UNIVERSITY OF OKLAHOMA

GRADUATE COLLEGE

INFORMATION CONTENT ANALYSIS OF COMBINED LIDAR-POLARIMETER RETRIEVALS TO IMPROVE AEROSOL REMOTE SENSING ACCURACY

A THESIS

SUBMITTED TO THE GRADUATE FACULTY

in partial fulfillment of the requirements for the

Degree of

MASTER OF SCIENCE IN METEOROLOGY

By

DAVID P. NOWICKI
Norman, Oklahoma
2021

INFORMATION CONTENT ANALYSIS OF COMBINED LIDAR-POLARIMETER

RETRIEVALS TO IMPROVE AEROSOL REMOTE SENSING ACCURACY

A THESIS APPROVED FOR THE
SCHOOL OF METEOROLOGY

BY THE COMMITTEE CONSISTING OF

Dr. Feng Xu, Chair

Dr. Greg McFarquhar

Dr. Jens Redemann

Acknowledgements

I would first like to thank my primary advisor Dr. Feng Xu for taking a chance by hiring me as his first master's student and having faith in my ability to learn and grow in a field in which I had no previous experience. The support and guidance offered by Dr. Xu and the entire Atmospheric Radiation and Remote Sensing group throughout my master's progress was indispensable for my success in completing this project and staying motivated throughout a painful pandemic and several difficult personal experiences. I feel I've found a research family more than a research group, and I would like to specifically thank the other members of ARRS — Dr. Lan Gao, Dr. Connor Flynn and Joe Huang — for all of their constructive advice in research and their caring words and support in all things outside research. I would also like to thank the School of Meteorology and NASA for their joint support in funding this research. Lastly within SoM, I would like to thank Shawn Riley for continued patience and technical support as well as Christie Upchurch for keeping my spirits up and helping me attack imposter syndrome when I wanted to give up on this whole process.

As for those that came before OU, I would be remiss not to thank my parents for continuing to trust me and believe that I can do what I set my mind to. I'd like to thank Grandmama for instilling a love of math and learning in me from a young age, as well as Grandma (Ann Washington) for raising me, loving me like one of her own grandchildren, and being my biggest supporter and cheerleader throughout all things, personal and professional. To my brothers, thank you for teaching me how to be myself confidently and love all parts of myself. You set the best examples for how to live fully and happily, and I look forward to joining you on many more adventures in life.

Throughout my life, several educators have been instrumental in molding me into the student and person I am today. My cello instructors Nancy Bateman and Sue Gaston, middle school teachers Suzanne Agent and JoJo Herring, tuba instructors Sonny Hill and Chuck Reynolds, and high school teachers Kathrina Pound, Brett Mayfield, Auston Edwards, Susan Shivers, Judy Jenkins, Jimmy Ainsworth and Kristi Ainsworth. During my time at the University of Mississippi, several professors took the time to invest in me personally, as well as academically. Specifically, I'd like to thank all of the following professors for their advice and support, without which I never would have been able to attend graduate school anywhere: Drs.

Michael Thomas Allen, Carrie Smith, Stefan Schulenberg, and Melinda Redding (psychology); Drs. Jeremy Clark, Micah Milinovich, Gerard Buskes, Sandra Spiroff, Erwin Miña-Díaz, and Sam Lisi (mathematics); Drs. Selim Giray and Thomas Peattie (music); and Dr. Tim Nordstrom (political science). Thank you to Alan Gerard and the whole Gerard family for introducing me to meteorology and being like family to me for many years. A thank you to Dr. Daphne LaDue and my entire REU family for helping me get my feet wet in meteorology research and showing me how fun, varied, and interesting research in this field can be. Moreover, I would like to thank my research advisors and research partner during my REU internship: John Hart, Ryan Bunker, and Drs. Ariel Cohen and Kim Klockow-McClain.

Lastly, I would like to thank the friends without whom I wouldn't be here today. For helping me survive undergrad, I acknowledge Elizabeth McDowell, Mackenzie Greene, and Starkey Morgan. From SoM, my roommate Tyler Young and my beautiful children Chadwick and Harriet for their patience and emotional support throughout my master's degree progress. As a whole, thank you to all the friends I've made within SoM. I have you to thank for helping me through classes, research struggles, and personal crises. More importantly, you've become a real family to me that has made me feel at home for the first time since childhood, which has made my time here more enjoyable and fulfilling than work or school could be anywhere else.

Table of Contents

Abstract

Aerosols are small particles suspended in the atmosphere which affect global climate change via scattering and absorbing sunlight and serving as cloud condensation nuclei. Moreover, the prevalence of some anthropogenic and toxic aerosols has a direct adverse impact on public health. At present, aerosols remain the greatest source of uncertainty for radiative forcing. Though accurate characterization of spatially and temporally varying aerosols is difficult, two powerful remote sensing techniques, polarimetric and lidar measurements, have the complementary strengths of resolving column-effective and vertical distributions of aerosol properties, respectively. Vigorous development of these instruments has sparked our interest in 1) their collocation on ground-based, airborne, and space-borne platforms and 2) developing more efficient, versatile forward and inverse models that make use of combined lidar and polarimetric data to enhance aerosol remote sensing accuracy. In our effort to address target 2, the information content analysis (ICA) method is adopted in this study. It uses *a priori* information of speciated aerosol characteristics and various assumptions of measurement uncertainties as input. Combined with the use of a light scattering model and a radiative transfer model, lidar and polarimetric signals are simulated. The retrieval uncertainties of a suite of aerosol-related geophysical variables (GVs) are output by ICA, which allows us to explore the dependence of GV retrieval capabilities on *a priori* knowledge of aerosol properties and instrument accuracy. We analyze results to 1) check whether aerosol GV uncertainty thresholds established by NASA's Aerosol and Cloud, Convection and Precipitation (ACCP) mission are met by different lidar and polarimetric instrument configurations and 2) identify a cost-effective, yet accurate, instrument combination for future missions. Five different methods are used to perturb *a priori* knowledge of state vector and measurement uncertainties across seven different lidar and/or polarimeter instrument

configurations. Calculated GV uncertainties output by our model are compared to threshold GV retrieval uncertainties established by ACCP's Science and Applications Traceability Matrix (SATM). Results show clear impact of *a priori* knowledge on retrieval capabilities as well as the importance of specific measurements to accurate retrievals of the GVs considered. Across results, the 2β 2α $1\beta_{\text{attn}}$ lidar and polarimeter (L6+POL) instrument combination consistently showed superior abilities of retrieving the GVs considered within SATM-prescribed uncertainty ranges.

Chapter 1. Introduction

In practice, remote sensing of aerosol properties can be difficult. As the technological capabilities of the instrumentation typically used in remote sensing grow, there is a continuous need to effectively use this instrumentation to its maximal effect. Though this is the case in many research fields, rapid improvement of both the efficiency and efficacy of aerosol remote sensing tools — a term encompassing both physical instruments and modeling algorithms — is of paramount importance not only to science but also to the general public. The National Academies of Sciences, Engineering and Medicine (NASEM) 2017 Decadal Survey, which is used to set forth long-term science objectives that benefit society, established the ACCP initiative, through which several large-scale objectives were set forth as a call to researchers across several related subfields of the atmospheric sciences. There are two main outcomes of improving our knowledge of aerosol content and atmospheric interactions among aerosols, clouds, and the ambient environment. First, better understanding of the relative abundances of aerosols that can be harmful to humans will lead to informed suggestions to governments, companies, etc., of how to improve of air quality, particularly in highly industrial areas. From a climate perspective, a better understanding of aerosol type and distribution — as well as how aerosols interact with clouds — help scientists quantify the amount and sign of radiative forcing (i.e., negative for net cooling and positive for net heating) that different aerosols produce. Better representation of aerosol radiative forcing in climatological models is of paramount importance to improving their predictive capabilities (NASA 2017).

During this period of history, communities around the world — particularly those in highly urbanized areas — are seeing declines in air quality that result in increased rates of respiratory illnesses, exacerbation of pre-existing health conditions and, far too often, death (Bowe et al. 2019; Brauer et al. 2019; Shiraiwa et al. 2017). Though several studies show evidence of natural causes

and exacerbators of high aerosol concentration (e.g., dust storms, volcanic eruptions), anthropogenic activities such as industrial combustion, mass use of fossil fuel-burning vehicles, and biomass burning also contribute to air quality decline across various parts of the world (Karagulian et al. 2015; Viana et al. 2008; Yang et al. 2017). Additionally, numerous studies during recent years have shown that aerosol contributions to climate change are subject to high uncertainties, with various sources disagreeing on not only the magnitude of different aerosols' contributions but also how these contributions will impact the climate, moving forward (Andreae et al. 2005; Intergovernmental Panel on Climate 2013). Though a full discussion of the intricacies surrounding this very active research area are beyond the scope of this research, rapid improvement and implementation of tools used to study the phenomenon of aerosol radiative forcing are needed if we are to better understand and prepare for the ways in which the climate is changing. Because of their key role in reducing air quality and increasing the rate of climate change, aerosols in the atmosphere must be analyzed and better understood in terms of prevalence, location and species before meaningful changes can be made to climatological models, public health directives, and the policies that govern and fund both efforts.

In an effort to meet these two broadly defined goals, the ACCP mission has set forth a number of standards that establish concrete criteria to define what level of accuracy must be reached in the aerosol remote sensing of various geophysical variables (GVs) to constitute progress that meaningfully advances the capabilities of the field (see Section 4.1 for more detail). These GVs are compiled in the Science and Applications Traceability Matrix (SATM), where the general descriptions of GVs of particular interest are given, along with the acceptable retrieval errors for each (NASA 2019). Any research moving forward as part of this broader initiative must first ensure that retrieval of these GVs is possible within the SATM-established margins of error for each GV.

Those GVs considered in this study are listed, along with short descriptions as well as their SATM requirements, in Table 1.

Table 1: Geophysical variables, their definitions, and their target uncertainties, as defined by ACCP's Science and Applications Traceability Matrix (NASA 2019).

GV	GV description	GV uncertainty in SATM*
SSA, column	Aerosol single scattering albedo - as placeholder for Aerosol Absorption Optical Depth (Column and BL)	±0.04
SSA, BL		
g_{eff}, column	Aerosol Asymmetry Parameter	±0.02
LR, column	Aerosol Extinction to Backscatter Ratio (Column and BL)	±25%
LR, BL		
τ_{fine}, column	Aerosol Fine Mode Optical Depth (Column and BL)	$\pm0.02\pm0.05\tau_{fine}$
τ_{fine}, BL		
τ, column	Aerosol Optical Depth (Column and BL)	$\pm0.02\pm0.05\tau$
τ, BL		
m_r, column	Aerosol Real Index of Refraction (Column and BL)	±0.025
m_r, BL		
α, profile	Aerosol Extinction (Profile)	Max of (0.02 km^{-1}, 20%)

*Uncertainties reported are for 50×50 km horizontal resolution and 250 m vertical resolution.

Viewing Table 1, one can see that there are several GVs for which retrieval capabilities must still be improved. The GVs considered in this study were chosen because reducing their retrieval uncertainties is a critically important first step to developing more advanced retrieval algorithms and eventually fabricating new instrumentation that can be used to address the ACCP mission's science goals. Additionally, polarimetric and lidar observations are well-suited to retrieve the environmental information necessary for the calculation of the GVs considered in this study. As mentioned previously, the primary underlying barrier to the improvement of GV retrieval

capabilities is the complexity of aerosols. Atmospheric aerosols not only vary greatly in terms of their intrinsic optical properties (e.g., complex refractive index) and microphysical properties (e.g., size and shape) but also in their spatial and temporal distributions (Dubovik et al. 2019). The sheer number of properties that must be considered to differentiate aerosol species make real-time monitoring of aerosol loading in the atmosphere quite difficult, especially on a large scale. Much money and time has been invested into the development of instruments that have more complete retrievals due to multi-angle and/or multi-spectral capabilities, and researchers often must balance the cost of instrumentation and experimentation with the benefit of additional data (Eloranta 2005; Hansen et al. 1995).

Though several missions utilizing polarimeters have been sent into orbit [e.g., see (Deschamps et al. 1994; Tanré et al. 2011)] and some *in situ* studies using aircraft have been performed, even recently (Knobelspiesse et al. 2020), there is often considerable monetary expense to these experiments. Additionally, it is difficult — even for those experiments considered successful in terms of producing useful data — for many operational instruments to reach the hoped-for retrieval accuracies obtained theoretically (Dubovik et al. 2019). There is need, therefore, to utilize more cost-effective ways to test retrieval accuracies of various instrument combinations prior to their design, fabrication, and eventual launch. Sensitivity studies like the one performed in this paper are one example of a way in which many different potential instrument combinations and retrieval techniques may be tested prior to instrument fabrication, such that cost of instrument development may be minimized from the outset and the specifications of the instrument may be calibrated to fit specific science goals (Mishchenko et al. 2007). To accomplish the ACCP-defined goals described earlier, this sensitivity study analyzes several combinations of two different kinds of instrument: polarimeter and lidar.

As regards this sensitivity study, the primary focus is to contribute to ACCP's broader mission by developing a tool that uses information content analysis (ICA) of combined polarimeter and lidar measurements to improve the accuracy of retrievals of aerosol properties in both the boundary layer as well as the entire vertical column of the atmosphere. This method is used to determine GV retrieval uncertainties given *a priori* knowledge of the environment contained within a state vector (see Section 3.1.1) and observational configurations (along with their associated errors, described in detail in Section 4.1). By altering these input data, one can test the impacts of *a priori* constraints on retrieved GV uncertainties. Because of the relatively simple design of this method, ICA has several strengths that make it a desirable and useful method, particularly in early stages of algorithm development. For example, ICA can be used — as it is in this study — to easily compare the relative abilities of different observational configurations. Additionally, this method allows for efficient calculation of a large number of different aerosol loading scenarios, which should prove useful in future research using this particular algorithm, particularly as regards improving the generalizability of our model.

Though the specifics of the exact polarimeter and lidar systems tested will be described in greater detail in later sections, the benefits of using each type of instrument must first be established. For the last several decades, numerous studies have shown the potential for use of polarimetric data in remote sensing of aerosols. Many polarimeters that have been deployed for operational use are passive sensors (i.e., they receive reflected light from the sun) that measure three (I, Q, U) of the four parameters contained within the Stokes' vector (Deschamps et al. 1994; Deuzé et al. 2001). Note that the fourth Stokes' parameter — V, which describes circular polarization — is typically ignored because of its relatively small magnitude, compared to those of the linear polarization Stokes' parameters Q and U (Coulson 1988). These three parameters

contain information regarding the intensity and depolarization of the incident light after a single scattering event — an improvement over previous methods that used intensity-only retrievals (Hasekamp and Landgraf 2005). It is important to note here that "intensity" will be used throughout the remainder of this paper as a shorthand for the physical quantity *intensity of polarized radiance*. Single-view polarimeters are generally able to retrieve well the spectral AOD, effective radius, and refractive index of different aerosol species (Hasekamp and Landgraf 2007; Mishchenko and Travis 1997). However, by using different view angles, more parameters that are species-specific (e.g., complex refractive index) can be retrieved to a higher accuracy due to the increase in available information content (Chen et al. 2017; Russell et al. 2014). Additionally, the use of multiple angles increases already present sensitivities to the shape and vertical distributions of aerosols, which can further help in increasing the accuracy of aerosol characterization (Dubovik et al. 2019).

To supplement polarimetric data, this sensitivity study also employs the use of lidar for aerosol characterization. Lidar (short for Light Detection and Ranging) instruments are active sensors (i.e., they both transmit and receive signals) that are useful for aerosol remote sensing because of the optical wavelengths (typically visible or near-visible wavelengths) at which they operate, where the sensitivities of these instruments to particulates in the atmosphere are strong (Collis and Russell 1976). The benefits of using lidar, which typically measures backscatter and/or extinction coefficients, include high spatial resolution, though it should be noted that increased spatial resolution comes with the trade-off of having smaller spatial coverage (Comerón et al. 2017). The greatest strength of lidar compared to other remote sensing methods (e.g., spectroscopy of emitted radiation) is its vertically resolved measurements that allow for better understanding of vertical distributions of aerosols as well as the relative locations of aerosol species to each other

and clouds (Zuev and Naats 2013). By adding spectral variability to lidar instruments (i.e., adding more wavelengths of transmitted light), researchers have been able to expand the capabilities of lidar over the past few decades to better discern optical properties (e.g., lidar ratio) that can be useful for aerosol characterization (Müller et al. 2000). Moreover, advanced lidar systems can use the additional information content provided by multi-spectral measurement to discern more microphysical properties (e.g., particle concentration, effective radius) than could be retrieved by earlier systems (Ansmann et al. 2012).

Because of the complementary nature of polarimetric and lidar measurements — the ability to discern vertical profiles of aerosol loading and species-specific aerosol properties, respectively — a combined retrieval algorithm that makes use of both types of instrument should improve model sensitivity to aerosol characteristics (Burton et al. 2016). So, to address the two broad goals of improving the ability to remote sense the optical and microphysical properties of aerosols — that is, to reduce the adverse public health impacts caused by the increasing presence of aerosols and to improve the ability to accurately represent aerosol uncertainties within climatological models — we set forth two hypotheses that align with the two primary goals of performing information content analysis in this study:

H1: To address Goal 1, which is finding out whether the GVs considered in this study may be adequately constrained by a combined lidar-polarimeter approach, we hypothesize that the instrument combination that is able to retrieve the most information content from its measurements will be able to constrain all of the GVs within their SATM threshold uncertainty requirements.

H2: To address Goal 2, which is to find (if possible) a cost-effective recommendation for which instrument combination to use, we hypothesize that there will be one instrument combination that will consistently outperform other configurations.

Chapter 2. Toy model of information content analysis

2.1 Error propagation method

Before directly performing ICA of combined retrieval methods, it was necessary to understand the concept of error propagation. To know how error will propagate through calculations of these measurements, one must establish *a priori* ranges of expected values of the observed variables and account for the error inherent in the measuring of these dimensions. By using Eq. (2.1), which is often used in error propagation studies, it is possible to analyze how changes in *a priori* knowledge of variables to be retrieved or changes in measurement error compound to affect the overall uncertainty of retrieved variables (or that of quantities derived from said variables).

$$\mathbf{S}_r{}^{-1} = \mathbf{J}^T \mathbf{S}_\varepsilon^{-1} \mathbf{J} + \mathbf{S}_a^{-1} \tag{2.1}$$

The boldface type in Eq. (2.1) indicates that the quantity is a matrix, while ^{T} and $^{-1}$ represent the transpose and inverse matrix operations, respectively. In this equation, $\mathbf{S}_a$ represents the *a priori* error covariance matrix, a square matrix with dimensions of ($n \times n$), where n is the number of variables observed. The diagonal of this matrix gives the variances correlated with *a priori* uncertainties of each variable to be observed. The off-diagonal terms in this matrix represent crosstalk uncertainties between variables, which are assumed to be zero. The $\mathbf{S}_\varepsilon$ term represents the observation (measurement) error covariance matrix, which is also square with dimensions ($m \times m$), where m is the number of measurements made for each observation. The resultant retrieved error covariance matrix, given here by $\mathbf{S}_r$, must be square of ($n \times n$) dimension. Its diagonal gives

the 1-σ uncertainty of retrieved parameters. The Jacobian matrix **J** shows how sensitive a forward model is to changes in the parameters to be retrieved by that model. It is necessarily of dimension ($m \times n$), and its components are calculated by the following formula,

$$J_{mn} = \frac{\partial F_m(\boldsymbol{x})}{\partial x_n} \approx \frac{F_m(\boldsymbol{x}') - F_m(\boldsymbol{x})}{x'_n - x_n}, \tag{2.2}$$

where the partial derivative of some measurement $\mathbf{F}(\mathbf{x})$ for a set of parameters within a state vector $\mathbf{x}$ is computed for each observation m and each parameter n (Knobelspiesse et al. 2012; Li et al. 2018). Small perturbations to the state vector (represented here by $\boldsymbol{x}'$) can be made to approximate the Jacobian at different levels of uncertainty for each parameter and observation.

In practice, the Jacobian must be recalculated repeatedly for small, incremental perturbations so that the linearity assumption of this calculation will remain valid over a larger domain — a method known as "finite differencing" (Knobelspiesse et al. 2012). This calculation is relatively easy for the table toy model case (and may even be found analytically), but when the number of variables or observations becomes large, as is the case when considering the radiative transfer model used in this study, calculation of the Jacobian becomes more complex.

2.2 Description of toy model

Prior to conducting ICA of the radiative transfer model as it relates to the combined instrument retrieval considered later in this paper, a toy model was designed to increase the readers' understanding of the link between the mathematics behind error propagation and the physical meaning of matrix components within Eq. (2.1). To this end, a simple experiment was performed to show how error propagation works in practice by using the example of measuring a table. First, we defined two parameters, length (L) and width (W) of the table, in meters. Also defined were three observational measurements that can be calculated from these parameters: area of the tabletop (A), perimeter of the tabletop (P) and parameter Q, defined as the ratio of the table's

dimensions. In terms of the toy model input variables, the formulae for the three measurements considered are as follows:

$$A = LW \tag{2.1}$$

$$P = 2(L + W) \tag{2.4}$$

$$Q = \frac{L}{W} \tag{2.5}$$

2.3 Calculations of component matrices

To determine components of $\mathbf{S}_a$, some *a priori* error and truth values must be assumed for both L and W. For this experiment, L = 3 meters and W = 2 meters were chosen as the true length values, and the relative error (ε) for both variables was set at a value of 3%. Multiplying each truth value with its associated error gives half the range of values one can reasonably expect a measurement to lie within. By adding and subtracting this value from the truth, the full range of possible values a given true value may equal in reality is given. For example, the length value may be calculated by finding L_{max} and L_{min} using the following:

$$L_{max} = L_{true} + L_{true}\varepsilon_L \tag{2.6}$$

$$L_{min} = L_{true} - L_{true}\varepsilon_L \tag{2.7}$$

The same method (this time, using ε_W) is used to calculate W_{max} and W_{min} values. Now, $\mathbf{S}_a$ may be written using values from Eqs. (2.6) and (2.7) as well as width calculations as:

$$\boldsymbol{S}_a = \begin{bmatrix} \left(\frac{L_{max} - L_{min}}{2}\right)^2 & 0 \\ 0 & \left(\frac{W_{max} - W_{min}}{2}\right)^2 \end{bmatrix}. \tag{2.8}$$

Here, off-diagonal terms are zero because we assume there is no correlation between the errors inherent in either variable's expected value. There is a one-half multiplicative factor that must be

taken into account after finding the whole range (the maximum/minimum difference) because the resultant maximum/minimum terms both have an error factor. By squaring the resultant quantity, the full variance of each component is found by squaring each scaled difference term.

Now that initial expected error of variables within the state vector has been accounted for, the other component matrices of Eq. (2.1) must be found. To find the components of the measurement error covariance matrix ($\mathbf{S}_\varepsilon$), individual measurements (y) are multiplied by their inherent errors (ε_y) to get the window of uncertainty (Δy) within which values may vary and still be considered truth, as follows:

$$y\varepsilon_y = \Delta y\,. \tag{2.9}$$

Once percentage errors have been turned into measurement error via Eq. (2.9), measurement variances are placed along the diagonal of the square matrix $\mathbf{S}_\varepsilon$. Each measurement is considered independent of all others, so crosstalk between measurement error is assumed to be negligible. Thus, $\mathbf{S}_\varepsilon$ is, for the toy case, a diagonal matrix of the form

$$\boldsymbol{S}_\varepsilon = \begin{bmatrix} \frac{1}{\Delta A^2} & 0 & 0 \\ 0 & \frac{1}{\Delta P^2} & 0 \\ 0 & 0 & \frac{1}{\Delta Q^2} \end{bmatrix}. \tag{2.10}$$

Once the Jacobian matrix is calculated as in Eq. (2.2), the toy model is able to output uncertainties for each parameter, based on error propagation of the *a priori* and measurement errors. If one adds a multiplicative scale factor to the error terms and runs this code iteratively, it is possible to watch as the resultant error in a given state variable grows larger before eventually saturating. That is, at some point, the amount of uncertainty in any variable will reach a maximum. For this toy model, the objective was to gain an understanding of the basic inputs to the error propagation method used, as well as to show the ideal behavior of error saturation curves when

either measurement errors or *a priori* error are well-characterized. By perturbing any one measurement or variable's initial uncertainty and running the code iteratively over many perturbations, one can see to what extent having knowledge about a specific measurement or variable's true value beforehand is necessary for constraining the uncertainties of other variables. This process will be expounded upon more within the methods section.

Chapter 3. Information content analysis for combined lidar-polarimeter remote sensing

Now that the simple case of the table toy model has been described, description of the inputs used in the ICA algorithm may proceed. Though this forward modeling process is more complicated than that described in Chapter 2, the terms involved are analogous. For example, the table length and width parameters are analogous to the GVs considered in the state vector (described later in this chapter). Likewise, measurements involving parameters (e.g., area, in the toy model) will also be taken into account in the ICA model. The only difference is that these measurements will come from the lidar and polarimetric instruments involved in retrieval. It follows that the Jacobian calculation, which is shown in its general form in Eq. (2.2), will also apply to the ICA model. In this study, the set of m observations will be the combined number of lidar and polarimeter measurements, as opposed to those given in Eqs. (2.3)-(2.5), and the n parameters within the state vector $\mathbf{x}$ will consist of speciated aerosol and surface properties. It is important to note a few key limitations of ICA, as it is used in this study. To use this analysis method, we must assume that our successful retrievals converge to a true solution. Additionally, we utilize a linearized model that results in more efficient calculations of iterative perturbations to the Jacobian matrix, at the cost of simplifying the model to the extent that some error (in relation to a more realistic setup) is inherently added to results from its use.

3.1 Forward model inputs

The inputs to the model include a) the preset truth of GVs and surface conditions contained in the state vector **x** and b) simulated polarimeter information in combination with different lidar signals. Scattering due to spherical particles is calculated using Mie theory (van de Hulst, 1981). However, calculating the Mie scattering properties that must be used as inputs to RT model and simulations of lidar signal is computationally inefficient. To avoid computing the Mie scattering properties within the algorithm itself, aerosol phase matrix elements $P_{jk,\mathrm{a}}^{(\mathrm{s})}$ and extinction [scattering] coefficients $K_{\mathrm{ext[sca],\,a}}^{(s)}$ for N_{bin} size bins and refractive indices are precalculated and saved in look up tables [LUTs, (Xu et al. 2021)]. The LUTs themselves are three-dimensional and, therefore, contain values calculated from all the possible bin combinations of three different state vector parameters: $m_r^{(s)}$, $m_i^{(s)}$, and $r_{\mathrm{m}}^{(s)}$. Specifics regarding bin size, number, etc. for these parameters are listed in Table 2.

Additionally, throughout the remainder of the paper — particularly in Sections 3.1.2 and 3.1.3 — there will be references to "truth values" of the GVs whose uncertainties are the primary focus of this study. Some of these GVs are members of the state vector (see Section 3.1.1), whereas many are calculated from single-scattering aerosol properties contained within the state vector (see Sections 3.1.2 and 3.1.3). The GVs considered in this study were listed earlier in Table 1, along with brief definitions of each and the methods by which their uncertainties are calculated, according to the current SATM requirements.

Table 2: Parameters used for calculating the Mie scattering properties saved in our LUTs are shown, along with their minimum and maximum range values, number of equally spaced bins used to span these ranges, and method of spacing between values within these ranges.

Parameter	Space	# of Bins	Range (min,max)
Refractive index, real (m_r)	Linear	22	(1.3, 1.7)
Refractive index, imaginary (m_i)	Linear in logarithmic space	5	(1×10^{-8}, 3×10^{-4})
		15	(5×10^{-4}, 0.5)
Mean radius (r_m)	Linear in logarithmic space	50	(5 [nm], 50 [μm])

3.1.1 State vector

Variables under sensitivity analysis are included in a state vector, which includes environmental information (i.e., surface reflectance and wind speed characteristics) as well as speciated aerosol properties pertaining to particle size, number, composition and distribution, which must be altered for each canonical case to change the aerosol loading for each specific case considered. Note that throughout this paper, "knowledge of the environment" will refer specifically to how well-characterized the *a priori* state vector is. Additionally, the uses of "canonical case" and "Case 61" used throughout are used for internal consistency with ongoing research within our research group. The use of "Case 61" specifically refers to the case considered for this thesis and is defined as such by the specific aerosol loading (type and location) depicted in Table 3.

The algorithm used is designed to retrieve multiple (N_s) aerosol species at once by using different combinations of lidar and polarimetric measurements, each of which provides different information content. This information is dependent on the specific lidar chosen, with regards to

the inclusion of extinction and backscatter/attenuated backscatter, the number of wavelengths included, and the vertical resolution. Additionally, information content from polarimetric data can be altered by the inclusion of different view angles and/or spectral channels. To handle multiple aerosol species at once, the state vector **x** includes the retrieval parameters for N_s species as

$$\mathbf{x} = \left[\mathbf{x}^{(1)};\ \mathbf{x}^{(2)};\ \cdots;\ \mathbf{x}^{(N_s)};\ \mathbf{x}^{(\text{ocean})}\right], \tag{3.1}$$

where N is the total number of retrieved properties and vectors $\mathbf{x}^{(s)}$ contain $N_{\text{aer}} = 6$ speciated aerosol properties for the s^{th} aerosol species, which can be expressed as

$$\mathbf{x}^{(s)} = \left[m_r^{(s)};\ m_i^{(s)};\ r_{\text{m}}^{(s)};\ \sigma_{\text{m}}^{(s)};\ C_{\text{v}}^{(s)};\ f_{\text{v}}^{(s)}\right]_{N_{\text{aer}}\times 1}. \tag{3.2}$$

For the s^{th} species in Eq. (3.2), m_{r} and m_{i} are layer-resolved real and imaginary parts of aerosol refractive index, respectively; r_{m} and σ_{m} contain layer-resolved geometric volume mean radius and logarithmic standard deviation of size distribution, respectively; C_{v} is the total column aerosol concentration (m^3/m^3); and f_{v} contains layer-resolved aerosol volume fractions (Xu et al. 2021). The remaining component of the state vector, given in Eq. (3.1) as $\mathbf{x}^{(\text{ocean})}$, includes environmental parameters that contribute to ocean surface reflectance, such as wind speed and chlorophyll-a concentration, which is included to account for the impact of biological materials within the ocean on radiance measurements returning from the surface. Additionally, empirically derived Lambertian surface reflectance terms are included as the last elements of $\mathbf{x}^{(\text{ocean})}$.

3.1.2 Aerosol single scattering properties for radiative transfer calculation

Having introduced the state vector **x** and its speciated components, we will now describe how those properties are calculated to get layer-resolved as well as column-effective values for use in the calculation of different species' single scattering quantities.

Before making calculations, we assume a lognormal volumetric size distribution $\frac{\mathrm{dv}_l^{(s)}(r)}{\mathrm{d}\ln r}$ for each of N_s different aerosol species (see Section 4.1). This quantity is calculated using two elements of **x**, the geometric volume mean radius $r_m^{(s)}$ and logarithmic standard deviation $\sigma_{m,l}^{(s)}$, for the s^{th} species in the l^{th} atmospheric layer according to the following:

$$\frac{\mathrm{dv}_l^{(s)}(r)}{\mathrm{d}\ln r} = \frac{1}{\sqrt{2\pi}\sigma_{\mathrm{m},l}^{(s)}} \exp\left[-\frac{(\ln r - \ln r_{\mathrm{m},l}^{(s)})^2}{2(\sigma_{\mathrm{m},l}^{(s)})^2} \right]. \tag{3.3}$$

Additionally, the speciated size distribution in Eq. (3.3) may be combined with speciated, layer-resolved values of volume concentration ($\mathrm{C}_{\mathrm{v},l}^{(s)}$) to get sub-column values of volume concentration for N_s species over l atmospheric layers by using

$$\frac{\mathrm{dV}_l(r)}{\mathrm{d}\ln r} = \sum_{s=1}^{N_s} \mathrm{C}_{\mathrm{v},l}^{(s)} \frac{\mathrm{dv}_l^{(s)}(r)}{\mathrm{d}\ln r}. \tag{3.4}$$

The speciated, layer-resolved volume concentration term in Eq. (3.4) is, itself, found for each species at the l^{th} layer by taking the product of the species' fractional volume at that level ($f_{\mathrm{v},l}^{(s)}$) and its total column volumetric concentration ($\mathrm{C}_{\mathrm{v}}^{\mathrm{tot},(s)}$), both of which are elements of **x**.

To calculate single scattering properties for use in our RT model, several of which are among the target GVs (e.g., AOD, SSA), it is necessary to account for differences in the sizes of aerosols whose properties are used in retrieval by defining N_{bin} number of size bins. Additionally, the algorithm assumes that the volume fraction of particles in the l^{th} layer ($f_{\mathrm{v},l}^{(s)}$) consists only of spherical aerosol particles. Once these steps are taken, the extinction[scattering] coefficients for particles in the i^{th} size bin ($K_{\mathrm{ext[sca],a},i}^{(s)}$) may be used to calculate volumetric extinction coefficient and volumetric scattering coefficient via

$$K_{\mathrm{ext,a},l}^{(s)} = \sum_{i=1}^{N_{\mathrm{bin}}} f_{\mathrm{v},i,l}^{(s)} K_{\mathrm{ext,a},i}^{(s)} \tag{3.5}$$

$$K_{\mathrm{sca,a},l}^{(s)} = \sum_{i=1}^{N_{\mathrm{bin}}} f_{\mathrm{v},i,l}^{(s)} K_{\mathrm{sca,a},i}^{(s)}\,, \tag{3.6}$$

where $f_{\mathrm{v},i,l}^{(s)}$ is the speciated volumetric fraction within the i^{th} size bin and is, itself, calculated using the central radius of the respective size bin (r_i) as well as the log-normal volumetric size distribution [from Eq. (3.3)] via

$$f_{\mathrm{v},i,l}^{(s)} = \frac{\mathrm{dv}_l^{(s)}(r_i)}{\mathrm{d}\ln r_i}. \tag{3.7}$$

Using the same binning procedure, individual elements ($P_{jk,\mathrm{a},l}$) of the aerosol phase matrix $\mathbf{P}_{\mathrm{a},l}$ are calculated using

$$P_{jk,\mathrm{a},l}^{(s)} = \frac{\sum_{i=1}^{N_{\mathrm{bin}}} f_{\mathrm{v},i,l}^{(s)} K_{\mathrm{sca,a},i}^{(s)} P_{jk,i}^{(s)}}{\sum_{i=1}^{N_{\mathrm{bin}}} f_{\mathrm{v},i,l}^{(s)} K_{\mathrm{sca,a},i}^{(s)}} \tag{3.8}$$

Now that aerosol particle size differences have been accounted for, layer-effective values of total extinction coefficient may be calculated by summing the contributions of all N_s species:

$$K_{\mathrm{ext,\,a},l} = \sum_{s=1}^{N_s} K_{\mathrm{ext,\,a},l}^{(s)}\ . \tag{3.9}$$

In a similar manner, the total SSA contributed by all species within a layer may be calculated by combining speciated extinction coefficient and speciated scattering coefficient values — from Eq. (3.5) and Eq. (3.6), respectively — with speciated volume fraction within the level ($f_{\mathrm{v},l}^{(s)}$) via

$$\omega_{0,\mathrm{a},\,l} = \frac{\sum_{s=1}^{N_s} f_{\mathrm{v},\,l}^{(s)} K_{\mathrm{sca,\,a},\,l}^{(s)}}{\sum_{s=1}^{N_s} f_{\mathrm{v},\,l}^{(s)} K_{\mathrm{ext,\,a},\,l}^{(s)}}. \tag{3.10}$$

Similarly, layer-effective values for elements of the aerosol phase matrix are given by

$$P_{jk,\mathrm{a},\,l} = \frac{\sum_{s=1}^{N_s} f_{\mathrm{v},\,l}^{(s)} K_{\mathrm{sca,\,a},\,l}^{(s)} P_{jk,\mathrm{a},\,l}^{(s)}}{\sum_{s=1}^{N_s} f_{\mathrm{v},\,l}^{(s)} K_{\mathrm{sca,\,a},\,l}^{(s)}}. \tag{3.11}$$

To find speciated values of AOD for the l^{th} level, speciated extinction coefficient values calculated using Eq. (3.5) are multiplied by their correlating layer volume concentration values:

$$\Delta\tau_{\mathrm{a},l}^{(s)} = \mathrm{C}_{\mathrm{v},l}^{(s)} K_{\mathrm{ext,a},l}^{(s)}. \tag{3.12}$$

Once again, we must sum the speciated contributions calculated in Eq. (3.12) to find the total AOD of the layer,

$$\Delta\tau_{\mathrm{a},l} = \sum_{s=1}^{N_s} \Delta\tau_{\mathrm{a},l}^{(s)}. \tag{3.13}$$

Moreover, total column[BL] values for AOD must be calculated via

$$\Delta\tau_{\mathrm{a,tot[BL]}} = \sum_{l=1}^{N_{\mathrm{tot[BL]}}} \Delta\tau_{\mathrm{a},l}, \tag{3.14}$$

where $N_{\mathrm{tot}} = 16$ and $N_{\mathrm{BL}} = 8$ are the number of layers considered when calculating total column and BL values, respectively. Now, column-effective[BL] values of SSA and m_{r}, as well as total column and BL values of fine-mode AOD may be calculated. For fine-mode AOD calculation, Eqs. (3.12)-(3.14) are used again. For SSA and m_{r} — as well as for asymmetry parameter, discussed later — AOD weighting of GVs in the total[BL] column was used in the following equations

$$\omega_{0,\mathrm{a,tot[BL]}} = \frac{\sum_{l=1}^{N_{\mathrm{tot[BL]}}} \Delta\tau_{\mathrm{a},l}\omega_{0,\mathrm{a},l}}{\Delta\tau_{\mathrm{a,tot[BL]}}} \tag{3.15}$$

$$m_{r,\mathrm{tot[BL]}} = \frac{\sum_{l=1}^{N_{\mathrm{tot[BL]}}} \Delta\tau_{\mathrm{a},l} m_{r,l}}{\Delta\tau_{\mathrm{a,tot[BL]}}} \tag{3.16}$$

to give truth values of SSA (ω_0) and real part of refractive index for error calculation and comparison. Though this weighting method is not perfect, it gives results that match well the results from previous retrieval simulations as well as real data analysis (Xu et al. 2017; Xu et al. 2016).

Last of the GVs calculated from RT model input is the asymmetry parameter (g), which is defined as the first moment of the scattering phase function $P(\Theta)$ (see Eq. [3.11]) and is calculated for a target volume of particles by

$$g = \frac{1}{2}\int_{-1}^{1} P(\Theta)\cos(\Theta) d(\cos\Theta)\,, \tag{3.17}$$

where the scattering angle is Θ and the scattering phase function is calculated using Mie theory. When considering column-effective asymmetry parameter, one must consider that 1) Mie scattering theory was used in pertinent calculations [e.g., for $P(\Theta)$] for this sensitivity study and 2) logarithmic size distributions of each species were included as input data to the model. Methodologically, these two factors should yield reasonable calculated "truth" values of the asymmetry parameter (Marshall et al. 1995). As with several GVs discussed previously, the truth value of column-effective asymmetry parameter (g_{eff}) is found by AOD weighting values via

$$g_{\mathrm{eff}} = \frac{\sum_{s=1}^{N_s}[g^{(s)} * \sum_{l=1}^{N_{\mathrm{tot}}} \Delta\tau_{\mathrm{a},l}^{(s)}]}{\Delta\tau_{\mathrm{a,tot}}}, \tag{3.18}$$

where $g^{(s)}$ is the speciated asymmetry parameter for each of N_s aerosol species considered (Andrews et al. 2006).

3.1.3 Aerosol single scattering properties for lidar equation

To utilize lidar measurements within the retrieval algorithm, all the single scattering aerosol properties that are used in the lidar equation (e.g., extinction and backscatter), must first be calculated (Liou 2002).

By using the speciated AOD values calculated in Eq. (3.12), one can calculate the speciated extinction due to aerosols $\alpha_{\mathrm{a},l}^{(s)}$ (units, m^{-1}) in a layer of thickness ΔH_l by

$$\alpha_{\mathrm{a},\,l}^{(s)} = \frac{\Delta\tau_{\mathrm{a},\,l}^{(s)}}{\Delta H_l} = \frac{\mathrm{C}_{\mathrm{v},\,l}^{(s)} K_{\mathrm{ext,a},\,l}^{(s)}}{\Delta H_l}, \tag{3.19}$$

and the total extinction coefficient for the l^{th} layer can be calculated by summing the values given by Eq. (3.19) over N_s aerosol species within the layer:

$$\alpha_{\mathrm{a},\,l} = \sum_{s=1}^{N_s} \alpha_{\mathrm{a},\,l}^{(s)}. \tag{3.20}$$

To calculate backscatter, the aerosol scattering coefficient $\beta_{\mathrm{sca},l}^{(s)}$ (units, m^{-1}) must be calculated for each species within a layer of thickness ΔH_l by

$$\beta_{\mathrm{sca},\,l}^{(s)} = \frac{\mathrm{C}_{\mathrm{v},\,l}^{(s)} K_{\mathrm{sca,a},\,l}^{(s)}}{\Delta H_l}, \tag{3.21}$$

from which speciated backscatter values may then be calculated using

$$\beta_{\mathrm{a},\,l}^{(s)} = \beta_{\mathrm{sca},\,l}^{(s)} \frac{P_{11,\,l}^{(s)}(180^\circ)}{4\pi}. \tag{3.22}$$

First noting that $\beta_{\mathrm{a},l}^{(s)}$ has units of $\mathrm{m}^{-1}\mathrm{Sr}^{-1}$ because of normalization over a solid angle in Eq. (3.22), total backscatter (same units) within the l^{th} layer can now be calculated by accounting for all N_s aerosol species within the layer:

$$\beta_{\mathrm{a},\,l} = \sum_{s=1}^{N_s} \beta_{\mathrm{a},\,l}^{(s)} \tag{3.23}$$

To find the lidar ratio (LR) total column[BL] values, the ratio of extinction to backscatter is used within

$$LR_{\mathrm{tot[BL]}} = \frac{\sum_{l=1}^{N_{\mathrm{tot[BL]}}} \alpha_{\mathrm{a},l}}{\sum_{l=1}^{N_{\mathrm{tot[BL]}}} \beta_{\mathrm{a},l}}, \tag{3.24}$$

where extinction and backscatter are calculated via Eq. (3.20) and Eq. (3.23), respectively.

3.2 Error propagation components

This information content analysis via linear-error propagation model is subject to two major assumptions: a) the retrieved solution is representative of the solution space and b) retrieval errors are linearly proportional to the measurement errors. These assumptions can be problematic in situations where the model and/or observation errors are large. Moreover, the observation covariance matrix in our current study assumes no correlations among observation elements so that its off-diagonal elements of the observation error covariance matrix $\mathbf{S}_\varepsilon$ are zero. In practice, however, observation errors can be highly correlated between neighboring spectral bands and/or angular measurements. More accurate characterization is necessary for an accurate information content analysis. In addition, the *a priori* covariance matrix also assumes no error dependency between state vector elements, so the off-diagonal terms of $\mathbf{S}_\mathrm{a}$ are also zero. These are potential error sources of our error estimate model. Without sufficient characterizations of the correlations between instrument errors and interference between cross terms in the *a priori* matrix, we tested several different error assessment methods.

Having addressed important assumptions made when using this ICA tool, we will use the rest of this section to a) describe the application of the error propagation formula discussed in

Chapter 2 to our ICA model and b) describe the recommended methods used to calculate GV uncertainties. Among the various error assessment methods tested before conducting this sensitivity study, those recommended within this section were found to best fit our physical understanding while agreeing well with results from direct combined retrieval simulations being developed in tandem with this tool (Xu et al. 2021).

3.2.1 Calculation of Jacobian matrix

To calculate the total Jacobian **J** as it appears in Eq. (2.1), one can use the general relation shown in Eq. (2.2), applying it to the retrieval model. Because there are multiple observations being considered in our algorithm, one can first break the total Jacobian into its observational components

$$\mathbf{J} = \left[\mathbf{J}_{\mathrm{L}}, \quad \mathbf{J}_{\mathrm{DOLP}}, \quad \mathbf{J}_{\alpha}, \quad \mathbf{J}_{\beta}\right]^{\mathrm{T}}, \tag{3.17}$$

where each Jacobian is computed by taking the partial derivative of each measurement with respect to all components of the state vector **x** in a manner similar to that given in Eq. (2.2). Each of the component Jacobians in Eq. (3.25) may be written in their full matrix forms. For example, the radiance Jacobian $\mathbf{J}_{\mathrm{L}}$ may be expanded as

$$\mathbf{J}_{\mathrm{L}} = \begin{bmatrix} \frac{\partial L_1}{\partial x_1} & \frac{\partial L_1}{\partial x_2} & \cdots & \frac{\partial L_1}{\partial x_N} \\ \frac{\partial L_2}{\partial x_1} & \frac{\partial L_2}{\partial x_2} & \cdots & \frac{\partial L_2}{\partial x_N} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial L_{N_L}}{\partial x_1} & \frac{\partial L_{N_L}}{\partial x_2} & \cdots & \frac{\partial L_{N_L}}{\partial x_N} \end{bmatrix}, \tag{3.18}$$

where N is the total number of elements of **x** and $N_L = N_{\lambda,\mathrm{L}} \times N_{\mathrm{view}}$, where the elements on the right-hand side of the equation represent number of spectral channels and number of view angles used by the polarimeter, respectively. One can also calculate the Jacobian for DOLP by using Eq. (3.26) and replacing "L" with "DOLP."

For the extinction and backscatter coefficient calculations, the Jacobian is slightly more complex in form because of the layer-resolved nature of these measurements. The extinction coefficient Jacobian may be written as

$$\mathbf{J}_{\alpha} = \begin{bmatrix} \mathbf{J}_{\alpha_{l=1}}(m_r^{(s)}) & \mathbf{J}_{\alpha_{l=1}}(m_i^{(s)}) & \cdots & \mathbf{J}_{\alpha_{l=1}}(C_{\mathrm{v}}^{(s)}) & \mathbf{J}_{\alpha_{l=1}}(f_{\mathrm{v},l=1}^{(s)}) & \mathbf{0} & \cdots & \mathbf{0} & \mathbf{0} \\ \mathbf{J}_{\alpha_{l=2}}(m_r^{(s)}) & \mathbf{J}_{\alpha_{l=2}}(m_i^{(s)}) & \cdots & \mathbf{J}_{\alpha_{l=2}}(C_{\mathrm{v}}^{(s)}) & \mathbf{0} & \mathbf{J}_{\alpha_{l=1}}(f_{\mathrm{v},l=2}^{(s)}) & \cdots & \mathbf{0} & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \mathbf{0} & \mathbf{0} \\ \mathbf{J}_{\alpha_{N_{\mathrm{layer}}}}(m_r^{(s)}) & \mathbf{J}_{\alpha_{N_{\mathrm{layer}}}}(m_i^{(s)}) & \cdots & \mathbf{J}_{\alpha_{N_{\mathrm{layer}}}}(C_{\mathrm{v}}^{(s)}) & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{J}_{\alpha_{N_{\mathrm{layer}}}}(f_{\mathrm{v},N_{\mathrm{layer}}}^{(s)}) & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} & \mathbf{0}_{\mathrm{ocean}} \end{bmatrix}, \tag{3.19}$$

where N_{layer} is the total number of vertical layers into which the atmosphere is resolved. For each layer considered, the Jacobian was calculated by taking partial derivatives of the extinction coefficient with respect to each element of **x**, as in Eq. (2.2), at each wavelength λ_x $(1 \leq x \leq N_\alpha)$ used, where N_α is the number of wavelengths used by the lidar in question. Consequently, the format of these layer-effective Jacobians is

$$\mathbf{J}_{\alpha,l}(f_{\mathrm{v}}^{(s)}) = \begin{bmatrix} \dfrac{\partial \alpha_l(\lambda_1)}{\partial f_{\mathrm{v},l}} \\ \dfrac{\partial \alpha_l(\lambda_2)}{\partial f_{\mathrm{v},l}} \\ \vdots \\ \dfrac{\partial \alpha_l(\lambda_{N_\alpha})}{\partial f_{\mathrm{v},l}} \end{bmatrix}. \tag{3.20}$$

To retrieve the Jacobians for backscatter, the same process used in Eq. (3.28) was used for each layer (just replace "α" with "β"). However, the overall backscatter coefficient Jacobian ignores the surface contribution considered in the calculation of $\mathbf{J}_\alpha$ in Eq. (3.27), yielding a slightly different Jacobian given by

$$\mathbf{J}_{\beta} = \begin{bmatrix} \mathbf{J}_{\beta_{l=1}}(m_r^{(s)}) & \mathbf{J}_{\beta_{l=1}}(m_i^{(s)}) & \cdots & \mathbf{J}_{\beta_{l=1}}(C_{\mathrm{v}}^{(s)}) & \mathbf{J}_{\beta_{l=1}}(f_{\mathrm{v},l=1}^{(s)}) & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{J}_{\beta_{l=2}}(m_r^{(s)}) & \mathbf{J}_{\beta_{l=2}}(m_i^{(s)}) & \cdots & \mathbf{J}_{\beta_{l=2}}(C_{\mathrm{v}}^{(s)}) & \mathbf{0} & \mathbf{J}_{\beta_{l=1}}(f_{\mathrm{v},l=2}^{(s)}) & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \mathbf{0} \\ \mathbf{J}_{\beta_{N_{\mathrm{layer}}}}(m_r^{(s)}) & \mathbf{J}_{\beta_{N_{\mathrm{layer}}}}(m_i^{(s)}) & \cdots & \mathbf{J}_{\beta_{N_{\mathrm{layer}}}}(C_{\mathrm{v}}^{(s)}) & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{J}_{\beta_{N_{\mathrm{layer}}}}(f_{\mathrm{v},N_{\mathrm{layer}}}^{(s)}) \end{bmatrix}. \tag{3.21}$$

3.2.2 Calculation of covariance matrices

For the observational covariance matrix $\mathbf{S}_\varepsilon$ described in Section 2.3 applied to our algorithm, it is assumed that correlation among signals is negligible such that off-diagonal "cross-talk" error terms may be ignored. The result is a diagonal covariance matrix that, like the total Jacobian described in the previous section, has individual components for each observation considered, resulting in

$$\mathbf{S}_\varepsilon = \begin{bmatrix} \mathbf{S}_{\varepsilon,\mathrm{L}} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{S}_{\varepsilon,\mathrm{DOLP}} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{S}_{\varepsilon,\alpha} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{S}_{\varepsilon,\beta} \end{bmatrix}. \tag{3.22}$$

By applying Eq. (2.9) to incorporate each j^{th} signal for each measurement (y_j) and its uncertainty (ε_j), each element for all y covariance matrices may be calculated by

$$\mathbf{S}_{\varepsilon,y,jj} = \varepsilon_j^2 y_j^2 \,. \tag{3.23}$$

To calculate the *a priori* covariance matrix, the general method shown in Eq. (2.8) was used. The only difference is that each subsequent element x_j ($1 \leq j \leq N$, where N is number of state vector parameters) along the diagonal of $\mathbf{S}_a$ is calculated relative to the range of the first parameter x_1 in the following manner:

$$\mathbf{S}_a = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & \dfrac{(x_{2,\max} - x_{2,\min})^2}{(x_{1,\max} - x_{1,\min})^2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \dfrac{(x_{N,\max} - x_{N,\min})^2}{(x_{1,\max} - x_{1,\min})^2} \end{bmatrix}, \tag{3.24}$$

where $[x_{j,\min},\ x_{j,\max}]$ are the minimum and maximum values of the j^{th} parameters, respectively, and are calculated using Eqs. (2.6) and (2.7).

3.2.3 Calculation of geophysical variable (GV) uncertainties

Before describing the various methods by which the ICA model was altered to retrieve GV uncertainties, it is necessary to describe how GV uncertainties were calculated. For these experiments, 7 GVs were considered. Five of the GVs were calculated for both the BL and the entire column: aerosol optical depth (AOD, represented by τ), fine mode AOD (τ_{fine}), single scattering albedo (SSA, represented by ω), extinction-to-backscatter ratio (or lidar ratio, abbreviated LR), and the real part of aerosol refractive index (m_r). One GV, aerosol asymmetry parameter (g), is only reported as a column-effective value. The final GV considered in this study, extinction (α), is reported as a layer-resolved value for each of the 16 layers into which the atmospheric column was divided for this algorithm. Each of these GVs, along with their descriptions and SATM retrieval uncertainty goals, are given in Table 1.

Because most of the GVs considered (e.g., SSA, LR) are not contained within the state vector itself but, rather, are derived from the directly retrieved elements of the state vector, the uncertainties of these GVs must account for error contributed by each component involved in their calculations. This process, known as error propagation, utilizes the chain rule to calculate GV uncertainties via

$$\sigma_y = \sqrt{\sum_{i=1}^{N}\sum_{j=1}^{N} \boldsymbol{S}_{\boldsymbol{r},ij} \frac{\partial y}{\partial x_i}\frac{\partial y}{\partial x_j}}, \tag{3.25}$$

where σ_y is the one-sigma uncertainty value for the y^{th} GV considered, y is a given GV, x_i and x_j are elements of the state vector $\mathbf{x}$, and $\boldsymbol{S}_{\boldsymbol{r},ij}$ are elements of the state error covariance matrix, representing the product of infinitesimal changes in the two state vector elements being considered (Burton et al. 2016).

With application of the above chain rule, the uncertainties of the derived GVs within this sensitivity study agree well with the uncertainty magnitudes from retrieval simulations-based uncertainty analysis as well as real data analysis (Xu et al. 2017; Xu et al. 2016). For the uncertainty calculations of asymmetry parameter (column-effective only) and real part of refractive index (column and BL), AOD weighting of uncertainties for all N_s species considered was performed according to Eq. (3.34) and Eq. (3.35), respectively, giving results that were more sensible in magnitude compared to the results of Xu et al. (2017); Xu et al. (2016). These AOD-weighted uncertainty analyses used the following equations:

$$\sigma_{g_{\text{eff}}} = g_{\text{eff}} \sqrt{\sum_{s=1}^{N_s} \left[\frac{\Delta\tau_{\text{a,tot}}^{(s)} \, \Delta g^{(s)}}{\Delta\tau_{\text{a,tot}}} \right]^2} \tag{3.26}$$

$$\sigma_{m_{\text{r,tot[BL]}}} = m_{\text{r,tot[BL]}} \sqrt{\sum_{s=1}^{N_{s,\text{tot[BL]}}} \left[\frac{\Delta\tau_{\text{a,tot[BL]}}^{(s)} \, \Delta m_{\text{r}}^{(s)}}{\Delta\tau_{\text{a,tot[BL]}}} \right]^2} . \tag{3.27}$$

When calculated according to the above equations, these two GVs' uncertainties are likely more realistic because of the relative gain of information content when AOD weighting is included.

Before delving into the specific methods used for this study, it is important to note one difference in the ways relative uncertainties for various GVs were calculated. Full error propagation including all elements of the Jacobian matrix was used in the calculation of AOD and extinction *only*. For all other GVs, only the diagonal elements were used to calculate error. The decision to (not) include certain elements of the Jacobian in calculation of each GV was made after several slightly different uncertainty calculation methods were attempted for each GV. Uncertainty calculation decisions concerning the Jacobian were then chosen based upon how well each agreed with previous results that had been internally validated by direct retrieval simulation and real data analysis methods.

Chapter 4. Methods

The ICA algorithm outlined above was developed together with direct combined retrieval simulation by Xu et al. (2021). In this section, we apply it to the analysis of a specific test case, which will be described within this chapter. Refer to Table 3 for specifics regarding the aerosol loading for the test case. Table 4 is also included at the beginning of this chapter to link the GV uncertainty calculations from Table 1, the "true" state vector parameters listed in Table 2, the case-specific information contained within Table 3, and results within Chapter 5, in which information from Table 4 has been used directly to generate all of the figures shown.

Table 3: Description of test case (Case 61). Aerosol types, their relative positions (FT and BL), and their concentrations (high/low) are provided. The number concentration (N_x, units of m^{-3}), median radius (r_x, units of μm), and logarithmic size distribution ($\ln\sigma_x$) for each type is given, where subscript $x = f$ [c] indicates fine [coarse] mode aerosol. Characteristics are color-coded for smoke (red) and marine (blue).

Species (Concentration)		Number concentration (N_f [m^{-3}])	Number concentration (N_f [m^{-3}])	Median radius r_f [μm]	Median radius r_c [μm]	Logarithmic size distribution $\ln\sigma_f$	Logarithmic size distribution $\ln\sigma_c$
Free Troposphere	Smoke (high)	8.993×10^{8}	9.251×10^{6}	0.12	0.36	0.40	0.45
Boundary Layer	Marine (low)	2.818×10^{7}	4.954×10^{6}	0.20	0.60	0.45	0.70

Table 4: GVs are listed in the left column. Their truth values, as calculated within the ICA algorithm via the equations detailed in Sections 3.1.2 and 3.1.3 are in the middle column. The column on the right gives the threshold (or maximum) GV uncertainty values allowed for the specific case considered in this study. These values were either given directly by the SATM matrix requirements or were calculated from GV truth values using the methods provided by SATM (see Table 1). These GV threshold uncertainties were used as the denominator when calculating the error ratios depicted on the ordinal axis of a majority of results shown (excepting extinction profiles). N.b. that two different values are given for the truth values and uncertainty thresholds of the extinction (α) profiles. The difference in truth values is due to the confinement of each aerosol species to either the BL or the FT. Threshold extinction uncertainty differences between the BL and FT occur because of the SATM calculation requirement for extinction profile (see Table 1).

GV	GV truth values	GV threshold uncertainties for case considered
SSA, column	0.89	±0.04
SSA, BL	~0.99	
g_{eff}, column	0.70	±0.02
LR, column	45.93	±11.48
LR, BL	24.09	±6.02
τ_{fine}, column	0.25	±0.03
τ_{fine}, BL	0.03	±0.02
τ, column	0.36	±0.04
τ, BL	0.10	±0.03
m_r, column	1.488	±0.025
m_r, BL	1.378	
α, profile, FT	129.20 [Mm^{-1}]	±25.84 [Mm^{-1}]
α, profile, BL	50.31 [Mm^{-1}]	±20.00 [Mm^{-1}]

4.1 Algorithm description and major assumptions

Within this section, some constant characteristics of the algorithm used are described, along with elaboration upon the key assumptions made during the algorithm's development. For

each case, the atmospheric column is divided into 20 layers, with the bottom 4 km of the atmosphere (i.e., the troposphere) making up the bottom 16 layers, spaced at 250 m vertical resolution. Within the uppermost four layers, Rayleigh scattering dominates, and no aerosols are present. Thus, these layers are largely ignored in the calculation of aerosol GVs. For this study, the boundary layer (BL) is operationally defined as the bottom eight layers of the atmospheric column (0-2 km above sea level). Free troposphere (FT), then, will describe the eight layers (2-4 km above sea level) between the BL and the Rayleigh-dominated layers. Though calculations for "total" column and column-effective values of the GVs include all layers, the vertical extinction profile is only given for the 16 layers in which aerosols are present, where Mie scattering theory applies. The view angle for all trials aligns with the principal plane, and three solar zenith angles (0º, 30º, 60º) could be considered so that any angular dependence of resultant GV uncertainties may be parsed out. After analyzing two different instrument combinations for our test case at all three solar zenith angles, it became clear that averaging the results from those three angles gave GV uncertainty values that were not significantly different than those obtained by only considering the 30º solar zenith angle. Thus, for all results reported within this paper, the 30º solar zenith angle was used as a proxy for the angle-averaged results to minimize computational time, thereby making the ICA tool more efficient.

Throughout all experimental trials, the same polarimeter (POL) was used. This polarimeter receives signals at seven wavelengths (360, 380, 410, 550, 670, 870, and 1650 nm) and 10 different angles between ±57º, inclusive (±57º, ±44º, ±32º, ±19º, ±6º). The proposed polarimeter's view angles and received wavelength channels were decided upon by combining strengths of earlier airborne MAPs SPEX airborne (Smit et al. 2019) and AirHARP (Puthukkudy et al. 2020). More specifically, these specifications closely mirror those of the proposed NASA PACE mission

(Cetinic et al. 2018a, 2018b), which will use three collocated satellite-borne instruments, including updated versions of the previously mentioned airborne MAPs, HARP2 and SPEXone (Fernandez Borda et al. 2018; Hasekamp et al. 2019; Martins et al. 2018). The proposed polarimeter has reference measurement uncertainty values of ΔI/I = 3% for intensity and ΔDOLP = 0.005 for degree of linear polarization.

For this study, three different instrument combinations were tested, each of which used polarimeter combined with one lidar. Throughout the remainder of this paper, the different lidars will be referred to as lidar 5 (L5), lidar 6 (L6), and lidar 9 (L9). Lidar 5 may be described as 1β 1α $1\beta_{\mathrm{attn}}$, for it measures backscatter and extinction at 532 nm as well as attenuated backscatter (β_{attn}) at 1064 nm, mimicking the specifications of the airborne HSRL instrument (Hair et al. 2008). In general, the number preceding a measurement within the full lidar description gives the number of wavelengths at which information is retrieved. Following this convention, L6 may be described as 2β 2α $1\beta_{\mathrm{attn}}$, where extinction and backscatter are measured at 355 and 532 nm and attenuated backscatter is measured at 1064 nm [see (Burton et al. 2016)]. Likewise, L9 may be described as $2\beta_{\mathrm{attn}}$, where only attenuated backscatter is measured at both 532 and 1064 nm, which is similar to the CALIOP lidar system used on the CALIPSO mission [see e.g., (Hunt et al. 2009; Winker et al. 2009)]. For all three lidars, the reference measurement uncertainties are Δα = 1.7e-5 [m^{-1}] for extinction and Δβ/β = 5% for backscatter. For ease of reference, specifications of the three lidar systems and polarimeter considered have been summarized in Table 5.

Table 5: Instruments used in this sensitivity study and their directly measured quantities. Truth measurement error values are listed for all four measurements, where δ_x and σ_x represent relative and absolute uncertainty, respectively. Wavelengths and view angles at which each instrument makes observations are also listed, where applicable.

Instrument	Retrieved/Calculated Measurements	Measurement Reference Uncertainty	Wavelengths [nm]		View angles
Polarimeter	Intensity	$\delta_I = 3\%$	360, 380, 410, 550, 670, 870, 1650		±57º, ±44º, ±32º, ±19º, ±6º
	DOLP	$\sigma_{DOLP} = 0.005$			
Lidar 5	Extinction (α)	σ_α = 1.7e5 [Mm^{-1}]	532		Nadir
	Backscatter (β)	$\delta_\beta = 5\%$	β	β_{attn}	
			532	1064	
Lidar 6	Extinction (α)	σ_α = 1.7e5 [Mm^{-1}]	355, 532		Nadir
	Backscatter (β)	$\delta_\beta = 5\%$	β	β_{attn}	
			355, 532	1064	
Lidar 9	Extinction (α)	σ_α = 1.7e5 [Mm^{-1}]	–		Nadir
	Backscatter (β)	$\delta_\beta = 5\%$	β	β_{attn}	
			–	532, 1064	

Having discussed characteristics of the ICA algorithm — including its resolution, internal calculation methods, and available instrumentation choices — it is necessary to describe additional assumptions made and requirements adhered to during the development of the forward model, as

well as the inherent uncertainties associated with each. Though the methods of calculating the maximum uncertainty thresholds for each GV considered in this study (see Table 1) are taken directly from the SATM requirements established by NASA's ACCP mission, a brief explanation of the body of research that has led to the development of these specific target uncertainties is necessary to legitimize the choices made during the development of this algorithm.

Since the nascence of climate studies, accurately quantifying net global radiative forcing (RF) has been of paramount importance. Because aerosols' numerous and widely varying interactions with clouds — as well as their relative abundances, spatiotemporal distributions, etc. [see Chapter 1] — greatly affect RF magnitude, a long-standing objective of the aerosol remote sensing community has been the improvement of observational and model capabilities such that species- and interaction-specific contributions may be parsed out from the already well-established total anthropogenic RF (IPCC 2014). To further this goal, many studies in recent years have focused on reducing the retrieval uncertainties of aerosol properties (e.g., the GVs in Table 1). As a part of NASA's Glory mission, Mishchenko et al. (2004) set forth ambitious retrieval uncertainty thresholds for several GVs (e.g., AOD, SSA, and m_r), which were to be accomplished by the proposed Aerosol Polarimetry Sensor (APS). The APS instrument ultimately failed to produce useful data due to its inability to reach orbit; however, advances in the accuracy and variety of observational and modeling capabilities remained continuous. As a result, NASA's subsequent Aerosol Cloud and Ecosystems (ACE) mission utilized the capabilities of lidar in addition to polarimetry to set even higher accuracy thresholds (i.e., lower GV retrieval uncertainty goals) for a larger suite of GVs (Loeb and Su 2010, among others).

Because of these (and many other) successful advances, the accuracy of estimated total aerosol and cloud RF was greatly improved between the IPCC's fourth (AR4) and fifth (AR5)

assessment reports. However, uncertainty values surrounding this estimate increased during this same period due to the addition of many different modeled and observationally derived data used in calculation of the estimate. As a result, aerosols, clouds, and their interactions remain the largest sources of uncertainty in RF estimation (IPCC 2014). To promote more efficient use of this proliferation of new modeling capabilities and rich observational data, the 2017 Decadal Survey called for a twofold reduction of the aerosol-generated RF uncertainties reported in AR5 (NASA 2017). NASA subsequently established the ACCP mission, which has allowed for greater multi-center collaboration on instrument and model development to reach this goal. Because the total RF uncertainty is essentially the result of a complex error propagation problem (see Section 2.1), preliminary work on the ACCP mission included solving the inverse problem for a large suite of aerosol GVs to find the maximum acceptable uncertainty ranges for each (NASA 2019). The GV uncertainty calculation methods listed in Table 1 (SATM threshold uncertainties) represent a select few of these ranges.

Now that the origin of the GV reference uncertainties in Table 1 — which are used as a metric for the relative success/failure of the various instrument combinations considered in this study — has been thoroughly explained, it is necessary to acknowledge a few major assumptions made within the ICA model used. Each assumption concerns the properties and/or interactions of any aerosols whose characteristics are used as model input.

By far the least potentially confounding assumption is that of a bimodal lognormal size distribution. In general, nearly any single-population (in this case, single-mode) distribution within a natural system may be approximated as a normal distribution, given a large enough sample size of similarly sized individual samples of the entire population. Because normal distributions are, by definition, centered at a mean value of zero, lognormal distributions must be used for any

physical quantity for which negative values are impossible. The assumption of a bimodal size distribution is, in a word, conventional. Many previous aerosol modeling studies have used this assumption, and its validity has been backed by numerous *in situ* studies. It should be noted that there have been case-by-case outlier instances in which trimodal size distributions have been discovered. However, these cases are exceedingly rare. As such, it is generally seen as safe to ignore such cases without losing model generality, significantly increasing model error, or impacting the accuracy of theory-based calculations.

Another major assumption of the ICA model regards the mixing state of aerosols used as model input. Though the model allows for separation of the column into 20 vertically resolved layers (of which only 16 are considered [see Section 4.1]), the version used for this study assumed two major layers (BL and FT), each of which contained only one aerosol species (see Table 3). Thus, mixing in each layer consisted only of internal mixing between fine and coarse mode particles. Moreover, this internal mixing was assumed to be both vertically and horizontally homogenous across the entire layer volume. Because the BL and FT were considered distinct, non-interacting layers, no external mixing was considered. The primary reason for these assumptions is that for every aerosol species considered, an effective set of aerosol properties is used as input so that measured signals may be reproduced reliably and used for more reliable comparison with results from real data analysis and inverse modeling results.

Because we assume homogeneity of aerosol distributions in these ways, it is likely that the calculated GV uncertainties of GVs which use lidar measurements for their calculations are not representing error accurately due to the uniformity of aerosol loading throughout each vertical layer. Moreover, it is likely that the impact of using vertically homogenous aerosol distributions results in an underrepresentation of GV relative error (particularly with extinction) due to the

ability of lidar to retrieve vertically resolved GVs well. That is, if the vertical homogeneity constraint within our model were to be relaxed, the calculated GV errors — which are calculated always in a sense relative to the true GV values — would be higher than shown in our current results, at least for those GVs that use lidar measurements and/or are vertically resolved.

Though the spatiotemporal diversity of aerosols is, in reality, typically quite different than the setup used within this model, the aforementioned benefits of making such an assumption are of greater importance at this stage of model development, when the overall validity of the model is still being established. Allowing for more realistic aerosol mixing at this point would add a considerable number of new error sources to the model, which would make establishing the internal validity of the model far more difficult.

Whereas the previous assumption concerned extrinsic properties of aerosols as well as their (non)interactions with aerosols of a different species, the last major assumption within the ICA model concerns a physical characteristic of each particle within the sample volume: sphericity. Again, it is known that in reality, many aerosols are not spherical. However, there are several reasons that the sphericity assumption is made in this and many other aerosol modeling studies, at least during initial trials. A few large factors limiting the inclusion of non-spherical particles in many studies are that most MAP-based retrievals of nonsphericity are based on the assumptions of 1) inclusion of a *single* but effective aerosol species and 2) spheroidal shape for all non-spherical particles. For the retrieval of *multiple* aerosol species from a combined use of lidar and polarimeter, some preliminary studies have to be performed first to identify 1) the valid species-specific non-spherical particle shapes that must be considered and 2) how to best improve algorithms accordingly to parse out various species identities, relative quantities, etc. from the complex data retrieved from more realistic sample volumes containing more than one aerosol species. Some very

recent studies have begun to include spheroidal particles within their models (e.g., Xu et al. 2021); however, including particles of one higher-order shape dimension adds a large amount of complexity to the scattering theory calculations, and we exclude them in this thesis for simplicity.

As is the case with nearly any assumption made regarding a computational model, each of the three assumptions described within this section are, to varying extents, over-simplifications of reality. However, two of the assumptions have significant precedence in the literature, particularly in studies concerning the early stages of new model development. The other major assumption — that aerosols typically obey a bimodal lognormal size distribution — is firmly established in the literature across modeling, remote sensing, and *in situ* studies.

4.2 Perturbation methods utilized

Having, now, an algorithm within which $\mathbf{S_a}$ and $\mathbf{S_\varepsilon}$ are easily altered, several analyses of the relative importance of information content within the algorithm were done using the same basic concepts learned from the toy model. The goal is simple: to find the extent to which the accuracy of measurements taken and knowledge of the environment *a priori* add to the successful retrieval of GV uncertainties. In any model, it is important to know the dependence of retrieval accuracy of GVs on the observational and *a priori* constraints imposed on aerosol properties. To see the impacts of *a priori*, the algorithm was run for several different scenarios. For each of the methods used, either quantities contained within the state vector or measurement errors were initialized at some uncertainty range, which were subsequently widened over 100 iterations of the model. In essence, for each different method, some set of quantities remains static and known while the uncertainty associated with either *a priori* knowledge of the environment or measurement precision is gradually increased. Because of the varying scales that must be used for these perturbations across the various methods utilized, each method must be described in more detail.

In the first scenario (i.e., altering state vector parameters' *a priori* uncertainties), the measurement errors are all left at their reference values, and the *a priori* state vector is initialized with minimum values set at truth. Initial maximum values for the state vector parameters $m_r^{(s)}$, $m_i^{(s)}$, $r_{\mathrm{m}}^{(s)}$, and $\sigma_{\mathrm{m}}^{(s)}$, were set at a value of $1.0001x_{\mathrm{truth}}$, where x_{truth} represents the "true" (i.e., case-specific data input directly into the model) *a priori* values of these parameters. The ultimate multiplicative scale factor (i.e., the factor used for the largest uncertainty window) for each of these four parameters has to be different to maintain a physically reasonable range of uncertainties. For example, having a scale factor that is a whole number quickly becomes unphysical for $m_r^{(s)}$, regardless of species or situation considered, for typical values of the real part of refractive index range from 1.33 – 1.7. Moreover, $r_{\mathrm{m}}^{(s)}$ may have values that reasonably range about an order of magnitude, and $m_i^{(s)}$ can easily range several orders of magnitude, depending on the species being considered. The final state vector parameters whose ranges vary with each iteration of the ICA code, $f_{\mathrm{v}}^{(s)}$ and $C_{\mathrm{v}}^{(s)}$, also make use of scale factors. However, the uncertainty ranges for these two parameters are calculated in a slightly different way than was done for the first four state vector parameters described. More specifically, the minimum and maximum values for $f_{\mathrm{v}}^{(s)}$ and $C_{\mathrm{v}}^{(s)}$ were determined using Eqs. (4.1) and (4.2):

$$x_{\mathrm{min}} = x_{\mathrm{truth}} / \Delta F_{\mathrm{min}} \tag{4.1}$$

$$x_{\mathrm{max}} = x_{\mathrm{truth}} \Delta F_{\mathrm{max}}\,, \tag{4.2}$$

where ΔF_{min}, the scale factor for the minimum value of the uncertainty range, is 1.0001, as it is for the other four parameters mentioned. It should be noted that using Eq. (4.1) does, indeed, reduce the values of the minimum uncertainty range to just below the truth; however, such a small

difference in value is, in practice, equivalent to setting the minimum value at truth, as was done for $m_r^{(s)}$, $m_i^{(s)}$, $r_{\mathrm{m}}^{(s)}$, and $\sigma_{\mathrm{m}}^{(s)}$.

As discussed previously, the scale factor values used to determine the maximum value of each parameter's uncertainty range are necessarily different. Therefore, Table 6 is included to show speciated state vector parameters for Case 61 whose *a priori* uncertainty range was perturbed along with its associated range of actual values associated with the range of each speciated state vector parameter value. Each parameter's set of scale factors was one of 100 evenly spaced values between ΔF_{min} and ΔF_{max}, in linear space. Running the model 100 times by iteratively increasing the multiplicative scale factors represents a continuous decrease in the *a priori* knowledge of the environment being considered. That is, as the multiplicative scale factors increase, one knows less prior information about aerosol and surface properties because of the use of a less accurate aerosol and surface climatology, meaning that the measurements themselves are being relied upon more to constrain the resultant GV uncertainties.

Table 6: Speciated parameter values for real part of refractive index (m_r), imaginary part of refractive index (m_i), median radius (r_m), logarithmic size distribution (σ_m), fractional volume (f_v), and volumetric concentration (C_v) are listed for the four species considered in our test case. The actual parameter values — or calculation methods used to obtain them — that are associated with the maximum scale factor value for the first perturbation and final perturbation are also given.

Parameter	$x_{\max,1}$	$x_{\max,100}$
m_r	$1.0001 m_{r,\text{truth}}$	1.7
mi	$1.0001 mi$,truth	0.1
r_m	$1.0001 r_{m,\text{truth}}$	$10 r_m$
σ_m	$1.0001 \sigma_{m,\text{truth}}$	$5 \sigma_m$
f_v	1×10^{-6}	1.0
$\left[C_{v,S_f} C_{v,S_c} C_{v,M_f} C_{v,M_c}\right]$	1×10^{-10} [0.27, 0.09, 0.05, 0.81]	1×10^{-8} [0.27, 0.09, 0.05, 0.81]

The other two methods involve iteratively increasing the measurement uncertainty ranges used in calculating GV uncertainties while leaving the *a priori* state vector information static. Because four different measurements (see Section 4.1) are used to calculate GVs in the combined retrieval setting — each of which must be altered while the other three measurements are held at their reference uncertainty values — a total of eight different sets of results are obtained from the two methods that involve increasing measurement uncertainties. Common to all of these trials is the use of a formula analogous to Eq. (4.2) to determine the maximum measurement error value for each subsequent iteration. As was the case for the *a priori* state vector parameters, the maximum values of measurement error scale factors may not be the same and still make physical sense. Moreover, of the four measurements utilized within the algorithm, backscatter and intensity report uncertainties as relative error (δ_x), whereas DOLP and extinction report absolute error (σ_x),

where x is the measurement in question. Initial measurement uncertainty values (reference values) are provided in Table 5 (see Section 4.1), and the value of $\Delta F_{\max}$ for the first iteration across all measurement error perturbation trials was set at a very small value of ($1+1\times 10^{-8}$).

Because a relative error of 1 indicates 100% error, the largest values of $\Delta F_{\max}$ for intensity and backscatter were chosen such that measurement uncertainty values of the last iteration are $\delta_x \approx 1$. Because DOLP represents the portion of incident wave that is polarized, a value of $\sigma_{\mathrm{DOLP}} = 1$ (complete polarization) was set as the maximum uncertainty value for DOLP. The maximum value of extinction uncertainty for these trials is $\sigma_\alpha = 85$ [Mm^{-1}]. For extinction and backscatter perturbations, the channels for all wavelengths considered are perturbed simultaneously.

Now that the method by which various measurement uncertainties are altered has been described in detail, the two different static *a priori* setups for elements of the state vector must be explained. Generally speaking, the two different setups may be described as "known environment" and "relaxed environment," each of which refers to the relative size of the *a priori* uncertainty ranges used for elements of the state vector. In other words, uncertainty ranges in the "known" cases were set such that one would have a reasonably good idea of the environmental and surface conditions of the volume being sampled, whereas parameters' uncertainty ranges for the "relaxed" cases are set sufficiently large that essentially nothing is known about the specifics of the environment within the sample volume. For both environmental setups, the minimum values of the uncertainty ranges for $m_r^{(s)}$, $m_i^{(s)}$, $r_{\mathrm{m}}^{(s)}$, and $\sigma_{\mathrm{m}}^{(s)}$ are set at truth. For the known environment setup, the maximum value of $m_r^{(s)}$ is 1.65 and the maximum values for $m_i^{(s)}$, $r_{\mathrm{m}}^{(s)}$, and $\sigma_{\mathrm{m}}^{(s)}$, are all set at $2(x_{\mathrm{truth}})$. The scale factors used to compute minimum and maximum values of $f_{\mathrm{v}}^{(s)}$ and $C_{\mathrm{v}}^{(s)}$ for the known environment setup are $\Delta F_{\min} = 1.0001$ and $\Delta F_{\max} = 10$, respectively. On the other

hand, the scale factors used to compute $f_{\mathrm{v}}^{(s)}$ and $C_{\mathrm{v}}^{(s)}$ in the relaxed environment setup are $\Delta F_{\min} = \Delta F_{\max} = 20$. Though these values do not seem, at first glance, to produce a larger range of uncertainty than those used in the known environment setup, one can quickly verify these relations using Eqs. (4.1) and (4.2). As far as the relaxed environment maximum values of the other state vector parameters are concerned: $m_{r,\max}^{(s)} = 1.7$, $m_{i,\max}^{(s)} = 0.1$, $r_{\mathrm{m,max}}^{(s)} = 10\left(r_{\mathrm{m,truth}}^{(s)}\right)$, $\sigma_{\mathrm{m,max}}^{(s)} = 5\left(\sigma_{\mathrm{m,truth}}^{(s)}\right)$. For a consolidated list of minimum/maximum state vector values used for the known/unknown environment settings, see Table 7 and Table 8.

Table 7: Speciated parameter values for real part of refractive index (m_r), imaginary part of refractive index (m_i), median radius (r_m), logarithmic size distribution (σ_m), fractional volume (f_v), and volumetric concentration (C_v) are listed for the four species considered in our test case. The actual parameter values — or calculation methods used to obtain them — that are associated with the static a priori minimum/maximum scale factor values for the state vector variables in the known environment case are also given.

Parameter	$x_{\min}$	$x_{\max}$
m_r	$m_{r,\mathrm{truth}}$	1.65
mi	$m_{i,\mathrm{truth}}$	$2m_{i,\mathrm{truth}}$
r_m	$r_{m,\mathrm{truth}}$	$2r_m$
σ_m	$\sigma_{m,\mathrm{truth}}$	$2\sigma_m$
f_v	0.125	1.000
$\left[C_{v,S_f} C_{v,S_c} C_{v,M_f} C_{v,M_c}\right]$	1×10^{-10} [0.27, 0.09, 0.05, 0.81]	1×10^{-9} [0.27, 0.09, 0.05, 0.81]

Table 8: Speciated parameter values for real part of refractive index (m_r), imaginary part of refractive index (m_i), median radius (r_m), logarithmic size distribution (σ_m), fractional volume (f_v), and volumetric concentration (C_v) are listed for the four species considered in our test case. The actual parameter values — or calculation methods used to obtain them — that are associated with the static a priori minimum/maximum scale factor values for the state vector variables in the unknown environment case are also given.

Parameter	$x_{\min}$	$x_{\max}$
m_r	$m_{r,\text{truth}}$	1.7
mi	$m_{i,\text{truth}}$	0.1
r_m	$r_{m,\text{truth}}$	$10r_m$
σ_m	$\sigma_{m,\text{truth}}$	$5\sigma_m$
f_v	0.0062	1.0000
$\left[C_{v,S_f} C_{v,S_c} C_{v,M_f} C_{v,M_c}\right]$	1×10^{-11} [0.13, 0.05, 0.02, 0.41]	1×10^{-8} [0.05, 0.02, 0.01, 0.16]

In all three methods used, the retrieval code was run for a total of 100 incrementally increasing *a priori* uncertainty ranges, which are controlled by the various multiplicative scale factors described previously. The resultant GV error ratios were plotted as a function of perturbation number or instrument error for the *a priori* state vector perturbation method and measurement error perturbation methods, respectively. For purposes of this paper, "error ratio" is defined as the ratio of the retrieved GV error (σ_{calc}) to the SATM target uncertainty (σ_{SATM}). Thus, an error ratio value of 1 means that the GV error is as large as allowed by the SATM requirement. For ease of analysis across GVs and cases, an asymptote at this error ratio value (denoted by a black, dashed line) has been included on each error saturation plot. Then, any error saturation curve exceeding this asymptote at all points shows that the GV in question's uncertainty cannot be well-constrained, even under the most ideal circumstances. If the error curve saturates completely at a value under the asymptote, the GV error is well-constrained for that particular setup. For instances

in which the curve crosses the asymptote, one may extrapolate from the intersection of these curves how large the *a priori* uncertainties of state vector parameters or of specific measurements may be before particular GVs can no longer be retrieved within an acceptable range of error.

Chapter 5. Results

As we begin to consider the various error perturbation methods, the general structure of the chapter must be described. Each subsection will follow a structure similar to that of Section 5.1, in which the results of the first perturbation method (i.e., increasing *a priori* error of state vector parameters) are discussed. Within Sections 5.2–5.5, we gauge the importance of each measurement accuracy to the ability of our model to constrain calculated GV uncertainties and analyze the impact of *a priori* knowledge on retrieval ability as each measurement accuracy declines by directly comparing results from the known and unknown environment methods described in Section 4.2. The following research questions are addressed throughout analyses within Chapter 5, and all conclusions we can draw by answering these questions are included in Chapter 6.

RQ1. Which, if any, of the GVs considered in this study may *not* be retrieved within their SATM-defined threshold uncertainties when *a priori* knowledge of state vector parameters is poorly characterized (i.e., "unknown environment" setup)?

RQ2. What is the impact of including more observational information content from lidar measurements (i.e., increasing the number of receiving channels) on GV retrieval abilities?

RQ3. Which observational configuration(s) tested is (are) able to constrain the greatest number of GVs in the most realistic sample environment (i.e., "unknown environment" setup)?

RQ4. Of the four measurements considered, which is most essential for our model's ability to adequately constrain GV uncertainty values within an "unknown" environment?

RQ5. Assuming measurement errors are minimized, to what extent must the a priori state vector parameters be known to guarantee adequate GV uncertainty constraint ability?

Before any analysis of results, we must check for internal consistency with our initial a priori perturbation method (see Section 5.1). Even though the starting uncertainty windows for a priori state vector parameters are slightly different (i.e., larger) than those used for the known environment method, the initial error ratio values in the known environment case (left column) should be similar to those values found within the first few a priori perturbations shown in Section 5.1.1. Because initial values appear to be within a reasonable range of previous results (and continue to do so throughout subsequent sections), we may conclude that the model has internal validity.

As a brief caveat, note well that much of the analysis within the subsections of Chapter 5 will be qualitative in nature, for the primary purpose of this study is to show the advantages of combined lidar-polarimeter retrievals of aerosol properties relative to those performed by lidar or polarimeters alone. However, tables have been included at the end of each subsection to provide — in a consolidated, easy-to-reference format — more specific, quantitative results. These data are useful to include for future research applications that may use this model. Additionally, including specific numeric results allows for researchers using slightly different models, methodologies, etc., to have something concrete against which past studies' findings may be compared. Moreover, inclusion of these data could impact the focus of future forward modeling and ICA studies by showing areas in which improvement can and must be made more starkly than could be accomplished through qualitative analysis alone.

5.1 Effects of relaxing a priori environmental knowledge

5.1.1 Conventional GV uncertainties

For AOD, there is an expectation the polarimeter will have reasonable ability to constrain column-integrated values of AOD because of MAPs' sensitivity to column-integrated values. Additionally, the range of view angles used by this polarimeter should work well for AOD retrieval because of oblateness of view angle (Xu et al. 2017). Another reasonable expectation, given Eq. (3.12), is that lidars with extinction measurement should also be able to constrain AOD values well. Thus, any combination of these should be able to constrain. There is little reason to think that L9 should be able to constrain AOD at all because of its backscatter-only measurement abilities.

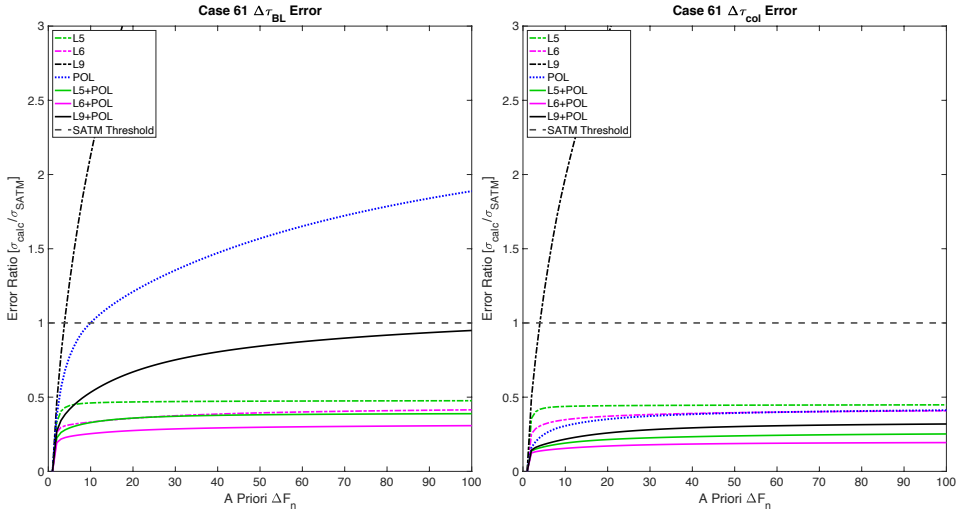

Figure 1: Test case total AOD ($\Delta\tau$) error ratio versus a priori state vector scale factor (ΔF_n) for boundary layer (BL, left panel) and total Mie scattering portion of atmospheric column (col, right panel), where n is the perturbation number (n=1,…,100). Results are shown for four single-instrument setups and three instrument combinations: L5 (green dash-dot), L6 (magenta dash-dot), L9 (black dash-dot), POL(blue dotted), L5+POL (green solid), L6+POL (magenta solid), and L9+POL (black solid)

Overall, the results in Fig. 1 seem to agree with expectations. The polarimeter-only case is unable to constrain within boundary layer likely because of interference by layers above (that is, the AOD of the FT). Because the additive uncertainty of AOD is hard to parse out, particularly because of the relatively poor ability of the polarimeter to give vertically resolved values of single-scattering properties). Additionally, the relatively lower number of particles within the BL sample

volume (compared to the column) likely reduces the polarimeter's ability to constrain AOD. This agrees with the inverse relationship between aerosol loading and AOD uncertainty described by Knobelspiesse et al. (2012). Because of polarimeter's relatively better ability to detect column-integrated values, it does show relatively greater ability to constrain total column uncertainty of AOD (though underperforming in BL case, as described previously). Though every instrument setup (other than L9) is able to constrain column AOD, all three instrument combinations perform better than any single-instrument setup. Moreover, though L9- and POL-only cases are unable to constrain AOD values within the BL, L9+POL is able to constrain, even when *a priori* environmental knowledge is scarce.

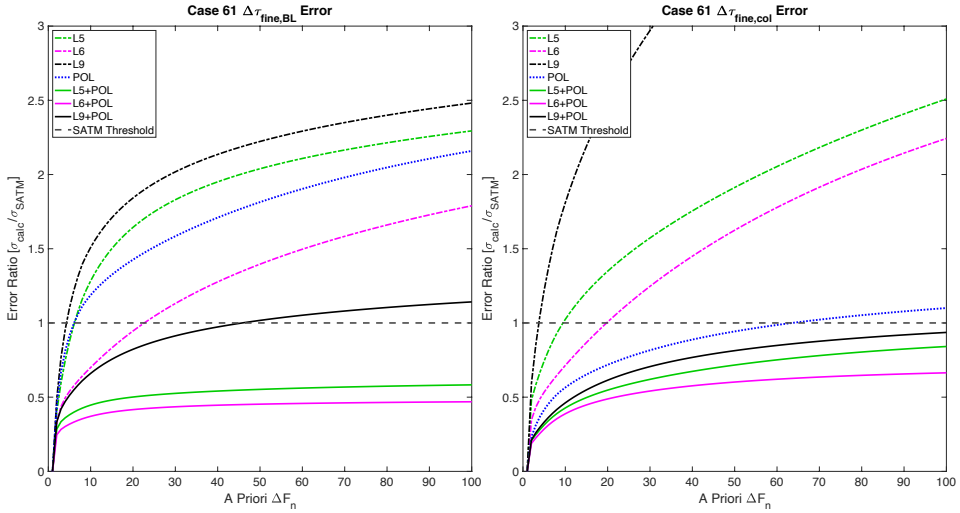

Figure 2: As in Fig. 1, except for fine mode AOD ($\Delta\tau_{fine}$) error ratio.

As regards the fine-mode AOD retrieval uncertainty, it should not be surprising that the overall ability to constrain is generally worse for polarimeter and combined instrument retrievals than it is for total AOD, for lower AOD values should yield higher relative AOD retrieval uncertainties (Knobelspiesse et al. 2012). It is clear in Fig. 2 that L9 and L5 are not viable options for AOD constraint, which makes sense, considering the lower amount of extinction coefficient information contained in these measurements versus L6. The reason for the relatively poor BL

performance of instrument combinations and polarimeter is the same as before: lower AOD yields higher AOD uncertainty. Regardless of vertical range considered, L5+POL and L6+POL combinations are able to constrain fine-mode AOD, even when the environment is unknown. Polarimeter sensitivity to fine-mode aerosols combined with the added information content of L5 and L6's vertically resolved extinction coefficient measurements explains why the AOD uncertainties are smallest for the L5+POL and L6+POL instrument combinations. As long as the environment can be somewhat well-characterized, the L9+POL combination would be able to constrain AOD, though this clearly limits that combination's comparative utility. Because of the POL-only inability to constrain well the fine-mode AOD within the BL, one can already see the limitations of polarimetric measurements, even when the environment is well-characterized. This shows clearly that combined instrument retrievals are superior to polarimeter- or lidar-only retrievals, at least as regards AOD.

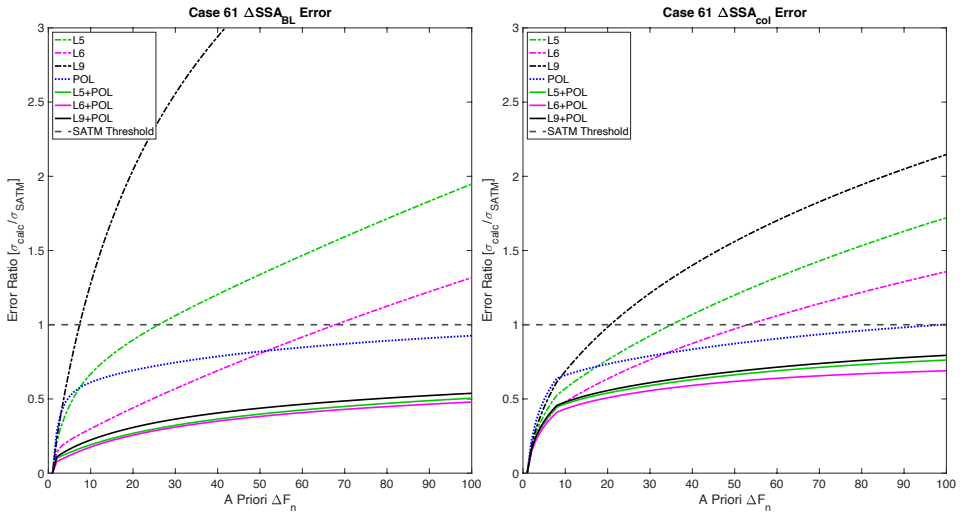

Figure 3: As in Fig. 1, but for single scattering albedo (ΔSSA) error ratio.

Referring to the SSA uncertainty plots, it is clear that as long as the environment is somewhat well-characterized, column values of SSA can be well-retrieved by all instrument choices, except perhaps L9. However, neither L9 nor L5 can adequately constrain SSA error within

the BL. The extra backscatter and extinction receivers added as one moves from L9 to L5 to L6 explains the relative success of those three instruments, as well as their respective combinations with the polarimeter measurements. Overall, the ability of the polarimeter to constrain well the SSA error makes sense, for MAPs tend to retrieve size distributions and shape information very well (Marshall et al. 1995).

Now one must consider why there is a relatively lower ability of the polarimeter to constrain the column SSA error. This is likely caused by the addition of the FT layer, which contains lower values of SSA because of the relatively higher absorption by smoke particles. Additionally, given a GV truth value and a window of error that is increasing, the magnitude of uncertainty relative to reported GV value gets larger faster for smaller truth values. Addition of polarimetric measurements (which are sensitive to the angular distribution of scattering) does help the combined instrument cases retrieve SSA better for the entire column versus the BL, which is likely attributable to a combination of factors. First, simply the addition of more information content could have helped improve the retrieval accuracy. Also, adding lidar adds vertically resolved extinction coefficient values which affect AOD values at every level. Since the SSA values used in this study were AOD-weighted, anything affecting the AOD should also affect the SSA.

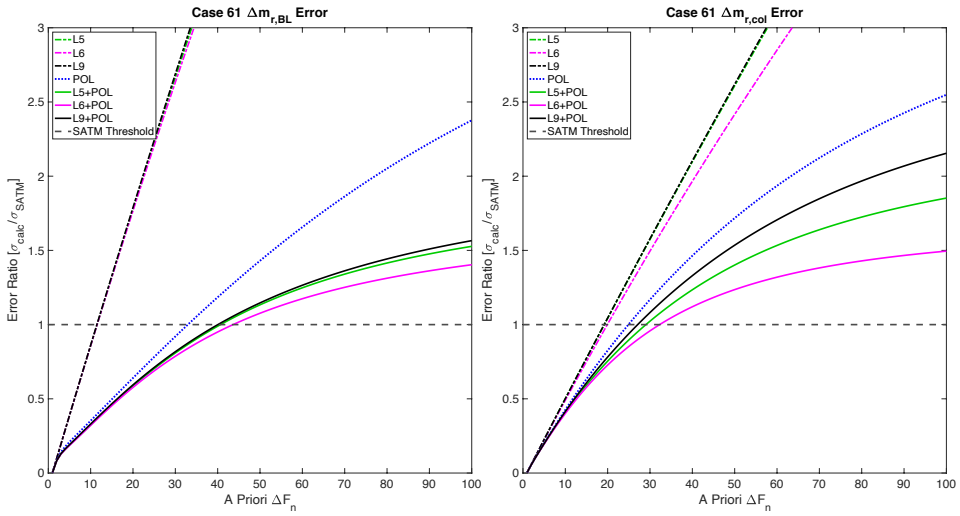

Figure 4: As in Fig. 1, but for real part of refractive index (Δm_r) error ratio.

Despite the sensitivity of MAPs to the complex refractive index, the polarimeter and all of the combined instrument cases constrain RRI error only as long as the environment is moderately well-known. However, considering that RRI is one of the *a priori* state vector parameters that is having its initial uncertainty window expanded to yield these results, one would naturally expect for this particular GV to be somewhat harder to constrain than the others. The most important takeaway from this set of figures is that the relative ability of all three combined instrument retrieval methods to constrain RRI is greater than that of any single-instrument method.

For this particular GV, it is important to note the differences between the BL and column results. Specifically, it seems counterintuitive that the polarimeter-only and combined instrument retrievals would perform better in the BL than for the entire column, considering the unavoidable two-way attenuation due to the FT layers that any signals would be subjected to. There are likely several contributing/competing factors that allow for relatively better ability to constrain RRI within the BL versus over the whole column. First, there is a relatively high concentration of smoke particles within the FT layers for Case 61. Though this adds a considerable information content, it also introduces, in this setup, two size modes of an entirely different species (i.e., smoke) with

different complex indices of refraction. So, one possible reason the BL values are more well constrained by the polarimeter and combined retrievals is that though there is less information contained in the BL retrievals, the amount of associated uncertainty is less because of the fewer number of different sizes/species of aerosols considered.

In addition, other competing factors may potentially add to this interesting relationship. For example, MAPs are sensitive to size distribution, and the FT is loaded with smoke particles that have smaller size distributions relative to those of the marine particles in the BL (Xu et al. 2017). This greater sensitivity should reduce uncertainties; however, the composition of the species confined to the FT may actually add enough extra uncertainty to overcome the added benefit of MAP sensitivity. That is, the widened *a priori* uncertainty ranges likely impact the species within the FT more so than those in the BL for Case 61, particularly due to the larger contribution of imaginary part of refractive index uncertainty to accurate retrievals of smoke particles relative to retrievals of salt particles. On the other hand, lidars are sensitive to particle concentration, so adding layers of high particle concentration as happens with Case 61 could give enough extra information content to help constrain the entire column better than the BL alone, which can be observed when looking at the lidar-only retrievals (Ansmann and Müller 2005; Burton et al. 2016). Though the exact relationships between all of these factors is not yet well-understood, further analysis of cases with different aerosol loading in different layers may shed some light on the underlying reasons for such a relationship to exist in this environmental setup.

Now considering the asymmetry parameter retrievals, one can see that the lidar-only retrievals are relatively poor in their ability to constrain retrieval of g_{eff}. Though the study used a scanning mobility particle sizer, not a polarimeter, Andrews et al. (2006) found that inclusion of aerosols with diameters greater than 0.15μm — which is the case for this study — within Mie

theory calculations result in the ability to capture 95% of the asymmetry parameter's true value. Because MAPs are sensitive to particle size distribution, it makes sense that the polarimeter-only setup was able to adequately constrain g_{eff}, even with almost no a priori environmental information. Though polarimeter alone is sufficient to constrain this particular GV, the comparative advantage of all three combined instrument retrievals should be noted.

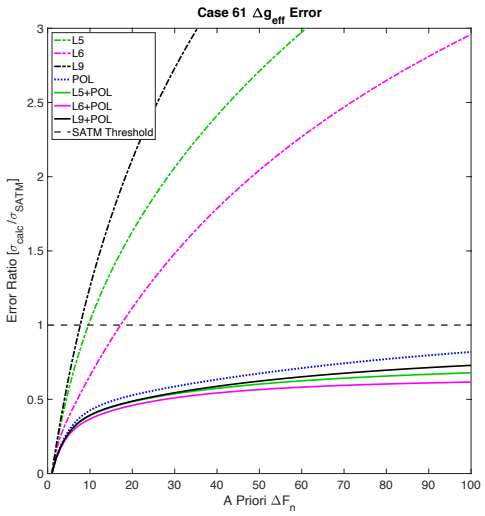

Figure 5: As in Fig. 1, but for column-effective asymmetry parameter (Δg_{eff}) error ratio.

5.1.2 Lidar-sensitive GV uncertainties

Of all the GVs considered in this sensitivity study, the only one reported as a vertically resolved value is extinction. The uncertainties associated with the various instrument combinations (including single-instrument cases) agree well with results from direct retrieval simulations and previous ICA performed using an earlier version of the retrieval algorithm used for this study (Lan Gao, personal communication). More so than for nearly any other GV, the vertically resolved extinction plots clearly show that polarimeter-only retrievals should not be considered sufficient for constraining several GVs. Shown in Figure 6 are profiled uncertainties at a few selected *a priori* uncertainty scales. Note that even small increases in the *a priori* uncertainties of state vector

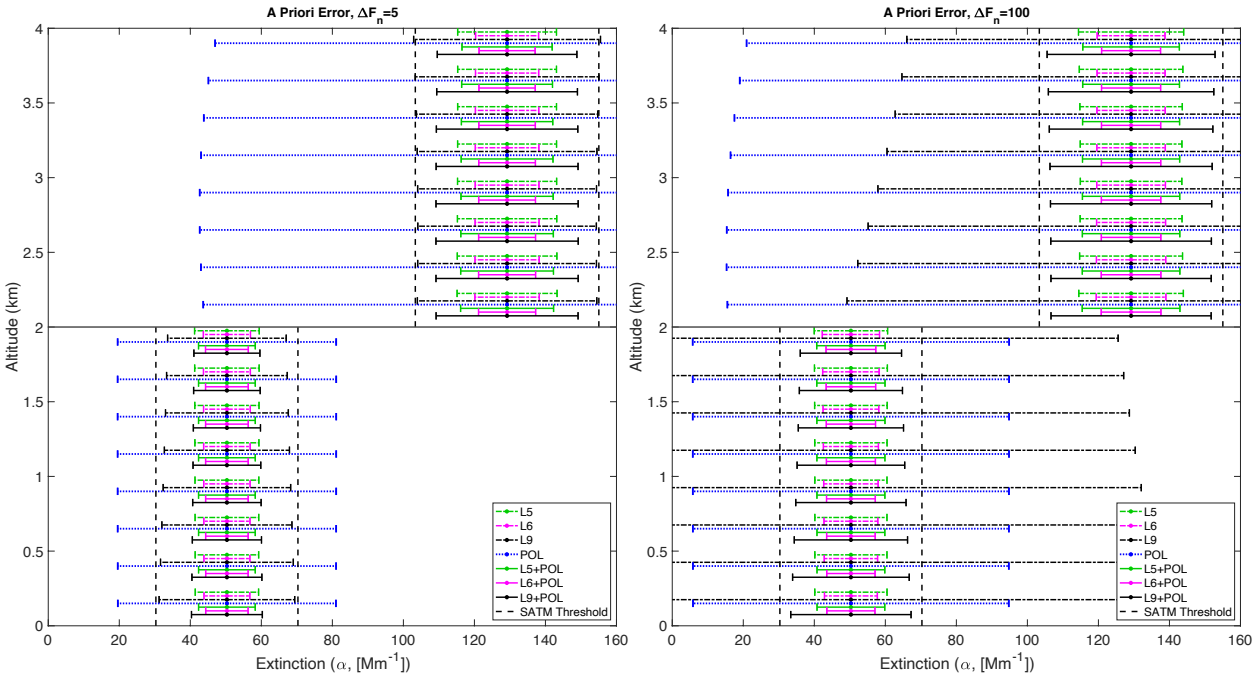

Figure 6: 250-meter vertically resolved truth values of extinction (colored circles) along with one-sigma error bars for a small state vector uncertainty window (left panel, n=5) and the maximum uncertainty window (right panel, n=100). Shown are four single-instrument setups and three instrument combinations: L5 (green dash-dot), L6 (magenta dash-dot), L9 (black dash-dot), POL (blue dotted), L5+POL (green solid), L6+POL (magenta solid; bolded), and L9+POL (black solid). The L6+POL combination is bolded for ease of identification. Vertical dashed black lines represent SATM one-sigma error requirement threshold. Horizontal solid black line shows boundary between boundary layer (BL) and free troposphere (FT).

parameters are enough for *a posteriori* extinction error to exceed SATM requirements. Nearly as poor at constraining this particular GV is L9, when used alone. As was the case for several other GVs considered, this fact should be somewhat obvious, considering L9 has no channels that receive extinction. It is important to note, however, that the complementary strengths of L9's ability to retrieve vertically resolved quantities and the polarimeter's ability to retrieve AOD well result in the L9+POL combination meeting the SATM requirements, regardless of how well-characterized the *a priori* environment is. Based on our results, the most predictive factor for relative ability to constrain extinction is number of α-channels used during retrieval. Inclusion of polarimetric data adds enough information content that — as mentioned previously — any of the combined instrument retrieval methods would be sufficient. However, unlike for several other GVs, lidar-only retrievals for L5 and L6 performed better than the L9+POL combination. Moreover, any method including L6 out-performed every other retrieval method, which lends

credence to the aforementioned positive impact on extinction retrieval abilities by adding more α-receiving channels to the lidar.

Before continuing, a few caveats regarding the extinction error profiles shown throughout this thesis must be addressed. One would expect extinction uncertainty to increase as altitude decreases; however, this is not a consistent trend in our results. Moreover, the author is aware that the opposite trend is present in vertically resolved extinction uncertainties for several of the instrument combinations used. Additionally, some instrument combinations' uncertainty values seem to remain static across all the various perturbations, in the FT, and in the BL. Closer analysis of these data has shown that the appearance of static extinction uncertainty values within the BL and FT is false. Moreover, the trend of increasing GV uncertainty as *a priori* state vector parameters' uncertainty increases does, in fact, hold true. As regards the seemingly unphysical trends in extinction error (many of which are more evident in extinction error profiles shown in later sections): finding a particular reason for this is not a focus of this thesis, nor is it within the scope of this thesis to discover. It is important to note that the differences level-to-level — even throughout one of the two major layers — are incredibly small for most instrument combinations, and for purposes of this thesis, the general trends that were expected are supported by the data.

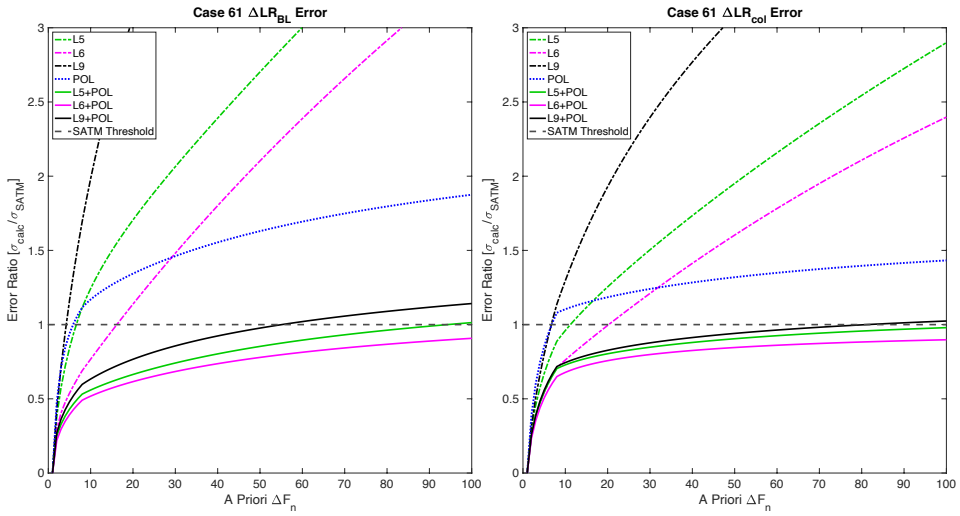

Figure 7: As in Fig. 1, except lidar ratio (ΔLR) error ratio.

Having discussed the extrinsic property of extinction, we now move to discussion of a derived intrinsic property, the LR. One can see relatively better ability to constrain over the whole column than in the BL, again likely related to the additional information content added by having twice as many lidar-provided extinction coefficient measurements in the column case. Additionally, the values of extinction are far higher in the FT than in the BL for Case 61, so the column values have more information content as well as higher values of extinction and lower values of backscatter, explaining the greater ability to constrain LR over the entire column. Although the polarimeter does not measure extinction or backscatter, its inclusion adds information content via its sensitivity to volume concentration. As long as the environment is somewhat well-characterized, any of the lidar-polarimeter combinations could adequately constrain LR error; however, for an entirely unknown environmental setup, using L9+POL could prove insufficient for accurate LR retrievals within the BL.

Table 9: Quantitative results to accompany the figures within Section 5.1. GVs considered are in the left-most column, and the seven different instrument combinations considered are listed across the top. Gray hatched boxes indicate instances in which an instrument combination was unable to constrain a particular GV over the entire perturbation domain (i.e., at a priori ΔFn, n=100). Green highlighted boxes indicate that the GV considered was adequately constrained over the entire domain, and the numbers give the GV uncertainty at the maximum perturbation.

GV	L5	L6	L9	POL	L5+POL	L6+POL	L9+POL
SSA, column					0.03	0.03	0.03
SSA, BL				0.04	0.02	0.02	0.02
g_{eff}, column				0.02	0.01	0.01	0.01
LR, column					11.25	10.31	
LR, BL						5.47	
τ_f, column					0.03	0.02	0.03
τ_f, BL					0.01	0.01	
τ, column	0.02	0.02		0.02	0.01	0.01	0.01
τ, BL	0.01	0.01			0.01	0.01	0.02
m_r, column							
m_r, BL							

5.2 Effects of increasing intensity measurement error

Before exploring the results gathered by increasing the error of the polarimeter's intensity measurement, we preemptively address one trend that remains true throughout this section. Because the measurement error being altered throughout this method does not factor into lidar calculations, all lines depicting lidar-only instrument choices within Section 5.2 should be constant. This result is to be expected and may, at first, seem trivial to include. However, these data were included so that direct comparisons can still be drawn across all methods used within this study, as well as between the relative efficacies of all instrument combinations. Moreover, the fact that the GV errors calculated using lidar-only setups are constant when intensity measurement error is altered does not preclude the ability of lidar-only retrievals to adequately constrain several

of the GVs considered. Knowledge of these lidar systems' capabilities will help inform future studies.

5.2.1 Conventional GV uncertainties

As in Section 5.1.1, we begin by analyzing AOD results. According to prior reasoning used in Section 5.1 and within Eqs. (3.12)-(3.14), we expect that instrument combinations including L5 and L6 will perform better than those with L9. This is due to L9's lack of channels which directly measure extinction. Additionally, we expect the polarimeter-only method to perform worse than all other instrument setups (except L9-only) because intensity is one of the two quantities measured by the polarimeter.

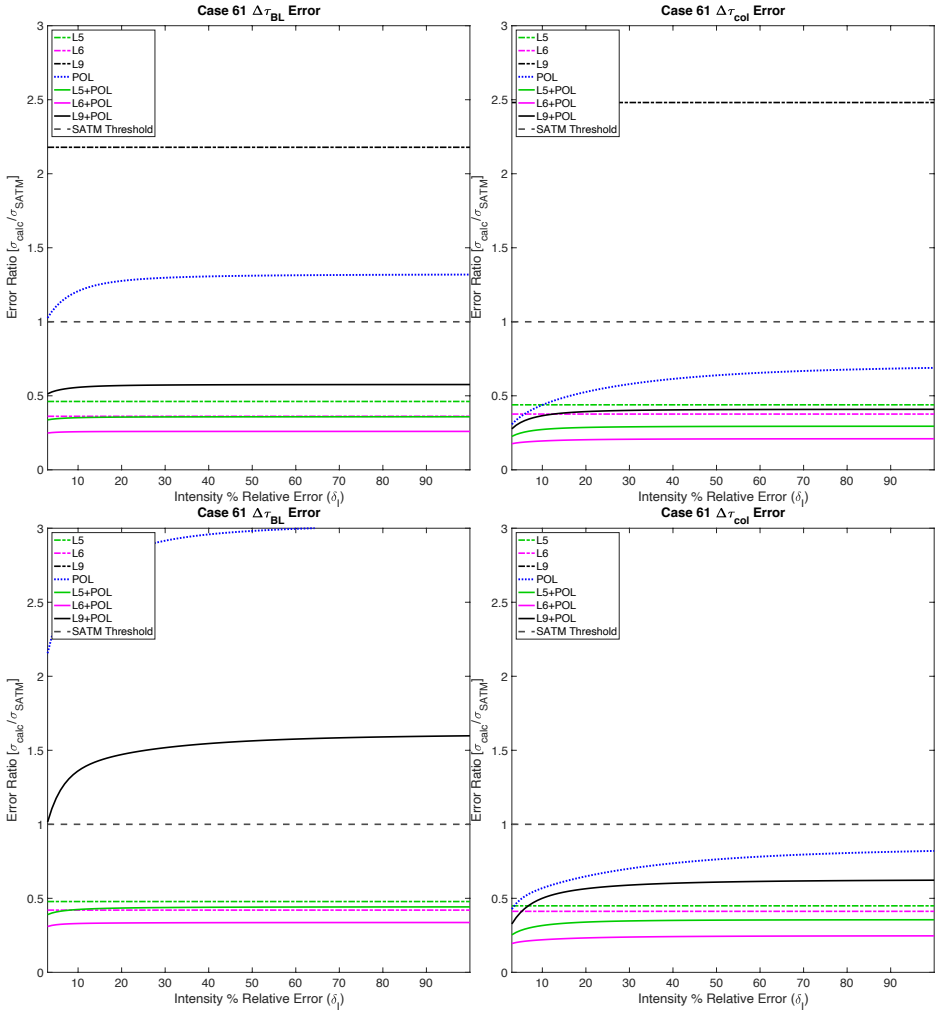

Figure 8: Test case total AOD (Δτ) error ratio versus percent relative error of intensity measurement (δ_I) for known environment setup boundary layer (BL, top left) and total column (col, top right) vertical ranges, as well as for unknown environment setup BL (bottom left) and col (bottom right) values. Results are shown for four single-instrument setups and three instrument combinations: L5 (green dash-dot), L6 (magenta dash-dot), L9 (black dash-dot), POL (blue dotted), L5+POL (green solid), L6+POL (magenta solid; bolded), and L9+POL (black solid).

At first glance, the impact of having better *a priori* knowledge of the environment (top row) is clearly noticeable for L9- and POL-only cases, as well as for the L9+POL setup. The expectation regarding L9's inability to constrain is confirmed, and the better performance of L5+POL and L6+POL relative to L9+POL is obvious. Additionally, the expectation that the polarimeter would be the most strongly impacted instrument when intensity measurement accuracy is reduced is confirmed by these results. Despite this, every instrument choice considered (except L9) is able to constrain column values of AOD, even with completely unreliable intensity

measurements and little specific knowledge about the environment being sampled. For BL retrievals, even very small increases in the intensity error cause the polarimeter-only setup to be unable to constrain AOD. The reasons for this likely have to do with the relatively lower aerosol number concentration in the BL as well as uncertainty added by two-way attenuation by the FT aerosols. It is notable that L5- and L6-only instrument choices are sufficient for constraining AOD in either spatial range, for this confirms our expectation that having receivers for extinction would give these instruments a higher capacity to constrain than L9.

Overall, the L5+POL and L6+POL combined instrument cases clearly perform better than others, when the environmental setup is known. Even when *a priori* knowledge is poor, these two combinations are among the best performers. However, L6+POL obviously performs best of all and is able to constrain AOD regardless of the reliability of intensity measurements or how well-characterized *a priori* knowledge is.

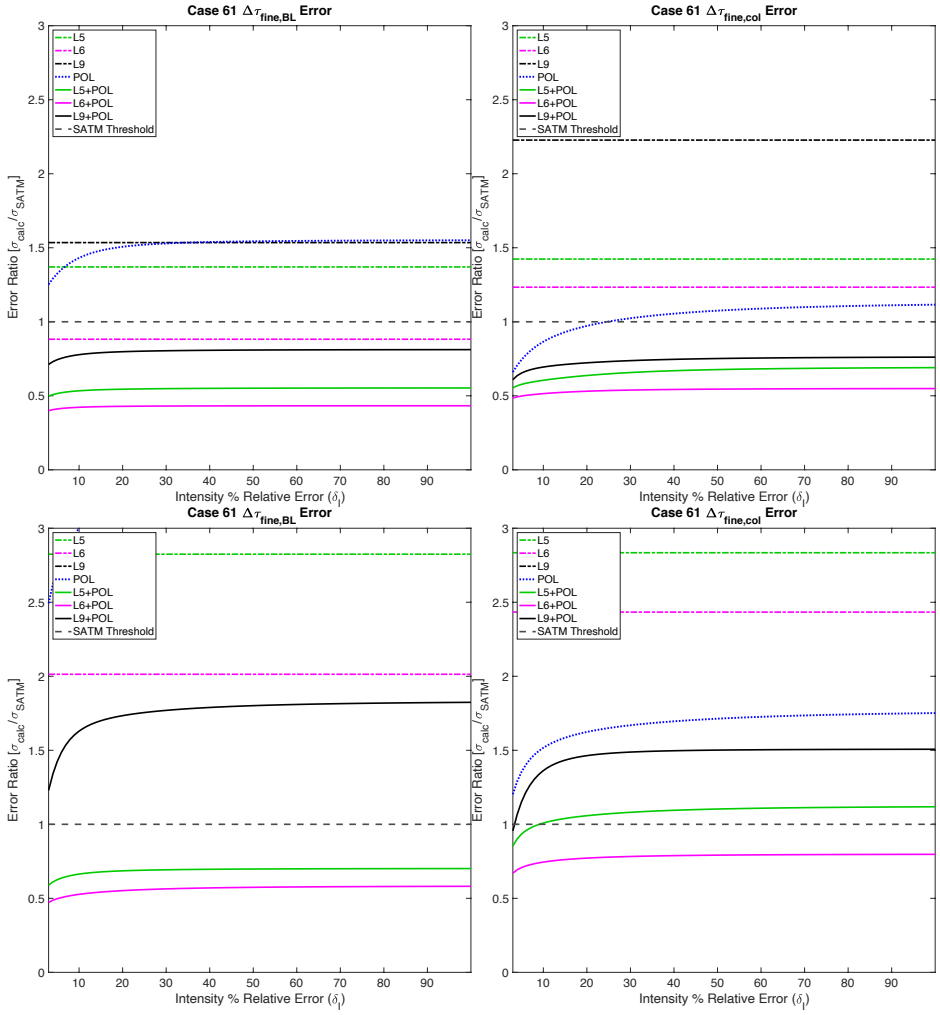

Figure 9: As in Fig. 8, except for fine mode AOD ($\Delta\tau_{fine}$) error ratio.

Considering the previous reasoning presented in Section 5.1.1 and the AOD results for all species included, the general relationships between different instruments' efficacies should not be surprising. When the environment is relatively well known, all three combined instrument methods can successfully constrain fine-mode AOD, regardless of the vertical range considered or whether any intensity information is available. More importantly, only the L6+POL method is able to constrain across any vertical range when the environmental setup and intensity measurements are unknown. Thus far, this is the first instance of the clear superiority of one instrument choice over all others tested.

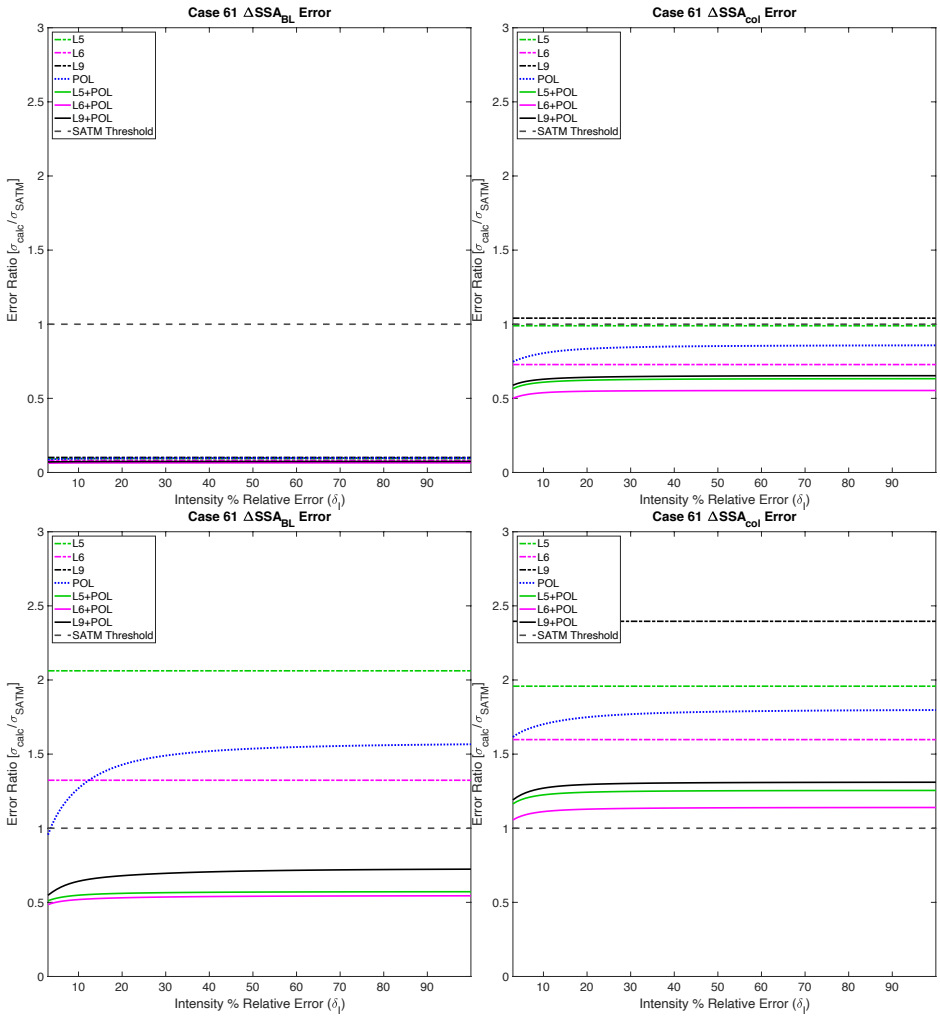

Figure 10: As in Fig. 8, except for single scattering albedo (ΔSSA) error ratio.

Though the effect of reducing *a priori* knowledge of the environment is obvious for several GVs investigated in this study, the SSA results show striking differences in the BL and total column effects, when comparing known and unknown environment cases. Once again, the L9-only retrieval is inferior to the retrieval abilities of all other instrument combinations. Because it is unable to constrain column-effective SSA even when intensity error is at a minimum and the environment is well-characterized, it bears no further discussion in this section.

Whereas the reasoning used in Section 5.1.1 mostly holds for these results, the abilities of the polarimeter are of particular interest when using this error perturbation method because it directly measures radiance. When the environment is known, the accuracy of intensity

measurements is unimportant for the constraint abilities of polarimeter-only and combined instrument methods. This implies that as long as DOLP measurements are within a reasonable margin of σ_{DOLP} (and environmental characteristics are well-known), SSA may be adequately constrained in both vertical ranges by all of the combined instrument configurations, as well as by the polarimeter only.

Having discussed relative constraint abilities when the environment is known, we turn now to comparing those results to the unknown environment case. In this more generalized case, all three combined instrument retrieval methods still perform well in the BL; however, no instrument(s) show any ability to constrain column values of SSA when the environment is unknown. This specific result is telling, in that it points to the possibility of even the best overall instrument combination (L6+POL) failing to meet the column SSA SATM requirement any time the environment is unknown, as it is defined within this study. Further research is needed to more precisely assess how well-characterized the environment must be for column-effective values of SSA to be adequately constrained without any radiance information so that specific maximum error thresholds of the polarimeter's radiance measurement capabilities may be reported.

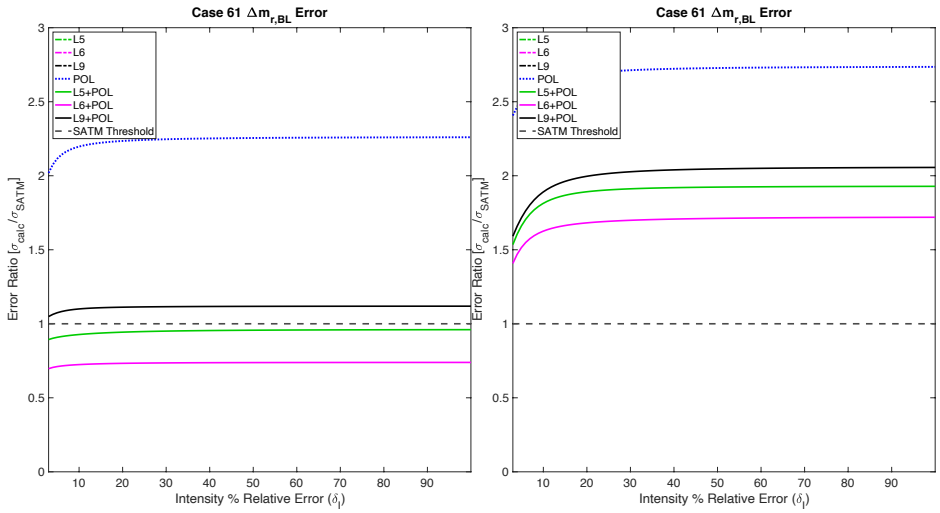

Figure 11: Test case real part of refractive index (Δm_r) error ratio versus percent relative error of intensity measurement (δ_I) for known environment setup boundary layer (BL, left panel) and unknown environment setup BL (right panel) ranges. Results are shown for four single-instrument setups and three instrument combinations: L5 (green dash-dot), L6 (magenta dash-dot), L9 (black dash-dot), POL (blue dotted), L5+POL (green solid), L6+POL (magenta solid; bolded), and L9+POL (black solid).

In analyzing Fig. 11, it is immediately obvious that only those combined instrument retrieval methods that include extinction channels in their lidar systems are able to constrain m_r within the BL. Moreover, this is only the case when the environment is well-characterized, which is clear when looking at the right panel of Fig. 11. As was the case with column-effective values of SSA, finding the extent to which one must know specific environmental parameters to adequately constrain m_r bears further research. Conspicuously missing from this section are the results for known/unknown environment results for total column m_r error. The reason for this is that no instrument configuration tested was able to achieve SATM requirements, even when the environment was well-known and the intensity measurement error was at its reference error level. These results show the importance of accurate radiance measurements as a necessity for adequately retrieving our entire suite of GVs.

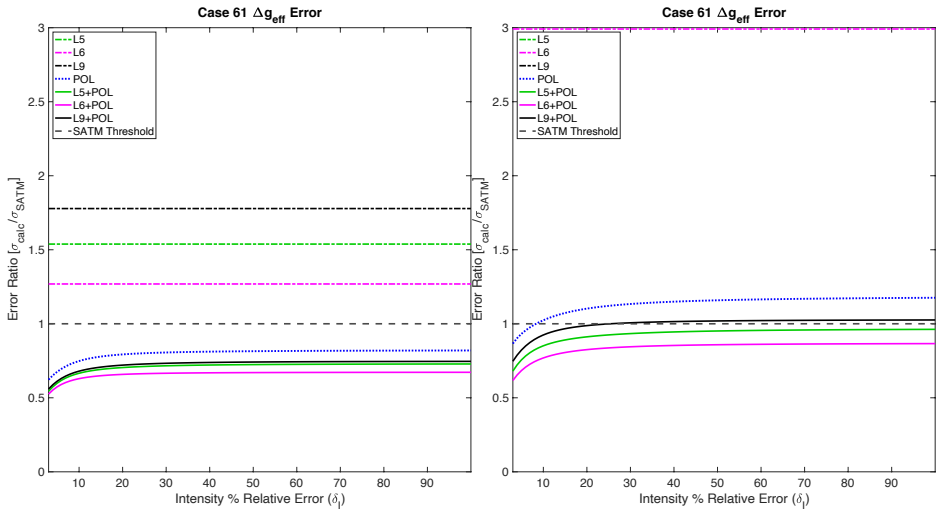

Figure 12: Column-effective asymmetry parameter (Δg_{eff}) error ratio versus percent relative error of intensity measurement (δ_I) for known (left panel) and unknown (right panel) environment setups.

Though we have already seen in Section 5.1.1 that *a priori* knowledge of the environment is unnecessary for the polarimeter and all of its combinations with lidar to adequately constrain g_{eff}, we have yet to parse out how lack of measurement accuracy may affect this ability. Viewing Fig. 12, the left panel shows good agreement with results from Section 5.1.1, in terms of ability to constrain g_{eff} when the environment is known. The right panel of the figure, on the other hand, reveals the importance of accurate radiance measurements for retrieval capabilities of the polarimeter-only and L9+POL instrument choices. An important takeaway is that neither environmental knowledge or intensity measurement is necessary for adequate constraint of g_{eff} by L6+POL or L5+POL combined retrieval methods.

5.2.2 Lidar-sensitive GV uncertainties

To begin analysis of the effects of increasing intensity measurement error on GVs used in the lidar equation (Liou 2002), we start by viewing Fig. 13, which gives the extinction profile for our test case with related uncertainties for the fifth ($n = 5$) and hundredth ($n = 100$) perturbations of intensity error in the known environment (top row) and unknown environment (bottom row) cases. We know from Section 5.1.2 that polarimeter-only and L9-only cases are unable to constrain

the entire extinction profile, even when the environment is rather well-characterized. The known environment results depicted in Fig. 13 bear out these findings. Overall, the impact of having little *a priori* environmental knowledge is apparent for L9-only, polarimeter-only, and the combination of these two instruments. All other instrument combinations (including L5-only and L6-only) easily constrain the entire vertically resolved column of extinction values, regardless of *a priori* knowledge. In other words, it appears that without significant knowledge of the environment or accurate intensity measurements, combined instrument retrievals which include a lidar system that has extinction receivers show clear advantage over other retrieval methods, in addition to meeting SATM requirements throughout the entire atmospheric column.

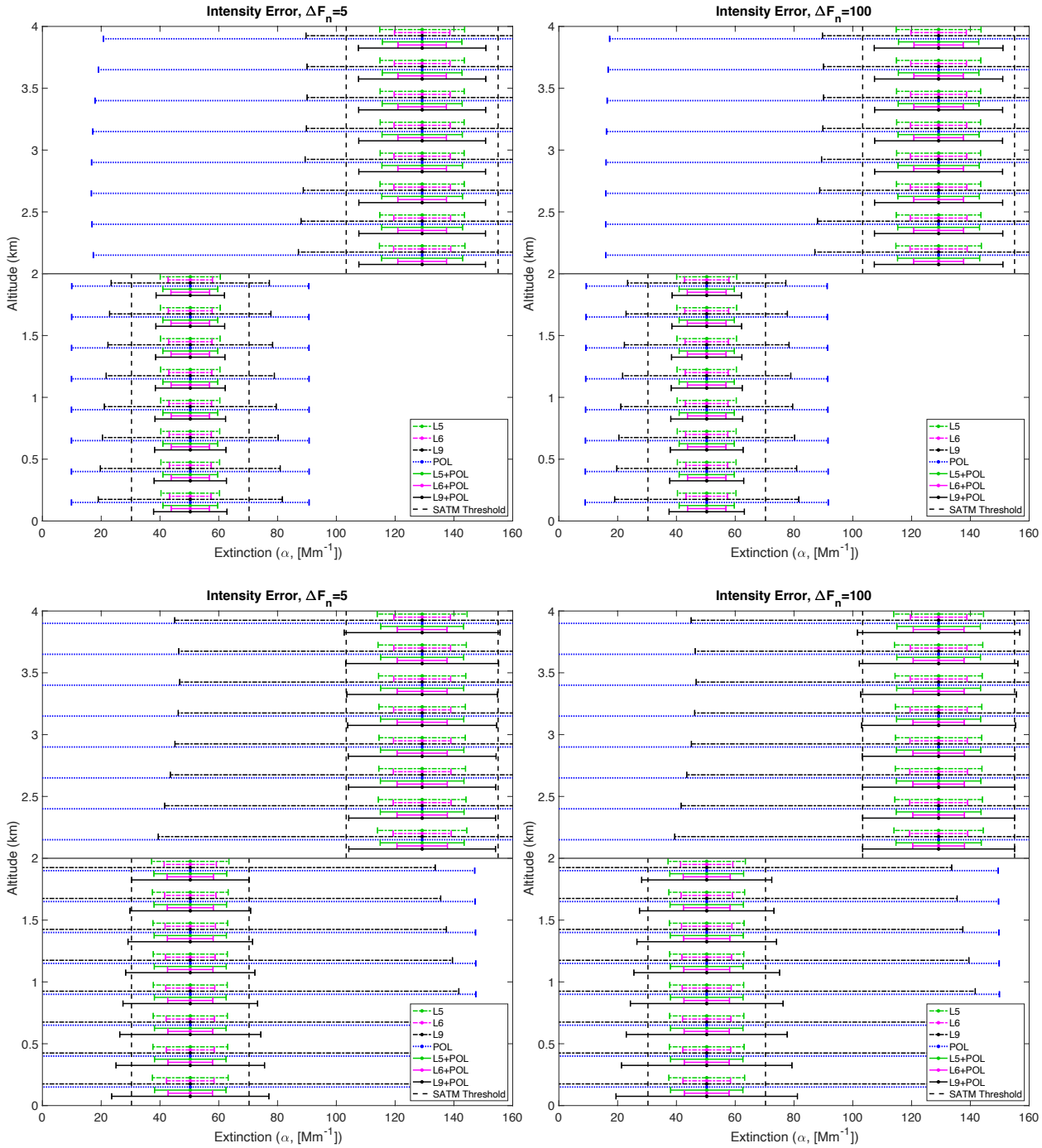

Figure 13: 250-meter vertically resolved truth values of extinction (colored circles) along with one-sigma error bars for known environment setup small intensity measurement uncertainty window (top left, n=5) and maximum uncertainty window (top right, n=100). Also shown are results for unknown environment setup small intensity measurement uncertainty window (bottom left) and maximum uncertainty window (bottom right). Results are given for four single-instrument setups and three instrument combinations: L5 (green dash-dot), L6 (magenta dash-dot), L9 (black dash-dot), POL (blue dotted), L5+POL (green solid), L6+POL (magenta solid; bolded), and L9+POL (black solid). The L6+POL combination is bolded for ease of identification. Vertical dashed black lines represent SATM one-sigma error requirement threshold. Horizontal solid black line shows boundary between boundary layer (BL) and free troposphere (FT).

Although the reasoning behind the differences for ability to constrain LR in the column and BL is somewhat more complex than is the case for several other GVs (see Section 5.1.2), the importance of accurate intensity measurements to retrieval of LR has not yet been discussed. The first thing to note is that results showing the unknown environment case are not shown here, for no instrument retrieval was able to constrain LR in the BL or total column. Additionally, the relative superiority of combined instrument retrievals is evident. From simply viewing the error saturation plots, it is clear that these combined retrievals adequately constrain LR when the environment is somewhat well-characterized; however, the shallowness of these curves show that the accuracy of intensity measurements is unimportant to the constraint ability of combined instrument retrieval of LR.

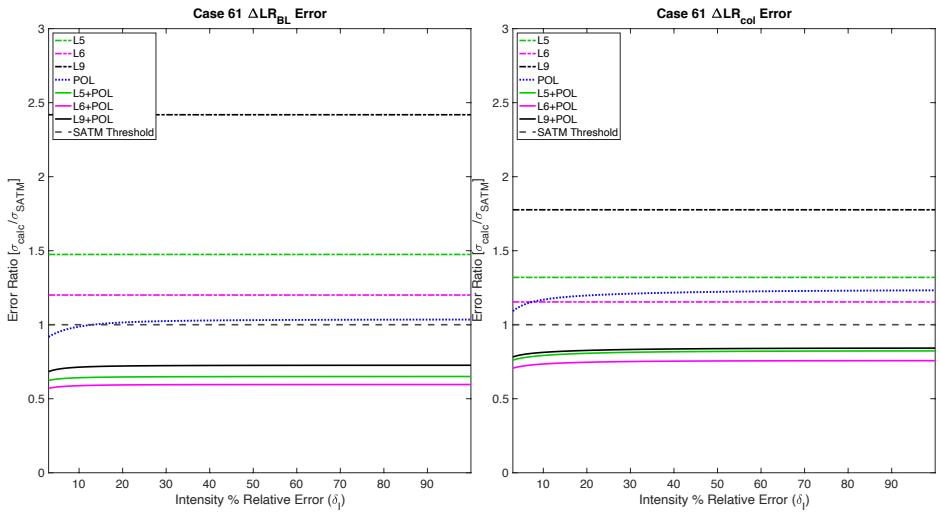

Figure 14: Test case lidar ratio (ΔLR) error ratio versus percent relative error of intensity measurement (δ_I) for known environment setup boundary layer (BL, left panel) and total column (col, right panel) ranges. Results are shown for four single-instrument setups and three instrument combinations: L5 (green dash-dot), L6 (magenta dash-dot), L9 (black dash-dot), POL (blue dotted), L5+POL (green solid), L6+POL (magenta solid; bolded), and L9+POL (black solid).

Table 10: Quantitative results to accompany the figures within Section 5.2 that depict "known environment" results. GVs considered are in the left-most column, and the seven different instrument combinations considered are listed across the top. Gray hatched boxes indicate instances in which the three measurements other than intensity were unable to constrain a particular GV at nearly any point over the entire perturbation domain (i.e., when $\delta_I > 3\%$). Green highlighted boxes indicate that the GV considered was adequately constrained over the entire domain, and the numbers give the GV uncertainty at the maximum perturbation ($\delta_{I,max}$). Yellow highlighted boxes indicate that an instrument combination was able to constrain GV uncertainty over a portion of the perturbation domain, and numbers within give the value of δ_I at which the GV may no longer be constrained by the other three measurements.

GV	L5	L6	L9	POL	L5+POL	L6+POL	L9+POL
SSA, column	0.04	0.03		0.03	0.03	0.02	0.03
SSA, BL	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01
g_{eff}, column				0.02	0.01	0.01	0.01
LR, column					9.45	8.69	9.67
LR, BL				12.80%	3.92	3.59	4.37
τ_f, column				24.55%	0.02	0.02	0.02
τ_f, BL		0.02			0.01	0.01	0.02
τ, column	0.02	0.01		0.03	0.01	0.01	0.02
τ, BL	0.01	0.01			0.01	0.01	0.01
m_r, column							
m_r, BL					0.02	0.02	

Table 11: As in Table 10, but for figures within Section 5.2 that depict "unknown environment" results.

GV	L5	L6	L9	POL	L5+POL	L6+POL	L9+POL
SSA, column							
SSA, BL				3.00%	0.02	0.02	0.03
g_{eff}, column				7.90%	0.02	0.02	24.55%
LR, column							
LR, BL							
τ_f, column					8.88%	0.03	3.00%
τ_f, BL					0.02	0.01	
τ, column	0.02	0.02		0.03	0.01	0.01	0.02
τ, BL	0.01	0.01			0.01	0.01	
m_r, column							
m_r, BL							

5.3 Effects of increasing DOLP measurement error

5.3.1 Conventional GV uncertainties

From the outset, it is clear that AOD uncertainties are sensitive to increases in DOLP error, though the added uncertainty saturates quickly. This implies that σ_{DOLP} is only useful as a constraint for AOD when its value is very small (i.e., close to the reference uncertainty value). As regards the different environmental scenarios, the impact of *a priori* knowledge is noticeable for both vertical ranges, but is far more noticeable in the BL. One takeaway is that L9+POL is nearly useless in the unknown environment BL case, making it the only combined instrument retrieval with issues constraining in this range. More importantly, neither accurate DOLP measurements nor *a priori* knowledge are necessary for accurate AOD retrievals. These trends, taken together, point to the primary importance of extinction measurement accuracy in the ability to constrain AOD, though this will be examined further in Section 5.3.

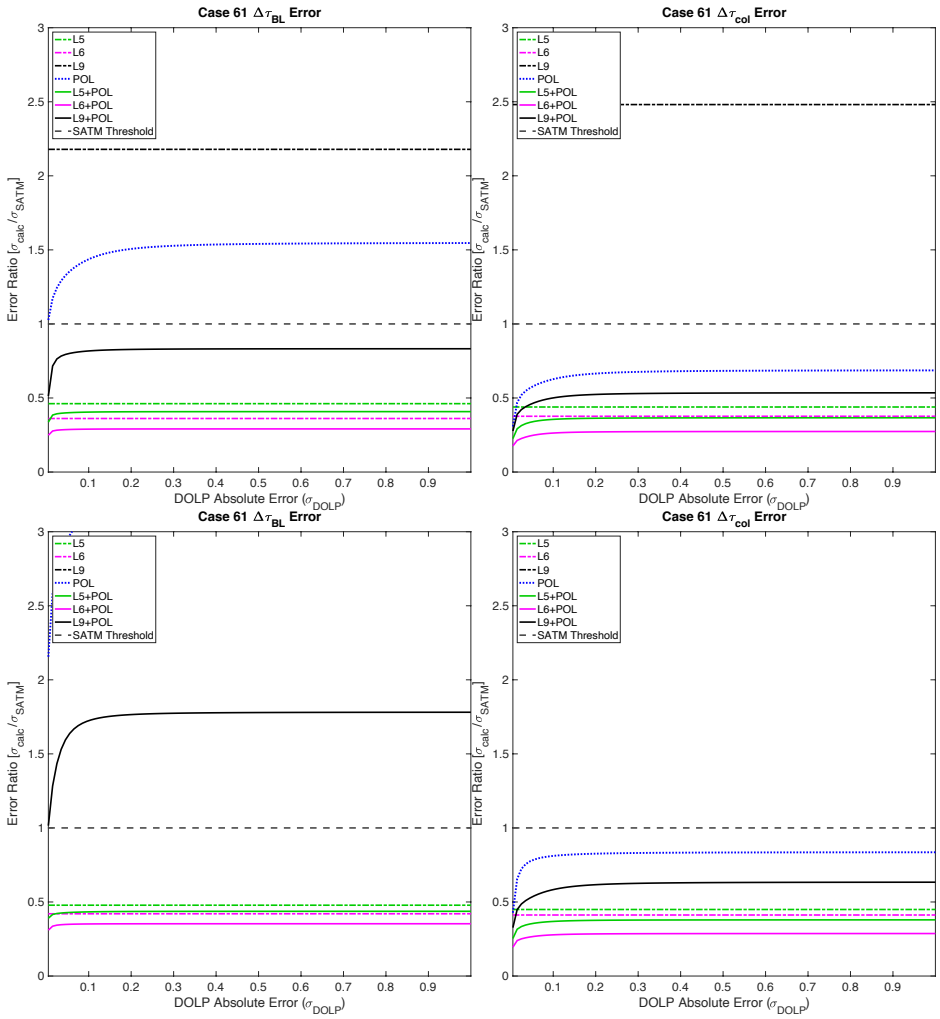

Figure 15: As in Fig. 8, except Δτ error ratio is plotted versus absolute error of degree of linear polarization measurement (σ_{DOLP}).

As far as fine-mode AOD is concerned, some trends remain the same. For example, the AOD error still saturates rather quickly, hinting at some sensitivity to σ_{DOLP}. Additionally, it is known that DOLP measurements have great sensitivity to fine-mode particles, which further explains the quick saturation of error. One factor that likely causes the fine-mode AOD to have worse retrievals than total AOD is the sensitivity of DOLP to AOD (Boesche et al. 2006). Because of the large differences in total versus fine-mode AOD, the overall abilities of our instrument retrievals appear to be worse at retrieving fine-mode AOD. Given Fig. 16, one can see that L5+POL and L6+POL combined retrievals perform best of any instrument choices; however, these

combinations are only able to constrain fine-mode AOD when the environment is relatively well-characterized. If the environment is unknown, σ_{DOLP} must also be near its reference uncertainty value to yield accurate fine-mode AOD retrievals. That is, DOLP accuracy is of major importance to the retrieval of this GV, which is perhaps one of the most important GVs to retrieve accurately (particularly in the BL) for purposes of improving air quality.

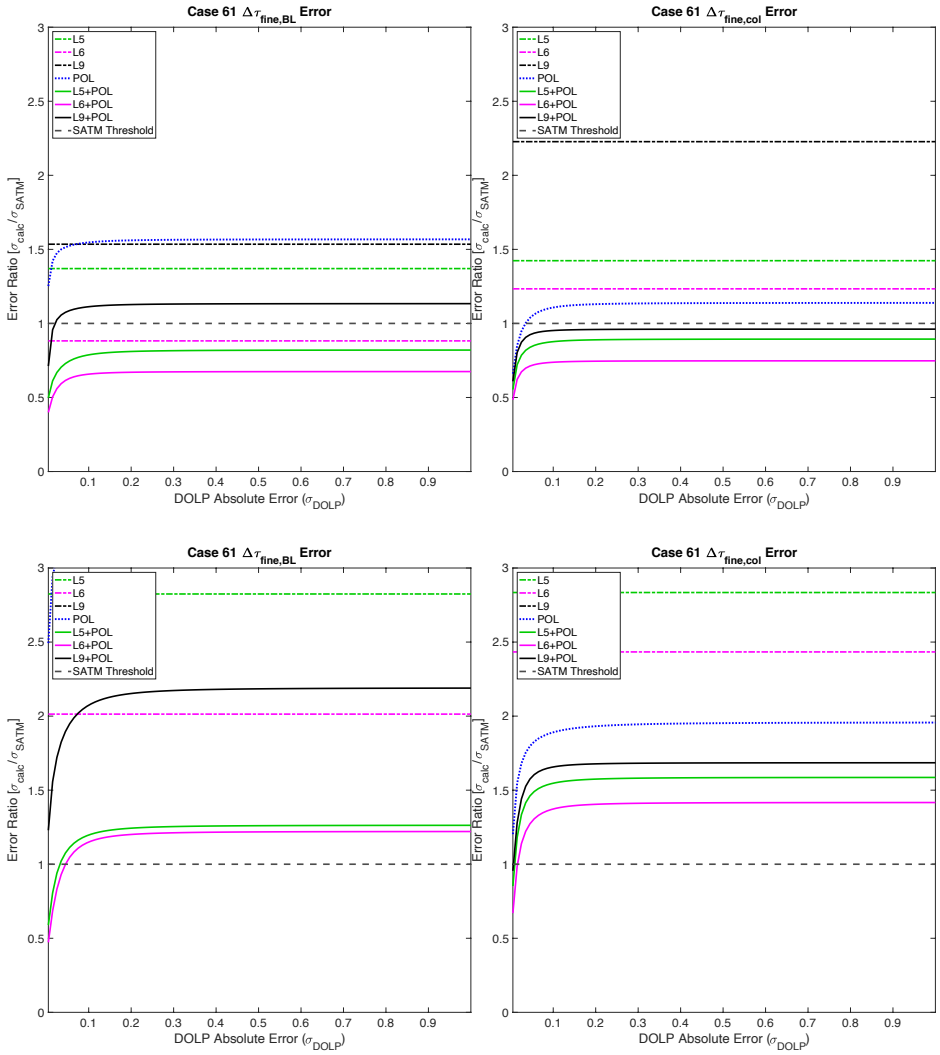

Figure 16: As in Fig. 15, except for fine mode AOD ($\Delta\tau_{fine}$) error ratio.

When analyzing Fig. 17, one notices many similar characteristics to Fig. 10. Again, we see a quick saturation of error for all instrument choices in all of the different methods shown. Several reasons cause this, including the sensitivity of DOLP to AOD, which is used in the calculation of

SSA, and real part of refractive index (Boesche et al. 2006). For these reasons, in the known environment case, every instrument combination is able to constrain SSA within the BL, and all but L9-only are able to constrain column-effective values of SSA. Once again, the impact of *a priori* on retrieval capabilities is marked.

Though all the combined instrument retrieval methods are able to constrain SSA within the BL in the unknown environment case, none of the instruments are capable of the same over the entire column. The reasons for this are likely somewhat interconnected in the impacts they have. First, DOLP is sensitive to the real part of refractive index, but it is the imaginary part of refractive index becomes significantly more important when the FT levels containing absorbing smoke particles are added in for the total column calculations. Additionally, the particle concentrations in the FT in Case 61 are much higher (at least one order of magnitude) for both size modes, adding larger quantities with associated uncertainties.

Overall, the impact of increasing σ_{DOLP} has relatively little effect on the ability of combined instrument methods to retrieve well values of SSA when the environment is well-known. That is, DOLP measurements can be essentially nonexistent in this instance, and as long as the other three measurements are operating at their reference uncertainties, SSA can be constrained, regardless of vertical depth considered. When less *a priori* is available, combined instrument retrievals appear to be sufficient to constrain BL values of SSA. However, further research including more of our Canonical Cases is necessary to determine what factor or combination of factors is causing the lack of ability to constrain SSA over the entire column when *a priori* isn't well-known. It is possible that these results are tied strongly to the particular aerosol loading chosen for Case 61, so the addition of more experimental cases in the future should offer more insight regarding the generalizability of the trends shown here.

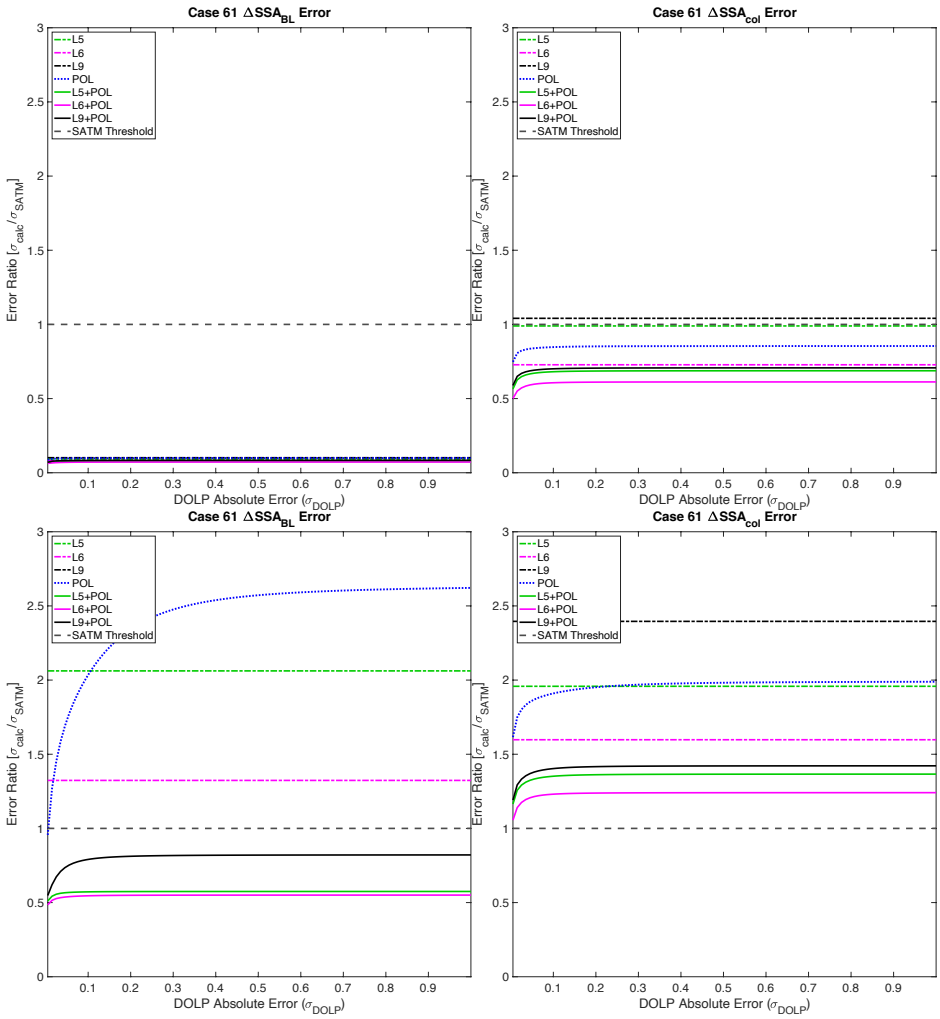

Figure 17: As in Fig. 15, except showing single scattering albedo (ΔSSA) error ratio.

Whereas the ability to constrain GVs has, thus far, required a bit of explanation, there is relatively little to say about the abilities of our various instrument configurations to constrain m_r when σ_{DOLP} is increased even slightly. As noted previously, DOLP is very sensitive to RRI (Boesche et al. 2006). Moreover, these results seem to show the necessity of minimizing σ_{DOLP} as much as possible to retrieve RRI with any kind of accuracy and, thereby, also show that radiance and lidar measurements (even in combination) are unable to adequately constrain RRI. It should be noted that Fig. 18 shows the known environment case, so again, more work must be done to

find out exactly how well-characterized the environment must be to allow reasonable retrieval capabilities that also allow for the measurement error of DOLP to vary more.

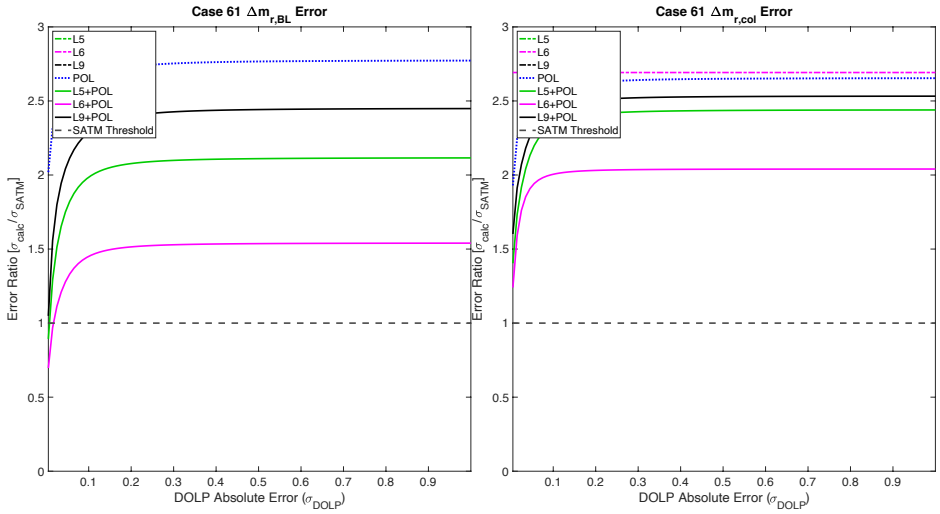

Figure 18: Real part of refractive index (Δm_r) error ratio versus absolute error of DOLP measurement (σ_{DOLP}) for known environment setup boundary layer (BL, left panel) and total column (col, right panel) ranges. Results are shown for four single-instrument setups and three instrument combinations: L5 (green dash-dot), L6 (magenta dash-dot), L9 (black dash-dot), POL (blue dotted), L5+POL (green solid), L6+POL (magenta solid; bolded), and L9+POL (black solid).

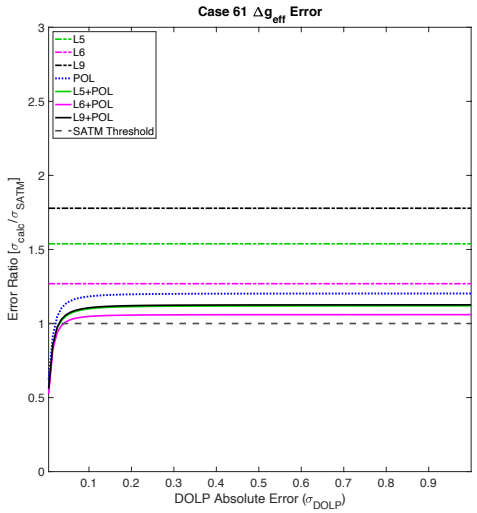

Figure 19: Column-effective asymmetry parameter (Δg_{eff}) error ratio versus absolute error of DOLP measurement (σ_{DOLP}) for the known environment setup.

When analyzing Fig. 19, it should be no surprise that the ability to constrain asymmetry parameter is greatly affected by the magnitude of DOLP measurement error. Asymmetry parameter is strongly tied to size distribution, and DOLP is sensitive to that characteristic. So, even if all three other measurements are within their reference errors, even small changes to DOLP

measurement accuracy have a large effect on the ability to constrain asymmetry parameter. Note that the only panel of Fig. 19 shows the known environment case, which accentuates even more the importance of minimizing σ_{DOLP}.

5.3.2 Lidar-sensitive GV uncertainties

As in previous sections, only a few examples of extinction error vertical profiles are shown here: one in which DOLP error is still close to the reference uncertainty and the most extreme case, in which the magnitude of σ_{DOLP} makes the DOLP measurement essentially useless. Viewing the top row of Fig. 20, which shows the results from the known environment case, one can see that there appears to be very little difference in the two plots, despite the large difference in σ_{DOLP} between the two retrieval trials. This implies that — at least when the environment is well-characterized — the accuracy of DOLP measurements is wholly unimportant to the ability of combined instrument retrieval methods to adequately constrain vertically resolved extinction profiles.

For reasons already explained in Section 5.1.2, there is little reason to expect the polarimeter-only retrievals to be able to constrain the vertically resolved extinction profile, even when there is a considerable amount of *a priori* knowledge. Additionally, for the same reasons as discussed in Section 5.2.2, L9-only should not be able to constrain extinction in any instance. The results bear out these expectations.

Now considering the bottom half of Fig. 20, which shows results from the same magnitude scale factor trials (but for the unknown environment setting), one can see similar trends as seen in Fig. 13 (see Section 5.2.2). One notable difference is that the L9+POL combination performs even worse when σ_{DOLP} is increased than it did when δ_L was increased. Though it is possible that

accurate DOLP measurements are more important to constraining extinction error than are accurate intensity measurements, it is more likely that this is due to the relatively larger amount σ_{DOLP} was increased from trial to trial. This conclusion is supported by the fact that the combined instrument retrievals that include extinction channels showed no significant difference in values between correlating plots depicting increases of σ_{DOLP} and δ_L.

To summarize, the magnitudes of the effects on extinction error caused by increasing σ_{DOLP} and δ_L may not be directly comparable. That is, in the absence of a true one-to-one scale factor comparison, this conclusion cannot yet be stated as fact, though future research should attempt to clarify this relationship. More importantly, the impact of *a priori* knowledge can still be seen in the widening of extinction error bars between the known and unknown cases for all three combined instrument methods. Additionally, the inferiority of L9+POL instrument combination relative to the other combined instrument retrieval methods has been shown, once again.

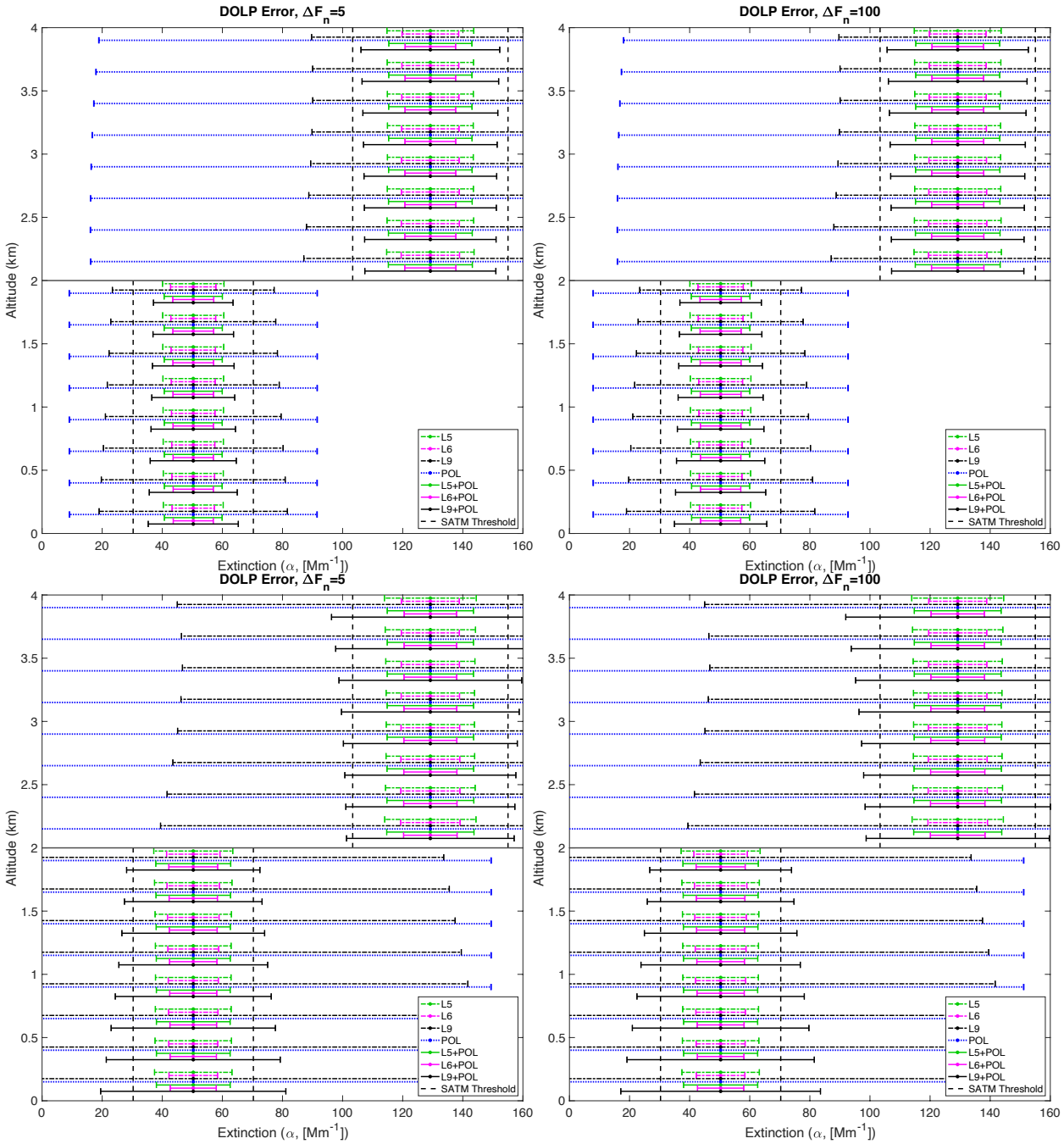

Figure 20: As in Fig. 13, except results are from perturbing a priori DOLP measurement error (σ_{DOLP}).

As we examine the effects of increasing σ_{DOLP} on the ability to retrieve LR, it should first be noted that, once again, only the results from the known environment case are shown, for even before DOLP measurement error is perturbed beyond its initial reference error, no instrument configuration is able to constrain LR in either vertical range. Looking at the known environment

case in Fig. 21, one can see that the only instrument configurations that can constrain LR irrespective of the depth considered are the combined instrument setups. Therefore, the superiority of these methods for LR retrieval — at least for the environment in Case 61 — is evident. It is not surprising that polarimeter-only retrievals are of relatively poor quality for LR. For one, the results of widening *a priori* error (see Section 5.1.2) have already shown that the environment must be relatively well-characterized for polarimeter to constrain LR by itself. By adding larger DOLP error, this ability suffers even more, especially considering that DOLP is measured by the polarimeter.

Considering the shallowness of the combined instrument retrieval curves — in addition to the fact that each of these methods is able to constrain LR with essentially no reliable DOLP measurement — show that σ_{DOLP} is unimportant to the LR constraint abilities of the combined instrument configurations. In other words, as long as the other three measurements (in particular, those for extinction and backscatter) are retrieving information at the reference uncertainty, LR can be constrained by combined instrument methods when the environment is somewhat well-characterized. The lack of constraint ability in the unknown environment setup shows the importance of having *a priori* knowledge and/or a polarimeter that has higher DOLP measurement precision.

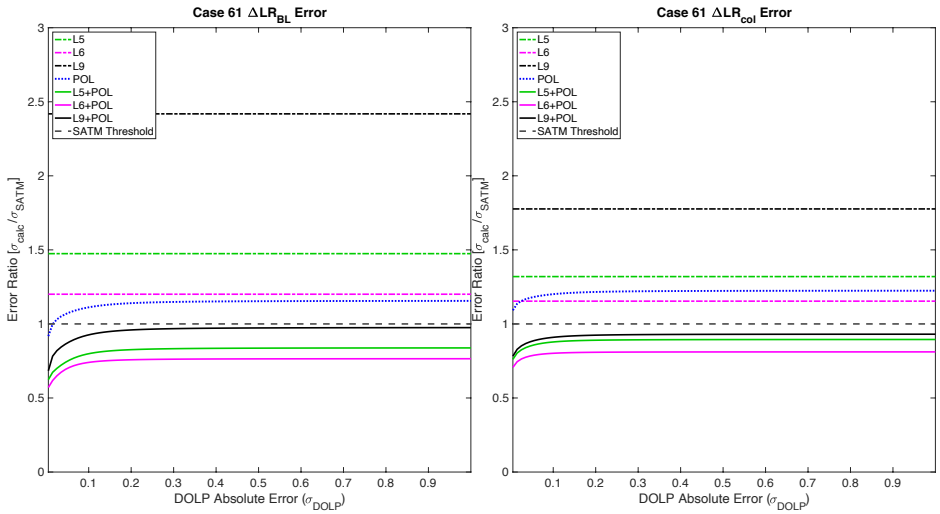

Figure 21: As in Fig. 14, except results are from perturbing a priori DOLP measurement error (σ_{DOLP}).

Table 12: Analogous to Table 10, but for figures within Section 5.3 depicting "known environment" results retrieved by perturbing σ_{DOLP}. It follows that GV uncertainties within green boxes show how well the three non-DOLP measurements can constrain the respective GVs. Additionally, yellow boxes give the magnitude of σ_{DOLP} at which a GV may no longer be constrained by the other three measurements.

GV	L5	L6	L9	POL	L5+POL	L6+POL	L9+POL
SSA, column	0.04	0.03		0.03	0.03	0.02	0.03
SSA, BL	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01
g_{eff}, column				0.015	0.025	0.035	0.025
LR, column					10.28	9.31	10.69
LR, BL				0.015	5.05	4.61	5.87
τ_f, column				0.025	0.03	0.02	0.03
τ_f, BL		0.02			0.02	0.01	0.015
τ, column	0.02	0.01		0.03	0.01	0.01	0.02
τ, BL	0.01	0.01			0.01	0.01	0.02
m_r, column							
m_r, BL					0.005	0.015	

Table 13: As in Table 12, except for figures within Section 5.3 depicting "unknown environment" results.

GV	L5	L6	L9	POL	L5+POL	L6+POL	L9+POL
SSA, column							
SSA, BL				0.005	0.02	0.02	0.03
g_{eff}, column				0.005	0.005	0.005	0.005
LR, column							
LR, BL							
τ_f, column					0.005	0.015	0.005
τ_f, BL					0.025	0.045	
τ, column	0.02	0.02		0.03	0.01	0.01	0.02
τ, BL	0.01	0.01			0.01	0.01	
m_r, column							
m_r, BL							

5.4 Effects of increasing extinction measurement error

As we move into discussion of the effects of increasing extinction measurement error on the ability of various instrument configurations to adequately constrain our suite of GVs, we set forth a few expectations for this section. In Sections 5.2-5.3, there was an expectation that lidar-only retrievals would have constant results for those perturbation methods, considering intensity and DOLP observations are measured by the polarimeter. Similarly, the expectation is that in Sections 5.4-5.5, polarimeter results will be constant, for extinction and backscatter are only measured by the lidar systems in question. Moreover, L9 should also show constant error ratio values within this section, because it uses only backscatter measurements.

5.4.1 Conventional GV uncertainties

First considering AOD, there are few surprises. Our expectations regarding L9 and POL are both confirmed, and both L5 and L6 are unable to constrain error because of their dependence on extinction channels. All combined instrument retrievals are clearly superior in constraint ability,

despite the lack of extinction measurement. Additionally, the shallowness of the error saturation curves for the combined instrument retrievals show a lack of sensitivity to extinction measurement error. As discussed in Section 5.3.1, this points to the importance of DOLP measurement in the ability to constrain AOD. Moreover, it is important to note that Fig. 22 is showing only the unknown environment setup, meaning that L5+POL and L6+POL are able to constrain AOD when extinction measurement is useless and *a priori* knowledge of the environment is minimal.

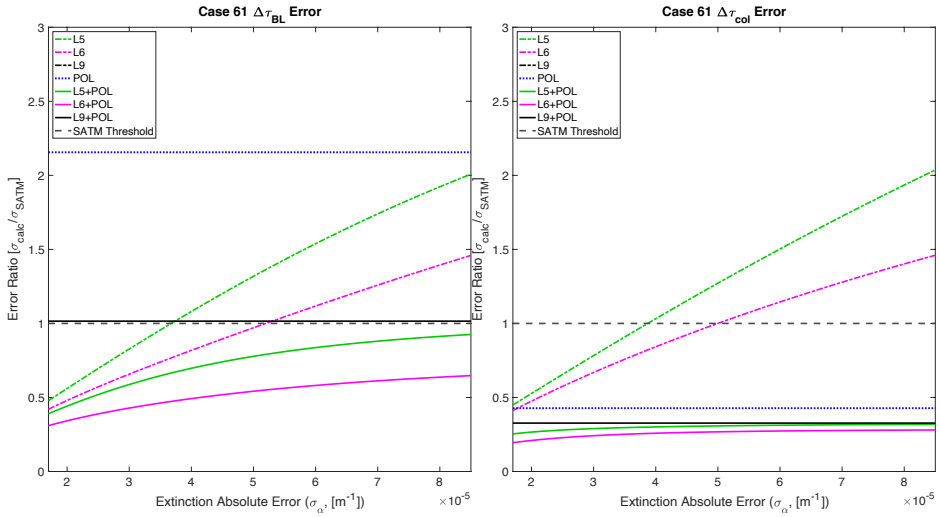

Figure 22: Total AOD ($\Delta\tau$) error ratio versus absolute error of extinction measurement (σ_{α}) for unknown environment setup boundary layer (BL, left panel) and total column (col, right panel) ranges. Results are shown for four single-instrument setups and three instrument combinations: L5 (green dash-dot), L6 (magenta dash-dot), L9 (black dash-dot), POL (blue dotted), L5+POL (green solid), L6+POL (magenta solid; bolded), and L9+POL (black solid).

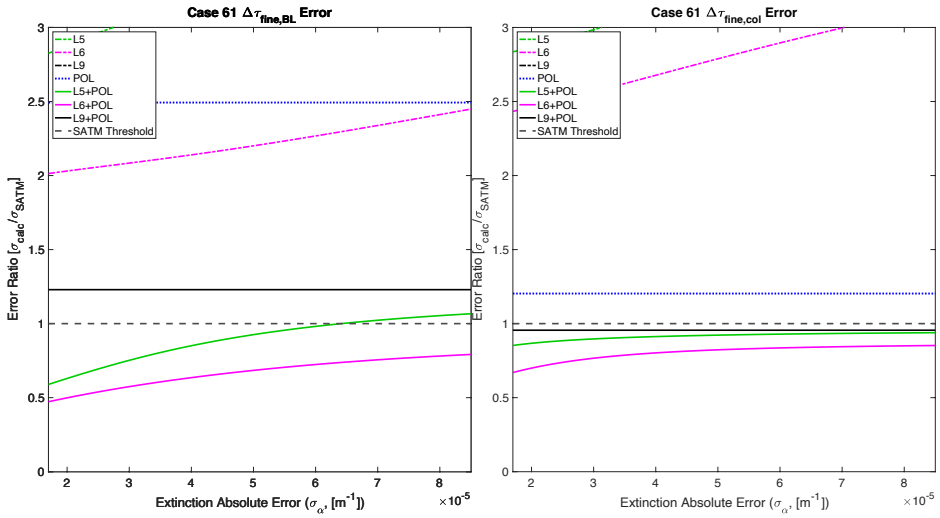

Figure 23: As in Fig. 22, except showing fine mode AOD ($\Delta\tau_{fine}$) error ratio.

Though Fig. 23 doesn't warrant much discussion, we will mention a few of the trends that are constant to further establish consistency of results and reasoning. First, we are again only showing the unknown environment setup, showing that *a priori* environmental knowledge is unnecessary for L6+POL to constrain AOD_{fine}, even when extinction measurements are not included. Moreover, this figure shows a clear advantage of L6+POL over all other instrument combinations, when considering both vertical ranges for which the variable must be constrained to adequately meet SATM requirements.

As we turn to discussion of the effects of increasing σ_α upon the retrieval uncertainties of SSA in the BL and column, first note that Fig. 24 includes the known and unknown environment methods. One would expect that the retrievals of BL aerosols would be relatively good, considering the BL is loaded with sea salt particles, which have a much higher value of SSA than the relatively darker smoke particles in the FT do. Even in the known environment case (top row), adding FT layers into the error propagation calculation greatly reduces the retrieval capabilities of all the instrument combinations considered. This is likely due to the dependence of SSA on extinction. By factoring in two-way attenuation of light through twice as much distance as in the BL case, it seems that the uncertainty of SSA would increase by virtue of the uncertainty of extinction increasing with every subsequent layer added. Moreover, the sample volume added by including the FT in this case contains smoke particles, which add to the magnitude of extinction values. The ability of sensing column-effective SSA becomes complicated, considering we are adding a large concentration of absorbing particles to a small concentration of particles with extremely high albedo. However, it does appear that the addition of a large concentration of absorbing particles, which would add to the magnitude of extinction, causes a corresponding increase in SSA retrieval uncertainty. However, this increase in uncertainty is not enough to

prevent several instrument configurations — including L6, which has been of particular interest throughout our analysis thus far — from adequately constraining SSA_{col}.

On the other hand, in the unknown environment case, the impact of having little *a priori* knowledge of state vector parameters is clear. There are similarities in the relative abilities to constrain when comparing BL to column, and retrieval capabilities are reduced when comparing known to unknown environment setups, regardless of vertical range considered. Of particular interest is the failure of all instrument combinations to constrain SSA_{col} when the environment is unknown, even when extinction receivers are operating at their reference uncertainties. It is clear that, in the environmental setup considered, the amount of *a priori* knowledge of the sample environment state is very important, even if all measurements are operating at low uncertainties.

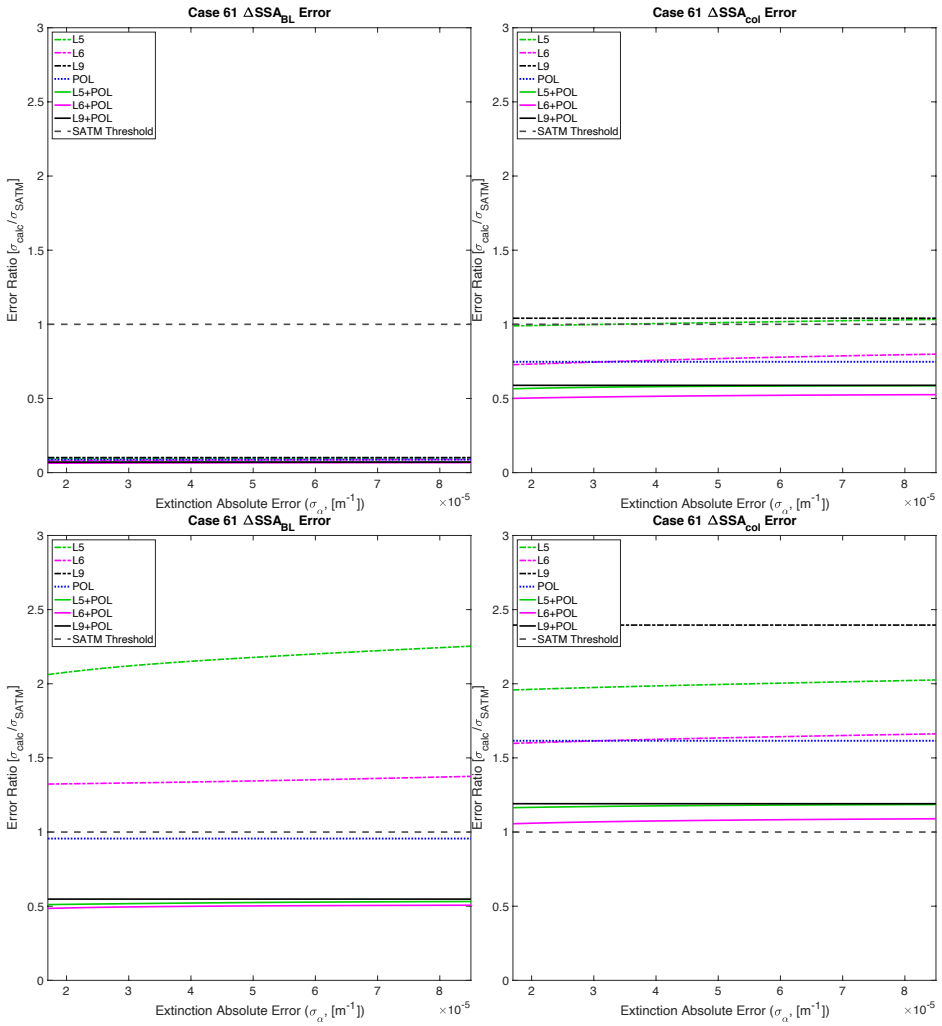

Figure 24: As in Fig. 8, except showing effect of perturbing absolute error of a priori extinction measurement (σ_{α}).

Continuing its pattern of being the hardest-to-constrain GV considered, the real part of refractive index is quite hard to constrain when increasing σ_{α}. Figure 25 shows the only instance in which m_r can be adequately constrained, given this perturbation method and aerosol loading. Not only must the environment be relatively well-characterized for any instrument configuration to constrain m_r retrieval uncertainty, but also the only instrument combination capable in this instance is the L6+POL combination that has been consistently the most capable configuration, across the board. Moreover, the only vertical range over which retrieval is possible given the *a priori* uncertainties provided for the known environment calculations is the BL.

In other words, even when the environment is somewhat well-characterized *a priori* and all the measurements are operating near their reference uncertainties, $m_{r,\mathrm{col}}$ retrieval error may not be adequately constrained by any instrument configuration. Because we know from Section 5.1.1 that $m_{r,\mathrm{col}}$ retrieval error may, in fact, be constrained in this same instance at a lower value of *a priori* perturbation, more analysis would need to be done to find out quantitatively to what extent the environment must be known *a priori* for $m_{r,\mathrm{col}}$ uncertainty for the total column to be constrained. However, considering that m_r uncertainty cannot be constrained in either vertical range in the more realistic scenario, it is unlikely that using any of the currently proposed instrument combinations would be sufficient for constraining this particular GV. This implies that the single most important factor to constraining real part of refractive index is the extent of *a priori* environmental knowledge.

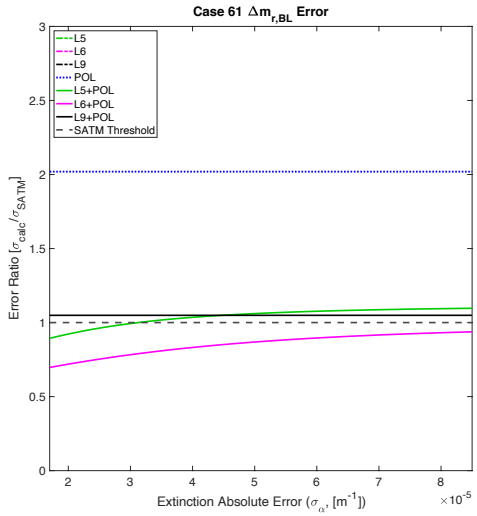

Figure 25: Real part of refractive index (Δm_r) error ratio versus a priori absolute error of extinction measurement (σ_α) for known environment setup and BL range.

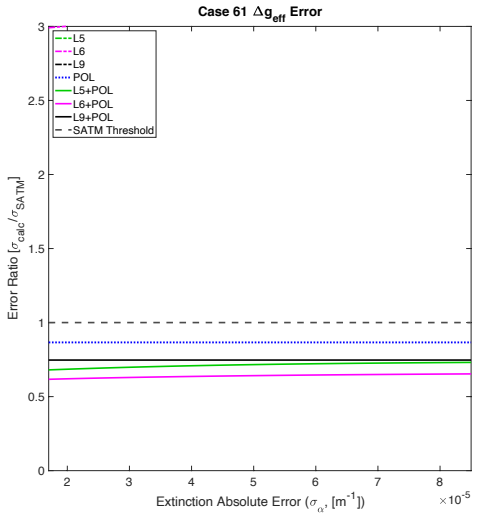

Figure 26: As in Fig. 25, except showing column-effective asymmetry parameter (Δg_{eff}) error ratio in unknown environment setup.

As we look at Fig. 26, which shows the retrieval uncertainty of asymmetry parameter after perturbing σ_α in the unknown environment case, a few things are obvious. First, only the polarimeter and combined instrument retrievals are able to constrain this GV. That the POL-only

case is sufficient shows how great of an impact there is on the ability of polarimetric instruments to accurately sense column-effective values. Also of note is the unimportance of extinction measurement accuracy to m_r retrieval capabilities, as long as polarimetric measurements are included.

5.4.2 Lidar-sensitive GV uncertainties

As before, we will first consider the effect of increasing measurement uncertainty — this time, uncertainty of extinction — on the vertically resolved values of extinction. One would likely assume that the effects of perturbing this specific measurement would have the greatest effect on the retrieval capabilities of extinction. This is, in fact, the case. We see in Fig. 27 in the top right panel that, even when the environment is well-known and extinction measurement error is relatively low, POL- and L9-only instruments are unable to constrain, which follows earlier patterns. Comparing this result to the top right, it is obvious that extinction error does have an impact on the ability to retrieve extinction, which was expected. Moreover, comparing each known (top row) to each corresponding unknown (bottom row), impact of *a priori* can be seen, as more instrument combinations fail when the environment is unknown, even when extinction measurement error is low. Most importantly, however, is that even when extinction measurement is essentially worthless and the environment is poorly characterized *a priori*, L6+POL is able to adequately constrain extinction throughout the column. Having said this, we must also note that L6-only configuration is unable to constrain extinction in the same instance, which shows the importance of polarimetric information content (specifically, the accuracy of DOLP measurement) to the ability to retrieve profiled extinction values.

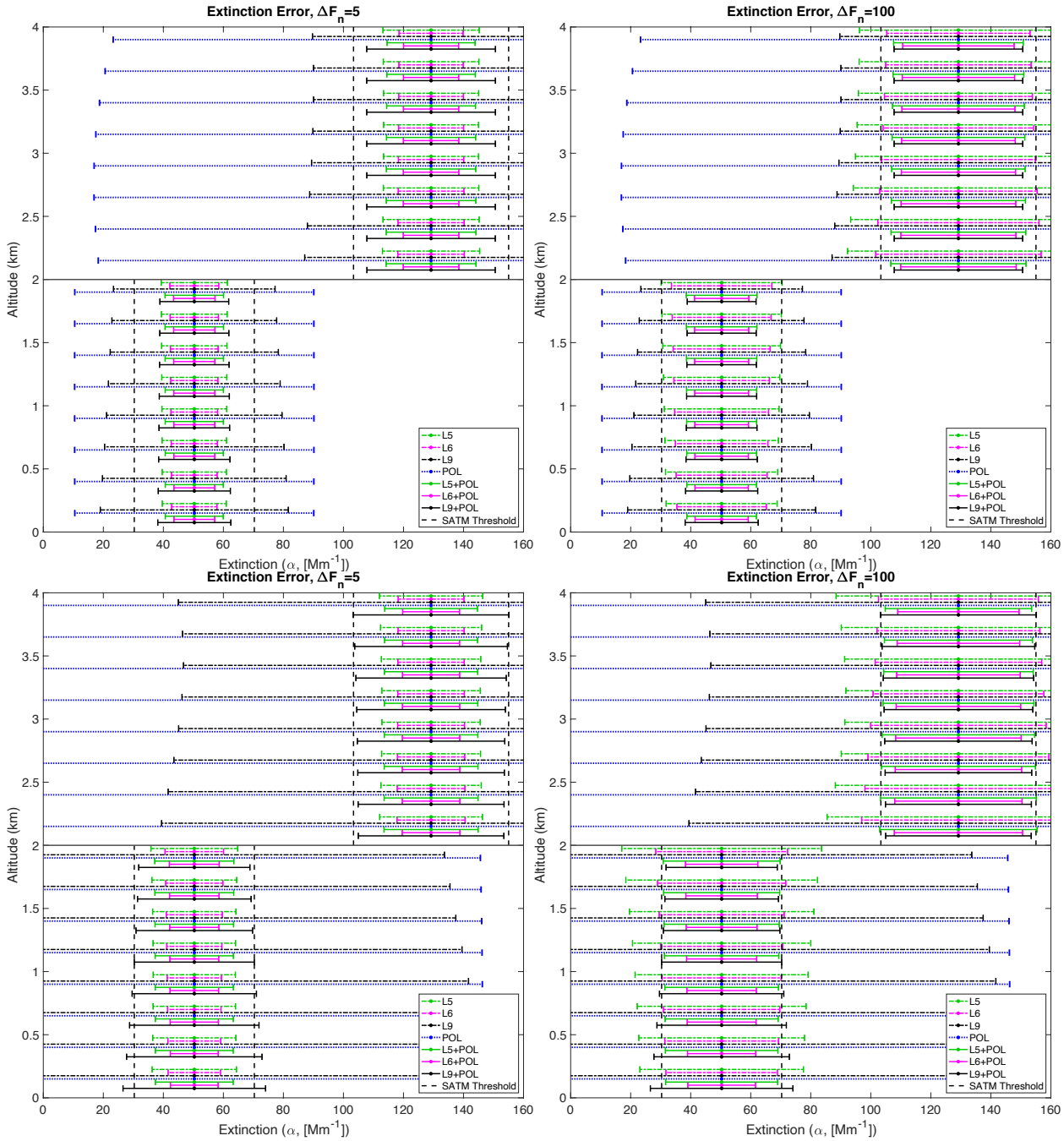

Figure 27: As in Fig. 13, except results are from perturbing a priori extinction measurement error (σ_α).

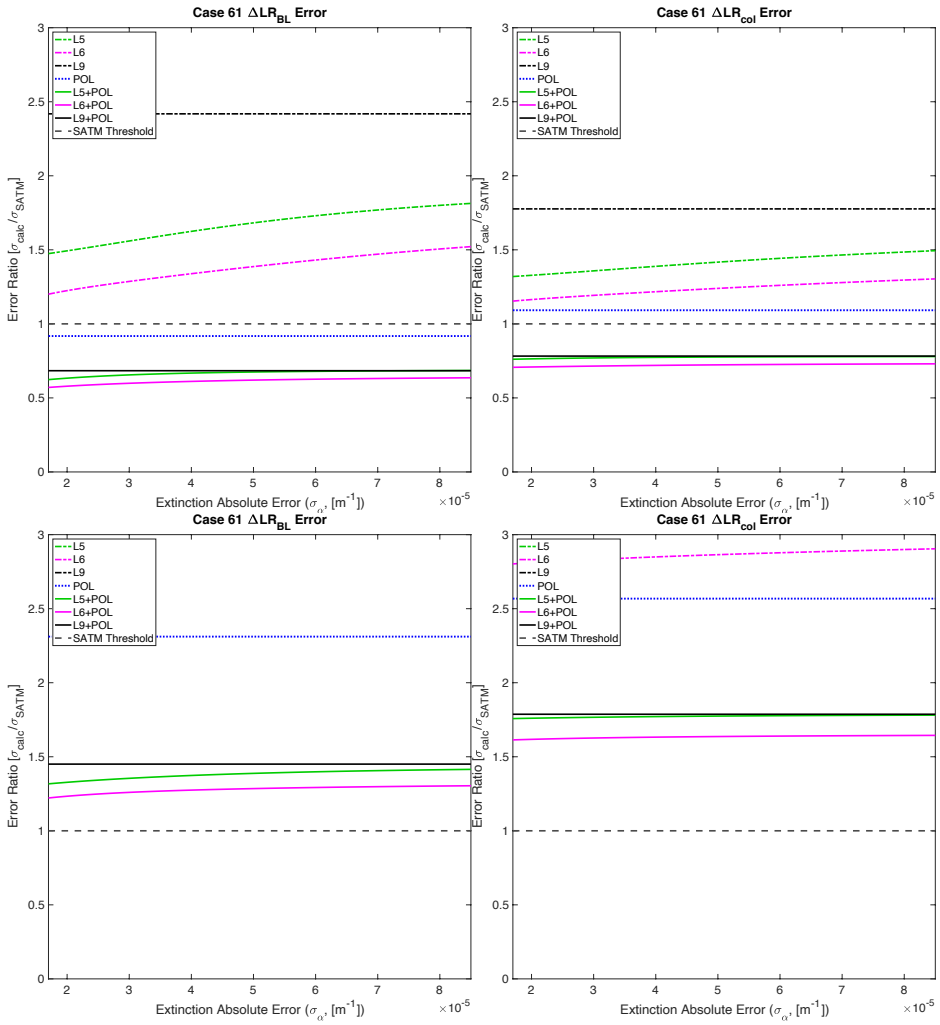

Figure 28: Lidar ratio (ΔLR) error ratio versus a priori absolute error of extinction measurement (σ_α) for known environment setup boundary layer (BL, top left) and total column (col, top right) ranges, as well as for unknown environment setup BL (bottom left) and col (bottom right) ranges.

Whereas the L6+POL combined instrument technique is sufficient to constrain the vertical profile of extinction error itself, even in the unknown environment case, this is not the case with LR. In the known environment setup, extinction measurement can be nearly ignored, and several instrument combinations can still constrain values of LR. However, when the environment is unknown *a priori*, no combination can meet the SATM requirement. The differences between these results and those seen for extinction only speak to the importance of error propagation to extinction and backscatter measurements, when they are used together. That is, the added ability of polarimetric measurements to determine column-effective distributions is not sufficient to offset

the addition of backscatter measurement uncertainty (even when small) in LR calculations when the environment is not well-characterized. This shows the importance of knowing *a priori* characteristics of the environment to accurate retrieval of LR, which has great impacts on the ability to retrieve much speciated information about the environment, considering LR is an intrinsic property of aerosols.

Table 14: Analogous to Table 10, but for figures within Section 5.4 depicting "known environment" results retrieved by perturbing σ_α (units of [Mm^{-1}]). It follows that GV uncertainties within green boxes show how well the three non-extinction measurements can constrain the respective GVs. Additionally, yellow boxes give the magnitude of σ_α at which a GV may no longer be constrained by the other three measurements.

GV	L5	L6	L9	POL	L5+POL	L6+POL	L9+POL
SSA, column	3.21×10^{-5}	0.03		0.03	0.02	0.02	0.02
SSA, BL	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01
g_{eff}, column				0.01	0.01	0.01	0.01
LR, column					8.95	8.38	8.98
LR, BL				5.53	4.13	3.83	4.12
τ_f, column				0.02	0.02	0.02	0.02
τ_f, BL		4.17×10^{-5}			0.02	0.01	0.02
τ, column	4.31×10^{-5}	1.00×10^{-4}		0.01	0.01	0.01	0.01
τ, BL	4.45×10^{-5}	1.00×10^{-4}			0.01	0.01	0.01
m_r, column							
m_r, BL					3.07×10^{-5}	0.02	

Table 15: As in Table 14, except for figures within Section 5.4 depicting "unknown environment" results.

GV	L5	L6	L9	POL	L5+POL	L6+POL	L9+POL
SSA, column							
SSA, BL				0.04	0.02	0.02	0.02
g_{eff}, column				0.02	0.01	0.01	0.01
LR, column							
LR, BL							
τ_f, column					0.03	0.03	0.03
τ_f, BL					1.00×10^{-5}	0.02	
τ, column	3.83×10^{-5}	5.00×10^{-5}		0.02	0.01	0.01	0.01
τ, BL	3.62×10^{-5}	1.00×10^{-4}			0.02	0.02	
m_r, column							
m_r, BL							

5.5 Effects of increasing backscatter measurement error

For Section 5.5 as a whole, the same expectations are made regarding the behavior of POL-only retrieval abilities as are made for Section 5.4. One important difference is that throughout Section 5.5, instrument configurations including L9 should show a much greater effect, considering that instrument contains only backscatter measurement capabilities.

5.5.1 Conventional GV uncertainties

As we have done with several figures before, we will show only the results for the unknown environment case for the ability to retrieve AOD while increasing δ_β, because several instrument combinations are still able to constrain AOD when the environment is poorly characterized. First, note that L6+POL remains the best option as far as constraint abilities are concerned. As regards the reasons that total column AOD values are being better constrained than BL values, one thought is that this result is mostly a factor of the aerosol loading considered in this specific case. By

including the smoke particles in the top layer, the backscatter values themselves, as a whole, should increase because of the relatively smaller size of smoke particles (both fine and coarse mode) than marine particles. Because smaller particles have the property of greater symmetry regarding forward and backward scattering, more backscatter information content is being added by sensing the contents of the entire column versus the BL alone. Most importantly, it is clear that backscatter measurements and *a priori* knowledge of the environment are unnecessary for many of the instrument combinations (in particular, L6+POL) to constrain AOD.

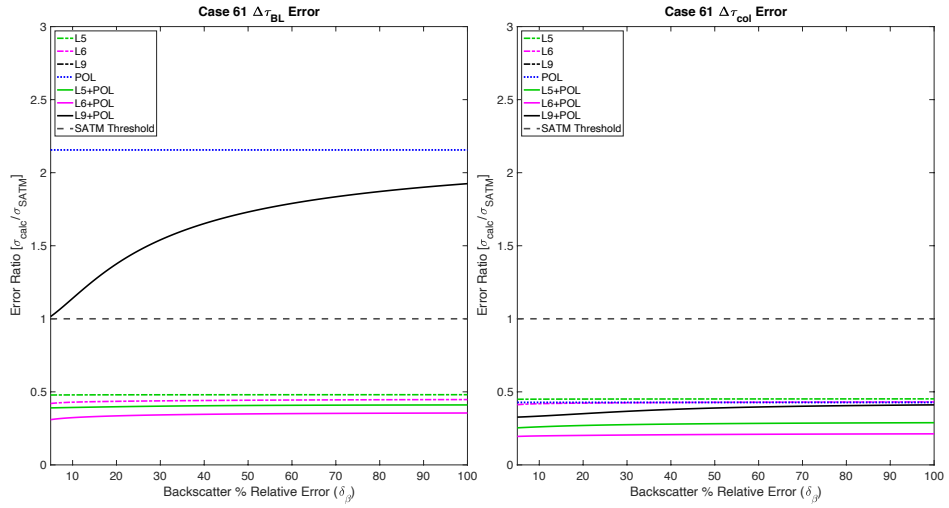

Figure 29: As in Fig. 22, except showing results from perturbing a priori percent relative error of backscatter measurement (δ_β).

Following previous reasoning, the abilities of our various instrument combinations to constrain fine mode AOD should be worse. That is, when not including coarse mode particles, the relative importance of backscatter will be greater because of the lack of interference by the strong forward scattering that larger particles produce. Considering Fig. 30, this expectation is borne out in the results. Again, results from the unknown environment setup are shown, meaning that even when *a priori* state vector and backscatter measurement errors are very large, the other three measurements are able to constrain fine mode AOD retrievals.

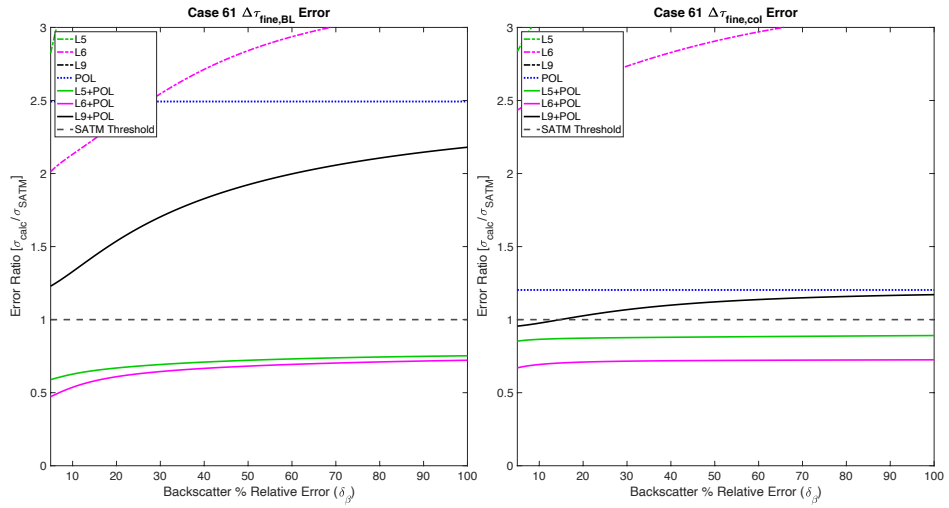

Figure 30: As in Fig. 29, except showing fine mode AOD ($\Delta\tau_{fine}$) error ratio.

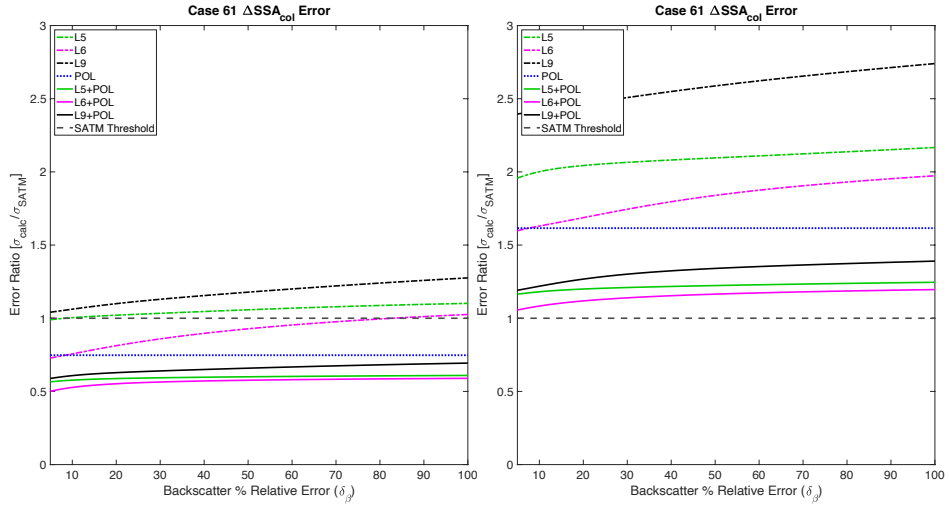

Figure 31: Single scattering albedo (ΔSSA) error ratio versus a priori percent relative error of backscatter measurement (δ_β). Results are shown for the known environment (left panel) and unknown environment (right panel) setup, both over the total column range.

In an effort to avoid repeating prior reasoning, the BL results showing the impact of increasing δ_β have not been included here. Even when the environment is not well-known, L6+POL is able to constrain SSA retrievals in the BL. In Fig. 31, the retrieval ability of column-effective SSA are shown for the known (left pane) and unknown (right pane) setups. The importance of *a priori* knowledge is clear, given the stark contrast between retrieval abilities of all instrument configurations when *a priori* uncertainties are increased. One reason for the relatively worse column-effective retrieval ability has to do with the calculation of SSA itself. By greatly

increasing the amount of absorption within the target volume (i.e., including a high concentration of low-SSA black carbon particles), the overall values of SSA are reduced, increasing the uncertainty values overall because of the added uncertainties of each FT layer.

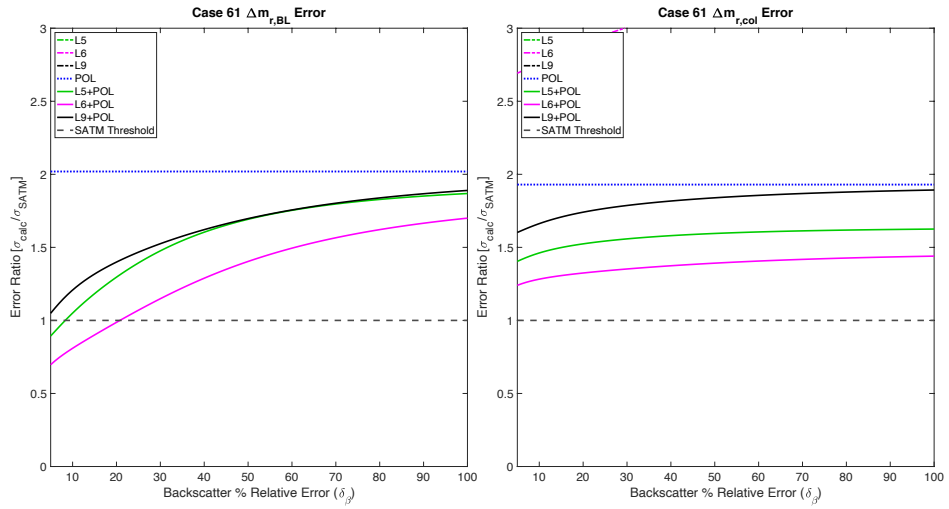

Figure 32: As in Fig. 18, except showing results for perturbing a priori percent relative error of backscatter measurement (δ_β).

As we return again to analyzing the ability to constrain real part of refractive index, the results are not promising. In Fig. 32, only those results from the known environment case are shown, for the only instance in which any of the instrument combinations can constrain m_r when δ_β is increased is when L6+POL combination is used to sample the BL. The importance of *a priori* knowledge of the environmental state cannot be overstated, in this instance, because even in the one instance (i.e., combination of method, vertical range, and instrument combination) in which there is some ability to constrain m_r, the possibility only exists when the value of δ_β is still relatively small.

For the asymmetry parameter, only the unknown environment case will be shown because of the ability of the POL-only (and all combined instrument techniques) to constrain this column-effective value. Though scattering angle is certainly important to the calculation of the asymmetry parameter, this GV may still be adequately constrained when there is little environmental

information and essentially no help is offered by the lidar backscatter channels. Especially considering that every lidar-only configuration has error ratios so large that they don't even show up in Fig. 33, these results show a strong dependence of accurate asymmetry parameter retrieval on accurate polarimetric measurements, which makes sense considering the relatively strong abilities of MAPs to constrain column-effective values.

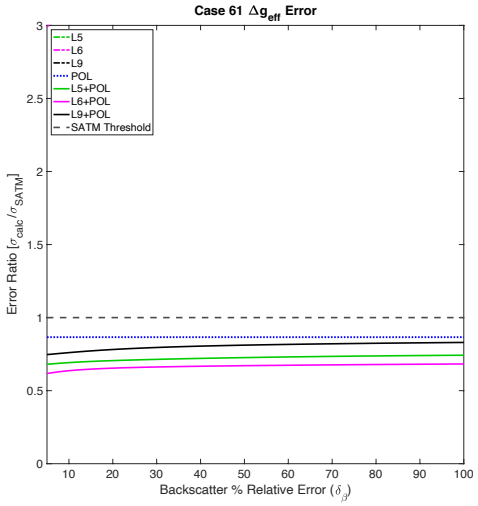

Figure 33: As in Fig. 26, except showing effects of perturbing a priori percent relative error of backscatter measurement (δ_β).

5.5.2 Lidar-sensitive GV uncertainties

Because there is little difference in the known/unknown environment setups when considering the retrieval abilities for extinction profiles, only selected runs of the unknown environment setup as δ_β is increased will be shown. In Fig. 34, it is clear that any configuration involving L9 has no ability to constrain extinction, even when δ_β is very small.

Though the retrieved *a posteriori* values of σ_α do grow larger when the environment is less well-characterized and when the accuracy of backscatter measurement is reduced, several instrument configurations — most importantly, L6+POL — are capable of constraining the entire extinction column. This means that lidar systems that include extinction receivers are still able to offer reasonable constraint even without backscatter measurement accuracy, likely by virtue of the

inclusion of well-operating extinction measurements that are able to offset the lack of information content that backscatter can provide.

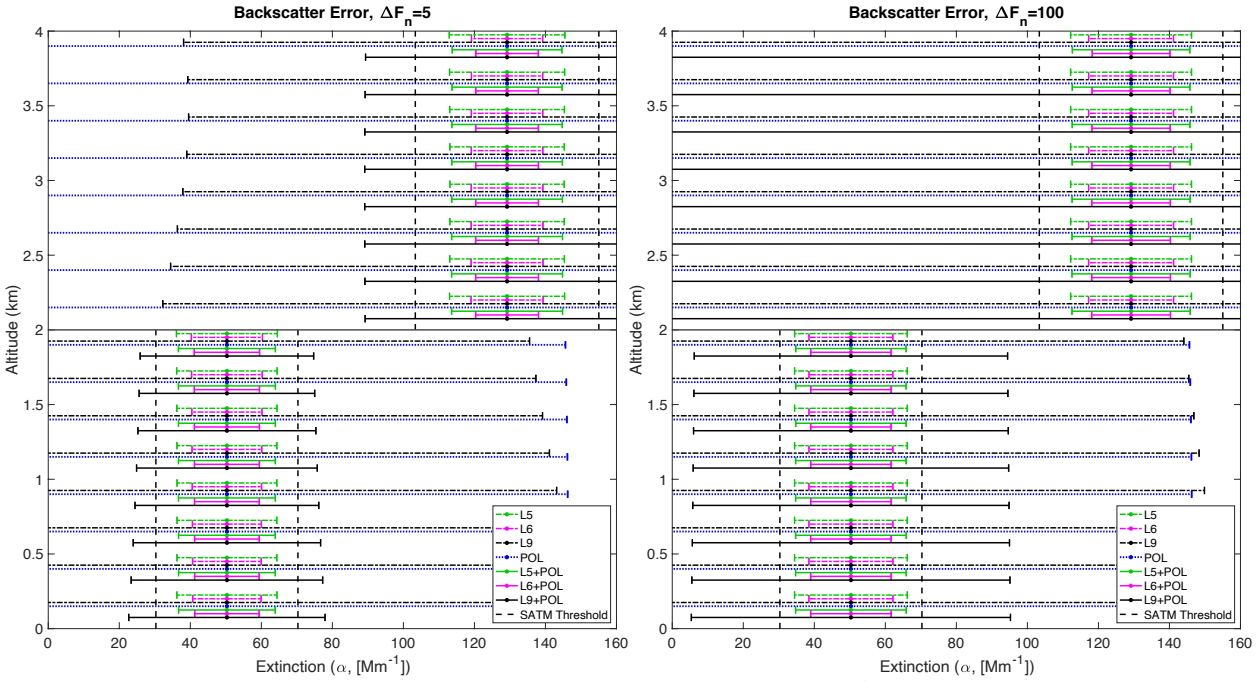

Figure 34: As in Fig. 27, except showing only the unknown environment setup results for perturbing a priori percent relative error of backscatter measurement (δ_β).

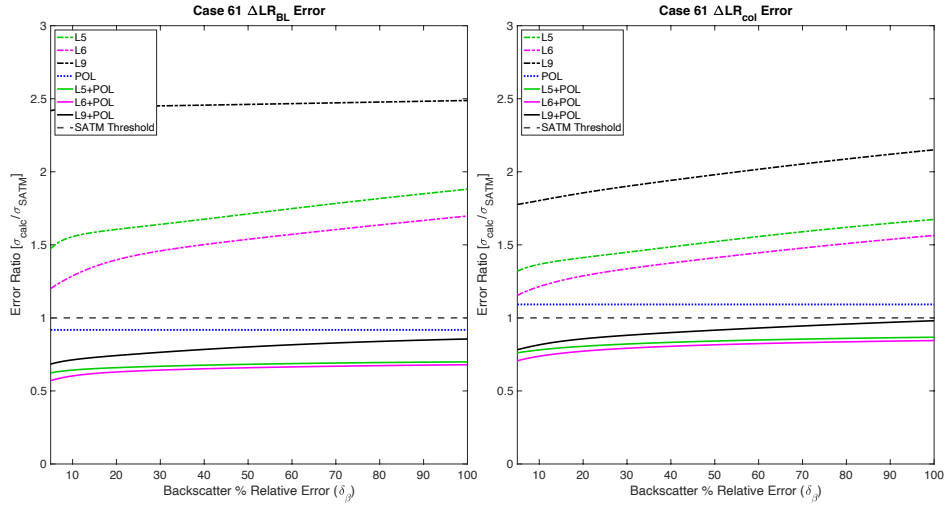

Figure 35: Lidar ratio (ΔLR) error ratio versus a priori percent relative error of backscatter measurement (δ_β) for known environment setup boundary layer (BL, left) and total column (col, right) ranges.

As was the case when σ_α was increased, when δ_β is increased, LR can only be constrained when, as in Fig. 35, the environment's *a priori* state is well-characterized. Moreover, the trends of relative constraint ability are the same as those seen in Section 5.5.1. We know from Section 5.1.2

that at least two combined instrument configurations are able to constrain LR, regardless of vertical range considered, when all four measurements are operating at their reference uncertainties. So, though the sensitivity of LR retrieval to changes in σ_α and δ_β seem relatively small, it is clear that both measurements must be very accurate for adequate LR constraint ability in the case of an unknown environment. If the environment is well-known, on the other hand, the superior instrument combination is, once again, the L6+POL combination.

Table 16: Analogous to Table 10, but for figures within Section 5.5 depicting "known environment" results retrieved by perturbing δ_β. It follows that GV uncertainties within green boxes show how well the three non-backscatter measurements can constrain the respective GVs. Additionally, yellow boxes give the magnitude of δ_β at which a GV may no longer be constrained by the other three measurements.

GV	L5	L6	L9	POL	L5+POL	L6+POL	L9+POL
SSA, column	7.88%	82.73%		0.03	0.02	0.02	0.03
SSA, BL	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01
g_{eff}, column				0.01	0.01	0.01	0.01
LR, column					9.96	9.70	11.26
LR, BL				5.53	4.21	4.09	5.15
τ_f, column				0.02	0.02	0.02	0.02
τ_f, BL		8.84%			0.01	0.01	50.10%
τ, column	0.02	0.02		0.01	0.01	0.01	0.01
τ, BL	0.01	0.01			0.01	0.01	0.02
m_r, column							
m_r, BL					7.88%	20.35%	

Table 17: As in Table 16, except for figures within Section 5.5 depicting "unknown environment" results.

GV	L5	L6	L9	POL	L5+POL	L6+POL	L9+POL
SSA, column							
SSA, BL				0.04	0.03	0.03	0.04
g_{eff}, column				0.02	0.01	0.01	0.02
LR, column							
LR, BL							
τ_f, column					0.03	0.02	14.60%
τ_f, BL					0.02	0.02	
τ, column	0.02	0.02		0.02	0.01	0.01	0.02
τ, BL	0.01	0.01			0.01	0.01	
m_r, column							
m_r, BL							

Chapter 6. Conclusions and future research implications

To be as thorough as possible in stating our conclusions, we will analyze whether H1 and H2 were confirmed in Section 6.1, then we will analyze how well our research questions were answered by our data analysis in Section 6.2. Finally, in Section 6.3 we will discuss implications of our current conclusions and how those should impact the direction and focus of future research.

6.1 Tested hypotheses

H1: This hypothesis is not supported by our analysis. Though L6+POL was able to constrain many GVs across several of our perturbation methods, there was failure of any instrument combination to adequately retrieve m_r (in BL or column) and g_{eff} when σ_{DOLP} was perturbed to be even slightly larger than its reference uncertainty.

H2: Though H1 wasn't entirely supported, the L6+POL instrument combination clearly outperformed all other combinations, regardless of GV, vertical range considered, and level of *a priori* knowledge of the environment or measurement accuracy.

6.2 Research questions

RQ1. Among the GVs considered, real part of refractive index is the only one which cannot be accurately retrieved within SATM-defined uncertainty limits by any of the observational configurations tested.

RQ2. Throughout all our results, there exists a consistent positive trend between increasing the number of lidar receiving channels and improved GV retrieval accuracy.

RQ3. Across all perturbation methods in which the *a priori* characteristics of state vector parameters were essentially unknown, the L6+POL configuration was consistently able to constrain the greatest number of GVs of any observational configuration tested. When perturbing DOLP error and backscatter error, the L5+POL configuration was able to constrain the same number of GVs, though the L6+POL configuration was able to constrain each GV within the same (or lower) uncertainty window.

RQ4. Whether viewing the error saturation figures in Section 5.3 or comparing the tables at the end of Sections 5.2–5.5, it appears that DOLP is the most essential measurement for adequate retrieval of GV uncertainties. That is, GV uncertainty constraint ability is rapidly reduced by relatively small increases in DOLP measurement error.

However, drawing any firm conclusions regarding the greater relative importance of DOLP measurement accuracy should be avoided at this early stage of analysis, for the criteria used within this study to define what constitutes an extremely large measurement error were somewhat subjective. This is particularly true for measurement error values

reported as absolute uncertainties (i.e., σ_{α} and σ_{DOLP}). The author suggests the development of more consistent "large error" criteria for absolute error values before drawing strong conclusions regarding the relative importance of the measurements considered.

RQ5. This particular research question is intentionally broad and was included with the hope that many useful, parameter-specific results could be gleaned from analysis of our results. However, early attempts to iteratively increase the *a priori* uncertainties of state vector parameters one-by-one resulted in error curves that, in most cases, became saturated over a negligibly small range of corresponding calculated GV uncertainties. It is likely that there is simply not enough information content contained within individual species-specific state vector parameters to yield useful results.

It is possible to backtrack through the algorithm results to find the *a priori* uncertainty values associated with every state vector parameter at the exact point at which calculated GV uncertainties exceed SATM requirements. However, within our current algorithm, the entire state vector is being perturbed at once by multiplicative scale factors that scale each individual parameter at different rates relative to the magnitudes of their truth values. Because of this, each potential result showing how well state vector parameters must be characterized *a priori* to guarantee adequate GV uncertainty retrievals would necessarily consist of an entire suite of species- and parameter-specific requirements. Any state vector uncertainty results stemming from this analysis method would be too complicated to use efficiently within our model. Moreover, each result would only be useful in highly specific environmental setups. Giving such results too much credence could potentially reduce the generalizability and/or potential uses of our model.

6.3 Discussion and future research

Overall, our first hypothesis was too broad to be fully supported. However, there is enough evidence to support our second hypothesis and conclude that, all things considered, the instrument configuration that is most adept at retrieving every GV considered in this study is the combined L6+POL configuration. This remains true regardless of which *a priori* uncertainties are increased or which vertical ranges are considered (i.e., BL, column, and vertically resolved).

Because we have now established which instrument configuration outperforms all the others, we will assess how well it is able to constrain specific GVs. AOD was able to be accurately retrieved in every setup considered. Moreover, fine mode AOD was also accurately retrieved by L6+POL, except when environmental knowledge was not well-characterized and σ_{DOLP} was greater than its reference uncertainty. This result is quite important because of the aforementioned risk this specific GV poses to human health. SSA was well-constrained across all perturbation methods within the BL; however, the only time column-effective SSA could be adequately retrieved was when all measurement uncertainties were at their reference uncertainties. Thus, it seems as though SSA's relationship with every measurement is important in some meaningful way, when smoke particles are included in the retrieval situation. The relatively low SSA and high concentration of these particles is likely the reason for the poor retrieval capabilities of this GV when considering the entire column. The vertically resolved extinction values were well constrained across all setups, with no exceptions. Asymmetry parameter was also well-constrained by the L6+POL configuration, except when DOLP uncertainty was larger than its reference uncertainty.

Both of the GVs that describe intrinsic properties of aerosol species were far more complicated in their reactions to different instrument measurement perturbations, even when the

most accurate instrument configuration was used. In particular, the real part of refractive index could not be accurately retrieved in any of the unknown environment cases, regardless of which measurement error was being altered. Moreover, the real part of refractive index is the only GV that could not be constrained by L6+POL when all measurements were kept at their reference uncertainties and *a priori* environmental parameters are the only inputs whose uncertainties were increased. The other intrinsic GV, lidar ratio, was able to be well-retrieved when only *a priori* state vector information was unknown, as well as when both state vector information and *a priori* measurement error of either intensity or DOLP were being increased. This makes sense, considering the LR relies on measurements made by the lidar (i.e., extinction and backscatter). When either of those measurement errors are perturbed, the ability to accurately retrieve LR is only possible when the environment is well-characterized.

Though best attempts have been made to explain the general trends of LR uncertainties over both vertical ranges considered throughout the results presented in Chapter 5, some of the results yielded from the algorithm in its current form are non-physical. This is evident in all of the LR uncertainty figures presented. For example, Figure 7 depicts the expected comparative advantages among the single instrument lidar systems and their respective polarimeter combinations; however, the same figure shows better performance of the polarimeter versus L5 and L6 instruments. Because L5 and L6 both directly measure extinction and the polarimeter does not, these results do not agree with the physical intuition stemming from basic knowledge of these instruments' measurement capabilities. Moreover, in LR uncertainty figures within Sections 5.2 and 5.3 — in which measurement error is increased only for measurements taken by the polarimeter — the single-instrument lidar results exceed uncertainty values expected using direct measurement uncertainties across the entire perturbation domain, which disagrees with results

presented in Section 5.1. However, the single versus multiple instrument results *are* correct in their relative constraint abilities, and the shape (flat) of the LR error saturation curves in Figs. 14 and 21 is consistent and matches physical expectations. It is likely, then, that the anomalous results are not being caused by misapplication of theory but, rather, reflect a magnitude issue caused by some systematic error present within the ICA model used, in its current form.

Though we have yet to find the specific cause of this error, the author suspects that one contributing factor may be the way in which extinction uncertainty is calculated (see Section 3.2.3). It is possible that uncertainty values of LR, which is the extinction-to-backscatter ratio, are larger because extinction is one of only two quantities for which full error propagation is used. That is, by including more sources of uncertainty in calculation of extinction versus those included in calculation of backscatter uncertainty, the numerator of the LR uncertainty (that of extinction) would be anomalously large compared to that of backscatter. This is just one possibility to explain the unreasonable single-lidar instrument results regarding the LR retrieval capabilities of this algorithm. Though finding the source of this major limitation of our algorithm is necessary to establish greater validity of the model, such a task proved to be beyond the scope of this study.

All things considered, the current recommendation as far as instrument combinations that should be considered a contender for fabrication and future use on an ACCP mission would be the L6+POL combination. However, even if we may assume that all four observations measured by the lidar-polarimeter instrument are operating at their reference uncertainties, future research should take a closer look into exactly how well the *a priori* state of the sample environment must be characterized (quantitatively) to adequately constrain the retrieval error of the real part of refractive index. If this range is within reason, then this instrument setup would be our recommendation for fabrication and use on a space-borne mission.

In the future, other ways to improve upon the current retrieval capabilities of the L6+POL instrument combination should be explored. Perhaps the most important first step to take is adding more test cases that use different initial aerosol setups. For example, changing the spatial distributions within the two broad layers (e.g., no longer assuming homogeneity of aerosol distribution throughout the BL) as well as the spatial relation (e.g., marine particles over smoke particles) could have an effect on retrieval capabilities. Additionally, new cases should consider entirely different species combinations, at least several of which include some level of non-sphericity among the particles within the sample volume. All of these steps would give a more complete picture of how generalizable the current algorithm is, as well as how useful, theoretically, this proposed instrument combination could be, in practice.

Another example of potential improvement to our model would be using our current instrumentation but adding data retrieved from results that stem from considering different solar zenith angles could potentially add to the information content available, resulting in improved retrieval capabilities. Another option, which involves adding hardware to the current recommended system, is adding more view angles to the lidar. Some research has shown that using lidar systems with more than one view angle has the result of adding enough information content to significantly improve the retrieval capabilities of multiple GVs (Alexandrov and Mishchenko 2017). Within the current design of the ICA algorithm, we also have the capability of including polarization in the lidar measurements, though during these somewhat preliminary stages of algorithm development, we are seeing how well the instruments combinations operate without such additional information content. Additionally, the algorithm used for this study could be altered to improve the multispectral capabilities of L6 or POL; however, these additions can become more expensive very quickly. Because an important part of our secondary goal is to

recommend a cost-effective instrument for fabrication, we suggest first looking into the other options mentioned before adding theoretical measurements that will result in real added costs.

References

Alexandrov, M. D., and M. I. Mishchenko, 2017: Information content of bistatic lidar observations of aerosols from space. *Optics express*, **25,** A134-A150.

Andreae, M. O., C. D. Jones, and P. M. Cox, 2005: Strong present-day aerosol cooling implies a hot future. *Nature*, **435,** 1187-1190.

Andrews, E., and Coauthors, 2006: Comparison of methods for deriving aerosol asymmetry parameter. *Journal of Geophysical Research: Atmospheres*, **111**.

Ansmann, A., and D. Müller, 2005: Lidar and atmospheric aerosol particles. *Lidar*, Springer, 105-141.

Ansmann, A., R. Neuber, P. Rairoux, and U. Wandinger, 2012: *Advances in Atmospheric Remote Sensing with Lidar: Selected Papers of the 18th International Laser Radar Conference (ILRC), Berlin, 22–26 July 1996.* Springer Science & Business Media.

Boesche, E., P. Stammes, T. Ruhtz, R. Preusker, and J. Fischer, 2006: Effect of aerosol microphysical properties on polarization of skylight: sensitivity study and measurements. *Applied Optics*, **45,** 8790-8805.

Bowe, B., Y. Xie, T. Li, Y. Yan, H. Xian, and Z. Al-Aly, 2019: Estimates of the 2016 global burden of kidney disease attributable to ambient fine particulate matter air pollution. *BMJ open*, **9,** e022450.

Brauer, M., and Coauthors, 2019: Mortality–Air Pollution Associations in Low-Exposure Environments (MAPLE): Phase 1. *Research Reports: Health Effects Institute*, **2019**.

Burton, S. P., and Coauthors, 2016: Information content and sensitivity of the 3β+ 2α lidar measurement system for aerosol microphysical retrievals. *Atmospheric Measurement Techniques*, **9,** 5555-5574.

Cetinic, I., C. R. McClain, and P. J. Werdell, 2018a: Pre-Aerosol, Clouds, and Ocean Ecosystem (PACE) Mission Science Definition Team Report. Volume 2; PACE Technical Report Series.

——, 2018b: PACE Technical Report Series, Volume 3: Polarimetry in the PACE Mission: Science Team Consensus Document.

Chen, X., D. Yang, Z. Cai, Y. Liu, and R. J. Spurr, 2017: Aerosol retrieval sensitivity and error analysis for the cloud and aerosol polarimetric imager on board TanSat: The effect of multi-angle measurement. *Remote Sensing*, **9,** 183.

Collis, R., and P. Russell, 1976: Lidar measurement of particles and gases by elastic backscattering and differential absorption. *Laser monitoring of the atmosphere*, Springer, 71-151.

Comerón, A., C. Muñoz-Porcar, F. Rocadenbosch, A. Rodríguez-Gómez, and M. Sicard, 2017: Current research in LIDAR technology used for the remote sensing of atmospheric aerosols. *Sensors*, **17,** 1450.

Coulson, K. L., 1988: *Polarization and Intensity of Light in the Atmosphere.* A Deepak Pub.

Deschamps, P.-Y., F.-M. Bréon, M. Leroy, A. Podaire, A. Bricaud, J.-C. Buriez, and G. Seze, 1994: The POLDER mission: Instrument characteristics and scientific objectives. *IEEE Transactions on geoscience and remote sensing*, **32,** 598-615.

Deuzé, J., and Coauthors, 2001: Remote sensing of aerosols over land surfaces from POLDER-ADEOS-1 polarized measurements. *Journal of Geophysical Research: Atmospheres*, **106,** 4913-4926.

Dubovik, O., and Coauthors, 2019: Polarimetric remote sensing of atmospheric aerosols: Instruments, methodologies, results, and perspectives. *Journal of Quantitative Spectroscopy and Radiative Transfer*, **224,** 474-511.

Eloranta, E. E., 2005: High spectral resolution lidar. *Lidar*, Springer, 143-163.

Fernandez Borda, R., J. Martins, B. McBride, L. Remer, and H. Barbosa, 2018: Capabilities of the HARP2 Polarimetric Sensor on the PACE Satellite. *AGU Fall Meeting Abstracts*, OS11D-1431.

Hansen, J., and Coauthors, 1995: Low-cost long-term monitoring of global climate forcings and feedbacks. *Climatic Change*, **31,** 247-271.

Hasekamp, O. P., and J. Landgraf, 2005: Retrieval of aerosol properties over the ocean from multispectral single-viewing-angle measurements of intensity and polarization: Retrieval approach, information content, and sensitivity study. *Journal of Geophysical Research: Atmospheres*, **110**.

——, 2007: Retrieval of aerosol properties over land surfaces: capabilities of multiple-viewing-angle intensity and polarization measurements. *Applied optics*, **46,** 3332-3344.

Hasekamp, O. P., and Coauthors, 2019: Aerosol measurements by SPEXone on the NASA PACE mission: expected retrieval capabilities. *Journal of Quantitative Spectroscopy and Radiative Transfer*, **227,** 170-184.

Hunt, W. H., D. M. Winker, M. A. Vaughan, K. A. Powell, P. L. Lucker, and C. Weimer, 2009: CALIPSO lidar description and performance assessment. *Journal of Atmospheric and Oceanic Technology*, **26,** 1214-1228.

Intergovernmental Panel on Climate, C., 2013: *Climate Change 2013 – The Physical Science Basis: Working Group I Contribution to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change.* Cambridge University Press.

Karagulian, F., C. A. Belis, C. F. C. Dora, A. M. Prüss-Ustün, S. Bonjour, H. Adair-Rohani, and M. Amann, 2015: Contributions to cities' ambient particulate matter (PM): A systematic review of local source contributions at global level. *Atmospheric Environment*, **120,** 475-483.

Knobelspiesse, K., and Coauthors, 2012: Analysis of fine-mode aerosol retrieval capabilities by different passive remote sensing instrument designs. *Optics express*, **20,** 21457-21484.

Knobelspiesse, K., and Coauthors, 2020: The Aerosol Characterization from Polarimeter and Lidar (ACEPOL) airborne field campaign. *Earth System Science Data*, **12,** 2183-2208.

Li, Z., and Coauthors, 2018: Directional Polarimetric Camera (DPC): Monitoring aerosol spectral optical properties over land from satellite observation. *Journal of Quantitative Spectroscopy and Radiative Transfer*, **218,** 21-37.

Liou, K.-N., 2002: *An introduction to atmospheric radiation.* Elsevier.

Loeb, N. G., and W. Su, 2010: Direct aerosol radiative forcing uncertainty based on a radiative perturbation analysis. *Journal of Climate*, **23,** 5288-5293.

Marshall, S. F., D. S. Covert, and R. J. Charlson, 1995: Relationship between asymmetry parameter and hemispheric backscatter ratio: implications for climate forcing by aerosols. *Applied Optics*, **34,** 6306-6311.

Martins, J. V., R. Fernandez-Borda, B. McBride, L. Remer, and H. M. J. Barbosa, 2018: The Harp Hype Ran Gular Imaging Polarimeter and the Need for Small Satellite Payloads with High Science Payoff for Earth Science Remote Sensing. *IGARSS 2018 - 2018 IEEE International Geoscience and Remote Sensing Symposium*, 6304-6307.

Mishchenko, M. I., and L. D. Travis, 1997: Satellite retrieval of aerosol properties over the ocean using polarization as well as intensity of reflected sunlight. *Journal of Geophysical Research: Atmospheres*, **102,** 16989-17013.

Mishchenko, M.I., and Coauthors, 2004: Monitoring of aerosol forcing of climate from space: analysis of measurement requirements. *Journal of Quantitative Spectroscopy and Radiative Transfer*, **88,** 149-161.

Mishchenko, M. I., and Coauthors, 2007: Accurate monitoring of terrestrial aerosols and total solar irradiance: introducing the Glory mission. *Bulletin of the American Meteorological Society*, **88,** 677-692.

Müller, D., F. Wagner, U. Wandinger, A. Ansmann, M. Wendisch, D. Althausen, and W. von Hoyningen-Huene, 2000: Microphysical particle parameters from extinction and backscatter lidar data by inversion with regularization: experiment. *Applied Optics*, **39,** 1879-1892.

NASA: Aerosol and Cloud, Convection and Precipitation (ACCP). [Available online at https://science.nasa.gov/earth-science/decadal-accp.]

——: Science and Applications Traceability Matrix. [Available online at https://science.nasa.gov/science-red/s3fs-public/atoms/files/ACCP_SATM_Rel_E_TAGGED.pdf.]

Puthukkudy, A., and Coauthors, 2020: Retrieval of aerosol properties from Airborne Hyper-Angular Rainbow Polarimeter (AirHARP) observations during ACEPOL 2017. *Atmos. Meas. Tech.*, **13,** 5207-5236.

Russell, P. B., and Coauthors, 2014: A multiparameter aerosol classification method and its application to retrievals from spaceborne polarimetry. *Journal of Geophysical Research: Atmospheres*, **119,** 9838-9863.

Shiraiwa, M., and Coauthors, 2017: Aerosol Health Effects from Molecular to Global Scales. *Environmental Science & Technology*, **51,** 13545-13567.

Smit, J. M., and Coauthors, 2019: SPEX airborne spectropolarimeter calibration and performance. *Applied Optics*, **58,** 5695-5719.

Tanré, D., and Coauthors, 2011: Remote sensing of aerosols by using polarized, directional and spectral measurements within the A-Train: the PARASOL mission. *Atmospheric Measurement Techniques*, **4,** 1383-1395.

Viana, M., and Coauthors, 2008: Source apportionment of particulate matter in Europe: a review of methods and results. *Journal of aerosol science*, **39,** 827-849.

Winker, D. M., and Coauthors, 2009: Overview of the CALIPSO mission and CALIOP data processing algorithms. *Journal of Atmospheric and Oceanic Technology*, **26,** 2310-2323.

Xu, F., G. van Harten, D. J. Diner, O. V. Kalashnikova, F. C. Seidel, C. J. Bruegge, and O. Dubovik, 2017: Coupled retrieval of aerosol properties and land surface reflection using the Airborne Multiangle SpectroPolarimetric Imager. *Journal of Geophysical Research: Atmospheres*, **122,** 7004-7026.

Xu, F., and Coauthors, 2016: Joint retrieval of aerosol and water-leaving radiance from multispectral, multiangular and polarimetric measurements over ocean. *Atmos. Meas. Tech.*, **9,** 2877-2907.

Xu, F., and Coauthors, 2021: A combined lidar-polarimeter inversion approach for aerosol remote sensing over ocean. *Frontiers in Remote Sensing.*, **2.** https://doi.org/10.3389/frsen.2021.620871.

Yang, Y., and Coauthors, 2017: Dust-wind interactions can intensify aerosol pollution over eastern China. *Nature communications*, **8,** 1-8.
Zuev, V. E., and I. E. Naats, 2013: *Inverse problems of lidar sensing of the atmosphere.* Vol. 29, Springer.