algorithm, covering the analysis of the target surface a priori climatology, the design of the conversion scheme, and a quantitative evaluation of the conversion quality. Section 4 details the two CIMAP retrieval configurations, discussing the principles of the correlation-based constraints, the specific algorithmic structure of each configuration, and systematic assessments of retrieval performances. Section 5 provides a summary of the key findings, a broader discussion of their scientific implications, and directions for future work.

2. Basic Structure of CIMAP Retrieval Framework

The general algorithmic structure common to the CIMAP configurations presented in this dissertation is introduced in this section. The core components of CIMAP include observational inputs, aerosol and surface properties as retrieved parameters, forward modeling of RT combined with light scattering and surface reflectance models, and an inversion module. This structure follows the foundational CIMAP framework of Xu et al. (2017).

Compared to the classical retrieval configurations such as those of Dubovik and King (2000) and Dubovik et al. (2011), CIMAP retrieves aerosol parameters in principal component (PC) space rather than operating directly in the conventional parameter space. By incorporating PC-based aerosol a priori constraints into the retrieval, CIMAP reduces the number of unknown parameters to be retrieved, enhances numerical stability of the least-squares solution, and provides improved retrieval efficiency.

In the atmospheric configuration adopted for CIMAP in this dissertation, the coupled atmosphere-surface system is decomposed into a layered structure from the top of the atmosphere to the surface, consisting of a pure Rayleigh scattering layer in the upper atmosphere, a mixed layer where aerosol particles coexist with Rayleigh scattering molecules in the lower atmosphere, and a land surface at the bottom. This layered decomposition captures the dominant radiative processes relevant to aerosol retrieval while remaining computationally tractable.

2.1 Aerosol properties

Aerosol microphysical properties are broadly classified into three fundamental categories: shape, size, and composition. Each of these categories plays a critical role in determining how aerosol particles interact with solar and terrestrial radiation, and together they form the essential inputs required for accurate RT modeling and aerosol retrieval.

Real atmospheric aerosols are mixtures of numerous chemical compounds including sulfates, nitrates, black carbon, organic carbon, mineral dust, and sea salt, each with distinct light-absorbing and light-scattering properties. Directly retrieving the full chemical composition of aerosols from column effective remote sensing observations is not feasible, as the measurements do not contain sufficient information to resolve individual chemical constituents. Instead, the complex refractive index, comprising a real part (RIR) that governs light scattering and an imaginary part (RII) that governs light absorption, is adopted as a proxy for aerosol composition in current retrieval algorithms and RT modeling frameworks [Dubovik and King 2000, Dubovik et al., 2011]. This is physically justified because the refractive index encapsulates the bulk optical response of the aerosol mixture to incident radiation and polarization, effectively summarizing the net radiative effect of the underlying chemical composition without requiring explicit knowledge of each constituent. In CIMAP, we describe the refractive index as a function of wavelength λ:

$$n(\lambda) = n_{\mathrm{r}}(\lambda) + i \times n_{\mathrm{i}}(\lambda) \quad (1)$$

where n_{r} is RIR and n_{i} is RII.

Aerosol size is characterized through the volumetric particle size distribution, which describes how the total aerosol volume is distributed across particles of different radii. Coarse-mode particles, such as mineral dust and sea salt, and fine-mode particles, such as combustion-derived aerosols, interact with radiation in fundamentally different ways. Accurately representing this size-dependent behavior is essential for computing aerosol optical properties. The volumetric size distribution ($\frac{dV(r)}{d\ln r}$) is therefore a central retrieved parameter in aerosol retrieval algorithms, typically represented as a bimodal lognormal distribution with separate fine and coarse mode components, each characterized by its own volume concentration ($C_{\mathrm{v,fine}}$ and $C_{\mathrm{v,coarse}}$), effective radius ($r_{\mathrm{m,fine}}$ and $r_{\mathrm{m,coarse}}$), and standard deviation (σ_{fine} and σ_{coarse}):

$$V(r) = \frac{dV(r)}{d\mathrm{In}r} = C_{\mathrm{v,fine}} \frac{1}{\sqrt{2\pi}\,\sigma_{\mathrm{fine}}} \exp\left[-\frac{(\mathrm{In}r - \mathrm{In}\, r_{\mathrm{m,fine}})^2}{2\sigma_{\mathrm{fine}}^2}\right] + C_{\mathrm{v,coarse}} \frac{1}{\sqrt{2\pi}\,\sigma_{\mathrm{coarse}}} \exp\left[-\frac{(\mathrm{In}r - \mathrm{In}\, r_{\mathrm{m,coarse}})^2}{2\sigma_{\mathrm{coarse}}^2}\right]. \quad (2)$$

In reality, atmospheric aerosol particles are highly irregular in shape. Embedding fully realistic, particle-by-particle shape representations into RT calculations is, however, computationally prohibitive, as it would require solving the electromagnetic scattering problem for an effectively infinite variety of irregular geometries across all particle sizes and orientations. As a practical and well-established approximation, a sphericity parameter ($f_{\mathrm{v,sphere}}$) is therefore adopted to represent particle shape in current aerosol retrieval frameworks. This parameter represents the fraction of particles that are modeled as spheres, with the remaining fraction represented as randomly oriented spheroids. Light-scattering properties of non-spherical but rotationally symmetric particles can be computed efficiently using T-matrix methods or pre-computed look-up tables [Dubovik et al., 2002, 2006]. By mixing spherical and spheroidal particles in a proportion determined by the retrieved sphericity parameter, the model can approximate the scattering behavior of irregular particles without the full computational burden of exact non-spherical RT calculations. This approach has been shown to provide a reasonable representation of dust and other non-spherical aerosol types, particularly in reproducing the depolarization of scattered light that is characteristic of irregular particles and that would be entirely absent in a purely spherical particle model [Dubovik et al., 2011].

One additional retrieved aerosol property is mean height of aerosol layer ($\overline{H}_{\mathrm{aer}}$), which describes the vertical distribution of aerosol concentration. CIMAP adopts a Gaussian vertical profile which is an approach widely used in aerosol retrieval algorithms (e.g., Dubovik et al., 2011; Xu et al., 2016, 2017, 2019), as it provides a practical approximation of aerosol vertical distribution for column-integrated retrievals when limited vertical information is available. We acknowledge that this assumption carries inherent limitations. In particular, it may not fully capture complex vertical structures such as multi-layer aerosol distributions, elevated dust plumes, or strong convective conditions. Nevertheless, for passive multi-angle polarimetric observations, which are primarily sensitive to column-integrated properties, the Gaussian profile offers a reasonable and computationally efficient parameterization. The Gaussian distribution ($c_{\mathrm{k}}\,(h)$) can be expressed as

$$c_{\mathrm{k}}\,(h) \sim \frac{\sqrt{\pi}\sigma_{\mathrm{k}}}{2}\left[\mathrm{erf}\left(-\frac{(h_{\max}-\overline{H}_{\mathrm{aer}})}{\sigma_{\mathrm{k}}}\right) - \mathrm{erf}\left(-\frac{(h_{\min}-\overline{H}_{\mathrm{aer}})}{\sigma_{\mathrm{k}}}\right)\right]\exp\left(-\frac{(h-\overline{H}_{\mathrm{aer}})^2}{{\sigma_{\mathrm{k}}}^2}\right) \tag{3}$$

where σ_{k} is the standard deviation characterizing the width of the aerosol layer; $h_{\max}$ and $h_{\min}$ are the maximum and minimum layers of assumed aerosol/Rayleigh mixed layer and $\int_{h_{\min}}^{h_{\max}} c_{\mathrm{k}}\,(h)\ dh = 1$. Consistent with Dubovik et al. (2011), σ_{k} is fixed at 1 km, leaving aerosol layer height as the sole retrieved parameter for the vertical concentration profile.

2.2 Observations

Observations are an essential component of the inversion framework, determined by the measurement principle, spectral range, resolution, and angular geometries of the sensors. Radiometers or polarimeters collect attenuated signals based on aerosol absorption and scattering at aerosol-sensitive wavelengths. RT modeling takes aerosol properties as inputs and simulates the corresponding observations, which are then compared to real observations to compute the residual used to inform the retrieval. As one of the Stokes parameters, light intensity I, commonly referred to as radiance, represents the total amount of radiative energy reaching the sensor, and is the standard radiometric component of measurements. For the polarimetric component, the other two Stokes parameters, Q and U, are commonly involved [Dubovik et al., 2019]. These are typically combined into the degree of linear polarization (DoLP) for use as retrieval observations, defined as:

$$\mathrm{DoLP} \approx \frac{\sqrt{Q^2 + U^2}}{I}. \tag{4}$$

2.3 Radiative Transfer Modeling

RT modeling is a key component in the aerosol retrieval process, serving as the forward model that simulates radiometric or polarimetric observations given a specific set of aerosol and surface properties as inputs. It provides the physical link between the aerosol microphysical and optical properties to be retrieved and the actual measurements recorded by the instrument.

The RT model adopted for all projects in this dissertation is the Markov Chain radiative transfer model (MarCh; Xu et al., 2011, 2012). The MarCh model achieves high computational efficiency through compact matrix operations which is an important advantage for iterative retrieval algorithms requiring repeated forward model evaluations. The MarCh model further supports coupling with different BRDF-based surface reflectance models, as well as pre-computed look-up tables for Rayleigh scattering online calculation, making it a flexible and efficient RT framework well-suited to the diverse aerosol retrieval applications explored in this dissertation.

2.4 Light Scattering and Absorption Calculation

In the MarCh RT model, the aerosol scattering and absorption properties are calculated. Aerosol light scattering properties are calculated using Mie theory [van de Hulst, 1957] for spherical particles and using the combined T-matrix and geometrical optics method for non-spherical particles (spheroids) [Dubovik et al., 2002, 2006]. The output includes extinction coefficient, scattering coefficient, and phase matrix. Aerosol scattering calculation uses refractive index and particle size distribution as input.

Accounting for aerosol size distribution, the extinction coefficient for particles ($K_{\mathrm{ext}}^{\mathrm{s}}$) is calculated by [Hansen and Travis, 1974],

$$K_{\mathrm{ext}}^{\mathrm{s}}(\lambda; n; r) = \int_{\ln r_{\min}}^{\ln r_{\max}} \frac{C_{\mathrm{ext}}^{\mathrm{s}}(\lambda; n; r)}{V(r)}\, d\ln r \tag{5}$$

where $C_{\mathrm{ext}}^{\mathrm{s}}(\dots)$ is the spectral extinction cross-section of an aerosol particle of radius r, refractive index real n, and volumetric particle size distribution $V(r)$ which is related to the number-weighted size distribution $N(r)$ by

$$V(r) = \frac{4\pi r^3}{3} N(r). \tag{6}$$

Using spheroids to approximate non-spherical atmospheric particles, the calculation of extinction coefficient ($K_{\mathrm{ext}}^{\mathrm{ns}}$) need to account for both size and aspect ratio distributions [Dubovik et al., 2002]:

$$K_{\mathrm{ext}}^{\mathrm{ns}}(\lambda; n; r) = \int_{\ln r_{\min}}^{\ln r_{\max}} \int_{\ln \varepsilon_{\min}}^{\ln \varepsilon_{\max}} \frac{C_{\mathrm{ext}}^{\mathrm{ns}}(\lambda; n; r)}{V(r)} \frac{dN(\varepsilon)}{d\ln \varepsilon}\, d\ln \varepsilon\, d\ln r \tag{7}$$

where $C_{\mathrm{ext}}^{\mathrm{ns}}$ is the extinction cross-section of a non-spherical particle; ε is the spheroid axis ratio. In the current study, the lower and upper bounds for the aspect ratio are set to be 0.33 and 3.0

respectively [Dubovik et al., 2006; Legrand et al., 2014]. The extinction coefficient is used to calculate aerosol optical depth (τ) via multiplying it with aerosol's column concentration, namely $\tau = C_{\mathrm{v}} \times K_{\mathrm{ext}}$. Accounting for the contribution by both spherical (τ_{s}) and non-spherical (τ_{ns}) particle AODs, the total aerosol optical depth is

$$\tau_{\mathrm{ext}}^{\mathrm{aer}}(\lambda) = \tau_{\mathrm{s}}(\lambda) + \tau_{\mathrm{ns}}(\lambda) \tag{8}$$

By introducing volumetric non-spherical fraction f_{s}, combining Eqs. ((5)), ((7)), and ((8)), $\tau_{\mathrm{ext}}^{\mathrm{aer}}(\lambda)$ can be derived as

$$\tau_{\mathrm{ext}}^{\mathrm{aer}}(\lambda) = \sum_{i=1}^{N_{\mathrm{bin}}} [\, f_{\mathrm{s}}\, K_{\mathrm{ext}}^{\mathrm{s}}(\lambda; n; r_i) + (1 - f_{\mathrm{s}}) K_{\mathrm{ext}}^{\mathrm{ns}}(\lambda; n; r_i)]\, \frac{dV(r_i)}{d\ln r} \tag{9}$$

where i denotes the bin index of the particle size distribution, which is discretized into N_{bin} size bins. In a similar manner, the scattering coefficient of spherical and non-spherical particles ($K_{\mathrm{sca}}^{\mathrm{s}}$ and $K_{\mathrm{sca}}^{\mathrm{ns}}$) are calculated by replacing C_{ext} by C_{sca} in Eqs. (5) and (7). Then the aerosol single scattering albedo (ω_0^{aer}) is obtained using

$$\omega_0^{\mathrm{aer}}(\lambda) = \frac{f_{\mathrm{s}} K_{\mathrm{sca}}^{\mathrm{s}} + (1 - f_{\mathrm{s}}) K_{\mathrm{sca}}^{\mathrm{ns}}}{f_{\mathrm{s}} K_{\mathrm{ext}}^{\mathrm{s}} + (1 - f_{\mathrm{s}}) K_{\mathrm{ext}}^{\mathrm{ns}}}. \tag{10}$$

Radiative transfer calculation uses the bulk aerosol properties of each atmospheric layer. The layer effective optical depth ($\tau_{\mathrm{ext}}^{\mathrm{layer}}$) and single scattering albedo ($\omega_0^{\mathrm{layer}}$), and phase matrix ($\mathbf{P}^{\mathrm{layer}}$) are contributed by both aerosols and molecules, namely:

$$\tau_{\mathrm{ext}}^{\mathrm{layer}}(\lambda) = \tau_{\mathrm{ext}}^{\mathrm{aer}}(\lambda) + \tau_{\mathrm{ext}}^{\mathrm{mol}}(\lambda), \tag{11}$$

$$\omega_0^{\mathrm{layer}}(\lambda) = \frac{\tau_{\mathrm{ext}}^{\mathrm{aer}}(\lambda)\omega_0^{\mathrm{aer}}(\lambda) + \tau_{\mathrm{ext}}^{\mathrm{mol}}(\lambda)\omega_0^{\mathrm{mol}}(\lambda)}{\tau_{\mathrm{ext}}^{\mathrm{total}}(\lambda)}, \tag{12}$$

$$\mathbf{P}^{\text{layer}}(\lambda;\theta_{\text{sca}}) = \frac{\tau_{\text{ext}}^{\text{aer}}(\lambda)\omega_0^{\text{aer}}(\lambda)\mathbf{P}^{\text{aer}}(\lambda;\theta_{\text{sca}}) + \tau_{\text{ext}}^{\text{mol}}(\lambda)\mathbf{P}^{\text{mol}}(\lambda;\theta_{\text{sca}})}{\tau_{\text{ext}}^{\text{aer}}(\lambda)\omega_0^{\text{aer}}(\lambda) + \tau_{\text{ext}}^{\text{mol}}(\lambda)} \,. \tag{13}$$

Informing MarCh RT with the layer-resolved optical properties and surface reflection, the TOA radiance field characterized by Stokes components I, Q and U are calculated and compared to the remote sensing data.

2.5 Surface Reflectance Calculation

Accounting for both non-polarimetric and polarimetric reflectance from the surface is also a major component of the forward modeling. In this dissertation, several Ross-Li-based and RPV-based surface models are employed. Although these models are based on different physical interpretations, they share the same functionality: reconstructing the angular Bidirectional Reflectance Factor (BRF) of the surface, and are all suitable for incorporation into the RT modeling and retrieval framework.

For the depolarizing surface reflection calculation, a total of 6 surface BRDF models are introduced in this section, each serving different purposes across the projects in this dissertation. In the two configurations of the CIMAP retrieval framework, the original Ross-Li [Ross 1981; Li 1992] and modified RPV (mRPV; Rahman et al., 1993, Martonchik et al., 1998; Engelsen et al., 1996) models are employed. In the conversion algorithm, the scaled Ross-Li (sRTLS; Lyapustin et al., 2025), regular and degenerate forms of RPV and modified RPV models are involved, where the sRTLS model serves as the conversion source and the remaining models serve as targets for generating a priori surface model parameters.

Both the Ross-Li and RPV related surface models have been widely adopted in interpretation of instrument measurements [Jones and Vaughan 2010], reconstruction of multispectral surface reflectance data [Hou et al., 2025], development of land classification algorithms [De Colstoun and Walthall 2006; Jiao et al., 2011], and for atmospheric correction [Lyapustin et al., 2018]. The Ross-Li model is used for generating the land surface products from the observations made by the MODIS [King et al., 1992], while a modified form of the RPV model is used for generating the land surface products from the observations made by the MISR [Martonchik et al., 1998].

2.5.1 The Ross-Li Model and Its Variants

The semi-empirical and linear kernel-driven Ross-Li model simulates the reflection of complex surfaces using three additive terms, each representing a distinct scattering mechanism: isotropic, volumetric, and geometric [Ross 1981; Li 1992; Wanner et al. 1995]. The isotropic kernel (ρ_{iso}) that represents a uniform and non-directional scattering component, the volumetric scattering kernel (ρ_{vol}) that represents the light scattering within a dense canopy, and the geometric scattering kernel (ρ_{geo}) that represents the geometric-optics effects such as shadows and occlusion from tree crowns or rough surfaces. The total surface BRF ($R_{\text{Ross-Li}}$) becomes:

$$R_{\text{Ross-Li}}\left(f_{\text{iso}}, f_{\text{geo}}, f_{\text{vol}}; \theta_0, \theta_{\text{v}}, \Delta\phi\right) \tag{14}$$
$$= f_{\text{iso}} \cdot \rho_{\text{iso}} + f_{\text{vol}} \cdot \rho_{\text{vol}}(\theta_0, \theta_{\text{v}}, \Delta\phi) + f_{\text{geo}} \cdot \rho_{\text{geo}}(\theta_0, \theta_{\text{v}}, \Delta\phi)$$

where θ_v and θ_0 represent the view and solar zenith angle respectively, $\Delta\phi$ is the azimuth angular difference between an incoming and reflected light ray, namely $\Delta\phi = \phi_{\text{v}} - \phi_0$. The relative azimuth angle is defined as $\Delta\phi = \phi_{\text{v}} - \phi_0$, where ϕ_0 and ϕ_{v} are the solar and viewing azimuth angles, respectively. In this work, $\Delta\phi = 0°$ corresponds to the forward-scattering direction and $\Delta\phi = 180°$ to the backscattering direction, consistent with the convention commonly used in radiative transfer theory. The multipliers ($f_{\text{iso}}, f_{\text{geo}}$ and f_{vol}) to these kernels control the weight of the relevant kernels that are often saved in some datasets such as MAIAC, the volumetric scattering

(ρ_{vol}) and geometric (ρ_{geo}) kernels capture more variety of angular distribution of surface reflection.

For the volume scattering kernel, the “thick” approximation option for volume scattering stands for the scenarios of high leaf area index (LAI) (a dense leaf canopy) [Roujean et al., 1992], where as the “thin” approximation option represents the low LAI scenarios (a thin leaf canopy) [Wanner et al., 1995]:

$$\rho_{\mathrm{vol}}(\theta_0, \theta_{\mathrm{v}}, \Delta\phi) = \begin{cases} \dfrac{(\pi/2 - \xi)\cos\xi + \sin\xi}{\cos\theta_0 + \cos\theta_{\mathrm{v}}} - \dfrac{\pi}{4}, & option = \text{"}thick\text{"} \\ \dfrac{(\pi/2 - \xi)\cos\xi + \sin\xi}{\cos\theta_0 \cos\theta_{\mathrm{v}}} - \dfrac{\pi}{2}, & option = \text{"}thin\text{"} \end{cases} \tag{15}$$

where $\cos\xi = \cos\theta_0\cos\theta_{\mathrm{v}} + \sin\theta_0\sin\theta_{\mathrm{v}}\cos\Delta\phi$. These two options are derived from two different approximations to a single-scattering radiative transfer theory by Ross 1981.

For the geometric scattering kernel, the “spare” option represents the effects of background shadows cast by sparse ensemble of surface objects, where as the “dense” option stands for the mutual shadowing effects from denser aggregates of sunlit crowns [Strahler et al., 1994; Wanner et al., 1995]:

$$\rho_{\mathrm{geo}}(\theta_0, \theta_{\mathrm{v}}, \Delta\phi) = \begin{cases} O(\theta_0, \theta_{\mathrm{v}}, t) - \sec\theta_0' - \sec\theta_{\mathrm{v}}' + \dfrac{1}{2}(1 + \cos\xi')\sec\theta_{\mathrm{v}}', & option = \text{"}sparse\text{"} \\ \dfrac{(1 + \cos\xi')\sec\theta_{\mathrm{v}}'}{\sec\theta_0' + \sec\theta_{\mathrm{v}}' - O(\theta_0, \theta_{\mathrm{v}}, t)} - 2, & option = \text{"}dense\text{"} \end{cases} \tag{16}$$

where

$$O = \frac{1}{\pi}(t - \sin t \cos t)(\sec \theta_0' + \sec \theta_\mathrm{v}') \tag{17}$$

$$\cos t = \frac{h}{b} \frac{\sqrt{D^2 + (\tan \theta_0' \tan \theta_\mathrm{v}' \sin \Delta\phi)^2}}{\sec \theta_0' + \sec \theta_\mathrm{v}'} \tag{18}$$

$$D = \sqrt{\tan^2 \theta_\mathrm{v}' + \tan^2 \theta_0' - 2 \tan \theta_\mathrm{v}' \tan \theta_0' \cos \Delta\phi} \tag{19}$$

$$\cos \xi' = \cos \theta_0' \cos \theta_\mathrm{v}' + \sin \theta_0' \sin \theta_\mathrm{v}' \cos \Delta\phi \tag{20}$$

$$\theta_0' = \tan^{-1}\left(\frac{b}{r} \tan \theta_0\right) \text{ and } \theta_\mathrm{v}' = \tan^{-1}\left(\frac{b}{r} \tan \theta_\mathrm{v}\right) \tag{21}$$

The $\frac{h}{b}$ and $\frac{b}{r}$ are dimensionless crown relative height and shape parameters, respectively. Both options are based on the geometric-optical mutual shadowing bidirectional reflectance distribution function model for canopies [Li and Strahler, 1992].

The combination of Ross-Thick and Li-Sparse model is suitable for many observations and therefore adopted by some prevalent algorithms, such as the Algorithm for MODIS Bidirectional Reflectance Anisotropics of the Land Surface (AMBRALS; Lucht, 2000) and MAIAC [Lyapustin et al. 2018].

To further enhance the modeling accuracy, a scaled RTLS (sRTLS) surface model has been proposed by Lyapustin et al. (2025), which constrained the model performance at high zenith angles. A hot-spot term is added to the volumetric kernel (thick option) in sRTLS model:

$$\rho_\mathrm{vol}(\theta_0, \theta_\mathrm{v}, \Delta\phi) = \frac{(\pi/2 - \xi) \cos \xi + \sin \xi}{\cos \overline{\theta_0} + \cos \overline{\theta_\mathrm{v}}} \left[1 + \left(1 + \frac{\xi}{\xi_0}\right)^{-1}\right] - \frac{\pi}{4}. \tag{22}$$

The geometric kernel (spare option) for sRTLS is given by

$$\rho_\mathrm{geo}(\theta_0, \theta_\mathrm{v}, \Delta\phi) = O(\theta_0, \theta_\mathrm{v}, t) - \sec \theta_0 - \sec \theta_\mathrm{v} + \frac{1}{2}(1 + \cos \xi') \sec \overline{\theta_\mathrm{v}} \sec \overline{\theta_0}\,. \tag{23}$$

A scaling strategy for the input angles (θ) is introduced to enhance the stability of kernels by weakening the rapid increase of $\cos\bar{\theta}$ when θ is larger than 60°. The $\cos\bar{\theta}$ is defined as

$$\cos\bar{\theta} = \cos\theta\,, \text{when } \cos\theta \geq 0.5, \tag{24}$$

$$\cos\bar{\theta} = w\cos\theta + (1-w)\sqrt{\cos\theta}\,, w = \frac{\cos\theta}{0.5}, \text{when } \cos\theta < 0.5 \tag{25}$$

The sRTLS model keeps the linear structure of the original RTLS model, which is still suitable for the conversion algorithm. It is adopted to generate MAIAC MCD19A3 dataset for global surface reflection.

2.5.2 The RPV Model and Its Variants

The RPV model captures optical and structural properties of the surface via the following expression:

$$\begin{aligned} &R_{\mathrm{RPV}}(r_{0,\lambda}, k_{\lambda}, g_{\lambda}; \theta_0, \theta_{\mathrm{v}}, \Delta\phi) \\ &\qquad = r_{0,\lambda} \cdot M(k_{\lambda}; \theta_0, \theta_{\mathrm{v}}) \cdot f_{\mathrm{HG}}(g_{\lambda}; \theta_{\mathrm{sca}}) \cdot h(r_{0,\lambda}; \theta_0, \theta_{\mathrm{v}}, \Delta\phi) \end{aligned} \tag{26}$$

where $r_{0,\lambda}$ is the spectral weight that controls the general magnitude of surface reflection, and k_{λ}, and g_{λ} are angular distribution parameters. The subscript λrepresents wavelength. Adopting the modified Minnaert function $M(k_{\lambda}; \theta_0, \theta_{\mathrm{v}})$ [Minnaert, 1941; Pinty and Ramond, 1986] to represent the overall angular shape of reflection, $M(k_{\lambda}; \theta_0, \theta_{\mathrm{v}})$ is expressed as

$$M(k_{\lambda}; \theta_0, \theta_{\mathrm{v}}) = \frac{\mu_{\mathrm{v}}^{k_{\lambda}-1}\mu_0^{k_{\lambda}-1}}{(\mu_{\mathrm{v}} + \mu_0)^{1-k_{\lambda}}} \tag{27}$$

where μ_0 is the cosine of solar zenith angle (θ_0); μ_{v} is the cosine of view zenith angle (θ_{v}). The model parameter k_{λ} determines whether the reflectance factors increase or decrease with increasing view zenith angle. When $k < 1$, the angular shape of BRF exhibits a generally “bowl” shape; when $k > 1$, the angular shape of BRF exhibits a generally “bell” shape; and when $k = 1$ the behavior is isotropic.

The $f_{\mathrm{HG}}(g; \theta_{\mathrm{sca}})$ term is the Henyey-Greenstein phase function [Henyey and Greenstein, 1941]. By deploying the asymmetry parameter g, it determines the dominance of either forward ($g > 0$) or backward scattering ($g < 0$) in the reflection pattern:

$$f_{\mathrm{HG}}(g_\lambda; \theta_{\mathrm{sca}}) = \frac{1 - {g_\lambda}^2}{(1 + 2g_\lambda \cos\theta_{\mathrm{sca}} + {g_\lambda}^2)^{1.5}} \tag{28}$$

where the scattering angle θ_{sca} is calculated from μ_0, μ_{v}, and $\Delta\phi$:

$$\cos\theta_{\mathrm{sca}} = -\mu_{\mathrm{v}}\,\mu_0 + \sqrt{(1 - \mu_{\mathrm{v}}^2)(1 - \mu_0^2)}\cos\Delta\phi \tag{29}$$

Additionally, the $h\left(r_{0,\lambda};\ \theta_0, \theta_{\mathrm{v}}, \Delta\phi\right)$ term describes the “hot spot,” which is a peaked increase in reflectance factor in the backscatter direction [Rahman et al., 1993]:

$$h(r_{0,\lambda};\ \theta_0, \theta_{\mathrm{v}}, \phi_{\mathrm{v}}, \phi_0) = 1 + R(G) = 1 + \frac{1 - r_{0,\lambda}}{1 + G} \tag{30}$$

where G is expressed as:

$$G(\theta_0, \theta_{\mathrm{v}}, \Delta\phi) = \sqrt{\tan^2\theta_{\mathrm{v}} + \tan^2\theta_0 + 2\tan\theta_{\mathrm{v}}\tan\theta_0\ \cdot\cos\Delta\phi} \tag{31}$$

As a variant of the RPV model, mRPV model uses the exponential term $e^{b_\lambda \cdot \theta_{\mathrm{sca}}}$ to describe asymmetry of angular distribution of BRF [Martonchik et al., 1998; Engelsen et al., 1996]:

$$R_{\mathrm{mRPV}}(r_{0,\lambda}, k_\lambda, b_\lambda; \theta_0, \theta_{\mathrm{v}}, \Delta\phi) = a_{0,\lambda} \cdot M(k_\lambda; \theta_0, \theta_{\mathrm{v}}) \cdot e^{b_\lambda \cdot \theta_{\mathrm{sca}}} \cdot h(r_{0,\lambda};\ \theta_0, \theta_{\mathrm{v}}, \Delta\phi) \tag{32}$$

where b_λ replaces g_λ as the corresponding model parameter.

Exclusion of the hotspot term in the Eq. (26) for the RPV model and Eq. (32) for the mRPV model leads to degenerate forms of RPV and mRPV models, denoted as d-RPV and d-mRPV, respectively:

$$R_{\mathrm{d-RPV}}(a_{0,\lambda}, k_{\lambda}, g_{\lambda}; \theta_0, \theta_{\mathrm{v}}, \phi_{\mathrm{v}}, \phi_0) = a_{0,\lambda} \cdot M(k_{\lambda}; \theta_0, \theta_{\mathrm{v}}) \cdot f_{\mathrm{HG}}(g_{\lambda}; \theta_{\mathrm{sca}}) \tag{33}$$

$$R_{\mathrm{d-mRPV}}\left(a_{0,\lambda}, k_{\lambda}, b_{\lambda}; \theta_0, \theta_{\mathrm{v}}, \phi_{\mathrm{v}}, \phi_0\right) = a_{0,\lambda} \cdot M(k_{\lambda}; \theta_0, \theta_{\mathrm{v}}) \cdot e^{b_{\lambda} \cdot \theta_{\mathrm{sca}}}. \tag{34}$$

As an advantage of degenerate forms, terms of d-mRPV and d-RPV model are linear and quasi-linear in logarithmic space.

2.5.3 Polarization of Surface Reflectance

In addition to the depolarizing component of surface reflection, the polarizing component of surface reflection is described by a polarized BRF (pBRF) model [Cox and Munk, 1954a; 1954b; Diner et al., 2012]:

$$\mathbf{R}_{\mathrm{surf,p}} = \varepsilon \frac{\boldsymbol{F}_{\mathrm{p}}\,(n_{\mathrm{surf}}\,; \theta_{\mathrm{sca}})}{4\mu_{\mathrm{n}}^4\,(\cos\theta_0 + \cos\theta_{\mathrm{v}})} \frac{1}{v_{\mathrm{s}}} \exp\left(-\frac{1-\mu_{\mathrm{n}}^2}{\mu_{\mathrm{n}}^2\, v_{\mathrm{s}}}\right) f_{\mathrm{sh}}(\theta_{\mathrm{sca}}) \tag{35}$$

where

$$f_{\mathrm{sh}}(\theta_{\mathrm{sca}}) = \left[\frac{1+\cos\left[k_{\gamma}(\pi-\theta_{\mathrm{sca}})\right]}{2}\right]^3, \tag{36}$$

$$\mu_{\mathrm{n}} = \frac{\mu_{\mathrm{v}}+\mu_0}{2}\left[\cos\left(\frac{\pi-\theta_{\mathrm{sca}}}{2}\right)\right]^{-1}, \tag{37}$$

In the above equation, ε accounts for the weight of polarized BRF, $\boldsymbol{F}_{\mathrm{p}}$ is the Fresnel reflection matrix [Hansen and Travis, 1974], n_{surf} is the refractive index of land surface (set to 1.5 in the current study), v_{s} is the slope variance, $f_{\mathrm{sh}}(\Omega)$ is the function account for the shadowing effect, k_{γ}

controls the width of the shadowing function, and μ_n is the cosine of tilt angle of the facet surface normal for a particular illumination and view geometry.

3. Surface a priori and Conversion of Surface Model Parameters

A major challenge of aerosol remote sensing over land using multi-angle and multi-spectral radiometry is to accurately model or correct for the contribution of anisotropic bidirectional surface reflection. When aerosol loadings are low or when the surface is bright, signals from aerosol scattering are typically less significant than those from the surface. This effect is especially pronounced in the satellite-borne scenario, where observations are made at the TOA and surface signals tend to dominate.

In aerosol retrieval, this surface dominance manifests as a nonlinear coupling between aerosol and surface components through radiative transfer processes, making accurate separation of the two particularly difficult. Reliable a priori surface reflectance information is therefore desirable to constrain and initialize surface modeling within the retrieval framework. To this end, a dedicated dataset encoding surface reflection properties is needed to serve as this initialization.

Suggested by the MAIA mission, the candidate dataset selected for the surface a priori is MAIAC. MAIAC surface reflectance product MCD19A3 provides spectral surface BRF parameters at 1 km spatial resolution, derived from MODIS measurements [King et al., 1992]. The MAIAC dataset includes coefficients for the isotropic, geometric, and volumetric scattering kernels of the sRTLS model, which together can generate an initial estimate of surface BRF to inform the retrieval. In MAIA's Level 2 aerosol retrieval framework, MAIAC MCD19A3 products serve as a priori parameters within an optimal estimation approach to characterize the radiometric contribution of surface reflection. This implementation further motivates us to incorporate MAIAC into the CIMAP retrieval framework.

As MAIAC surface products are derived using the Ross-Li model and some aerosol retrievals use the RPV formulation, differences in surface parameterization across retrieval systems complicate the transfer and reuse of surface reflectance constraints. To address this issue, we develop a computationally efficient algorithm that converts Ross-Li BRF parameters into their RPV counterparts. In principle, the conversion can be achieved by fitting BRFs generated from the Ross-Li model parameters to derive RPV model parameters via suitably initialized nonlinear least squares optimization. In this dissertation, we introduce a more efficient and stable way of achieving this conversion that capitalizes on the RPV model's quasi-linearity in logarithmic space and does not require any parameter initialization. Conversions from Ross-Li to both RPV and the mRPV were explored. However, the conversion introduced in this work is intended as a numerical mapping between the Ross–Li and RPV formulations in BRF space, rather than implying physical equivalence between the two surface models. Although the converted RPV parameters basically reproduce the BRF angular behavior and preserve hemispherical reflectance, their physical interpretation should not be considered identical to that of parameters in the original RPV formulation. Instead, the converted parameters primarily serve as an effective numerical representation that enables interoperability between the two BRF models within retrieval frameworks.

3.1 Quality Evaluation of MAIAC Surface Reflectance Data

Prior to implementation into the retrieval framework, a quality assessment of MAIAC MCD19A3 products is conducted to evaluate the angular and spectral quality of the reconstructed surface reflection properties. To this end, a cross-validation is performed against MISR Level 2 land surface products (MIL2ASLS; Diner et al., 1999), which provides model parameters for the mRPV model.

The study region selected for this assessment is the southwest coast area of the United States (SWCAUS), which also serves as a primary target area (PTA) for MAIA aerosol research. A total of 12 collocated cases are drawn from this region: pixels 1–3 are located along the western edge of the Mojave Desert, pixels 4–6 in the Coast Range adjacent to the Central Valley, pixels 7–9 spanning a north–south transect across the Sierra Nevada, and pixels 10–12 in the Great Basin Desert. The collocation is done on 03/13/2017. Figure 1 presents the map of the SWCAUS region with the 12 collocated cases.

Based on these collocations, two surface reflection properties are computed from both MAIAC and MISR products for comparison: spectral surface directional-hemispherical reflectance (DHR) and spectral invariance.

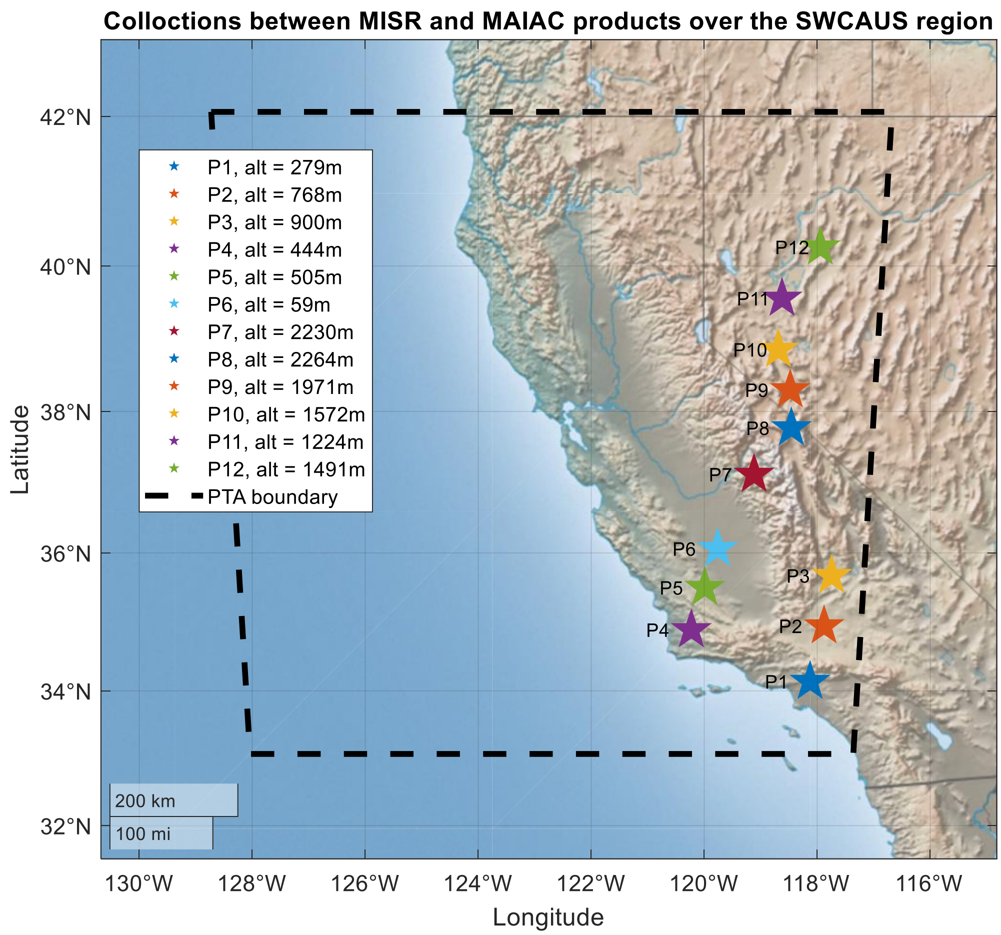


Figure 1. Geolocations of collocated pixels in SWCAUS region. 12 pixels are marked with star symbols, with their altitudes labeled. The black dashed line delimits the SWCAUS region, which is also one of the primary target areas of the MAIA mission.

3.1.1 Comparisons of Spectral DHR

The DHR or surface albedo, equals to the ratio of reflected flux to the direct solar irradiance, ignoring the diffuse skylight component [Wanner et al., 1997; Lucht et al., 2000]. It can be computed by integrating the surface BRF R_{model} over the entire hemisphere, namely

$$w_{0,\text{model}} = \frac{1}{\pi}\int_0^{2\pi}\int_0^{\pi/2} R_{\text{model}} \cos\theta_{\text{v}} \sin\theta_{\text{v}}\, d\theta_{\text{v}}\, d(\Delta\phi). \tag{38}$$

DHR represents the portion of surface-reflected signals that directly influence aerosol property retrieval. An accurate initial estimate of surface DHR minimizes errors in the retrieved aerosol properties and accelerates convergence across retrieval iterations.

Figure 2 demonstrates comparisons of surface DHR derived from MISR and MAIAC surface products for 12 collocated pixels over the SWCAUS region at a solar zenith angle (SZA) of 10°. Figure 3 and Figure 4 show the same comparisons at SZAs of 45° and 60°, respectively. MISR DHR is presented at 446, 558, 672, and 867 nm, while MAIAC DHR is presented at 465.5, 553.5, 644.9, and 855.6 nm. Strong agreement is found for most pixels; however, pixel 7 shows a notable discrepancy. Pixel 7 is located in the Sierra Nevada Mountain range (Lat: 37.156°, Lon: -119.057°) at an elevation of 2,230 m, and was possibly snow-covered at the time of collocation (03/13/2017). The elevated surface DHR is captured by the MISR product, but not by MAIAC. This discrepancy is likely attributable to differences in the retrieval principles underlying the two products. MISR surface products are derived simultaneously with the observations, making them more likely to capture near-real-time land surface conditions. In contrast, MAIAC surface products are derived from an 8-day average of observations, which may dilute short-term amplifications of DHR caused by snow cover. Overall, the strong agreement between MISR and MAIAC across most pixels suggests that the MAIAC dataset is sufficiently reliable for use in a priori generation. Nevertheless, caution should be exercised when applying MAIAC products in regions where surface properties change rapidly.

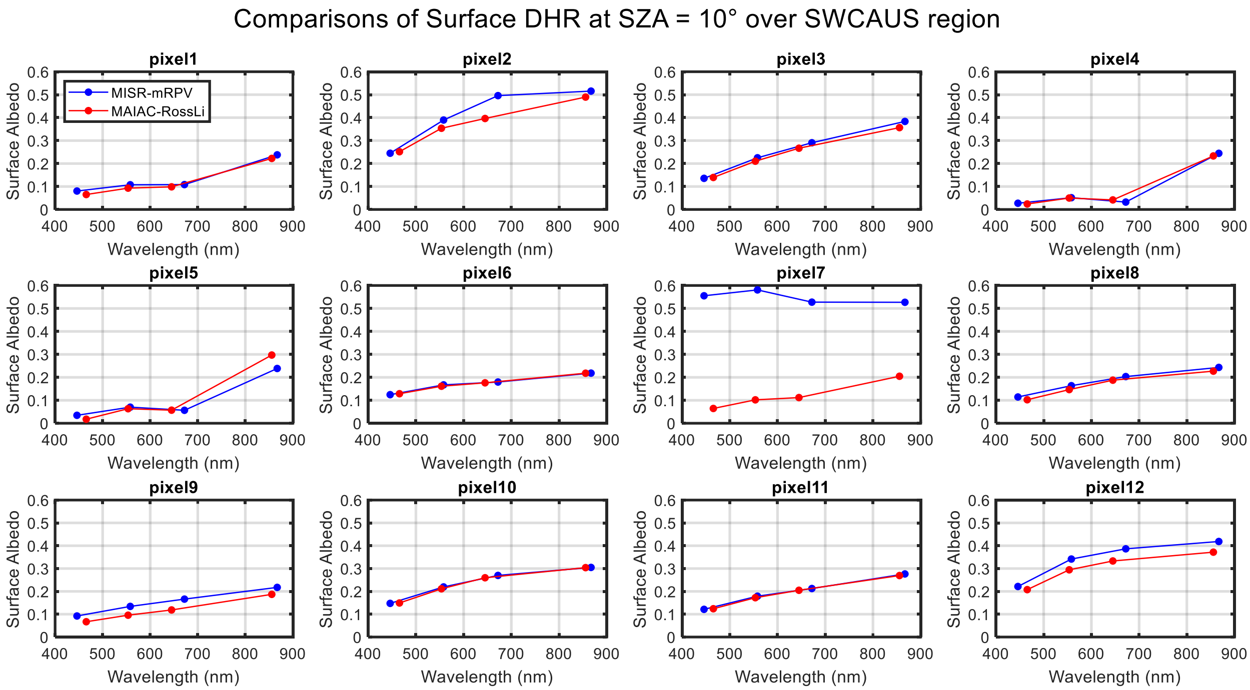


Figure 2. Comparisons of surface DHR across 12 collocated pixels over SWCAUS region at SZA = 10°. MISR DHR (blue dotted lines) are presented at 446, 558, 672, and 867 nm, while MAIAC DHR (red dotted lines) is presented at 465.5, 553.5, 644.9, and 855.6 nm.

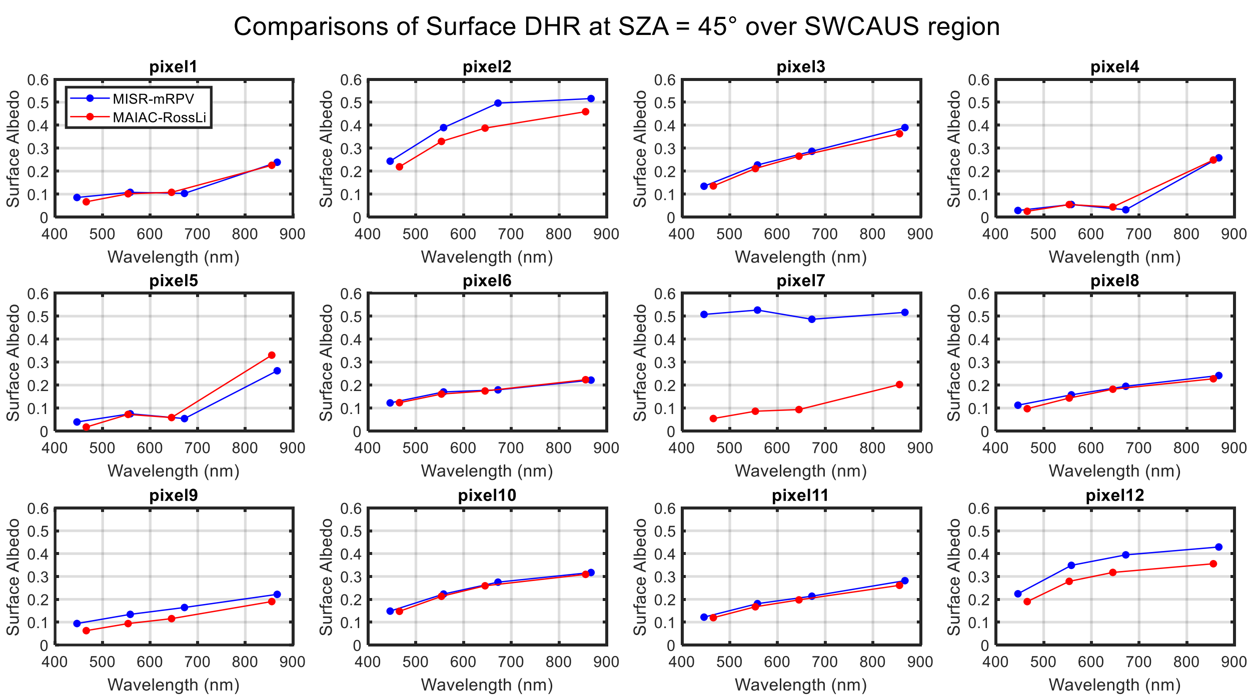


Figure 3. Comparisons of surface DHR across 12 collocated pixels over SWCAUS region at SZA = 45°. MISR DHR (blue dotted lines) are presented at 446, 558, 672, and 867 nm, while MAIAC DHR (red dotted lines) is presented at 465.5, 553.5, 644.9, and 855.6 nm.

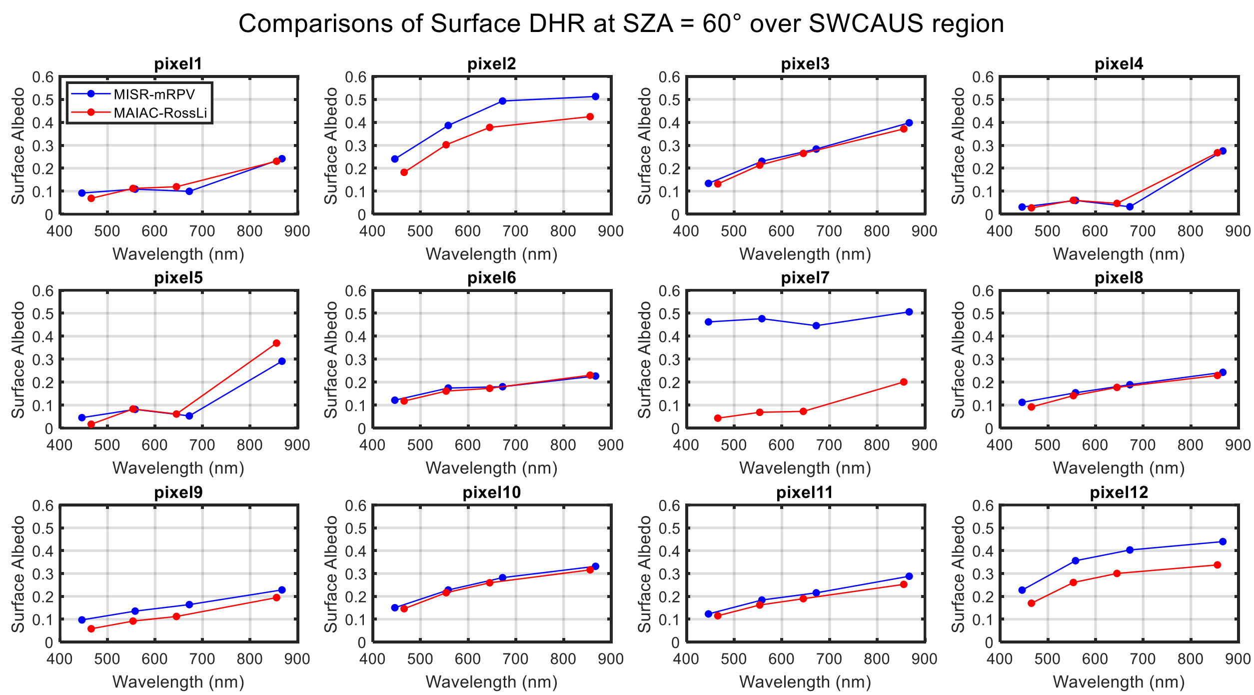


Figure 4. Comparisons of surface DHR across 12 collocated pixels over SWCAUS region at SZA = 60°. MISR DHR (blue dotted lines) are presented at 446, 558, 672, and 867 nm, while MAIAC DHR (red dotted lines) is presented at 465.5, 553.5, 644.9, and 855.6 nm.

3.1.2 Comparisons of Spectral Invariance of Surface BRF

In MISR's operational surface retrieval [Diner et al., 1999], the surface directional reflectance factor is assumed to be nearly spectrally invariant across view angles. This assumption arises because surface scattering objects are large relative to the observed wavelengths, in contrast to atmospheric scattering, where aerosol particle sizes are comparable to the wavelength of light [Flowerdew and Haigh, 1995]. As a consequence of this spectral invariance, the magnitude and shape of the normalized spectral BRF are nearly identical across visible wavelengths, meaning the BRF angular structure is largely independent of wavelength.

This assumption carries significant implications for aerosol retrieval. Since the spectral scattering behavior of the surface is similar across wavelengths, the corresponding model parameters can be treated as spectrally shared rather than independently retrieved at each wavelength. This reduces the number of unknowns in the state vector, which in turn mitigates the ill-posedness of the inversion problem and improves retrieval stability.

Building on this, this assessment compares MAIAC and MISR surface products using normalized BRF to evaluate the degree of spectral invariance across the selected wavelengths. The normalization is performed by dividing the BRDFs of the MISR mRPV model and the MAIAC sRTLS model by their respective spectral weights, as given by:

$$R_{\mathrm{MISR\_shape}} = \frac{R_{\mathrm{MISR}}}{a_{0,\lambda}} = M(k_\lambda; \theta_0, \theta_\mathrm{v}) \cdot e^{b_\lambda \cdot \theta_{\mathrm{sca}}} \tag{39}$$

$$R_{\mathrm{MAIAC\_shape}} = \frac{R_{\mathrm{MAIAC}}}{f_{\mathrm{iso}}} = 1 + \frac{f_{\mathrm{vol}} \cdot \rho_{\mathrm{vol}}(\theta_0, \theta_\mathrm{v}, \Delta\phi) + f_{\mathrm{geo}} \cdot \rho_{\mathrm{geo}}(\theta_0, \theta_\mathrm{v}, \Delta\phi)}{f_{\mathrm{iso}}}. \tag{40}$$

This normalization isolates the relative angular shapes of the BRF across all view angles. One simplification introduced for the MISR BRF calculation is that the mRPV model is replaced by the d-mRPV model. This is necessary because the spectral weight in the mRPV model is embedded within the hotspot term, where simple normalization cannot fully remove the spectral ratio introduced by $a_{0,\lambda}$. The d-mRPV model circumvents this issue by separating the spectral and angular components, allowing a cleaner normalization.

Figure 5–7 present comparisons of normalized BRF across 12 collocated pixels derived from MAIAC and MISR surface products. Figure 5 shows results in the principal plane, while Fig. 6 and Fig. 7 show comparisons at 45° and 60° off the principal plane, respectively. In each subplot, colored lines represent the normalized BRF of individual pixels, with an offset of 0.5 applied to each successive pixel starting from the second to prevent overlap and improve visual clarity. Error bars at each view angle represent the standard deviation of BRF across the four wavelengths, serving as a direct indicator of spectral invariance: smaller error bars indicate that the normalized BRF is more consistent across wavelengths, while larger error bars suggest stronger spectral dependence. For both products and across all three scattering planes, solar zenith angles of 0°, 45°, and 60° are compared.

Across all scattering planes (Figs. 5-7), the results reveal a consistent pattern of differences between the two products in terms of both spatial variability and spectral invariance. Regarding spatial variability, certain MISR pixels (e.g., pixels 1, 4, 5, and 7) exhibit notably different normalized BRF magnitudes and angular shapes compared to the remaining pixels, indicating significant spatial heterogeneity in the MISR surface products. In contrast, MAIAC products display more uniform magnitudes and angular shapes across all pixels, suggesting that MAIAC surface retrievals are spatially smoother and relatively less sensitive to local surface variations. This difference likely reflects the distinct retrieval strategies of the two products: as discussed earlier, MISR retrievals are observation-specific and thus more responsive to instantaneous surface conditions, whereas MAIAC retrievals are temporally averaged, which suppresses spatial variability.

Regarding spectral invariance, the MAIAC normalized BRF consistently exhibits smaller error bars across view angles compared to MISR, as clearly illustrated in pixels 4 and 5. This indicates that the MAIAC BRF maintains more similar magnitudes across wavelengths, demonstrating stronger spectral invariance. Conversely, the larger error bars in MISR suggest greater spectral variation in its normalized BRF.

Taken together, these results suggest that while MISR products are better suited to capturing spatially and temporally specific surface conditions, MAIAC products exhibit stronger spectral invariance and greater spatial consistency across pixels. Given that spectral invariance is a key assumption underlying aerosol retrieval, the stronger spectral invariance of MAIAC normalized BRF supports its suitability as a reliable a priori surface input. However, the reduced spatial variability in MAIAC products should be noted as a potential limitation in heterogeneous or rapidly changing surface conditions, consistent with the findings discussed in the DHR comparison above.

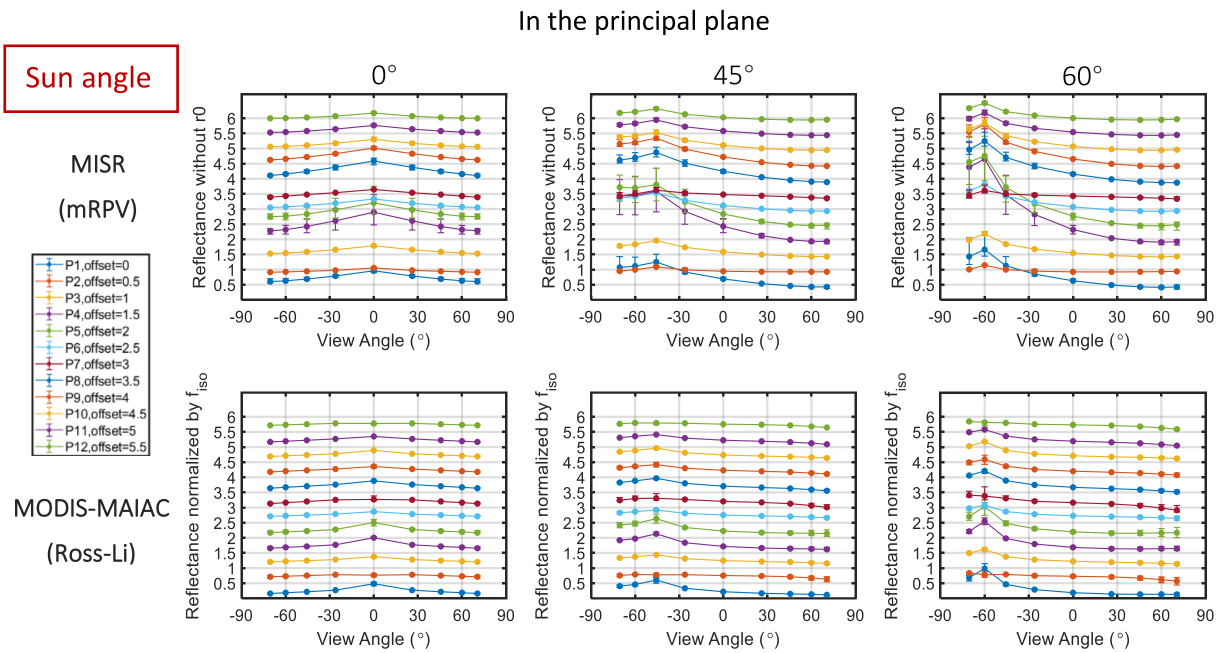


Figure 5. Comparisons of normalized BRF across 12 collocated pixels derived from MAIAC and MISR surface products at SZA = 0°, 45°, and 60° in the principal plane. Colored lines represent the normalized BRF of individual pixels, with an offset of 0.5 applied to each successive pixel. Error bars at each view angle represent the standard deviation of BRF across the four wavelengths.

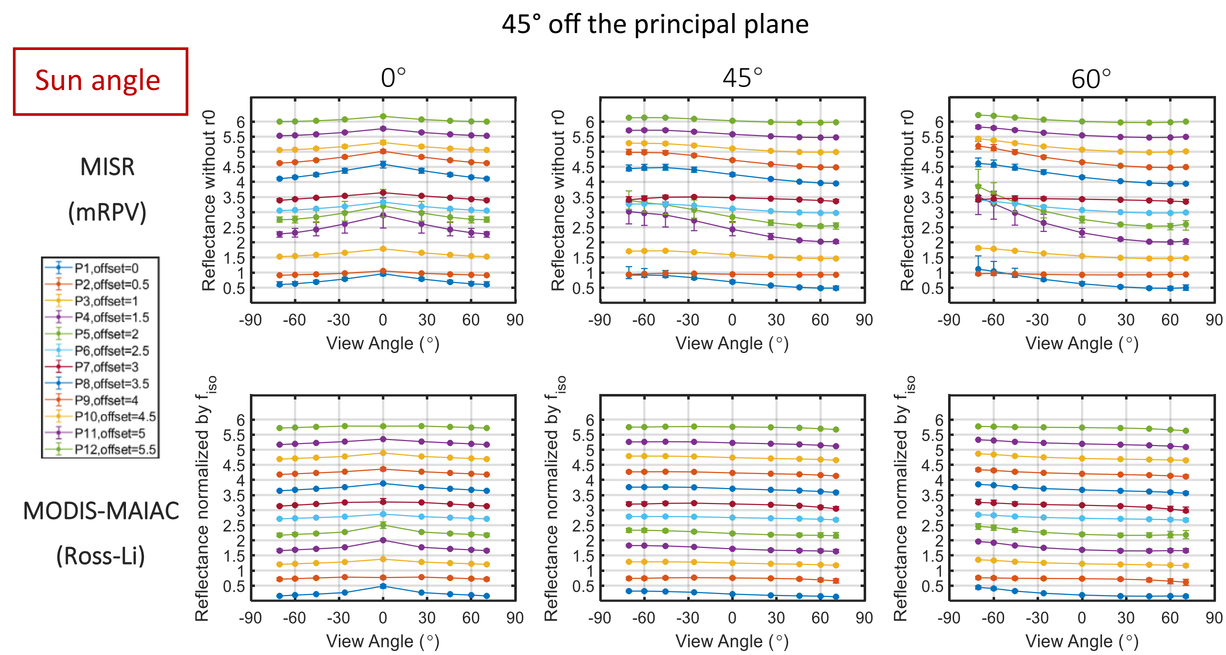


Figure 6. Comparisons of normalized BRF across 12 collocated pixels derived from MAIAC and MISR surface products at SZA = 0°, 45°, and 60°, 45° off the principal plane. Colored lines represent the normalized BRF of individual pixels, with an offset of 0.5 applied to each successive pixel. Error bars at each view angle represent the standard deviation of BRF across the four wavelengths.

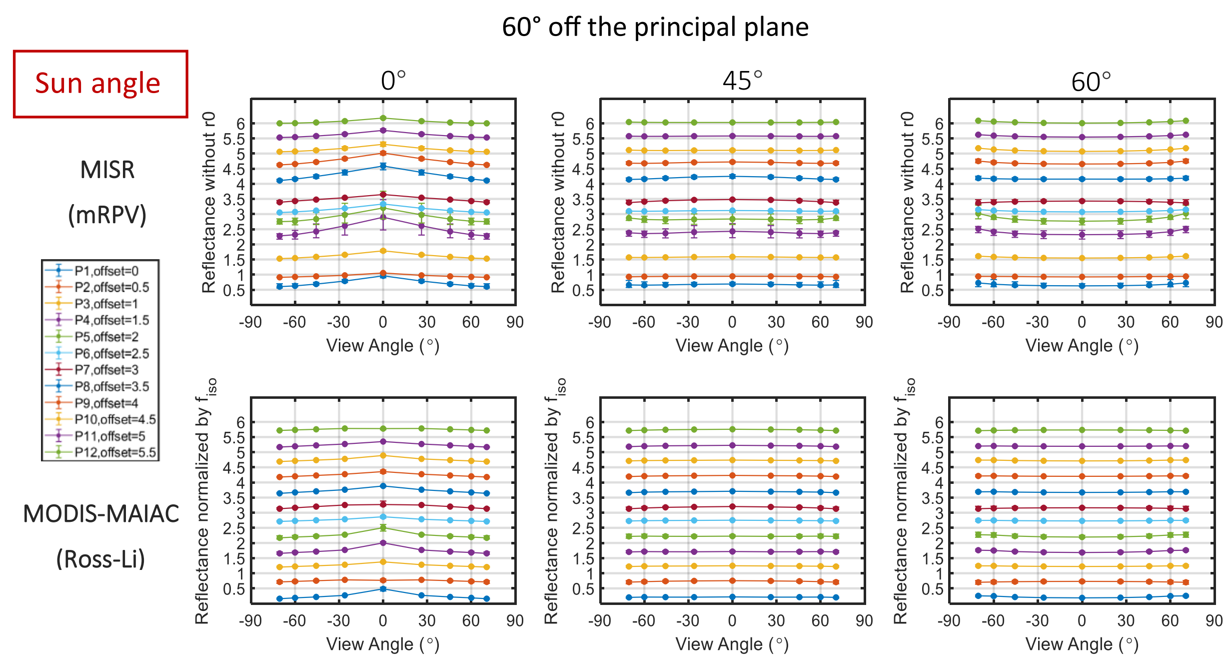


Figure 7. Comparisons of normalized BRF across 12 collocated pixels derived from MAIAC and MISR surface products at SZA = 0°, 45°, and 60°, 60° off the principal plane. Colored lines represent the normalized BRF of

individual pixels, with an offset of 0.5 applied to each successive pixel. Error bars at each view angle represent the standard deviation of BRF across the four wavelengths.

3.1.3 Summary

The above tests demonstrate that the MAIAC dataset is of high quality and suitable for use as a surface a priori for aerosol retrieval, consistent with expectations from the analysis of the MAIA mission. In the spectral DHR tests, MAIAC shows strong agreement with the MISR reference, while in the spectral invariance tests, MAIAC exhibits a high degree of spectral invariance. However, due to the underlying data generation principles of MAIAC, its surface products may not capture rapid variations in regions where surface properties change quickly, introducing potential uncertainty in retrievals over such surface scenarios.

3.2 Model parameter conversion

In this section, the conversion algorithm for surface model parameters is introduced. The general structure of our algorithm is shown in Fig. 8. Two options are described for the RPV model and its variants, with (Option 1) or without (Option 2) requiring the angular BRF shape to be spectrally independent [Diner et al., 2005; 2011], while both options share the same algorithm structure. Specifically, Option 1 allows both spectral weight and angular shape parameters to depend on wavelength. Option 2 only allows the spectral weight to have wavelength dependence. For both Option 1 and 2, four RPV variants are discussed: d-mRPV, d-RPV, mRPV, and original RPV. The conversion starts with simulating surface BRF for a variety of view geometries and a given solar zenith angle, using Ross-Li model parameters informed from MAIAC. The details of paths and options are introduced in the following subsections.

It should be noted that in strict sense, surface BRF model parameters are intrinsic surface properties and are not expected to depend on illumination or viewing geometry. In practice, however, satellite observations such as those from MISR and MODIS are acquired under a given solar zenith angle. Surface BRF model conversion is therefore performed separately for the prescribed illumination condition. For each solar zenith angle, a corresponding set of effective RPV parameters is obtained such that the converted model reproduces the RTLS BRF under the same illumination condition. This procedure does not introduce any additional degree of freedom in the RPV model. Rather, the purpose is to provide an illumination-conditioned parameterization that improves the representation of the angular reflectance structure for the geometry of interest. Such a parameterization is then used as an initial guess to be further refined, together with the optimization of aerosol properties.

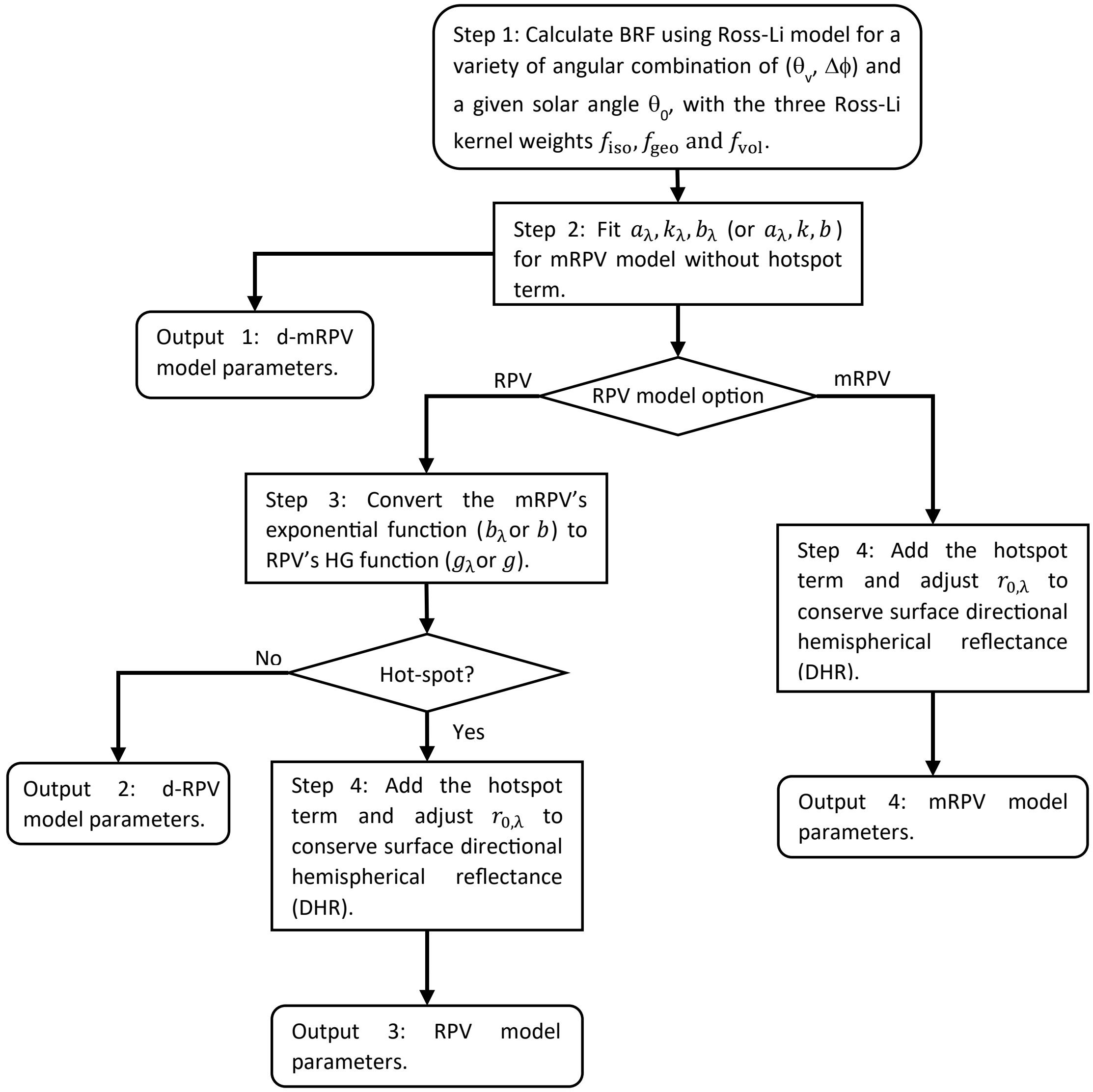


Figure 8. Algorithm flow chart for converting Ross-Li to RPV model parameters. Note that Option 1 (k_λ and b_λ are wavelength dependent) and option 2 (k and b are wavelength independent) share the same algorithm structure. The surface hemispherical reflectance (DHR) is computed by integrating the surface BRF "R_{model}" over the entire hemisphere, namely $\frac{1}{\pi}\int_0^{2\pi}\int_0^{\pi/2} R_{\mathrm{model}} \cos\theta_{\mathrm{v}} \sin\theta_{\mathrm{v}}\, d\theta_{\mathrm{v}}\, d(\Delta\phi)$.

3.2.1 Conversion of degenerate forms of RPV-based models

First, we demonstrate the conversion to d-mRPV model. For a given sun angle θ_0, the surface BRF is calculated by using Ross-Li model for a set of observation geometries that combine N_{v} view zenith angles between 0º and 65º and N_{az} relative azimuth angles between 30º and 165°. The coverage excludes an angular region near the hotspot (backscatter) direction, which occurs when the viewing direction is approximately 180º from the illumination direction (e.g., $165º \leq \theta_{\text{sca}} \leq 180º$). The $N_{\text{v}} \times N_{\text{az}}$ cases are then used to solve for the d-mRPV model parameters $a_{0,\lambda}$, k_{λ}, and b_{λ} by performing linear least-squares fitting in logarithmic space:

$$\ln(\mathrm{R}_{\text{Ross-Li}}) = \ln(\mathrm{R}_{\text{d-mRPV}}) = \ln(a_{0,\lambda}) + (k_{\lambda} - 1)\ln[(\mu_{\text{v}}\mu_0)(\mu_{\text{v}} + \mu_0)] + b_{\lambda} \cdot \theta_{\text{sca}} \quad (41)$$

The MAIAC parameters of the Ross-Li model are spectrally dependent. We first demonstrate the Ross-Li to d-mRPV conversion that also allows $a_{0,\lambda}$, k_{λ}, and b_{λ} to be spectrally dependent (Option 1). These parameters are obtained by solving Eq. (41) in the following matrix form:

$$\mathbf{R}_{\lambda} = \mathbf{A}\,\mathbf{X}_{\lambda} \quad (42)$$

where $\mathbf{R}_{\lambda}$ is a $N \times 1$ column vector that contains the natural logarithm of the BRF calculated using the Ross-Li model for a variety of view geometries; the $N \times 3$ $\mathbf{A}$ matrix is calculated from the geometric basis functions of the d-mRPV model; $\mathbf{X}_{\lambda}$ is the 3×1 solution vector; and $N = N_{\text{v}} \times N_{\text{az}}$. The matrices $\mathbf{R}_{\lambda}$ and $\mathbf{A}$ are constructed from the N combinations of view zenith angles and relative azimuth angles for each wavelength. Specifically, let the submatrix $\mathbf{R}_{\lambda,i}$ be written as

$$\mathbf{R}_{\lambda,i} = \begin{bmatrix} \ln\left[R_{\text{Ross-Li},\lambda}(\theta_0, \theta_i, \Delta\phi_1)\right] \\ \vdots \\ \ln\left[R_{\text{Ross-Li},\lambda}\left(\theta_0, \theta_i, \Delta\phi_{N_{\text{az}}}\right)\right] \end{bmatrix} \quad (43)$$

Then incorporating the calculated BRFs at all N_v view angles, we have

$$\mathbf{R}_{\lambda} = \begin{bmatrix} \mathbf{R}_{\lambda,1} \\ \vdots \\ \mathbf{R}_{\lambda,N_{\text{v}}} \end{bmatrix} \quad (44)$$

Similarly, $\mathbf{A}$ is composed of submatrices $\mathbf{A}_i$ defined as

$$\mathbf{A}_i = \begin{bmatrix} 1 & \ln[(\mu_{v,i}\mu_0)(\mu_{v,i}+\mu_0)] & \theta_{sca}(\mu_{v,i},\mu_0,\Delta\phi_1) \\ & \vdots & \\ 1 & \ln[(\mu_{v,i}\mu_0)(\mu_{v,i}+\mu_0)] & \theta_{sca}(\mu_{v,i},\mu_0,\Delta\phi_{N_{az}}) \end{bmatrix} \tag{45}$$

from which the overall matrix **A** can be obtained from

$$\mathbf{A} = \begin{bmatrix} \mathbf{A}_1 \\ \vdots \\ \mathbf{A}_{N_v} \end{bmatrix} \tag{46}$$

Finally, the solution vector $\mathbf{X}_\lambda$ is

$$\mathbf{X}_\lambda = \begin{bmatrix} \ln(a_{0,\lambda}) \\ k_\lambda - 1 \\ b_\lambda \end{bmatrix} \tag{47}$$

The solution to Eq. (47) for each wavelength is obtained by taking the pseudoinverse:

$$\mathbf{X}_\lambda = (\mathbf{A}^{\mathrm{T}}\mathbf{A})^{-1}\mathbf{A}^{\mathrm{T}}\mathbf{R}_\lambda \tag{48}$$

Note that the least-squares system defined by the above equation is overdetermined, as the number of BRF samples across viewing geometries greatly exceeds the number of unknown parameters, which contributes to numerical stability in the parameter estimation. In addition, BRF values are ensured to be positive within the angular domain considered, so that the logarithmic transformation does not introduce numerical singularities.

Option 2 allows only the spectral weight $a_{0,\lambda}$ to be spectrally dependent whereas the angular model parameters k and b are wavelength independent. To this end, the linear system of equations to be solved for $(a_{0,\lambda}, k, b)$ includes all wavelengths simultaneously. Letting M be the number of wavelengths, the system takes the following matrix form:

$$\mathbf{R} = \mathbf{A}\,\mathbf{X} \tag{49}$$

where **R** is a (M×N)×1 column vector that contains the natural logarithm of the BRF calculated using the Ross-Li model for a variety of view geometries and wavelengths; **A** is a (M×N)×(M+2) matrix calculated from the geometric basis functions of the d-mRPV model; and $\mathbf{X}_\lambda$ is the (M+2)×1 solution vector. The matrices **R** and **A** are constructed from the N combinations of view zenith angles and relative azimuth angles for all M wavelengths. To this end, all elements in Eq. (43) need to be updated to be a vector so that the equation is re-written as

$$\mathbf{R}_i = \begin{bmatrix} \ln[\mathbf{R}_{\text{Ross-Li}}(\theta_0, \theta_i, \Delta\phi_1)] \\ \vdots \\ \ln\left[\mathbf{R}_{\text{Ross-Li}}\left(\theta_0, \theta_i, \Delta\phi_{N_{\text{az}}}\right)\right] \end{bmatrix} \tag{50}$$

where the column vector $\mathbf{R}_{\text{Ross-Li}}(\theta_0, \theta_i, \Delta\phi_1)$ includes all wavelengths, namely

$$\mathbf{R}_{\text{Ross-Li}}\left(\theta_0, \theta_i, \Delta\phi_j\right) = \begin{bmatrix} \mathbf{R}_{\text{Ross-Li},\lambda_1}\left(\theta_0, \theta_i, \Delta\phi_j\right) \\ \vdots \\ \mathbf{R}_{\text{Ross-Li},\lambda_M}\left(\theta_0, \theta_i, \Delta\phi_j\right) \end{bmatrix} \tag{51}$$

Incorporating all view angles, we have the complete vector

$$\mathbf{R} = \begin{bmatrix} \mathbf{R}_1 \\ \vdots \\ \mathbf{R}_{N_{\text{v}}} \end{bmatrix} \tag{52}$$

Matrix **A** is constructed by first establishing the submatrix to account for the wavelength-dependent spectral weights and wavelength-independent angular shape parameters:

$$\mathbf{A}_{i,j} = \begin{bmatrix} 1 & 0 & \cdots & 0 & \ln\left[\left(\mu_{\text{v},i}\mu_0\right)\left(\mu_{\text{v},i} + \mu_0\right)\right] & \theta_{\text{sca}}\left(\mu_{\text{v},i}, \mu_0, \Delta\phi_j\right) \\ 0 & 1 & \cdots & 0 & \ln\left[\left(\mu_{\text{v},i}\mu_0\right)\left(\mu_{\text{v},i} + \mu_0\right)\right] & \theta_{\text{sca}}\left(\mu_{\text{v},i}, \mu_0, \Delta\phi_j\right) \\ \vdots & \vdots & \vdots & & \vdots & \\ 0 & \cdots & 1 & 0 & \ln\left[\left(\mu_{\text{v},i}\mu_0\right)\left(\mu_{\text{v},i} + \mu_0\right)\right] & \theta_{\text{sca}}\left(\mu_{\text{v},i}, \mu_0, \Delta\phi_j\right) \\ 0 & \cdots & 0 & 1 & \ln\left[\left(\mu_{\text{v},i}\mu_0\right)\left(\mu_{\text{v},i} + \mu_0\right)\right] & \theta_{\text{sca}}\left(\mu_{\text{v},i}, \mu_0, \Delta\phi_j\right) \end{bmatrix} \tag{53}$$

Next, the matrices $\mathbf{A}_i$ are constructed as

$$\mathbf{A}_i = \begin{bmatrix} \mathbf{A}_{i,1} \\ \vdots \\ \mathbf{A}_{i,N_{\text{az}}} \end{bmatrix} \tag{54}$$

from which the complete matrix $\mathbf{A}$ is given by Eq. (46), using Eq. (54) as the formulation for $\mathbf{A}_1, \ldots, \mathbf{A}_{N_\text{v}}$.

Finally, the solution vector is expressed as

$$\mathbf{X} = \begin{bmatrix} \ln(a_{0,\lambda_1}) \\ \ln(a_{0,\lambda_2}) \\ \vdots \\ \ln(a_{0,\lambda_M}) \\ k-1 \\ b \end{bmatrix} \tag{55}$$

The solution for $\mathbf{X}$ follows the same methodology described by Eq. (48).

3.2.2 Conversion to d-RPV model

Following the conversion of Ross-Li to d-mRPV model parameters, additional steps are included to achieve the conversion to d-RPV model parameters. Recall that the d-RPV model can be expressed by Eq. (33). It can be observed that the mRPV and RPV models share the same components except for the functions of scattering, namely $\exp(b_\lambda \cdot \theta_{\text{sca}})$ for mRPV vs. $f_{\text{HG}}(g_\lambda; \theta_{\text{sca}})$ for RPV. Once the conversion from Ross-Li to mRPV is obtained, the optimal value of g_λ is determined by fitting $\exp(b_\lambda \cdot \theta_{\text{sca}})$ to find the value of g_λ that minimizes the differences between the two functions of scattering angle:

$$\min_{g_\lambda} \sum_{i=1}^{N} \left| f_{\text{HG}}(g_\lambda; \theta_{\text{sca},i}) - e^{b_\lambda \cdot \theta_{\text{sca},i}} \right|^2 \tag{56}$$

Applying Eq. (56) to each wavelength provides the d-RPV model parameter g_λ for Option 1. Modifying Eq. (56) by dropping the wavelength dependence provides the solution of g for Option 2.

3.2.3 Adding Hot Spot Function into the RPV and mRPV Models

The hotspot has the maximum contribution when the viewing direction equals to the illumination direction. In this case, $\theta_\mathrm{v} = \theta_0$ and the relative azimuth angle $\Delta\phi = \phi_\mathrm{v} - \phi_0 = 180°$, corresponding to backscatter (θ_sca =180°). In the RPV model, the hotspot corresponds to a peak in the backscatter direction due to $G(\theta_\mathrm{v}, \theta_0, \Delta\phi)$ =0. By requiring $R_\mathrm{Ross-Li}(\theta_\mathrm{v}, \theta_0, 180º) = R_\mathrm{RPV}(\theta_\mathrm{v}, \theta_0, 180º)$, an updated value of $r_{0,\lambda}$ can be derived to enforce this condition by solving the following equation that is second order in $r_{0,\lambda}$:

$$R_\mathrm{Ross-Li}(\theta_\mathrm{v}, \theta_0, 180º) = r_{0,\lambda} \cdot M(k_\lambda; \theta_\mathrm{v}, \theta_0) \cdot f_\mathrm{HG}(g_\lambda; 180º) \cdot (2 - r_{0,\lambda}) \tag{57}$$

The solution is:

$$r_{0,\lambda} = 1 + [1 - R_\mathrm{Ross-Li}(\theta_\mathrm{v}, \theta_0, 180º)/B]^{1/2} \tag{58}$$

where $B = M(k_\lambda; \theta_\mathrm{v}, \theta_0) \cdot f_\mathrm{HG}(g_\lambda; 180º)$.

Equation 57 ensures that the value of the BRF hotspot is same as that calculated using the Ross-Li model. However, this does not guarantee that the surface albedo calculated from Ross-Li model is conserved. Here the surface albedo, DHR, equals to the ratio of reflected flux to the direct solar irradiance, ignoring the diffuse skylight component [Wanner et al., 1997; Lucht et al., 2000]. To conserve DHR, the value of $r_{0,\lambda}$ needs to be adjusted in a different way. The Ross-Li surface DHR ($w_{0,\mathrm{Ross-Li}}$) can be calculated from the known f_iso, f_geo, and f_vol, namely:

$$w_{0,\text{Ross-Li}} = f_{\text{iso}} + \frac{f_{\text{vol}}}{\pi}\int_0^{2\pi}\int_0^{\pi/2}\rho_{\text{vol}}(\theta_0,\theta_{\text{v}},\Delta\phi)\cos\theta_{\text{v}}\sin\theta_{\text{v}}\,d\theta_{\text{v}}\,d(\Delta\phi) \tag{59}$$

$$+\frac{f_{\text{geo}}}{\pi}\int_0^{2\pi}\int_0^{\pi/2}\rho_{\text{geo}}(\theta_0,\theta_{\text{v}},\Delta\phi)\cos\theta_{\text{v}}\sin\theta_{\text{v}}\,d\theta_{\text{v}}\,d(\Delta\phi)\,.$$

781 For the RPV model, surface DHR ($w_{0,\text{RPV}}$) contributed by the fitted parameters $a_{0,\lambda}, k_\lambda, b_\lambda$ can be

782 calculated by

$$w_{0,\text{RPV}} = \frac{1}{\pi}\int_0^{2\pi}\int_0^{\pi/2} R_{\text{RPV}}\left(a_{0,\lambda},k_\lambda,b_\lambda;\theta_0,\theta_{\text{v}},\Delta\phi\right)\cos\theta_{\text{v}}\ \sin\theta_{\text{v}}\,d\theta_{\text{v}}\,d(\Delta\phi). \tag{60}$$

783 Expanding the hot spot term of RPV based equations, $w_{0,\text{RPV}}$ can be split into two orders:

$$w_{0,\text{RPV}} = w_{0,\text{RPV}-1} + w_{0,\text{RPV}-2} \tag{61}$$

784 where

$$w_{0,\text{RPV}(1)} = r_{0,\lambda}\cdot E_1 \tag{62}$$

$$= r_{0,\lambda}\cdot\frac{1}{\pi}\int_0^{2\pi}\int_0^{\frac{\pi}{2}} M\cdot f_{\text{HG}}\cdot\left(1+\frac{1}{1+G}\right)\cos\theta_{\text{v}}\sin\theta_{\text{v}}\,d\theta_{\text{v}}d(\Delta\phi),$$

$$w_{0,\text{RPV}(2)} = r_{0,\lambda}^2\cdot E_2 \tag{63}$$

$$= r_{0,\lambda}^2\cdot\frac{1}{\pi}\int_0^{2\pi}\int_0^{\frac{\pi}{2}} M\cdot f_{\text{HG}}\cdot\left(-\frac{1}{1+G}\right)\cos\theta_{\text{v}}\sin\theta_{\text{v}}\,d\theta_{\text{v}}d(\Delta\phi).$$

785 Forcing $w_{0,\text{Ross-Li}} = w_{0,\text{RPV}}$, $r_{0,\lambda}$ can be derived from the following quadratic equation

$$w_{0,\text{RPV}(1)} + w_{0,\text{RPV}(2)} = r_{0,\lambda}^2\cdot E_2 + r_{0,\lambda}\cdot E_1 = w_{0,\text{Ross-Li}}. \tag{64}$$

786 Two candidate solutions of $r_{0,\lambda}$ can be derived from solving the above equation. We select 787 the one that is closer to $a_{0,\lambda}$, which represents the adjusted spectral weight after accounting for the 788 hotspot term. This way, we obtain the RPV model parameters as solutions. Applying the same 789 process to mRPV model, we can adjust the spectral weight of mRPV model parameters and obtain

the final solution. Implementing the same procedure to the wavelength independent mRPV and RPV model, we will obtain the solutions for Option 2.

3.3 Numerical study of conversion quality

3.3.1 Conversion Quality of Surface Reflection Properties

To validate the quality of the conversion, we calculate the surface BRF and DHR with the RPV-based model and compare them with those calculated from sRTLS model. We utilize the MAIAC MCD19A3D dataset generated from MODIS observations which provides spectral sRTLS BRF parameters f_{iso}, f_{geo}, and f_{vol} at 1 km spatial resolution. They are used to derive the mRPV and original RPV model parameters using the approach introduced in the previous section. Additional comparison is made against MISR's land surface product MIL2ASLS at 446, 558, 672, and 867 nm [Diner et al., 1999]. We demonstrate the conversion for three surface scenarios, including an urban surface area (lat = 34.13°, lon = −118.13°), a desert surface area (lat = 38.85°, lon = −118.69°), and a vegetated surface area (lat = 35.50°, lon = −119.98°).

3.3.1.1 Evaluation of Converted DHR

Figure 9-11 show a comparison of the surface DHR calculated from the obtained mRPV, RPV, d-mRPV, and d-RPV model parameters for the violet to shortwave infrared spectral range covered by MODIS. The three panels show the results for SZA of 10°, 45°, and 65°. The surface DHR curves for RPV and mRPV are not separated in these figures due to the enforcement of DHR conservation so that their values are same as that modeled by sRTLS. Without the enforcement of DHR conservation, however, we can find better agreement of d-mRPV with the sRTLS model than does d-RPV. When solar zenith angle increases to be off-nadir, however, the accuracy of d-RPV's DHR increases. This is possibly due to that the nadir hotspot is stronger in magnitude than an off-nadir hotspot. Since it contributes more to DHR when the Sun is close to zenith, larger bias is observed when hotspot is excluded in the d-RPV model. It can also be observed that MAIAC's sRTLS model-based surface DHRs at 466, 554, 645, and 856 nm are generally consistent with MISR's surface DHRs for all three surface types. The best DHR agreement between MODIS and MISR is found for the desert surface. Larger differences are observed for the urban and vegetated surfaces, possibly as a result of the differences in the MODIS and MISR angular sampling strategies (primarily cross-track for MODIS and along-track for MISR).

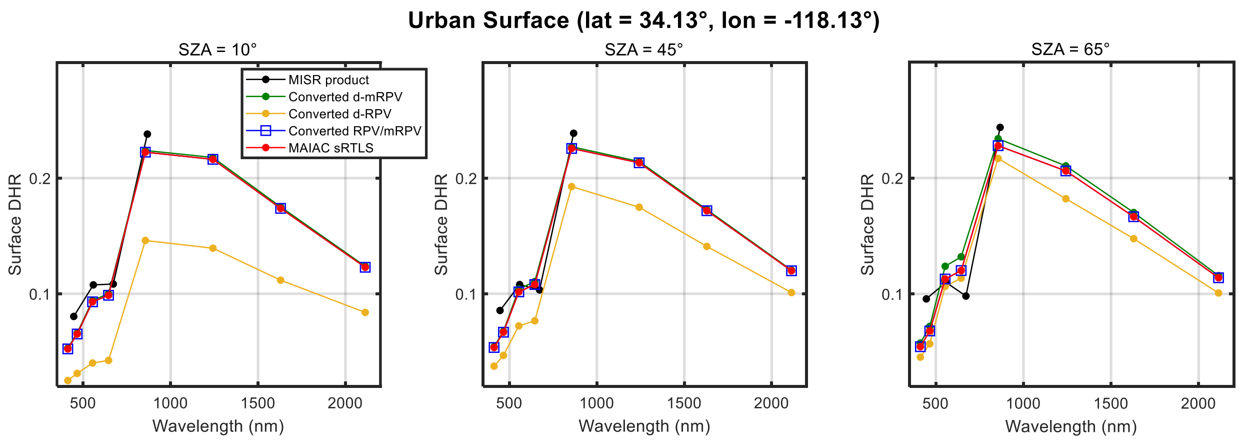


Figure 9. Comparisons of surface DHR of an urban surface area calculated from the d-mRPV, d-RPV, mRPV/RPV, MISR mRPV and MAIAC sRTLS model. The parameters of these models are converted from the sRTLS model parameters saved in the MODIS surface product MCD19A3D dataset, generated using MAIAC algorithm. The MISR surface products are indicated in black dotted curve. The three panels correspond to the solar zenith angles 10°, 45°, and 65°.

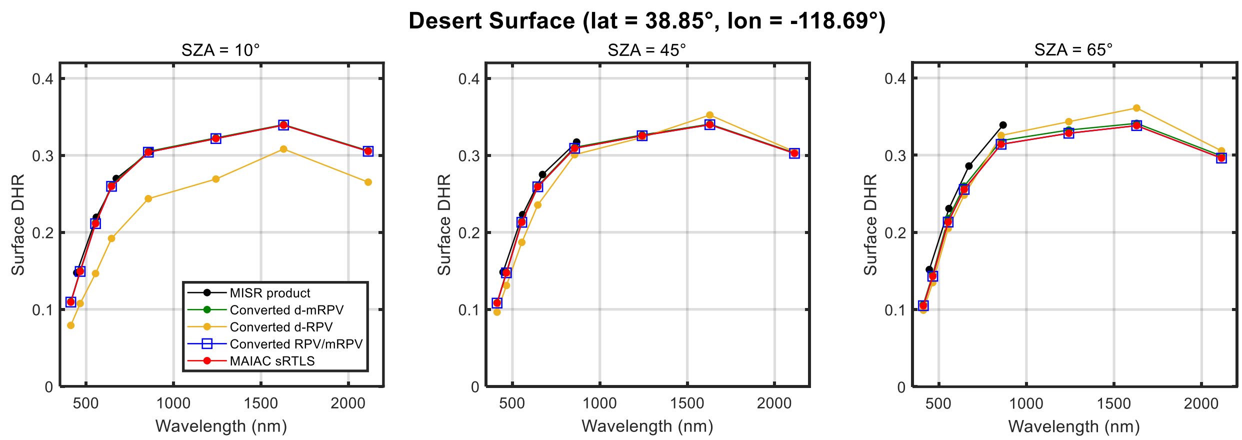


Figure 10. Similar plot as Fig. 9, but for the desert surface area.

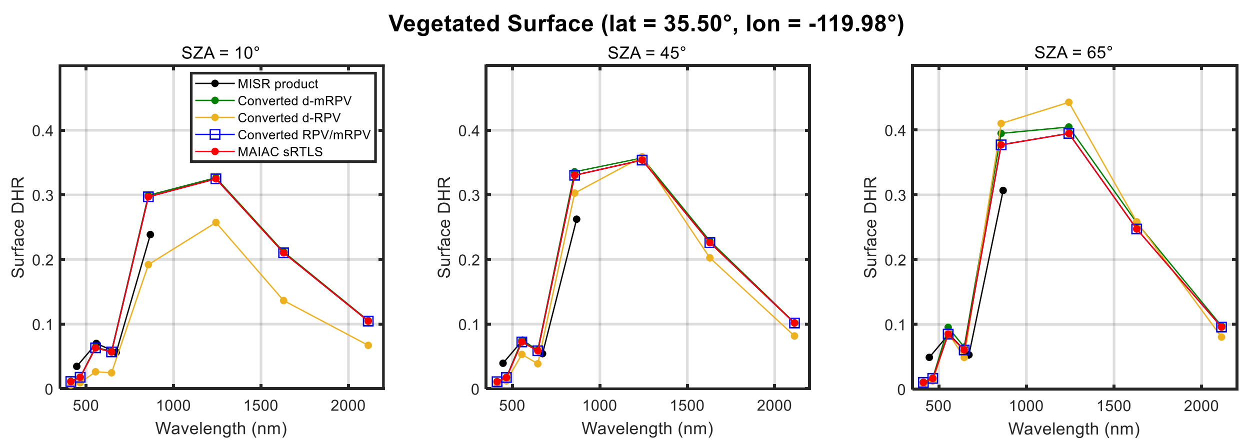


Figure 11. Similar plot as Fig. 9, but for the vegetated surface area.

3.3.1.2 Evaluation of converted BRF

We then compare the angular distribution of surface BRF from two converted models (mRPV and RPV) with the original sRTLS model as well as MISR surface product. The three rows of panels in Fig. 12 show the surface BRFs (solid lines with dots) in the principal plane for the urban area for solar zenith angles 10°, 45°, and 65°, respectively. The first two columns show the BRFs calculated by use of mRPV and RPV models which are converted from the Ross-Li model. The last two columns show the MAIAC and operational MISR products. In each plot, we also added the differences in BRF for the converted RPV, converted mRPV, and MISR datasets relative to the MAIAC's sRTLS reference. The differences are indicated using dashed lines. Figure 13 and 14 show similar BRF comparisons but for the azimuthal planes 45° and 90° from the principal plane, respectively. The grey vertical dash lines indicate the thresholds of $\theta_{\mathrm{v}} = +/- 65°$, representing the recommended limits for the conversion. For extreme case testing, the BRF simulations have been extended to $\theta_{\mathrm{v}} = 70°$.

While the magnitudes and general angular shapes of the converted BRFs in the principal plane are similar to the original MAIAC's sRTLS model values, some differences exist, reflecting the impact of intrinsic modeling errors: even with the enforcement of DHR conservations, the differences still exist, primarily in the angular region ($|\theta_0 - \theta_{\mathrm{v}}| < 15°|$), indicating the impact of the hotspot effect. For the BRF at extreme viewing or incidence angles ($\theta_{\mathrm{v}} > 80°$ or $\theta_0 > 80°$, not shown here), larger deviations are observed. Indeed, due to intrinsic differences in the functional forms of the Ross-Li and RPV-based BRF models, a trade-off exists between accurately reproducing the hotspot peak and minimizing angular deviations across the full BRF domain. In the proposed conversion, hotspot matching is enforced together with DHR conservation to maintain radiometric consistency, which may lead to increased deviations in some angular regions near the hotspot. Simplified variants that omit the hotspot term reduce angular deviations in non-hotspot regions but cannot reproduce the hotspot peak, illustrating this inherent compromise. The purpose of the conversion, however, is not to achieve a perfect BRF representation, but to provide a physically consistent initialization for subsequent optimization-based retrievals, in which surface parameters are further refined.

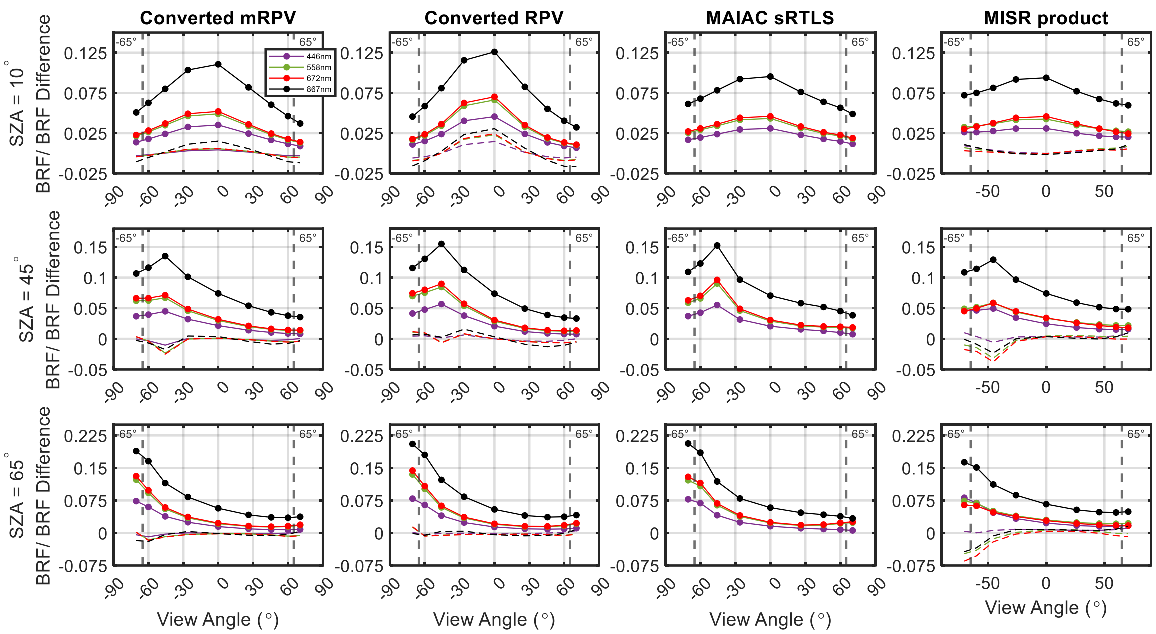


Figure 12. Surface BRFs (solid lines with dots) of an urban surface area in the principal plane at 446, 558, 645, and 867 nm, along with the corresponding differences (dashed lines) for the converted RPV, converted mRPV, and MISR datasets relative to the MAIAC reference. The first column is the surfacer BRF calculated from the converted mRPV model; the second column is calculated from the converted RPV model; the third column shows the MAIAC BRFs; and the fourth column shows products of MISR BRFs. The three rows are for solar zenith angles of 10 °, 45°, and 65°.

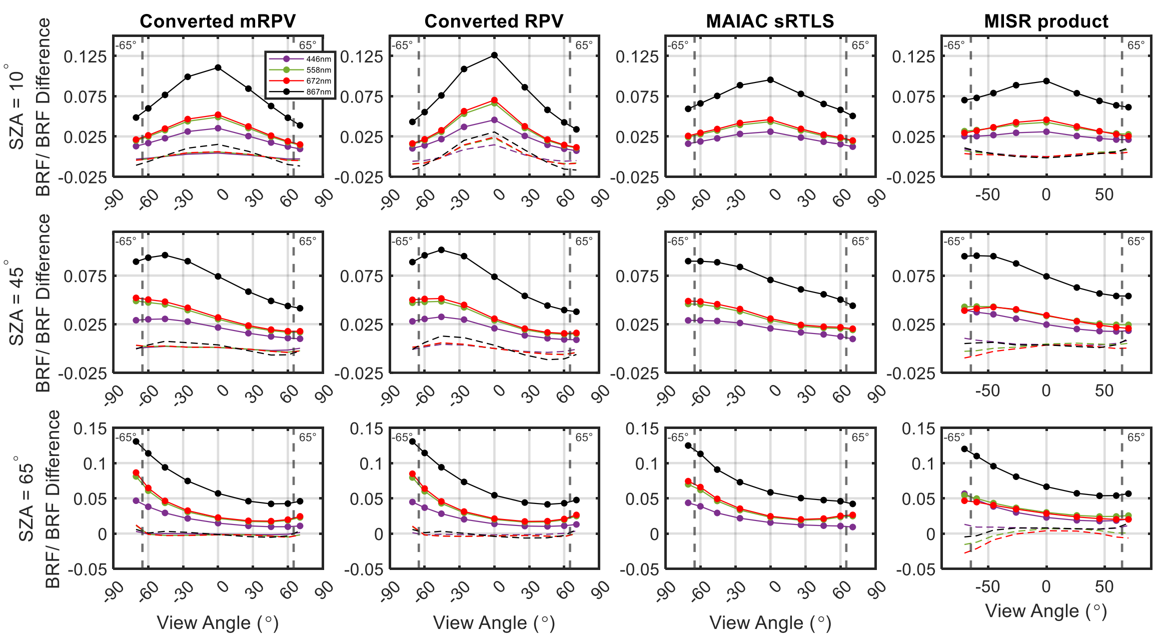


Figure 13. Same as Fig. 12, but for comparisons at 45° off the principal plane.

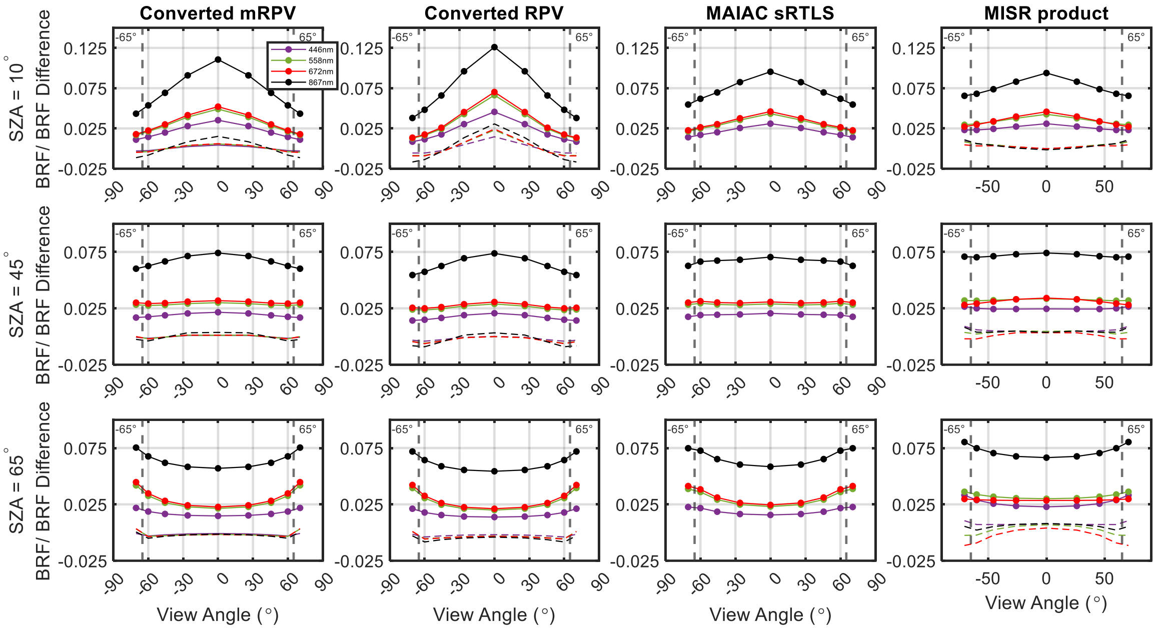
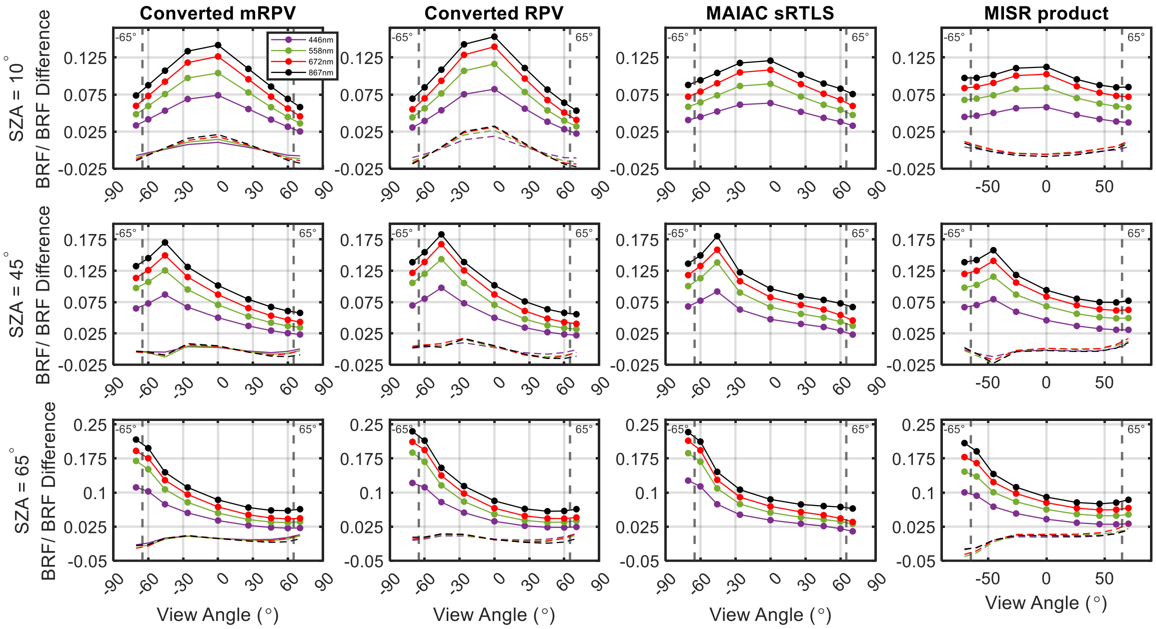


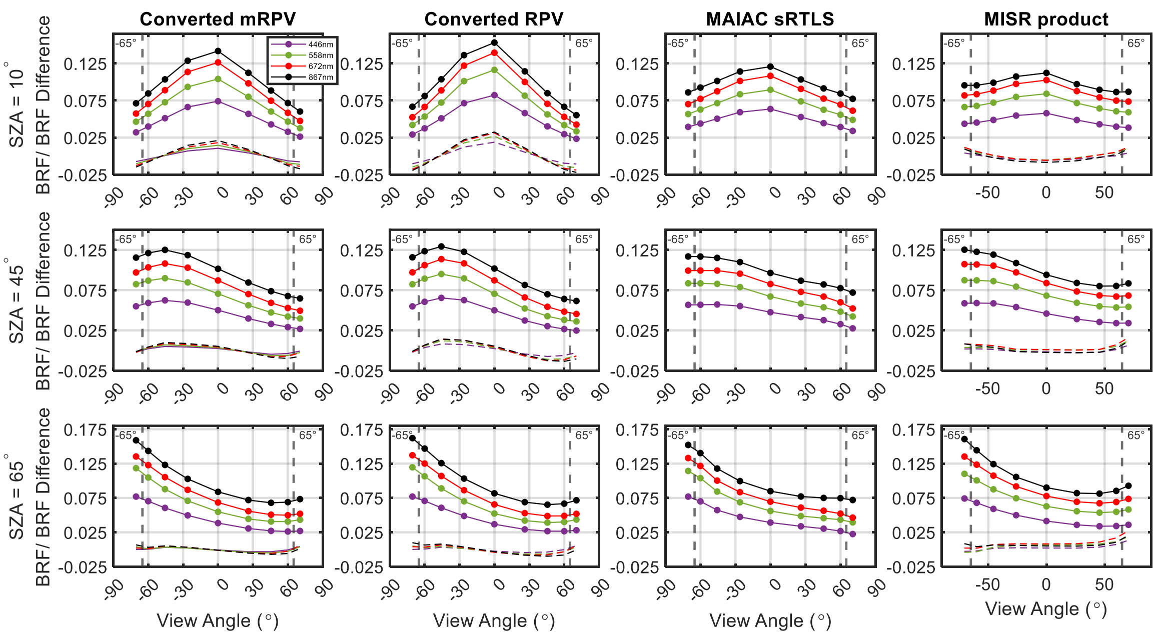
Figure 14. Same as Fig. 12, but for comparisons at 90° off the principal plane.

Figure 15-17 are similar to Figs. 12-14, but for a desert area. Most converted BRFs are consistent with the those from original MAIAC-generated Ross-Li values. The largest differences are found near the hotspot angular region for the RPV model. When SZA = 45°, The mRPV- and RPV-based BRFs in this azimuth plane exhibit a bell-shaped pattern similar to that produced by the sRTLS model, whereas the MISR surface BRF product exhibits a bowl-shaped pattern at 865nm (plotted as a black dotted curve) or less curved at 446, 558, and 645nm (plotted in other colors). This indicates challenges still exist with BRF modeling for a desert surface. Moreover, the MISR and MODIS surface products are derived independently using observations from the two instruments. The limited number and range of viewing geometries may result in insufficient information content to fully constrain the surface BRF at all angles. In this content, a combination of MISR and MODIS observations can potentially provide improved constraints on surface reflectance.



879

880 Figure 15. Same as Fig. 12, but for a desert area.



881

882 Figure 16. Same as Fig. 13, but for a desert area.

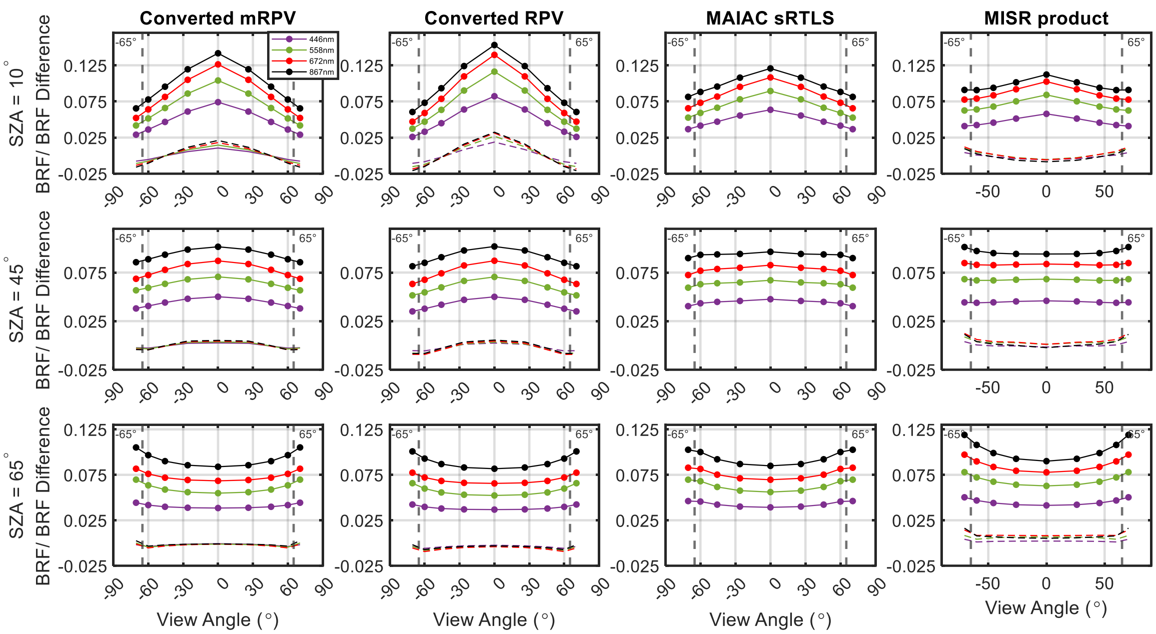


Figure 17. Same as Fig. 14, but for a desert area.

Figure 18 shows the comparison of surface BRF differences relative to sRTLS model derived from the assumptions of wavelength-independent (Option 2) and wavelength-dependent (Option 1) surface BRF angular shape. Same as Fig.Figure 12, calculations are performed for the urban surface area in the principal plane. Associated with Option 2, wavelength dependent set of angular shape parameter (k, b) are not allowed to be used for fitting the sRTLS surface BRF in each band. Therefore, compared with Option 1, fitting error increases at most angles when mRPV model is used to account for the hotspot effect. When using RPV model to account for the hotspot effect, fitting errors are still found increased at the spectral band of 867 nm for the option of wavelength-independent BRF angular shape (Option 2). This is expected in the presence of strong spectral difference of surface BRF angular shape in near-infrared and visible bands in the original RTLS model. However, this option does not cause increment of fitting errors in the principal plane for all bands in RPV model.

Indeed, the conversion from the sRTLS to the RPV model is constrained by simultaneously enforcing agreement in both the DHR and the hotspot region. The DHR represents the integrated contribution of surface BRFs over all viewing geometries, whereas the hotspot is confined to a limited angular range near the principal plane. As a result, these two constraints cannot always be satisfied simultaneously. Across the visible bands, enforcing agreement in the hotspot region tends to improve the fitting across the principal plane. In contrast, when the hotspot is less pronounced (e.g., at 867 nm), the DHR constraint becomes more dominant, which does not necessarily always guarantee the best fit within the principal plane for RPV model, but tends to provide a better representation over the overall angular range.

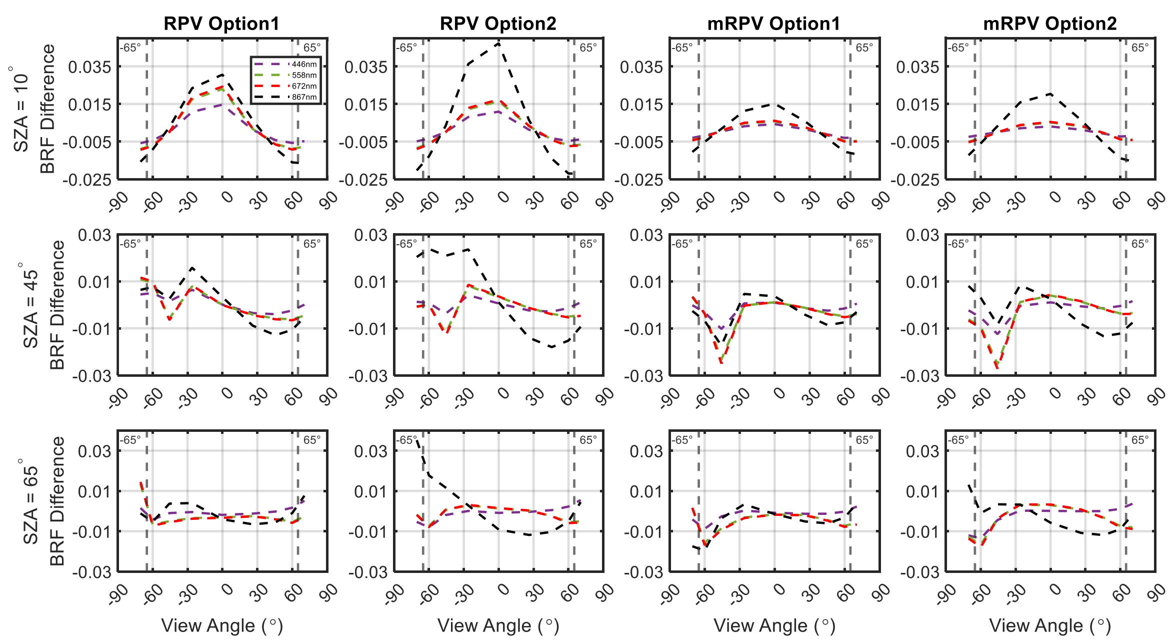


Figure 18. Same as the urban surface area in Fig. 12, but the BRF differences relative to sRTLS model are compared for two conversion options: wavelength dependent (Option 1) and wavelength independent (Option 2) angular shape parameters of surface BRFs. The first and second columns correspond to the Option 1 and Option 2 of converted RPV models, respectively; the third and fourth columns correspond to the Option1 and Option2 of converted mRPV models, respectively. The three rows correspond to solar zenith angles of 10°, 45°, and 65°, respectively.

3.3.2 Influence on Aerosol Retrieval

To perform the conversion impacts on aerosol retrieval, an aerosol retrieval test is conducted to compare the aerosol properties before and after conversions. Using the converted model parameters of RPV model and the original sRTLS model from MAIAC MCD19A3 dataset as initial guesses, we retrieve aerosol properties from a set of combined MISR–MODIS observations while simultaneously refining the surface reflection and compare the results with those obtained using the sRTLS model directly. A total of 967 globally distributed matchups between the combined MISR-MODIS dataset and AERONET are used in the evaluation, covering a broad AOD range up to 4. The combined dataset includes the radiances at the wavelengths of 412, 446, 558, 672, 748, 867, 1240, 1640, and 2130 nm. At the four MISR wavelengths (446, 558, 672, and 867 nm), MISR radiances from up to nine view angles are used. At the remaining wavelengths, MODIS radiances are used, which are available only at the nadir view angle. Among the 967 collocated datasets, 347 cases have AOD values smaller than 0.3, while 620 cases have AOD values larger than 0.3. AERONET also provides SSA and size distribution for 165 cases, which are used for comparison.

The retrieval algorithm is based on an optimal estimate approach developed for JPL's AirMSPI [Xu et al., 2017], with the addition of a priori of aerosol amount informed by a lookup table search [Diner et al., 2008; Zhang et al., submitted]. The a priori constraints of wavelength dependent (Option 1) and spectral invariance of surface BRF shape (Option 2) are imposed to constrain two separate retrieval tests. Retrieval output includes aerosol particle concentration, bi-modal log-normal particle size distribution parameters and refractive index, effective height and width of vertical aerosol profile, and a set of surface model parameters that govern the RPV or sRTLS models. To better retrieve surface reflection using the RPV model, the sRTLS model parameters informed by the MAIAC product are first converted to an equivalent set of RPV parameters based on Option 1 and Option 2, using the approach introduced in Section 3.2. These converted parameters are then used as the initial guess for the RPV surface parameters in the aerosol-surface joint retrieval algorithm [Zhang et al., submitted]. During the retrieval process, the surface parameters are further adjusted through the optimization procedure together with aerosol parameters to fit the combined MISR-MODIS observations. In addition, we calculate the AOD, aerosol SSA, and effective radius of fine and coarse mode particles from the retrieval outputs. It should be noted that this retrieval configuration differs from the general CIMAP framework described earlier. This configuration was adopted to prioritize evaluation results that are intended to inform the development of MAIA's operational aerosol retrieval algorithms.

Figure 19 demonstrates the comparison of retrieved AODs at 550 nm using RPV Option 1, RPV Option 2, and the sRTLS model against AERONET reference data, along with an intercomparison with MISR AOD. Both RPV Option 1 and RPV Option 2 retrievals exhibit strong correlations with AERONET AOD, with a coefficient of determination (R^2) of approximately 0.8 and a slope of approximately 0.9. RPV and sRTLS model-based AOD retrievals indicate a similar quality for the entire range $0 < AOD < 4$: the mean absolute difference (MAD) of AOD is 0.123, 0.125, and 0.125, whereas the mean differences (MD) are -0.017, -0.019 and -0.025 for RPV Option 1, RPV Option 2and sRTLS, respectively. When all AOD cases are divided into low-to-moderate (< 0.3) and moderate-to-high (> 0.3) categories, the RPV model-based retrieval yields results comparable to those obtained using sRTLS for low-to-moderate AODs. However, for moderate-to-high AODs, the RPV-based retrieval exhibits better quality than the sRTLS-based retrieval (MD ≈ −0.03, −0.03 vs. −0.04) when the MD metric is used to quantify differences relative to AERONET data. The higher moderate-to-high AODs retrieved by the RPV model also contribute to improved all-case AOD statistics. Moreover, compared with the optimization-based AOD retrievals, the MISR AOD product exhibits a larger underestimation of AOD (MD = −0.126, as indicated in Table 1) than the optimization-based retrievals, which show much smaller biases (MD = −0.017 for the RPV model Option 1 and MD = −0.025 for the sRTLS model). This is mainly caused by MISR's underestimated AOD by 0.204 for moderate-to-high AODs (AOD > 0.3). For low-to-moderate AODs (AOD < 0.3), however, MISR AODs exhibit substantially improved accuracy, with the MD ≈ 0.013 and MAD ≈ 0.034, comparable to the optimization retrievals (MD ≈ 0.002, 0.007 and MAD ≈ 0.04, 0.04).

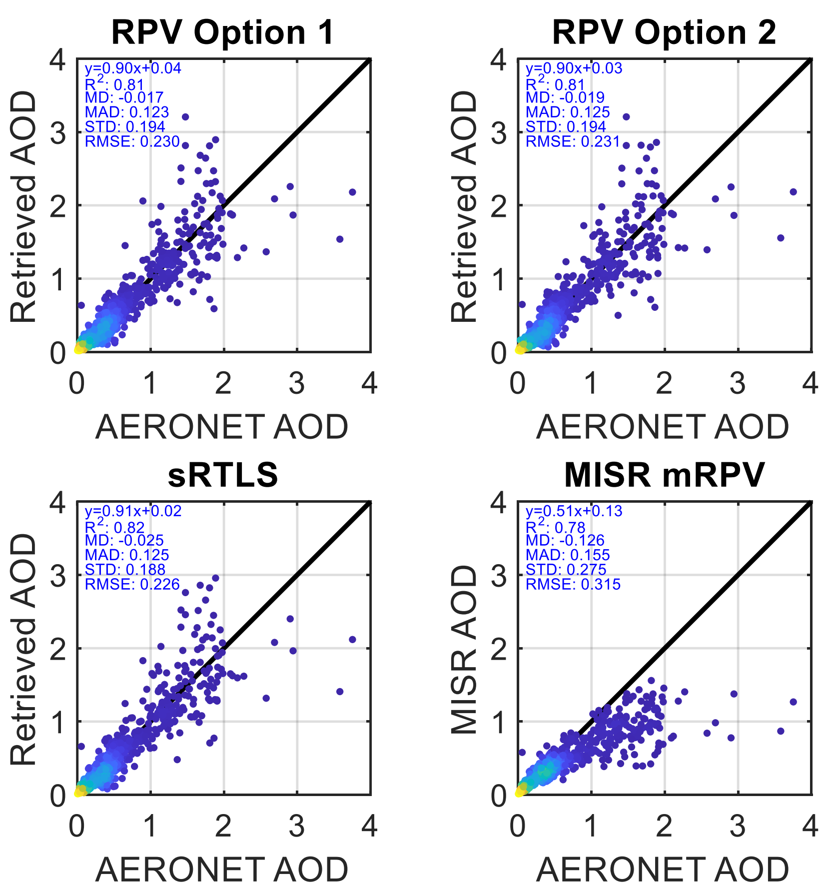


Figure 19. Comparisons of AOD retrievals at 550nm. Upper left panel: Optimization-based retrieval with RPV model with the initial guess of model parameters converted from MAIAC dataset of sRTLS model parameters. Option 1 is selected for this retrieval. Upper right panel: Similar to upper left panel but the retrieval uses Option 2, namely wavelength independent assumption of surface BRF angular shape. Low left panel: Optimization-based retrieval with

sRTLS model with its initial guess directly informed by using MAIAC dataset. Lower right panel: MISR retrieval. Color indicates point density, with warmer colors (yellow) representing higher data density and cooler colors (blue) indicating lower data density. Statistics are labeled in the upper left corners.

The comparison of the retrieved SSA against AERONET products is demonstrated in Fig. 20. The sRTLS constrained aerosol SSA retrievals show slightly better agreement with AERONET data, namely 0.036 (sRTLS) vs. 0.034 (RPV Option1), 0.037 (RPV Option2) for MD and 0.043 (sRTLS) vs. 0.046 (RPV Option 1), 0.051 (RPV Option 2) for MAD. Note that among the 165 cases of SSA comparison, 162 cases are associated with AOD > 0.3, therefore little difference is observed in MAD or MD for all-case SSA and SSA for AOD > 0.3. However, both RPV and sRTLS adopted optimization-based retrievals indicate better SSA accuracy than MISR's SSA values indicated in Table 1 (0.04, 0.05 vs. 0.07 difference than AERONET reference data). This could be due to a) the combined use MISR and MODIS data to constrain aerosol retrieval extend the lower end of spectral coverage to 412 nm, shorter than MISR's 445 nm, which favors absorbing aerosol retrieval; and b) the search of aerosol solution in a continuous aerosol parameter space than in a lookup table consisting of a limited number (74) of mixtures in MISR algorithm which may lack sufficiency to represent the optical and microphysical properties of globally distributed aerosols.

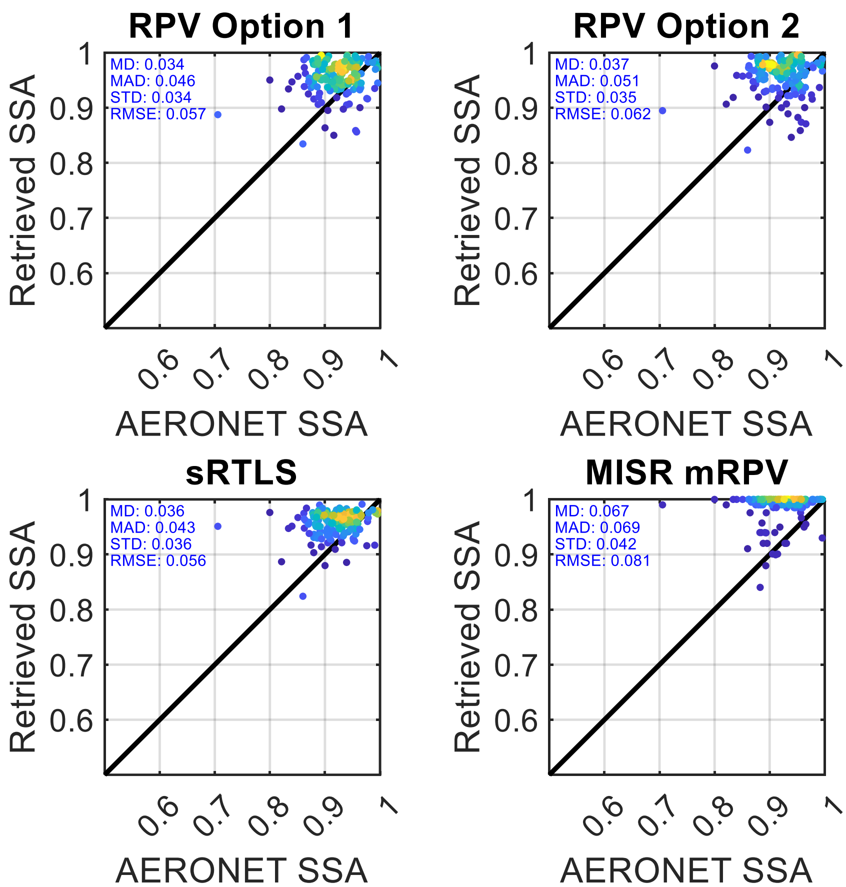


Figure 20. Similar to Fig. 19 but for comparison of SSA retrievals at 550nm.

Comparison of the retrieved effective radii of fine and coarse mode particles against AERONET products is demonstrated in Fig. 21. For the fine mode, the MD indicates better performance for the RPV model-based retrieval (0.024 µm, 0.015 µm difference relative to AERONET) than for the sRTLS model–based retrieval (0.023 µm difference relative to AERONET), whereas the MAD indicates better performance for the sRTLS model (0.046 µm difference relative to AERONET) than for the RPV model (0.06 µm, 0.058 µm difference relative to AERONET). For the coarse mode, however, both the MD and MAD metrics indicate better performance for the sRTLS model: the MD is −0.021µm for sRTLS compared with 0.036 µm and −0.107 µm for RPV, and the MAD is 0.381µm for sRTLS compared with 0.373 µm and 0.479µm for RPV, relative to the AERONET reference. Moreover, the sRTLS approach yields a more tightly clustered distribution around the AERONET reference data, with a standard deviation (STD) of 0.290 μm, smaller than 0.3 μm and 0.358 μm for RPV. However, note should be paid to that, though AEROENT SSA and particle size are used as the reference for comparison, they are also inversion products and therefore subjected to uncertainties of 0.03-0.07 for SSA and 15%-100% for particle size, depending on the aerosol loading and particle type [Dubovik et al., 2000].

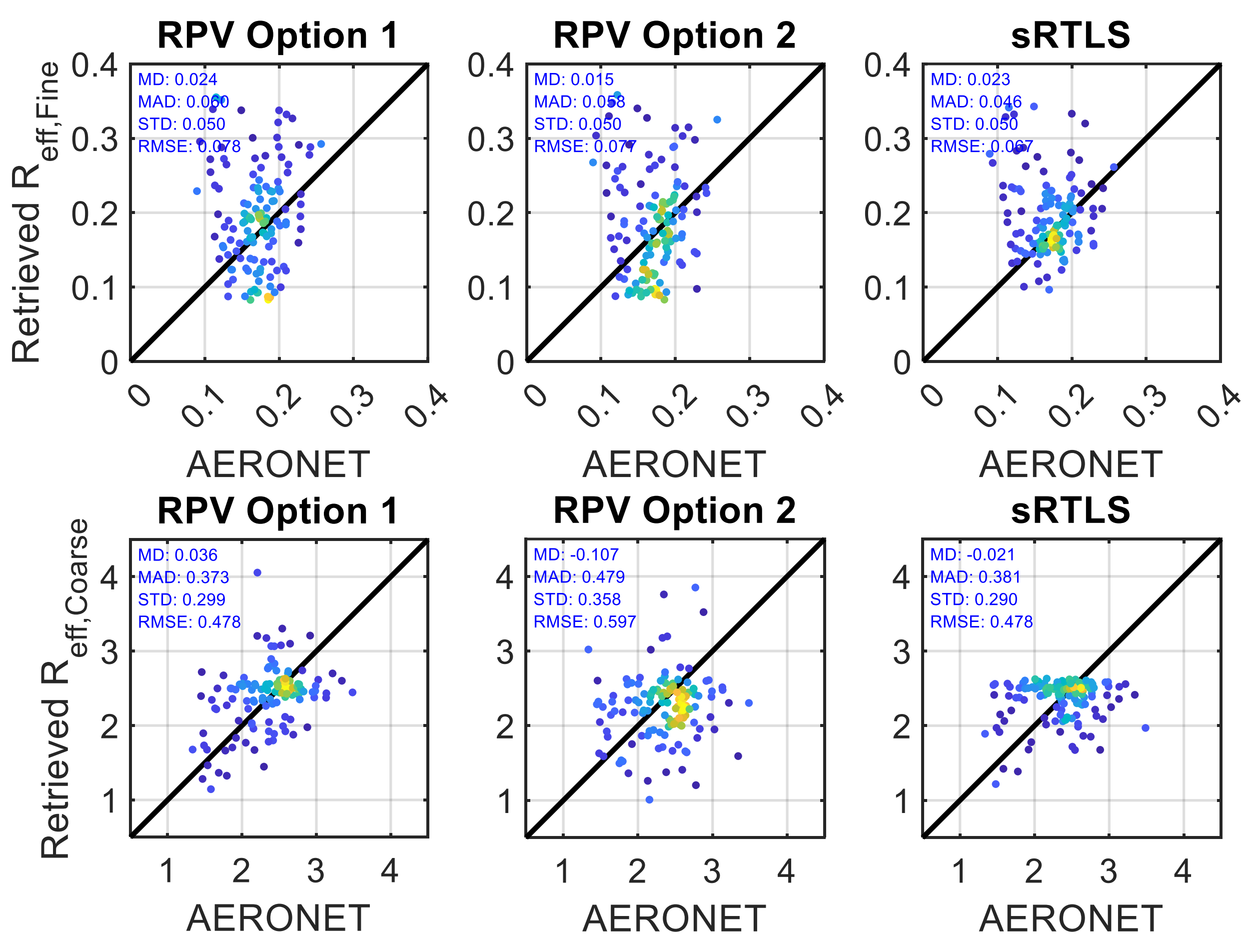


Figure 21. Comparison of retrieved effective radius against AERONET reference data. The retrievals use RPV option 1 (left panels), RPV option 2 (middle panels), and sRTLS (right panels) models to calculate surface BRFs. Upper panels: fine mode effective radius. Lower panels: coarse mode effective radius. Color indicates point density, with warmer colors (yellow) representing higher data density and cooler colors (blue) indicating lower data density. Statistics are labeled in the upper left corners.

The statistics of the optimization-based retrievals using the sRTLS and RPV (both Options) surface models, and the comparison to the MISR v23 aerosol product, are summarized in Table 1. Between the sRTLS and RPV model-based retrieval results, the numbers indicating better agreement with the AERONET reference are highlighted in grey. Overall, the performances of the sRTLS and RPV models are comparable, with their relative differences depending on the specific aerosol quantities and evaluation metrics. The RPV based models exhibit slightly improved performance for AOD retrievals, whereas the sRTLS model demonstrates superior performance for SSA and coarse-mode aerosol size retrievals. Although converting the surface representation from sRTLS to RPV may introduce angle- and surface-type dependent BRF differences, the optimization-based retrieval allows the surface BRF parameters to be adjusted within physically reasonable ranges during the iteration-based process to fit MISR-MODIS data with an optimal solution and minimized residue. In this case, the observed differences between retrievals using the sRTLS and RPV models are more likely attributable to a) their underlying physical formulations in representing the angular anisotropy of BRF of different surface types; and b) the sufficiency of information content to disentangle SSA, particle size, and AOD.

We have also tested the performance of RPV and sRTLS models for two generally classified surface using the metric Normalized Difference Vegetation Index (NDVI). NDVI is calculated using nadir-view radiances measured in the red and near-infrared bands, namely NDVI= $(L_{NIR} - L_{Red})/(L_{NIR} + L_{Red})$. The higher NDVI values (> 0.6) are typically associated with dense vegetation canopies, for which multiple scattering dominates surface reflectance. For lower NDVI values (NDVI < 0.6), surface reflectance typically reflects mixed or non-canopy scattering regimes, with more contributions from soil background, sparse vegetation, or heterogeneous surface components.

Within the combined MISR and MODIS dataset, we have identified 872 cases with NDVI > 0.6 and 95 cases with NDVI < 0.6. Table 2 and Table 3 summarize the retrieval comparison of AOD for NDVI < 0.6 and NDVI > 0.6 respectively. Due to the missing AERONET reference data of SSA and effective particle radius, no statistics is made for these quantities for NDVI > 0.6. For NDVI < 0.6, the numbers for SSA and effective particle radius are same as those in Table 1. The sRTLS- and RPV-based AOD retrievals exhibit generally better performance over vegetation-dominated surfaces (NDVI > 0.6) than over surfaces characterized by NDVI < 0.6. This is mainly due to that for both RTLS and RPV, higher surface BRF modeling accuracy is generally achieved over darker surfaces (NDVI > 0.6). With the increased surface brightness (decreased NDVI) and heterogeneity, modeling accuracy degrades. Between the performance of RPV and sRTLS model-based retrievals, however, the retrieval accuracy is consistent with our observation from all-case analysis (summarized in Table 1): namely RPV yields slightly better AOD retrievals whereas sRTLS yields superior performance for SSA and coarse-mode aerosol size retrievals. Moreover, for the majority of AOD retrievals associated with NVDI < 0.6, the optimization-based retrievals that deploy either RPV or sRTLS have better performance than MISR retrievals, as characterized by smaller MD, MAD and STD. It should be noted that higher accuracy for NDVI > 0.6 than for NDVI < 0.6 is also observed from comparing MISR retrieval statistics in two NDVI regimes in Tables 2 and 3. For dense vegetated surfaces (NDVI > 0.6), however, MISR AODs have slightly smaller MD, MAD and STD than those of optimization-based retrievals that deploy RPV or sRTLS model. Indeed, MISR aerosol retrieval capitalizes on the surface contrast to get the path radiance through empirical orthogonal function analysis of multi-angle observations, which avoids the deployment of any specific surface model [Diner et al., 1999]. Consequently, it has less subjection to the surface modeling errors. Regardless the MISR or optimization-based retrieval, the retrieval errors increase when surfaces transition from vegetation dominated (NDVI > 0.6) to less or non-vegetated surfaces (NDVI < 0.6). For NDVI < 0.6 (brighter surface), however, deployment of a RPV or sRTLS model helps more to constrain AOD retrieval than does the MISR approach.

Generally, the statistical differences between optimizations using the sRTLS model and both RPV Options are minimal, indicating limited impact from the surface model conversion on the aerosol retrieval algorithm. It is worth noting that RPV Option 1 shows slightly better performance than RPV Option 2, particularly for SSA and effective radius. However, given the limited number of cases available for SSA and effective radius comparison, further investigation is needed before a definitive conclusion can be drawn regarding the optimal Option selection.

Table 1. Statistics of optimization-based retrievals with RPV and sRTLS surface model and comparison to MISR products (Numbers indicating better agreement with the AERONET reference are highlighted in grey.)

Aerosol Quantity	**Optimization with sRTLS model**	**Optimization with RPV model Option1**	**Optimization with RPV model Option2**	**MISR product (v23)**
$MD_{AOD\_all\_cases}$*	-0.025	-0.017	-0.019	-0.126
$MD_{AOD\leq0.3}$	0.002	0.007	0.002	0.013
$MD_{AOD>0.3}$	-0.040	-0.031	-0.031	-0.204
$MAD_{AOD\_all\_cases}$**	0.125	0.123	0.125	0.150
$MAD_{AOD\leq0.3}$	0.040	0.041	0.039	0.034
$MAD_{AOD>0.3}$	0.173	0.169	0.173	0.223
$STD_{AOD\_all\_cases}$***	0.188	0.194	0.194	0.275
$MD_{SSA\_all\_cases}$	0.036	0.034	0.037	0.067
$MD_{SSA\_AOD\leq0.3}$	0.027	0.024	0.040	0.056
$MD_{SSA\_AOD>0.3}$	0.036	0.034	0.037	0.067
$MAD_{SSA\_all\_cases}$	0.043	0.046	0.051	0.069
$MAD_{SSA\_AOD\leq0.3}$	0.038	0.034	0.047	0.057
$MAD_{SSA\_AOD>0.3}$	0.043	0.047	0.052	0.069
$STD_{SSA\_all\_cases}$	0.036	0.034	0.035	0.042
MD_{Reff_fine}	0.023	0.024	0.015	N/A
MAD_{Reff_fine}	0.046	0.060	0.058	N/A
STD_{Reff_fine}	0.050	0.050	0.050	N/A
MD_{Reff_coarse}	-0.021	0.036	-0.107	N/A
MAD_{Reff_coarse}	0.381	0.373	0.479	N/A
STD_{Reff_coarse}	0.290	0.299	0.358	N/A

*: The mean difference (MD) is calculated as: $MD = \frac{1}{N_{cases}} \sum_{n=1}^{N_{cases}} (x_{Retrieved,n} - x_{AERONET,n})$, where x represents

the retrieval quantity under analysis and N_{cases} is number of retrieval cases under analysis;

**: The mean absolute difference (MAD) is calculated as: $MAD = \frac{1}{N_{cases}} \sum_{n=1}^{N_{cases}} |x_{Retrieved,n} - x_{AERONET,n}|$;

***: The standard deviation (STD) is calculated as: $STD = \left[\frac{1}{N_{cases}-1} \sum_{n=1}^{N_{cases}} (x_{Retrieved,n} - x_{AERONET,n})^2\right]^{1/2}$.

Table 2. Same as Table 1, but for the 872 retrieval cases over the surfaces of NDVI<0.6

Aerosol Quantity	**Optimization with sRTLS model**	**Optimization with RPV model Option1**	**Optimization with RPV model Option2**	**MISR product (v23)**
$MD_{AOD\_all\_cases}$	-0.028	-0.021	-0.021	-0.140
$MD_{AOD\leq 0.3}$	0.001	0.005	0.003	0.012
$MD_{AOD>0.3}$	-0.042	-0.034	-0.033	-0.215
$MAD_{AOD\_all\_cases}$	0.132	0.129	0.131	0.168
$MAD_{AOD\leq 0.3}$	0.041	0.040	0.039	0.034
$MAD_{AOD>0.3}$	0.177	0.172	0.176	0.233
$STD_{AOD\_all\_cases}$	0.195	0.201	0.202	0.286

Table 3. Same as Table 1, but for 95 retrieval cases over the surfaces of NDVI>0.6

Aerosol Quantity	**Optimization with sRTLS model**	**Optimization with RPV model Option1**	**Optimization with RPV model Option2**	**MISR product (v23)**
$MD_{AOD\_all\_cases}$	0.002	0.002	-0.003	0.002
$MD_{AOD\leq 0.3}$	0.005	0.016	-0.003	0.018
$MD_{AOD>0.3}$	-0.005	0.020	-0.002	-0.027
$MAD_{AOD\_all\_cases}$	0.058	0.040	0.068	0.040
$MAD_{AOD\leq 0.3}$	0.034	0.046	0.037	0.032
$MAD_{AOD>0.3}$	0.101	0.117	0.125	0.055
$STD_{AOD\_all\_cases}$	0.072	0.094	0.081	0.032

3.4 Summary

An algorithm has been developed to convert from Ross-Li to several RPV model parameterizations of surface bidirectional reflectance. Optional constraints, including DHR conservation, hotspot conservation and spectral invariance on the BRF angular shape can be imposed. Numerical studies of the conversion were performed for urban, desert, and vegetated surfaces. The results show good agreement in both the spectral shape and magnitude of the BRF at most viewing angles. Relatively larger deviation shows up in near hotspot angular region and at extreme view angles. As one application, the Ross-Li model-based surface BRF dataset is used to inform aerosol inversions employing the RPV model. Aerosol retrieval practice using the combined MISR and MODIS observations shows that RPV model-based retrieval yield comparable aerosol optical depth statistics to those obtained using an sRTLS surface model directly. The RPV model-based retrieval exhibits slightly enhanced AOD retrievals, whereas the sRTLS model-based retrieval shows improved agreement with AERONET for SSA and coarse-mode particle size. These differences reflect the distinct physical representations of surface angular anisotropy in the two models, together with the sufficiency of information content in the remote sensing observations for accurately disentangling SSA, particle size, and AOD. To refine our analysis, more comprehensive surface-type-stratified assessment of aerosol retrieval performance are required, which necessitates a robust land-cover characterization, sufficient sampling, control of confounding factors such as aerosol loading and observation geometry, and reliable ground truth (e.g. for SSA and particle size) for comparison.

4. Correlation-based Inversion Algorithm for Aerosol Microphysical Properties

While the previously developed inversion strategies and constraints have demonstrated considerable success in stabilizing the retrieval, further development of the aerosol inversion algorithm is needed to explore opportunities for improvement. Physically, the spatial, temporal, and spectral variations of aerosol properties are governed by a limited set of surface sources, transportation pathways, and atmospheric environments characteristic of each region, which are mathematically correlated with one another. These region-specific correlations encode the dominant aerosol types and their characteristic property co-variations, effectively serving as a regional fingerprint of the local aerosol climatology. Investigation of aerosol climatology from AERONET products has revealed significant correlations between aerosol properties, including refractive index and size distribution, that are both robust and regionally coherent [Xu et al., 2019; Huang et al., 2021, 2026b]. By incorporating these correlations as an additional constraint, the retrieval gains access to region-specific aerosol information that complements and reinforces the constraints already in use.

Building on this conception, the correlation constraint was proposed in Xu et al. (2019), which further constrains aerosol retrieval by exploiting the correlations among intrinsic aerosol properties. With the implementation of this constraint, the resulting algorithm framework is referred to as the CIMAP [Xu et al., 2019; Huang et al., 2021, 2026b]. Within this framework, principal component analysis (PCA) is employed to capture the correlations among aerosol fields, where the first few principal components (PCs) represent the dominant modes of co-variation. By adopting a selected number of PCs as retrieved parameters, the retrieval parameter space is significantly reduced and the associated computational burden is relieved. Crucially, the correlation constraints operate synergistically alongside existing constraints, such as smoothness constraints, within the classic retrieval framework of Dubovik et al. (2011). This combined use of constraints ensures that the retrieved solution remains physically plausible in a general sense while remaining specifically consistent with the aerosol climatology of the region under study.

CIMAP was initially tested using airborne measurements from the AirMSPI [Xu et al., 2019]. Comparison of AOD and SSA derived from 27 AirMSPI datasets against reference data from AERONET yielded MADs of approximately 0.029 and 0.038, respectively. This level of accuracy is comparable to that achieved by non-correlation-based multi-pixel inversion approaches that solve for aerosol properties in the regular parameter space, but with significant acceleration of computational speed, demonstrating the effectiveness of the correlation constraint without sacrificing retrieval quality.

Building on the success of Xu et al. (2019), the CIMAP algorithm has since been extended across different configurations, constraints, and observational inputs. In this dissertation, two CIMAP applications are introduced: CIMAP-A, which retrieves aerosol properties from ground-based AERONET observations, and CIMAP-P, which retrieves aerosol properties from satellite-borne POLDER observations. The purpose of these two CIMAP extensions is to: 1) complete the deployment of the CIMAP algorithm to common instruments with different measurement principles; 2) retrieve the target aerosol properties with low uncertainties for aerosol studies; 3) accelerate retrieval speed; and 4) provide additional experience and information for implementation in further missions or field campaigns, and for further algorithm improvement and constraint development. While the two applications share a similar algorithmic structure, they differ in their specific configurations. Notably, the MAIAC a priori is adopted only in CIMAP-P, as CIMAP-A was developed prior to the implementation of the MAIAC conversion algorithm.

4.1 Correlations among Aerosol Properties

Aerosols often exhibit strong seasonal and spatial dependence, with their distribution and variability governed by a few prevailing aerosol types, as widely reported in observational and modeling studies [Malm et al., 1994, 2004; Hand et al., 2024; Blanchard et al., 2013]. For example, in urban regions aerosol scattering and absorption coefficients typically show enhanced variability during winter and summer [Laj et al., 2020; Chan et al., 2018]. In mountainous regions, seasonal variability in some aerosol properties can be even more pronounced due to fluctuations in boundary layer height, complex terrain-induced circulation, and vertical transport processes [Laj et al., 2020]. At the same time, the Earth's surface serves as the primary source of aerosol particles, resulting in strong links between surface characteristics and aerosol composition and optical properties. Marine environments are dominated by sea salt aerosols generated through wind and wave activity [Monahan et al., 1986], while arid and semi-arid regions contribute mineral dust through ideal soil conditions and wind velocity [Ginoux et al., 2001]. Anthropogenic emissions from urban and industrial areas release black carbon, sulfates, and nitrates, whereas biomass burning in rural regions contributes significant amounts of carbonaceous aerosols [Akagi et al., 2011; Andreae et al., 2019]. In addition, forested regions emit biogenic volatile organic compounds (BVOCs), which undergo oxidation to form secondary organic aerosols [Kanakidou et al., 2005; Hallquist et al., 2009]. Correlation in aerosol properties arises because, in a given region or season, aerosol types and mixtures are not arbitrary but prevailed by a limited number of sources and atmospheric processes. As a result, aerosol properties tend to vary in a coordinated manner, which can be efficiently captured by a small number of PCs [Xu et al., 2019; Huang et al., 2021, 2026b]. Note that the PCs do not directly correspond to specific aerosol sources. Instead, they represent the dominant modes of correlated variability in aerosol parameters.

In the CIMAP algorithm, the inherent correlations in the optical and microphysical properties of aerosols, informed by aerosol climatology, can be captured using PCA and constructed into an a priori constraint. PCA identifies the few dominating PCs that contribute to the variations in aerosol properties. Retrieval then proceeds in a significantly reduced parameter space by determining the few dominating PC-related quantities such as PC weights. Consequently, the algorithm achieves higher computational efficiency and improved numerical stability [Xu et al., 2019].

Mathematically, a correlated aerosol property (e.g. refractive index across wavelengths, particle concentration across size bins) are expressed as a function of PC weight **w**, PC matrix **v** and spatiotemporal mean ($\bar{\mathbf{x}}$):

$$\log(\mathbf{a}_{\mathrm{aerosol}}) = \bar{\mathbf{x}} + \mathbf{v} \times \mathbf{w} \tag{65}$$

where the PC matrix **v** constitutes of the dominating PCs. Assume the first few PCs dominates the spatiotemporal variations of a correlated aerosol property over a target area, they are used to inform aerosol inversion. Under such a context, the retrieval becomes to determine an optimal set of weights (**w**) of the dominating PCs. During an optimization process, the updated aerosol properties in regular parameter space can be calculated from the updated PC weights using Eq. (65). Combined with the input of uncorrelated aerosol properties and surface reflectance properties, the total signals are simulated using the MarCh model [Xu et al., 2011, 2012] to fit satellite observations.

4.2 CIMAP for Ground-Based AERONET Observations

Based on the application of CIMAP to AirMISPI, a continued configuration of CIMAP-A is adopted for AERONET sky almucantar observations. AERONET sites are globally distributed, providing extensive records for regional aerosol studies as well as datasets for PCA and retrieval intercomparison. By virtue of its ground-based measurement geometry, AERONET observations offer several advantages: the relative uncertainty of radiance measurement is relatively low, assumed to be 5%, and the absolute uncertainty of AOD measurement is estimated at 0.01–0.02

(conservative estimates from Holben et al., 1998, 2006). Furthermore, the AERONET sensor captures a downward-measured signal, which requires simpler RT modeling compared to sensors that capture two-path TOA signals. By incorporating these AERONET characteristics, CIMAP-A aims to: 1) leverage AERONET measurements to demonstrate its value for accurately retrieving aerosol properties; 2) accelerate retrieval speed; and 3) provide insights for future applications as an algorithm test prior to deployment on satellite-borne measurements.

CIMAP-A applies PCA to capture correlations among aerosol properties and construct aerosol a priori constraints. By retrieving the PC weights relevant to aerosols rather than the full set of aerosol parameters, CIMAP-A maintains retrieval accuracy comparable to AERONET inversion products while providing computational acceleration relative to retrievals performed in the full parameter space. To evaluate its performance, three geographically distinct regions are included in the test. The first is the southern California area (CA; ~150 km domain), centered on the AERONET site at Caltech (lat: 34.14°N, lon: 118.13°W). The second is an area in northern China (~1,500 km domain), centered on the AERONET site at Beijing (BJ; lat: 39.977°N, lon: 116.381°E). The third is a southern Africa area (NB; ~1,400 km domain), centered around the AERONET site at Namibe (lat: 12.178°S, lon: 15.159°E). These three regions were selected to encompass diverse aerosol regimes with region-specific characteristics, and the retrieval tests aim to assess CIMAP-A's performance across varying aerosol compositions and surface conditions.

4.2.1 Algorithm Structure and Configurations of CIMAP-A

CIMAP-A is an optimization-based algorithm that determines aerosol properties by fitting ground-based passive measurements in AERONET spectral bands. Figure 22 illustrates its general algorithmic framework for retrieving AERONET almucantar measurements. The AERONET climatology of aerosol fields, including refractive indices and volumetric particle size distributions, is compiled to construct a database for computing the PCs. The retrieval adopts the MarCH RT model and incorporates a light scattering model to generate modeled signals that are fit to the AERONET measurements of sky radiance (L_λ) and spectral AOD (τ_λ) with the lowest fitting residual. Following an iterative optimization process, the PC weights are adjusted until the optimal solution is achieved.

To adapt CIMAP for eventual implementation with satellite measurements, configuration of CIMAP-A is simplified in this test. Specifically, the Level 2.0 AERONET inversion products are adopted both as the PCA dataset and for validation, owing to their rigorous data quality filtering criteria applied relative to Level 1.5 products. At the same time, the only constraint applied within the retrieval framework is the correlation constraint derived from the PCA. For parameters specific to CIMAP-A, a subscript "A" will be added to distinguish them from other CIMAP configurations.

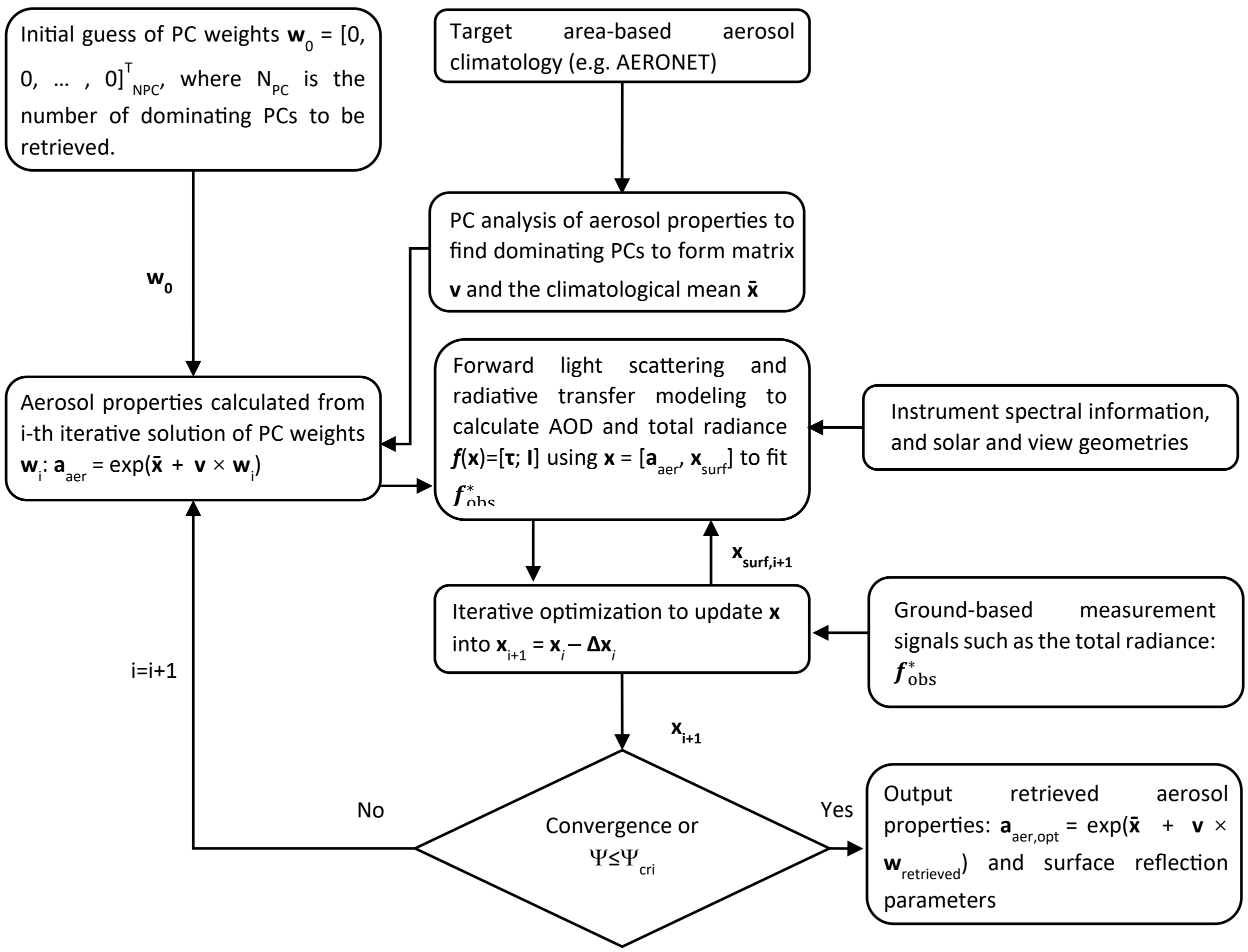


Figure 22. General structure of the CIMAP algorithm for retrieving AERONET measurements.

4.2.1.1 Aerosol A Priori

The correlation in aerosol properties can be captured by PC analysis of the spatially and temporally varying aerosol parameters. The aerosol climatology collected from AERONET data is utilized to generate a training dataset for PC analysis. Incorporating these aerosol products into the training dataset, we perform PC analysis of it in logarithmic space to leverage the difference of magnitudes in aerosol parameters. Recall the general equation Eq. (65), every single record of aerosol fields can be represented by PCs, namely:

$$\mathbf{x}_{\mathrm{A},i} = \log\left(\mathbf{a}_{\mathrm{A},i}\right) = \bar{\mathbf{x}}_{\mathrm{A}} + \mathbf{v}_{\mathrm{A}} \times \mathbf{w}_{\mathrm{A},i} \tag{66}$$

where $\mathbf{a}_{\mathrm{A},i}$ is the i^{th} record of $\mathbf{a}_{\mathrm{aerosol}}$ in the training dataset, $\bar{\mathbf{x}}_{\mathrm{A}}$ is a vector containing the spatial and temporal mean of aerosol parameters under PC analysis, $\mathbf{v}_{\mathrm{A}}$ denotes the PC matrix (each column is associated with a single PC vector), and $\mathbf{w}_{\mathrm{A}}$ is the matrix containing PC weights to multiply with PC vectors. Indeed, performing PC analysis in logarithmic space ensures non-negativity of the solution after reconstructing aerosol properties $\mathbf{a}_{\mathrm{A}}$ from taking the exponential of $\mathbf{x}_{\mathrm{A}}$ in Eq. 66. Moreover, the matrix $\mathbf{v}_{\mathrm{A}}$ in Eq. 66 can be truncated to contain only a desired number of PCs (N_{PC}) to reconstruct aerosol properties ($\mathbf{a}_{\mathrm{A}}$), namely

$$\mathbf{v}_{\mathrm{A}} = \left[\mathbf{v}_{\mathrm{A},1}\,, \mathbf{v}_{\mathrm{A},2}\,, \ldots\ \mathbf{v}_{\mathrm{A},N_{\mathrm{PC}}}\right] \in \mathbb{R}^{L_{\mathrm{aero}} \times N_{\mathrm{PC}}} \tag{67}$$

where $\mathbf{v}_{\mathrm{A},k}$ is k^{th} PC vector containing a total of L_{aero} elements and

$$\mathbf{v}_{\mathrm{A},k} \in \mathbb{R}^{L_{\mathrm{aero}}} \tag{68}$$

Taking p^{th} aerosol property saved in i^{th} aerosol record as an example, the row vector containing dominating PC weights are presented as

$$\mathbf{w}_{\mathrm{A},i} \in \mathbb{R}^{N_{\mathrm{PC}}} \tag{69}$$

We excluded non-spherical particles in our retrieval. Therefore, PC analysis of aerosol climatology was performed only over the cases with volumetric spherical particle fraction more than 95%.

To demonstrate effect of reconstruction of aerosol properties, we perform a PC analysis of the AERONET aerosol climatology reported by multiple sites in southern California. PCA is conducted on a dataset compiled from 21 AERONET sites spanning 670 km across southern California. The northernmost site is NASA Ames, located at 122.058°W, 37.42°N, and the southernmost site is La Jolla, located at 117.250°W, 32.871°N.

The Fig. 23a demonstrates the original aerosol products from the training dataset, including volumetric size distribution resolved into 22 size bins, and refractive index at 440, 670, 870, and 1020 nm. The Fig. 23b demonstrates the mean of aerosol parameters (in the black curve) and the first seven (dominating) PCs (in different colors). Figure 23c demonstrates captured variance associated with 1-10 PCs. It is observed that 78% and 97% of variance are captured by three and seven dominating PCs, respectively. Including more PCs will capture more variance in the original fields. However, due to the uncertainties of these retrieved aerosol properties which originate from measurement uncertainties, adding more PCs does not necessarily guarantee more fidelity to the true correlation in aerosol parameters. To be demonstrated later, we determine the optimal number of PCs based on the cost function which accounts for the measurement uncertainties and on a user specified threshold of percentage variance captured by PCs. Since most variance in the training set can be captured by a small number of dominating PCs, CIMAP is based on retrieval in PC space. In this study, PCs and the mean vector are pre-determined from target-specific training datasets. Retrieval is then implemented to resolve the weights of the few PCs that dominate the correlation.

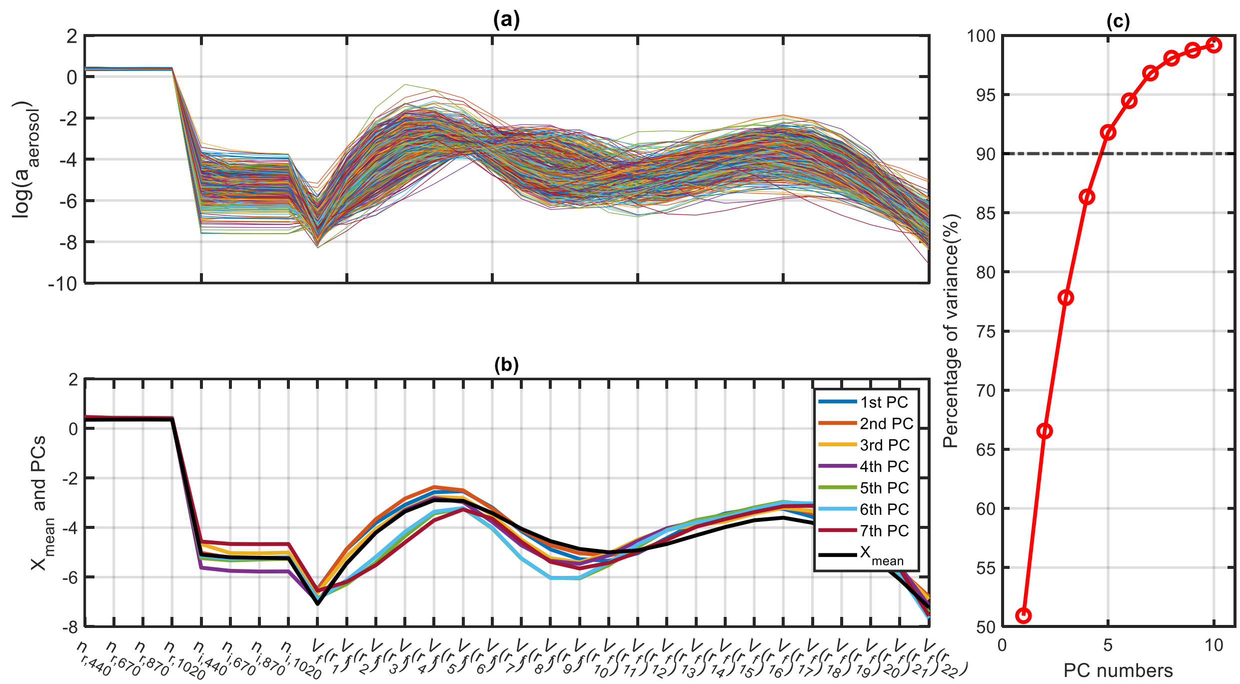


Figure 23. a) An example of AERONET data records in logarithmic scale under PC analysis. The data are reported by multiple AERONET sites in southern California. Every record includes real and imaginary parts of refractive index at four AERONET wavelengths 440, 670, 870, and 1020nm and volumetric size distribution discretized into 22 size bins. b) The spatial and temporal mean of the data record and the reconstructed aerosol properties using leading PCs. c) Percentage of data variance captured by the first 1-10 PCs.

4.2.1.2 State Vector

In the CIMAP-A algorithm, several simplifications are made to facilitate parameterization and convergence. The state vector is restricted to the PC weights ($\mathbf{w}_A$) for the correlated aerosol parameters. The PC weights are initialized to be 0, namely $\mathbf{w}_A = [0,0,\ldots,0]$, so that the climatology mean will be set as the aerosol properties for the RT modeling and light scattering calculation in the 1st iteration. The aerosol layer height is fixed at 2,000 m. Although the CIMAP framework allows for more complex surface reflection to be modeled via the Ross-Li BRDF model (Eq. (14)), its application to AERONET data retrieval assumes a Lambertian surface, such that the spectral albedos reported in the AERONET aerosol inversion products are used directly. This simplification is supported empirically that tests using the BRDF retrieved by the GRASP algorithm [Dubovik et al., 2011] indicate only marginal improvement across all cases in the current study.

4.2.1.3 AERONET Observations

The CIMAP-A uses AERONET radiance at six aerosol and surface bands centered at 440, 670, 870, and 1020nm. The observational vector consists of spectral AODs (τ_{A}) and sky radiances (L) measured by AERONET at multiple angles in almucantar scan configuration. In this case, $\mathbf{f}^{*}_{\mathrm{A,obs}}$ is expressed as

$$\begin{aligned}\mathbf{f}^{*}_{\mathrm{A,obs}} = \log\,[&\tau_{\mathrm{A}}(\lambda_1);\ L(\lambda_1;\theta_1;\phi_1);\ L(\lambda_1;\theta_2;\phi_2);\dots;\ L\big(\lambda_1;\theta_{N_\theta};\phi_{N_\theta}\big);\\ &\tau_{\mathrm{A}}(\lambda_2);\ L(\lambda_2;\theta_1;\phi_1);\ L(\lambda_2;\theta_2;\phi_2);\dots;\ L\big(\lambda_2;\theta_{N_\theta};\phi_{N_\theta}\big);\\ &\tau_{\mathrm{A}}(\lambda_3);\ L(\lambda_3;\theta_1;\phi_1);\ L(\lambda_3;\theta_2;\phi_2);\dots;\ L\big(\lambda_3;\theta_{N_\theta};\phi_{N_\theta}\big);\\ &\tau_{\mathrm{A}}(\lambda_4);L(\lambda_4;\theta_1;\phi_1);\ L(\lambda_4;\theta_2;\phi_2);\ \dots;\ L\big(\lambda_4;\theta_{N_\theta};\phi_{N_\theta}\big)]\end{aligned} \tag{70}$$

where in λ_1 = 440nm, λ_2=670nm, λ_3=870nm, and λ_4=1020nm, N_θ is the number of view angles at which sky radiances are measured. To evaluate the weighting matrix, the absolute uncertainty of AOD measurement is assumed to be 0.015 and the relative uncertainty of radiance measurement is assumed to be 5%. In the above equation, the radiance is converted into logarithmic space to leverage the differences in the magnitudes of signals.

4.2.1.4 Inversion Formulation

A numerical inversion module is used to adjust aerosol parameters so that the modeled spectral sky radiances fit observations with the lowest residue. As the kernel of our inversion module, we use the Levenberg-Marquardt (LM) algorithm [Levenberg, 1944; Marquardt, 1963]. The LM algorithm combines the strengths of the Gauss-Newton method and gradient descent method by incorporating a regularization term into the normal equation system. The optimal solution is approached in an iterative way, namely the increment of solution ($\Delta\mathbf{w}_{A,q}$) following q^{th} iteration is derived by solving the following linear system of equations in matrix form:

$$\left[\mathbf{J}_A^T\,\mathbf{W}_A\,\mathbf{J}_A + \lambda_q\,\mathrm{diag}(\mathbf{J}_A^T\,\mathbf{W}_A\,\mathbf{J}_\mathbf{A})\right]\Delta\mathbf{w}_{A,q} = \mathbf{J}_A^T\,\mathbf{W}_A\left[\mathbf{f}_A(\mathbf{w}_{A,q}) - \mathbf{f}_{A,obs}^*\right] \tag{71}$$

where $\mathbf{J_A}$ is the Jacobian matrix with its element evaluated by $J_{A,ij} = \partial f_{A,i}(\mathbf{x}_{A,q})/\partial x_j$ and $\mathbf{W_A}$ is a weighting matrix with its $(i, i)^{th}$ diagonal element evaluated by

$$\mathrm{W}_{A,ii} = c \cdot \sigma_{A,ii}^{-2} \tag{72}$$

In the above equation, c is a constant number equals to $N_\tau/2$ and $N_L/2$ for AERONET AOD and radiance signals, respectively, where N_τ and N_L are the total numbers of AOD and radiance signals, respectively. Assuming no correlations among errors in observational signals, the off-diagonal terms of the weighting matrix is zero, $\mathbf{f}_{A,obs}^*$ is the observational vector containing all measurements including AODs and sky radiances, $\mathbf{f}_A(\mathbf{w}_{A,q})$ is the model fit with q^{th} iterative solution of PC weights $\mathbf{w}_{A,q}$ which is associated with the aerosol solution $\mathbf{a}_{A,q}$. This essentially means the LM algorithm is similar to the steepest descent method. The matrix $\mathrm{diag}(\mathbf{J}_A^T\,\mathbf{W}_A\,\mathbf{J}_\mathbf{A})$ in Eq. (71) was suggested by Fletcher (1971). It scales each component of the gradient according to the curvature to ensure larger movement along the directions where the gradient is smaller. The scaling operation avoids slow convergence in the direction of a small gradient. There are many discussions about the proper choice of the damping parameter λ_q. We followed the recommendation of Marquardt [Marquardt, 1963] by initializing the damping factor with a value λ_0 and a scale factor $v > 1$ and computing the residual cost function after an iteration. We then compare the cost functions of two successive iterations. If the cost function is reduced with the updated solution

$$\mathbf{w}_{\mathrm{A},q+1} = \mathbf{w}_{\mathrm{A},q} - \Delta\mathbf{w}_{\mathrm{A},q} \tag{73}$$

then λ_0 is updated to be λ_0/v for the next iteration; otherwise, λ_0 is updated to be $\lambda_0 v$. In our current study, the scale factor v is empirically determined as 10. As a measure of the degree of fitting, the Chi-square function form (χ^2_{A}) of the cost function is defined with respect to the measurement uncertainties associated the solution from q^{th} iteration:

$$\chi^2_{\mathrm{A}}(q) = \left[\mathbf{f}_{\mathrm{A}}(\mathbf{w}_{\mathrm{A},q}) - \mathbf{f}^*_{\mathrm{A,obs}}\right]^{\mathrm{T}} \mathbf{W}_{\mathrm{A}} \left[\mathbf{f}_{\mathrm{A}}(\mathbf{w}_{\mathrm{A},q}) - \mathbf{f}^*_{\mathrm{A,obs}}\right]. \tag{74}$$

4.2.2 Retrieval Implementation and Demonstration for CIMAP-A

4.2.2.1 Data

To evaluate CIMAP-A's retrieval performance across distinct aerosol regimes, we test its retrieval of anthropogenic aerosols, particulate air pollutants, biomass burning aerosols using a total of 60 sets of AERONET measurements drawn from three regions: a southern California area (CA), a northern China area around Beijing (BJ), and a southern Africa area (NB). Twenty aerosol scenarios are selected for each region, collectively spanning a broad range of aerosol loading, absorption, and size distribution conditions. Figure 24 displays the geographical locations of the AERONET sites from which the 60 measurements are collected, with the CA, BJ, and SA areas shown, respectively.

For the CA area, the selected AERONET sites are Caltech, UCLA, Santa Monica, El Segundo, and MISR-JPL. For BJ area, the sites are Beijing, XiangHe, AOE Baotou, Xinglong, Yulin, and Yufa PEK. For NB area, the sites are Namibe, Mwinilunga, Lubango, Windpoort, Tsumkwe, and Huambo. Each region is treated separately to preserve the spatial correlations required by the PCA.

It should be noted that the training dataset described in Section 4.2.1.1 and shown in Fig. 23 differs from the CA dataset used here for CIMAP-A retrieval demonstration. Specifically, the latter is a spatial subset of the training dataset, reduced in climatological scope to target a more precise area around Los Angeles (lon: 118.28°W, lat: 34.038°N).

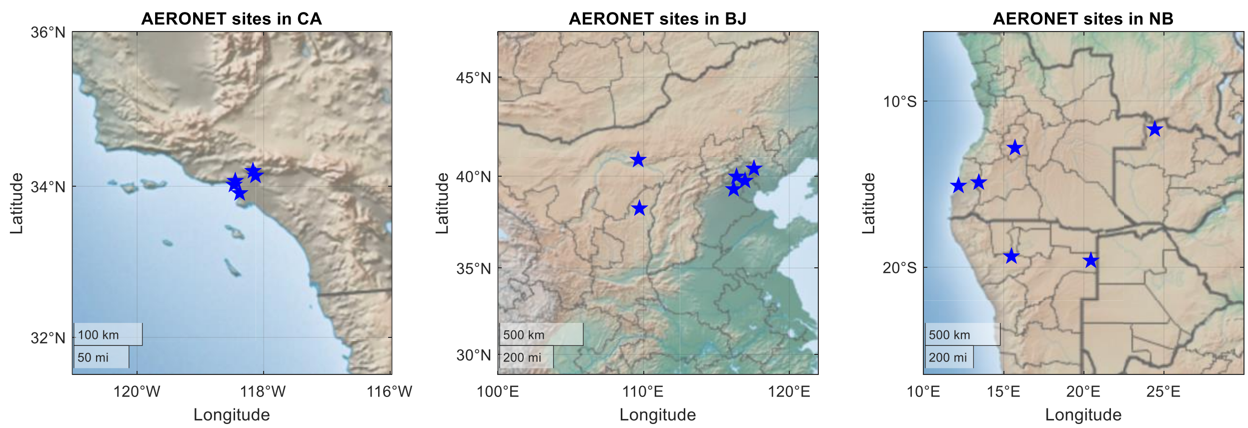


Figure 24. a) Geographical locations of selected CA region including AERONET site Caltech, UCLA, Santa Monica, El Segundo, and MISR-JPL (~150km domain). b) Geographical locations of selected BJ region including AERONET site Beijing, XiangHe, AOE Baotou, Xinglong, Yulin, and Yufa PEK (~1400km domain). c) Geographical locations of selected NB region including AERONET site Namibe, Mwinilunga, Lubango, Windpoort, Tsumkwe, and Huambo (~1400km domain).

4.2.2.2 Determination of Optimal Number of PCs

The N_{PC} adopted by CIMAP-A is a critical configurational parameter that directly governs the balance between retrieval efficiency and the strength of the correlation constraint. In the extreme case where the number of PCs equals the number of aerosol parameters (i.e., $N_{PC} = L_{aero}$), the algorithm essentially degenerates to assuming no correlations in aerosol properties. This is because an arbitrary set of aerosol parameters can be constructed from L_{aero} mutually orthogonal PCs, rendering the retrieval parameter space identical to that in the traditional AERONET optimization and yielding no gain in retrieval efficiency. On the other hand, if too few PCs are involved in the retrieval, the imposed correlation constraint becomes too strong, making the retrieval more susceptible to errors in the pre-determined PC vectors from the training dataset for the targeted area. These two extremes together highlight that the determination of an optimal N_{PC} must carefully balance the capitalization of correlations in aerosol properties and the maximization of inversion efficiency. To this end, we perform a test of CIMAP retrieval using different numbers of PCs and investigate the dependence of the fitting residual (evaluated by Eq. (74)) on the number of PCs. Based on this analysis, CIMAP-A adopts two criteria for choosing the optimal N_{PC}: 1) the fitting residual of the retrieval utilizing NPC falls below the instrumental measurement uncertainties (i.e., $\leqslant 1$); and 2) more than 95% of the aerosol variance over the targeted area is captured by the PCs.

Figure 25 demonstrates these criteria across the three study regions. The upper row represents the average fitting residual over radiance and AOD, where the vertical bars indicate the standard deviation of the residual from the statistics of 20 selected retrieval cases. The lower row represents the captured variance from the first 10 PCs. For the CA region, using more than 6 PCs brings the fitting residual within measurement uncertainties, while using 7 PCs keeps the captured variance above 95%. Satisfying both criteria, 7 PCs are therefore adopted for CIMAP-A of AERONET observations over the CA region. For the BJ and NB regions, the first 8 and 4 PCs ensure a fitting residual of $\chi_{\mathrm{A}}^2 \leq 1$, respectively. A similar retrieval test indicates that the first 7 PCs capture more than 95% of the aerosol variance in both the BJ and NB regions. Taking the larger of the values satisfying both criteria, the first 8 and 7 PCs are therefore adopted as the optimal PC numbers for CIMAP retrieval over the BJ and NB regions, respectively. The selected N_{PC} will be used to conduct CIMAP-A retrievals for the detailed retrieval quality assessment.

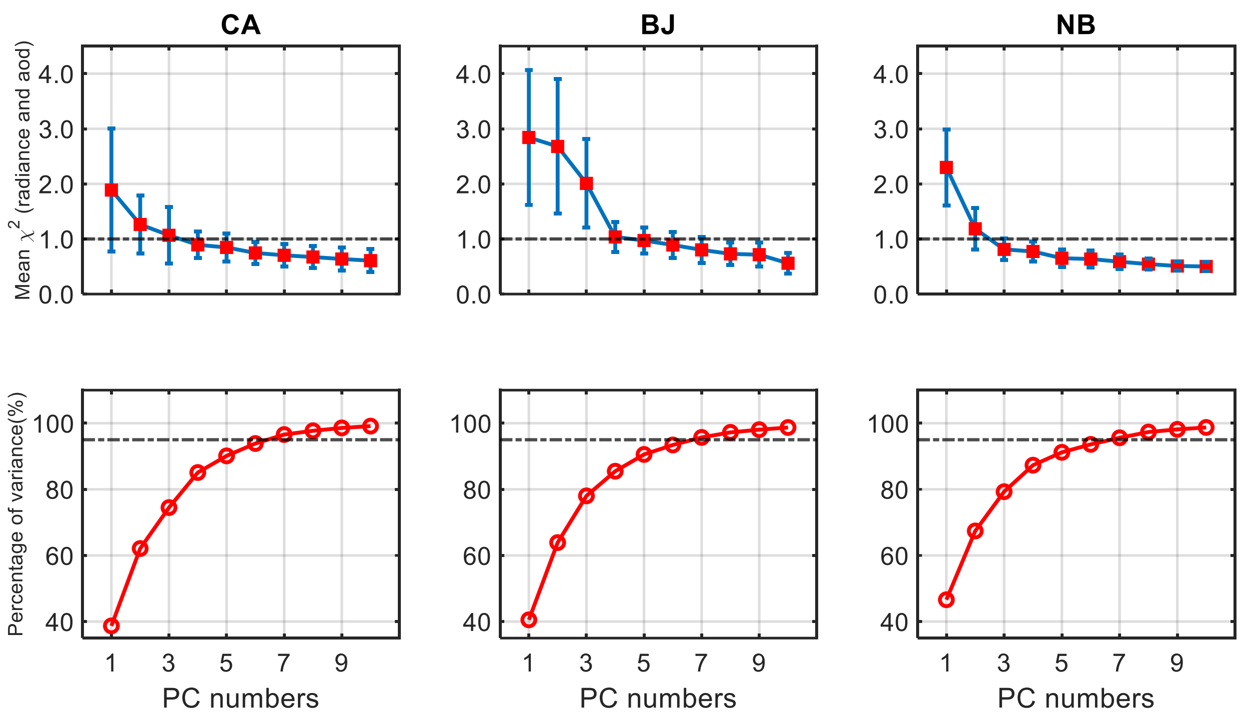


Figure 25. Upper panels: Chi-square (χ^2) fitting errors as a function of numbers of PCs adopted in CIMAP retrieval. The square dots are the average of a total of 20 randomly selected retrievals of AERONET measurements in the CA, BJ, and NB regions. The vertical bar indicates the spread (1s) of the fitting errors. Bottom panels: Percentage of variance captured by the specific PCs CA, BJ, and NB regions. The black dash lines indicate the threshold of 95 % captured variances.

4.2.2.3 Single-Case Analysis

We begin by presenting a single-case CIMAP-A retrieval for each selected region. Figure 26 shows CIMAP-A retrievals of aerosol properties derived from AOD and sky radiance measurements acquired at the AERONET UCLA site (lon: 118.45°W, lat: 34.07°N) at 00:15:16 UTC on October 10, 2006. In the upper two panels, CIMAP-A retrievals of the real and imaginary parts of the refractive index at the four AERONET wavelengths (440, 670, 870, and 1020 nm) are compared against AERONET reference products. The middle panel compares the retrieved volumetric aerosol size distributions, and the bottom panel compares CIMAP-A retrievals of spectral AOD and SSA against AERONET reference data. In all panels, AERONET and CIMAP-A results are plotted in black and red, respectively.

The retrieval results show good overall agreement with the AERONET reference. The differences in the real and imaginary parts of the refractive index are less than 0.02 and 0.015, respectively. For the bimodal volumetric size distribution, CIMAP-A's fine-mode radius differs slightly to the AERONET (by approximately 0.05 μm), while both indicate a peak coarse-mode aerosol concentration near 3 μm. For spectral AOD, CIMAP-A retrievals agree closely with the AERONET measurements, with absolute differences falling within the measurement uncertainty of 0.015. For SSA, shown in the right panel, the agreement is likewise reasonable: the differences remain within the AERONET-reported uncertainty of 0.03 for non-absorbing aerosols (AOD at 440 nm > 0.2) and 0.04 for absorbing aerosols (AOD ⩾ 0.5) [Dubovik et al., 2000; Holben et al., 2006].

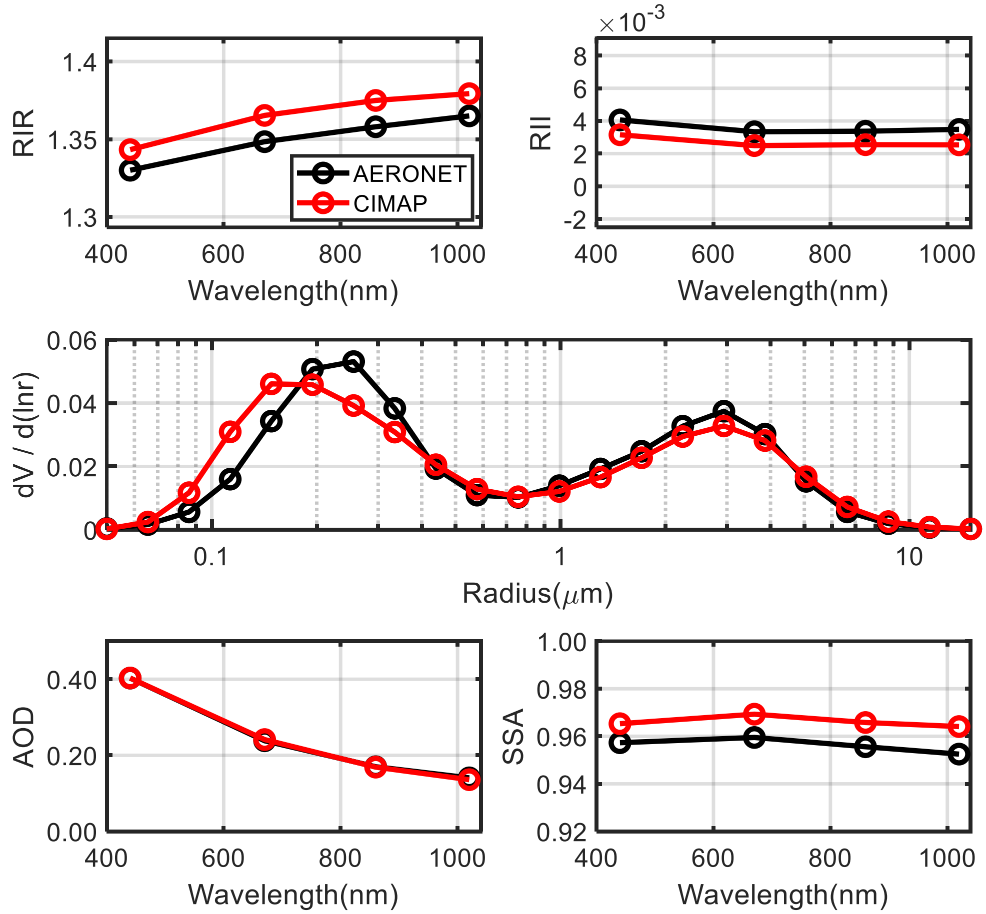


Figure 26. The aerosol refractive index, size distribution, AOD, and SSA retrieved from AOD and sky radiance measurements (almucantar mode) acquired at the AERONET UCLA site. The UTC time of acquisition is 00:15:16 of October 10 of 2006. CIMAP results are plotted in red curve with empty circles. AERONET operational retrievals are plotted in black curve with empty circles. Upper left panel: RIR at the four AERONET wavelengths 440, 670, 870, and 1020nm; Upper right panel: RII; Middle panel: volumetric aerosol size distribution; Bottom left panel: AOD at the four AERONET wavelengths; Bottom right panel: SSA at the four AERONET wavelengths.

Using the retrieved aerosol abundance and properties, we calculate the modeled sky radiances and compare them to the almucantar measurements in Fig. 27. Reasonable agreement is found across all four AERONET wavelengths, indicating that the retrieved aerosol state is physically consistent with the observed angular radiance distribution. The relative differences between modeled and measured radiances, though not plotted here, remain within the assumed measurement uncertainty of 5%, further confirming the fidelity of the forward model fit and the robustness of the retrieval solution.

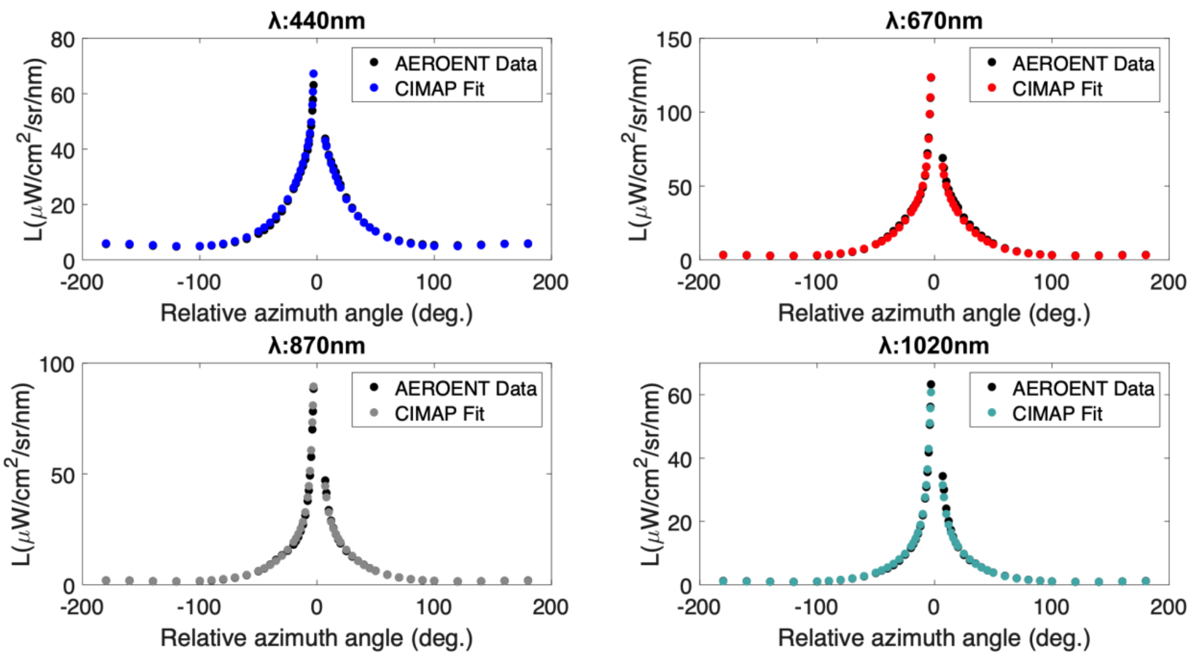


Figure 27. Comparison of modeled sky radiances to the measurements by the almucantar. The modeled radiances are calculated from the retrieved solution. The comparison of radiances in upper left, upper right, bottom left and bottom right panels are for AERONET wavelength 440, 670, 870, and 1020 nm, respectively.

4.2.2.4 Retrieval Validation Across the Full Dataset

Following the case study, we demonstrate the CIMAP-A retrieval of all cases in three areas against AERONET. AOD, SSA, and effective radius are included in the comparison. Figure 28 demonstrates the comparison of AOD at 440, 670, 870, and 1020 nm. Linear regression fitting, MAD, STD, R^2, and root mean square error (RMSE) are reported as statistics. All three regions show strong consistency compared to AERONET. Most of the scatter points lie on the 1:1 line, with a MAD of 0.005, which is within the measurement uncertainties. Figure 29 demonstrates the comparison of SSA at four AERONET wavelengths. Similar to the AOD quality, the retrieved SSA shows strong agreement with AERONET, with a MAD of 0.019. The MAD is within the AERONET retrieval uncertainties of SSA (0.0325), which are averaged from the uncertainties estimated for weakly absorbing aerosols (0.025) and absorbing aerosols (0.04) in [Dubovik et al., 2000] when AOD at 440 nm is larger than 0.5.

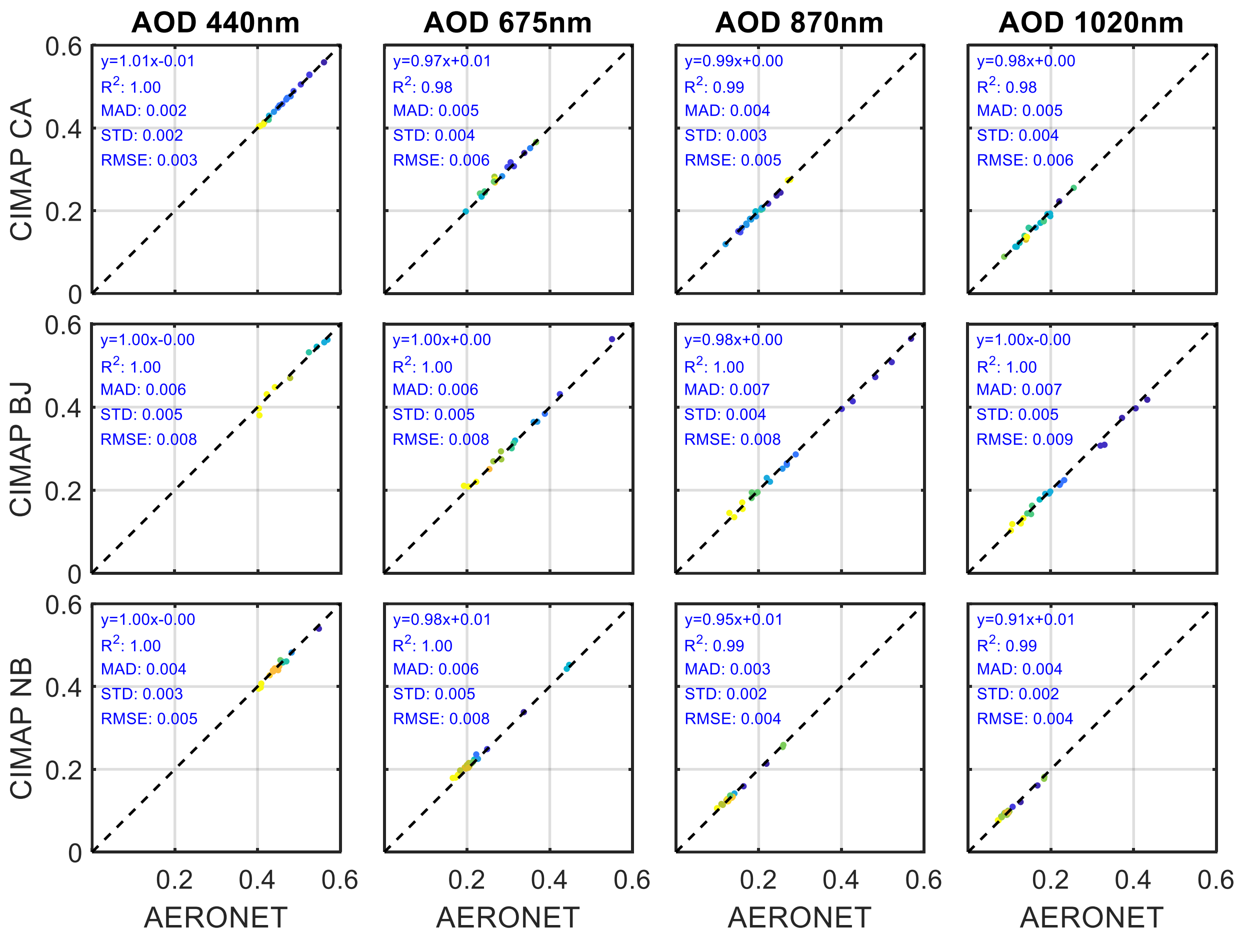


Figure 28. Comparisons of retrieved AOD at 440, 670, 870, and 1020nm. Statistics are labeled in the upper left corner of each subplot. Comparisons of CIMAP-A AOD to AEORNET data in CA, BJ, and NB regions are present in the first, the second, and the third row, respectively. The color represents point density, where warmer color (yellow) indicates higher concentrations of data and cooler color (blue) represent lower density. A total of 20 matchups is displayed in each region.

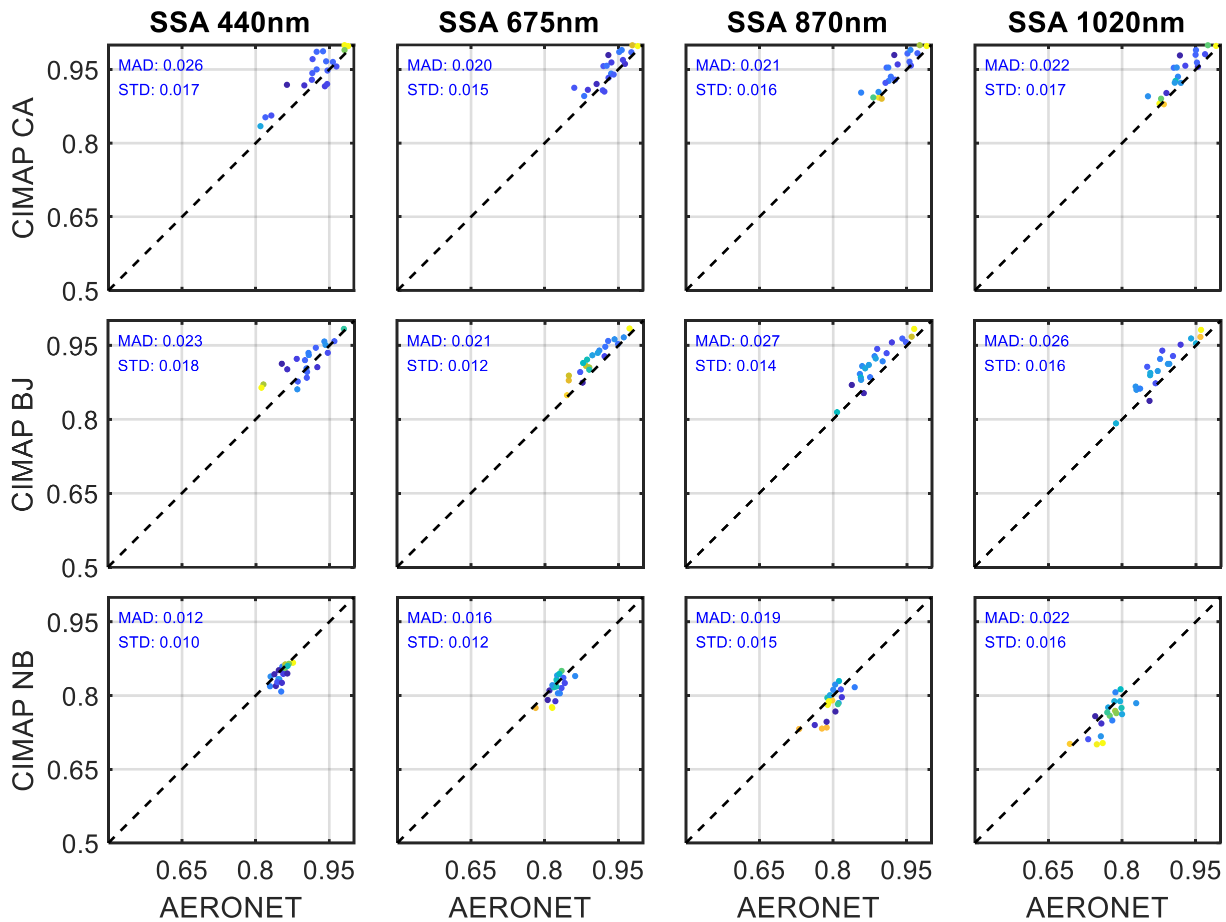


Figure 29. Comparisons of retrieved SSA at 440, 670, 870, and 1020nm. Statistics are labeled in the upper left corner of each subplot. Comparisons of CIMAP-A SSA to AEORNET data in CA, BJ, and NB regions are present in the first, the second, and the third row, respectively. The color represents point density, where warmer color (yellow) indicates higher concentrations of data and cooler color (blue) represent lower density. A total of 20 matchups is displayed in each region.

To measure the quality of the retrieved aerosol size, we further calculate the effective radii of fine mode aerosols ($r_{\text{eff,fine}}$), coarse mode aerosols ($r_{\text{eff,coarse}}$), and compare them to AERONET reference data. The equation for effective radius calculation is:

$$r_{\text{eff}} = \frac{\int_{\mathrm{In}r_{\min}}^{\mathrm{In}r_{\max}} \frac{dV(r)}{d\ln r} d\ln r}{\int_{\mathrm{In}r_{\min}}^{\mathrm{In}r_{\max}} \frac{1}{r} \frac{dV(r)}{d\ln r} d\ln r} \tag{75}$$

where the quadrature bounds, $r_{\min}$ and $r_{\max}$, depend on the particle size mode. Fine and coarse mode aerosols are distinguished by the relative minimum of size distribution frequency within the range of 0.439 to 0.992 µm of size interval.

Figure 30 demonstrates the comparison of fine and coarse mode effective radius. Most scatter points lie closely to the 1:1 line. Few outliers are found. For example, the coarse mode effective radius of the BJ region has a larger MAD. This may be due to the complex aerosol scenario in the BJ region, as CIMAP itself is also subject to retrieval uncertainties. Overall, the MAD is 0.013 for fine mode and 0.2 for coarse mode effective radius, which indicates high retrieval quality.

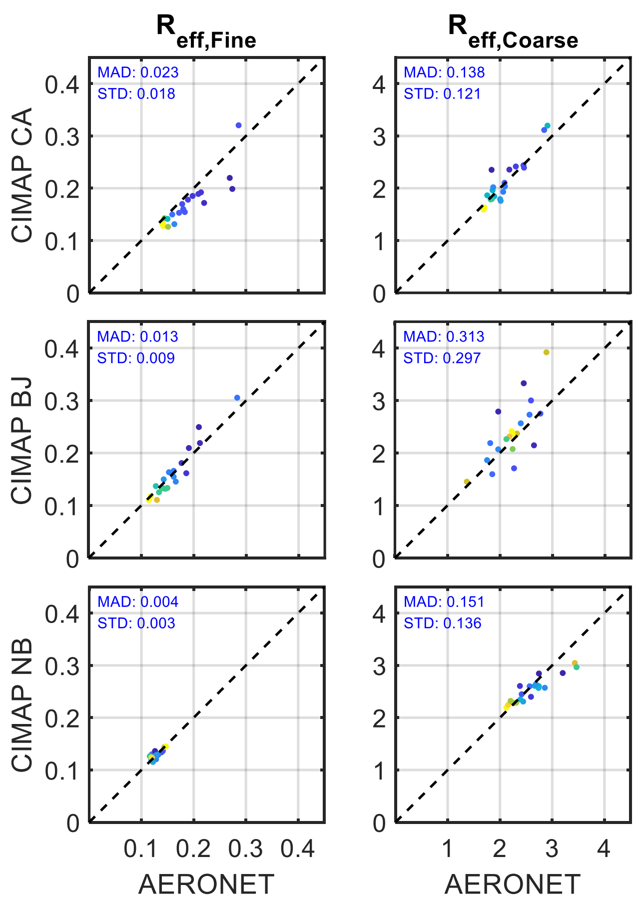


Figure 30. Comparisons of retrieved Effective radii for fine and coarse mode volumetric size distribution. Statistics are labeled in the upper left corner of each subplot. Comparisons of CIMAP-A AOD to AEORNET data in CA, BJ,

and NB regions are present in the first, the second, and the third row, respectively. The color represents point density, where warmer color (yellow) indicates higher concentrations of data and cooler color (blue) represent lower density. A total of 20 matchups is displayed in each region.

Overall, the comparison of CIMAP and AERONET retrievals for all three areas indicates similar retrieval accuracies of the two algorithms. Table 4 summarizes the MADs of CIMAP-A retrievals and AERONET reference data of AOD, SSA, real and imaginary parts of refractive index, and effective radius of both modes of size distribution. The MADs are found to be 0.005, 0.019, 0.039, 0.003 and 0.013 μm and 0.02 μm for these quantities, respectively, which are mostly within or close to AERONET measurement or retrieval uncertainties.

Table 4. Mean absolute differences of CIMAP and AERONET retrieval aerosol quantities.

Aerosol quantities	Southern California	Southern Africa	North China	All-case & all-site mean
AOD (440 nm)	0.002	0.004	0.006	0.005 (incl. wavelength mean)
AOD (670 nm)	0.005	0.006	0.006	
AOD (870 nm)	0.004	0.003	0.007	
AOD (1020 nm)	0.005	0.004	0.007	
SSA (440 nm)	0.026	0.012	0.023	0.019 (incl. wavelength mean)
SSA (670 nm)	0.020	0.016	0.021	
SSA (870 nm)	0.021	0.019	0.027	
SSA (1020 nm)	0.022	0.022	0.026	
RIR (440 nm)	0.035	0.047	0.052	0.039 (incl. wavelength mean)
RIR (670 nm)	0.025	0.046	0.042	
RIR (870 nm)	0.022	0.049	0.039	
RIR (1020 nm)	0.019	0.051	0.039	
RII (440 nm)	0.004	0.005	0.005	0.003 (incl. wavelength mean)
RII (670 nm)	0.002	0.004	0.003	
RII (870 nm)	0.002	0.004	0.003	
RII (1020 nm)	0.002	0.004	0.003	
Effective radius (fine mode, μm)	0.023	0.004	0.013	0.013
Effective radius (coarse mode, μm)	0.138	0.151	0.313	0.200

Comparing to traditional direct retrieval of 30 aerosol parameters in regular parameter space, CIMAP capitalizes on the aerosol correlation and retrieves 7-8 PC weights only. Therefore, it improves the retrieval efficiency. To get insight into CIMAP's gain of retrieval efficiency, we implemented an extra retrieval using the maximum number (30) of PCs. Considering the mutual orthogonality of PC vectors, 30 PC based retrieval essentially means no correlation is invoked so that the retrieval is considered as a proxy of traditional retrieval in regular parameter space. We then calculate the ratio (R) of retrieval time cost (t) from $R=t_{nPC}/t_{30PC}$, where N_{PC}=7, 8, and 7 for CIMAP retrievals of measurements over CA, BJ, and NB, respectively.

Figure 31 demonstrates the "R" values for all 60 cases. The mean retrieval time of CIMAP was 13%, 20%, and 21% of that of the non-correlation-based retrievals for the three study areas, respectively, yielding an average reduction of ~82% in computational cost. Taken together with the retrieval quality demonstrated in the preceding sections, these results indicate that CIMAP-A achieves both high stability and strong computational efficiency.

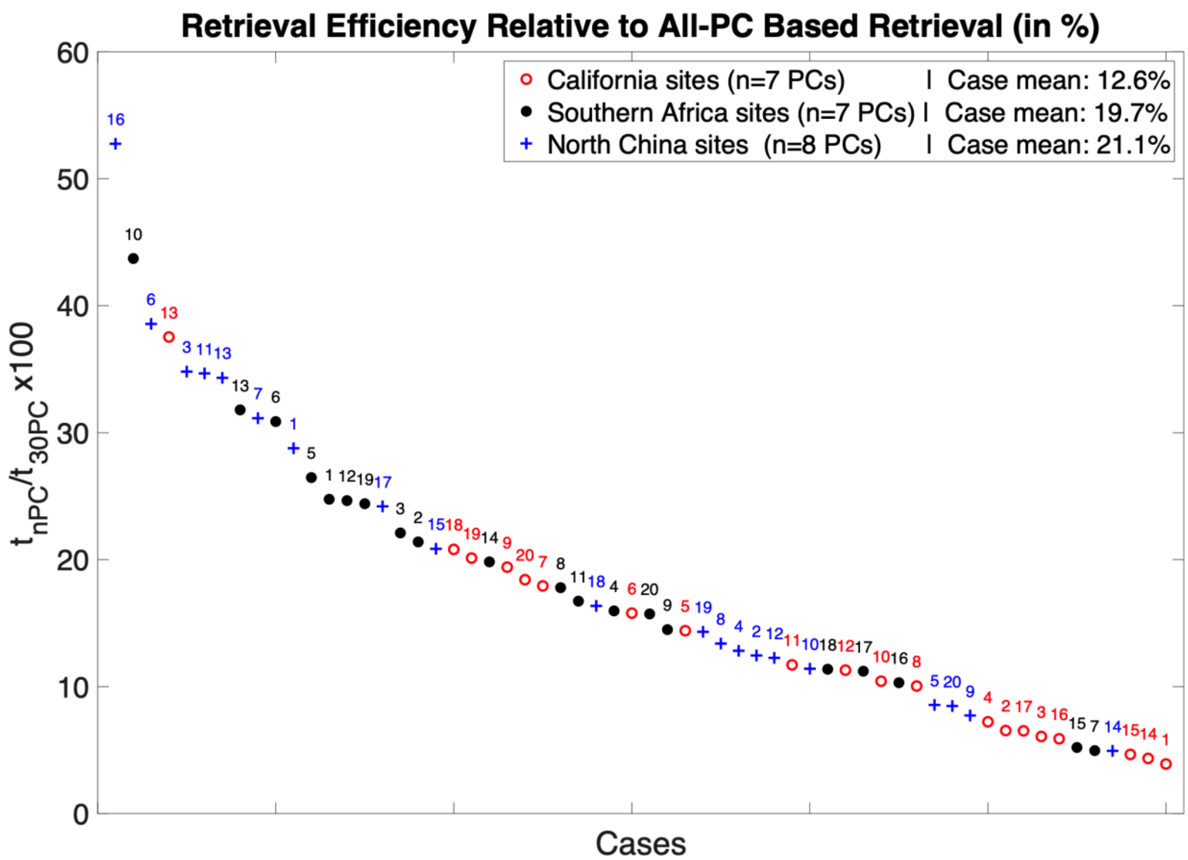


Figure 31. The ratio of retrieval time cost from using 7 or 8 PCs against using 30 PCs. Red, black and blue markers and text correspond to the selected retrievals of AERONET measurements in southern California, southern Africa, and north China, respectively. The numbers above markers are retrieval case numbers in each region.

4.2.3 Summary for CIMAP-A

By capturing the correlation in aerosol properties, a principal-component based inversion approach – CIMAP-A, is developed to retrieve aerosol properties including refractive index and volumetric size distribution from AERONET measurements of AODs and sky radiances. With the adoption of 7-8 PCs, CIMAP retrievals of AERONET measurements in several selected areas in CA, BJ and NB show that it significantly reduces the retrieval parameter space and saves the retrieval time cost by >80%. The MAD between CIMAP retrievals and AERONET products are 0.005, 0.019, 0.039, 0.003 and 0.013 μm and 0.02 μm for AOD, SSA, real and imaginary parts of refractive index, and effective radius of both modes of size distribution, respectively. The differences in CIMAP and AERONET retrievals are generally within or close to AERONET measurement or retrieval uncertainties.

CIMAP's retrievals of ground-based measurements by AERONET in this work and airborne observations by AirMSPI in an earlier work [Xu et al., 2019] form a stepstone toward retrieving satellite-borne measurements which involves the inversion of a large number of imaging pixels, but less view angles and complex radiative transfer conditions. In this case, both retrieval accuracy and efficiency are important concerns. In addition to the retrieval efficiency gain by reducing retrieval parameter space, CIMAP also offers a higher degree of flexibilities to interface with diverse sources of a priori information about aerosols (e.g. climatology, transport models, reanalysis, etc) as long as these sources can inform the retrieval with a proper initialization of PCs. Under such a context, CIMAP is expected to mitigate the ill-posedness which arises from retrieving a large set of aerosol parameters (such as aerosol species abundance and properties) in regular parameter space when observation alone contains insufficient information.

While CIMAP-A adopted several simplifications in its current configuration, future CIMAP development will target more physically complete retrievals. Specifically, aerosol layer height and particle sphericity are expected to be included in the state vector, and the anisotropy of surface reflection will need to be accounted for simultaneously. Additional constraints are also anticipated. For instance, many retrieval algorithms such as operational AERONET aerosol retrieval impose smoothness constraints on the spectral variation of the aerosol refractive index and on the variation of aerosol volume concentration across 22 size bins [Dubovik and King, 2000]; same smoothness constraints have been formulated within the correlation-based retrieval framework and validated using airborne observations [Xu et al., 2019]. In the next stage of CIMAP design, we will evaluate the effect of these additional constraints on the accuracy and efficiency of CIMAP.

4.3 CIMAP for Satellite-Borne POLDER Observation

The preliminary successes in retrieving ground-based and airborne data provide a strong foundation for extending the CIMAP algorithm to satellite-borne remote sensing observations. Ground-based and airborne platforms, while valuable, are inherently limited in their spatial coverage and temporal sampling, making it difficult to capture the global distribution of aerosols at the scales required for climate studies. Satellite-borne platforms, by contrast, offer continuous, large-scale observations that are critical for monitoring aerosol properties across diverse geographic regions and atmospheric conditions. As a result, launching satellite-borne platforms equipped with radiometry and polarimetry instruments has become a growing trend in recent and future instrumental implementation. Given this trajectory, applying CIMAP to retrieve satellite-borne observations is not merely an optional extension, but rather an essential and logical next step, both for keeping pace with the evolution of remote sensing instrumentation and for demonstrating the algorithm's broader capability to process current and future observations at a global scale.

While satellite remote sensing provides broad spatial coverage that is of high value for constraining the climate impacts of aerosols, these observations often suffer from limited information content to simultaneously constrain the large set of aerosol and surface parameters required for accurate retrievals [Dubovik et al., 2021]. Incorporating a priori information into the retrieval process is therefore a natural and effective strategy to compensate for this information deficit. By systematically introducing constraints derived from prior knowledge of aerosol properties, the CIMAP framework is well-suited to mitigate this challenge and stabilize the retrieval toward physically meaningful solutions.

In this section, a second CIMAP configuration, CIMAP-P, is demonstrated. CIMAP-P retrieves aerosol properties from the satellite-borne POLDER instrument, which acquired both radiometric and polarimetric measurements at multiple angles and wavelengths. Spanning 2005–2013, the POLDER record constitutes the longest continuous archive of spaceborne polarimetric observations, whose polarimetric component supplies additional retrieval constraints beyond those available from radiance alone, and is therefore expected to mitigate the ill-posedness in the aerosol retrieval.

This study selected the Southwestern Coast of the US (SWCAUS; latitude: 33.07°N–42.06°N, longitude: 128.73°W–116.68°W) as the study region, where observations, the PC-based a priori, MAIAC surface a priori, and validation data are all collocated in this area. Unlike CIMAP's previous applications, which capitalized on the correlations among different types of aerosol properties using a unified set of PCs, this study configures CIMAP-P to capture the correlations within each individual aerosol property separately. This refined configuration allows the internal structure of each aerosol property to be more faithfully represented, enabling the algorithm to leverage more targeted and precise a priori constraints, and thereby improving satellite-borne aerosol retrieval accuracy.

4.3.1 Algorithm Structure and Configurations of CIMAP-P

CIMAP-P is an optimization-based algorithm that determines aerosol properties by fitting satellite-borne passive measurements in aerosol-sensitive spectral bands. Figure 32 illustrates its general algorithmic framework for retrieving multi-angle polarimetric measurements. Both aerosol and surface a priori constraints are incorporated into the retrieval. The AERONET climatology of aerosol fields including refractive indices and volumetric particle size distributions, is used to construct a database for computing the PCs for each type of aerosol property. In addition to using aerosol a priori through PCA, CIMAP has also built in an option of adopting surface a priori such as those from MAIAC [Lyapustin et al. 2018] to improve the aerosol retrieval accuracy. The MarCh model [Xu et al., 2011, 2012] is used for the forward modeling. The inversion then solves for the state vectors iteratively by minimizing the Chi-square calculated between the observations and the modeled signals under the convergence is achieved. For parameters specific to CIMAP-P, a subscript "P" will be added to distinguish them from other CIMAP configurations.

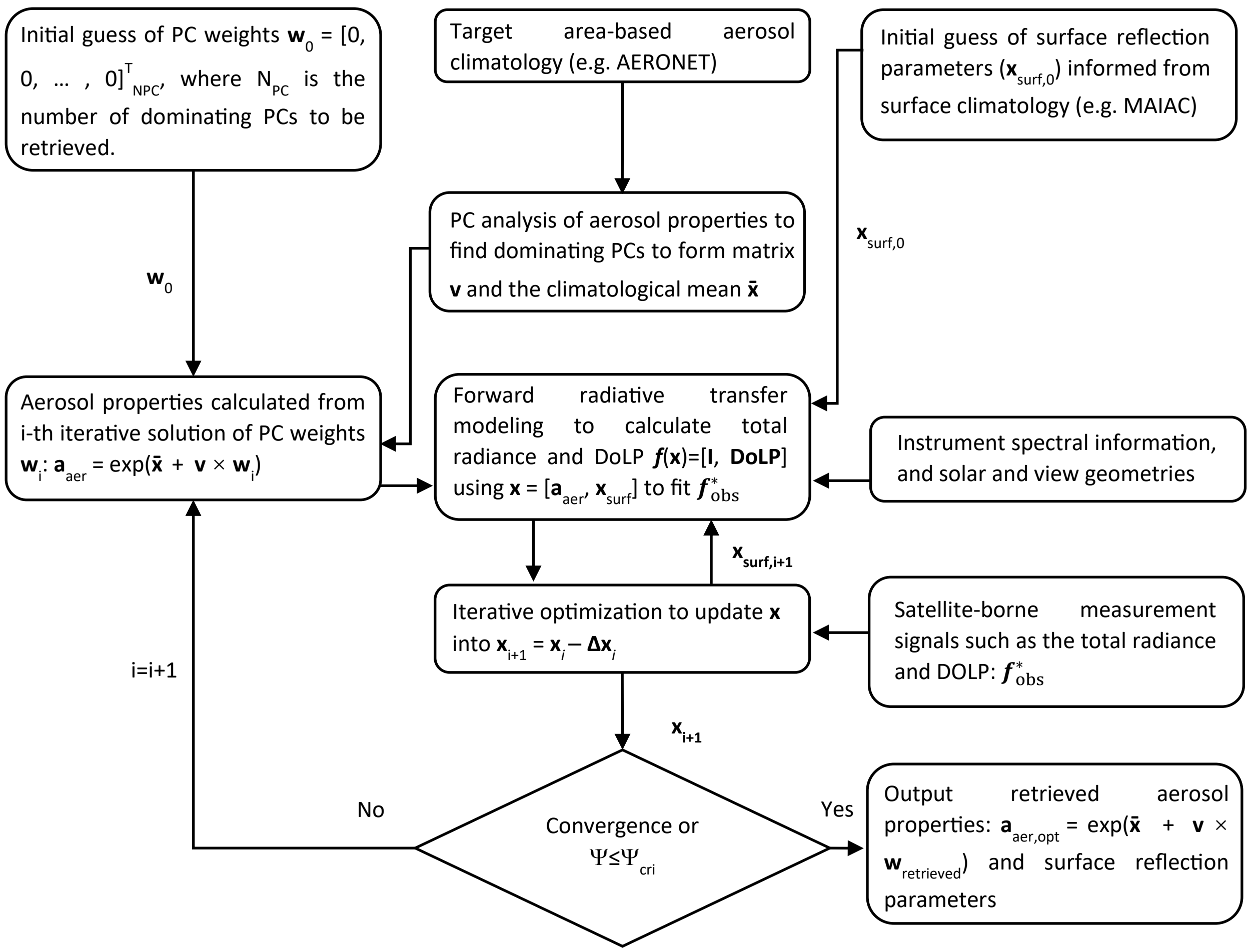


Figure 32. General structure of the CIMAP-P algorithm.

4.3.1.1 Aerosol A Priori

In this study, we selected the SWCAUS as the target region to capitalizes on representative aerosol records available from AERONET climatology. This enables a comprehensive spatial and temporal coverage of aerosol properties.

To effectively represent the correlations in aerosol properties, aerosol climatology is classified across a variety of surface types and seasonal cycles for the target area. To account for the seasonal cycles, we choose spring (March–May), summer (June–August), fall (September–November), and winter (December–February). Then for each season, the classification is refined referring to the underlying surface type. For this purpose, the Normalized Difference Vegetation

Index (NDVI), as a proxy for surface typing, is employed. NDVI leverages the distinct spectral signatures of various surface types derived from visible and near-infrared reflectance [Tucker 1979]. We adopt the NDVI data provided by MOD13C2 NDVI dataset derived from the products of MODIS [King et al. 1992]. The dataset provides monthly NDVI with a resolution of 0.05 degree [Huete and Justice, 1999; Didan and Barreto-Muñoz, 2019]. In our study, three NDVI thresholds are set for land surface categorization: low NDVI with values ≤0.2 (e.g. barren areas), moderate NDVI with values between 0.2 and 0.6 (e.g. shrub and grassland), and high NDVI with values greater than 0.6 (e.g. forest terrain) [Pettorelli, 2013].

Following aerosol classification, PCA is implemented for every combination of season and surface type to capture correlations in aerosol properties. In our study, AERONET version 3 level 1.5 inversion product [Dubovik and King, 2000; Dubovik et al., 2000; Dubovik et al., 2002] from 1994 to 2022 are used as aerosol climatology. The aerosol properties include real and imaginary parts of aerosol refractive index, and volumetric PSD of fine and coarse mode particles. For i^{th} aerosol record we have $=\mathbf{a}_{\mathrm{P},i} = \left[\mathbf{a}_{\mathrm{P},i}^{(\mathrm{RIR})}, \mathbf{a}_{\mathrm{P},i}^{(\mathrm{RII})}, \mathbf{a}_{\mathrm{P},i}^{(\mathrm{PSD_{fine}})}, \mathbf{a}_{\mathrm{P},i}^{(\mathrm{PSD_{coarse}})}\right]$, where $\mathbf{a}_{\mathrm{P}}^{(p)}$ represents a specific type of aerosol property. Specifically, we have

$$\mathbf{a}_{\mathrm{P}}^{(\mathrm{RIR})} = \left[\mathrm{n_r}(\lambda_1), \mathrm{n_r}(\lambda_2), \ldots, \mathrm{n_r}\left(\lambda_{N_\lambda}\right)\right] \tag{76}$$

$$\mathbf{a}_{\mathrm{P}}^{(\mathrm{RII})} = \left[\mathrm{n_i}(\lambda_1), \mathrm{n_i}(\lambda_2), \ldots, \mathrm{n_i}\left(\lambda_{N_\lambda}\right)\right] \tag{77}$$

for real and imaginary parts of aerosol refractive index respectively, and

$$\mathbf{a}_{\mathrm{P}}^{(\mathrm{PSD_{fine}})} = \left[\frac{dV(r_1)}{d\ln r_1}, \frac{dV(r_2)}{d\ln r_2}, \ldots, \frac{dV(r_{N_{bin}})}{d\ln r_{N_{bin}}}\right]_{\mathrm{fine}} \tag{78}$$

$$\mathbf{a}_{\mathrm{P}}^{(\mathrm{PSD_{coarse}})} = \left[\frac{dV(r_1)}{d\ln r_1}, \frac{dV(r_2)}{d\ln r_2}, \ldots, \frac{dV(r_{N_{bin}})}{d\ln r_{N_{bin}}}\right]_{\mathrm{coarse}} \tag{79}$$

for fine and coarse mode particle size distribution respectively. Note that, to account for the differences in the spectral bands of the AERONET and POLDER instruments, the refractive indices at the four AERONET wavelengths are interpolated to match the POLDER wavelengths before implementing PCA. The log-normal size distribution is parameterized by volume median radius, standard deviation, and volume concentration for both fine and coarse modes. The bi-modal size distribution is reconstructed by Eq.(2). In our implementation of PCA, the mode parameters (r_{m}, σ, C_{v}) of both fine and coarse mode particles are adopted from AERONET inversion products. Then the fine and coarse mode PSD are constructed and discretized over 22 AERONET size bins (namely N_{bin}=22 in Eqs. (78)-(79)). PCA is applied to determine the dominating few PCs for the fine and coarse particles respectively.

To avoid the appearance of negativity of the constructed aerosol properties from PCs and to leverage the differences of magnitudes in aerosol parameters, PCA of a correlated aerosol property is performed in logarithmic space, namely:

$$\mathbf{x}_{\mathrm{P},i}^{(p)} = \log\left(\mathbf{a}_{\mathrm{P},i}^{(p)}\right) \approx \bar{\mathbf{x}}_{\mathrm{P}}^{(p)} + \sum_{\boldsymbol{k}=\mathbf{1}}^{N_{\mathrm{PC}}^{(p)}} \mathbf{v}_{\mathrm{P},k}^{(p)} \times \mathbf{w}_{\mathrm{P},k,i}^{(p)} \tag{80}$$

where $\mathbf{a}_{\mathrm{P},i}^{(p)}$ represents p^{th} aerosol property saved in i^{th} aerosol record; $\bar{\mathbf{x}}_{\mathrm{P}}^{(p)}$ represents its climatological mean; $\mathbf{w}_{\mathrm{P},k,i}^{(p)}$ is the weight multiplied with k^{th} PC; and **v** is the PC matrix that contain a certain number ($N_{\mathrm{PC}}^{(p)}$) of PC vectors. A specific aerosol record can be recovered to the full accuracy by adopting a maximum number of PCs, namely $N_{\mathrm{PC}}^{(p)}$=$L_{\mathrm{aero}}^{(p)}$, where $L_{\mathrm{aero}}^{(p)}$ is the length of p^{th} aerosol property in regular parameter space. Grouping the dominating PCs into the matrix **v**, we have

$$\mathbf{v}_{\mathrm{P},k}^{(p)} = \left[\mathbf{v}_{\mathrm{P},1}^{(p)}, \mathbf{v}_{\mathrm{P},2}^{(p)}, \dots\ \mathbf{v}_{\mathrm{P},N_{\mathrm{PC}}}^{(p)}\right] \in \mathbb{R}^{L_{\mathrm{aero}}^{(p)} \times N_{\mathrm{PC}}^{(p)}} \tag{81}$$

where $\mathbf{v}_{\mathrm{P},k}^{(p)}$ is k^{th} PC vector containing a total of $L_{\mathrm{aero}}^{(p)}$ elements of p^{th} property and

$$\mathbf{v}_{\mathrm{P},k}^{(p)} \in \mathbb{R}^{L_{\mathrm{aero}}^{(p)}} \tag{82}$$

In Eq. (80), $\mathbf{w}_{\mathrm{P},k,i}^{(p)}$ is a row vector containing dominating PC weights. Taking p^{th} aerosol property saved in i^{th} aerosol record as an example, we have

$$\mathbf{w}_{\mathrm{P},i}^{(p)} \in \mathbb{R}^{N_{\mathrm{PC}}^{(p)}} \tag{83}$$

To quantify the PCA quality, we then demonstrate the selection and reconstruction of PCA from AERONET aerosol climatology in the SWCAUS. This region includes a total of 76 available AERONET sites, providing 200,672 individual records of AERONET products. Figure 33a illustrates the domain of the target area. The stars mark the location of AERONET sites. The dark green, moderate green, and light green indicate the terrain of forest (NDVI>0.6), shrub/grassland (0.2<NDVI≤0.6) and barren/desert (NDVI≤0.2) respectively. We did not include water and ocean surface surfaces in PC analysis due to lack or insufficiency of AERONET data. The four panels in Figure 33b shows the percentage of variance of four aerosol properties, including refractive index (real part denoted by "RIR" and the imaginary part denoted by "RII"), and fine and coarse mode aerosol size distributions, captured by using different number of PCs. In each panel, the three colored curves represent three types of surfaces, and the error bars indicate the standard deviation due to seasonal influences. The dashed black line indicates the fraction metric of 95% the variance. For all types of underlying surfaces and seasons three PCs are found to capture ⩾95% variance, meaning that there are strong correlations in these four type of aerosol properties. From the retrieval perspective, it is only necessary to determine the first few PCs' weights to capture the

major variations in aerosol properties. This results in a substantial reduction in the dimensionality of retrieval parameter space. Taking fine and coarse mode particle size distribution as an example, AERONET uses 22 size bin-based aerosol concentrations to represent a full PSD. By capturing correlation using PCs, only 3 PCs are needed to achieve the same purpose. Within CIMAP inversion framework, the PC weights for the fine and coarse modes are retrieved separately and used to construct the overall PSD using Eq. (2).

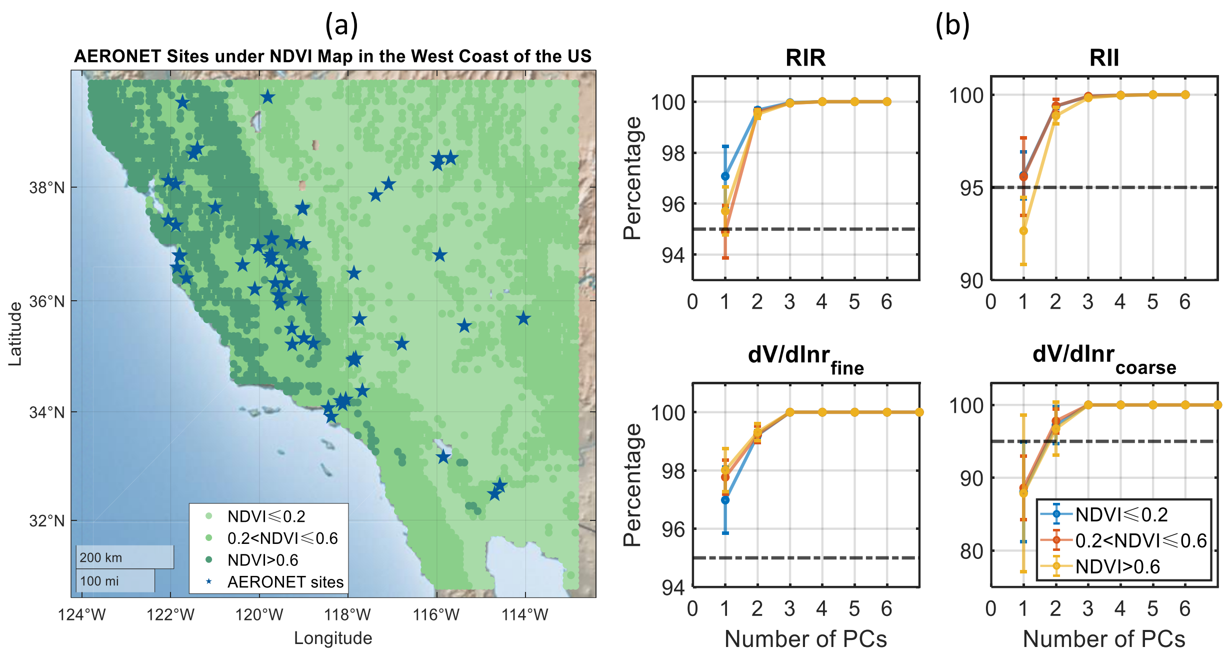


Figure 33. (a) NDVI map and the located AERONET sites in the SWCAUS - a region of 1000 by 1000 km^2 area. Pixels with dark green, moderate green, and light green scatters indicate forest terrain (NDVI>0.6), shrub and grassland (0.2<NDVI≤0.6), and barren surface (NDVI≤0.2) areas, respectively. Blue stars mark AERONET sites in the region. (b) Percentages of variance of real and imaginary part of refractive index, and volumetric particle size distribution in fine and coarse modes captured by using a different number of PCs. Blue lines, orange lines, and yellow lines indicate PCs for forest terrain, shrub and grassland, and barren surfaces, respectively. Error bars are the 1-s deviation due to seasonal dependence of aerosol PCs.

The number of PCs (N_{PC}) used for retrieving a correlated aerosol property is configurable in CIMAP, with upper limits defined by the number of spectral bands N_λ for refractive index and the number of size bins N_{bin} for size distribution. Theoretically, including all PCs to reconstruct aerosol properties will recover exactly same amount of information as that could be provided by regular parameters. However, such an option essentially eliminates the effect of correlation constraints, increases the number of retrieval parameters, and reduce the retrieval efficiency. Due to insufficiency of polarimetric measurements to cover all details of aerosol properties, using a maximum number of PCs will also cause the retrieval to be more sensitive to measurement and modeling errors. On the other hand, if N_{PC} is too small, then the spatial and temporal variations of aerosol properties in the target area cannot be well represented. In such a context, an optimal N_{PC} is needed to achieve a balance between retaining sufficient representability of aerosols and maintaining a compact retrieval parameter space to improve the retrieval accuracy and efficiency.

In our implementation, we find three PCs ($N_{PC} = 3$) is the optimal setting as it captures more than 95% of the spatial and temporal variance in the correlated aerosol properties. To access the quality of reconstructed aerosol properties by using $N_{PC} = 3$, we calculate the aerosol properties in regular parameter space by Eq. (80) and compare the results to the original aerosol records reported by AERONET. Figure 34 a-d demonstrates the comparison of the RIR, RII, fine mode and coarse mode particle concentrations in four representative size bins, respectively for winter and moderate NDVI values between 0.2 and 0.6. The complete dataset demonstrated in Figure 34 includes 13,140 aerosol records. The regression demonstrates that the reconstructed aerosol properties derived from the PCs align well with the original data, with only a limited number of outliers, primarily resulting from the reduced representativeness associated with using a subset of PCs to reconstruct the original aerosol properties. Outliers also show up in the RII regression plot. Indeed, out of the 13,140 records only 185 records exhibit RII values exceeding 0.1, and 37 records exceed 0.2, representing a marginal fraction of the dataset.

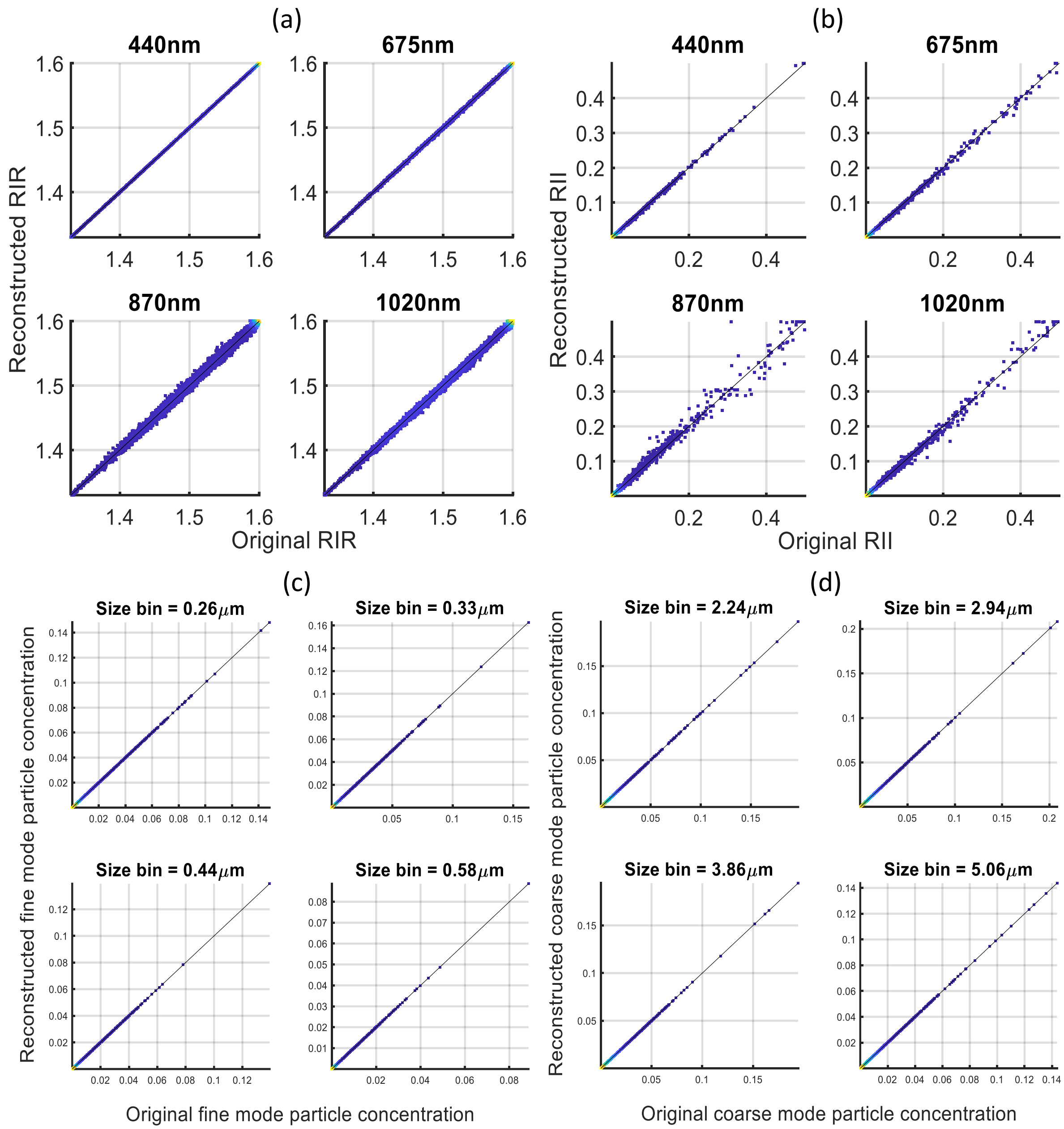


Figure 34. a) Reconstructed real part of spectral refractive index using three PCs and comparison to the original AERONET data in winter and surface of moderate NDVI values. b) Same as Fig.Figure 34a but for imaginary of refractive index. c) Same as Fig.Figure 34a but for fine mode size bins centered at 0.26, 0.33, 0.44, and 0.58μm. d). Same as Fig.Figure 34a but for coarse mode size bins centered at 2.24, 2.94, 3.86, and 5.06μm.

4.3.1.2 Surface A Priori

CIMAP allows a priori surface constraints to be used, enabling the retrieval to be initialized with physically meaningful surface reflectance estimates rather than arbitrary starting values. In this study, we use the surface reflection dataset of the MAIAC [Lyapustin et al., 2018] to generate an initial guess of surface BRF and DHR to inform the retrieval. The MAIAC MCD19A3 dataset provides spectral surface BRF parameters at 1 km spatial resolution, derived from measurements by the MODIS [King et al., 1992]. Specifically, the dataset includes kernel coefficients for the sRTLS [Lyapustin et al., 2025] model, which decomposes surface reflectance into isotropic, geometric, and volumetric scattering components [Wanner et al., 1995].

According to the preliminary conclusion from Section 3, the MAIAC dataset is suitable for providing a priori surface information to inform the aerosol retrieval. For every observed pixel, the MAIAC dataset is queried to retrieve kernel coefficients at the geographically corresponding location. These coefficients are then used to initialize the surface reflectance calculation within CIMAP, as described in the algorithm structure (Figure 32).

However, a direct application of these coefficients is not straightforward, because CIMAP employs a different surface reflectance model than MAIAC. While MAIAC adopts the sRTLS model, CIMAP uses the mRPV model to parameterize spectral surface reflectance. To bridge this difference and provide CIMAP with a reliable initial guess of mRPV model parameters, the conversion algorithm introduced in Section 4 is applied to map MAIAC sRTLS parameters into the mRPV parameter space.

As part of this conversion, we adopt the spectral invariance assumption [Diner et al., 2012], which posits that the angular shape of surface BRFs is spectrally invariant. Under this assumption, only the spectrally-dependent scaling parameter $r_{0,\lambda}$ varies with wavelength, while the angular shape parameters k and b remain wavelength-independent. Because the kernel coefficients from MAIAC are defined at MODIS wavelengths, the converted mRPV parameters are initially obtained at those same wavelengths. An interpolation is subsequently performed to map these values onto the spectral bands of the target instrument.

A practical consideration also arises from the spatial resolution mismatch between the MAIAC MCD19A3 surface product (~1 km) and POLDER pixels (~6 km). To address this, the nearest MAIAC pixel to each POLDER pixel is selected to provide the initial surface parameter estimates. Importantly, since the surface BRFs initialized from MAIAC are jointly optimized with aerosol parameters during the retrieval, this co-optimization provides a degree of tolerance to mismatches arising from spatial resolution differences and surface modeling errors.

4.3.1.3 State Vector

In CIMAP-P algorithm, the state vector $\mathbf{x}_\mathrm{P}$ includes the PC weights ($\mathbf{w}_\mathrm{P}$) for correlated aerosol parameters, including the PC weights for each aerosol properties, the uncorrelated aerosol parameters including the volumetric spherical particle fraction ($f_{\mathrm{v,sphere}}$) and the mean height of aerosol layer ($\overline{H}_{\mathrm{aer}}$) that conforms to Gaussian profile [Xu et al., 2017] as well as the regular surface reflection parameters including ($\mathbf{r}_{0,\lambda}, k, b$) for BRF and ($\varepsilon, v_\mathrm{s}$) for pBRF, namely:

$$\mathbf{x}_\mathrm{P} = \left[\log\left(\mathbf{w}_\mathrm{P}, f_{\mathrm{v,sphere}}, \overline{H}_{\mathrm{aer}}, \mathbf{r}_{0,\lambda}, k, b, \varepsilon, v_\mathrm{s}\right)\right]^\mathrm{T} \tag{84}$$

where $\mathbf{r}_{0,\lambda}$ in bold is a vector that contains a set of spectral values for all wavelengths, and $\mathbf{w}_\mathrm{P}$ includes the weights of the PCs that dominate the correlations in aerosol refractive index and in particle size distribution, namely:

$$\mathbf{w}_\mathrm{P} = \left[\mathbf{w}_\mathrm{P}^{(\mathrm{RIR})}, \quad \mathbf{w}_\mathrm{P}^{(\mathrm{RIR})}, \quad \mathbf{w}_\mathrm{P}^{(\mathrm{PSD_{fine}})}, \quad \mathbf{w}_\mathrm{P}^{(\mathrm{PSD_{coarse}})}\right]^\mathrm{T} \tag{85}$$

Table 5 lists the initial guesses, ranges, and Lagrange multipliers of all retrieval parameters.

Table 5. Range, initial guess and Lagrange multipliers for the CIMAP-P retrieval parameters

Retrieved parameters	Range	Vector Size	Initial Guess	Lagrange multiplier
$w_{\mathrm{P},1-3}^{\mathrm{refr,r}}$	Informed from PCA of AERONET aerosol climatology	3	0	1
$w_{\mathrm{P},1-3}^{\mathrm{refr,i}}$		3	0	1
$w_{\mathrm{P},1-3}^{\mathrm{fine}}$		3	0	0
$w_{\mathrm{P},1-3}^{\mathrm{coarse}}$		3	0	0
$f_{\mathrm{v,sphere}}$	[0.05, 0.99]	1	0.9	-
$\bar{H}_{\mathrm{aer}}$	[0.05, 10]	1	1	-
$r_{0,\lambda}$	$[0.3, 3] \times r_{0,\lambda}$	6	Converted from MAIAC	0.01
k	$[0.5, 2.5] \times k$, if $k > 0$; [-0.2, -1.7] $\times k$, if $k < 0$	1	Converted from MAIAC	-
b	$[0.5, 2.5] \times b$, if $b > 0$; [-0.2, -1.7] $\times b$, if $b < 0$	1	Converted from MAIAC	-
ε	[0.001, 1]	1	0.01	-
v_{s}	[5, 50]	1	15	-

4.3.1.4 POLDER Observations

The CIMAP-P uses POLDER's screened data at all view angles, including the radiances in six aerosol and surface bands centered at 440, 490, 565, 675, 870, and 1020nm and the degree of linear polarization (DoLP) data at 2$^{\mathrm{nd}}$, 4$^{\mathrm{th}}$ and 5$^{\mathrm{th}}$ bands. The DoLP is calculated from the measured Stokes parameters I, Q, and U by Eq. (4).

The observational vector $\mathbf{f}_{\mathrm{P}}^{\mathrm{obs},*}$ include spectral measurement in all bands and view angles $\mathbf{\Omega}(\theta, \phi)$, namely

$$\mathbf{f}_{\mathrm{P}}^{\mathrm{obs},*} = \log[\mathbf{I}(\lambda_1), \mathbf{I}(\lambda_2), \ldots\ \mathbf{I}(\lambda_6), \mathbf{DoLP}(\lambda_2), \mathbf{DoLP}(\lambda_4), \mathbf{DoLP}(\lambda_5)\ ]^{\mathrm{T}}, \qquad (86)$$

where

$$\mathbf{I}(\lambda_i) = \left[\mathrm{I}(\lambda_i, \mathbf{\Omega}_1), \mathrm{I}(\lambda_i, \mathbf{\Omega}_2), \dots, \mathrm{I}\left(\lambda_i, \mathbf{\Omega}_{N_\Omega}\right)\right] \tag{87}$$

and

$$\mathbf{DoLP}(\lambda_i) = \left[\mathrm{DoLP}(\lambda_i, \mathbf{\Omega}_1), \mathrm{DoLP}(\lambda_i, \mathbf{\Omega}_2), \dots, \mathrm{DoLP}\left(\lambda_i, \mathbf{\Omega}_{N_\Omega}\right)\right] \tag{88}$$

Note that $\mathbf{f}_{\mathrm{P}}^{\mathrm{obs},*}$ is transformed into logarithmic space to mitigate the difference in magnitude of the measured signals. The weighting matrix of the observations is calculated by referring to the relative measurement uncertainty 0.03 for radiance and absolute measurement uncertainty of 0.01 for DoLP. In the process of optimization, retrieval terminates when the residuals between the simulated and observed measurements, as captured by the cost function, drop below measurement uncertainties or when convergence is achieved; namely, when the relative change in the fitting residuals between two successive iterations falls below 2%.

4.3.1.5 Inversion Formulation

The inversion module in CIMAP-P deploys a multi-constraint least-square method to fit observations. It adjusts the aerosol and surface parameters iteratively until the fitting residue converges or drops below a threshold value. The inversion is set to solve the following equations [Dubovik et al., 2011]:

$$\begin{cases} \boldsymbol{f}_{\mathrm{P}}^{\mathrm{obs},*} = \boldsymbol{f}(\mathbf{x}_{\mathrm{P}}) + \Delta\boldsymbol{f}_{\mathrm{P}}^{\mathrm{obs}} \\ \boldsymbol{f}_{\mathrm{P}}^{\mathrm{reg},*} = \mathbf{S}_{\mathrm{P}}^{\mathrm{reg}}\mathbf{x} + \Delta\boldsymbol{f}_{\mathrm{P}}^{\mathrm{reg}} \\ \boldsymbol{f}_{\mathrm{P}}^{\mathrm{corr},*} = \mathbf{S}_{\mathrm{P}}^{\mathrm{PC}}\mathbf{x} + \overline{\mathbf{x}_{\mathrm{P}}} + \Delta\boldsymbol{f}_{\mathrm{P}}^{\mathrm{corr}} \\ \mathbf{x}_{\mathrm{P,a}} = \mathbf{x}_{\mathrm{P}} + \Delta\mathbf{x}_{\mathrm{P,a}} \end{cases} \tag{89}$$

where $\boldsymbol{f}_{\mathrm{P}}^{\mathrm{obs},*}$ is the vector contains the measurement data, $\mathbf{x}_{\mathrm{P}}$ is the state vector, and $\boldsymbol{f}(\mathbf{x}_{\mathrm{P}})$ is the modeled observations, and $\Delta\boldsymbol{f}_{\mathrm{P}}^{\mathrm{obs}}$ is a vector containing the measurement uncertainties. The second and third lines of Eq. (89) correspond to the imposed a priori constraints that ensure smooth variations of regular and correlated parameters as a function of wavelength or particle size, respectively. The smooth variation is ensured by forcing the m^{th}-order of derivative of relevant retrieval parameters to approach zero, namely $\boldsymbol{f}_{\mathrm{P}}^{\mathrm{reg},*} = \mathbf{0}^*$ and $\boldsymbol{f}_{\mathrm{P}}^{\mathrm{corr},*} = \mathbf{0}^*$ within the uncertainties of $\Delta\boldsymbol{f}_{\mathrm{P}}^{\mathrm{reg}}$and $\Delta\boldsymbol{f}_{\mathrm{P}}^{\mathrm{corr}}$, respectively. In the fourth line of Eq. (89), $\mathbf{x}_{\mathrm{P,a}}$ is the a priori estimate of the state vector and $\Delta\mathbf{x}_{\mathrm{P,a}}$ is the uncertainties of a priori constraints.

The cost function integrates multiple complementary constraints within a unified least-squares framework. The observational term ensures consistency with measured radiometric and polarimetric signals, while the smoothness constraints regularize the variations of a set of retrieval parameters (e.g. across wavelengths or size bins) to mitigate non-physical oscillations. The correlation-based constraint further enforces physically consistent relationships in aerosol parameters, and the a priori term provides additional stabilization based on climatological knowledge. Together, these constraints balance measurement fidelity and physical realism, leading to a stable and well-posed inversion. Under such a context, Eq. (89) is solved iteratively by minimizing the following quadratic form of cost function [Dubovik et al., 2011]:

$$\Psi\left(\mathbf{x}_{\mathrm{P},i}\right) = \sum_{t=1}^{M} \Psi_t(\mathbf{x}_{\mathrm{P},i}) \tag{90}$$

where $\Psi_t(\mathbf{x}_{\mathrm{P},i})$ represents the t^{th} type of constraint that contributes to the cost function, associated with i^{th} iterative solution $\mathbf{x}_{\mathrm{P},i}$. We account for a total of 4 types of constraints. Referring to Eq. (89), t=1 is associated with observational constraint calculated by comparing the model ($\boldsymbol{f}(\mathbf{x}_{\mathrm{P}})$) and remotely measured signals ($\boldsymbol{f}_{\mathrm{P}}^{\mathrm{obs},*}$); t=2 is associated with the smoothness constraint on regular parameters (imposed via matrix $\mathbf{\Omega}^{\mathrm{reg}}$); t=3 is associated with the smoothness constraint on correlated parameters (imposed via matrix $\mathbf{\Omega}^{\mathrm{corr}}$), and t=4 is associated with the a priori constraint. These components are evaluated by

$$
\Psi_t(\mathbf{x}_{\mathrm{P},i}) = \begin{cases} \gamma_t\left(\Delta \boldsymbol{f}_{\mathrm{P},i}\right)^{\mathrm{T}} \mathbf{W}_f^{-1} \Delta \boldsymbol{f}_{\mathrm{P},i}, & t = 1 \\ \gamma_t\left(\mathbf{S}^{\mathrm{reg}} \mathbf{x}_{\mathrm{P},i}\right)^{\mathrm{T}} \left(\mathbf{S}^{\mathrm{reg}} \mathbf{x}_{\mathrm{P},i}\right), & t = 2 \\ \gamma_t\left(\mathbf{S}^{\mathrm{corr}} \mathbf{x}_{\mathrm{P},i} + \overline{\mathbf{x}}\right)^{\mathrm{T}} \left(\mathbf{S}^{\mathrm{corr}} \mathbf{x}_{\mathrm{P},i} + \overline{\mathbf{x}_{\mathrm{P}}}\right), & t = 3 \\ \gamma_t\left(\mathbf{x}_{\mathrm{P},i} - \mathbf{x}_{\mathrm{P,a}}\right)^{\mathrm{T}} \mathbf{W}_a^{-1} \left(\mathbf{x}_{\mathrm{P},i} - \mathbf{x}_{\mathrm{P,a}}\right), & t = 4 \end{cases} \tag{91}
$$

where $\Delta \boldsymbol{f}_{\mathrm{P},i}$ is the difference between modeled and observed measurements ($\Delta \boldsymbol{f}_{\mathrm{P},i} = \boldsymbol{f}\left(\mathbf{x}_{\mathrm{P},i}\right) - \boldsymbol{f}_{\mathrm{P}}^{\mathrm{obs},*}$), $\mathbf{W}_f$ is the corresponding weighting matrix calculated from the measurement errors, $\mathbf{W}_a$ is weighting matrix calculated from the estimated of range of a retrieval parameter [Dubovik et al., 2011], and γ_t is the Lagrange multiplier that controls the strength of a certain type of constraint (see Table 5 for their initial value). γ_t is dynamically updated during iteration referring to the residue of fitting, namely [Dubovik et al., 2011]

$$
\gamma_{i+1} = \gamma_i \, \frac{\Psi\left(\mathbf{x}_{\mathrm{P},i}\right)}{\Psi\left(\mathbf{x}_{\mathrm{P},i+1}\right)}. \tag{92}
$$

The solution is approached iteratively via

$$
\mathbf{x}_{\mathrm{P},i+1} = \mathbf{x}_{\mathrm{P},i} - \Delta \mathbf{x}_{\mathrm{P},i+1} \tag{93}
$$

where the increment of solution $\Delta \mathbf{x}_{\mathrm{P},i}$ is obtained by solving the following system of equations:

$$
\mathbf{A}_{\mathrm{P},i} \, \Delta \mathbf{x}_{\mathrm{P},i} = \nabla \Psi\left(\mathbf{x}_{\mathrm{P},i}\right) \tag{94}
$$

$$
\mathbf{A}_{\mathrm{P},i} = \mathbf{J}_{\mathrm{P},i}^{\mathrm{T}} \, \mathbf{W}_{\mathrm{P}} \, \mathbf{J}_{\mathrm{P},i} + \gamma_{\mathrm{s}} \boldsymbol{\Omega}_i^{\mathrm{reg}} + \gamma_{\mathrm{s}} \boldsymbol{\Omega}_i^{\mathrm{corr}} + \gamma_{\mathrm{a}} \mathbf{W}_{\mathrm{P,a}}^{-1} \tag{95}
$$

$$
\begin{aligned} \nabla \Psi\left(\mathbf{x}_{\mathrm{P},i}\right) = {} & \mathbf{J}_{\mathrm{P},i}^{\mathrm{T}} \, \mathbf{W}_{\mathrm{P}} \, \Delta \boldsymbol{f}_{\mathrm{P},i} + \gamma_{\mathrm{s}} \, \boldsymbol{\Omega}^{\mathrm{reg}} \mathbf{x}_{\mathrm{P},i} + \gamma_{\mathrm{s}} \left[\boldsymbol{\Omega}^{\mathrm{corr}} \mathbf{x}_{\mathrm{P},i} + \left(\mathbf{S}^{\mathrm{corr}} \right)^{\mathrm{T}} \mathbf{S}^{\mathrm{reg}} \boldsymbol{f}_{\mathrm{P}}(\mathbf{x}_{\mathrm{P}}) \right] \\ & + \gamma_{\mathrm{a}} \mathbf{W}_{\mathrm{P,a}}^{-1} \left(\mathbf{x}_{\mathrm{P},i} - \mathbf{x}_{\mathrm{P,a}}\right) \end{aligned} \tag{96}
$$

where $\mathbf{J}_{\mathrm{P}}$ is the Jacobian matrix, $\mathbf{W}_{\mathrm{P}}$ is the weighting matrix of observations, and $\boldsymbol{\Omega}$ is the smoothness matrix, and $\boldsymbol{f}_{\mathrm{P}}(\mathbf{x}_{\mathrm{P}}) = \boldsymbol{f}_{\mathrm{P}}\left(\mathbf{x}_{\mathrm{P},1}\right), \boldsymbol{f}_{\mathrm{P}}\left(\mathbf{x}_{\mathrm{P},2}\right), \dots, \boldsymbol{f}_{\mathrm{P}}\left(\mathbf{x}_{\mathrm{P},N_{\mathrm{t}}}\right)$, with $f_{\mathrm{P}}(\mathbf{x}_i) = \mathbf{x}_{\mathrm{P},i}$ if the i^{th} aerosol property is a regular parameter and $f_{\mathrm{P}}(\mathbf{x}_{\mathrm{P},i}) = \overline{\mathbf{x}}_{\mathrm{P},i}$ if the i^{th} aerosol property is correlated parameter. To impose the smoothness constraint on regular type aerosol properties, $\mathbf{\Omega}^{\mathrm{reg}}$ is calculated by accounting for the smoothness constraints imposed on all N_{t} types of retrieval parameters, namely

$$\mathbf{\Omega}^{\mathrm{reg}} = \begin{bmatrix} \mathbf{\Omega}_1^{\mathrm{reg}} & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & \mathbf{\Omega}_2^{\mathrm{reg}} & \dots & \mathbf{0} \\ \cdots & \cdots & \ddots & \dots \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{\Omega}_{N_{\mathrm{t}}}^{\mathrm{reg}} \end{bmatrix} \tag{97}$$

where for a regular individual type of retrieval parameters, $\mathbf{\Omega}_j^{\mathrm{reg}}$ is calculated from differentiation matrix:

$$\mathbf{\Omega}_j^{\mathrm{reg}} = \left(\mathbf{S}_j^{\mathrm{reg}}\right)^{\mathrm{T}} \mathbf{S}_j^{\mathrm{reg}}. \tag{98}$$

The differentiation matrix $\mathbf{S}^{\mathrm{reg}}$ is constructed by discretizing the m^{th} order of derivative of function $\mathbf{f}(\lambda)$ or $\mathbf{f}(\mathrm{N_{bin}})$ over the wavelength or size bin grids. Taking m=1 and m=2 as example, we have the following for j^{th} type of retrieval parameter of length $L(j)$ [Dubovik et al., 2011; Xu et al., 2019]:

$$\mathbf{S}_j^{\mathrm{reg}}(m=1) = \begin{bmatrix} \frac{-1}{\Delta_1} & \frac{1}{\Delta_1} & 0 & \cdots & 0 \\ \vdots & \frac{-1}{\Delta_2} & \frac{1}{\Delta_2} & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & 0 \\ 0 & 0 & \cdots & \frac{-1}{\Delta_{L(j)-1}} & \frac{1}{\Delta_{L(j)-1}} \end{bmatrix}, \tag{99}$$

and

$$\mathbf{S}_j^{\mathrm{reg}}(m=2) \\ = \begin{bmatrix} \frac{2}{\Delta_1(\Delta_1+\Delta_2)} & \frac{-2}{(\Delta_1\Delta_2)} & \frac{2}{\Delta_2(\Delta_1+\Delta_2)} & 0 & \cdots & 0 \\ \vdots & \frac{2}{\Delta_2(\Delta_2+\Delta_3)} & \frac{-2}{(\Delta_2\Delta_3)} & \frac{2}{\Delta_3(\Delta_2+\Delta_3)} & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & 0 \\ 0 & 0 & \cdots & \frac{2}{\Delta_{L(j)-1}\left(\Delta_{L(j)-1}+\Delta_{L(j)}\right)} & \frac{-2}{\left(\Delta_{L(j)-1}\Delta_{L(j)}\right)} & \frac{2}{\Delta_{L(j)}\left(\Delta_{L(j)-1}+\Delta_{L(j)}\right)} \end{bmatrix}. \tag{100}$$

where $\Delta_l = z_{l+1} - z_l$, and $L(j)$ is the length of grids. If no smoothness constraint is imposed on a certain type of retrieval property or the index refers to a correlated type of retrieval parameter, then we set $\boldsymbol{\Omega}_j^{\mathrm{reg}} = 0$.

To impose the smoothness constraints on correlated aerosol parameters, we use

$$\boldsymbol{\Omega}^{\mathrm{corr}} = \begin{bmatrix} \boldsymbol{\Omega}_1^{\mathrm{corr}} & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & \boldsymbol{\Omega}_2^{\mathrm{corr}} & \dots & \mathbf{0} \\ \cdots & \cdots & \ddots & \dots \\ \mathbf{0} & \mathbf{0} & \cdots & \boldsymbol{\Omega}_{N_\mathrm{t}}^{\mathrm{corr}} \end{bmatrix}, \tag{101}$$

where $\boldsymbol{\Omega}_j^{\mathrm{corr}}$ is calculated in the following way for correlated aerosol properties, by

$$\boldsymbol{\Omega}_j^{\mathrm{corr}} = \left(\mathbf{S}_j^{\mathrm{corr}} \right)^{\mathrm{T}} \mathbf{S}_j^{\mathrm{corr}}, \tag{102}$$

where $\mathbf{S}_j^{\mathrm{corr}}$ is evaluated by

$$\mathbf{S}_j^{\mathrm{corr}} = \mathbf{S}_j^{\mathrm{reg}} \mathbf{v}_\mathrm{P}^{\mathrm{T}}, \tag{103}$$

where $\mathbf{v}_\mathrm{P}^{\mathrm{T}}$ is the PC matrix for j^{th} type of aerosol correlated parameters. If j^{th} type of retrieval parameter is uncorrelated, we set $\boldsymbol{\Omega}_j^{\mathrm{corr}} = 0$.

4.3.2 Retrieval Implementation and Demonstration for CIMAP-P

4.3.2.1 Data

A dataset of POLDER observations during 2008 to 2011 over all available AERONET sites in the SWCAUS are generated for testing CIMAP-P. A total of 345 collocated data is found, including 6 for spring and low NDVI, 36 for spring and moderate NDVI, 9 for summer and low NDVI, 54 for fall and low NDVI, 128 for fall and moderate NDVI, 30 for winter and low NDVI, and 82 for winter and moderate NDVI. The retrieval covers low to moderately high AOD (~0.5) scenarios. AERONET inversion products were collocated with POLDER measurements within a

20-minute temporal window of the satellite overpass. The choice of this matching window represents a strategic balance between maximizing the number of valid collocations and maintaining sufficient temporal consistency between satellite and ground-based observations.

CIMAP-P retrievals are also compared to the products of the GRASP [Dubovik et al., 2011]. GRASP is an optimization-based algorithm and has been applied to retrieve POLDER/ PARASOL observations using a variety of configurations, including GRASP-Component (or “GRASP-C”), GRASP High-Precision (or “GRASP-HP”), and GRASP-Models (or “GRASP-M”). Comparison of CIMAP to GRASP products from these configurations and to AERONET data allows us to assess relative performance across different algorithmic approaches using a common reference. Compared to GRASP-HP, GRASP-C and GRASP-M demonstrate better agreement of AOD with AERONET reference [Li et al. 2019, 2022; Chen et al. 2020]. Conceptually, CIMAP, GRASP-C, and GRASP-M reduce the dimensionality of the aerosol retrieval problem by constraining aerosol properties within a correlated subspace derived from climatological information. However, there are fundamental differences in aerosol representation schemes. CIMAP represents aerosol microphysical properties in a continuous space using PCA. RIR, RII, and PSD are reconstructed from PC weights. In contrast, the GRASP-C and GRASP-M products approximate aerosols as mixtures of several prescribed components. In GRASP-M, aerosol components have fixed size distributions and optical properties, and only their concentrations are retrieved [Chen et al., 2020], whereas in GRASP-C the size distributions are not fixed, and the mixing fractions of the components are retrieved instead [Li et al., 2019]. AOD from GRASP-HP is not recommended for comparison due to biases in low-AOD scenarios. GRASP-M, on the other hand, are not suggested for SSA and PSD comparisons, given the limitations of assuming a mixture of aerosol models with fixed optical properties [Chen et al., 2020]. Nevertheless, all GRASP products are included in this study for completeness, in order to provide a comprehensive intercomparison across different retrieval frameworks. This allows us to assess the consistency and variability among algorithms with differing assumptions and parameterizations, and to place the CIMAP results within a broader methodological context. Note that while GRASP-M and GRASP-C are typically recommended for AOD comparison [Chen et al., 2020], we have also included GRASP-HP for comparison. This is because, in CIMAP, the retrieved aerosol microphysical properties including refractive index and particle size distributions constructed from PC vary in a continuous space. This is more closely aligned with the assumptions in the GRASP-HP which also retrieves aerosol properties in a continuous parameter space.

4.3.2.2 Single-Case Analysis: Comparisons of Spatial Maps

To evaluate CIMAP-P's performance, we first applied it to retrieve aerosols from a POLDER image over the SWCAUS. As indicated in Fig. 35, POLDER's observation of this area on November 15, 2010 was selected as an example because the measurements were taken under nearly cloud-free conditions and one of the cases in the full dataset is collocated with this observation swath. This area covers a variety of terrains, including urban areas, deserts, forests, barrens, and mountains.

Figure 36 displays the retrieved AOD and SSA maps for the study area. The spatial distributions of these quantities exhibit notable heterogeneity across the region: relatively higher AOD values are observed over the inland Southern California valley. The surrounding mountains act as physical barriers so that aerosols and their precursors are trapped within the valley, limiting their dispersion [Smith et al., 1981a,b]. Lower AOD values are observed toward the coastal and ocean-facing portions of the scene. SSA values are higher (close to 1) over the valley regions, indicating a prevalence of scattering aerosols, and lower (around 0.85) in other areas, reflecting a greater presence of absorbing aerosols. Note that these spatial patterns cannot be well captured by AERONET measurements alone. However, The AERONET climatology, constructed from multi-site and multi-temporal records, provides regional statistical constraints by capturing correlations in aerosol properties.

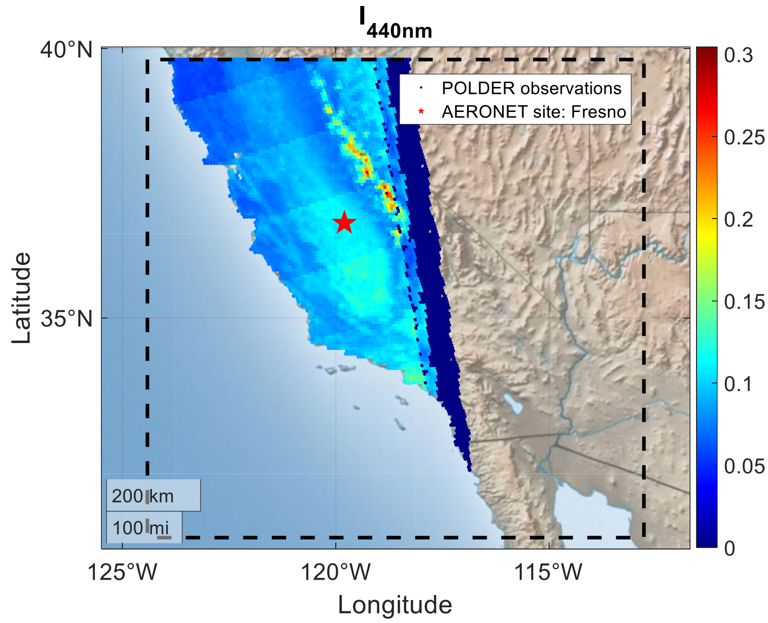


Figure 35. POLDER radiance at the 440nm-band in the 7th angular view of the SWCAUS on November 15, 2010. The red star marks the location of the AERONET Fresno site.

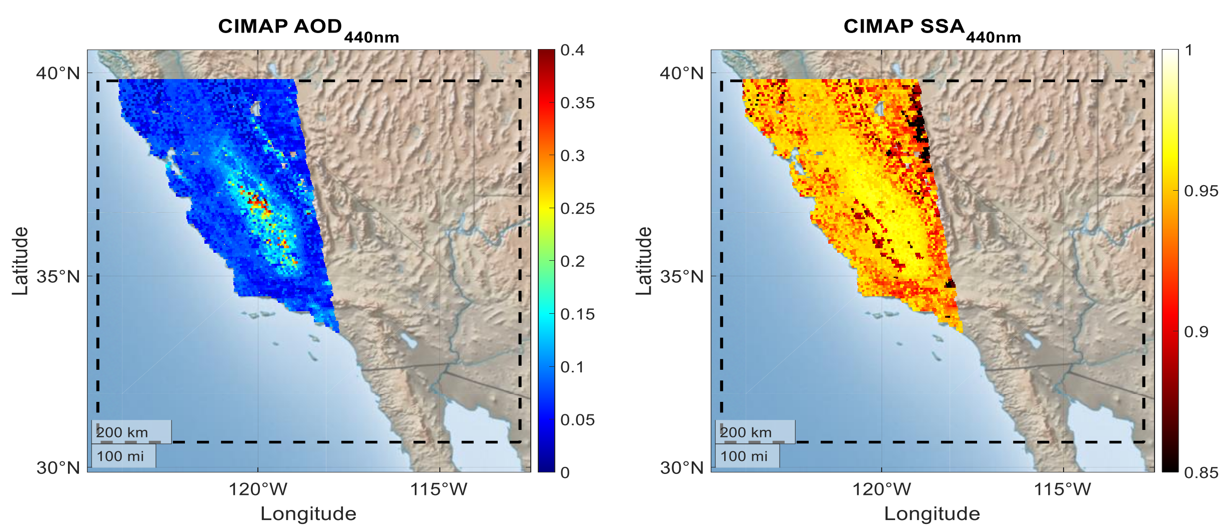


Figure 36. Maps of CIMAP retrieved AOD (left) and SSA (right) at the reference wavelength 440nm in the Los Angeles centered area in the SWCAUS. The retrieval uses the POLDER observation made on Nov. 15. 2010.

To further evaluate the pixel-to-pixel spatial consistency between GRASP products and CIMAP-P, we collocated GRASP data from November 15, 2010, with CIMAP-P retrievals at 440 nm, represented by all data points shown in Fig. 36. Three GRASP configurations, GRASP-C, GRASP-HP, and GRASP-M, are all included in this comparison. The collocation yielded 1,477 common points for the GRASP-C product, 1,717 for GRASP-HP, and 1,028 for GRASP-Model, reflecting the different spatial coverage resulting from the filtering criteria applied to each configuration.

Figure 37 shows the comparison between CIMAP-P and the three GRASP configurations. Overall, CIMAP-P tends to overestimate AOD relative to both GRASP-C and GRASP-M at lower AOD values, while it tends to underestimate AOD under moderate to high AOD conditions. In comparison, CIMAP-P achieves reasonable agreement with GRASP-HP under low AOD conditions (0–0.15), but similarly tends to underestimate at higher AOD values. It should be noted that due to the different filtering criteria applied to each GRASP configuration, the spatial distributions of the collocated GRASP AODs differ across the three comparisons, which should be considered when interpreting the results. Although no ground-based reference is available for these comparisons, the pixel-to-pixel evaluation is nonetheless valuable as it reveals possible biases and strengths of CIMAP-P relative to an independent retrieval algorithm, providing useful directions for future improvement.

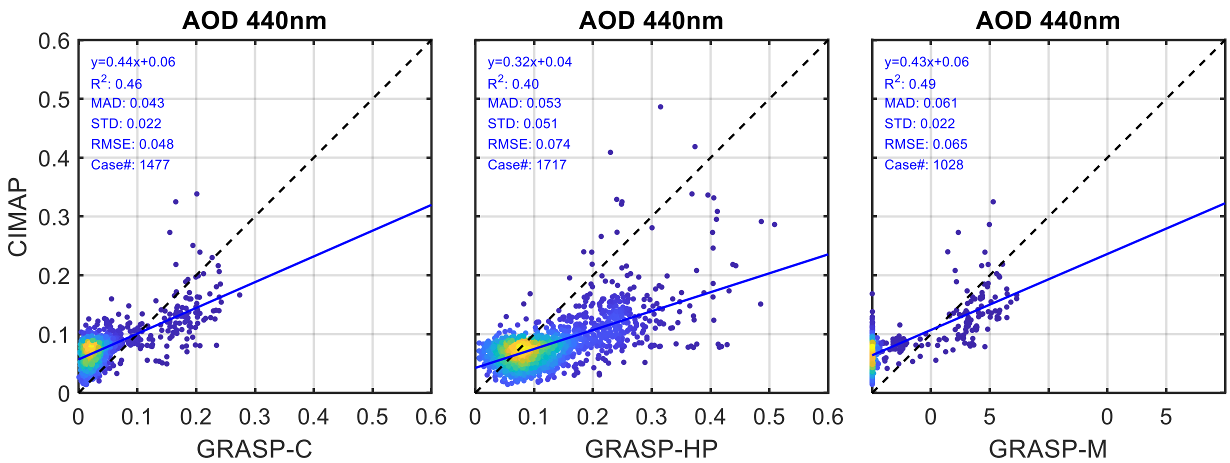


Figure 37. Comparisons between CIMAP with GRASP-Component, GRASP- High-Precision, and GRASP-Model of AOD at 440nm from the POLDER observation on November 15, 2010. 1,477 collocations are found for the GRASP-Component product, 1,717 for GRASP-HP, and 1,028 for GRASP-Model. The color represents point density, where warmer color (yellow) indicates higher concentrations of data and cooler color (blue) represent lower density. Statistics are presented on the left corner.

In the comparisons presented in Fig. 38, 39, and 40, the AOD and SSA spatial distributions of GRASP-C, GRASP-HP, and GRASP-M are also provided as a reference for evaluating the spatial consistency of GRASP products. It should be noted that the number of collocated SSA data points from the GRASP products is insufficient in this scenario for a statistically meaningful comparison, and therefore the GRASP SSA products are not compared against CIMAP-P SSA retrievals in this analysis.

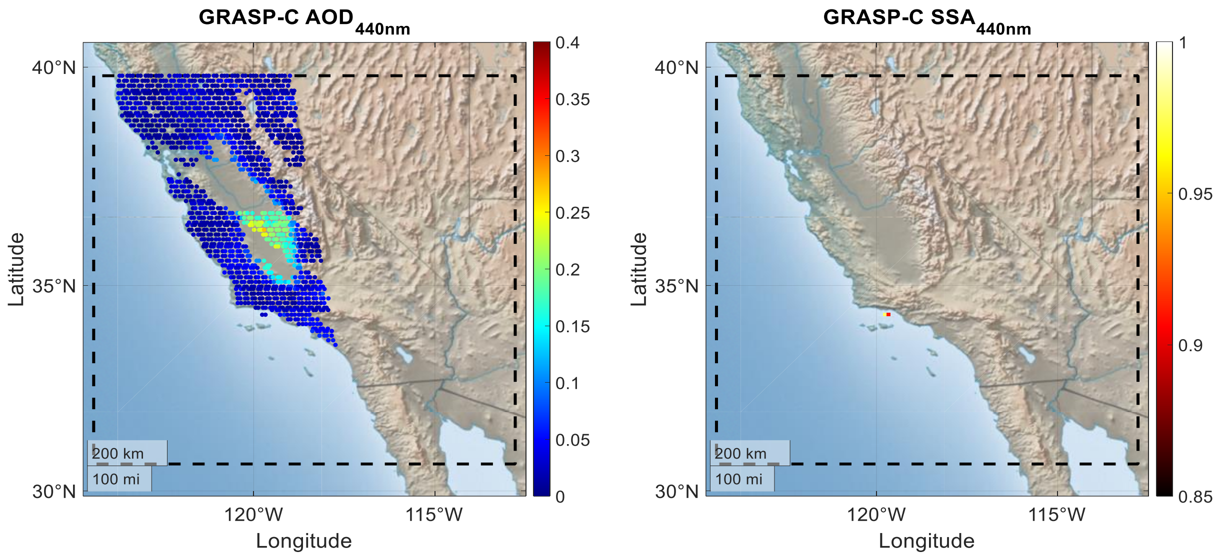


Figure 38. Spatial maps of GRASP-C retrieved AOD (left) and SSA (right) at the reference wavelength 440nm in the Los Angeles centered area in the SWCAUS.

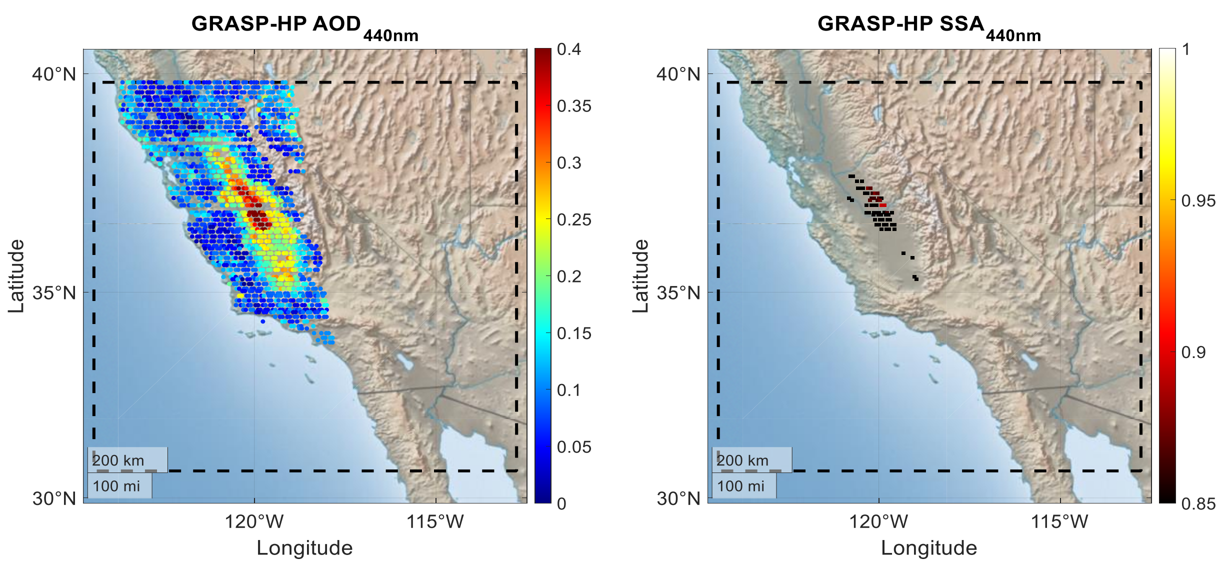


Figure 39. Spatial maps of GRASP-HP retrieved AOD (left) and SSA (right) at the reference wavelength 440nm in the Los Angeles centered area in the SWCAUS.

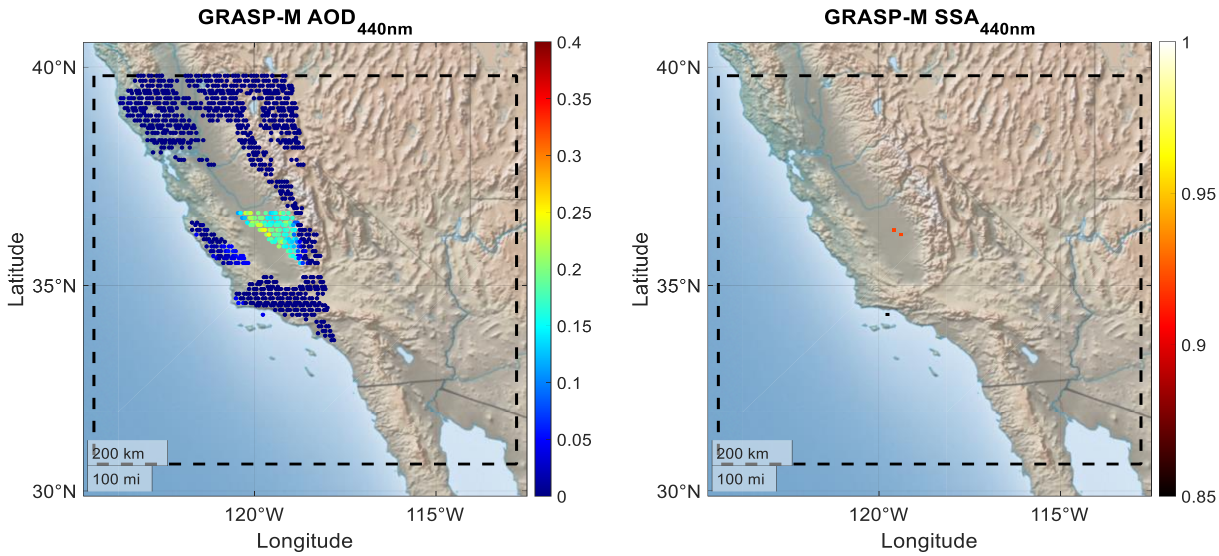


Figure 40. Spatial maps of GRASP-M retrieved AOD (left) and SSA (right) at the reference wavelength 440nm in the Los Angeles centered area in the SWCAUS.

4.3.2.3 Single-Case Analysis: Comparisons of Aerosol Properties at a Collocated Pixel

Marked by the red star in Fig. 35, the AERONET Fresno site by POLDER provides reference data of AOD, SSA, and volumetric PSD for comparing to CIMAP-P retrievals. The Fresno site is of interest due to its aerosol characteristics. Influenced by agricultural and rural emissions, the Fresno site frequently exhibits high aerosol loading, making it an ideal scenario for retrieval validation. The two top panels of Fig. 41 demonstrate the comparison of AOD and SSA. It can be observed that CIMAP retrieval agrees well with AERONET data, with the absolute difference of 0.04-0.08 for AOD and 0.01-0.02 for SSA, depending on wavelengths. The middle panel of Fig. 41 demonstrates the comparisons of volumetric particle size distribution. The peak of fine and coarse mode particle concentration differs by approximately 0.15 and 1 µm respectively. Associated with the retrieval, the two lowest panels of Fig. 41 show fittings of the modeled radiance and DoLP using the retrieved solution to the measurements. The mean differences of fitting are found 1.9% for radiance and 0.008 for DoLP.

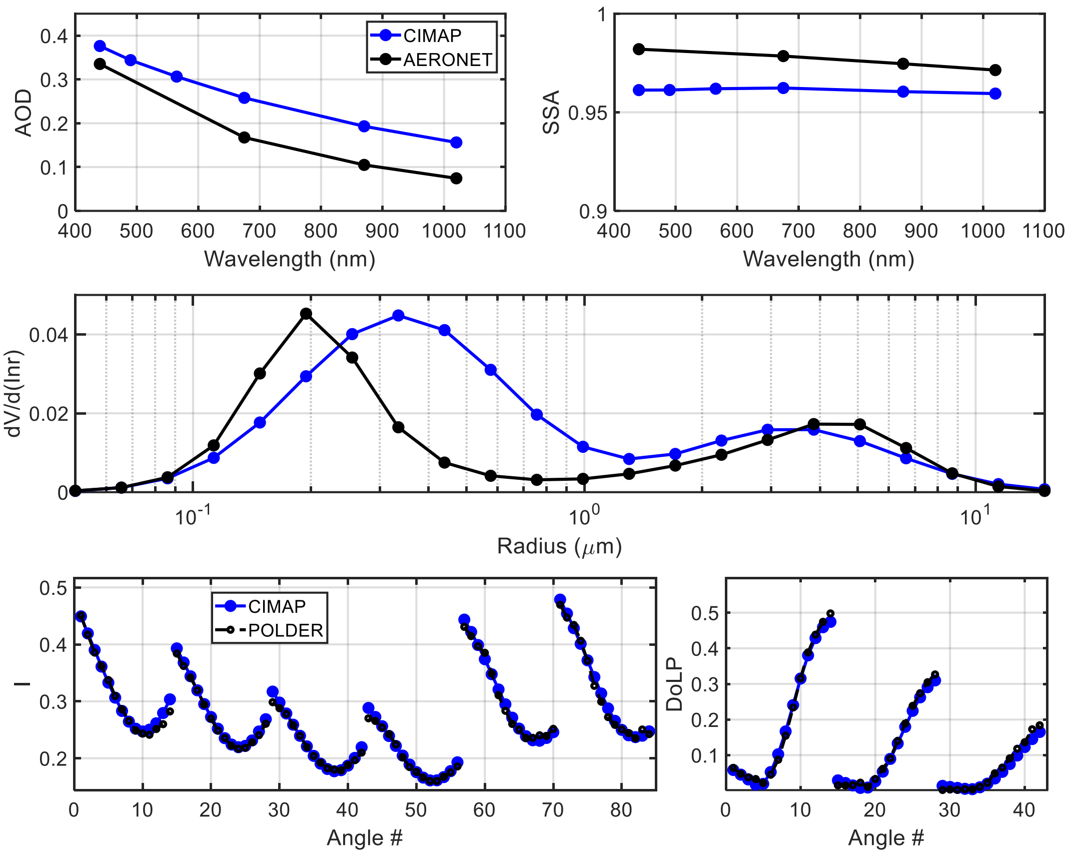


Figure 41. Upper left panel: Comparison of CIMAP-P retrieved AODs to AERONET reference data at 440, 675, 870, and 1020nm. Upper right panel: Same plot as upper left panel but for SSA comparison. Middle panel: Same comparisons but for volumetric particle size distribution. Lower left panel: Comparison of POLDER measured and modeled radiances using the retrieved solution. Lower right panel: Comparison of POLDER measured and modeled DoLP.

4.3.2.3 Retrieval Validation Across the Full Dataset

To gain more statistics about CIMAP-P's performance in retrieving aerosol properties, we use it to retrieve the full dataset collocated with AERONET and GRASP. The MAD, R^2, STD, and RMSE are used for statistical comparison with AERONET.

We first evaluate the retrieval quality of CIMAP-P AOD using the initial guess of surface reflectance informed by the MAIAC dataset. A total of 273 data points were identified from the full dataset for which all three GRASP configurations and CIMAP-P report AOD at the collocated AERONET sites. The first row of panels in Fig. 42 shows the CIMAP-P AOD and AERONET

AOD regression for the four AERONET wavelengths including 440, 675, 870 and 1020 nm. The second, third, and fourth rows of Fig. 42 show the comparison of GRASP-C, GRASP-HP, and GRASP-M AODs against AERONET AODs, respectively. The retrieved AODs at POLDER wavelengths are interpolated at four AERONET wavelengths to make the comparison. Regression against AERONET shows that CIMAP-P exhibits lowest MAD, STD, and RMSE at 675, 870, and 1020nm. At the 440 nm wavelength, the GRASP-C and GRASP-M maintain the highest performance of accuracy as indicated by all four metrics. Measured by regression slopes, the GRASP-M and GRASP-C products exhibit superior performance. The lower regression slope observed in the CIMAP-P retrieval is primarily attributed to increased amount of scattered data when AOD is larger than 0.15 or less than 0.05. CIMAP-P's gain of smaller MAD, STD, and RMSE of AODs in red-to-near-infrared is largely driven by the imposition of a priori constraints, including (a) PCs of aerosol properties derived from AERONET climatology, and (b) initial guesses of surface reflection from the MAIAC dataset.

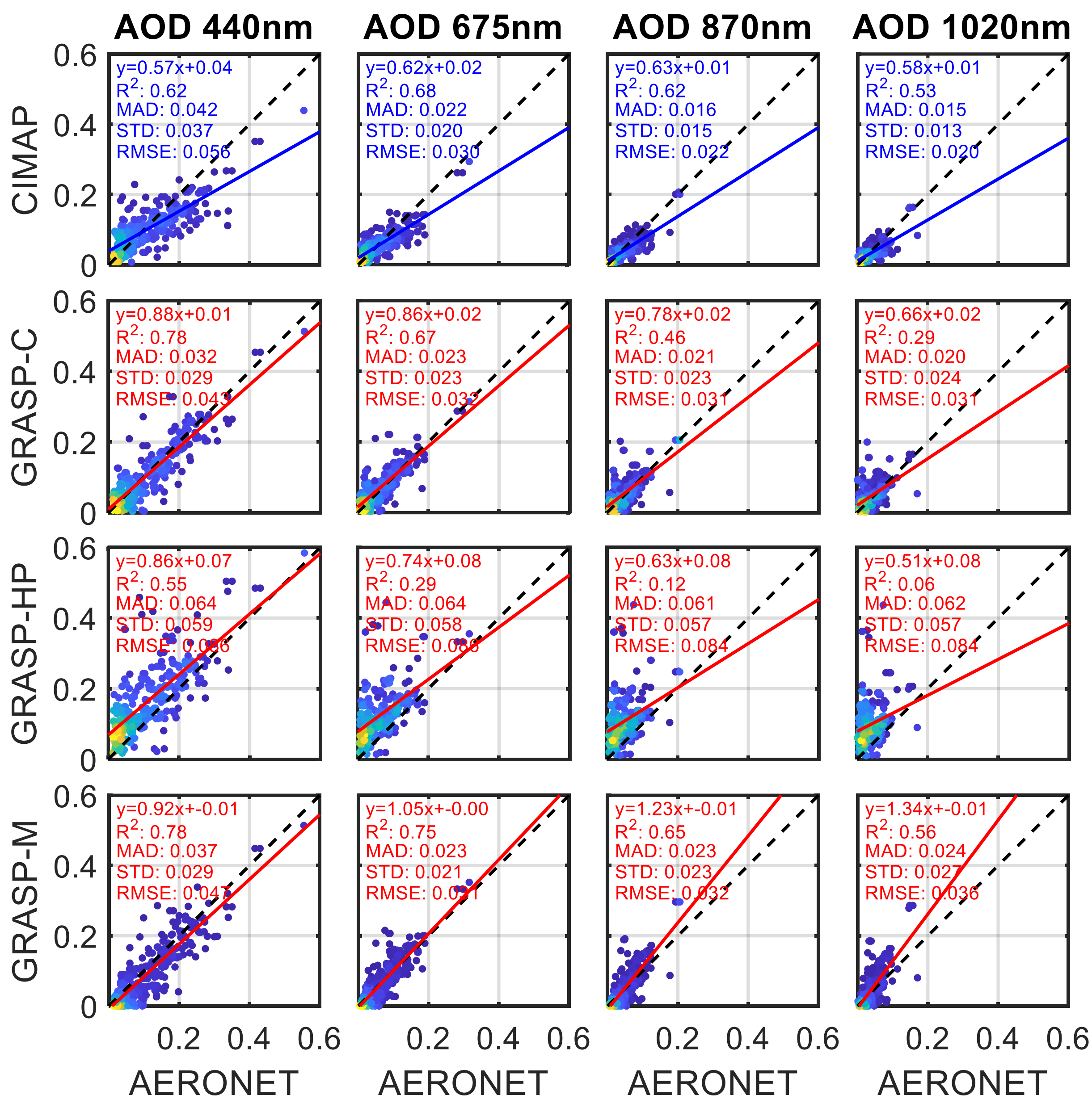


Figure 42. Comparisons of retrieved AODs at 440, 675, 870, and 1020nm. The solid blue or red lines indicate the linear regression. Statistics are labeled in the upper left corner of each subplot. The first row: Comparisons of CIMAP-P AOD to AEORNET data. The second row: Comparisons of GRASP-Component AOD to AEORNET data. The third row: Comparisons of GRASP- High-Precision AOD to AEORNET data. The fourth row: Comparisons of GRASP-Models AOD to AEORNET data. The color represents point density, where warmer color (yellow) indicates higher concentrations of data and cooler color (blue) represent lower density. A total of 273 GRASP-CIMAP-AERONET matchups are displayed here.

To assess the impact of the MAIAC surface a priori, we also performed a CIMAP-P retrieval without using the MAIAC data (hereafter referred to as the "CIMAP-W" configuration) and demonstrate the AODs in Figure 43. Compared to the retrieval using the MAIAC surface a priori (Fig. 42, first row), CIMAP-W exhibited slightly increased MAD, and RMSE. In addition, increased spread of AOD was observed due to the appearance of more outliers. However, the general performance remains reasonable, reflecting the benefit of using MAIAC product as an initial guess of surface reflection.

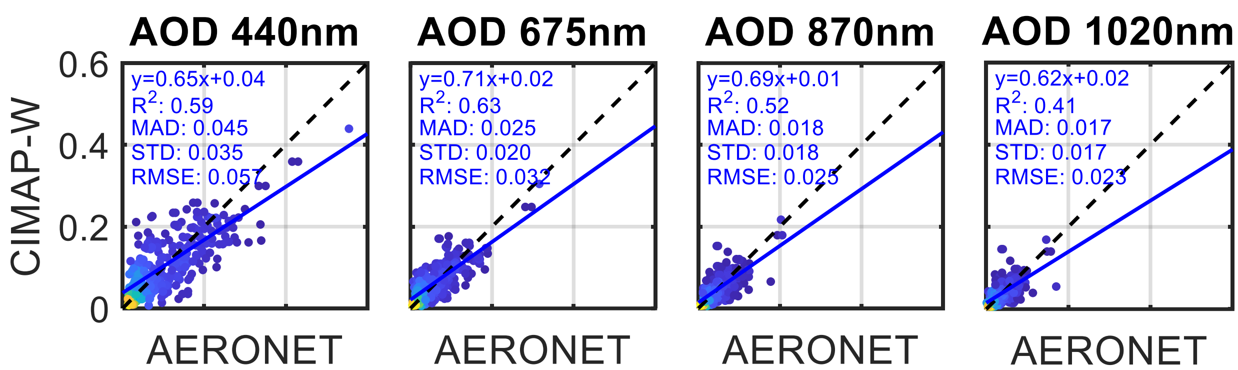


Figure 43. Comparisons of CIMAP retrieved AODs without applying MAIAC surface a priori (namely CIMAP-W configuration) against AERONET data. Same as in Figure 42, a total of 273 matchups is displayed here.

We further compare CIMAP-P and GRASP retrievals of PSD and SSA against AERONET reference data. Given that the number of GRASP-AERONET collocations depend on GRASP configurations, we have identified GRASP-configuration dependent matchups with AERONET for comparison. For PSD comparison, we use the effective radius (r_{eff}) for fine and coarse mode particles, which is calculated by Eq. (75). The separation point between the fine and coarse mode particles is determined by identifying the bin where the minimum volume concentration occurs within the size interval ranging from 0.439 to 0.992 μm, referring to AERONET Inversion Products (Version 3).

Figure 44 illustrates the comparison of CIMAP and GRASP retrieved effective radii to AERONET reference data. The upper two rows present the comparisons of CIMAP and GRASP-Component results to AERONET reference, respectively. The third and fourth rows display the comparisons for GRASP-HP and GRASP-M results to AERONET reference, respectively. For both fine and coarse mode particles, CIMAP reveals a reasonable agreement with AERONET r_{eff}, with a majority of data points clustered closely along the 1:1 dashed line. CIMAP achieves MAD values of 0.032 and 0.402 μm, alongside STD values of 0.028 and 0.3 for the fine and coarse particles, respectively. Compared to GRASP-C, CIMAP demonstrates better performance in terms of both MAD and STD metrics. This is primarily attributed to the imposition of correlation-based a priori of PSD, which effectively captures the relationship inherent in aerosol size distributions. Moreover, these results highlight CIMAP's improvement of retrieving coarse mode PSD it leads to smallest MAD. It should be noted that due to the limited number of GRASP-HP-AERONET matchups and GRASP-M-AERONET matchups for PSD comparison (less than 30), the MAD and STD used to assess GRASP's performance in retrieving PSD is likely subjected to bias.

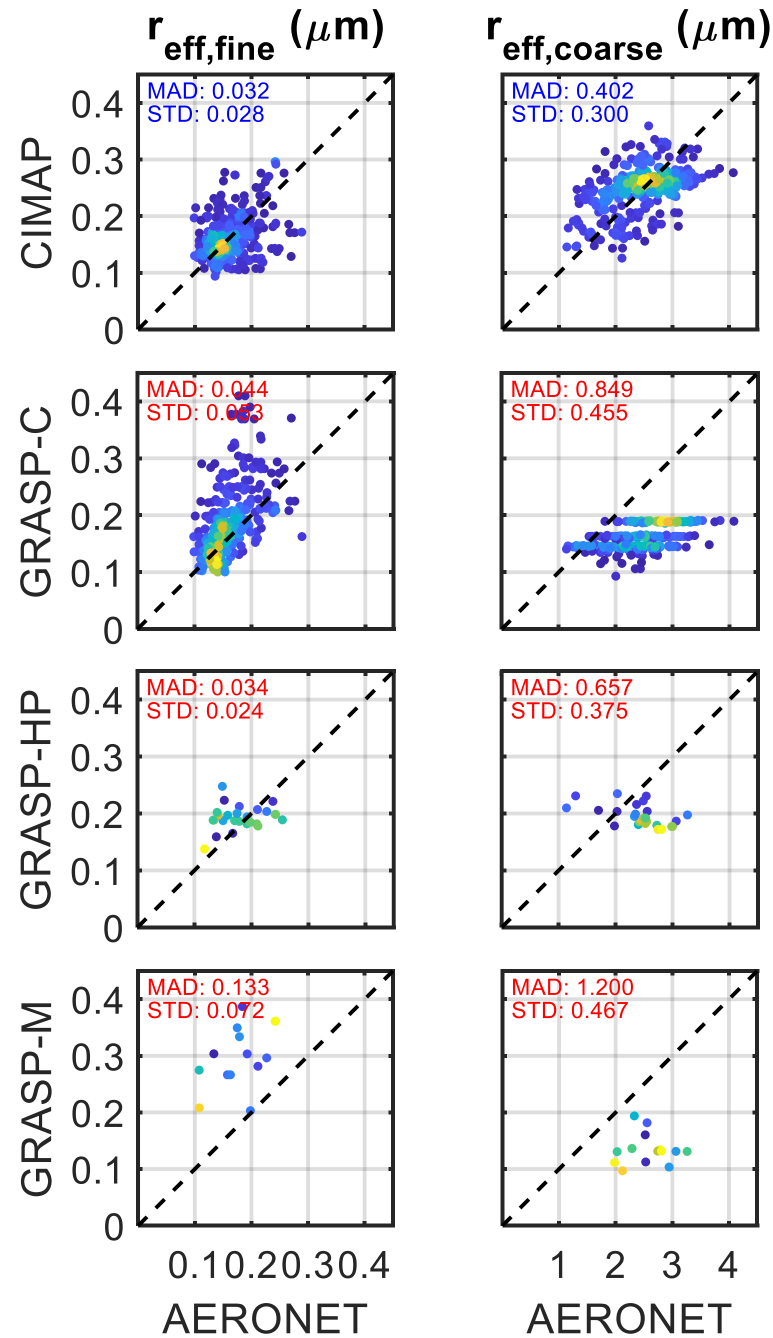


Figure 44. Comparisons of CIMAP and GRASP retrievals of effective radius of fine (left column) and coarse (right column) mode PSD to the AERONET data. The first row: CIMAP versus AERONET data, totaling 324 points. The second row: GRASP-Component versus AEORNET data, totaling 324 points. The third row: GRASP- High-Precision versus AEORNET data, totaling 16 points. The fourth row: GRASP-Models versus AEORNET data, totaling 15 points.

The comparisons of SSA retrievals are demonstrated in Fig. 45. A total of 345 CIMAP-AERONET matchups were found, and the comparison is made in the first row. The second and third rows break the SSA comparison into two regimes, namely AOD<0.3 (331 matchups) and AOD>0.3 (14 matchups), respectively. The fourth, fifth and sixth rows display the available GRASP products from Component (16 matchups), High-Precision (26 matchups), and Models (15 matchups) configurations, respectively. CIMAP-P retrieval of SSA yields MADs of 0.048, 0.052, 0.064, and 0.070 across the four wavelengths. Moreover, CIMAP-P’s SSA retrieval error is around 0.05 for low or moderately high aerosol loadings (AOD<0.30) and less than 0.02 for higher aerosol loadings (AOD>0.30) at the visible wavelengths. As expected, the SSA error increases when for AOD reduces. However, even for low or moderately high aerosol loadings, the SSA retrieval error can still retain a level of 0.05 from CIMAP retrieval. Due to that GRSAP has stringent data selection criteria [see https://www.grasp-open.com/products/polder-data-release], less SSA products are available for comparison against AERONET reference data.

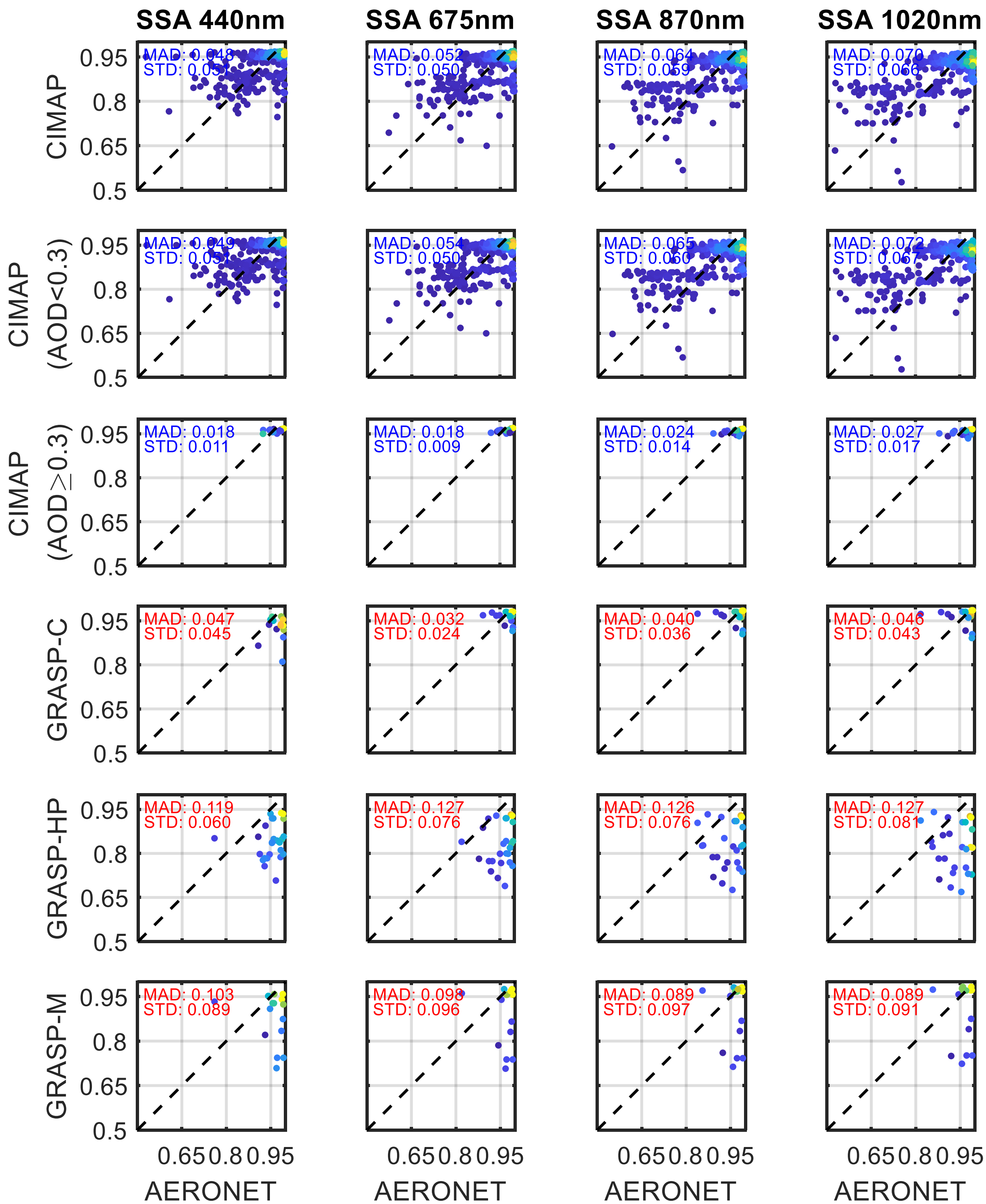


Figure 45. Comparisons of CIMAP-P and GRASP retrievals of SSA at 440, 675, 870, and 1020nm to the AERONET data. Statistics are labeled in the upper left corner of each subplot. The first row: CIMAP-P versus AERONET data, totaling 345 points. The second row: Comparisons of CIMAP-P SSA to AEORNET data for AOD<0.3, totaling 331 matchups. The third row: Comparisons of CIMAP-P SSA to AEORNET data for AOD>0.3, totaling 14 matchups. The fourth row: GRASP-Component versus AEORNET data, totaling 16 matchups. The fifth row: GRASP-High-

Precision versus AEORNET data, totaling 26 matchups. The sixth row: GRASP-Models versus AEORNET data, totaling 15 matchups.

The metrics used to assess the performance of CIMAP-P, CIMAP-W and GRASP retrievals of AOD, SSA, and effective radius against AERONET reference data are summarized in Table 6–Table 10, including regression slopes, R^2, MAD, STD, and RMSE. Generally, CIMAP-P achieves comparable performance to the GRASP-C products of AOD and SSA, though the two approaches differ in algorithm basis. Benefiting from utilizing correlation in aerosol properties, CIMAP-P gain high retrieval accuracy for effective radii of fine and coarse mode PSDs. Moreover, PC based CIMAP retrieval reduces the dimensionality of the retrieval parameters, leading to improved retrieval efficiency and effective mitigation of the ill-posedness in aerosol inversion. Despite the benefits, retrieval errors still exist. Possible causes include the surface BRF modeling errors, radiometric and polarimetric measurement errors, the assumptions about aerosol particle shape (e.g. sphere and spheroid) in light scattering simulation and aerosol vertical profiles (namely assuming "Gaussian" distribution), as well as the small residual error associated with representing aerosol variability using a limited number of PCs, despite capturing the dominant modes of variation across the target area.

Table 6. AOD retrieval accuracy assessment for CIMAP-P and GRASP, using AERONET data as the reference. Sample size: 273. AOD reference wavelengths: 440 nm and 675 nm. Values in shadow denote the algorithm of best performance.

	AOD_{440nm}					AOD_{675nm}				
	CIMAP-P	CIMAP-W	GRASP-C	GRASP-HP	GRASP-M	CIMAP-P	CIMAP-W	GRASP-C	GRASP-HP	GRASP-M
Slope	0.57	0.65	0.88	0.86	0.92	0.62	0.71	0.86	0.74	1.05
R^2	0.62	0.59	0.78	0.55	0.78	0.68	0.63	0.67	0.29	0.75
MAD	0.042	0.045	0.032	0.064	0.037	0.022	0.025	0.023	0.064	0.023
STD	0.037	0.035	0.029	0.059	0.029	0.020	0.020	0.023	0.058	0.021
RMSE	0.053	0.057	0.043	0.086	0.047	0.030	0.032	0.032	0.086	0.031

Table 7. Same as Table 6, but for AODs at the reference wavelengths of: 870nm and 1020nm.

	AOD_{870nm}					AOD_{1020nm}				
	CIMAP-P	CIMAP-W	GRASP-C	GRASP-HP	GRASP-M	CIMAP-P	CIMAP-W	GRASP-C	GRASP-HP	GRASP-M
Slope	0.63	0.69	0.78	0.63	1.23	0.58	0.62	0.66	0.51	1.34
R^2	0.62	0.52	0.46	0.12	0.65	0.53	0.41	0.29	0.06	0.56
MAD	0.016	0.018	0.021	0.061	0.023	0.015	0.017	0.020	0.062	0.024
STD	0.015	0.018	0.023	0.057	0.023	0.013	0.017	0.024	0.057	0.027
RMSE	0.022	0.025	0.031	0.084	0.032	0.020	0.023	0.031	0.084	0.036

Table 8. SSA retrieval accuracy assessment for CIMAP-P and GRASP, using AERONET data as the reference SSA reference wavelengths: 440 nm and 675 nm. Sample size: 345 for CIMAP-P and CIMAP-W, and 16, 26, and 15 for GRASP-C, GRASP-HP, and GRASP-M, respectively. Values in shadow denote the algorithm of best performance.

	SSA_{440nm}					SSA_{675nm}				
	CIMAP-P	CIMAP-W	GRASP-C	GRASP-HP	GRASP-M	CIMAP-P	CIMAP-W	GRASP-C	GRASP-HP	GRASP-M
MAD	0.048	0.063	0.047	0.119	0.103	0.052	0.056	0.032	0.127	0.098
STD	0.051	0.076	0.045	0.060	0.089	0.050	0.056	0.024	0.076	0.096

Table 9. Same as Table 8, but for SSAs at the reference wavelengths of 870 and 1020nm.

	SSA_{870nm}					SSA_{1020nm}				
	CIMAP-P	CIMAP-W	GRASP-C	GRASP-HP	GRASP-M	CIMAP-P	CIMAP-W	GRASP-C	GRASP-HP	GRASP-M
MAD	0.064	0.061	0.040	0.126	0.089	0.070	0.063	0.046	0.127	0.089
STD	0.059	0.050	0.036	0.076	0.097	0.066	0.055	0.043	0.081	0.091

Table 10. Assessment of retrieval accuracy of effective radius of fine and coarse mode particles for CIMAP-P and GRASP, using AERONET data as the reference. Sample size: 324 for CIMAP-P, CIMAP-W, and GRASP-C, and 16 and 15 for GRASP-HP and GRASP-M, respectively. Values in shadow denote the algorithm of best performance.

	$r_{\mathrm{eff,fine}}$					$r_{\mathrm{eff,coarse}}$				
	CIMAP-P	CIMAP-W	GRASP-C	GRASP-HP	GRASP-M	CIMAP-P	CIMAP-W	GRASP-C	GRASP-HP	GRASP-M
MAD	0.032	0.042	0.044	0.034	0.133	0.402	0.491	0.849	0.657	1.200
STD	0.028	0.031	0.053	0.024	0.072	0.300	0.365	0.455	0.375	0.467

4.3.3 Summary for CIMAP-P

Building on the AERONET implementation (CIMAP-A) described above, we developed the correlation-based inversion algorithm CIMAP-P for satellite-borne POLDER observations over the SWCAUS region. Like its predecessor, CIMAP-P capitalizes on the spatial, temporal, and spectral correlations in aerosol optical and physical properties using a small number of PCs. Retrieval of the weights of these few dominant PCs substantially reduces the dimensionality of the state vector, alleviates the ill-posedness of the inversion, and improves retrieval stability and accuracy. In this configuration, each type of aerosol property is constructed into PC space with spatial and seasonal classifications as the implementation of the correlation constraint. In addition, CIMAP-P allows surface reflectance parameters to be initialized using the MAIAC dataset, further enhancing aerosol retrieval performance over land surfaces. CIMAP-P has been implemented to retrieve aerosol properties from POLDER satellite observations. Comparison with collocated

AERONET reference data yields MAD values of 0.042, 0.022, 0.016, and 0.015 for AOD, and 0.048, 0.052, 0.064, and 0.070 for SSA at the reference wavelengths of 440, 675, 870, and 1020 nm, respectively. Furthermore, the MADs of the retrieved effective radius are 0.032 μm and 0.402 μm for the fine- and coarse-mode aerosols, respectively. These results demonstrate that CIMAP-P provides competitive retrieval accuracy, with performance comparable to that of three GRASP configurations (GRASP-C, GRASP-HP, and GRASP-M). Comparison of the AOD retrievals with and without the use of MAIAC surface a priori information (the latter denoted CIMAP-W) further shows that incorporating the MAIAC surface a priori improves retrieval accuracy.

Taken together, these preliminary results establish a foundation for the CIMAP framework to be applied to additional satellite-borne measurements, to incorporate and develop further correlation constraints, and ultimately to provide reliable, long-term aerosol records for climate and air quality research.

5. Summary and Outlook

To investigate the impacts of aerosol-radiation and aerosol-cloud interactions, stable and robust aerosol retrieval algorithms are needed to derive aerosol microphysical properties from instruments embedded on different platforms. In recent decades, successive generations of instruments have been designed and implemented to obtain aerosol-related measurements, with an increasing trend toward satellite-borne observations equipped with polarimetry. Correspondingly, aerosol retrieval algorithms have evolved as the direct methodology for deriving aerosol properties from these measurements, with ongoing improvements through the incorporation of various constraints. In particular, constraints built upon well-characterized a priori information are well suited to aerosol retrievals operating under limited information, as they help mitigate the inherent ill-posedness of the inversion problem. Motivated by this need, this dissertation introduces two applications intended to improve current state-of-the-art retrieval techniques through the use of a priori information drawn from well-tested datasets.

The first application is an algorithm that converts Ross-Li-based surface model parameters into RPV-based parameters. The purpose of this design is to bridge a priori information from well-established Ross-Li-based datasets to retrievals that employ RPV-based surface models, by providing converted parameters as surface a priori for retrieval initialization. The Ross-Li-based surface model parameters can be converted into the d-RPV, d-mRPV, RPV, and mRPV models, under either of two options: assuming all model parameters are wavelength-dependent, or assuming spectral invariance. The MAIAC dataset was evaluated and selected as the source for this conversion, as it provides globally covered, high-resolution sRTLS model parameters. The quality of the conversion was assessed through comparisons of surface DHR and BRF, yielding errors within acceptable bounds. Further comparisons of retrieved aerosol properties, using an optimization-based, two-step retrieval (MAIA-like) initialized with surface a priori either before or after conversion, show that the converted model parameters introduce only minimal impact, demonstrating the reliability of this application. As such, this conversion algorithm is expected to benefit aerosol retrievals by supplying surface a priori information, offering an improvement over retrievals lacking such information.

The second application is an aerosol inversion framework, CIMAP, which adopts correlation constraints to improve the accuracy and efficiency of the retrieval. The implementation of correlation constraints utilizes the intrinsic correlation among aerosol properties to construct an aerosol a priori from an aerosol dataset by means of PCA. This approach is able to preserve a selected level of correlation through a chosen number of PCs while substantially reducing the parameter space of the state vector, thereby achieving both retrieval robustness and acceleration.

Two CIMAP configurations, CIMAP-A and CIMAP-P, are introduced in this dissertation. CIMAP-A serves as the baseline configuration, leveraging the strong information content of ground-based AERONET observations to retrieve aerosol properties in three study regions: CA, BJ, and NB. Correlation constraints are imposed on the retrieval alongside simplifying assumptions for aerosol and surface properties. Comparisons with AERONET inversion products demonstrate the high accuracy of CIMAP-retrieved aerosol properties across the three regions using only 7-8 PCs (out of 30 available). The MAD between CIMAP retrievals and AERONET products are 0.005, 0.019, 0.039, and 0.003 for AOD, SSA, RIR, and RII, respectively, and 0.013 μm and 0.02 μm for the effective radii of the fine and coarse modes, respectively, while achieving over 80% acceleration in processing speed.

Building on this baseline, the second configuration, CIMAP-P, represents an extended version designed to test the framework's applicability to a wide range of future satellite-borne observations. CIMAP-P utilizes POLDER observations as a representative model of current satellite-borne measurements, characterized by global coverage, multi-angle viewing, multi-spectral channels, and polarization. In this configuration, CIMAP-P retains the correlation constraints while relaxing certain simplifying assumptions, retrieving the full set of aerosol and surface properties and incorporating smoothness constraints. The MAIAC dataset is used to establish the surface a priori, with the first application's conversion algorithm employed to convert the sRTLS parameters into the mRPV model. Comparison with collocated AERONET reference data yields AOD MAD of 0.042, 0.022, 0.016, and 0.015, and SSA MAD of 0.048, 0.052, 0.064, and 0.070 at the reference wavelengths of 440, 675, 870, and 1020 nm, respectively. Furthermore, the retrieved effective radius MADs are 0.032 μm and 0.402 μm for the fine- and coarse-mode PSD, respectively. Overall, these results demonstrate reliable agreement with AERONET inversion products, with retrieval performance approaching that of three GRASP configurations (GRASP-C, GRASP-HP, and GRASP-M).

Several assumptions made in the CIMAP approach merit further investigation. For instance, a Gaussian profile is assumed for the vertical distribution of aerosols [e.g., Dubovik et al., 2011; Xu et al., 2017]. While computationally convenient, this assumption has limitations in fully representing complex vertical structures, such as multi-layer aerosol distributions, elevated dust plumes, or strong convective conditions. Similarly, non-spherical particles are represented by spheroids [Dubovik et al., 2006], which may lack fidelity in some realistic scenarios. When lidar data becomes available, a combined lidar-polarimetric retrieval [e.g., Xu et al., 2021; Regmi et al., 2026] could help mitigate these limitations by resolving the vertical profile of aerosol abundance and providing additional insight into particle non-sphericity.

As ongoing work, CIMAP-P retrievals are being extended to global POLDER observations in order to cover a broader range of surface and aerosol regimes. To this end, the CIMAP-P framework is under extension to construct region-specific PCs as a priori information, drawing on available aerosol records. In regions with representative aerosol records, the approach can be applied directly, as demonstrated in this dissertation. In regions with sparse or no ground-based measurements, alternative data sources, such as satellite climatology, reanalysis products, or field campaign data, can instead be used to perform the PCA. Additionally, incorporating further PC-related quantities (e.g., weights, elements, and spatiotemporal means) as extra degrees of freedom in the retrieval [Xu et al., 2019] would allow the retrieval to better capture regional deviations from climatological correlations and mitigate biases, particularly in regions where prior information is limited or unrepresentative. Beyond this, the incorporation of both inner- and cross-property correlations among aerosol and surface variables is being investigated. This will be achieved by extending the current parameter-wise PCA framework to a joint representation, enabling correlations to be captured both within a given type of aerosol parameter (e.g., among size distribution and refractive index variables) and across different types of retrieval parameters (e.g., aerosol-surface interactions). Such an approach would impose more physically meaningful constraints on the aerosol inversion.

In parallel, the introduction of spatial smoothness constraints on aerosol variables across neighboring pixels is also being explored. This would be implemented through additional regularization terms that penalize large spatial gradients, thereby enforcing spatial consistency in the retrieval field. Such constraints have proven effective in aerosol retrievals from airborne polarimetric measurements [Xu et al., 2019], where they reduce pixel-level variability in aerosol properties under low-signal or high-uncertainty conditions.

It should be noted that retrieval speed performance has not been formally evaluated for CIMAP-P, since the retrievals are performed on a supercomputing resource (the OU Supercomputing Center for Education & Research; OSCER), where performance varies depending on the cores and nodes allocated, precluding a consistent speed evaluation. One potential solution is to replace the forward modeling module with a machine-learning-based module (e.g., a feedforward neural network [Gao et al., 2021a, 2021b, 2022, 2023; Fan et al., 2019; Stamnes et al., 2023]), which would allow direct comparison between retrievals using the full parameter set (all PCs) and a reduced parameter set (fewer PCs) based on the acceleration afforded by the machine-learning-based module. This change in forward modeling is not intended to replace the physical model but rather to facilitate exploration of correlation-based constraint development [Redemann and Gao, 2023].

Looking forward, the CIMAP framework holds particular promise for application to current and future multi-angle polarimetric missions such as PACE, MAIA, and 3MI. These missions provide enhanced spectral, angular, and polarization measurements, resulting in increased information content for aerosol retrieval. By leveraging correlation structures among aerosol parameters, the CIMAP algorithm reduces the dimensionality of the retrieval parameter space while integrating multiple constraints within a unified inversion framework. As a result, it offers an efficient means of exploiting this increased information content while maintaining numerical stability. The flexibility of the CIMAP framework further allows its adaptation to different instrument configurations and observational geometries, positioning it as a strong candidate approach for global aerosol remote sensing inversion.

6. Reference

Ackerman, A., Toon, O., Stevens, D., Heymsfield, A., Ramanathan, V., & Welton, E. (2000). Reduction of tropical cloudiness by soot. Science, 288, 1042–1047.

Akagi, S. K., Yokelson, R. J., Wiedinmyer, C., Alvarado, M. J., Reid, J. S., Karl, T., Crounse, J. D., & Wennberg, P. O. (2011). Emission factors for open and domestic biomass burning for use in atmospheric models. Atmospheric Chemistry and Physics, 11(9), 4039–4072.

Albrecht, B. (1989). Aerosols, cloud microphysics, and fractional cloudiness. Science, 245, 1227–1230.

Andreae, M. O. (2019). Emission of trace gases and aerosols from biomass burning – an updated assessment. Atmospheric Chemistry and Physics, 19(13), 8523–8546.

Blanchard, C. L., Hidy, G. M., Tanenbaum, S., Edgerton, E. S., & Hartsell, B. E. (2013). The Southeastern Aerosol Research and Characterization (SEARCH) study: Temporal trends in gas and PM concentrations and composition, 1999–2010. Journal of the Air & Waste Management Association, 63(3), 247–259.

Bodhaine, B. A., Wood, N. B., Dutton, E. G., & Slusser, J. R. (1999). On Rayleigh optical depth calculations. Journal of Atmospheric and Oceanic Technology, 16(11), 1854–1861.

Brown de Colstoun, E. C., & Walthall, C. L. (2006). Improving global scale land cover classifications with multi-directional POLDER data and a decision tree classifier. Remote Sensing of Environment, 100(4), 474–485.

Chan, E. A. W., Gantt, B., & McDow, S. (2018). The reduction of summer sulfate and switch from summertime to wintertime PM2.5 concentration maxima in the United States. Atmospheric Environment, 175, 25–32.

Chen, C., Dubovik, O., Litvinov, P., et al. (2024). Extended aerosol and surface characterization from S5P/TROPOMI with GRASP algorithm. Part II: Global validation and intercomparison. Remote Sensing of Environment, 313, 114327.

Chylek, P., & Coakley, J. (1974). Aerosols and climate. Science, 183, 75–77.

Cox, C., & Munk, W. (1954a). Statistics of the sea surface derived from sun glitter. Journal of the Optical Society of America, 44, 838–850.

Cox, C., & Munk, W. (1954b). Measurement of the roughness of the sea surface from photographs of the sun's glitter. Journal of the Optical Society of America, 44, 838–850.

Deschamps, P. Y., Bréon, F. M., Leroy, M., Podaire, A., Bricaud, A., Buriez, J. C., et al. (1994). The POLDER mission: Instrument characteristics and scientific objectives. IEEE Transactions on Geoscience and Remote Sensing, 32, 598–615.

Didan, K., & Barreto-Muñoz, A. (2019). MODIS vegetation index user's guide (MOD13 series). University of Arizona.

Diner, D. J., Abdou, W. A., Ackerman, T. P., Crean, K., Gordon, H. R., Kahn, R. A., Martonchik, J. V., McMuldroch, S., Paradise, S. R., & Pinty, B. (2008). Level 2 aerosol retrieval algorithm theoretical basis (JPL Publication D-11400, Rev. G). Jet Propulsion Laboratory, California Institute of Technology.

Diner, D. J., Beckert, J. C., Reilly, T. H., Bruegge, C. J., Conel, J. E., Kahn, R. A., et al. (1998). Multi-angle Imaging SpectroRadiometer (MISR) instrument description and experiment overview. IEEE Transactions on Geoscience and Remote Sensing, 36, 1072–1087.

Diner, D. J., Boland, S. W., Brauer, M., Bruegge, C., Burke, K. A., Chipman, R., et al. (2018). Advances in multiangle satellite remote sensing of speciated airborne particulate matter and association with adverse health effects: From MISR to MAIA. Journal of Applied Remote Sensing, 12, 042603.

Diner, D. J., Martonchik, J. V., Borel, C., Gerstl, S., Gordon, H. R., Knyazikhin, Y., Myneni, R., Pinty, B., & Verstraete, M. M. (1999). Level 2 surface retrieval algorithm theoretical basis. Jet Propulsion Laboratory, California Institute of Technology.

Diner, D. J., Martonchik, J. V., Kahn, R. A., Pinty, B., Gobron, N., Nelson, D. L., & Holben, B. N. (2005). Using angular and spectral shape similarity constraints to improve MISR aerosol and surface retrievals over land. Remote Sensing of Environment, 94(2), 155–171.

Diner, D. J., Xu, F., Garay, M. J., Martonchik, J. V., Rheingans, B. E., Geier, S., et al. (2013). The airborne multiangle SpectroPolarimetric Imager (AirMSPI): A new tool for aerosol and cloud remote sensing. Atmospheric Measurement Techniques, 6, 2007–2025.

Diner, D. J., Xu, F., Martonchik, J. V., Rheingans, B. E., Geier, S., Jovanovic, V. M., Davis, A., Chipman, R. A., & McClain, S. C. (2012). Exploration of a polarized surface bidirectional reflectance model using the ground-based multiangle spectropolarimetric imager. Atmosphere, 3(4), 591–619.

Dubovik, O., & King, M. D. (2000). A flexible inversion algorithm for retrieval of aerosol optical properties from Sun and sky radiance measurements. Journal of Geophysical Research: Atmospheres, 105, 673–696.

Dubovik, O., Herman, M., Holdak, A., Lapyonok, T., Tanré, D., Deuzé, J. L., et al. (2011). Statistically optimized inversion algorithm for enhanced retrieval of aerosol properties from spectral multi-angle polarimetric satellite observations. Atmospheric Measurement Techniques, 4, 975–1018.

Dubovik, O., Holben, B. N., Lapyonok, T., et al. (2002). Non-spherical aerosol retrieval method employing light scattering by spheroids. Geophysical Research Letters, 29(10).

Dubovik, O., Lapyonok, T., Litvinov, P., Herman, M., Fuertes, D., Ducos, F., Torres, B., Derimian, Y., Huang, X., Lopatin, A., Chaikovsky, A., Aspetsberger, M., & Federspiel, C. (2014). GRASP: A versatile algorithm for characterizing the atmosphere. SPIE Newsroom.

Dubovik, O., Li, Z., Mishchenko, M. I., Tanré, D., Karol, Y., Bojkov, B., Cairns, B., Diner, D. J., Espinosa, W. R., Goloub, P., Gu, X., Hasekamp, O., Hong, J., Hou, W., Knobelspiesse, K. D., Landgraf, J., Li, L., Litvinov, P., Liu, Y., Lopatin, A., Marbach, T., Maring, H., Martins, V., Meijer, Y., Milinevsky, G., Mukai, S., Parol, F., Qiao, Y., Remer, L., Rietjens, J., Sano, I., Stammes, P., Stamnes, S., Sun, X., Tabary, P., Travis, L. D., Waquet, F., Xu, F., Yan, C., & Yin, D. (2019). Polarimetric remote sensing of atmospheric aerosols: Instruments, methodologies, results, and perspectives. Journal of Quantitative Spectroscopy and Radiative Transfer, 224, 474–511.

Dubovik, O., Schuster, G. L., Xu, F., Hu, Y., Bösch, H., Landgraf, J., & Li, Z. (2021). Grand challenges in satellite remote sensing. Frontiers in Remote Sensing, 2, 619818.

Dubovik, O., Sinyuk, A., Lapyonok, T., et al. (2006). Application of spheroid models to account for aerosol particle nonsphericity in remote sensing of desert dust. Journal of Geophysical Research: Atmospheres, 111, D11208.

Dubovik, O., Smirnov, A., Holben, B. N., King, M. D., Kaufman, Y. J., Eck, T. F., et al. (2000). Accuracy assessments of aerosol optical properties retrieved from Aerosol Robotic Network (AERONET) Sun and sky radiance measurements. Journal of Geophysical Research: Atmospheres, 105, 9791–9806.

Engelsen, O., Pinty, B., Verstraete, M. M., & Martonchik, J. V. (1996). Parametric bidirectional reflectance factor models: Evaluation, improvements and applications. Office for Official Publications of the European Communities.

Fan, C., Fu, G., Di Noia, A., Smit, M., Rietjens, J. H., Ferrare, R. A., Burton, S., Li, Z., & Hasekamp, O. P. (2019). Use of a neural network-based ocean body radiative transfer model for aerosol retrievals from multi-angle polarimetric measurements. Remote Sensing, 11, 2877.

Fletcher, R. (1971). A modified Marquardt subroutine for non-linear least squares (Tech. Rep. AERE-R 6799). Atomic Energy Research Establishment.

Flowerdew, R. J., & Haigh, J. D. (1995). An approximation to improve accuracy in the derivation of surface reflectances from multi-look satellite radiometers. Geophysical Research Letters, 22(13), 1693–1696.

Fougnie, B., Marbach, T., Lacan, A., Lang, R., Schlüssel, P., Poli, G., et al. (2018). The multi-viewing multi-channel multi-polarisation imager—Overview of the 3MI polarimetric mission for aerosol and cloud characterization. Journal of Quantitative Spectroscopy and Radiative Transfer, 219, 23–32.

Gao, M., Franz, B. A., Knobelspiesse, K., Zhai, P.-W., Martins, V., Burton, S., Cairns, B., Ferrare, R., Gales, J., Hasekamp, O., Hu, Y., Ibrahim, A., McBride, B., Puthukkudy, A., Werdell, P. J., & Xu, X. (2021a). Efficient multi-angle polarimetric inversion of aerosols and ocean color powered by a deep neural network forward model. Atmospheric Measurement Techniques, 14, 4083–4110.

Gao, M., Franz, B. A., Zhai, P.-W., Knobelspiesse, K., Sayer, A. M., Xu, X., Martins, J. V., Cairns, B., Castellanos, P., Fu, G., Hannadige, N., Hasekamp, O., Hu, Y., Ibrahim, A., Patt, F., Puthukkudy, A., & Werdell, P. J. (2023). Simultaneous retrieval of aerosol and ocean properties from PACE HARP2 with uncertainty assessment using cascading neural network radiative transfer models. Atmospheric Measurement Techniques, 16, 5863–5881.

Gao, M., Knobelspiesse, K., Franz, B. A., Zhai, P.-W., Martins, V., Burton, S. P., Cairns, B., Ferrare, R., Fenn, M. A., Hasekamp, O., Hu, Y., Ibrahim, A., Sayer, A. M., Werdell, P. J., & Xu, X. (2021b). Adaptive data screening for multi-angle polarimetric aerosol and ocean color remote sensing accelerated by deep learning. Frontiers in Remote Sensing, 2, 46.

Gao, M., Knobelspiesse, K., Franz, B. A., Zhai, P.-W., Sayer, A. M., Ibrahim, A., Cairns, B., Hasekamp, O., Hu, Y., Martins, V., Werdell, P. J., & Xu, X. (2022). Effective uncertainty quantification for multi-angle polarimetric aerosol remote sensing over ocean. Atmospheric Measurement Techniques, 15, 4859–4879.

Ginoux, P., Chin, M., Tegen, I., Prospero, J. M., Holben, B., Dubovik, O., & Lin, S. J. (2001). Sources and distributions of dust aerosols simulated with the GOCART model. Journal of Geophysical Research: Atmospheres, 106(D17), 20255–20273.

Hallquist, M., Wenger, J. C., Baltensperger, U., Rudich, Y., Simpson, D., Claeys, M., Dommen, J., Donahue, N. M., George, C., Goldstein, A. H., Hamilton, J. F., Herrmann, H., Hoffmann, T., Iinuma, Y., Jang, M., Jenkin, M. E., Jimenez, J. L., Kiendler-Scharr, A., Maenhaut, W., … Wildt, J. (2009). The formation, properties and impact of secondary organic aerosol: Current and emerging issues. Atmospheric Chemistry and Physics, 9(14), 5155–5236.

Hand, J. L., Prenni, A. J., & Schichtel, B. A. (2024). Trends in seasonal mean speciated aerosol composition in remote areas of the United States from 2000 through 2021. Journal of Geophysical Research: Atmospheres, 129.

Hansen, J. E., & Travis, L. D. (1974). Light scattering in planetary atmospheres. Space Science Reviews, 16, 527–610.

Hasekamp, O., Litvinov, P., & Butz, A. (2011). Aerosol properties over the ocean from PARASOL multi-angle photopolarimetric measurements. Journal of Geophysical Research: Atmospheres, 116, D14204.

Henyey, L. G., & Greenstein, J. L. (1941). Diffuse radiation in the galaxy. The Astrophysical Journal, 93, 70–83.

Holben, B. N., Eck, T. F., Slutsker, I., Smirnov, A., Sinyuk, A., Schafer, J., Giles, D., & Dubovik, O. (2006). AERONET's Version 2.0 quality assurance criteria. In Remote sensing of the atmosphere and clouds (Vol. 6408, p. 64080Q). SPIE.

Holben, B. N., Eck, T. F., Slutsker, I., Tanré, D., Buis, J. P., Setzer, A., et al. (1998). AERONET—a federated instrument network and data archive for aerosol characterization. Remote Sensing of Environment, 66, 1–16.

Hoogeveen, R., Kohlhaas, R., Campo, J., Rietjens, J., Eigenraam, A., et al. (2023). SPEXone, the multi-angle spectropolarimeter for PACE, adjusted and improved for new space applications. In ICSO 2022 (Vol. 12777, pp. 1914–1923). SPIE.

Hou, W., Liu, X., Wang, J., Chen, C., & Xu, X. (2025). Multispectral land surface reflectance reconstruction based on non-negative matrix factorization: Bridging spectral resolution gaps for GRASP TROPOMI BRDF product in visible. Remote Sensing, 17(6), 1053.

Hsu, N. C., Lee, J., Sayer, A. M., Kim, W., Bettenhausen, C., & Tsay, S. C. (2019). VIIRS deep blue aerosol products over land: Extending the EOS long-term aerosol data records. Journal of Geophysical Research: Atmospheres, 124, 4026–4053.

Hsu, N. C., Tsay, S. C., King, M. D., & Herman, J. R. (2004). Aerosol properties over bright-reflecting source regions. IEEE Transactions on Geoscience and Remote Sensing, 42, 557–569.

Huang, T., Xu, F., Gao, L., Flynn, C., & Dubovik, O. (2021). A Correlation-based Inversion Method for Aerosol Property (CIMAP) Retrieval from AERONET Measurements. Journal of Quantitative Spectroscopy and Radiative Transfer, 272, 107808.

Huang, T., Zhang, W., Xu, F., Diner, D. J., Kalashnikova, O. V., Lyapustin, A. I., & Wang, Y. (2026a). Conversion from Ross-Li to Rahman-Pinty-Verstraete model parameterizations of surface bidirectional reflectance. Journal of Quantitative Spectroscopy and Radiative Transfer, 362, 109998.

Huang, T., Xu, F., Kalashnikova, O., Lyapustin, A. (2026b, manuscript under review). A Correlation-based Inversion Approach for Satellite-borne Aerosol Remote Sensing Journal of Quantitative Spectroscopy and Radiative Transfer.

Huete, A., & Justice, C. (1999). MODIS vegetation index (MOD 13) algorithm theoretical basis document.

Jiao, Z., Woodcock, C., Schaaf, C. B., Tan, B., Liu, J., Gao, F., et al. (2011). Improving MODIS land cover classification by combining MODIS spectral and angular signatures in a Canadian boreal forest. Canadian Journal of Remote Sensing, 37(2), 184–203.

Jones, H. G., & Vaughan, R. A. (2010). Remote sensing of vegetation: Principles, techniques, and applications. Oxford University Press.

Kahn, R. A., & Gaitley, B. J. (2015). An analysis of global aerosol type as retrieved by MISR. Journal of Geophysical Research: Atmospheres, 120(8), 4248–4281.

Kahn, R. A., Gaitley, B. J., Garay, M. J., Diner, D. J., Eck, T. F., Smirnov, A., & Holben, B. N. (2010). Multiangle Imaging SpectroRadiometer global aerosol product assessment by comparison with the Aerosol Robotic Network. Journal of Geophysical Research: Atmospheres, 115(D23), D23209.

Kanakidou, M., Seinfeld, J. H., Pandis, S. N., Barnes, I., Dentener, F. J., Facchini, M. C., … Wilson, J. (2005). Organic aerosol and global climate modelling: A review. Atmospheric Chemistry and Physics, 5, 1053–1123.

King, M. D., Kaufman, Y. J., Menzel, W. P., & Tanre, D. (1992). Remote sensing of cloud, aerosol, and water vapor properties from the moderate resolution imaging spectrometer (MODIS). IEEE Transactions on Geoscience and Remote Sensing, 30, 2–27.

Knobelspiesse, K., Cairns, B., Mishchenko, M., Chowdhary, J., Tsigaridis, K., van Diedenhoven, B., et al. (2012). Analysis of fine-mode aerosol retrieval capabilities by different passive remote sensing instrument designs. Optics Express, 20, 21457–21484.

Koch, D., & Del Genio, A. D. (2010). Black carbon semi-direct effects on cloud cover: Review and synthesis. Atmospheric Chemistry and Physics, 10, 7685–7696.

Kokhanovsky, A. A., Davis, A. B., Cairns, B., Dubovik, O., Hasekamp, O., et al. (2015). Space-based remote sensing of aerosols: The multi-angle spectropolarimetric frontier. Earth-Science Reviews, 145, 85–116.

Laj, P., Bigi, A., Rose, C., Andrews, E., Lund Myhre, C., Collaud Coen, M., et al. (2020). A global analysis of climate-relevant aerosol properties retrieved from the network of Global Atmosphere Watch (GAW) near-surface observatories. Atmospheric Measurement Techniques, 13(8), 4353–4392.

Legrand, M., Dubovik, O., Lapyonok, T., & Derimian, Y. (2014). Accounting for particle non-sphericity in modeling mineral dust radiative properties in the thermal infrared. Journal of Quantitative Spectroscopy and Radiative Transfer, 149, 219–240.

Levenberg, K. (1944). A method for the solution of certain non-linear problems in least squares. Quarterly of Applied Mathematics, 2, 164–168.

Levy, R. C., Mattoo, S., Munchak, L. A., Remer, L. A., Sayer, A. M., Patadia, F., & Hsu, N. C. (2013). The Collection 6 MODIS aerosol products over land and ocean. Atmospheric Measurement Techniques, 6(11), 2989–3034.

Li, L., Dubovik, O., Derimian, Y., Schuster, G. L., Lapyonok, T., Litvinov, P., Ducos, F., Fuertes, D., Chen, C., Li, Z., Lopatin, A., Torres, B., & Che, H. (2019). Retrieval of aerosol components directly from satellite and ground-based measurements. Atmospheric Chemistry and Physics, 19(21), 13409–13443.

Li, X., & Strahler, A. H. (1992). Geometric-optical bidirectional reflectance modeling of the discrete crown vegetation canopy: Effect of crown shape and mutual shadowing. IEEE Transactions on Geoscience and Remote Sensing, 30, 276–292.

Li, Z., Blarel, L., Podvin, T., Goloub, P., Buis, J.-P., & Morel, J.-P. (2008). Transferring the calibration of direct solar irradiance to diffuse-sky radiance measurements for CIMEL Sun-sky radiometers. Applied Optics, 47(10), 1368–1377.

Limbacher, J. A., Kahn, R. A., & Lee, J. (2022). The new MISR research aerosol retrieval algorithm: A multi-angle, multi-spectral, bounded-variable least squares retrieval of aerosol particle properties over both land and water. Atmospheric Measurement Techniques, 15, 6865–6887.

Litvinov, P., Fuertes, D., Lopatin, A., Dubovik, O., Marbach, T., Lacan, A., Vandenbussche, S., Stammes, P., Wang, P., Tilstra, L. G., Chimot, J., De Smedt, I., Lerot, C., Tagliabue, G., Adcock, C., Loyola, D., García-Alvarez, S., Lemmens, B., & Compernolle, S. (2024). Extended aerosol and surface characterization from S5P/TROPOMI with GRASP algorithm. Part I: Conditions, approaches, performance and new possibilities. Remote Sensing of Environment, 312, 114306.

Lopatin, A., Dubovik, O., Chaikovsky, A., Goloub, P., Lapyonok, T., Tanré, D., & Litvinov, P. (2013). Enhancement of aerosol characterization using synergy of lidar and sun-photometer coincident observations: The GARRLiC algorithm. Atmospheric Measurement Techniques, 6(8), 2065–2088.

Lopatin, A., Dubovik, O., Fuertes, D., Stenchikov, G., Lapyonok, T., Veselovskii, I., Wienhold, F. G., Shevchenko, I., Hu, Q., & Parajuli, S. (2021). Synergy processing of diverse ground-based remote sensing and in situ data using the GRASP algorithm: Applications to radiative budget and aerosol microphysical properties retrieval. Atmospheric Measurement Techniques, 14(4), 2575–2614.

Lucht, W., Schaaf, C. B., & Strahler, A. H. (2000). An algorithm for the retrieval of albedo from space using semiempirical BRDF models. IEEE Transactions on Geoscience and Remote Sensing, 38(2), 977–998.

Lyapustin, A., & Wang, Y. (2018). MCD19A3 MODIS/Terra+Aqua BRDF model parameters 8-day L3 global 1km SIN Grid V006 [Data set]. NASA Land Processes Distributed Active Archive Center.

Lyapustin, A., Wang, Y., Korkin, S., & Huang, D. (2018). MODIS Collection 6 MAIAC algorithm. Atmospheric Measurement Techniques, 11(10), 5741–5765.

Lyapustin, A., Wang, Y., Korkin, S., Schaaf, C., Wang, W., & Wang, Z. (2025). Scaled RTLS BRDF model extended to high zenith angles. Frontiers in Remote Sensing, 6, 1533803.

Malm, W. C., Schichtel, B. A., Pitchford, M. L., Ashbaugh, L. L., & Eldred, R. A. (2004). Spatial and monthly trends in speciated fine particle concentration in the United States. Journal of Geophysical Research: Atmospheres, 109, D03306.

Malm, W. C., Sisler, J. F., Huffman, D., Eldred, R. A., & Cahill, T. A. (1994). Spatial and seasonal trends in particle concentration and optical extinction in the United States. Journal of Geophysical Research: Atmospheres, 99(D1), 1347–1370.

Marquardt, D. W. (1963). An algorithm for least-squares estimation of nonlinear parameters. Journal of the Society for Industrial and Applied Mathematics, 11, 431–441.

Martonchik, J. V., Diner, D. J., Kahn, R. A., Ackermann, T. P., Verstraete, M. M., Pinty, B., & Gordon, H. R. (1998). Techniques for the retrieval of aerosol properties over land and ocean using multiangle imaging. IEEE Transactions on Geoscience and Remote Sensing, 36, 1212–1227.

McBride, B. A., Sienkiewicz, N., Xu, X., Puthukkudy, A., Fernandez-Borda, R., et al. (2024). In-flight characterization of the Hyper-Angular Rainbow Polarimeter (HARP2) on the NASA

PACE mission. In Sensors, systems, and next-generation satellites XXVIII (Vol. 13192, pp. 113–122). SPIE.

Meister, G., Knuble, J. J., et al. (2024). The Ocean Color Instrument (OCI) on the Plankton, Aerosol, Cloud, ocean Ecosystem (PACE) mission: System design and prelaunch radiometric performance. IEEE Transactions on Geoscience and Remote Sensing, 62, 5517418.

Meyer, K., Platnick, S., & Zhang, Z. (2015). Simultaneously inferring above-cloud absorbing aerosol optical thickness and underlying liquid phase cloud optical and microphysical properties using MODIS. Journal of Geophysical Research: Atmospheres, 120, 5524–5547.

Minnaert, M. (1941). The reciprocity principle in lunar photometry. The Astrophysical Journal, 93, 403–410.

Mishchenko, M. I., Yatskiv, Y. S., Rosenbush, V. K., & Videen, G. (Eds.). (2011). Polarimetric detection, characterization, and remote sensing. Springer.

Monahan, E. C., Spiel, D. E., & Davidson, K. L. (1986). A model of marine aerosol generation via whitecaps and wave disruption. In E. C. Monahan & G. Mac Niocaill (Eds.), Oceanic whitecaps (pp. 167–174). Springer.

Pachauri, R., & Meyer, L. (Eds.). (2014). Climate change 2014: Synthesis report. Contribution of Working Groups I, II and III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change. IPCC.

Patadia, F., Levy, R. C., & Mattoo, S. (2018). Correcting for trace gas absorption when retrieving aerosol optical depth from satellite observations of reflected shortwave radiation. Atmospheric Measurement Techniques, 11, 3205–3219.

Petro, S., Pham, K., & Hilton, G. (2020). Plankton, Aerosol, Cloud, ocean Ecosystem (PACE) mission integration and testing. In 2020 IEEE Aerospace Conference (pp. 1–20). IEEE.

Pettorelli, N. (2013). The normalized difference vegetation index. Oxford University Press.

Pinty, B., & Ramond, D. (1986). A simple bidirectional reflectance model for terrestrial surfaces. Journal of Geophysical Research, 91, 7803–7808.

Rahman, H., Pinty, B., & Verstraete, M. M. (1993). Coupled surface-atmosphere reflectance (CSAR) model. 2. Semiempirical surface model usable with NOAA Advanced Very High Resolution Radiometer data. Journal of Geophysical Research, 98, 20791–20801.

Redemann, J., & Gao, L. (2024). A machine learning paradigm for necessary observations to reduce uncertainties in aerosol climate forcing. Nature Communications, 15, 8343.

Regmi, G., Espinosa, W. R., Martins, J. V., Puthukkudy, A., Saito, M., Dubovik, O., Momoi, M., & Lopatin, A. (manuscript under review). Combined lidar-polarimeter dust retrievals using spheroidal and hexahedral particle shape models. EGUsphere.

Remer, L. A., & Kaufman, Y. J. (1998). Dynamical aerosol model: Urban/industrial aerosol. Journal of Geophysical Research: Atmospheres, 103(D13), 13859–13871.

Remer, L. A., Kaufman, Y. J., Tanré, D., Mattoo, S., Chu, D. A., Martins, J. V., Li, R.-R., Ichoku, C., Levy, R. C., Kleidman, R. G., & Holben, B. N. (2005). The MODIS aerosol algorithm, products and validation. Journal of the Atmospheric Sciences, 62(4), 947–973.

Ross, J. (1981). The radiation regime and architecture of plant stands. Dr. W. Junk Publishers.

Roujean, J. L., Leroy, M., & Deschamps, P. Y. (1992). A bidirectional reflectance model of the Earth’s surface for the correction of remote sensing data. Journal of Geophysical Research, 97(D18), 20455–20468.

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

Smith, T. B., Lehrman, D. E., Reible, D. D., & Shair, F. H. (1981a). The origin and fate of airborne pollutants within the San Joaquin Valley: Extended summary and special analysis topics (Vol. 1). Prepared for California Air Resources Board. Meteorology Research Inc./California Institute of Technology.

Smith, T. B., Lehrman, D. E., Reible, D. D., & Shair, F. H. (1981b). The origin and fate of airborne pollutants within the San Joaquin Valley: Extended summary and special analysis topics (Vol. 2). Prepared for California Air Resources Board. Meteorology Research Inc./California Institute of Technology.

Stamnes, S., Jones, M., Allen, J. G., Chemyakin, E., Bell, A., Chowdhary, J., Liu, X., Burton, S. P., Van Diedenhoven, B., Hasekamp, O., Hair, J., Hu, Y., Hostetler, C., Ferrare, R., Stamnes, K., & Cairns, B. (2023). The PACE-MAPP algorithm: Simultaneous aerosol and ocean polarimeter products using coupled atmosphere-ocean vector radiative transfer. Frontiers in Remote Sensing, 4, 1–16.

Strahler, A. H. (1994). Vegetation canopy reflectance modeling—Recent developments and remote sensing perspectives. In Proceedings of the 6th International Symposium on Physical Measurements and Signatures in Remote Sensing (pp. 593–600).

Tikhonov, A. N., & Arsenin, V. Y. (1977). Solutions of ill-posed problems (F. John, Trans.). V. H. Winston & Sons.

Twomey, S. (1974). Pollution and planetary albedo. Atmospheric Environment, 8, 1251–1256.

van de Hulst, H. C. (1957). Light scattering by small particles. John Wiley and Sons.

Wanner, W., Li, X., & Strahler, A. H. (1995). On the derivation of kernels for kernel-driven models of bidirectional reflectance. Journal of Geophysical Research: Atmospheres, 100, 21077–21089.

Witek, M. L., Garay, M. J., Diner, D. J., Bull, M. A., & Seidel, F. C. (2018). New approach to the retrieval of AOD and its uncertainty from MISR observations over dark water. Atmospheric Measurement Techniques, 11, 429–439.

Xu, F., Davis, A. B., Sanghavi, S. V., Martonchik, J. V., & Diner, D. J. (2012). Linearization of Markov chain formalism for vector radiative transfer in a plane-parallel atmosphere/surface system. Applied Optics, 51, 3491–3507.

Xu, F., Davis, A. B., West, R. A., Martonchik, J. V., & Diner, D. J. (2011). Markov chain formalism for vector radiative transfer in a plane-parallel atmosphere overlying a polarizing surface. Optics Letters, 36, 2083–2085.

Xu, F., Diner, D. J., Dubovik, O., & Schechner, Y. (2019). A correlated multi-pixel inversion approach for aerosol remote sensing. Remote Sensing, 11, 746.

Xu, F., Dubovik, O., Zhai, P.-W., Diner, D. J., Kalashnikova, O. V., Seidel, F. C., Litvinov, P., Bovchaliuk, A., Garay, M. J., van Harten, G., & Davis, A. B. (2016). Joint retrieval of aerosol and water-leaving radiance from multispectral, multiangular and polarimetric measurements over ocean. Atmospheric Measurement Techniques, 9, 2877–2907.

Xu, F., Gao, L., Redemann, J., Flynn, C. J., Espinosa, W. R., da Silva, A. M., Stamnes, S., Burton, S. P., Liu, X., Ferrare, R., Cairns, B., & Dubovik, O. (2021). A combined lidar-polarimeter inversion approach for aerosol remote sensing over ocean. Frontiers in Remote Sensing, 2, 620871.

Xu, F., van Harten, G., Diner, D. J., Kalashnikova, O. V., Seidel, F. C., Bruegge, C. J., et al. (2017). Coupled retrieval of aerosol properties and land surface reflection using the airborne multiangle spectropolarimetric imager. Journal of Geophysical Research: Atmospheres, 122, 7004–7026.

Yang, P., Meyer, K., Levy, R., Li, D., Xu, F., Rapp, A. D., & Zhang, Z. (2026, In preparation). Satellite optical remote sensing of clouds and aerosols: From particle single-scattering and gaseous absorption through radiative transfer to retrieval products.

Zhang, W., Xu, F., Chen, B., Huang, T., Diner, D. J., Bull, M. A., McDuffie, J. L., Witek, M. L., Kalashnikova, O. V., & Lyapustin, A. (2026, In preparation). MAIA Level 2 aerosol retrieval algorithm test and improvement using MISR and MODIS measurements.

Zhang, X., Li, L., Chen, C., Zheng, Y., Dubovik, O., Derimian, Y., Lopatin, A., Gui, K., Wang, Y., Zhao, H., Liang, Y., Holben, B., & Che, H. (2022). Extensive characterization of aerosol optical properties and chemical component concentrations: Application of the GRASP/Component approach to long-term AERONET measurements. Science of the Total Environment, 812, 152553.

7. CV and Publication

EDUCATION

School of Meteorology, University of Oklahoma Aug 2021- Jul 2026

- Ph.D., Meteorology

School of Electrical and Computer Engineering, University of Oklahoma Aug 2019- Jun 2021

- MS, Electrical Engineering

School of Electrical and Computer Engineering, University of Oklahoma Aug 2016- May 2019

- BS, Electrical Engineering, Minor in Mathematics
- Cumulative GPA: 3.39/4.0

PUBLICATION

Taozhong Huang, Feng Xu, Lan Gao, Connor J. Flynn, Oleg Dubovik, A correlation-based inversion method for aerosol property retrieval from AERONET measurements. *Journal of Quantitative Spectroscopy and Radiative Transfer*. 272, 107808, 2021.

Taozhong Huang, Wenzhi Zhang, Feng Xu, David J. Diner, Olga V. Kalashnikova, Alexei I. Lyapustin and Yujie Wang, Conversion from Ross-Li to Rahman-Pinty-Verstraete model parameterizations of surface bidirectional reflectance. *Journal of Quantitative Spectroscopy and Radiative Transfer*. 362, 109998, 2026.

Taozhong Huang, Feng Xu, Olga V. Kalashnikova, Alexei I. Lyapustin, A Correlation-based Inversion Approach for Satellite-borne Aerosol Remote Sensing. *Journal of Quantitative Spectroscopy and Radiative Transfer*. 110054; 0022-4073, 2026.

Wenzhi Zhang, Feng Xu, Benting Chen, **Taozhong Huang**, David J. Diner, Michael A. Bull, James L. McDuffie, Marcin L. Witek, Olga V. Kalashnikova, Alexei I. Lyapustin, MAIA Level 2 aerosol retrieval algorithm test and improvement using MISR and MODIS measurements. *IEEE Transactions on Geoscience & Remote Sensing*. Under major revision, 2026.

CONFERENCE/ PRESENTATION

Taozhong Huang, Benting Chen, Olga Kalashnikova, Oleg Dubovik, Alexei Lyapustin, and Feng Xu. A Correlation-based Inversion Approach for Aerosol Remote Sensing: Test Using POLDER Data. **Presented at: 2026 AMS Annual Meeting**; Houston, TX, USA.

Taozhong Huang, Wenzhi Zhang, Benting Chen, David J. Diner, Olga Kalashnikova, James L. McDuffie, Michael A. Bull, Marcin L. Witek, Oleg Dubovik, Alexei Lyapustin, and Feng Xu. Comparing MAIAC and MISR Surface Bidirectional Reflectane Factors: Implications for MAIA Aerosol Retrievals. **Presented at: 4th Advancement of POLarimetric Observations: instruments, calibration, and improved aerosol and cloud retrievals (APOLO) Conference**; November 2024; Kyoto, Japan.

Taozhong Huang, Feng Xu, Jens Redemann, Lan Gao, Meng Gao, and Oleg Dubovik. A Correlation-based Inversion Method for Aerosol Remote Sensing and Test Using AERONET and POLDER Measurements. **Presented at: American Geophysical Union (AGU) 2022**; December 2022; Chicago, IL, USA.

Taozhong Huang, Benting Chen, Lan Gao, Connor J. Flynn, Oleg Dubovik and Feng Xu. A Correlation-based Inversion Method for Aerosol Property (CIMAP) Retrieval: A Case Study Using AERONET and POLDER Measurements. **Presented at: American Geophysical Union (AGU) 2021**; December 2021; New Orleans, LA, USA.

Feng Xu, **Taozhong Huang**, Meng Gao, Kirk Knobelspiesse, David Diner, Lan Gao, Connor Flynn, Jens Redemann, Oleg Dubovik. Aerosol Remote Sensing Inversion: Improvement of Retrieval Efficiency. **Presented at: AGU Fall Meeting 2021**, December 2021, New Orleans, LA, USA.

Yixin Wen, John Krause, Brian Kahn, and **Taozhong Huang**. The use of AIRS and polarimetric radar to monitor severe local storms: creating synergy between hyperspectral sounders and ground radars. **Presented at: NASA**

Sounder Science Team Meeting 2019, September 2019, Hyattsville, MD, USA.