Retrieving Aerosols, Ozone, and NO_2 Using MFRSR, RSS, and CIMEL Data

Scott Michael Gianelli

Submitted in partial fulfillment of the
requirements for the degree
of Doctor of Philosophy
in the Graduate School of Arts and Sciences

COLUMBIA UNIVERSITY

2004

ABSTRACT

Retrieving Aerosols, Ozone, and NO_2 Using MFRSR, RSS, and CIMEL Data

Scott Michael Gianelli

The ability to measure gas and aerosol amounts, and aerosol particle sizes, from data obtained by sun photometers needs improvement. The retrieval strategies for two widely used sun photometers, the Multi-Filter Rotating Shadowband Radiometer (MFRSR) and the CIMEL Electronique 318A Spectral Radiometer, disagree on whether to measure gases simultaneously with aerosols or simply use climatological values, and on how many aerosol modes are necessary to explain the aerosol extinction. A crucial issue for the MFRSR is that errors in the retrieval of gas amounts result in errors in the retrieved aerosol particle size. The Rotating Shadowband Spectroradiometer (RSS) has much higher spectral resolution than either the MFRSR or the CIMEL. This results in superior gas measurement, but the Ring effect can interfere with gas measurements if ignored. This study uses data taken between July 1999 and July 2000 from an MFRSR, RSS, and CIMEL device co-located at the Southern Great Plains (SGP) site in Oklahoma.

An Empirical Orthogonal Function (EOF) analysis is performed on all three data sets, resulting in five conclusions. First, the information obtainable from these devices is limited. Second, the aerosol size distribution is bimodal. Third, the RSS can best separate gases from aerosols, but cannot otherwise retrieve aerosol information significantly better. Fourth, the fine aerosol mode is measurable by all three devices, but

additional infrared wavelength ranges would better define the coarse mode. Fifth, subtle defects within the data show up in the EOF analysis.

Retrievals are performed, using existing techniques for the MFRSR and CIMEL, and newly devised techniques for retrieving aerosols and NO_2 with RSS data. The assumption of unimodal size distributions causes large overestimations of NO_2 amounts and the aerosol effective radius, and cannot explain the spectral curves in the data. Bimodal retrievals must not overanalyze the data, or important seasonal cycles could get masked. Finally, the accuracy of MFRSR retrievals can be improved by replacing the filter at 670 nm with another at a different wavelength, and by altering the assumptions inherent in the retrieval strategy.

Table of Contents

List of Tables

List of Figures

Acknowledgements

I would first like to thank my advisor, Dr. Barbara Carlson, for her patient guidance over the past seven years. Many other scientists at the NASA Goddard Institute for Space Studies have furthered my development as well, including Andrew Lacis, both as a teacher and as a secondary advisor; Mikhail Alexandrov, who allowed me to use his algorithm for aerosol retrievals using MFRSR data; Anthony Del Genio and David Rind, as instructors; and William Rossow, who gave copious of advice on how to critically analyze scientific papers and effectively present my own research. In addition, I would like to thank the faculty and staff of the Department of Applied Physics and Applied Mathematics at Columbia University for accepting me into their doctoral program and shepherding me through the program to its conclusion, especially the departmental administrator Marlene Arbo, and the three department members who served as readers on my thesis committee, Lorenzo Polvani, Mark Cane, and Adam Sobel. I would also like to thank the fifth member of my committee, Dr. Martin Visbeck of the Lamont Doherty Earth Observatory and Columbia's Department of Earth and Environmental Sciences.

I also wish to thank the physics faculties at St. John's University in Jamaica, New York and Adelphi University in Garden City, New York, where I obtained my B. S. and M. S., respectively. Particular gratitude belongs to Dr. Robert Finkel at St. John's and Dr. G. N. Rao at Adelphi, without whom none of my subsequent work in science would have been possible.

Several people outside of GISS supplied me with critical information and support during my seven years of graduate study. Lee Harrison and Joseph Michalsky of the

State University of New York at Albany helped design and develop both the MFRSR and RSS, and provided me with the RSS data and insight into how the devices work. James Slusser, Gwen Scott, and the late David Bigelow of the USDA UV-B monitoring network provided me with MFRSR data. The CIMEL data was obtained from the website of AERONET. Marco Vountas of the University of Bremen in Germany generously and patiently answered all my queries concerning the Raman scattering portion of the SCIATRAN radiative transfer code and the nature of the Ring effect. SCIATRAN is the property of the Institute of Environmental Physics (IUP) / Institute of Remote Sensing (IFE), University of Bremen, Germany. My research was supported in part by the Atmospheric Radiation Measurement (ARM) program of the U.S. Department of Energy and NASA's Atmospheric Radiation Sciences Program.

Over the past seven and half years, I have had the great pleasure of sharing my graduate student experience at GISS with many intelligent, thoughtful, and enjoyable fellow students. I feel honored to have studied and socialized with them, and I wish them all the best of luck and the happiest of futures. I would also like to thank all of my friends and family for their encouragement and many happy moments. In particular, I would like to acknowledge and remember my grandparents, Arthur and Florence Gianelli and James and Ethel Scott. I know they're smiling somewhere.

Finally, I need to express the greatest amount of gratitude to my parents, Arthur and Barbara Gianelli, for all their love, support, and patience. This work is dedicated to them.

Chapter 1: Monitoring Aerosols and Gas Absorbers Using Ground-Based Devices

1.1: What Is Being Monitored

This study examines data from ground-based instruments called sun photometers, which measure the intensity of sunlight at different wavelengths. From the spectral variation of the intensity values, one can obtain information about the aerosols and gases in the atmosphere. Aerosols, tiny airborne particles or droplets suspended in the air, are an important part of the global climate picture, in that they can both absorb sunlight and reflect it back into space, while also serving as cloud condensation nuclei. However, the magnitude of their effect on global climate has yet to be quantified adequately. Aerosols are generally categorized by composition, source, and whether they are produced by natural or anthropogenic processes. Anthropogenic, or man-made, processes that produce aerosols not only include industrial emissions, but also more subtle and harder to quantify activities that alter the strength of natural sources. For example, intense agriculture can add to the amount of windblown dust in a region. The major categories of aerosols are sulfates, dust, sea salt, soot, and organic aerosols. Sulfate aerosols are the end product when sulfuric compounds are released into the atmosphere. Because industrial processes readily produce sulfate aerosols, the amount of sulfates in the atmosphere has increased substantially from pre-industrial levels. Sulfates do not absorb solar radiation, but as scatterers, they add to the amount of sunlight that gets reflected back into space. Dust aerosols can come from any soil, but are most commonly

associated with deserts, especially the large dust storms that originate in the Sahara and blow over the Atlantic. While dust does absorb sunlight, the degree of this absorption is a subject of some controversy (Kaufman *et al*. 2001). Also, dust tends to be non-spherical. Soot and organic aerosols result from many of the same processes, although the regular biological processes of certain plants can also create organic aerosols. Soot is distinguished from organic aerosols in that it consists of strongly absorbing elemental carbon (Lacis and Mishchenko 1995). The absorption of sunlight due to anthropogenic soot could potentially offset the total cooling effect due to scattering by all anthropogenic aerosols. Therefore, the role of soot in the climatic effects of aerosols needs to be quantified to a clearer degree than it is at present.

A number of gases absorb sunlight in the visible portion of the spectrum, but some of these are more easily monitored with sun photometers than others. For example, water vapor absorbs sunlight the most strongly of any of the gases due to a series of absorption bands in the visible and near infrared, beginning with a very weak band centered at 442 nm and culminating with a very strong band peaking at 935 nm. The combination of the strength and complexity of the absorption bands prevents a simple, direct measurement of water vapor amounts using sun photometers with broadband filters, however. Column water vapor is still measurable by these devices, but not in a way that can be conveniently incorporated into a coupled aerosol retrieval. Oxygen also has some significant absorption bands, including a very strong peak called the A-band at 760 nm, a moderate peak at 680 nm, and a weak line at about 628 nm. In addition, the collisions between pairs of O_2 molecules produce a continuum in the visible region of the spectrum, including several significant peaks centered at 370, 445, 475, 530, 575, and

630 nm (Greenblatt *et al*. 1990). The variations in oxygen abundance correspond to variations in air pressure, though, and can therefore be monitored more easily by other means.

Two gases that can be monitored simultaneously with aerosols, due to the way they absorb visible solar radiation, are ozone and nitrogen dioxide. Ozone (O_3) is an important gas to monitor for three significant reasons. First and foremost, stratospheric ozone shields the earth's surface from the sun's harmful ultraviolet radiation. Most of the ozone in an atmospheric column is located in the stratosphere, and ozone has been sufficiently well monitored in recent decades to determine that stratospheric ozone concentrations are, in fact, declining. The ozone "holes" may be located over the polar regions, but even the middle latitudes have recently experienced a decadal decline of about 2% (IPCC 2001). Second, ozone is also a significant greenhouse gas, with absorption bands in the infrared wavelengths at which the earth emits radiation out to space. Finally, ozone produced in the troposphere through combustion and industrial processes can impair human respiratory health.

The Total Ozone Mapping Spectrometer (TOMS) and the Europe-based Global Ozone Monitoring Experiment (GOME) provide global satellite coverage of column ozone amounts, while networks of Dobson and Brewer spectrometers monitor ozone from the ground. A recent comparative study of ozone retrieval devices and algorithms showed that, despite some systematic errors in the satellite data, the satellite and ground-based devices generally agree within a few percent (Bramstedt *et al*. 2002). However, systematic biases do exist; for example, the ozone amounts retrieved by TOMS consistently exceed those retrieved by the Dobson devices, and the GOME results

fluctuate around the Dobson results in a seasonal cycle. In the middle latitudes, ozone amounts generally fluctuate around 300 Dobson Units (DU), where a Dobson Unit is the thickness, in thousandths of a centimeter, that a particular gas would have if the entire amount contained within an atmospheric column were isolated and held at standard temperature and pressure (Stephens 1994).

Like ozone, nitrogen dioxide (NO_2) can adversely affect human respiratory health. At face value, NO_2 might not match ozone in toxicity, but its presence in the atmosphere in the wake of industrial and combustion processes leads to the formation of ozone. Satellite observations of forest fires have shown elevated NO_2 levels leading to similar increases in ozone as well (Burrows *et al*. 1999). NO_2 is also directly involved in the chemical processes that create and remove ozone from the stratosphere. The recent global analysis of tropospheric NO_2 amounts conducted by GOME (Velders *et al*. 2001) suggests that typical tropospheric NO_2 column amounts can range from $2x10^{15}$ molecules/cm^2 (approximately 0.075 DU) over the ocean, to higher than $2x10^{16}$ molecules/cm^2 (0.75 DU) in highly polluted areas like central Europe or the northeastern United States. However, it is possible that the GOME satellites miss some NO_2 absorption near the surface, so the tropospheric values might be a bit low. Typical stratospheric amounts, as determined by the same study, tend to be in the neighborhood of 0.075 DU, although the amount increases with decreasing latitude towards the polar region where it's summer. Therefore, in pristine areas the bulk of NO_2 absorption in an atmospheric column may take place in the stratosphere.

1.2: Aerosols and Climate

1.2.1: The Direct and Indirect Effects

The ways in which aerosols influence climate are generally divided into two categories. The first category, called the direct effect, involves the absorption and scattering by the aerosols themselves. The direct effect is the measure of the contribution of aerosols to the planetary albedo, or the fraction of sunlight incident on the earth that gets reflected back to space. Aerosols can affect the amount of outgoing reflected radiation in a number of different ways, depending on the circumstances. For example, over low reflecting surfaces, such as the ocean, aerosols will increase the total albedo and cool the planet, but over higher reflecting surfaces on land, the effect of aerosols is more ambiguous. Furthermore, aerosols that absorb significantly, such as soot and dust, will tend to lower the albedo (Herman and Browning 1975). Aerosol absorption also warms the neighboring air, which will eventually circulate to the surface and bring the added warmth with it. According to models, the increase in carbon dioxide from pre-industrial levels has generated a corresponding increase in the global mean surface radiative flux that has been calculated to be approximately 2 W/m^2. In order to similarly quantify the radiative forcing produced by aerosols, a great number of aerosol modeling studies have been conducted, where estimations of global aerosol properties are used as input and their simulated effects are then calculated and analyzed.

The second effect caused by aerosols, the indirect effect, refers to their influence on the number of droplets in a given cloud, which in turn affects the cloud's overall

reflectivity, precipitation, and lifetime. A water droplet below a certain size will shrink from evaporation faster than it can expand from condensation. Therefore, in order to produce cloud droplets, supersaturated air requires the presence of other particles, namely aerosol particles, around which water can accumulate. Increasing the number of CCN, given the same amount of water, will produce smaller droplets. Satellite-based retrievals of cloud droplet size have shown that cloud droplets are indeed smaller where more CCN are present (Han *et al*. 1994), and a recent study of global values for cloud droplet radii showed that cloud particles tend to be largest over remote oceans and smallest over polluted land areas (Bréon *et al*. 2002). However, the number of CCN in a given air mass can reach a saturation level, where additional aerosols beyond this level do not significantly alter the number of CCN or the size of cloud particles. A recent study by Menon *et al*. (2002) analyzing the relationship between sulfates and CCN concluded that the number of CCN has a non-linear relationship to sulfate mass in general, being less sensitive when sulfate levels are high.

Adding aerosols to saturated air produces two effects. First, the total cloud droplet surface area will increase for the same amount of liquid water, thereby increasing the reflectance and optical thickness of the cloud (Twomey 1977a). The second part of the indirect effect involves changes in precipitation. Albrecht (1989) suggested that the reduction in the number of large droplets in marine clouds would lead to a reduction in drizzle. This prolongs the life of the cloud, enabling it to reflect even more sunlight back to space. On a tangible level, the indirect effect makes its presence felt in changes in weather trends. A detailed analysis of daily maximum and minimum temperatures in

both hemispheres (Karl *et al*. 1993) showed that the minimum temperature has increased since 1950 at three times the rate of the maximum temperature.

1.2.2: The Need for Improved Aerosol Monitoring

The 2001 IPCC report indicated that the role of anthropogenic aerosols in climate change remains, in a quantitative sense, quite vague. The combined negative forcing due to the scattering of sulfates and soot is not presently believed to exceed 1 W/m^2, with significant margins of error. Absorption by soot, on the other hand, produces a positive forcing which at least partially cancels out the scattering. Uncertainties in quantifying the effects of mineral dust are such that even the sign of the net forcing cannot presently be determined. The indirect effect clearly produces a negative forcing, but the strength of this forcing remains an extremely open question. The IPCC rated the level of scientific understanding of the direct effect of sulfates as "low." All the other aerosol effects, including the indirect effect, were rated as "very low." In other words, much more work needs to be done in the study of every facet of aerosols. This includes improving the techniques used to monitor aerosols, both from above and below. Not only do aerosol amounts need to be measured better, but their physical and radiative properties do as well. For one thing, absorption needs to be better quantified. Just as importantly, though, the aerosol size distribution needs to be more accurately determined. This will not only improve understanding of the direct effect, but the indirect effect as well; the number of aerosol particles, and therefore potential cloud-condensing nuclei, in a given air mass can be inferred only if the aerosol extinction and size distribution have both been accurately

determined. In addition, since gases like ozone and nitrogen dioxide absorb at wavelengths relevant to the analysis of aerosols, these gases also need to be measured with as much accuracy as possible. A complete assessment of aerosol properties is only possible to the degree that the respective contributions of aerosols and gases to the total extinction can be separated.

1.3: Monitoring Aerosols and Gases from the Ground

1.3.1: Comparing Ground-Based Devices with Satellite Detectors

Some existing strategies to retrieve aerosol and gas amounts involve satellite instruments, while others incorporate national and global networks of ground-based devices. Satellites provide complete global coverage, but their measurement strategies are complicated by the need to analyze reflected sunlight and to separate aerosol and surface contributions. Ground-based devices, by contrast, measure the intensity of light coming directly from the sun on an unimpeded path to the detector. The direct beam is far simpler to analyze mathematically than scattered radiation, providing ground-based aerosol monitoring devices with a huge advantage over their satellite counterparts. As a result, networks of ground-based devices produce more reliable aerosol measurements at a given site than a satellite instrument overlooking the site can do. The accuracy of the satellite measurements can therefore be tested, by comparing satellite results with the "ground truth" obtained from sun photometer devices.

1.3.2: Sources of Data for This Study

The three devices used in this study are all ground-based sun photometers. Two widely used sun photometers are the Multi-Filter Rotating Shadowband Radiometer (MFRSR) and the CIMEL Electronique 318A Spectral Radiometer. The Multi-Filter Rotating Shadowband Radiometer (MFRSR) has six filters of different wavelengths (Harrison *et al*. 1994). A number of networks use MFRSR devices, and nearly a hundred MFRSR's are in operation in the United States alone (Alexandrov *et al*. 2002b). The CIMEL Electronique 318A spectral radiometer is the device of choice for the most significant global aerosol monitoring network, AERONET (Holben *et al*. 1998). Like the MFRSR, the CIMEL is a low-resolution device, with filters at seven different wavelengths. While some MFRSR filters are located where gas absorption is strong, for the purpose of simultaneously retrieving gas amounts and aerosol properties, the CIMEL wavelengths are arranged to minimize the interference of gaseous absorption, in order to focus exclusively on aerosol measurement. This means that the CIMEL and the MFRSR have very different retrieval strategies, whose relative strengths and weaknesses can be analyzed. The MFRSR algorithms also assume that the size distribution of the aerosols can be described adequately as a single mode, but the CIMEL algorithm assumes a bimodal distribution, meaning that the aerosols are separated into two distinct size groups.

A third device, called the Rotating Shadowband Spectroradiometer (RSS), contains a high-resolution array of over 1000 channels (Harrison *et al*. 1999). The increased resolution improves the separation of the extinctions due to gases and aerosols,

making the RSS a superior device at face value over the MFRSR. However, it is also significantly more expensive, and a small network of MFRSR devices can be set up for the price of one RSS. Only two RSS devices are operating at the present time. In fact, no scheme for retrieving the aerosol size distribution existed for the RSS prior to this study. Similarly, as the resolution of the RSS enables superior separation of gases and aerosols relative to the MFRSR, an algorithm to retrieve column amounts of nitrogen dioxide utilizing the higher resolution of the RSS data has also been designed and implemented as part of this study, and is presented here as well. The RSS can also be used to produce MFRSR and CIMEL "equivalent" data, by defining the "equivalent" data to correspond to the central wavelengths of the MFRSR and CIMEL filters, but with the higher resolution of the RSS measurements. This provides a direct comparison of these instruments, and their measurement potential, in a way that is not subject to calibration differences.

The Southern Great Plains (SGP) atmospheric research site in Oklahoma is the one place where an MFRSR, an RSS, and a CIMEL device are co-located, and a comparative analysis of the data from all three devices can be made. Indeed, such an analysis has been conducted once before in 1997, when the Atmospheric Radiation Measurement (ARM) program held an Intensive Observation Period to study aerosols (Schmid *et al*. 1999). The Schmid *et al*. study showed that despite different approaches to obtaining and analyzing the data, the devices showed agreement in aerosol optical depth values to within 0.02 at all wavelengths, within the accuracy of the calibration. Even with optimized calibration and the best possible optical depth measurements, though, differences persist with regards to how to convert these optical depth values into

a complete analysis of the aerosol optical depth size distribution, and how to best account for the absorption of ozone and nitrogen dioxide.

1.4: How Sun Photometers Monitor Aerosols and Gases

1.4.1: Atmospheric Extinction and Optical Depth

Two processes directly affect incoming and outgoing radiation. In the first case, the gases and particles in the atmosphere absorb the radiation. In the second, these gases and particles scatter the radiation, sometimes within the atmosphere, sometimes towards the surface, and sometimes back into space. The combination of absorption and scattering taking place as photons of sunlight make their way from the top of the atmosphere to the ground is collectively referred to as extinction. Before conducting a detailed analysis of the data and retrieving aerosol properties and gas amounts from them, the degree of extinction, or optical depth, needs to be determined for each channel of the device. Usually represented by the Greek letter τ, optical depth is calculated by means of the Beer-Lambert formula:

$$I = I_0 * e^{-\tau * m}.$$

I is the intensity of the direct solar beam which reaches the surface without being scattered, and I_0 is the intensity of incoming solar radiation at the top of the atmosphere. The sun emits radiation approximately like a blackbody at a temperature of 6000 K, with a peak in intensity at 470 nm (Liou 1992). The spectral dependence of the solar intensity

is shown in Figure 1.1. The emitted radiation is not a smooth function of wavelength, because a number of different elements absorb radiation within the sun's atmosphere. Detectors on the ground can measure I directly, provided that the detector has been well calibrated. The airmass m is the amount of atmosphere that the sunlight travels through, relative to directly overhead (a solar zenith angle of 0° gives an airmass value of 1). For zenith angles less than 60°, this term can be accurately approximated by the inverse of the cosine of the zenith angle. Optical depth is a particularly useful quantity to work with because it is additive; that is, the total optical depth at a given wavelength equals the sum of the optical depth amounts of each gas and aerosol at that wavelength.

1.4.2: The Absorption of Sunlight by Ozone and Nitrogen Dioxide

The measurement of the amounts of different absorbing gases present in the atmosphere is possible because each type of gas molecule has its own characteristic absorption spectrum, by which the strength of the absorption varies with wavelength. A number of gases in the earth's atmosphere absorb significant amounts of incoming solar radiation in the near ultraviolet, visible, and near infrared parts of the spectrum. The optical depth of each gas at a given wavelength can be expressed as the product of the total column concentration of the gas and a proportionality constant, called the absorption coefficient, which varies with wavelength. Figure 1.2 shows the contributions to the atmospheric transmission and optical depth of the various gas absorbers, and the molecular, or Rayleigh, scattering, from 370 to 1050 nm. The humidity is assumed to be relatively low, ozone absorption is calculated for a typical amount of 300 DU, and

nitrogen dioxide is calculated for a value of 1 DU. The spectrum of ozone (O_3) has a moderately strong absorption band, called the Chappuis band, located in the visible region to go with the exceptionally strong absorption bands in the ultraviolet, called the Hartley and Huggins bands (Liou 1992). The Chappuis band is much smoother and broader than the visible absorption bands for water vapor and nitrogen dioxide. Ozone absorption remains fairly significant throughout most of the visible spectrum, peaking at 610 nm. The absorption spectrum of nitrogen dioxide is considerably detailed, although overall NO_2 absorption tends to be relatively small. Schroeder and Davies (1987) demonstrated that NO_2 absorption is too significant to ignore if aerosol optical depth is to be measured accurately. NO_2 absorption peaks in the blue region of the spectrum at 415 nm, where the MFRSR has its lowest wavelength. As a result, improperly accounting for NO_2 can cause serious errors in the MFRSR and CIMEL retrievals of aerosol size distribution. Overestimating NO_2 amounts will lead to retrieved particle sizes that are too large, while underestimating NO_2, or ignoring it entirely like the CIMEL algorithms do, will make the particles seem smaller than they actually are.

1.4.3: The Scattering of Sunlight Due to Gases

The molecules of gas in the atmosphere scatter the incoming solar radiation as well. Lord Rayleigh first studied the scattering of electromagnetic radiation by molecules, and consequently the phenomenon bears his name. A number of methods have been developed to evaluate the optical depth due to Rayleigh scattering (Hansen and Travis 1974, Bodhaine *et al*. 1999). The two defining characteristics of Rayleigh

extinction are that it varies very nearly to the inverse of the fourth power of the wavelength ($\tau \approx C/\lambda^4$), and that it is linearly proportional to the atmospheric pressure at the surface. In principle, this makes calculating Rayleigh optical depth simple, and the Rayleigh contribution is therefore usually subtracted from the total optical depth before the analysis of ground-based data begins. However, the recent technique developed by Bodhaine *et al*. (1999) underscores the need for care in this regard, as the calculations derived from this technique produce significant differences with older formulas at wavelengths below 400 nm.

Another type of scattering in the atmosphere, too conspicuous in high-resolution data to be ignored, is Raman scattering, through which the scattered light emerges at a slightly different wavelength than the incident light. The effect of Raman scattering on the measured intensity of the Fraunhofer lines, or lines where elements absorb significant amounts of radiation within the solar atmosphere and thereby sharply reduce the outgoing solar flux at that wavelength, was first observed by Grainger and Ring (1962). Grainger and Ring discovered that the drop-off in measured intensity at the wavelengths of the Fraunhofer lines was less severe in scattered sunlight than it was in moonlight. Some additional light reached the detector in the scattered sunlight measurements, making the Fraunhofer line appear to be partially "filled in." This so-called filling in has since become known as the Ring effect. A number of modeling studies have been conducted to analyze various features of the Ring effect, most significantly regarding how it alters the measured abundance of atmospheric gases such as nitrogen dioxide. Fish and Jones (1995) reported that the Ring effect lowered retrieved NO_2 amounts by 7%. Vountas *et*

al. (1998) confirmed this work, and incorporated the Ring effect into the publicly available GOMETRAN and SCIATRAN radiative transfer models.

The significant details of the Ring effect spectrum are readily visible in high-resolution devices like the RSS. Figure 1.3 shows the average optical depth between the wavelengths of 390 and 450 nm that the RSS measured on the thirty days within the data set that have the highest humidity. Humidity itself does not directly influence the Ring effect, but it was in this particular graph where the peculiar features of the phenomenon first became apparent. The wavelengths where the Ring effect is strongest are 393.4 and 396.8 nm, the locations of Fraunhofer absorption lines for calcium. Because the incoming solar intensity is so low at these wavelengths, much more radiation gets scattered into those wavelengths than is scattered out, and the apparent optical depth is significantly reduced. Other significant features include a smaller Fraunhofer line due to calcium at 422.7 nm, and a complicated pattern resulting from the combination of Fraunhofer lines due to iron and calcium at 430.8 nm and hydrogen at 434.0 nm.

1.4.4: The Extinction Due to Aerosols

Aerosols have many different physical and chemical properties. Some are natural, and some are put into the air through man-made, or anthropogenic, processes. Unlike gases, which absorb at very specific wavelengths, the extinction due to aerosols follows a smooth, continuous curve across the spectrum. The size distribution of the aerosols determines how rapidly the extinction changes with wavelength. In 1908, Gustav Mie applied Maxwell's equations to the scattering caused by spherical dielectrics.

As most liquid water droplets and aerosol particles are assumed to be spherical, Mie scattering has been the principal means of analyzing the wavelength dependence of cloud and aerosol scattering. Some particles, especially desert dust, are non-spherical, but Mishchenko *et al*. (1997) showed that where the direct beam at least is concerned, the error in calculating aerosol extinction that results from an assumption of spherical particles does not exceed 1%. Hansen and Travis (1974) identified three specific properties of aerosols that need to be measured in order for Mie scattering to completely describe the extinction of a given mass of aerosols as a function of wavelength. One of these, the single-scattering albedo, is the ratio of the extinction due to scattering to the total extinction. Strongly absorbing aerosols like soot can have values for the single-scattering albedo significantly less than one. The remaining properties are the effective radius and effective variance. The effective radius is defined by the following formula:

$$r_{eff} = \frac{\int_0^{\infty} r \cdot r^2 n(r) dr}{\int_0^{\infty} r^2 n(r) dr},$$

where $n(r)\,dr$ is the fraction of particles between the radii r and $r+dr$, or the fractional size distribution of the aerosols. The effective variance is then defined as

$$v_{eff} = \frac{\int_0^{\infty} (r - r_{eff})^2 \cdot r^2 n(r) dr}{r_{eff}^2 \int_0^{\infty} r^2 n(r) dr}.$$

Effective variance is a measure of the spread of the particles around a central size. In an aerosol mass with a small effective variance, the particles will be roughly the same size, while a large effective variance means a broader range of sizes.

The size distribution does not always follow a simple, bell-shaped curve, however. A number of different mathematical forms have been applied to aerosol size distributions, including gamma, lognormal, and power law. When the effective variance is small, the difference in extinction caused by altering the assumed shape of the aerosol size distribution is negligible. Furthermore, even when the effective variance becomes large, limits to the accuracy of retrieved values of the effective radius and variance, coupled with the inability to retrieve more than two or three items of independent aerosol information from the data, make a clear determination of the shape of the aerosol size distribution impossible. The gamma size distribution is used in the GISS radiative transfer models, because the absence of broad "tails" in the size distribution simplifies the analysis of multiple aerosol modes, or distinct groups of aerosols with their own values for the effective radius and variance. As the determination of the presence of multiple aerosol modes is one of the key objectives of this study, all the Mie scattering calculations performed for this study assume a gamma distribution. The established size distribution retrievals for the MFRSR and CIMEL are based on conflicting premises, because the existing aerosol retrieval scheme for the MFRSR (Alexandrov *et al*. 2002a) assumes a single-mode size distribution, but the retrieval used for the CIMEL devices (Dubovik and King 2000) assumes a bimodal distribution.

The relationship between aerosol extinction and wavelength is sometimes expressed in terms of a quantity called the Ångstrom coefficient. Ångstrom devised a relatively simple empirical formula relating aerosol extinction and wavelength,

$$\tau = C * \lambda^{-\alpha},$$

where τ is the optical depth, λ is wavelength, and α is the Ångstrom coefficient. This quantity can be calculated by taking the natural logarithm of both sides of the equation, and then performing a linear fit to ln τ vs. ln λ. The Ångstrom relationship is an approximation, but it can provide a useful rough estimate of the behavior of the aerosol size distribution when an exact measurement is either not needed or not possible. The greater the Ångstrom coefficient, the smaller the particles; a value of 4 for a would imply molecule-sized particles and Rayleigh scattering.

According to Mie scattering, the extinction due to aerosol scattering varies significantly as the effective radius and effective variance change. Figure 1.4 illustrates how the Mie extinction coefficient of the aerosols, normalized to equal one at 550 nm, varies with size for a low variance of 0.1. For particles smaller than 0.3 μm, extinction steadily decreases with wavelength. In fact, the smaller the particle, the more closely the extinction curve resembles the $1/\lambda^4$ dependence of Rayleigh scattering. However, if the effective radius exceeds 0.3 μm, the extinction will increase with wavelength in the violet and ultraviolet regions of the spectrum. As the particle size increases beyond 0.3 μm, the extinction coefficient peaks at an increasingly longer wavelength. The changing response of aerosols to wavelength with the particle size makes the size distribution of the particles measurable in principle. However, since the minimum wavelength of the MFRSR is located where NO_2 absorption peaks, separating the gas and aerosol contributions to

extinction at this wavelength can prove to be quite difficult for this device. Figure 1.5 shows how the Mie extinction curves vary with wavelength when the size distribution has a high variance of 0.5. The curves are qualitatively similar to those in Figure 1.4, but the extinction does not change nearly as sharply with wavelength. In addition, the extinction coefficient peaks at lower wavelengths for a given effective radius. Figures 1.4 and 1.5 also show that Mie curves of a low variance can qualitatively and quantitatively resemble curves with a higher variance, but a lower effective radius value. The curve for an effective radius of 0.25 μm and a variance of 0.1 is very similar to the curve for an effective radius of 0.2 and a variance of 0.5, for example. In practice, this makes the effective radius and variance difficult to measure uniquely in aerosol retrievals.

1.4.5: Calibrating Sun Photometers

The most significant source of discrepancies in optical depth measurements between devices is calibration. Calibrating a sun photometer means determining what voltage signal produced by the instrument corresponds to I_0, the incoming solar flux at the top of the atmosphere. For a well-calibrated sun photometer, the optical depth at a given wavelength and time of day can be easily and accurately determined using the Beer-Lambert equation. I_0 cannot be measured directly by a detector on the earth's surface, however. Instead, a number of methods have been developed to indirectly infer I_0, and the most widely used of these is the method of Langley regression. Taking the logarithm of both sides of the Beer-Lambert law gives the following equation:

$$\ln I = \ln I_0 - \tau * m$$

I is measured, and the airmass m can be easily calculated as a function of solar zenith angle. The natural logarithm of I varies linearly with m, provided that the optical depth does not change significantly over the course of the day (Shaw *et al*. 1973). Therefore, plotting ln I vs. m ideally yields a straight line, whose slope is the negative of the optical depth, and whose y-intercept equals the logarithm of I_0. A typical series of Langley plots, from the data obtained by the MFRSR device at SGP on the afternoon of March 20, 2000, is shown in Figure 1.6.

Unfortunately, calibration by the Langley technique has more than its share of pitfalls. The most significant flaw with Langley plots is that aerosol amounts usually do vary over the course of a day, often in such a manner that the resulting plot still produces a straight line, but the calculated intercept differs significantly from the actual ln I_0. Unless the sun photometer is located in an area where the air is pristine, like Mauna Loa or some other high-altitude site, Langley-retrieved I_0 values will fluctuate substantially on a day-to-day basis. Viable options are unfortunately limited. Michalsky *et al*. (2001) devised a calibration procedure which looks at Langley regressions obtained over long stretches of time, and have used this procedure to calibrate their MFRSR and RSS devices, including the RSS used in this study. Modifications to the Langley approach have been attempted by Soufflet *et al*. (1992) and Forgan (1994), based on the premise that the aerosol size distribution remains steadier over the course of the day than the optical depth, and therefore serves as a more stable basis of calibration. Alexandrov *et al*. (2002a) utilized this premise when developing a new calibration technique for the MFRSR. This new technique also exploits the ability of the MFRSR to measure direct

and diffuse intensities with the same detector. Using this calibration method, large sets of data have been examined at a number of different sites (Alexandrov *et al*. 2002b).

AERONET attempts to reduce the uncertainty in the calibration of its devices by periodically bringing the CIMEL detectors to Mauna Loa. The high altitude of Mauna Loa provides the most pristine environment possible to calibrate sun photometers, and Langley regressions obtained in the mornings at Mauna Loa produce extremely consistent values for I_0. The problem with this strategy, though, is that filters in a sun photometer can change characteristics very abruptly, and practicality prevents the detectors from being calibrated more than twice a year. This is less of an issue with newer, ion-deposition filters, but the filter transmission can still be altered, for example, by changes in temperature. One of the significant structural differences between the MFRSR and CIMEL devices is that the MFRSR is temperature-controlled, while the CIMEL is not.

Using standard lamps as an alternative to calibration by the Langley method has also been attempted, but the intensities of the lamps themselves show greater variability than do Langley plots from a typical device (Schmid *et al*. 1998). Furthermore, Michalsky *et al*. (1998) showed that even NIST standard lamps do not always agree with each other within the stated accuracy, and that the output of some of these lamps varies significantly over time.

1.4.6: Retrieving Gas Amounts and Aerosol Properties

Traditionally, solving for the aerosol size distribution entails the use of an inversion algorithm based on Mie scattering calculations. Twomey (1977b) identified the aerosol size distribution problem as a Fredholm integral equation of the first kind, and discussed some of the methods used to solve these problems. For their MFRSR analysis, Alexandrov *et al*. (2002a) instead set up a look-up table using a broad range of effective radius values, and five allowable values for the effective variance of a gamma distribution. For a given variance, the radius is incremented until the best fit to the data is determined. Aerosol extinction, ozone, and nitrogen dioxide are solved for simultaneously, along with the calibration of the five non-vapor channels. The Alexandrov *et al*. study assumes that the aerosol size distribution can be adequately described by a single mode. If the distribution happens instead to be bimodal, then this assumption could lead to serious errors in the determination of the nitrogen dioxide and ozone column amounts, which could in turn negatively influence the rest of the retrieval. In particular, overestimating the amount of nitrogen dioxide will lead to an overestimation of the aerosol particle sizes.

The sharp detail of the RSS spectrum between 400 and 450 nm, where NO_2 absorption is strongest, should ensure superior accuracy to the MFRSR in measuring NO_2, and an algorithm was developed for this study to exploit the high resolution of the RSS to determine column amounts of nitrogen dioxide. This should not only result in an improvement of RSS aerosol retrievals, but also serve as a test of the ability of the MFRSR to produce reliable coupled gas/aerosol retrievals. The aerosol retrievals

developed for the RSS also incorporate look-up tables, with the addition of more wavelengths to exploit the greater spectral detail in the RSS device.

The first of the CIMEL device's two detectors measures optical depth from the direct radiance, but the second detector, which measures diffuse sky radiance, is also utilized in the retrieval of the aerosol size distribution (Dubovik and King 2000). This distribution is expressed in the form of the derivative of the total volume of particles of a given radius, with respect to the logarithm of the radius (dV/d ln r), and is calculated at 22 different radius values between 0.05 μm and 15 μm. The units given for this quantity are in μm^3 of total particle volume at a given particle radius per μm^2 of horizontal cross-sectional area in an atmospheric column, and the results of these retrievals are available on the AERONET website. Unlike the size distribution retrieval of Alexandrov *et al.* (2002a), the Dubovik and King algorithm allows for bimodal size distributions. CIMEL devices are not used to retrieve other quantities besides aerosol; instead, assumed values based on climatology are subtracted from the total optical depth in each channel. The Rayleigh optical depth comes from the formulas given in Penndorf (1957), and ozone amounts are obtained from a look-up table (London *et al.* 1976). These ozone amounts are then multiplied by the absorption coefficients from (Vigroux 1953) to determine the ozone optical depth. Nitrogen dioxide absorption is considered to be negligible at the CIMEL wavelengths.

1.5: Retrieval Results from Previous Studies

Alexandrov *et al*. (2002b) examined several years worth of MFRSR data at a number of different sites, including New York City; Albany, NY; Howland, ME; the SGP site in Oklahoma; the Central Plains Experimental Range in Colorado; Davis, CA; and Eugene, OR. Most of the sites showed a strong seasonal cycle in aerosol optical depth, with maximum values in the summer and minimum values in winter or late autumn. For the three sites in the northeastern United States, the annual mean retrieved effective radius was 0.43 μm, with higher values in winter than in summer. The California and Oklahoma sites produced similar results, with slightly higher mean retrieved effective radius values. The ozone cycles measured by the MFRSR showed good qualitative agreement with TOMS satellite data, although the MFRSR-retrieved values in Albany were systematically smaller. Retrieved column NO_2 values tended to be on the order of several Dobson units, and showed a strong correlation with aerosol optical depth values at each site.

The AERONET aerosol climatology (Holben *et al*. 2001) includes Ångstrom coefficient data from the SGP site. Using data from 1994 to 1999, Holben *et al*. found lower values for α in late winter and early spring at SGP, and found larger values for α in late summer and the fall. They calculated the highest values for α at the SGP site in December, but with a limited number of days relative to the other months in the sample. However, the lowest monthly value of α still exceeded 1, meaning that the aerosol extinction steadily decreased with wavelength even for the largest aerosols observed. The effective radius values reported in the Alexandrov *et al*. (2002b) MFRSR

climatology for the SGP site, being large enough to cause an increase in extinction with wavelength at least in the violet part of the spectrum, do not agree with this result.

Prior to this study, no algorithm for retrieving aerosol properties or nitrogen dioxide amounts using RSS data had been developed. A high-resolution retrieval of ozone amounts, though, had been attempted by Harrison (private communication, 2003). The results of this retrieval are included within the data set of optical depth values, averaged over a morning or afternoon, which are used in this study.

1.6: Using EOF's to Analyze the Data

As a first step in this analysis, the Empirical Orthogonal Functions (EOF's) of the optical depth data obtained from the MFRSR, RSS, and CIMEL are calculated. These EOF's facilitate the analysis of the data sets as wholes, as well as one day at a time, offering the potential to develop enhanced retrieval techniques based on objective analysis of the data. The effect that changes in the wavelength combinations used in the data sets will have on the retrieval results can be predicted through a careful examination of the EOF's, and calibration-related differences in the measured optical depth values can be isolated to produce an objective analysis of the different data sets. Finally, the EOF analysis establishes limits to the number of pieces of information that can be independently retrieved from the data, and provides clues as to what these items are.

1.6.1: Calculating Empirical Orthogonal Functions

Mathematically speaking, empirical orthogonal functions are the set of orthonormal vectors that most efficiently represent a given set of data (Peixoto and Oort 1992). For example, say that **F** is an M x N matrix, where M equals the number of mornings and afternoons for which RSS optical depth data exists, and N equals the number of channels in the RSS data. The first step towards determining the EOF's is to calculate the covariance matrix **R**, an M x M matrix which is equal to $(\mathbf{F}*\mathbf{F}^{T})/N$. The covariance matrix can then be diagonalized, and its eigenvalues and corresponding eigenvectors calculated. These eigenvectors are the empirical orthogonal functions. In EOF analysis, the first eigenvalue is always the largest, with the subsequent eigenvalues steadily decreasing in amount. Each of the M eigenvalues explains a fraction of the total variance within the data set, as shown in the formula

$$V_m = \frac{\lambda_m}{\sum_{i=1}^{M} \lambda_i}.$$

Typically, most of the variance can be accounted for in the first few eigenvalues, implying that a few eigenvectors can explain essentially all the variance in the data set. The set of M eigenvectors can be represented as an M x M square matrix **E**, and each eigenvector can be treated as a function of time, or more specifically, of the days in the data set. In addition, each eigenvector has a corresponding set of coefficients of projection, the set of which can be obtained from matrix multiplication:

$$\mathbf{C} = \mathbf{E}^{T} * \mathbf{F}.$$

C, like **F**, is an M x N matrix. Each row of **C** can be treated as a function of wavelength, with the first row corresponding to the first eigenvector, the second row to the second eigenvector, and so on. The full data set **F**, then, can be expressed as the matrix product of **E** and **C**, which algebraically means that **F** is the sum of the matrix products of each eigenvector with its corresponding coefficients.

Empirical orthogonal functions are ideally suited to the analysis of multidimensional data, especially if one of the dimensions is time. They also enable the simultaneous analysis of the complete set of data. All prior studies of aerosols using sun photometer data analyzed the data one day at a time. Most importantly, though, EOF's substantially reduce the number of independent variables in a large set of data. Even the data from a thousand-channel array like the RSS, for example, can be expressed as the sum of a small number of functions.

1.6.2: Objective Analysis of Optical Depth Data

The primary benefit of applying the EOF analysis to the optical depth data obtained from the MFRSR, RSS, and CIMEL is that the EOF's can serve as an objective means of comparing the aerosol properties observed by the three devices. Differences in the optical depth measurements could conceivably mask some common characteristics not only in the aerosols being observed, but in the gases as well. These commonalities, if they exist, will emerge in the EOF analysis. For the aerosols, this is especially true when the EOF analysis is performed after the mean optical depth values for each wavelength

have been subtracted from the data. Any consistent difference in optical depth between the devices, due to calibration or something else, will not affect these EOF's. In addition, as the optical depth contributions of ozone and NO_2 do not vary on the same scale as those for aerosols and water vapor do, ozone and NO_2 essentially disappear from the EOF's with the mean subtracted. EOF's can also serve as a diagnostic tool in the data analysis. For example, if a problem exists in one of the devices, and the magnitude of the optical depth error caused by this problem changes over the course of the data set, then the problem will emerge very distinctly as one of the principal eigenvectors in the EOF analysis.

Box *et al*. (1996) used EOF analysis on synthetic data to set quantitative limits for how much aerosol information can be retrieved from a ground-based multi-spectral device, given a certain level of noise, and to determine what wavelength combination would produce optimal results. Each wavelength combination contained at least seven channels, beginning at 368 nm and extending at least as far as 1030 nm – roughly as broad a range as the RSS, and already significantly broader than that of the MFRSR. With a relative value of 1 assigned to the first eigenvalue, it was concluded that only three eigenvalues in this minimum wavelength range studied would have a relative value greater than 0.01, and only four greater than 0.001. It was also concluded that adding channels within this wavelength range would not increase the available information on the aerosols being observed. This would mean, if verified, that the superior resolution of the RSS would not provide any theoretical advantage over the MFRSR and the CIMEL in measuring aerosol properties, beyond better separation of the gaseous absorption and aerosol extinction. Given seven channels to choose for an aerosol monitoring device so

as to maximize the number of significant eigenvalues, Box *et al*. (1996) chose the wavelengths of 368, 500, 675, 862, 1030, 1725, and 2230 nm.

The Box *et al*. (1996) analysis may have been theoretical, but the question of whether or not the RSS can outperform the other devices in terms of retrieving aerosol information is testable on actual data, and the results will have significant implications for ground-based aerosol monitoring. The size distribution of a given aerosol mode can be described with two quantities, the effective radius and effective variance. The optical depth at a given wavelength of this mode would be a third item of information. If the actual aerosols being observed are unimodal, as the MFRSR algorithm of Alexandrov *et al*. (2002a) assumes, then the limits imposed by the analysis of Box *et al*. would not present a problem. However, the coefficients of projections of the eigenvectors can show if a single-mode distribution is sufficient to explain the aerosol extinction, or if it is indeed necessary to assume a bimodal distribution. If bimodality must be accepted, three important questions need to be addressed. First, what aerosol properties can be retrieved given a limited amount of information? Second, how does the assumption of bimodality affect the retrievals of ozone and nitrogen dioxide? Third, what can be reasonably assumed about the aerosol properties that can't be clearly measured? Finally, to what degree does the accuracy of the retrievals of the gas absorbers in the MFRSR and RSS depend on the initial assumptions about the aerosol properties?

1.7: Summary of Subsequent Chapters

In Chapter 2, the different design philosophies of the MFRSR, RSS, and CIMEL are discussed in detail. The existing caibration and retrieval strategies for each device are placed in the context of this study, and the new algorithms developed for this study are explained explicitly. Chapter 2 also presents a comparison of the optical depth data, from these instruments as functions of both time and wavelength. The results of the EOF analysis, presented in Chapter 3. lead to five important conclusions. First of all, there is a limit to the amount of information obtainable from any of these devices, including the RSS, as previously predicted in an EOF analysis of theoretical data (Box *et al*. 1996). Second, the aerosol size distribution is in fact bimodal. Third, the resolution of the RSS gives an advantage over the other two devices for separating gas absorption from the aerosol extinction, but does not provide any additional aerosol information beyond a clearer separation of the aerosol size modes. Fourth, despite differences in the retrievals based on optical depth, the aerosols observed by all three detectors have similar values for the effective radius of the smaller, or fine, aerosol mode. Thus, the fine mode effective radius is one of the independent quantities that the devices can accurately measure. Fifth, even subtle problems with the measurements will show up in the EOF analysis clearly and completely, and the EOF's can therefore serve as a diagnostic tool to predict when the retrieval will produce flawed results.

These EOF-based results are then further investigated through more detailed radiative transfer analysis of the data. The retrievals assuming a single-mode aerosol size distribution are presented in Chapter 4. The spectral detail provided by the RSS confirms

that retrieving aerosols and gases simultaneously causes large overestimations in the amount of nitrogen dioxide, with corresponding overestimations of the aerosol effective radius. Likewise, ozone absorption cannot be uniquely distinguished from aerosol extinction, either. When gas absorption is accounted for as accurately as possible, it becomes clear that one aerosol mode can not reproduce the aerosol curves observed in the data. Therefore, the retrievals are re-performed, this time assuming a bimodal distribution, and the results are presented in Chapter 5. As predicted by the EOF analysis, the bimodal retrievals need to take into account the limited amount of aerosol information obtainable from the data; otherwise, attempt to retrieve too many independent quantities could obscure important seasonal patterns in the aerosol size distribution.

Given that the current set of filters for the MFRSR cannot uniquely separate the gas and aerosol extinction at shorter wavelengths, the insight provided by the EOF analysis and supported by the radiative transfer analysis is used to investigate if a superior filter set exists. The "MFRSR equivalent" data from the RSS is re-examined in Chapter 6, to select a better filter set for the MFRSR measurements. In this part of the analysis, it is shown that by replacing one channel in the "MFRSR equivalent" retrieval with another at a different wavelength for which filters already exist, it is shown that the existing MFRSR devices can be improved significantly by simply replacing one filter and altering the assumptions inherent in the retrieval strategy. Chapter 6, and the dissertation as a whole, then conclude with a discussion of the results of this study and their implications, and the future research options laid open by these results.

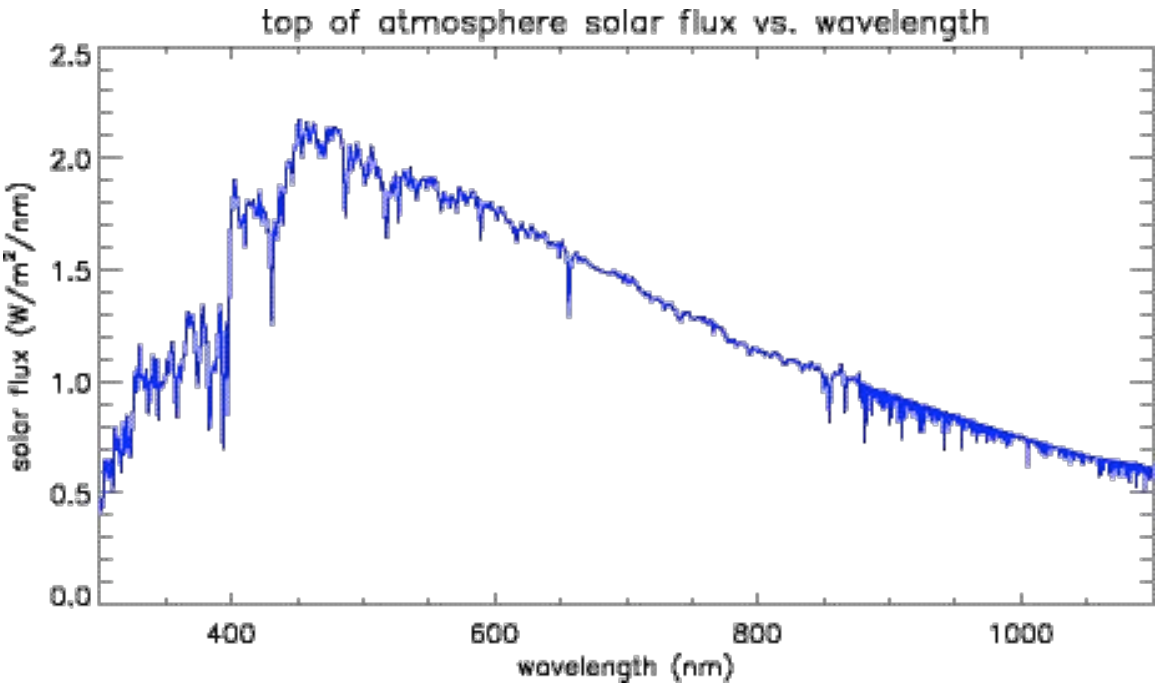

Figure 1.1: The intensity of solar radiation reaching the top of the earth's atmosphere, as a function of wavelength.

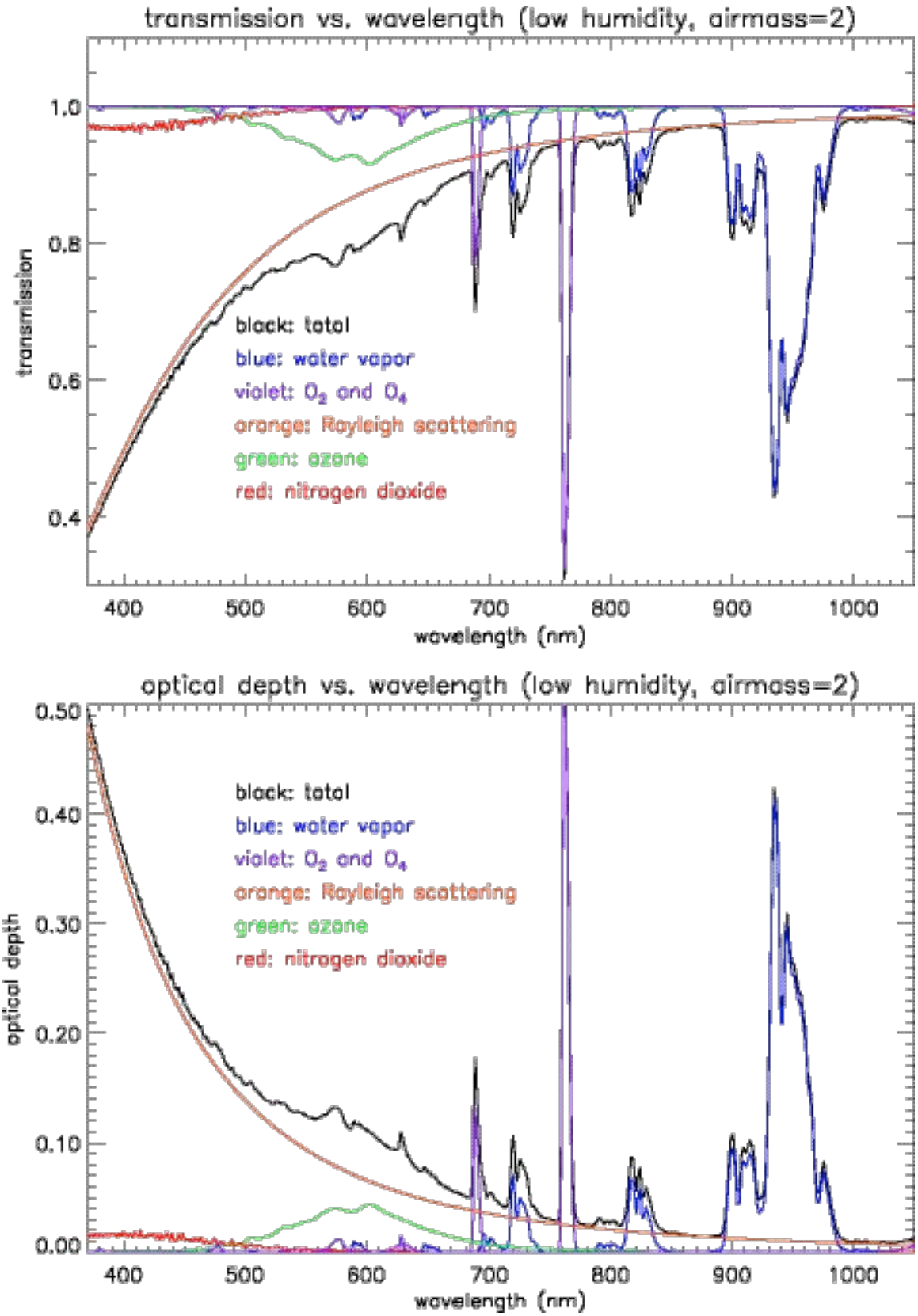

Figure 1.2: The respective contributions to the atmospheric transmission and optical depth of the various gas absorbers and Rayleigh scattering.

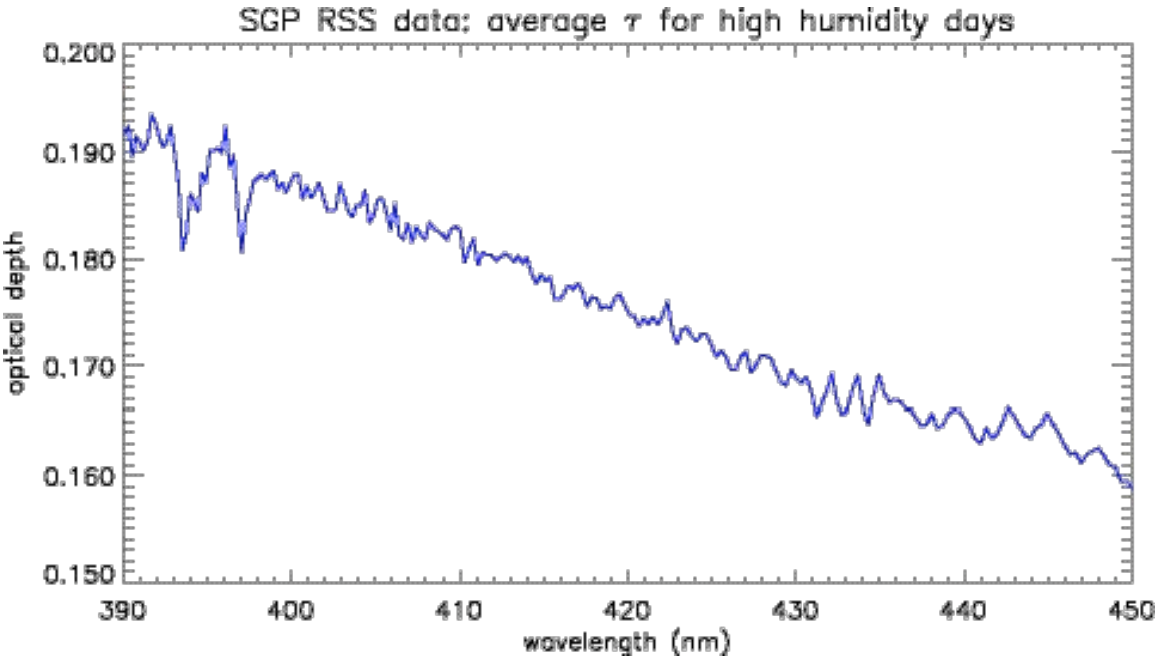

Figure 1.3: The RSS optical depth, averaged over 30 high humidity days, between the wavelengths of 390 and 450 nm. Several prominent features of the Ring effect are apparent in this graph.

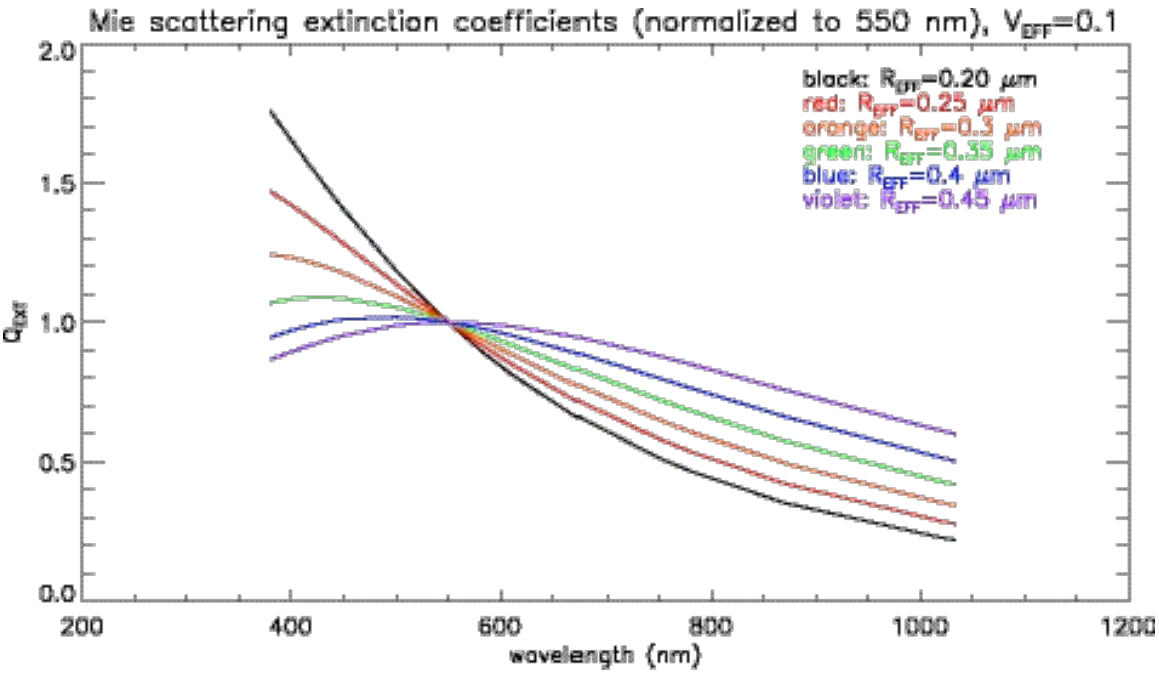

Figure 1.4: The Mie extinction curves as a function of wavelength for a series of different effective radius values, assuming a low effective variance of 0.1.

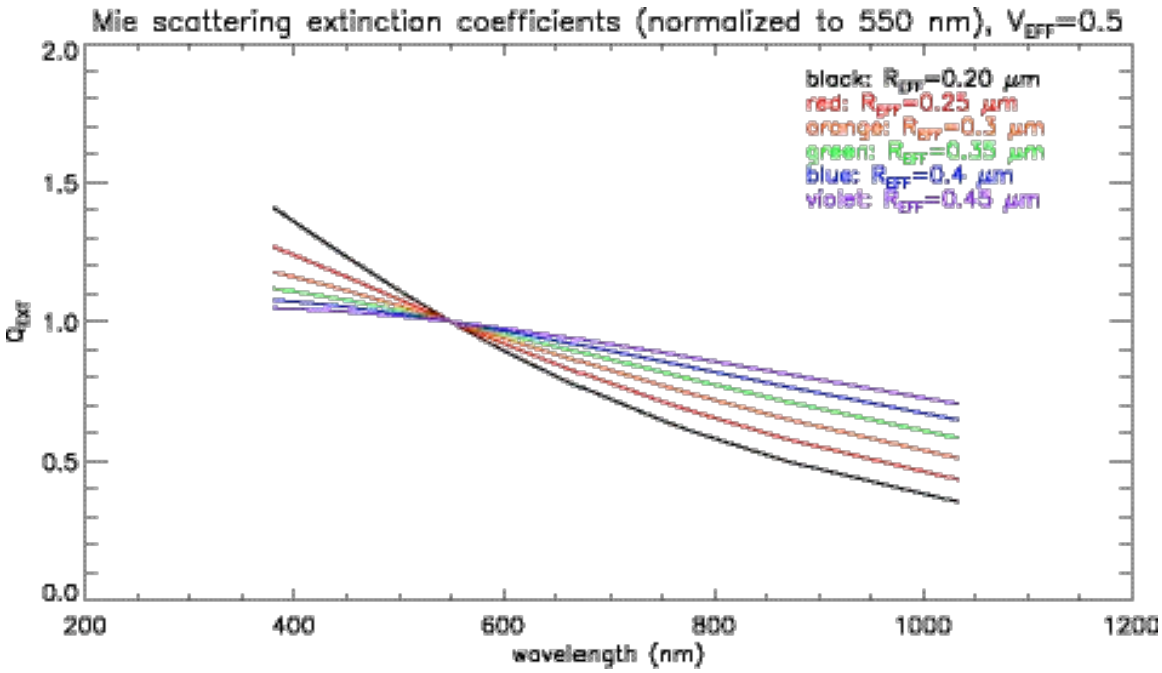

Figure 1.5: The Mie extinction curves as a function of wavelength for a series of different effective radius values, assuming a low effective variance of 0.5.

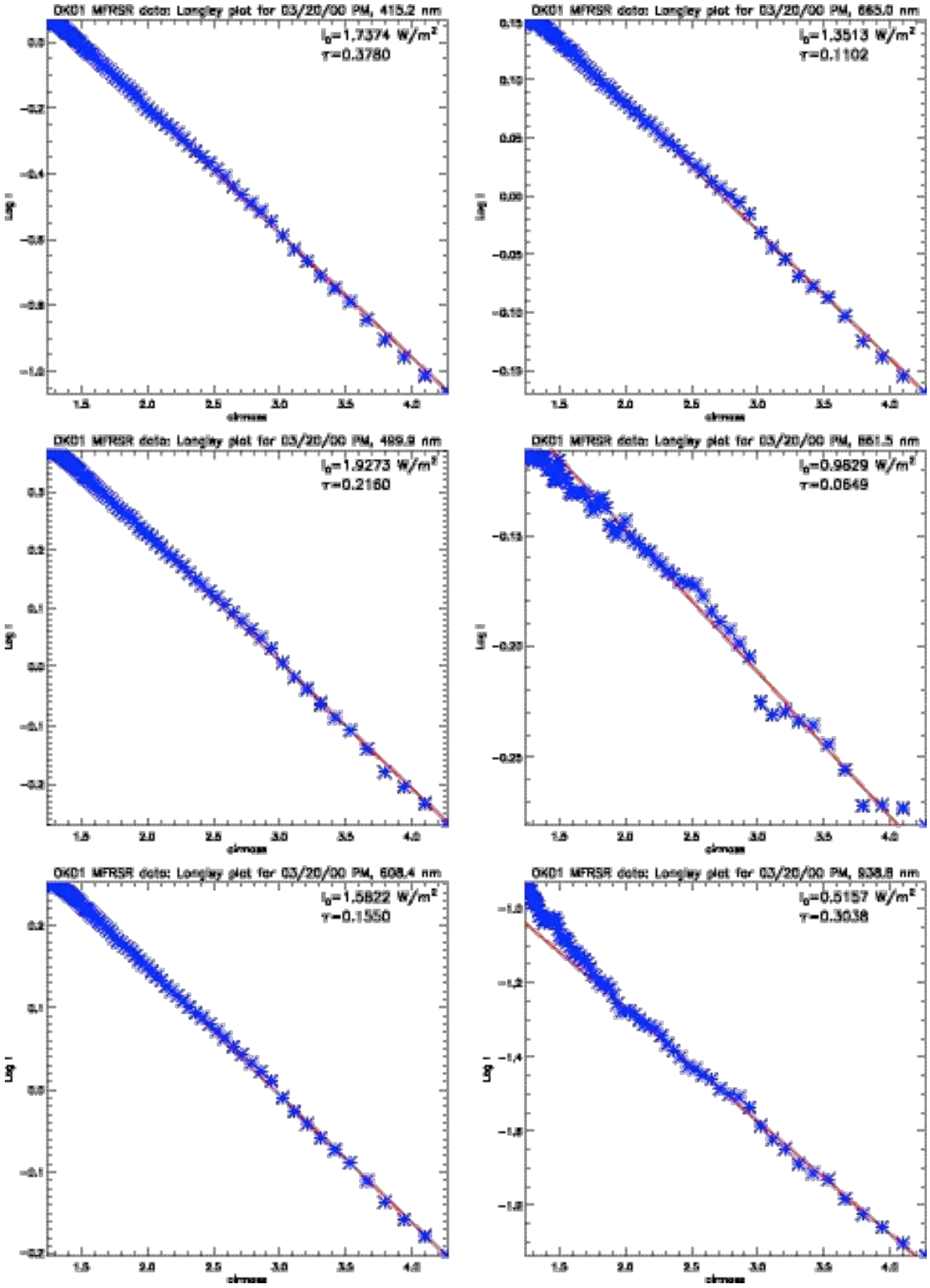

Figure 1.6: Langley plots for the afternoon of March 20, 2000, as measured by the MFRSR device at the SGP site.

Chapter 1 References

Albrecht, *Science* **245**, 1227-1230 (1989).

Alexandrov, M. D., Lacis, A. A., Carlson, B. E., and Cairns, B., *Journal of the Atmospheric Sciences* **59**, 524-543 (2002a).

Alexandrov, M. D., Lacis, A. A., Carlson, B. E., and Cairns, B., *Journal of the Atmospheric Sciences* **59**, 544-566 (2002b).

Bodhaine, B. A., Wood, N. B., Dutton, E. G., and Slusser, J. R, *Journal of Atmospheric and Oceanic Technology* **16**, 1854-1861 (1999).

Box, G. P., Box, M. A., and Krücker, J., *Journal of Geophysical Research* **101**, 19211-19214 (1996).

Bramstedt, K., Gleason, J., Loyola, D., Thomas, W., Bracher, A., Weber, M., and Burrows, J. P., *Atmospheric Chemistry and Physics Discussions* **2**, 1131-1157 (2002).

Bréon, F., Tanré, D., and Generoso, S., *Science* **295**, 834-838 (2002).

Burrows, J. P., Weber, M., Buchwitz, M., Rozanov, V., Ladstätter-Wei□enmayer, A., Richter, A., DeBeek, R., Hoogen, R., Bramstedt, K., Eichman, K., and Eisinger, M., *Journal of the Atmospheric Sciences* **56**, 151-175 (1999).

Eck, T. F., Holben, B. N., Reid, J. S., Dubovik, O., Smirnov, A., O'Neill, N. T., Slutsker, I., and Kinne, S., *Journal of Geophysical Research* **104**, 31333-31349 (1999).

Fish, D. J., and Jones, R. L., *Geophysical Research Letters* **22**, 811-814 (1995).

Forgan, B. W., *Applied Optics* **33**, 4841-4850 (1994).

Grainger, J. F., and Ring, J., *Nature* **193**, 762 (1962).

Green, R. O., *Journal of Geophysical Research* **106**, 17443-17452 (2001).

Greenblatt, G. D., Orlando, J. J., Burkholder, J. B., and Ravishankara, A. R., *Journal of Geophysical Research* **95**, 18577-18582 (1990).

Han, Q., Rossow, W. B., and Lacis, A. A., *Journal of Climate* **7**, 465-497 (1994).

Hansen, J. E., and Travis, L. D., *Space Science Reviews* **16**, 527-610 (1974).

Harrison, L., Michalsky, J., and Berndt, J., *Applied Optics* **33**, 5118-5125 (1994).

Harrison, L., Beauharnois, M., Berndt, J., Kiedron, P., Michalsky, J., and Min, Q., *Geophysical Research Letters* **26**, 1715-1718 (1999).

Hayasaka, T., Meguro, Y., Sasano, Y., and Takamura, T., *Applied Optics* **37**, 961-970 (1998).

Herman, B. M., and Browning, S. R., *Journal of the Atmospheric Sciences* **32**, 1430-1445 (1975).

Higurashi, A., Nakajima, T., Holben, B. N., Smirnov, A., Frouin, R., and Chatenet, B., *Journal of Climate* **13**, 2011-2027 (2000).

Holben, B. N., Eck, T. F., Slutsker, I., Tanré, D., Buis, J. P., Setzer, A., Vermote, E., Reagan, J. A., Kaufman, Y. J., Nakajima, T., Lavenu, F., Jankowiak, I., and Smirnov, A., *Remote Sensing of Environment* **66**, 1-16 (1998).

Holben, B. N., Tanré, D., Smirnov, A., Eck, T. E., Slutsker, I., Abuhassan, N., Newcomb, W. W., Schafer, J. S., Chatenet, B., Lavenu, F., Kaufman, Y. J., Vande Castle, J., Setzer, A., Markham, B., Clark, D., Frouin, R., Halthore, R., Karneli, A., O'Neill, N. T., Pietras, C., Pinker, R. T., Voss, K., and Zibordi, G., *Journal of Geophysical Research* **106**, 12067-12097 (2001).

Intergovernmental Panel on Climate Change, *Climate Change 2001: The Scientific Basis*, Eds. Houghton, J. T., Ding, Y., Griggs, D. J., Noguer, M., van der Linden, P. J., Dai, X., Maskell, K., and Johnson, C. A., Cambridge University Press (2001).

Karl, T. R., Jones, P. D., Knight, R. W., Kukla, G., Plummer, N., Razuvayev, V., Gallo, K. P., Lindseay, J., Charlson, R. J., and Peterson, T. C., *Bulletin of the American Meteorological Society* **74**, 1007-1023 (1993).

Kaufman, Y. J., Tanré, D., Dubovik, O., Karnieli, A., and Remer, L. A., *Geophysical Research Letters* **28**, 1479-1482 (2001).

King, M. D., Byrne, D. M., Herman, B. M., and Reagan, J. A., *Journal of the Atmospheric Sciences* **35**, 2153-2167 (1978).

Lacis, A. A., and Mishchenko, M. I., *Aerosol Forcing of Climate*, ed. R. J. Charlson and J. Heintzenberg, J. Wiley & Sons, pp. 11-42 (1995).

London, J., Bojkov, R. D., Oltmans, S., and Kelly, J. I., *Atlas of the Global Distribution of Total Ozone July 1957-1967*, NCAR Technical Note 133+STR, National Center for Atmospheric Research, Boulder, CO (1976).

Menon, S., Saxena, V. K., Durkee, P., Wenny, B. N., and Nielsen K., *Atmospheric Research* **61**, 169-187 (2002).

Michalsky, J. J., Berndt, J. L., Kiedron, P. W., and Harrison, L. C., *Proceedings of the Seventh ARM Science Team Meeting*, U. S. Dept. of Energy, CONF-9703365, pp. 471-476 (1998).

Michalsky, J. J., Schlemmer, J. A., Berkheiser, W. E., Berndt, J. L., Harrison, L. C., Laulainen, N. S., Larson, N. R., and Barnard, J. C., *Journal of Geophysical Research* **106**, 12099-12107 (2001).

Mishchenko, M. I., Travis, L. D., Kahn, R. A., and West, R. A., *Journal of Geophysical Research* **102**, 16831-16847 (2003).

Penndorf, R., *Journal of the Optical Society of America* **47**, 176-182 (1957).

Schmid, B., Spyak, P. R., Biggar, S. F., Wehrli, C., Sekler, J., Ingold, T., Mätzler, C., and Kämpfer, N., *Applied Optics* **37**, 3923-3941 (1998).

Schmid, B., Michalsky, J., Halthore, R., Beauharnois, M., Harrison, L., Livingston, J., Russell, P., Holben, B., Eck, T., and Smirnov, A., *Geophysical Research Letters* **26**, 2725-2728 (1999).

Schroeder, R., and Davies, J. A., *Atmosphere Ocean* **25(2)**, 109-114 (1987).

Shaw, G. E., Reagan, J. A., and Herman, B. M., *Journal of Applied Meteorology* **12**, 374-380 (1973).

Soufflet, V., Deveaux, C., and Tanré, D., *Applied Optics* **31**, 2154-2162 (1992).Stephens, G. L., *Remote Sensing of the Lower Atmosphere: An Introduction*, Oxford University Press (1994).

Taylor, K. E., and Penner, J. E., *Nature* **369**, 734-737 (1994).

Twomey, S., *Journal of the Atmospheric Sciences* **34**, 1149-1152 (1977a).

Twomey, S., *Introduction to the Mathematics of Inversion in Remote sensing and Indirect Measurements*, Developments in Geomathematics 3, Elsevier (1977b).

Velders, G. J. M., Granier, C., Portmann, R. W., Pfeilsticker, K., Wenig, M., Wagner, T., Platt, U., Richter, A., and Burrows, J. P., *Journal of Geophysical Research* **106**, 12643-12660 (2001).

Vigroux, E., *Annales Physicae* **8**, 709 (1953).

Vountas, M., Rozanov, V. V., and Burrows, J. P., *Journal of Quantitative Spectroscopy and Radiative Transfer* **60**, 943-961 (1998).

Chapter 2: The Data, and the Methods of Analysis

2.1: Sun Photometers

The three instruments used in this study belong to a category of devices called sun photometers. Simply put, sun photometers measure the intensity of sunlight. While some sun photometers also measure diffuse, or scattered, solar radiation, sun photometers primarily measure the direct beam, or the portion of the solar radiation that travels from the sun to the detector on an unaltered path. Some sun photometers measure sunlight at one wavelength, or perhaps take one measurement over the full spectrum, while others use filters to look at several or many different wavelengths. In addition, some sun photometers are hand-held, while others are securely fastened to a fixed location and automated. The devices in this study measure the intensity of sunlight at multiple wavelengths at a fixed location. In the case of the MFRSR and the CIMEL, a handful of filters isolate sections of the spectrum approximately 10 nm in width. The RSS, by contrast, incorporates a diffraction grating and an array of over 1000 channels to produce high-resolution spectral data. Either way, measuring sunlight at multiple wavelengths makes possible the analysis of the aerosol size distribution and the measurement of column amounts of absorbing gases like ozone and nitrogen dioxide.

The science of sun photometry originated in the nineteenth century with Ångstrom (1894). Ångstrom placed two identical strips of metal near each other, and when one strip was exposed to solar radiation and thereby heated, the temperature imbalance between the two strips caused a current to flow from the hot strip to the cold

one. As this current could be easily measured, the energy difference that caused it could be calculated as well. The first hand-held, multi-channel sun photometers were developed by Volz (1957). By the sixties, these sun photometers were measuring turbidity, or aerosol concentrations, in networks of stations situated across the United States and Western Europe (Volz 1969). Flowers *et al*. (1969) constructed a climatology of turbidity using five years of data from the American network of Volz sun photometers. Among other things, the results indicated that the lowest turbidity occurred in the western plains while the highest occurred in the east, and that turbidity follows an annual cycle, which peaks in the summer and reaches a minimum in winter. By this time, it had already become clear that aerosol size distributions needed to be measured in addition to turbidity, and that fully reliable aerosol measurements required greater precision over longer time periods than were possible with hand-held devices. Shaw *et al*. (1973) developed an automated device called a filter-wheel radiometer, which could record data at multiple wavelengths when left unattended for long stretches of the day. The first device physically similar to the MFRSR and RSS was called the dial radiometer (Wesely 1982). Developed at Argonne National Laboratory, the dial radiometer used a thin, rotating shade to sweep a circle around the device at 5-minute intervals, enabling direct and diffuse intensities to be measured by the same detector. Unfortunately, the device underestimated the diffuse and overestimated the direct, and corrections needed to be introduced. An Italian group (Guzzi *et al*. 1985) also developed a ground-based spectroradiometer that used a rotating shadow band to separate direct and diffuse intensities.

2.2: The Multi-Filter Rotating Shadowband Radiometer

2.2.1: The Design Philosophy

The Multi-Filter Rotating Shadowband Radiometer (MFRSR) is a ground-based, automated sun photometer that measures sunlight at six different wavelengths (Harrison *et al*. 1994). The MFRSR measures the total and diffuse intensities with the same detector, eliminating any calibration-related discrepancies between the two quantities that might arise in other instruments. From these two measurements, the direct beam can then be calculated. The total and diffuse intensities are separated from each other through the use of the rotating shadow band, which rotates to four different positions during a measurement interval. The detector itself does not change its position. In step one, the shadow band is situated below the detector, so that the detector measures the total incoming solar flux. The next three steps are made in sequence, where the band blocks out the sun in the middle one and is positioned on either side of the sun in the other two. The two side-of-sun measurements are averaged, and then subtracted from the total, to give the portion of the diffuse radiation that hits the detector from the direction of the sun. This quantity is then added to the sun-blocked measurement to yield the diffuse sky radiance. Subtracting the diffuse from the total gives the vertical component of the direct beam intensity. The magnitude of the direct intensity can then be calculated by dividing its vertical component by the cosine of the solar zenith angle. By enabling the measurement of multiple quantities with one detector, the shadow band makes the

MFRSR a highly cost-effective device. The CIMEL, by comparison, costs about three times as much.

The six filters of the MFRSR have central wavelengths at approximately 415, 500, 610, 670, 870, and 935 nm, and bandwidths of 10 nm full width, half maximum. 415 nm is the location of the absorption peak of nitrogen dioxide, which allows, at least in principle, for the best possible calculation of column NO_2 amounts given a low-resolution device. However, as the device does not have any channels located in the ultraviolet region of the spectrum, where NO_2 absorption weakens, the individual extinctions of both small particle aerosols and NO_2 peak in the same filter. As noted by Alexandrov *et al*. (2002a), this makes the respective contributions of NO_2 and aerosols difficult to accurately separate in practice, and can have significant bearing on the quality of the retrieved aerosol size distribution. The 500 nm filter is at a spectral location where the absorption of NO_2 starts to weaken, while absorption in the Chappuis band of ozone begins to get strong. Ozone absorption peaks at 610 nm, where the next filter is situated, and declines from the peak by about half at 670 nm. The 870 nm channel is the most important channel for the specific analysis of aerosols, because gaseous absorption at that wavelength is entirely negligible. For this reason, and because there are no calibration differences between the direct and diffuse intensities in the MFRSR, Alexandrov *et al*. used the direct/diffuse ratio of the 870 nm as the starting point in their calibration algorithm. The 935 nm channel is used to measure column amounts of water vapor, which absorbs especially strongly at that wavelength.

Figure 2.1 shows the direct and diffuse intensities, taken over the course of March 20, 2000, measured by the MFRSR device located at the SGP site. For clear-sky

conditions, the diffuse will always be lower than the direct at MFRSR wavelengths. However, the addition of scattering particles will decrease the direct intensity while increasing the diffuse. The sharp, $1/\lambda^4$ wavelength dependence of Rayleigh extinction accounts for the amount of diffuse radiation, relative to the direct, being substantially larger at the shorter wavelengths.

The MFRSR device was developed for use in the Atmospheric Radiation Measurement (ARM) program, and a number of these devices have since been used at sites across the country. They are primarily used to retrieve aerosol optical depth values. Multiyear data sets of aerosol optical depth exist for several locations. This data has been used for a number of purposes. In addition to the climatologies reported by Alexandrov *et al*. (2002b), MFRSR's have served as inputs to radiative transfer models to test their accuracy against measured radiance values (Michalsky *et al*. 2001). The MFRSR has also been used to retrieve column amounts of water vapor and ozone (Michalsky *et al*. 1995), and cloud optical depth (Leontieva and Stamnes 1996). The ultraviolet version of the device, called the UV-MFRSR, has been used specifically to retrieve column ozone amounts (Slusser *et al*. 1999) as part of the U. S. Department of Agriculture's UV-B measurement program.

2.2.2: The MFRSR Retrieval Algorithm

Retrievals were obtained from the MFRSR data using the Alexandrov *et al*. (2002a) algorithm. This algorithm is based on the assumptions that the aerosol size distribution has only one mode, and that due to the lack of independent information

present in the channel at 670 nm, the effective variance cannot be specified. The retrievals were made assuming a gamma size distribution, using effective variance values of 0.01, 0.1, 0.2, 0.3, and 0.4. It is postulated that the actual effective variance lies in between the extreme values, and consequently the actual values for the column amounts of the gases and the aerosol optical depth and effective radius fall between the values retrieved for the extreme variance cases. Because of the nature of the Alexandrov *et al*. algorithm, the calibration values of the MFRSR channels are adjusted for each variance in a way that eliminates any residual optical depth that might otherwise appear in the data. Mean values of the aerosol optical depth, effective radius, column ozone, and column nitrogen dioxide are calculated for each variance during each clear morning or afternoon. As a general rule, the retrieved values for each quantity vary monotonically with variance, either steadily increasing or steadily decreasing. This study will focus on the results for the retrievals assuming the highest effective variance used, 0.4, since these have been found to provide the best agreement with correlative data.

2.3: The Rotating Shadowband Spectroradiometer

2.3.1: Properties and Applications

The Rotating Shadowband Spectroradiometer (RSS) is somewhat similar to the MFRSR in that it too has a shadow band that enables simultaneous calculations of total, direct, and diffuse radiances (Harrison *et al*. 1999). However, the RSS is designed for high-resolution spectral analysis. Current versions of the RSS have 1016 active channels,

ranging from 360 to 1150 nm in wavelength. While the channels in the filter devices have bandwidths of 10 nm, the RSS channels have bandwidths as small as 0.3 nm at the lower wavelengths. The bandwidth increases steadily with wavelength, but does not exceed 3 nm even at the maximum wavelength. The utility of a high-resolution approach is illustrated in Figure 2.2. The average optical depth over an afternoon where column NO_2 amounts were particularly high is shown between 415 and 500 nm, the wavelengths that correspond to the first two MFRSR channels. The degree of spectral detail makes the signal of nitrogen dioxide emerge much more clearly in the RSS than it possibly could in the MFRSR. However, this higher spectral resolution comes at a price, as the noise level of the measurements can limit their utility, as shown in Figure 2.3. Very little definitive can be said about the column NO_2 amount for this day, other than that it cannot be very high. In fact, the most significant feature present within the data, the jagged pattern extending from about 430 to 435 nm, is attributable to the filling in of solar absorption lines resulting from the Ring effect. Nevertheless, the higher resolution of the RSS makes it generally possible to uniquely separate aerosol and gas effects at the shorter wavelengths. This advantage is tempered, though, by a much higher cost and a lack of compactness and portability.

In addition to the comparative aerosol study of Schmid *et al.* (1999), RSS data have also been used to retrieve column water vapor amounts by analyzing spectral data in the 940 nm absorption band and matching the measurements with output from a radiative transfer model (Kiedron *et al.* 2001). The results show qualitative agreements with four different microwave radiometers, but significant quantitative differences persist in the measurements, which have yet to be fully reconciled. Otherwise, with only two devices

currently in operation and a substantial cost to produce additional ones, the full potential of the RSS has yet to be realized. No published attempt to determine aerosol size distributions or column NO_2 amounts using RSS data exists, prior to this study.

2.3.2: Single-mode Aerosol Retrievals

Given that the RSS data were already provided in optical depth form, no additional calibration was necessary. However, as no algorithm designed specifically to retrieve aerosol size distributions from RSS data previously existed, one needed to be designed. For starters, the Mie scattering look-up tables used in the Alexandrov *et al.* algorithm are expanded, with some additional wavelengths incorporated to accommodate the greater number of channels, plus a greater range of effective radius values and a high variance of 0.5 added to accommodate the possibility of a bimodal distribution. The wavelengths used in the full RSS aerosol retrieval are 380, 400, 415, 440, 450, 500, 525, 550, 585, 610, 670, 750, 780, 870, 1020, and 1034 nm. All the MFRSR and CIMEL wavelengths are included, along with additional wavelengths located at gaps, or "windows," in the water vapor absorption spectrum. Windows such as those at 750 and 780 nm are broad enough that an RSS channel located within will not be affected by any stray water vapor absorption, but too narrow to place filters of 10 nm bandwidth like those of the MFRSR and CIMEL.

Because optical depth is additive, and the aerosol and gas optical depths can be expressed as the products of a coefficient and the column amount, the total optical depth at a given wavelength can be defined by the simple linear equation

$$\tau_\lambda = q_\lambda \tau_{a550} + \alpha_\lambda * NO_2 + \beta_\lambda * O_3,$$

where q_λ is the aerosol extinction coefficient normalized to unity at 550 nm; τ_{a550} is the aerosol optical depth at 550 nm; and α_λ and β_λ are the absorption coefficients of nitrogen dioxide and ozone, respectively. For a given aerosol effective radius and variance, the extinction coefficient can be determined with the lookup table, and the gas absorption coefficients are known functions of wavelengths. The full set of the optical depth equations can be collectively expressed as a matrix equation of the form

$$\mathbf{A*x=b}.$$

A, in this case, is a matrix containing the extinction coefficients of the aerosols and the absorption coefficients of nitrogen dioxide and ozone. The vector **b** is comprised of the optical depth values measured by the RSS. The solution vector **x** consists of the aerosol optical depth at 550 nm, along with the column amounts of ozone and nitrogen dioxide, and can be calculated by means of a least-squares regression. In the cases where single mode distributions are assumed, a series of iterations is performed for effective variance values of 0.01, 0.1, 0.2, 0.3, and 0.4, and 0.5. The effective radius is steadily increased, with the Mie scattering coefficients adjusted accordingly. For each effective radius increment, the least squares technique is performed, and the residual optical depth values are calculated by subtracting the components of **b**, the measured optical depths, by the components of **A*x**, the calculated optical depths. The RMS residual is calculated for each radius value, and the iteration continues until the RMS residual for the new radius is greater than that for the previous radius, indicating that the minimum residual for that variance has been located. The solution vector that produces this minimum residual is

considered the best fit for the given effective variance. After the process is completed for all six variance values, the solution vector and corresponding variance that produce the smallest minimum residual are output as the retrieved solution.

2.3.3: Isolated Retrieval for Nitrogen Dioxide

One of the advantages of a high-resolution device like the RSS is that it is not necessary to retrieve gas amounts by looking at a handful of points covering the entire spectral range of the device. Instead, spectral regions particularly sensitive to the gas in question can be isolated, and retrievals can be performed using the detailed information contained within this region. This is essentially the approach used with ground-based devices called Brewer and Dobson spectrometers to retrieve column amounts of ozone and nitrogen dioxide. The primary difference between the RSS and the Brewer and Dobson devices is that the RSS measures light primarily in the visible spectrum, while the Brewer and Dobson devices measure ultraviolet light. The specific technique used to measure gases with the Brewer and Dobson devices is called Differential Optical Absorption Spectroscopy, or DOAS (Platt *et al*. 1979). DOAS looks at wavelengths where the absorption spectrum of a gas varies particularly rapidly relative to the smooth continuum of aerosol extinction, and uses the difference in intensity measured in this spectral region to calculate the amount of that gas present in an atmospheric column.

The mean RSS-measured optical depth values for each morning or afternoon in the data set were accompanied by a retrieved value for ozone, based on an analysis of the details of the Chappuis band (Harrison, private communication, 2003). Unfortunately for

visible spectrum devices, the Chappuis absorption band lacks the both strength and the sharp variation of the ultraviolet ozone absorption bands. This makes the measurement of ozone using only visible spectral data rather difficult, and the ability to separate the extinction due to ozone from that due to aerosols depends largely on the accuracy of the assumptions made about the aerosols. Nitrogen dioxide values, on the other hand, were not included with the data. Given the substantial detail in the NO_2 absorption spectrum easily observable by the RSS, a high-resolution NO_2 retrieval would serve as a valuable test of the accuracy of the low-resolution, coupled gas/aerosol retrievals. This, in turn, would provide a means of assessing the quality of the lower-resolution MFRSR retrievals. Therefore, a retrieval algorithm specifically for measuring nitrogen dioxide with the RSS was designed and then implemented for this study.

The first step in retrieving nitrogen dioxide using the RSS is to account for the significant amount of random noise inherent in the device. In general, the degree of random noise is large enough to substantially compromise the accuracy of a straightforward DOAS retrieval. Figure 2.4 shows the optical depth measured by the RSS on the morning of June 23, 2000 between the wavelengths of 366 and 466 nm, where NO_2 absorption is most significant. The plot shows three distinct noise regions. The first region, where the data is noisiest, extends to 397 nm, which also happens to be the location of a significant Fraunhofer line. The next region, between 397 and 436 nm, shows an intermediate amount of noise, while the final region has a relatively small amount of noise. In general, it is only in the third region where the absorption signal of NO_2 rises noticeably above the noise. However, when the mean optical depth values for the entire data set are plotted, as in Figure 2.5, the details of the NO_2 spectrum show up

much more clearly throughout this portion of the spectrum. This suggests that most of the noise, excepting the Ring effect, is random in nature and can be reduced through averaging. However, since averaging adjacent data points would eliminate the detail in the nitrogen dioxide spectrum, a different approach needs to be taken.

To optimize the accuracy of the retrieval, the points strongly affected by the Ring effect are removed from consideration. The rest of the procedure owes more to the k-distribution method of Lacis and Oinas (1991) than it does to DOAS techniques. For each noise region, the wavelengths are grouped by their value of the NO_2 absorption coefficient. For example, in the low noise region, the six points with the highest values of the NO_2 absorption coefficient are averaged together, followed by the next six points, and so on. This means that the data points, already averaged over time, are now averaged spectrally as well. Twelve new data points are created from the averaged values in the low noise region, along with sixteen in the medium region and two in the high region, for thirty total. These data points are then used in a least squares regression. The curve for the aerosol extinction in this region is expressed as a third order polynomial function, but not of wavelength directly. To illustrate the need for this distinction, consider that a third-order aerosol fit with wavelength would produce the following optical depth equation at a given channel:

$$\tau_{\lambda} = \sigma_{\lambda} N + c_0 + c_1 \lambda + c_2 \lambda^2 + c_3 \lambda^3 .$$

Averaging this equation over multiple, non-adjacent wavelengths produces the following equation:

$$\overline{\tau_\lambda} = \overline{\tau_\lambda} N + c_0 + c_1 \overline{\lambda} + c_2 \overline{\lambda^2} + c_3 \overline{\lambda^3}.$$

The mathematically correct way to set up the matrix requires the first-order term to be the average of the corresponding wavelengths, the second-order term to be the average of the squares of the wavelengths (not the square of the average wavelength), and the third-order term to be the average of the cubes of the wavelengths. The least-squares regression then solves for c_0, c_1, c_2, c_3, and N. Figure 2.6 shows the averaged optical depth points, with the fitted aerosol extinction amounts subtracted out, plotted against the averaged NO_2 absorption coefficients for the morning of January 22, 2000. The initial retrieved NO_2 amount is given, along with an estimated error derived from the deviations of the plotted points from the line of best fit. In this particular example, a couple of points seemed to deviate from the line significantly. As this proved to be a common occurrence within the data set, an additional step was added to the algorithm, where the least squares regression is repeated after points located more than two standard deviations from the best-fit line are removed. The new, and final, fit for the morning of January 22 is shown in Figure 2.7.

2.3.4: MFRSR and CIMEL "Equivalent" Retrievals

Besides the aerosol retrievals using many channels, and the NO_2 retrieval utilizing the full resolution, subsets of the RSS data can produce MFRSR and CIMEL "equivalent" retrievals. Wavelengths corresponding to the five non-water filters of the MFRSR, or the six CIMEL filters within the wavelength range of the RSS, can be

isolated from the rest of the RSS data. The retrieval algorithms devised for sixteen channels need to be modified only slightly to accommodate fewer channels. Assuming that the retrieval with the largest number of channels produces the best results, the "equivalent" retrievals will show how much error is introduced just by using a smaller number of wavelengths, without needing to refer to the other devices directly. The primary advantages of this approach are that it eliminates calibration differences as a source of retrieval error, and the higher resolution of the RSS channels eliminates the possibility of extraneous gaseous absorption adding to the total optical depth in a given channel.

Since one of the goals of this study is to assess and improve the accuracy of the MFRSR retrievals, the "MFRSR equivalent" approach can be combined with the broader spectral range of the RSS to see if a different set of filters would improve the accuracy of the MFRSR retrievals, especially regarding the ability to break the uniqueness issues between small particle aerosol extinction and NO_2 absorption. This study therefore concludes with a re-examination of the "MFRSR equivalent" retrievals, this time assuming that the aerosol size distribution is bimodal. The revised set of assumptions should improve the quality of the retrievals using existing MFRSR data, relative to the assumptions used in Alexandrov *et al*. (2002a). The differences between this retrieval and the 16-channel RSS retrieval indicate the limits to the accuracy that can be obtained using the MFRSR with its current combination of filters. The "MFRSR equivalent" retrievals are then altered further, by replacing the 670 nm channel first with one at 375 nm, and then with one at 1034 nm, to see which one of these wavelengths best improves the agreement with the 16-channel retrieval. As filters with these approximate

wavelengths already exist, making such an alteration to existing MFRSR devices would be minor and inexpensive.

2.4: The CIMEL Electronique 318A Spectral Radiometer

2.4.1: The CIMEL Design Philosophy

The CIMEL Electronique 318A spectral radiometer, used in the AERONET network (Holben *et al*. 1998), differs structurally from the rotating shadow band devices in two ways. First, the direct beam is measured, along with the solar aureole (the diffraction ring around the sun), by means of a detector that tracks the sun across the sky and has a field of view of 1.2°. Second, sky radiance is also measured, at a series of specific angular distances from the sun, using an additional detector. These multi-angle measurements do not provide a complete picture of the total and diffuse radiation, but they do provide additional information with which to more precisely define aerosol optical properties.

The wavelengths used in the CIMEL devices are selected to optimize the retrievability of aerosol amounts and properties specifically, by minimizing the contributions of ozone and nitrogen dioxide. The CIMEL devices used by AERONET have eight filters, with central wavelengths of 340, 380, 440, 500, 670, 870, 940, and 1020 nm. In contrast to the MFRSR filter selection, the wavelengths of the CIMEL filters are chosen to avoid regions of strong gaseous absorption. With the exception of retrieving water vapor from the 940 nm channel data, the CIMEL cannot be used to

retrieve gas abundances. Each filter of the CIMEL has a bandwidth of approximately 10 nm full width, half maximum. Direct measurements are taken at all wavelengths, and all channels except the 940 nm channel, where water vapor absorption is very strong, are used in the aerosol retrievals. By avoiding absorption peaks, it is assumed that any errors caused by using a climatological value for ozone or by neglecting nitrogen dioxide will be negligible. The multi-angle sky radiances are measured at 440, 670, 870, and 1020 nm.

A potential weakness of the multi-filter devices, especially when the wavelength range extends well into the ultraviolet like the CIMEL does, is that the filters located at different wavelengths are physically dissimilar. For example, Figure 2.8 shows the top-of atmosphere solar flux, with the locations of the MFRSR and CIMEL filters marked by asterisks. If the devices were located at the top of the atmosphere, the intensity passing through the filters at either end of the solar spectrum would be about half what it is at the peak. However, much less intensity reaches the ground at the ultraviolet wavelengths, due to the spectral dependence of Rayleigh scattering. Therefore, the filters at these wavelengths need to be more sensitive, with the gains adjusted accordingly, introducing an additional source of error affecting the measured intensity in the low-wavelength filters. Obviously, errors in the extreme wavelengths used in an aerosol retrieval will lead to the larger errors in the aerosol particle size than errors in the central wavelengths will.

NASA developed the Aerosol Robotic Network, or AERONET, to augment its satellite monitoring systems, on the grounds that “the prospect of fully understanding aerosols’ influence on climate forcing is small without validation and augmentation by

ancillary ground-based observations as can be provided by radiometers historically known as sun photometers (Holben *et al*. 1998)." Consisting of over twenty American sites and dozens of additional sites around the world, AERONET aims to provide as comprehensive a global aerosol climatology as can be accomplished using ground-based devices. AERONET's instrument of choice is the CIMEL device. These devices are not used to retrieve other quantities besides aerosol; instead, assumed values are subtracted from the total optical depth in each channel. Dubovik and King (2000) developed the algorithm used by AERONET to retrieve aerosol optical properties, such as the size distribution, single-scattering albedo, and index of refraction, using these optical depth values combined with the multi-angle sky radiances also measured by the CIMEL device. Holben *et al*. (2001) presented the AERONET climatology results up to that point. Their values for the aerosol optical depth at the Southern Great Plains (SGP) site in Oklahoma showed the same seasonal cycle for aerosol optical depth observed in Alexandrov *et al*. (2002b), even though their values for the Ångstrom parameter retrieved with the CIMEL device do not show good quantitative or qualitative agreement with the MFRSR-retrieved size distributions.

2.4.2: Optical Depth Retrievals

The AERONET retrieval algorithm of Dubovik and King (2000) uses the optical depth and sky radiances from four of the channels, but for the sake of more direct comparison with the MFRSR and RSS, retrievals using only optical depth values were developed in this study, incorporating the same general procedures as the retrievals

designed for the RSS. Minor modifications to the program were necessary to account for the removal of ozone and the fact that the CIMEL retrieval ignores the spectral contribution from NO_2 absorption. Unfortunately, a number of problems exist in the CIMEL data. The most obvious of these problems occurs in the 670 nm channel. Figure 2.9 shows the mean measured aerosol optical depth values for the morning of February 27, 2000. This day, like many other days in the data set, shows an unnatural dip in the optical depth at 670 nm. As these dips are far too large to be explained by an overestimation of ozone absorption, it was concluded that the 670 nm channel is in some way defective, and the channel was disregarded in all of the retrievals. In addition, many days showed a slight bulge in the measured optical depth at 440 nm. This is shown clearly in Figure 2.10, with data taken on the afternoon of November 2, 1999, a day in which the aerosol extinction was especially low. Perhaps this bulge in the data is also caused by a defect in the channel, but the omission of NO_2 absorption could also produce this result. The 340 nm channel was also disregarded for the retrievals, not because it was defective, but because the RSS does not extend to that low a wavelength, and the results of the two devices are more easily comparable without it.

2.4.3: The Almucantar Retrievals

The CIMEL device has two detectors, the second of which measures diffuse sky radiance (Holben *et al*. 1998). The sky radiance detector operates at the wavelengths of 440 nm, 670 nm, 870 nm, and 1020 nm. It also has two different observation sequences, called almucantar and principal plane. The almucantar measurements are taken at the

same solar zenith angle as the sun, at a series of different scattering angles. These measurements, combined with the corresponding optical depth values, are then input into the algorithm of Dubovik and King (2000), through which the aerosol size distribution is calculated. This distribution is expressed in the form of the derivative of the total volume of particles of a given radius, with respect to the logarithm of the radius (dV/d ln r), and is calculated at 22 different radius values between 0.05 μm and 15 μm. The units given for this quantity are in μm^3 of total particle volume at a given particle radius per μm^2 of horizontal cross-sectional area in an atmospheric column. The results of these retrievals are available on the AERONET website. The almucantar algorithm produces 22 values of output, though. While the additional angles may increase the amount of aerosol information retrievable beyond the limit of three set by Box *et al*. (1996) on direct beam data in the CIMEL wavelength range, it is not likely that there are 22 independent items of information obtainable from the data. The algorithm does, however, make a number of assumptions in order to constrain the shape that the size distribution can take. One of these assumptions, based on the findings of Remer and Kaufman (1998), is that the aerosol size distribution is bimodal at a minimum, and that other modes may possibly exist as well. This clearly contrasts with the assumption made in the MFRSR retrievals of Alexandrov *et al*. (2002a). In this study, these assumptions will be tested, to determine if one approach provides superior results to the other. In order to compare the almucantar retrievals more directly to the retrievals from the CIMEL and RSS optical depth data, the size distributions were converted to coarse and fine mode optical depth, effective radius, and effective variance values using numerical integration, based on equations in Lacis and Mishchenko (1995).

2.5: Calibration Strategies

With the exception of the pristine environment at high-altitude sites like Mauna Loa, aerosol amounts usually change significantly over the course of a morning or afternoon, affecting the results of the Langley regressions than can be performed on the data (Shaw *et al*. 1973). As a result, optimizing the accuracy of sun photometer calibrations requires going beyond the assumption that a given day's Langley regression produces the most accurate value for I_0 obtainable for that day's data. The three devices used in this study each have their own strategy for calibrating the data.

Alexandrov *et al*. (2002a) developed a new method for calibrating MFRSR devices, not involving Langley regression, based on two assumptions. First, the fifth channel (~865 nm), which does not contain any significant amounts of gaseous absorption, can be more accurately calibrated by comparing the day's direct/diffuse ratios, whose values do not depend on the calibration, with the results predicted by a computer simulation. The second assumption, based on the studies of Soufflet *et al*. (1992) and Forgan (1994), is that the aerosol size distribution remains more constant during the day than the aerosol optical depth does, and therefore provides a more stable means of calibration than Langley regressions do. The algorithm of Alexandrov *et al*. determines the calibration of the filters while simultaneously calculating the aerosol amount, the effective radius and variance, and column amounts of ozone and nitrogen dioxide. This calibration technique has shown greater stability – that is, less deviation – than standard Langley calibrations over the same data set. However, the algorithm assumes that the aerosol size distribution has a single mode, introducing a potentially

significant source of error. The MFRSR data in used in this study was calibrated using the Alexandrov *et al*. algorithm.

Michalsky *et al*. (2001) devised a multi-step calibration procedure for the ARM RSS and MFRSR devices that begins with obtaining all the Langley regressions for a data set, as before. To find the precise calibration on a given day, the 20 nearest Langley-retrieved I_0 values are collected. From these, the days which provide the 10 median values of the ratio $I_0(500\ nm)/I_0(862\ nm)$, considered to be an indicator of stability, are selected, and their calibration values averaged. These averaged values are then plotted on a graph spanning the entire dataset. The curve that best fits the data is then used as the final set of calibration values. The RSS data used in this study consists of morning and afternoon averages of optical depths obtained using this calibration strategy.

Instead of relying on in situ calibrations, every filter head used in AERONET's CIMEL devices is periodically removed from its device and shipped to Mauna Loa in Hawaii to be calibrated, while another recently calibrated head is inserted in its place in the device. Mauna Loa, being located at a very high altitude, is the least polluted research site in the United States, and the Langley regressions made there are far more consistent than any that can be made elsewhere. Unfortunately, the properties of a set of filters used in a given device can sometimes change very rapidly, and the calibrations obtained at Mauna Loa may only be accurate for a short period of time.

2.6: The Data

2.6.1: The Data Sets Used in This Study

The SGP site in Oklahoma houses a wide variety of monitoring devices, making comparisons between these devices relatively simple to perform. The data used in this study come from an MFRSR device belonging to the USDA UV-B Monitoring and Research Program (Bigelow *et al.* 1998), an RSS belonging to the Atmospheric Radiation Measurement (ARM) program of the Department of Energy, and a CIMEL device belonging to AERONET. The data sets span a year, from July 1999 to July 2000. The direct intensities for the five non-water channels of MFRSR data were simultaneously calibrated and processed into retrievals for ozone, nitrogen dioxide, and the aerosol optical depth and size distribution on a day-to-day basis, using the technique of Alexandrov *et al.* (2002a). The RSS data consists of morning or afternoon-averaged optical depth values. This data had previously been calibrated by the Michalsky *et al.* (2001) method. In addition, the total, direct, and diffuse intensities were examined for a couple of days, primarily to examine the influence of the Ring effect on the RSS data. The CIMEL data consists of level 2 quality-screened aerosol optical depth measurements available through the AERONET website. Rayleigh and ozone optical depths have already been subtracted from this data. In addition to the optical depth data, this study also examines aerosol size distributions obtained using the optical depth values and almucantar measurements at 440, 670, 870, and 1020 nm.

2.6.2: Optical Depth Comparisons between the Sets

Similar to the case with the Schmid *et al.* (1999) study, the optical depth values between the three devices in this study generally agree within 0.02, the limit of calibration accuracy. However, the difference in optical depth between the MFRSR and the other two devices gets larger with increasing optical depth. This is shown in Figure 2.11, which plots the optical depth vs. day for each of the three devices at the channel closest to 870 nm. The red points in the graph correspond to common mornings or afternoons between all three data sets. As no gases absorb significantly at 870 nm, the optical depth comes almost entirely from aerosols. The RSS and CIMEL agree quite a bit more with each other than either does with the MFRSR, from having lower optical depth values in general to having strong peaks on two days, June 23 and July 5, 2000. Neither has any particularly high optical depth days between day 460 (the beginning of April 2000) and 500 (the middle of May), while the MFRSR has several. However, the data sets have a number of common elements as well. All show a seasonal cycle, with the largest optical depth values in summer and the lowest in winter. They also show a series of high optical depth days between days 270 and 300, or October 1999. In addition, the aerosol baseline appears to be elevated in the final two months of the data set, most likely a result of the forest fires raging in the Southwest during the spring of 2000. Figure 2.12 shows the mean optical depths, plotted vs. wavelength, for each of the three devices. The CIMEL has the sharpest curve of the three devices, suggesting that the effective radius will be the lowest in the CIMEL.

The differences in optical depth between the three devices vary with wavelength, a factor that will impact the retrieved values. Figure 2.13 shows the plot of morning and afternoon averaged optical depth values, for the five non-water vapor MFRSR channels, vs. the corresponding averaged values for the RSS channels nearest in wavelength. The least squares linear fit for the two quantities is given for each wavelength, along with the correlation coefficient. The plots show a number of significant features. First, each fit has a significant positive intercept, indicating that on low optical depth days, the optical depth values measured by the MFRSR tend to be significantly higher, relatively speaking, than the corresponding RSS values. In addition, the slopes are mostly greater than one, and in a few channels much greater. This means that the absolute differences in optical depth tend to increase with the optical depth itself. However, the slope for the 415 nm channel is less than one, meaning that as optical depth increases overall, the difference between the MFRSR and RSS optical depth values at that wavelength decreases while increasing at the other channels. This will have a serious impact on the retrieved aerosol size distributions, such that the effective radius values retrieved by the MFRSR can be expected to be much higher than those for the RSS, even before considering any errors caused by the inclusion of NO_2 absorption in the retrieval. Furthermore, the high slopes at 610 and 670 nm will probably cause ozone retrievals to be significantly higher for the MFRSR than for the RSS. Much of the systematic discrepancy is likely caused by a difference in the calibration methods, but some of it could possibly also be due to other differences in the devices. Some of the random errors, by contrast, probably result from the optical depth values for the two devices being averaged over different lengths of time.

This was unavoidable, because the RSS optical depth values were obtained as morning and afternoon means to begin with.

The plots of the optical depths for the CIMEL channels vs. the nearest RSS wavelengths are presented in Figure 2.14. As the CIMEL channels have already had the optical depth due to ozone subtracted from them, the same procedure is performed for the RSS channels as well, using the high-resolution ozone values which accompanied the data set. Again, the intercept values are all positive, so that for low optical depth days, the CIMEL and MFRSR devices are more likely to agree with each other than with the RSS. In addition, the intercept values are largest at the smaller wavelengths, which would suggest lower retrieved values for the effective radius, at least when the optical depth is low. Since the AERONET retrieval scheme does not subtract nitrogen dioxide from its measured optical depth, any NO_2 present will add to the optical depth at the lower wavelengths and cause a further reduction in retrieved particle size. For the CIMEL vs. RSS plots, the slopes all happen to be less than one, and far less than one for the smaller wavelengths. This confirms that for all the channels, the measured optical depth values for the CIMEL will diverge from those of the MFRSR, but converge with those of the RSS, as the optical depth increases. The highest optical depth days at the lower wavelengths in the CIMEL do not always correspond with high optical depth days for the RSS, though. As the two devices operate under different cloud screening procedures, it is quite possible that the CIMEL includes some points that are discarded as cloudy in the MFRSR and RSS, and vice versa.

2.7: Retrievals Assuming a Bimodal Distribution

To increase the flexibility of the RSS and CIMEL retrievals, the size distribution of the aerosols is permitted to be bimodal. One of the major objectives of this study is to determine if the assumption of a single-mode aerosol size distribution produces acceptable retrieval results, or if instead the assumption of a bimodal aerosol distribution is necessary. For the bimodal retrievals, three additional items of information are retrieved, as the aerosol optical depth, effective radius, and effective variance are now required for both coarse mode and fine mode components. The optical depth equation now includes a second aerosol term, corresponding to the coarse aerosol mode. The iterative process contains a nested loop. For a given fine mode effective radius and variance and coarse mode effective variance, the coarse mode effective radius is incremented from 0.5 μm to 5.0 μm in steps of 0.1 μm. The minimum residual is determined, and the process repeats for a new value of the fine mode effective radius. This is done for each combination of coarse and fine mode effective variance, and the minimum overall residual is considered the best fit. This process takes much longer than the single mode process, but never more than a few minutes for a given day of data. The bimodal retrieval uses the gas amounts retrieved separately from the aerosols using high spectral resolution. However, the limits in obtainable information suggested by the Box *et al*. (1996) study call into question the ability of even the RSS to adequately constrain the effective variance of either the coarse or the fine mode. Careful attention must therefore be paid to whether the retrieved effective variance values show any sort of

dependence on one of the other retrieved quantities, implying that the quantities cannot be retrieved independently of each other.

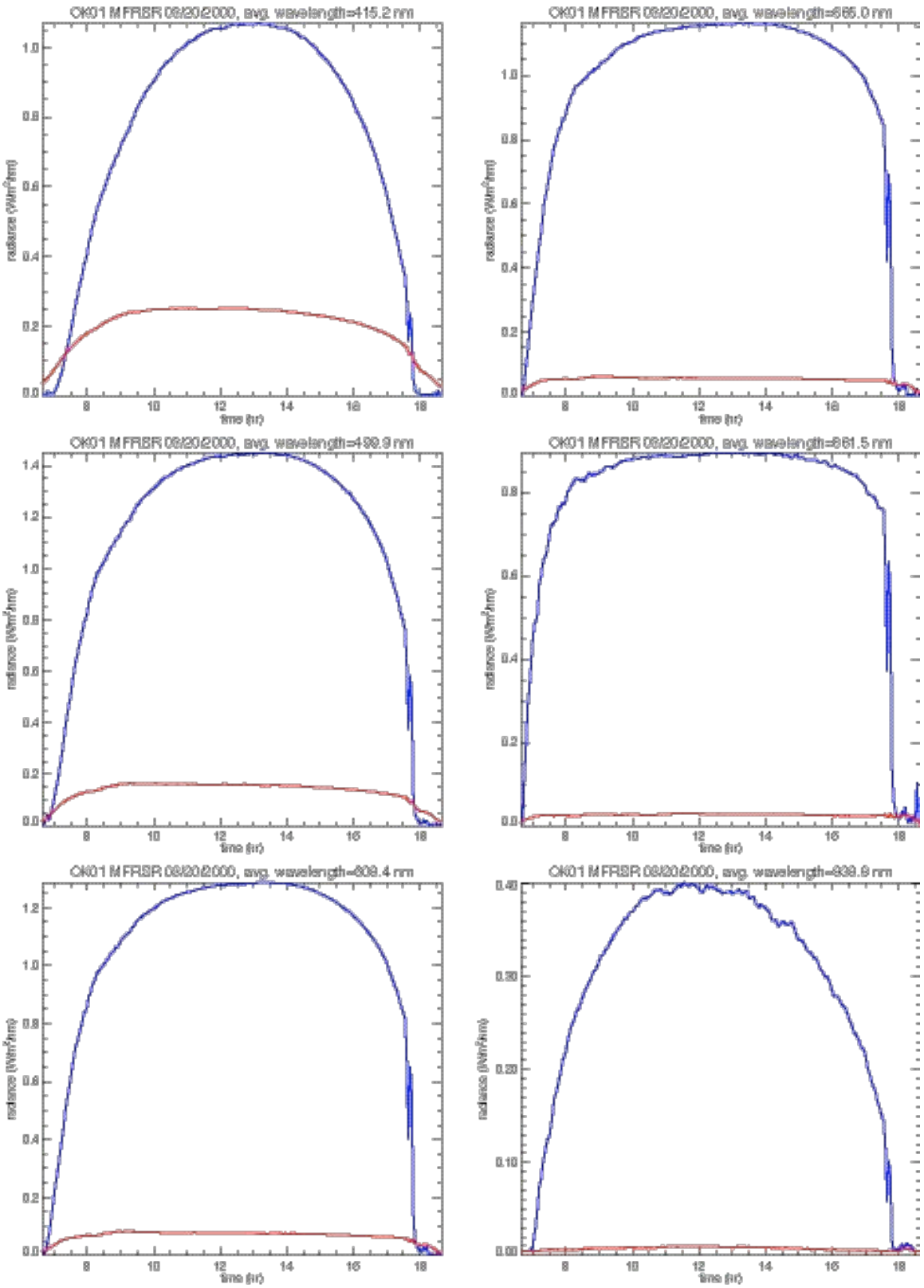

Figure 2.1: The direct (blue) and diffuse (red) intensities measured over the course of a day by the MFRSR located at the Southern Great Plains (SGP) site in Oklahoma.

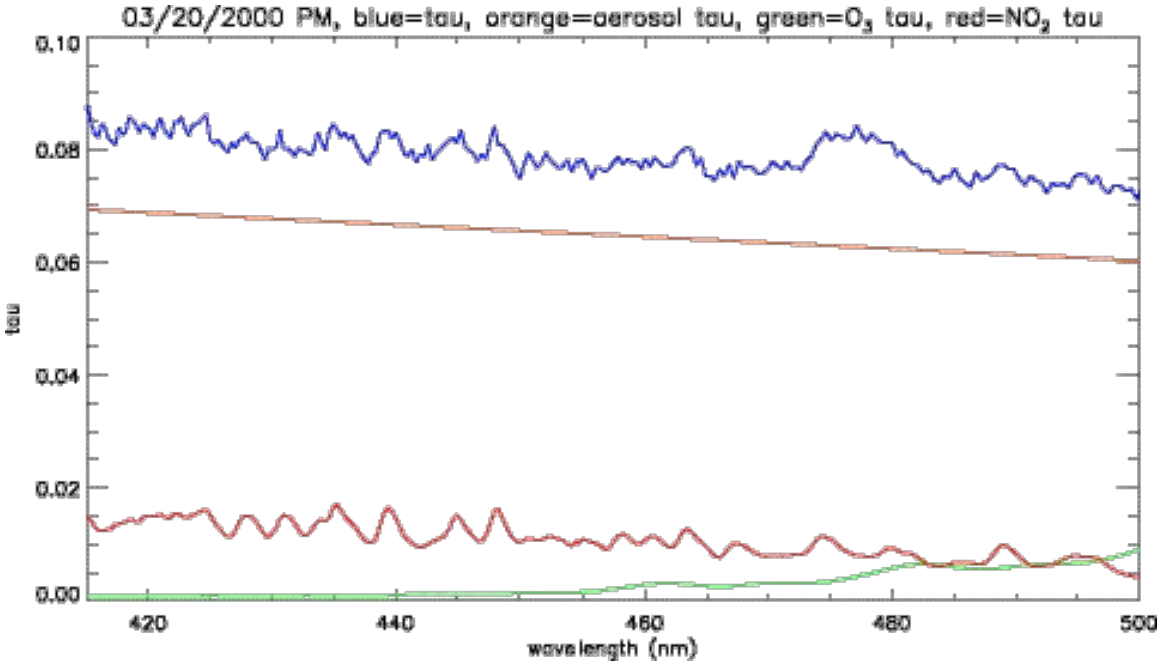

Figure 2.2: The optical depth between 415 nm and 500 nm measured by the RSS on the afternoon of March 20, 2000. Note that despite some noise, the details of the nitrogen dioxide spectrum are readily apparent.

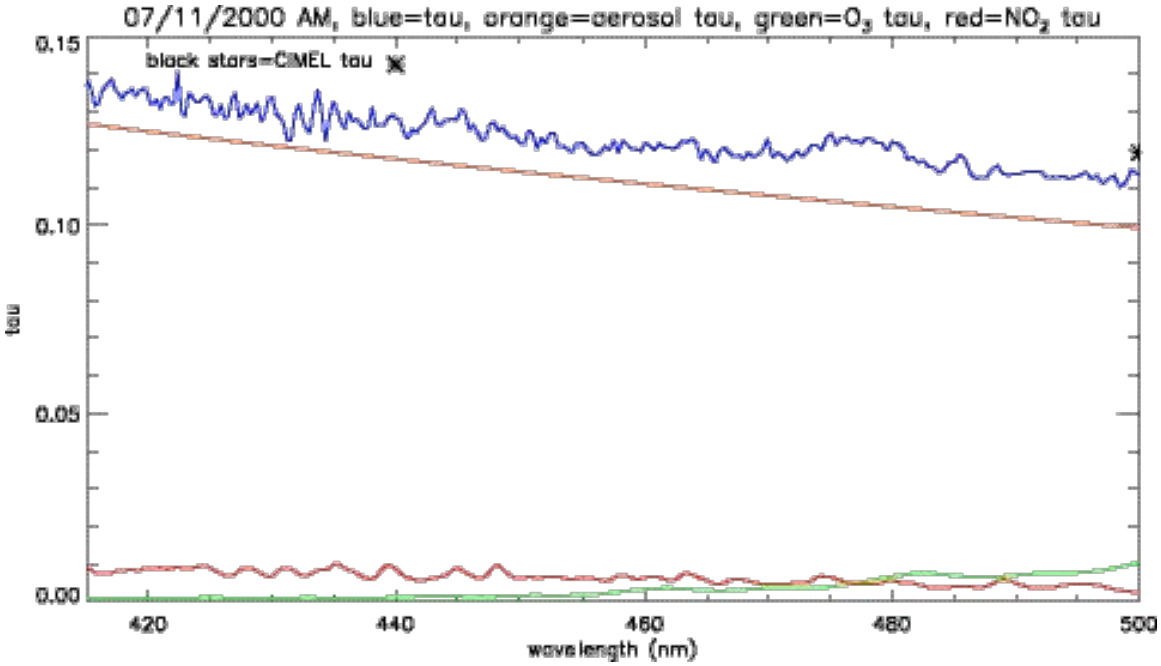

Figure 2.3: The optical depth between 400 nm and 500 nm measured by the RSS on the morning of July 11, 2000. Here, whatever NO_2 signal exists is drowned out by the noise. Note also a sharp pattern present between 430 and 435 nm; this is caused by the Ring effect.

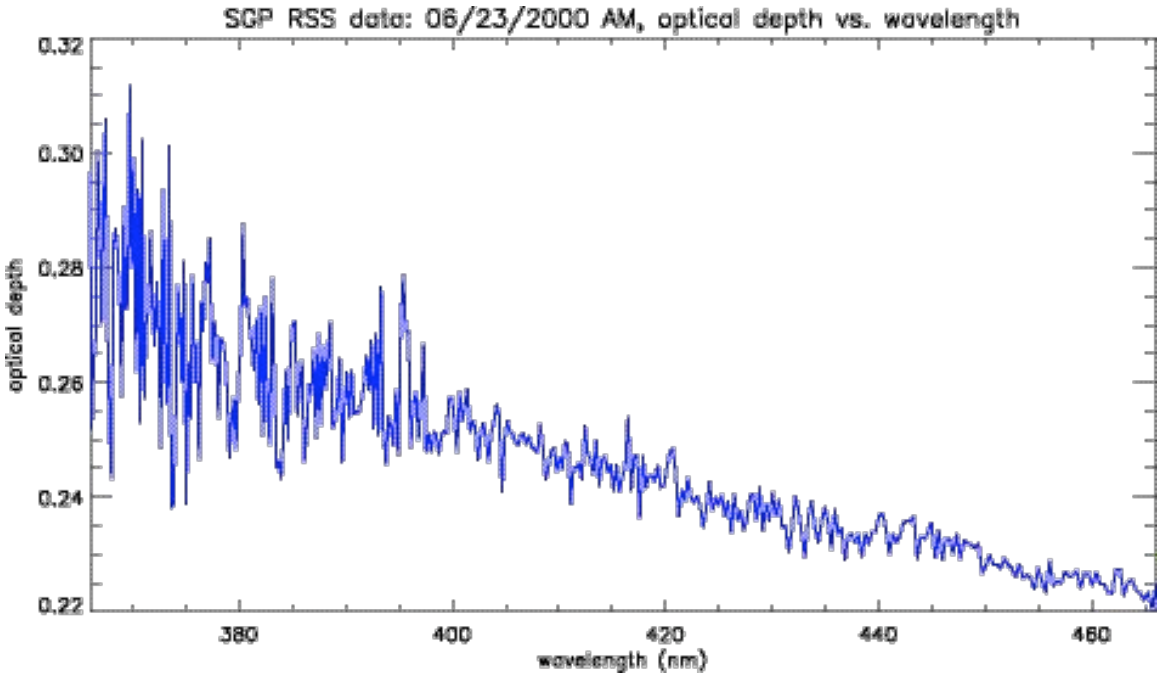

Figure 2.4: A plot of RSS-measured optical depth vs. wavelength between 366 and 466 nm, showing three distinct regions of noise, for the morning of June 23, 2000.

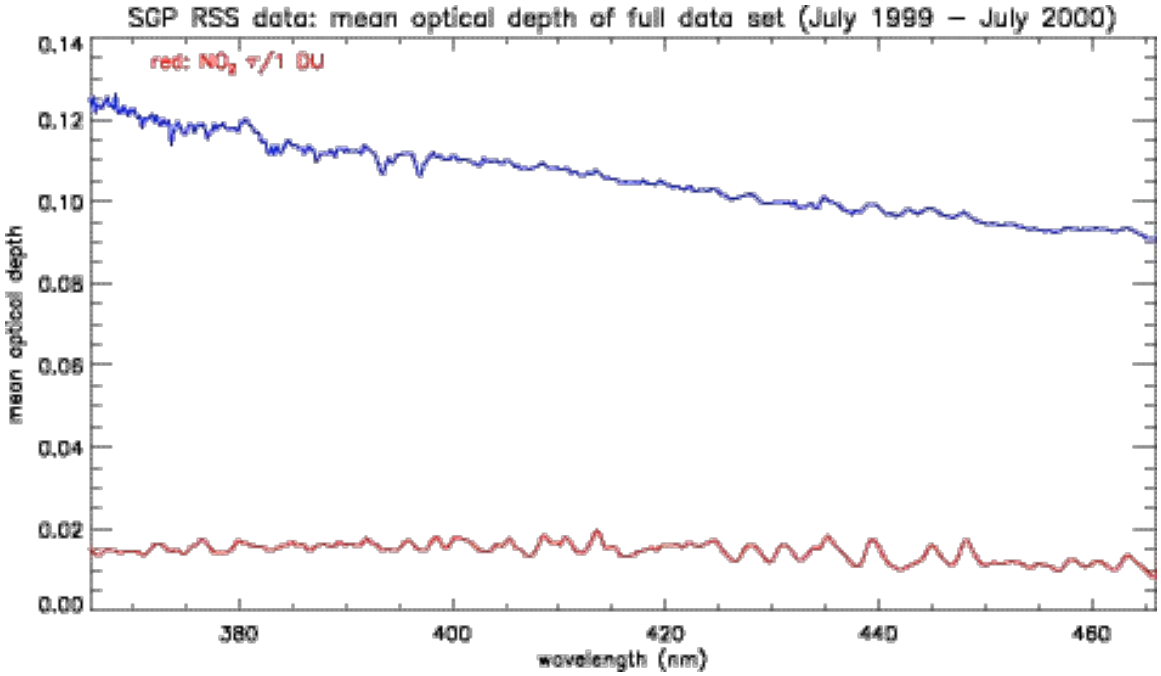

Figure 2.5: The mean optical depth for the full RSS data set between 366 and 466 nm. The NO_2 signal shows up far more clearly for the mean than on most individual days, indicating that despite the presence of the Ring effect, most of the noise in the data can be treated as random.

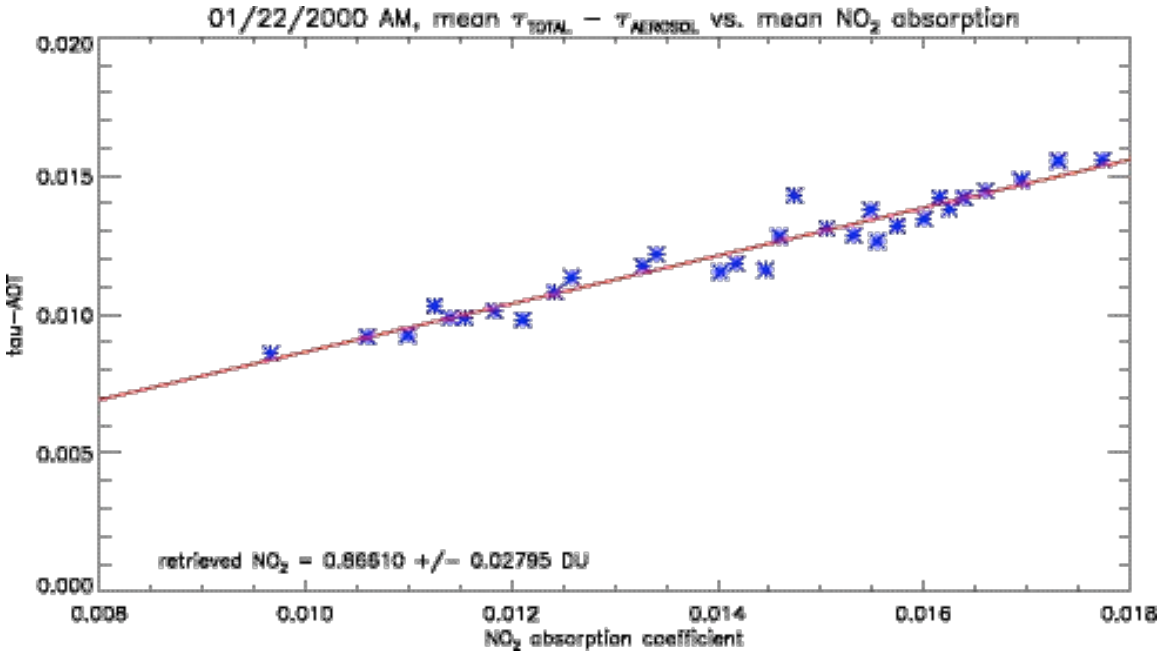

Figure 2.6: The result of the first least squares regression in the NO_2 retrieval algorithm for the morning of January 22, 2000, after the retrieved aerosol extinction values have been subtracted out.

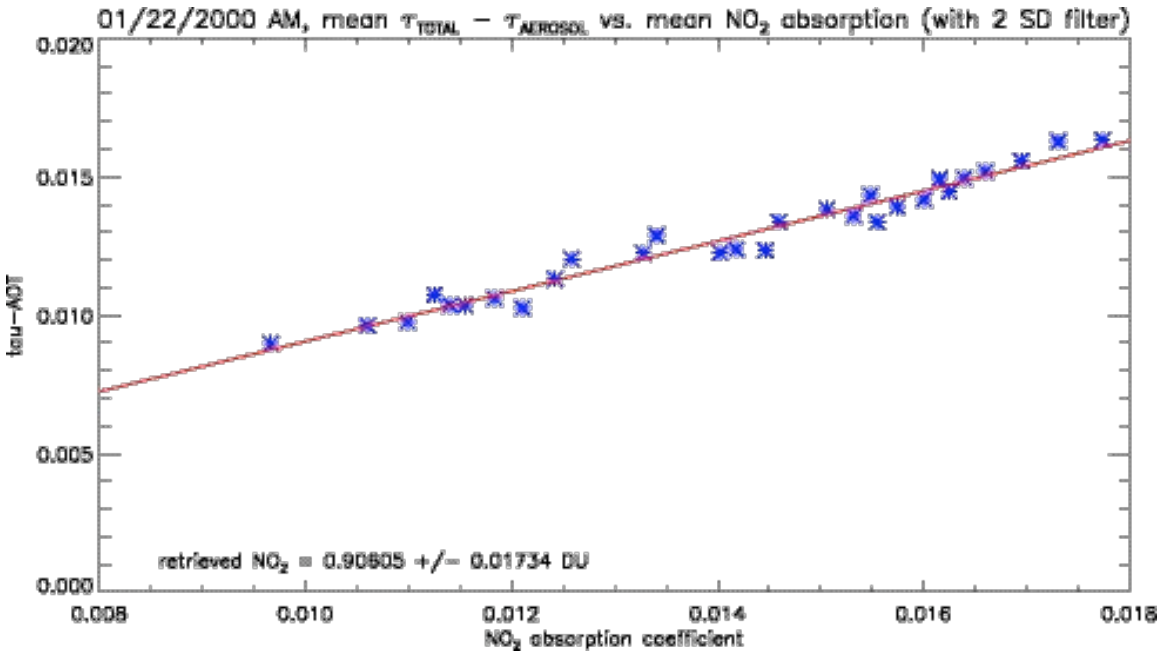

Figure 2.7: The final result, after points beyond two standard deviations of the line in Figure 2.6 have been subtracted out and the process repeated.

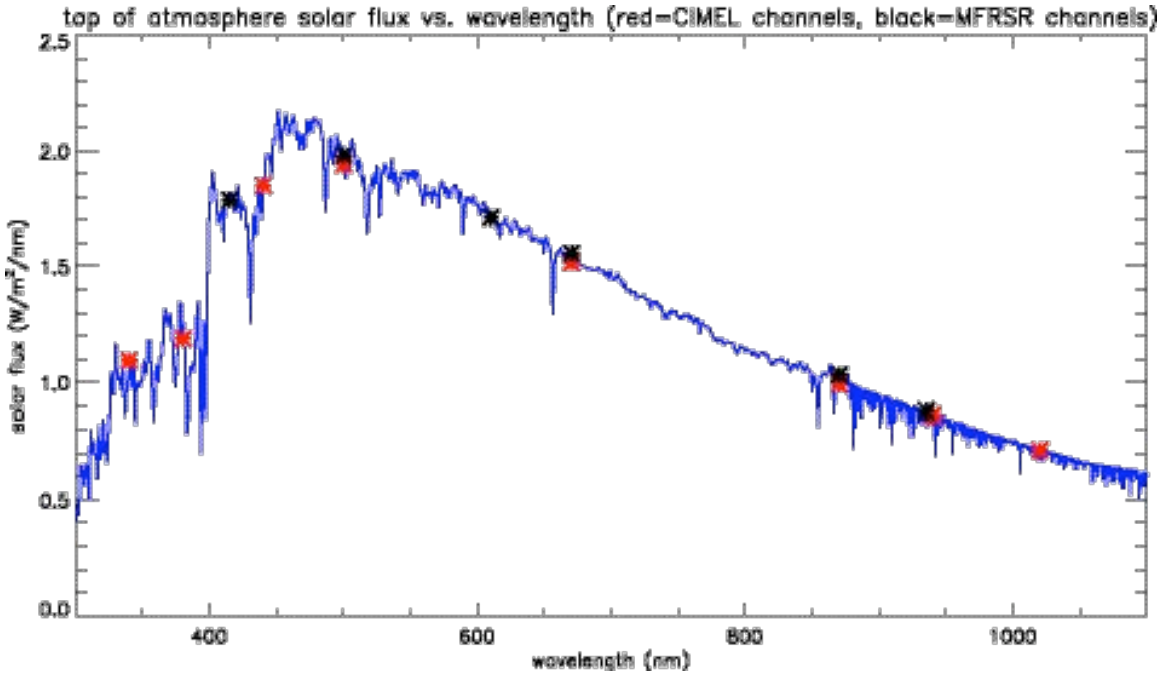

Figure 2.8: The plot of top-of-atmosphere solar flux vs. wavelength. The locations of the CIMEL filters are denoted by red asterisks, while those of the MFRSR are denoted by black asterisks. The combination of low solar flux and high Rayleigh optical depth requires greater sensitivity from the ultraviolet filters in the CIMEL than from the visible ones.

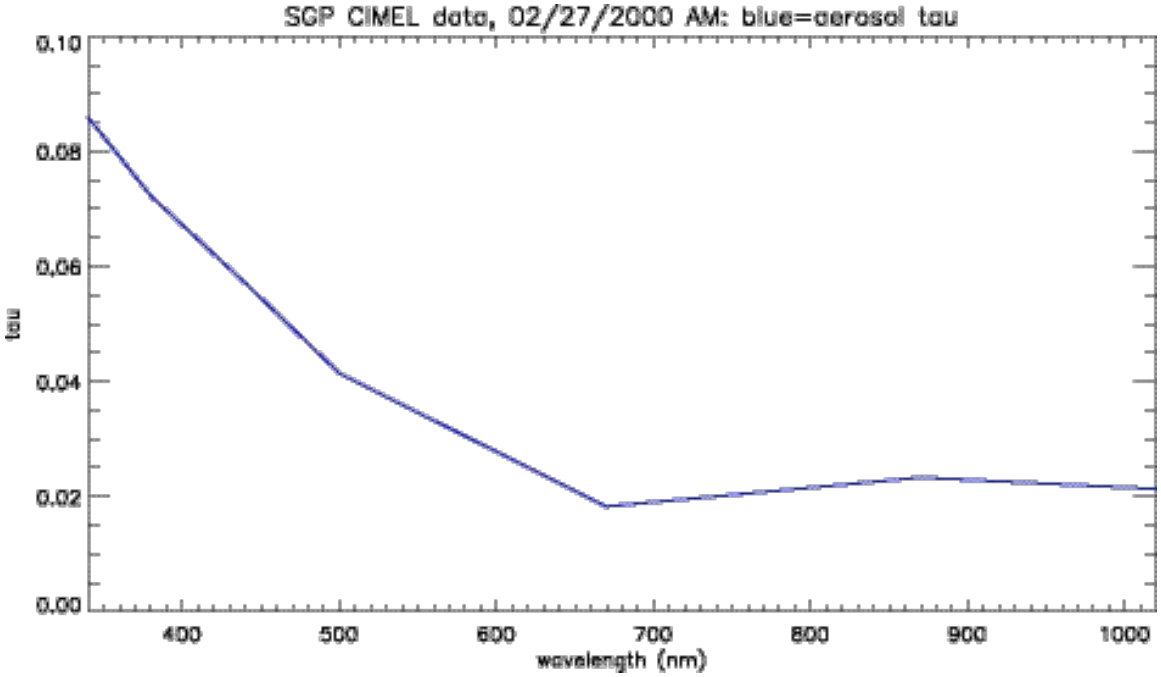

Figure 2.9: The averaged aerosol optical depth values measured by the CIMEL on the morning of February 27, 2000. This day shows an unnatural dip in the optical depth at 670 nm.

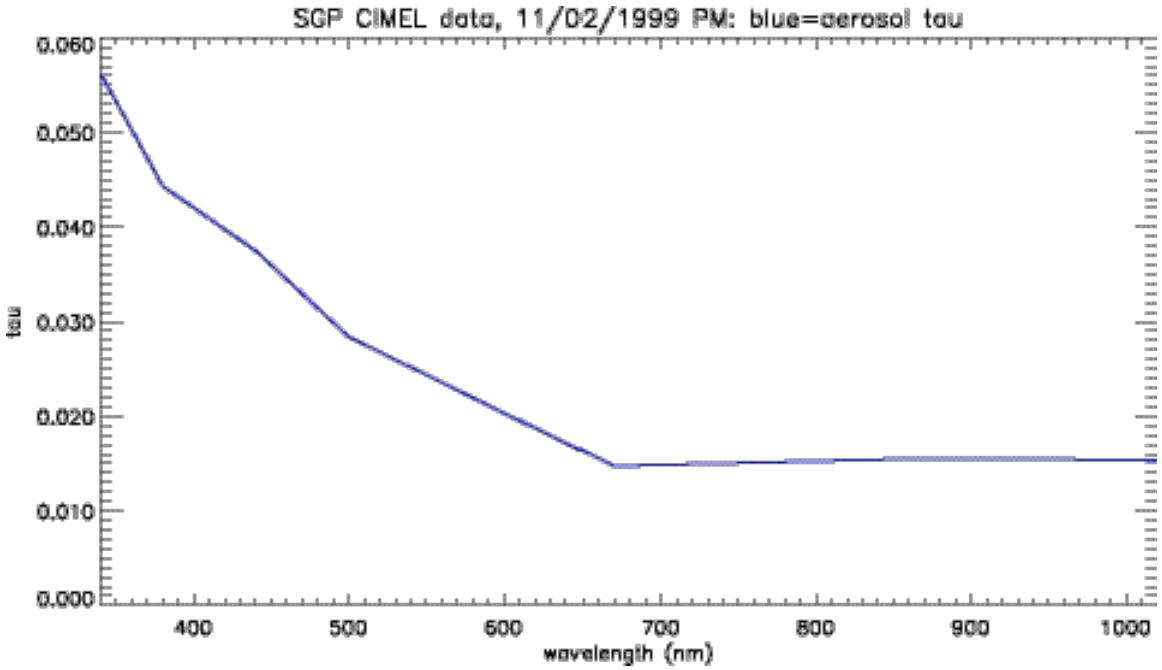

Figure 2.10: The averaged aerosol optical depth values for the CIMEL on the afternoon of November 2, 1999. The slight bulge in the optical depth at 440 nm might be due to absorption by nitrogen dioxide, or it might be the result of the same defect that causes the dip at 670 nm.

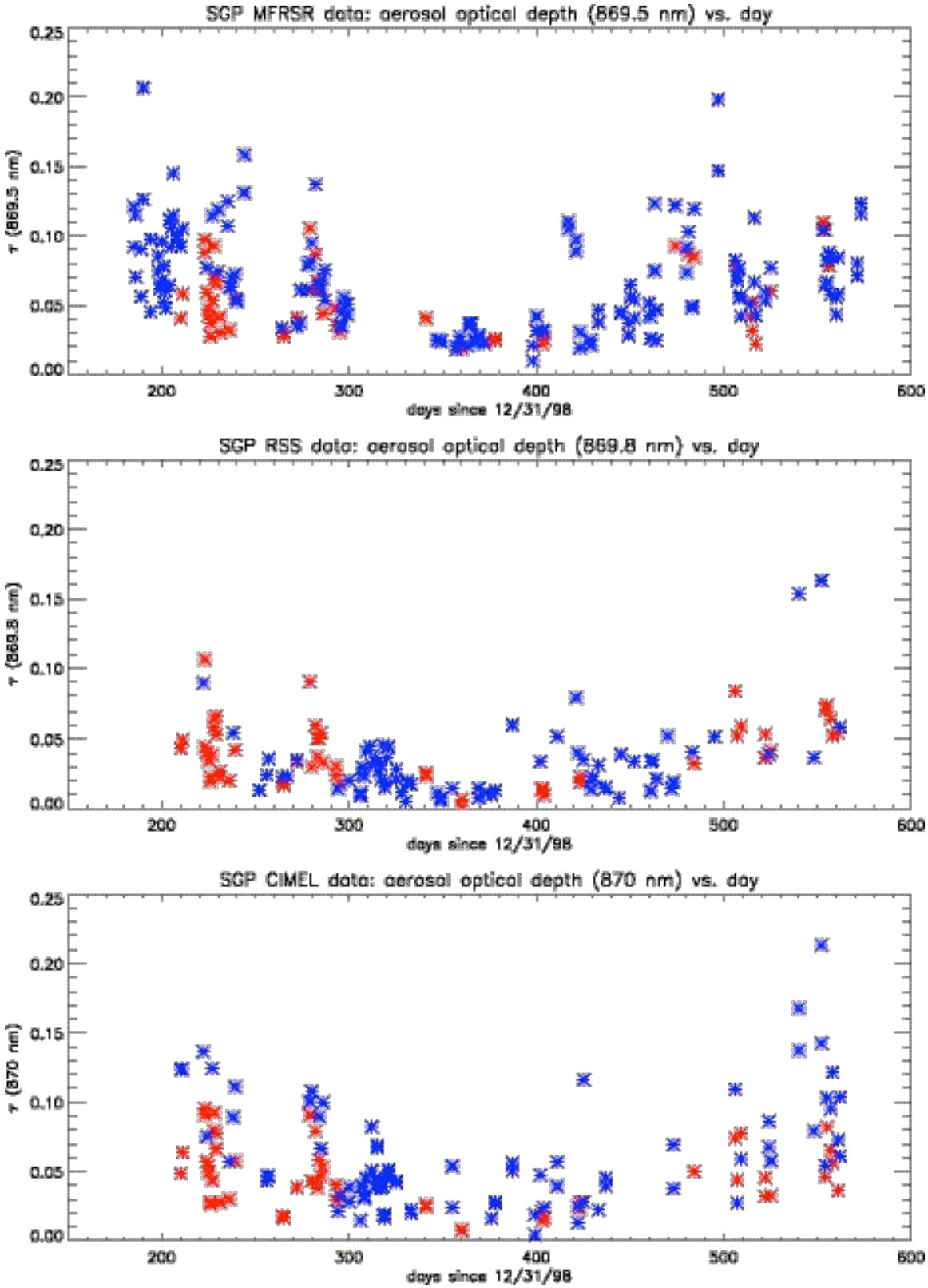

Figure 2.11: The plots of optical depth vs. day, for all three devices, at the wavelength closest to 870 nm, where gas absorption is negligible. The red points signify mornings or afternoons for which all three devices produced data.

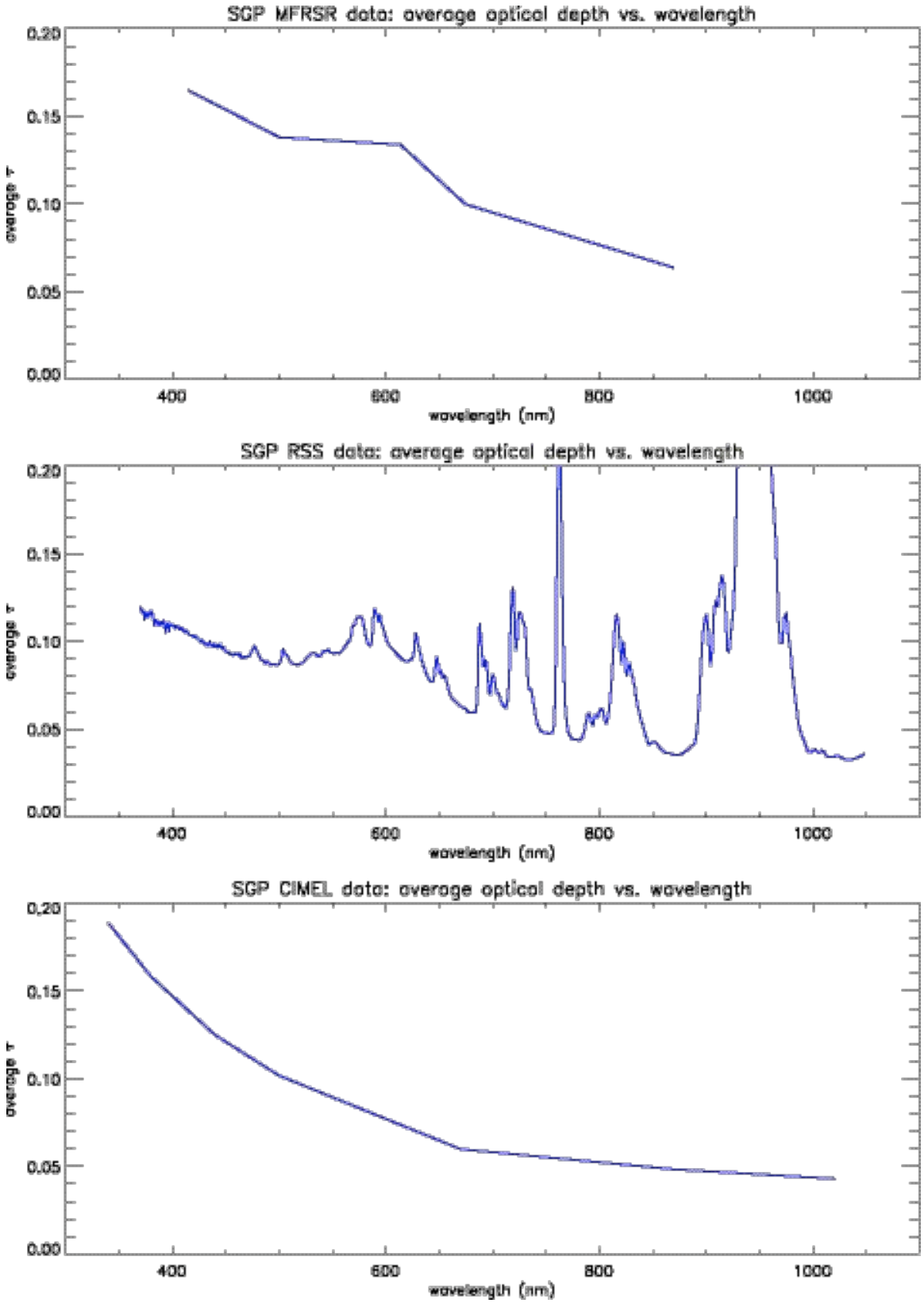

Figure 2.12: The plots vs. wavelength of the average optical depths in the data set at each channel for each of the three devices.

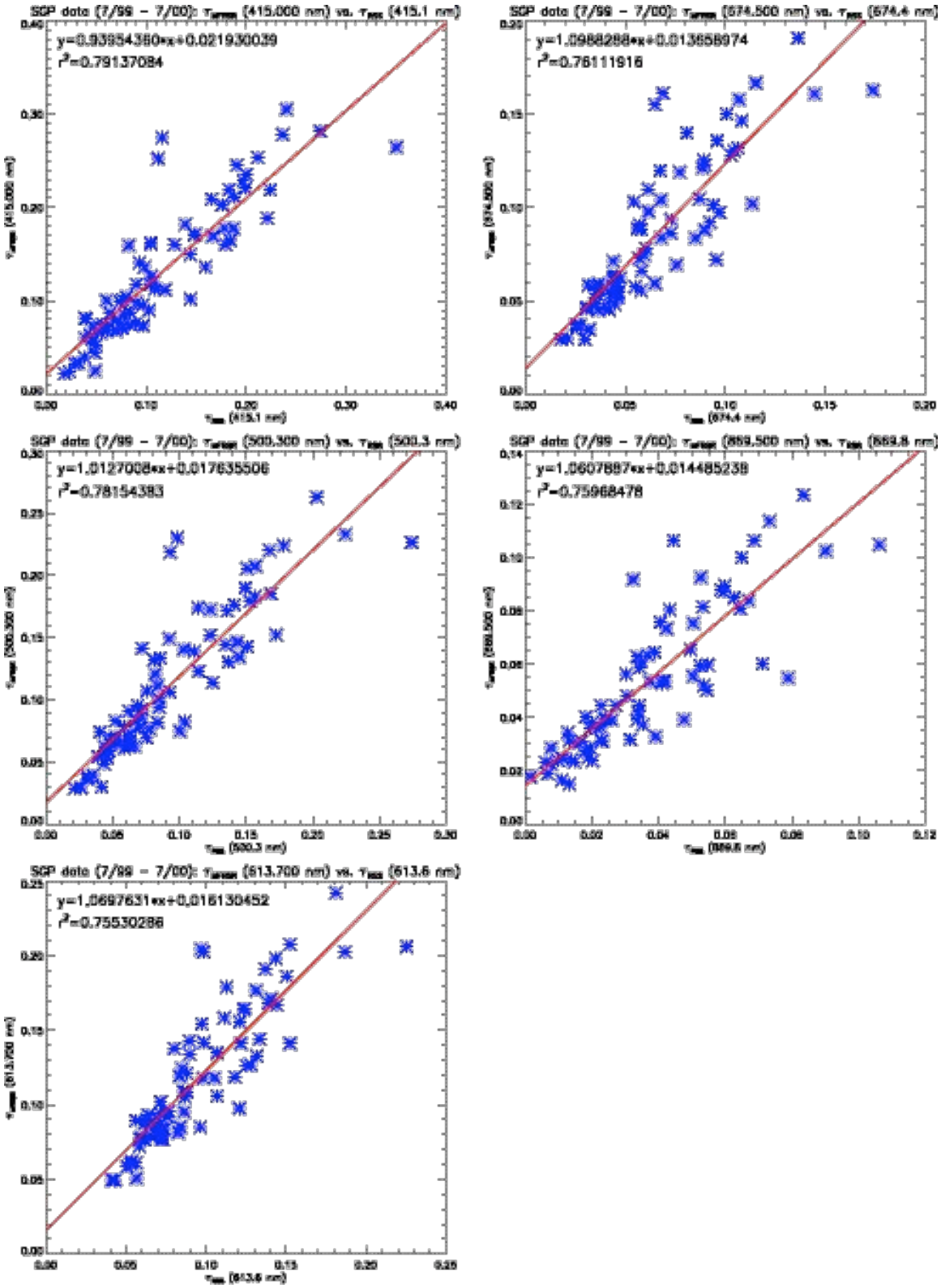

Figure 2.13: Comparative plots of the mean morning or afternoon optical depths at the five MFRSR channels vs. those of the closest RSS channels.

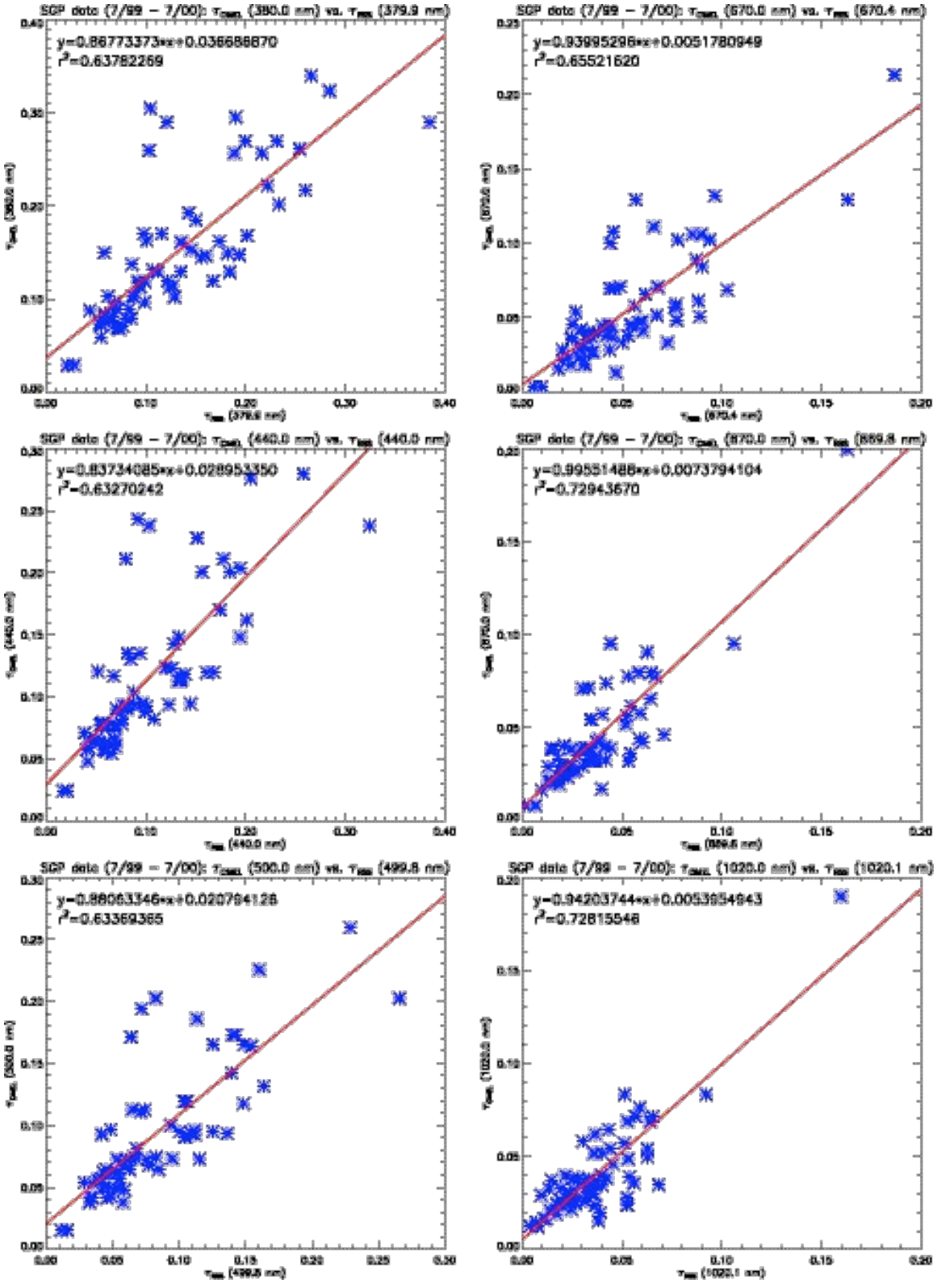

Figure 2.14: The same as Figure 2.13, except for the CIMEL instead of the MFRSR.

Chapter 2 References

Alexandrov, M. D., Lacis, A. A., Carlson, B. E., and Cairns, B., *Journal of the Atmospheric Sciences* **59**, 524-543 (2002a).

Alexandrov, M. D., Lacis, A. A., Carlson, B. E., and Cairns, B., *Journal of the Atmospheric Sciences* **59**, 544-566 (2002b).

Ångstrom, *Physical Review* **1**, 365-371 (1894).

Bigelow, D. S., Slusser, J. R., Beaubien, A. F., and Gibson, J. H., *Bulletin of the American Meteorological Society* **79**, 601-615 (1998).

Dubovik, O., and King, M. D., *Journal of Geophysical Research* **105**, 20673-20696 (2000).

Flowers, E. C., McCormick, R. A., and Kurfis, K. R., *Journal of Applied Meteorology* **8**, 955-962 (1969).

Guzzi, R., Maracci, G. C., Rizzi, R., and Siccardi, A., *Applied Optics* **24**, 2859-2864 (1985).

Harrison, L., Michalsky, J., and Berndt, J., *Applied Optics* **33**, 5118-5125 (1994).

Harrison, L., Beauharnois, M., Berndt, J., Kiedron, P., Michalsky, J., and Min, Q., *Geophysical Research Letters* **26**, 1715-1718 (1999).

Holben, B. N., Eck, T. F., Slutsker, I., Tanré, D., Buis, J. P., Setzer, A., Vermote, E., Reagan, J. A., Kaufman, Y. J., Nakajima, T., Lavenu, F., Jankowiak, I., and Smirnov, A., *Remote Sensing of Environment* **66**, 1-16 (1998).

Holben, B. N., Tanré, D., Smirnov, A., Eck, T. E., Slutsker, I., Abuhassan, N., Newcomb, W. W., Schafer, J. S., Chatenet, B., Lavenu, F., Kaufman, Y. J., Vande Castle, J.,

Setzer, A., Markham, B., Clark, D., Frouin, R., Halthore, R., Karneli, A., O'Neill, N. T., Pietras, C., Pinker, R. T., Voss, K., and Zibordi, G., *Journal of Geophysical Research* **106**, 12067-12097 (2001).

Kiedron, P., Michalsky, J., Schmid, B., Slater, D., Berndt, J., Harrison, L., Racette, P., Westwater, E., and Han, Y., *Journal of Geophysical Research* **106**, 24007-24016 (2001).

Lacis, A. A., and Mishchenko, M. I., *Aerosol Forcing of Climate*, ed. R. J. Charlson and J. Heintzenberg, J. Wiley & Sons, pp. 11-42 (1995).

Leontieva, E., and Stamnes, K., *Journal of Applied Meteorology* **35**, 2011-2022 (1996).

London, J., Bojkov, R. D., Oltmans, S., and Kelly, J. I., *Atlas of the Global Distribution of Total Ozone July 1957-1967*, NCAR Technical Note 133+STR, National Center for Atmospheric Research, Boulder, CO (1976).

Michalsky, J. J., Liljegren, J. C., and Harrison, L. C., *Journal of Geophysical Research* **100**, 25995-26003 (1995).

Penndorf, R., *Journal of the Optical Society of America* **47**, 176-182 (1957).

Platt, U., Perner, D., and Patz, H. W., *Journal of Geophysical Research* **84**, 6329-6335 (1979).

Remer, L. A., and Kaufman, Y. J., *Journal of Geophysical Research* **103**, 13859-13871 (1998).

Schmid, B., Michalsky, J., Halthore, R., Beauharnois, M., Harrison, L., Livingston, J., Russell, P., Holben, B., Eck, T., and Smirnov, A., *Geophysical Research Letters* **26**, 2725-2728 (1999).

Shaw, G. E., Reagan, J. A., and Herman, B. M., *Journal of Applied Meteorology* **12**, 374-380 (1973).

Vigroux, E., *Annales Physicae* **8**, 709 (1953).

Volz, F. E., *Pure and Applied Geophysics* **36**, 138-147 (1957).

Volz, F. E., *Tellus* **21**, 625-629 (1969).

Wesely, M. L., *Journal of Applied Meteorology* **21**, 373-383 (1982).

Chapter 3: The EOF Analysis of the Optical Depth Data

3.1: Exploration of the Information Content in the Data

Box *et al*. (1996) investigated the information content and wavelength selection for multi-spectral radiometers by performing an EOF analysis on synthetic optical depth data for twelve different wavelengths. These wavelengths corresponded to the potential locations of sun photometer filters, and spanned the range between 368 and 2230 nm. In their theoretical analysis, they found that the interdependence of the measurements at different wavelengths allows measurements at other wavelengths to be predicted. They conclude by recommending a seven-channel instrument with filters located at 368, 500, 675, 862, 1030, 1055, and either 1725 or 2230 nm. Box *et al*. defined the noise level above which an eigenvector yielded significant information as the square root of the variance that eigenvector contributes relative to the first eigenvector. At a noise level of 10%, corresponding to an eigenvector contributing 1% as much of the total variance as the first eigenvector does, the filter combination containing 1725 nm yields three pieces of information, while the set that includes 2230 nm will give four pieces. As the noise level is reduced, the amount of obtainable information increases for both filter combinations, meaning that noise inherent in the data can mask information that would otherwise have been obtainable. This can be seen in the RSS data, where a considerable amount of random noise in the short wavelengths leads to a large margin of error in the NO_2 retrievals, and consequently limits the accuracy of the retrieved aerosol size distribution as well. Indeed, the presence of gaseous absorption in general can add to the

"noise" level in the retrieval of specific aerosol properties when the assumed or measured values of gas amounts turn out to be inaccurate.

The wavelength combination in the Box *et al*. analysis that extends only to 1030 nm, like the RSS and the CIMEL essentially do, just barely yields a third item of information at the 10% noise level, but the third eigenvector for this combination accounts for well less than half the amount of variance that the third eigenvector in the combination extending to 1725 nm accounts for. Furthermore, the addition of extra channels within the established wavelength range, beyond the minimum amount necessary to yield a certain number of items of information, only slightly enhances the contribution of the eigenvectors lower than the first to the total variance. This implies that for a given wavelength range, the additional information obtainable from the data barely increases with the number of channels.

In practice, the question of whether or not the RSS can outperform the other devices in terms of retrieving aerosol information, with its higher resolution and greater number of channels, is testable on actual data. If it turns out that the aerosol information retrievable is indeed limited, this will have significant implications for ground-based aerosol monitoring. The size distribution of a given aerosol mode can be described with two quantities, the effective radius and effective variance. The optical depth at a given wavelength of this mode would be a third item of information. If the aerosols being observed turn out to be unimodal, then the limits imposed by the analysis of Box *et al*. would not be problematic. If the aerosol extinction in the data cannot be adequately described by a single-mode size distribution, on the other hand, the question then becomes what aerosol properties in a bimodal distribution can be retrieved, given an

incomplete amount of information. As the coefficients of projection of a given eigenvector depend on wavelength, it is possible that a set of EOF's will explicitly reveal the aerosol information inherent in the data sets.

In this study, the results for a given EOF are presented in two graphs. The top graph shows the particular eigenvector plotted vs. day, or time. The bottom graph shows the coefficients of projection of the eigenvector plotted vs. wavelength. For some of the plots of the EOF's of RSS data, an additional plot is added, showing an isolated portion of the plot of the coefficients of projection vs. wavelength in closer detail. This facilitates the identification of the presence of NO_2 absorption or the Ring effect in a given eigenvector.

3.2: The EOF's of the Base Data Sets

3.2.1: EOF's of the MFRSR Data

The study begins with a principal component analysis of the MFRSR optical depth data. Excluding the water vapor channel, the MFRSR has only five channels available for the analysis of aerosols, ozone, NO_2. This limits the number of EOF's produced to five, but only the first two eigenvectors contribute more than 1% to the total variance. Figure 3.1 shows that the coefficients of the first EOF of the full MFRSR set of optical depth data qualitatively represent the mean optical depth for the full data set; the pattern in the coefficients very closely resembles the graph of the mean optical depth vs. wavelength, previously shown in Figure 2.12. The eigenvector itself, when plotted vs.

day, shows the established features of the seasonal dependence of the aerosol optical depth, including the annual cycle and the relative peaks in October 1999 seen in all the data sets (see Figure 2.11). Ozone absorption clearly shows up above the aerosol baseline curve in the coefficients, but the signal due to the absorption by NO_2 is a lot less obvious. The coefficients of the second EOF, shown in Figure 3.2, are in fact dominated by ozone absorption. This second EOF accounts for 3% of the total variance, indicating that while this eigenvector contributes much less to the total variance than the first, it yields a second item of significant information. The pattern of the eigenvector is largely an inversion of the first. Two conclusions can be drawn from this. First, if the first eigenvector were to be used as an approximation of the actual optical depth, ozone absorption would be overestimated on the high optical depth days and underestimated on the low days. Second, and more interestingly, if the eigenvectors are combined in such a way that the aerosol dependence gets canceled out, the resulting plot vs. day would primarily represent the variation in ozone amounts. There is also a large baseline offset in the coefficients of the second EOF, though, which increases slightly with wavelength. This curve could be providing some information about aerosol extinction, suggesting that changes in aerosol size distribution are also discernible in the second EOF. Unfortunately, these two EOF's indicate that the extinctions due to ozone and aerosols are so intertwined in the data that a completely accurate separation of the two may prove to be very difficult. The third eigenvector accounts for less than 0.2% of the total variance, still significant for only a 4% noise level or lower. Given the increasingly evident difficulty in producing accurate gas measurements, and the resulting error in

aerosol properties, it is nearly certain that the noise exceeds 4%, confirming that only two pieces of aerosol information are provided by in the MFRSR data.

3.2.2: EOF's of the RSS Data

The RSS has a broader wavelength range than the MFRSR, along with much higher spectral resolution. Given these two facts, an EOF analysis of the RSS optical depth data will produce different results than what was observed in the MFRSR data. Hopefully, some important insights can be gleaned from how, and where, these two sets of EOF's differ. Figure 3.3 shows the first eigenvector, along with its corresponding coefficients of projections, for the EOF of the full RSS data set. The eigenvector itself shows essentially the same pattern as the optical depth vs. day plot in Figure 2.11, although the peaks at June 23 (day 540) and July 5 (day 552) are curiously absent. The prominent absorption features of water vapor, ozone, and oxygen are all present in the coefficients of projection. The NO_2 absorption features are present as well, although this is better illustrated in the more detailed graph on the bottom, which focuses on the wavelengths between 390 and 450 nm. A couple of additional features emerge within this more limited spectral interval, including the Ring effect between 393 and 397 nm and 431 and 435 nm, and the weak water vapor absorption band at 443 nm.

A sharp drop-off at low wavelengths, actually starting at 600 nm, emerges in the second EOF, shown in Figure 3.4. The second EOF contributes 10% to the total variance, quite a bit more than the corresponding EOF for the MFRSR data. Significant water vapor and oxygen absorption are present in the coefficients as well. While the

nearly diagonal decrease does not resemble the spectral dependence associated with any specific aerosol size distribution, it is clear from the large magnitude of the coefficients that changes in air pressure on a day-to-day basis, with their corresponding effect on the optical depth due to Rayleigh scattering, could not account for this dramatic spectral dependence. Changes in aerosol extinction, however, can account for this. The negative values for the baseline curve mean that on days when the EOF value is positive, the ratio of the aerosol extinction in the blue and violet regions of the spectrum to the extinction at larger wavelengths tends to be low. Conversely, the days where the EOF value is negative have a higher ratio. As this ratio depends on the size of the aerosols observed, the result suggests that the second EOF provides some information on the size distribution. Positive EOF values correspond to larger particles, while negative values correspond to smaller particles. The water vapor and oxygen absorption bands are both present as well, so this EOF suggests that particle size varies on the same time scale as humidity and pressure. Like the second EOF of the MFRSR data, the pattern in the second EOF resembles the inverse of the aerosol optical depth. The two conspicuous exceptions, once again, are June 23 and July 5, the days in which the aerosol optical depth at 870 nm was highest. These two days do not have the highest EOF values, but they do not have low values like the other high aerosol days do, either. This suggests that the aerosol behavior on these two days somehow differs from the general behavior. The first EOF, taken by itself, would underestimate the aerosol extinction at the short wavelengths on all the other high aerosol days and overestimate the extinction at short wavelengths on the low aerosol days.

The third EOF accounts for 1.6% of the total variance, making it more significant than the third MFRSR eigenvector, and indicating that an additional piece of aerosol information is more clearly present in the RSS data. However, the RSS data includes a couple of gas absorbers, namely water vapor and oxygen, not present in the five MFRSR channels used in the EOF. For a more direct comparison with the MFRSR EOF analysis, then, the EOF's were recalculated, omitting spectral intervals that contain significant H_2O and O_2 absorption, and subtracting out the O_2-O_2 absorption. Figure 3.5 shows the first eigenvector, which again resembles the mean of the data set. The bottom panel of the figure, which focuses on the more limited spectral range between 390 and 450 nm, clearly shows the contributions of both nitrogen dioxide and the Ring effect to the total spectrum. This confirms that the absorption of NO_2 is actually present in the first MFRSR eigenvector as well, although the limited spectral resolution of the MFRSR makes the association much more tenuous. The curve of the eigenvector very closely mirrors the one seen in Figure 3.3; again, June 23 and July 5 do not emerge as peak values the way they do for the aerosol optical depth. This eigenvector contributes less to the total variance than the first eigenvector of the MFRSR data, but more than the first eigenvector of the RSS data with water vapor and oxygen bands included. The difference with the MFRSR data implies that the greater spectral detail inherent in the RSS data yields additional information beyond what can be seen in the coarser-resolution MFRSR data alone. On the other hand, the difference resulting from the exclusion of the water vapor and oxygen bands probably results from aerosol extinction correlating less with column amounts of ozone than with column amounts of water vapor.

The second EOF, shown in Figure 3.6, contains both ozone and NO_2 in sharp detail, along with the Ring effect. The presence of a relatively flat baseline curve in the coefficients of the projections, with a significant ozone contribution superposed on it, implies that aerosols contribute here in a similar manner to the second EOF of the MFRSR data set, although here the baseline curve slightly decreases with wavelength. Again, the pattern in the eigenvector itself inverts the pattern of the aerosol optical depth. As in Figure 3.4, the values of the eigenvector components corresponding to June 23 and July 5 are not conspicuously higher than the baseline curve, but not as low as the values on most of the other high aerosol days. The third eigenvector contributes only 0.3% to the total variance. Predictably, from examining the first eigenvector, this value is larger than the variance associated with the corresponding EOF for the MFRSR analysis, but smaller than the variance of the third eigenvector when water vapor is included. In fact, as Figure 3.7 shows for several different EOF's using RSS data, the third eigenvector rarely accounts for much more than 1% of the total variance, and usually contributes less. This suggests that even with the added resolution of the RSS, the limited wavelength range of the device allows for only two items of aerosol information to be retrieved unambiguously, with a third item barely retrievable in only a few cases. It is not clear, though, why the inclusion of water vapor spectral bands increases the significance of the third eigenvector. This dependence did not have anything to do with the spacing of channels in the RSS. The channels are much more narrowly spaced at the lower wavelengths, but additional EOF's produced using evenly spaced wavelengths did not deviate significantly from those calculated using the full data. Nor are the basic features of the eigenvectors changed by removing the strongest absorption lines of water vapor;

the inclusion of even weak water vapor absorption lines produces a stronger third eigenvector than the inclusion of no absorption lines. It is possible that the presence of gas absorption in the data brings out some aerosol properties in the EOF analysis that may otherwise remain hidden. With or without water vapor, the result suggests that the RSS can do at least a slightly better job at obtaining one additional piece of information about the aerosols than the MFRSR. The fourth eigenvector deals very specifically with the Ring effect, and does not appear to contain any spectral information pertinent to aerosols or any of the gaseous absorbers. Subsequent eigenvectors account for the random noise that affects the short wavelengths on most of the days in the RSS data set.

3.2.3: EOF's of the CIMEL Data

Without the presence of any gas absorbers in the CIMEL data, the variance in the EOF's is almost completely dominated by the first eigenvector, presented in Figure 3.8. The second eigenvector contributes 10% in the RSS data where water vapor is present, 4% in the RSS data where ozone is the primary absorber, 3% in the MFRSR where ozone again is the primary absorber, but less than 1% in the CIMEL data where no absorber is significantly present. It is not clear if the gaseous absorption alone creates the difference in variance, or if ozone and water vapor correlate with aerosols in a way that causes the contribution of the aerosols to be likewise separated into multiple eigenvectors. Otherwise, the first EOF of the CIMEL data does show some similarities to the other first eigenvectors, and generally follows the pattern of the optical depth values presented in Figure 2.11. Once again, the high optical depth values on June 23 and July 5, which were

observed by the CIMEL as well as the RSS, are not reflected in the first eigenvector. The second eigenvector, by contrast, makes a negligible contribution to the total variance, and the pattern of the coefficients of projection does not seem to show anything consistent with changes in aerosol properties. In addition, the pattern in the eigenvector, while correlating well with aerosol optical depth, is not inverted the way the same eigenvector is in the EOF's of the MFRSR and RSS data.

3.3: EOF's of the RSS and MFRSR with the Means Subtracted

3.3.1: The Benefits of Subtracting the Means

The determination of Box *et al.* (1996) that two or three items of aerosol information are obtainable from data in the wavelength range of these three devices appears to be borne out by the EOF analysis, but only when wavelengths sensitive to gas absorption are included. Furthermore, if there really are only three items of information obtainable from the data, it is not clear from the preliminary analysis what the additional pieces of aerosol information, beyond optical depth, are. Constraints on the amount of obtainable aerosol information would not significantly impact the retrieval if a single-mode aerosol size distribution can be safely assumed, but would require assumptions to be made about the specific details of a bimodal size distribution. The EOF's of the RSS data, when the wavelengths sensitive to the absorption of water vapor are included, only hint at information about the aerosol sizes, and do not provide a reason to assume either a single-mode or a bimodal distribution. For the sake of a more uniform comparison, the

EOF's of the MFRSR data could be calculated with the water vapor channel included. However, a clearer picture of the aerosol properties would more likely emerge by also recalculating the EOF's of the MFRSR and RSS data sets after the mean optical depth value at each wavelength has been subtracted. Water vapor varies to a strong enough degree that its absorption signal will remain in these EOF's, but the signals of ozone and nitrogen dioxide will essentially vanish, and any spectral pattern observed in the plot of the coefficients of projection vs. wavelength over the regions not sensitive to water vapor absorption will result from aerosols only. In addition, the significant difference in measured optical depth values between the MFRSR and RSS will vanish with the subtraction of the mean values, allowing for a more objective comparison between the two data sets than the EOF analysis of the unaltered optical depth data can provide. Whatever correlation exists between the aerosol size distribution and water vapor amounts should therefore appear in a less diluted form when the data are analyzed after subtracting the mean optical depth values.

Because water vapor is avoided in the Alexandrov *et al.* (2002a) algorithm, the optical depth values for the 935 nm channel of the MFRSR were obtained directly from the Langley regressions, for the purpose of introducing that channel into the EOF analysis. The Rayleigh extinction is subtracted out from the Langley optical depths before the EOF analysis is performed. Unlike nitrogen dioxide and ozone, water vapor absorption does not vary linearly with column amount, so optical depth values derived from Langley regressions for channels sensitive to water vapor absorption will result in some degree of error. However, the purpose of this exercise is not to quantify column water vapor amounts, but to take advantage of the fact that correlations between aerosol

extinction and water vapor absorption appear to reveal information about the aerosol properties in an EOF analysis that could not be obtained as clearly without including channels sensitive to water vapor absorption. Without having the optical depth information from the water vapor channel in the CIMEL data, results from that device are not included in this section of this study. Still, with ozone, nitrogen dioxide, and the calibration-related offset in optical depth between the MFRSR and RSS largely removed from the data by subtracting the means, these EOF's should provide the most objective comparison possible of the aerosol information contained within the data of these two devices.

3.3.2: The First Eigenvectors

Figures 3.9 and 3.10 show the first EOF's, respectively, of the MFRSR and RSS data sets with the means subtracted and water vapor channels included. In both cases, the gaseous absorption bands of ozone and NO_2 are not present, and the oxygen absorption lines show up only barely, but water vapor is strongly present. Both eigenvectors show a strong seasonal dependence, consistent with the aerosol optical depth measured by the two devices, and also with the first eigenvectors of the unaltered data sets. However, much of the detail appears to be missing in the first eigenvector of the MFRSR data, especially for the high optical depth days in October 1999. By contrast, the first eigenvector of the RSS data maintains the features seen in Figures 3.3 and 3.5, including the absence of high values on June 23 and July 5. The first eigenvector of the MFRSR data contributes 89% to the total variance, while the first eigenvector of the RSS

contributes only 75% to the total. Not surprisingly, this indicates that the RSS continues to produce more information beyond the first eigenvector than the MFRSR does. The contribution of the first eigenvector to the total variance is, in both cases, significantly lower with the means subtracted than it had been before, however, confirming that the secondary eigenvectors contain more information, regardless of the device, when the EOF's are produced after the mean optical depth in each channel has been subtracted.

In addition, the coefficients of both EOF's have clearly defined baseline curves, which decrease sharply with wavelength at shorter wavelengths and more gradually at longer wavelengths. The spectral behavior of these curves strongly resembles the behavior of Mie scattering curves for small particle sizes. For the sake of comparison, the coefficients were input into modified aerosol retrieval algorithms, to determine what effective radius and effective variance values best fit the curves. For each size distribution, a least squares linear fit is applied. The x-axis terms are the Mie coefficients, for each channel used in the retrieval, normalized to 550 nm. The y-axis terms are the values of the coefficients of projection for the channels used in the retrieval. Figure 3.11 shows the results of the Mie fit for the first EOF of the MFRSR data with the mean subtracted, while Figure 3.12 shows the corresponding fit for the RSS. The effective radius value retrieved for the MFRSR is 0.138 μm, with a variance of 0.4. By contrast, the effective radius value retrieved for the RSS is 0.160, with a variance of 0.1. Both cases exhibit a positive offset, meaning a baseline greater than zero, resulting from the intercept in the linear fit. If these first eigenvectors really do represent a mode of small aerosol particles, then this result indicates, on a quantitative level, where the effective radius of a fine aerosol mode can be expected to lie. From a strictly physical

standpoint, though, the results do not agree completely with each other, because the MFRSR result suggests a very broad range of small particles, while the RSS results suggests a narrow range of somewhat larger particles. However, as noted in Chapter 1 (see section 1.5, and Figures 1.5 and 1.6), the effective radius and variance are hard to measure uniquely, because the curve for a higher radius value with a low variance can often resemble the curve for a somewhat lower radius value with a high variance. In other words, the measured values in Figures 3.11 and 3.12 are not as different as they might appear at face value. In fact, the EOF analysis might be suggesting that the fine mode effective radius and effective variance simply cannot both be retrieved uniquely, thus confirming limitations to the amount of obtainable information contained within the data from either device. Despite the optical depth difference between the two devices, the MFRSR and RSS seem to observe similarities in the behavior of the small aerosol particles. Furthermore, if the coefficients of these EOF's do in fact represent the Mie curve of the fine mode aerosols being observed, then the two effective radius values presented in Figures 3.11 and 3.12 provide at least a rough estimate of what the mean of the retrieved values for the effective radius of the fine mode should be. The accuracy of these values can then be tested using the retrieval algorithms on the full data sets.

3.3.3: The Second Eigenvectors

The second EOF's of the MFRSR and RSS data, with water vapor included and the mean optical depths subtracted, are shown in Figures 3.13 and 3.14. These eigenvectors contribute 10% to the total variance of the MFRSR EOF's, and 22% to the

RSS EOF's. For the MFRSR, the inverted aerosol optical depth pattern observed in the second eigenvector of the unaltered data set emerges in the second eigenvector with the means subtracted as well. The second eigenvector for the RSS, on the other hand, shows an upright seasonal aerosol cycle. The high aerosol days in October appear to be inverted, but June 23 and July 5 do figure prominently in the positive direction in this eigenvector. This further reinforces the notion that these two days, which had the two highest optical depth values at 870 nm in the entire data set, differ in a significant way from the other high aerosol days. The baseline curves in the coefficients of projection steadily increase with wavelength for both devices, although much more sharply at lower wavelengths for the MFRSR than for the RSS. Curiously, the water vapor line points downward in the coefficients of this EOF for the MFRSR, but upward, as was the case in the first EOF for both devices, for the RSS. The second EOF's for the two devices clearly do not show the same thing. Judging from the relative contributions to the total variance, the second eigenvector of the RSS data set contains more information than the corresponding eigenvector of the MFRSR data set.

The coefficients of projection of the second EOF of the RSS data show a soft increase with wavelength. To determine if this curve is consistent with the Mie extinction curve of a larger, or coarse, aerosol mode in a bimodal size distribution, a Mie fit is applied to these coefficients as well, and presented in Figure 3.15. An effective radius of 1.5 μm, with a corresponding effective variance of 0.4, produces the best fit. The negative offset needed to produce this fit, though, is fairly large. Since the offset in the first eigenvector is positive, the results for the two eigenvectors somewhat balance each other out, but not enough to justify drawing firm quantitative conclusions about the

aerosol size distribution. The second EOF of the MFRSR data is even more difficult to interpret, given the lower variance and the inverted water vapor signal. The coarse mode Mie fit, shown in Figure 3.16, produces an offset that is far too large to justify interpreting the eigenvector as a coarse aerosol mode, confirming that the second eigenvectors of the RSS and the MFRSR data with the mean subtracted indeed do not represent the same physical feature. Only the difference in the number and range of wavelengths used by the two devices can account for this clear difference in observation.

The existing aerosol information appears to define the fine mode better than the coarse mode, but it appears unlikely that the two modes can be uniquely separated from each other, even in the RSS data. Qualitatively, the results of the EOF analysis of the RSS data do indicate that the aerosol size distribution can be expressed as the sum of two curves, the first of which decreases fairly sharply with wavelength, and the second of which increases gently with wavelength. This strongly suggests that the aerosol distribution is bimodal, and that any accurate aerosol retrieval needs to take bimodality into account. The two eigenvectors might not represent an exact, one-to-one correspondence with the fine and coarse aerosol modes over the entire data set, but if the aerosol properties changed during this time period, there would be no reason to expect such uniformity. The second eigenvector of the MFRSR data does not show anything that can be easily interpreted as an additional, independent piece of aerosol information, on account of the MFRSR's smaller wavelength range and less spectral detail relative to the RSS.

3.4: The MFRSR and CIMEL "Equivalent" EOF's

The MFRSR and CIMEL EOF's can be directly compared and contrasted with the EOF's of the full RSS data set, but more information can be obtained by producing EOF's using only the RSS channels closest to the wavelengths of the other two devices, thereby creating MFRSR and CIMEL "equivalent" EOF's. For one thing, since the "equivalent" data sets are part of the full RSS data, optical depth differences are no longer an issue in making comparisons with the RSS, and any differences in the resulting EOF's stem solely from the different combinations of wavelengths used. Changes in the shapes of the eigenvectors and the coefficients of projection, and the contribution of the eigenvectors to the total variance, can be examined as a function of the wavelength combination using one set of data. Also, as the MFRSR and CIMEL use several broadband filters while the RSS uses a high-resolution array, the possibility of water vapor absorption affecting one of the non-vapor channels of the MFRSR or the CIMEL can be disregarded for the "equivalent" RSS channels. Because the spectral response is only examined over narrow intervals, out-of-band contributions may arise in the MFRSR and CIMEL data and be responsible for some of the differences between these two devices and the RSS. Finally, given the absence of data from the water vapor channel of the CIMEL device, and the problems already observed with the 670 nm channel, the "CIMEL equivalent" data set, with the water vapor channel added, should most clearly how the aerosol size distribution manifests itself in the EOF's using the wavelength combination employed by the CIMEL.

3.4.1: The "Equivalent" Sets, without Water Vapor

The coefficients of the first two EOF's of the full set (that is, without subtracting the means) of "MFRSR equivalent" data, presented in Figures 3.17 and 3.18, qualitatively resemble those of Figures 3.1 and 3.2 very closely. The first set of coefficients resembles the mean optical depth for the data set, while the second shows ozone absorption above a mostly flat baseline. However, the coefficients in Figure 3.18, while showing a similar baseline curve and ozone signal to the ones in Figure 3.2, have much lower values, and a slightly more pronounced curve to them. Given that the optical depth values for the MFRSR were significantly larger in general than those for the RSS, this is not a terribly surprising result, but it is possible that this difference will affect the way ozone is measured in the retrieval algorithms for the RSS and MFRSR. The pattern in the first eigenvector reflects the temporal variation of the aerosol optical depth, just as before. Even in the "MFRSR equivalent" data set, though, the high optical depth values on June 23 and July 5 are not reflected in this eigenvector, indicating that whatever feature distinguishes these two days from the rest of the data set does not become lost when only the five MFRSR wavelengths are examined. The second eigenvector accounts for 5% of the total variance, about 2% more than the corresponding MFRSR eigenvector presented in Figure 3.2. This difference must result from the optical depth differences between the two devices, although the reason for it is unclear. Still, the fact that the variance and the spectral pattern of the coefficients of projection for the second eigenvector of the "MFRSR equivalent" set agree most closely with that of the MFRSR data indicates that these quantities are largely by-products of the chosen wavelength

combination, and do not depend very much on the specific data set used. The inverted optical depth pattern in the eigenvector mirrors the pattern seen in Figure 3.2.

The first EOF for a "CIMEL equivalent" isolation of the RSS data is shown in Figure 3.19. As was the case for the EOF analysis of the CIMEL data itself, the first eigenvector accounts for over 99% of the total variance. This confirms that the degree of variance contributed by the secondary eigenvectors, in any of these data sets, depends on the presence of gas absorbers. The stronger the absorption, the greater the amount of variance in the secondary eigenvectors, and the greater the separation of the aerosol extinction into distinct parts in the coefficients of projection of these eigenvectors. However, at least some element of the aerosol extinction remains independent of the first eigenvector of the RSS and CIMEL data, regardless of the wavelength combination used. As Figure 3.19 shows, the high aerosol optical depth days of June 23 and July 5 once again do not show up in the first eigenvector.

3.4.2: With Water Vapor, Subtracting the Mean Optical Depths

Figures 3.20 and 3.21 show the first and second eigenvector for the "MFRSR equivalent" data set, with a water vapor channel included and the mean optical depths subtracted out before the EOF's are calculated. The variance values very closely resemble those of the first two eigenvectors of the MFRSR data, with the mean subtracted. Predictably, the patterns of the eigenvectors resemble those seen in all the other EOF's of the RSS data. June 23 and July 5 emerge very conspicuously above the rest of the data in the second eigenvector, even though the high optical depth days in

August and October are inverted in the graph. The coefficients of projection show the same wavelength dependence observed in the corresponding EOF's of the MFRSR data (see Figures 3.9 and 3.13). The water vapor line points down in the coefficients of projection of the second eigenvector. As this agrees with the MFRSR EOF's, but not the EOF's of the full RSS data set, the observed pattern must depend on the combination of wavelengths used, and not on the device.

The first two EOF's of the "CIMEL equivalent" data are presented in Figures 3.22 and 3.23. The second eigenvector contributes 16% to the total variance, squarely in between the contributions of the second eigenvectors obtained from the full RSS and the "MFRSR equivalent" data sets. Although the seasonal cycle does not emerge strongly from the components of the second eigenvector, the patterns in the two eigenvectors otherwise resemble those from the "MFRSR equivalent" data set. The most distinguishing feature of Figure 3.23, though, is the complete absence of water vapor in the coefficients of projection. The water vapor absorption signal pointed up in the full RSS data, and pointed downward in the "MFRSR equivalent" data, but vanished in the "CIMEL equivalent" data. Evidently, the intermediate amount of information contained in the CIMEL wavelength range produces this result. Furthermore, the baseline curve in the coefficients of projection is less pronounced that the corresponding curve in the "MFRSR equivalent" data, but more pronounced than what was seen using the full spectral data of the RSS. If the second EOF of the RSS data with the mean subtracted truly represents a coarse aerosol mode, while the MFRSR and "MFRSR equivalent" data show something more ambiguous, then the "CIMEL equivalent" data, and by analogy the data from the CIMEL itself, at least obtain a closer approximation of the large mode than

the MFRSR does. The "CIMEL equivalent" data does not completely approximate the CIMEL data, though, because the wavelength range of the RSS does not extend to 340 nm. Therefore, the CIMEL likely contains more information than the EOF analysis indicates

When a small mode fit is applied to the first eigenvectors of the "equivalent" data sets, as shown in Figures 3.24 and 3.25, the results agree well with what was already observed in the MFRSR and full RSS EOF's (see Figures 3.11 and 3.12). Not surprisingly, the effective radius obtained from the "CIMEL equivalent" data, 0.156 μm, more closely matches the value from the full RSS spectral data than does the effective radius obtained from the "MFRSR equivalent" data, 0.152 μm. For the three EOF's using RSS data, the fit does not change significantly when the wavelength combination is changed. If the aerosol size distribution is indeed bimodal, and the first eigenvectors in the data sets where the mean optical depth values have been subtracted do represent the extinction due to fine mode aerosols, then this would indicate that all three devices are capable of measuring, at the very least, the effective radius of the fine mode. Having said that, it does appear that the MFRSR would be somewhat less capable than the other two devices of defining the coarse and fine modes uniquely, due to a narrower wavelength range, and consequently will not produce as accurate results, even if a bimodal distribution is assumed.

3.5: Detecting Defects in the Data with EOF Analysis

In addition to establishing the limits to obtainable information contained within a data set, and determining what information can be retrieved within these limits, the EOF analysis of a set of sunphotometer data can also reveal minor defects, or flaws, within the data that might not be immediately noticeable in a cursory glance of the data from individual days. For example, Figure 3.26 shows the fifth eigenvector of the RSS optical depth data, using the full RSS wavelength range, and not subtracting the mean optical depths beforehand. Long-term wavelength stability has been as issue for the RSS (Harrison *et al*. 2003); the central wavelength of a given channel in the device can fluctuate slightly over time. In an EOF analysis, such fluctuations will manifest themselves, in the spectral patterns of the coefficients of projections, as first derivatives of the absorption coefficients of the gases. Such first-derivative patterns are clearly seen in Figure 3.26, and appear in some of the other subsidiary eigenvectors as well. The days on which the absolute value of the EOF is high are the days when the wavelength shift is most significant.

The third eigenvector of the CIMEL data, shown in Figure 3.27, suggests that something affected every channel of data throughout the course of the data set, with potentially serious implications for the CIMEL aerosol retrievals. The eigenvector components show a very sharp downward trend until about day 500, at which point they rise sharply, indicating an apparent discontinuity in the CIMEL data. This means that some device problem or defect had been skewing the CIMEL device for most of the duration of this data set, and some sort of correction or adjustment was made to the

device around day 500. Indeed, the ARM device log for the CIMEL at the SGP site indicates that a problem with a chip in the sensor head was identified on April 17 (day 473) and that a replacement chip was sent to SGP, but the logs do not indicate specifically when the new chip arrived and was installed. The low contribution to the variance for this eigenvector implies that its overall effect on the output of the device was more subtle than dramatic, but the shape of the eigenvector clearly shows that the problem interfered with the full data set. Furthermore, the plot of the coefficients vs. wavelength confirms that the output of all the channels was altered, but no channel more so than the 670 nm channel. The strong value of the coefficient at 670 nm implies that this EOF reflects the problems already observed in this channel (see Figure 2.9). However, the peak in the coefficient at 670 nm is accompanied by a minimum at 440 nm, suggesting that what had appeared to possibly be an absorption signal from nitrogen dioxide (see Figure 2.10) might also actually be the result of the same device-related error. It is already clear that the 670 nm channel cannot be used in any CIMEL retrievals for this data set, but if the problem in this channel affects the other channels as well, as this EOF implies, the retrievals could conceivably show a trend in the results resembling the pattern in this eigenvector. In the optical depth data, the trend is too subtle to emerge without the EOF analysis, but retrievals of the aerosol size distribution, both using only the optical depth data and the combined optical depth and almucantar data, might bear the mark of this defect in the device.

3.6: Conclusions

The variance values observed in the EOF's of the MFRSR, RSS, and CIMEL data support the conclusions of the theoretical analysis of Box *et al*. (1996), that at most three independent pieces of aerosol information are obtainable from data in this wavelength range. Still, it remains to be determined which pieces of information are retrievable, and which are not. The limitations in the data would not be important if the aerosol size distribution can be expressed as a single mode, but the EOF's calculated after the mean optical depth values are subtracted suggest that the size distribution is in fact bimodal. If this is true – and it still needs to be verified by the retrieval algorithms – then there is not enough information contained within the wavelength range of any of these devices to uniquely define all of the aerosol properties. Compounding this problem is the inherent difficulty in separating the contributions of aerosols and the different gases to the total extinction. In particular, the lack of spectral resolution in the MFRSR makes distinguishing NO_2 absorption from small-particle aerosol extinction in the 415 nm channel virtually impossible. The first step in establishing the best retrieval algorithm, then, is to optimize the measurements of the gas amounts. Once this has been done, the unimodality or bimodality of the aerosol size distribution can be readily determined.

In addition to suggesting that the aerosol size distribution is bimodal, The EOF analysis has also indicated that there is insufficient information in the data to measure the fine mode effective variance independently of the effective radius. If true, an assumption would have to be made about the variance to retrieve consistent effective radius values. Furthermore, the EOF analysis with the mean optical depths subtracted out sets the value

for the fine mode effective radius in the area of 0.16 μm if the variance is assumed to be low, or 0.13 μm if the variance is assumed to be high. This conclusion can be tested using both the bimodal retrieval algorithms in place for AERONET (Dubovik and King 2000), and those developed for use in this study. First, though, the retrievals assuming a single-mode aerosol size distribution need to be evaluated closely, to determine if they can adequately measure both the gases and the aerosols, or if, as the EOF analysis performed here suggests, an assumption of a single aerosol mode will necessarily result in error.

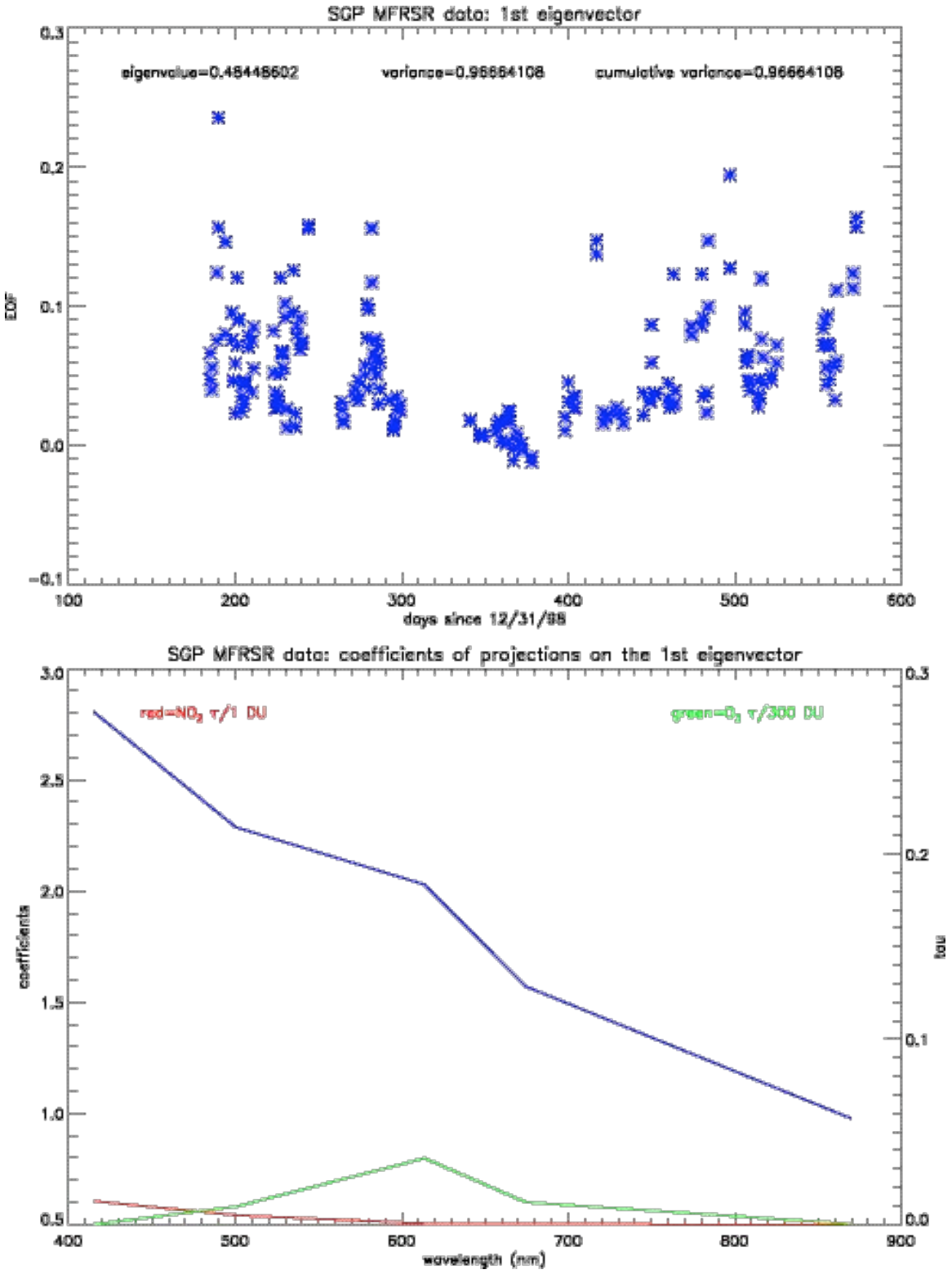

Figure 3.1: The first EOF of the full MFRSR data set.

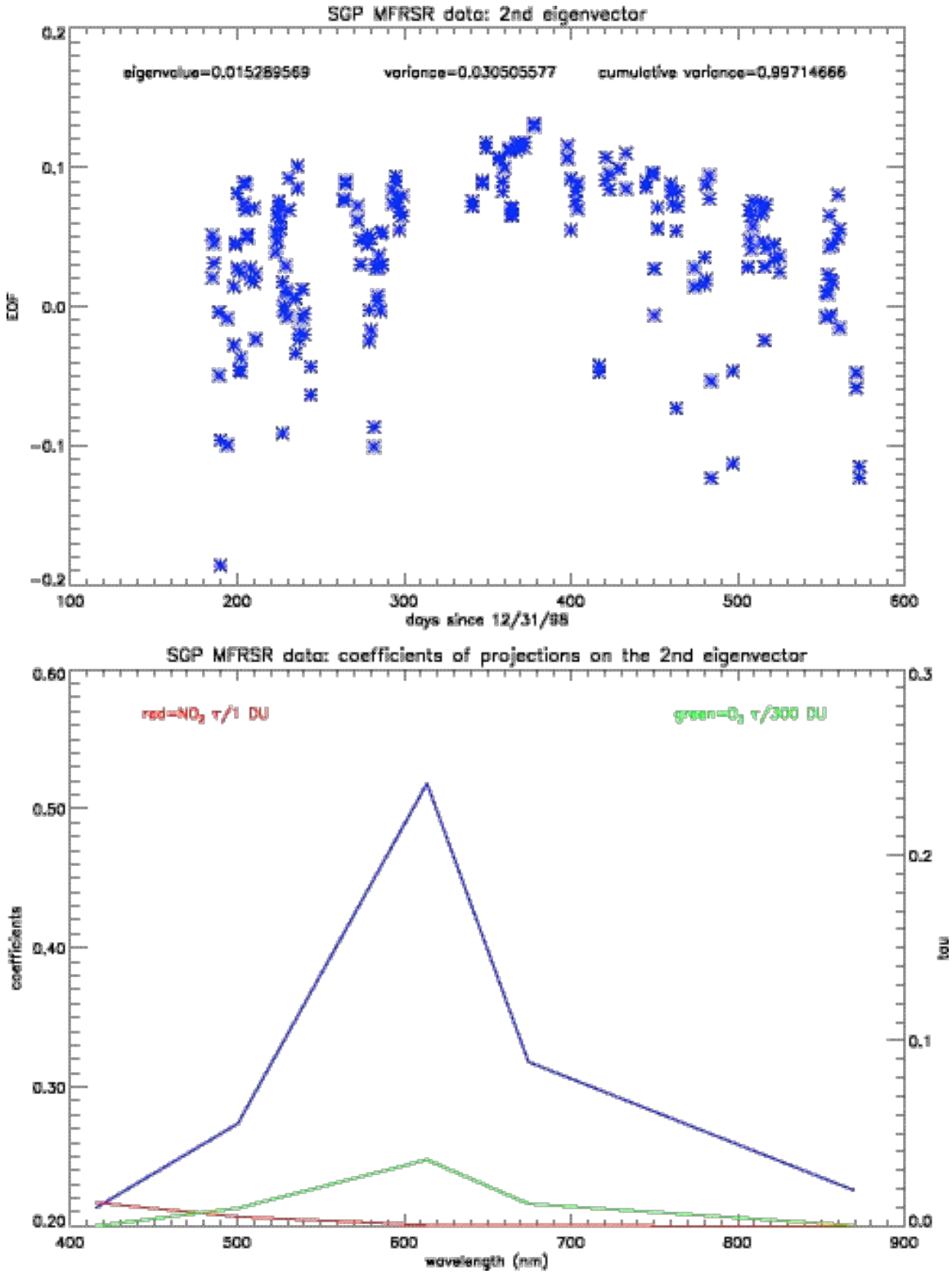

Figure 3.2: The second EOF of the full MFRSR data set. Ozone clearly dominates, although there is a slight baseline curve as well.

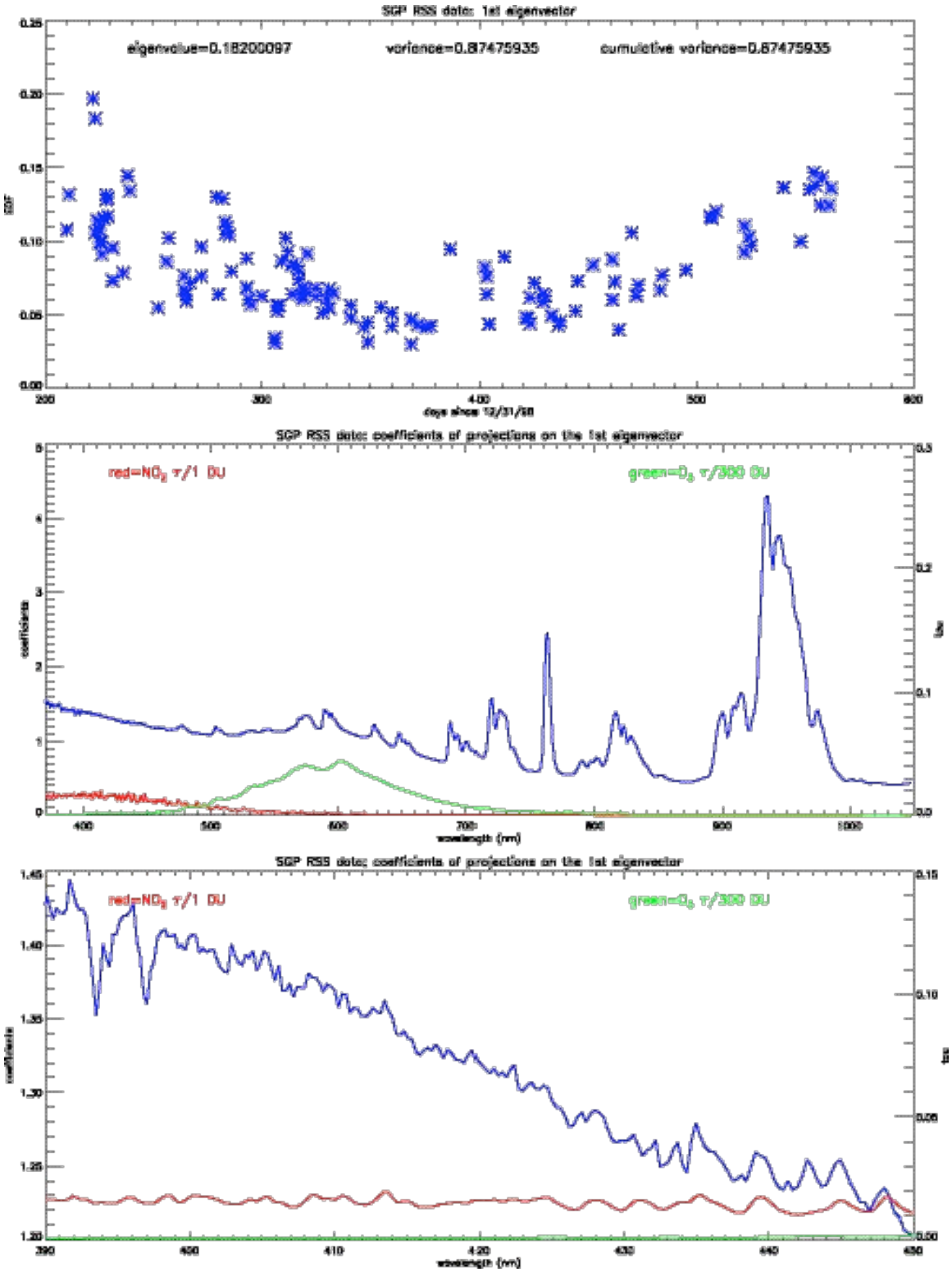

Figure 3.3: The first EOF of the full RSS data set, plotted vs. day. The corresponding coefficients of projections are plotted vs. wavelength on the middle graph, and in closer detail between 390 and 450 nm on the bottom graph.

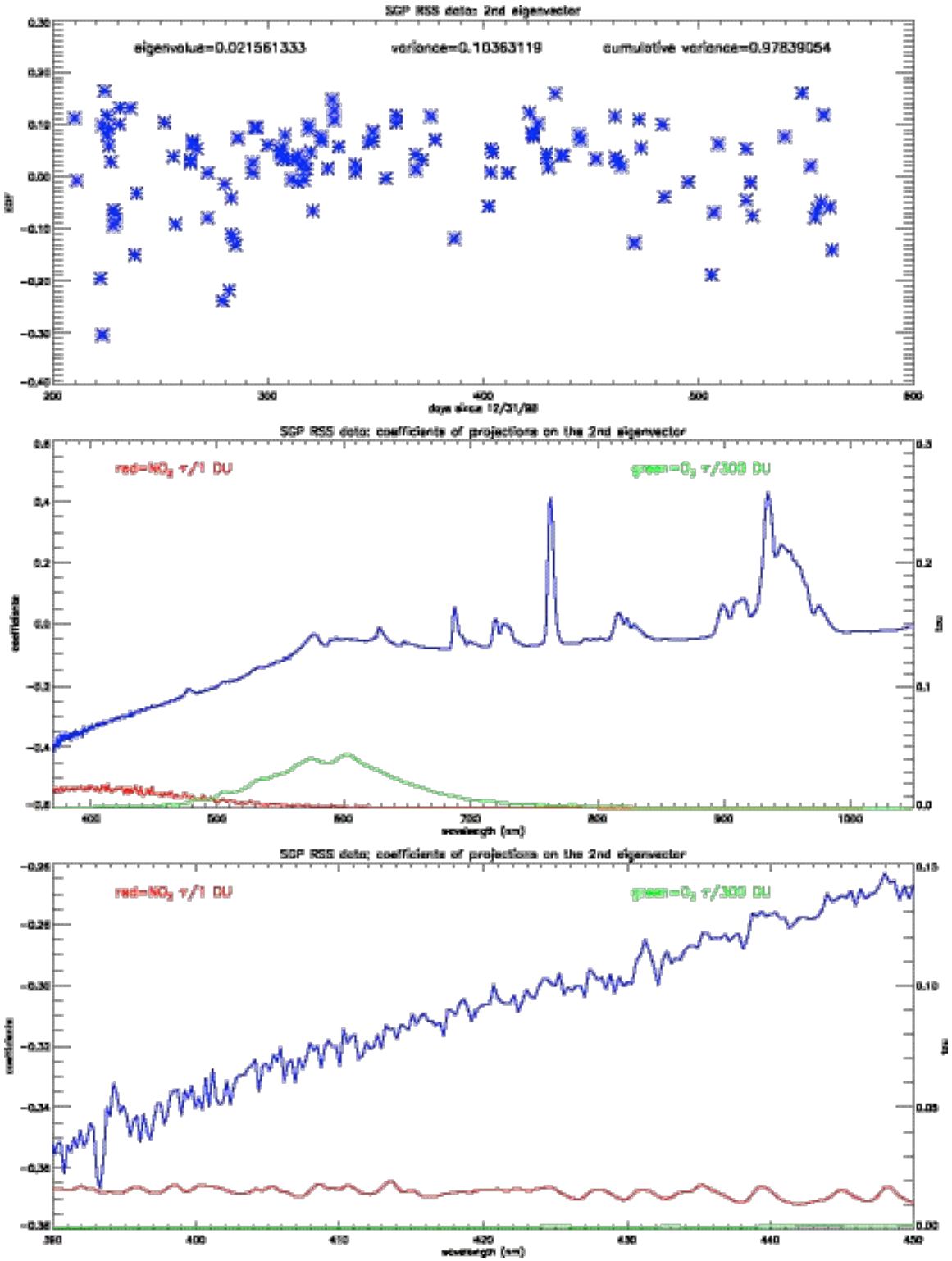

Figure 3.4: The second EOF for the full RSS data set. The water vapor and oxygen absorption lines are present, along with a sharp drop below 600 nm.

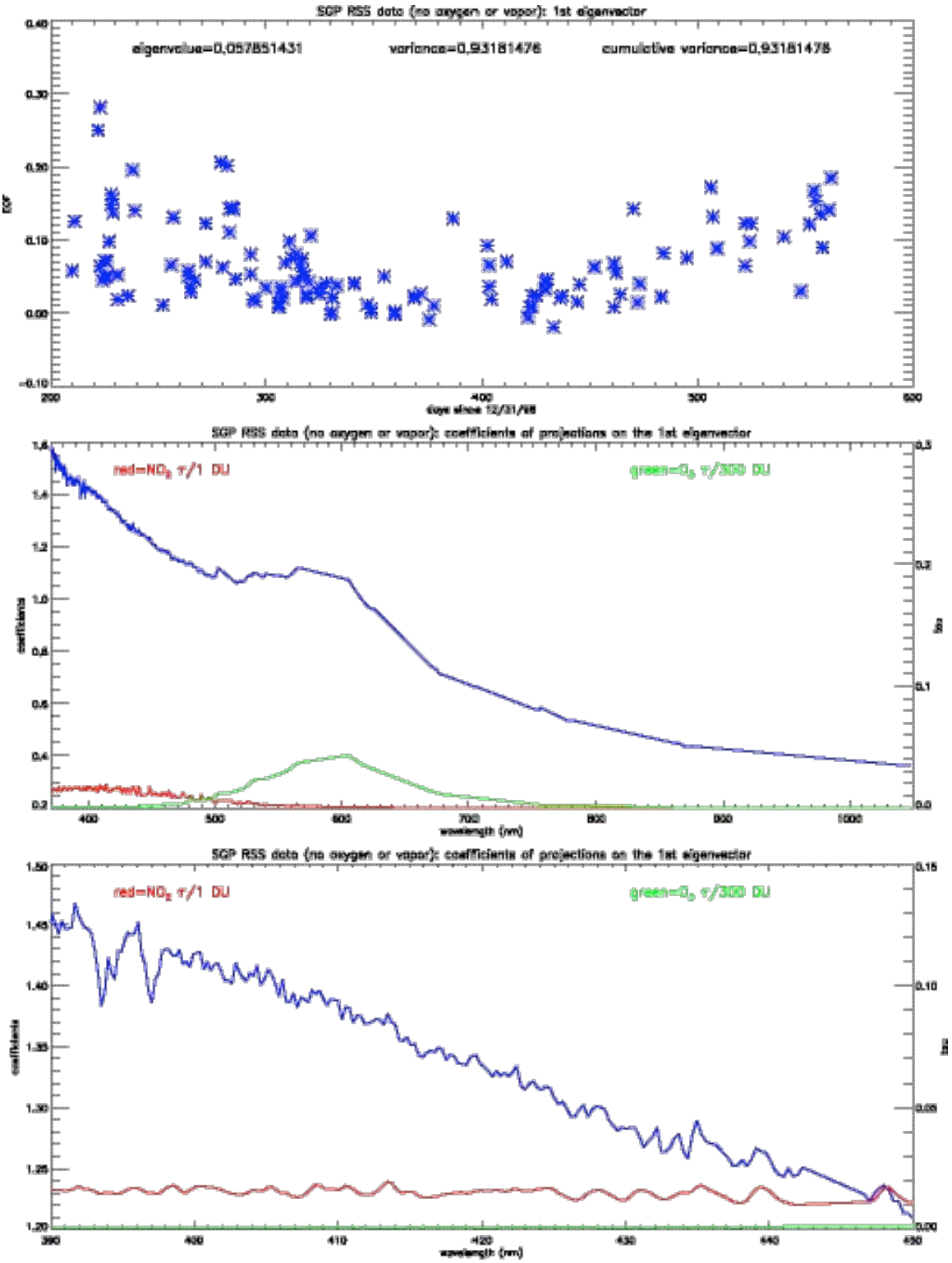

Figure 3.5: The first eigenvector of the RSS data set, without any of the oxygen and water vapor absorption wavelengths. The contributions of NO_2 and the Ring effect are clearly discernible in the lowest graph.

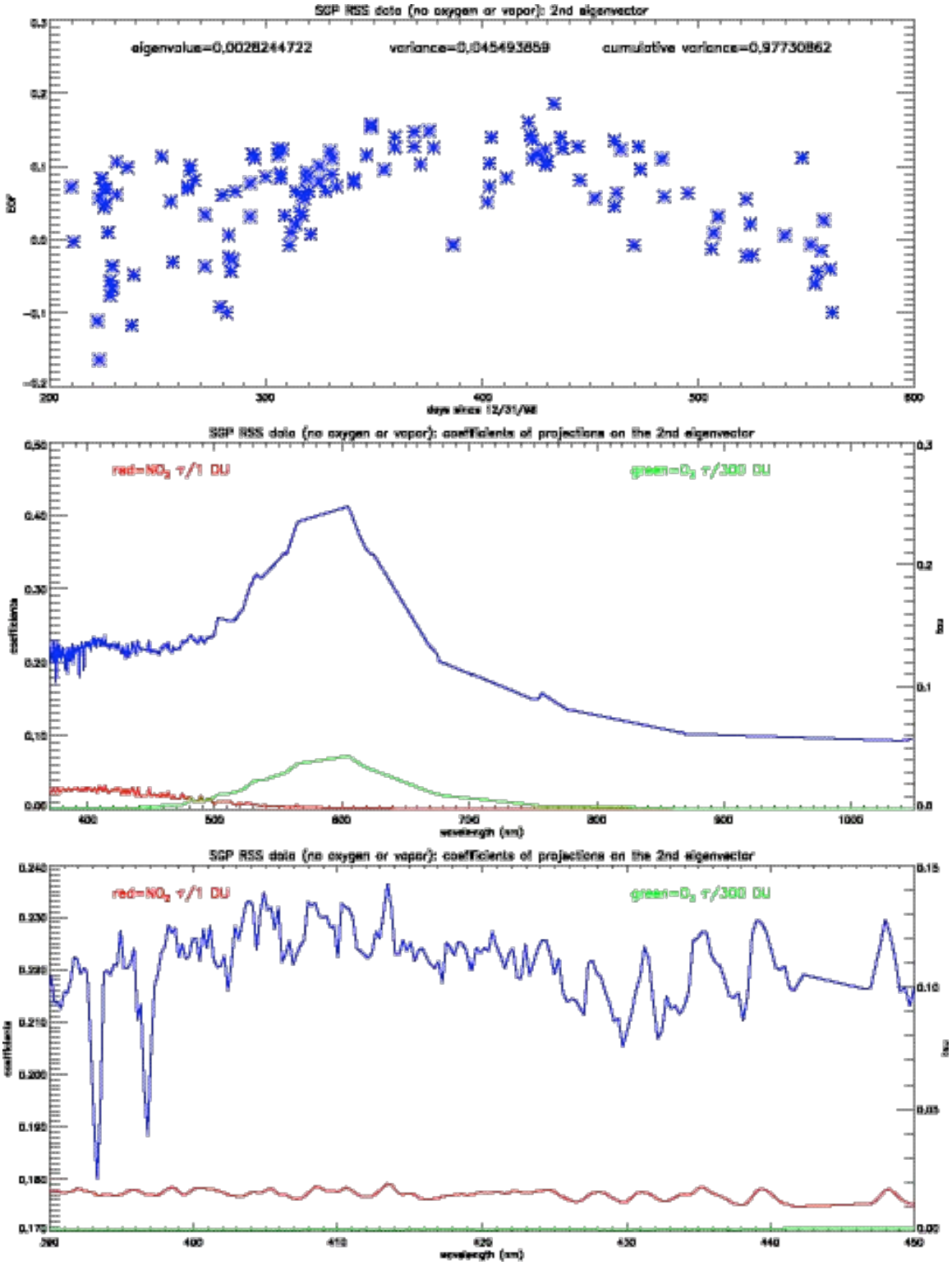

Figure 3.6: The second EOF for the RSS data set without water vapor or oxygen.

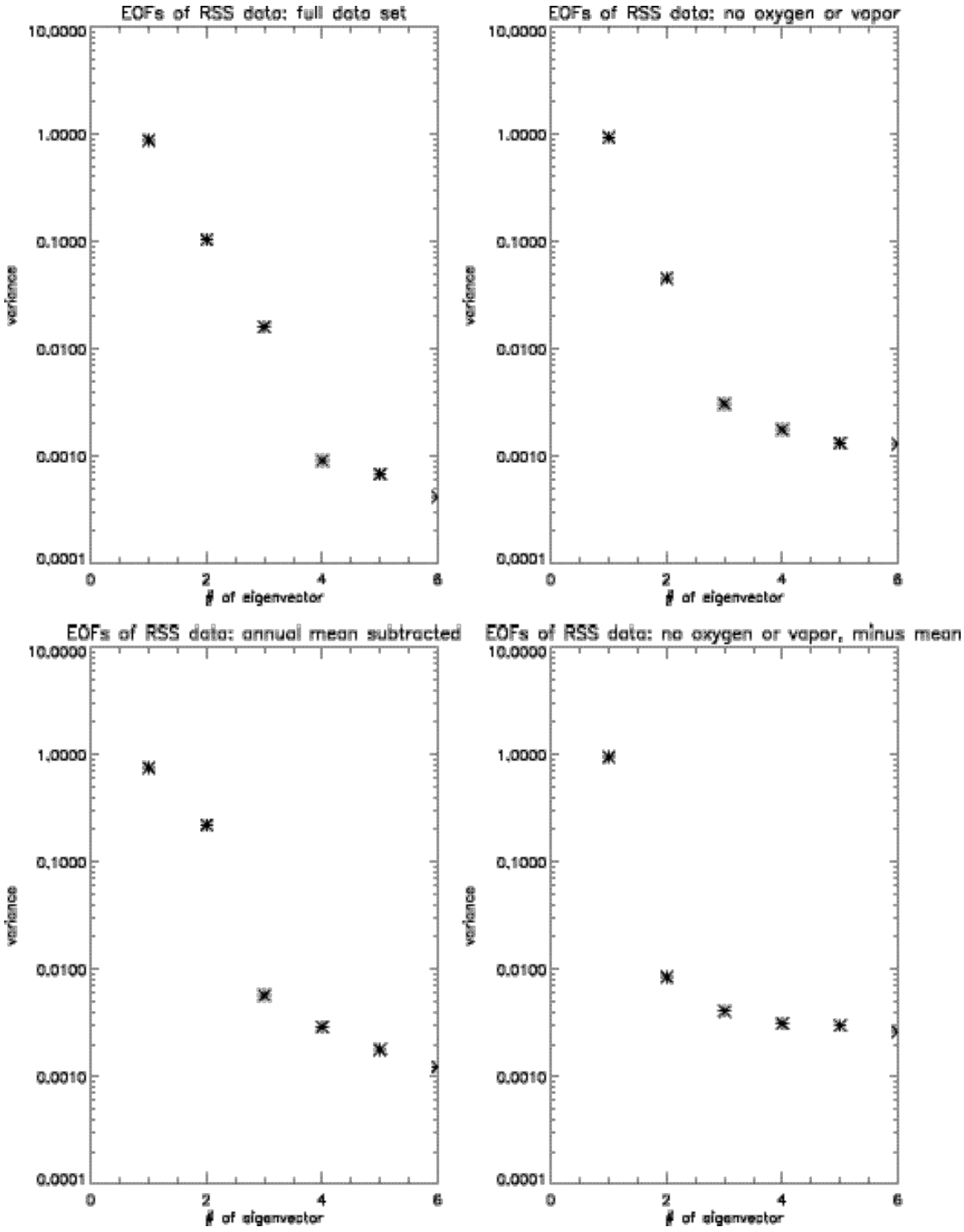

Figure 3.7: Logarithmic plots of the contributions to the total variance for several sets of eigenvectors using RSS data.

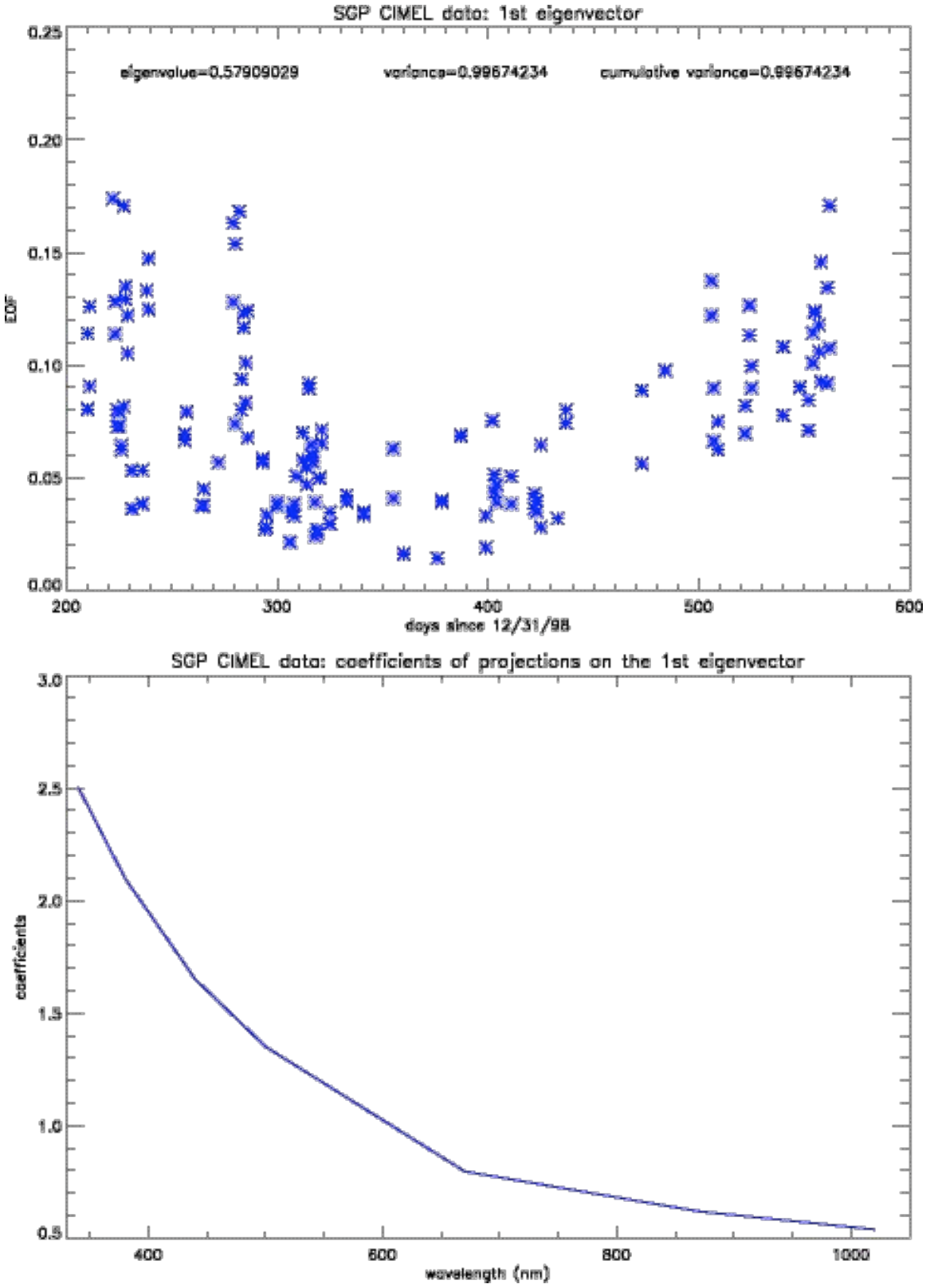

Figure 3.8: The first EOF of the full CIMEL data set.

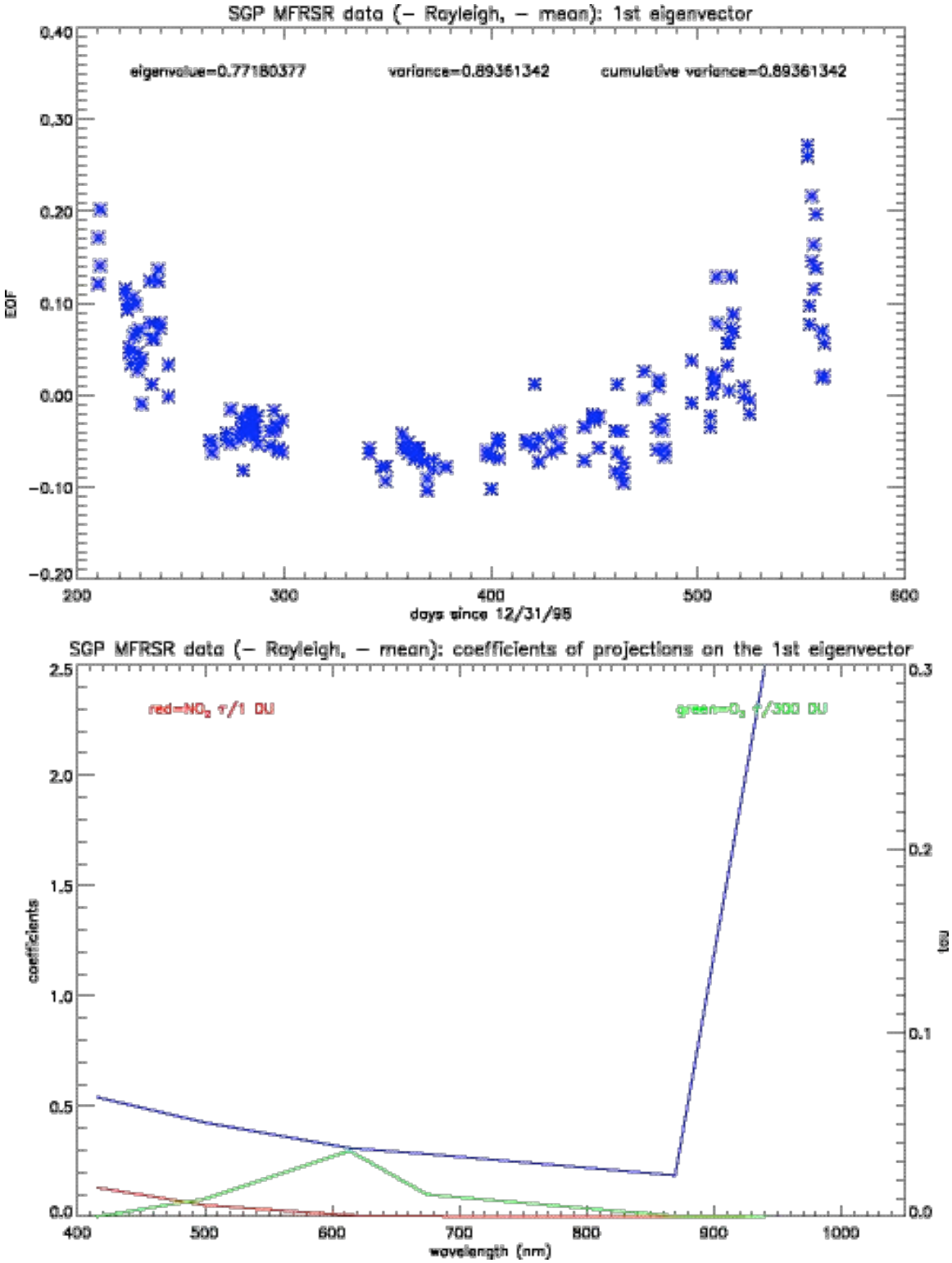

Figure 3.9: The first EOF of the MFRSR data, with the water vapor channel included and the means subtracted.

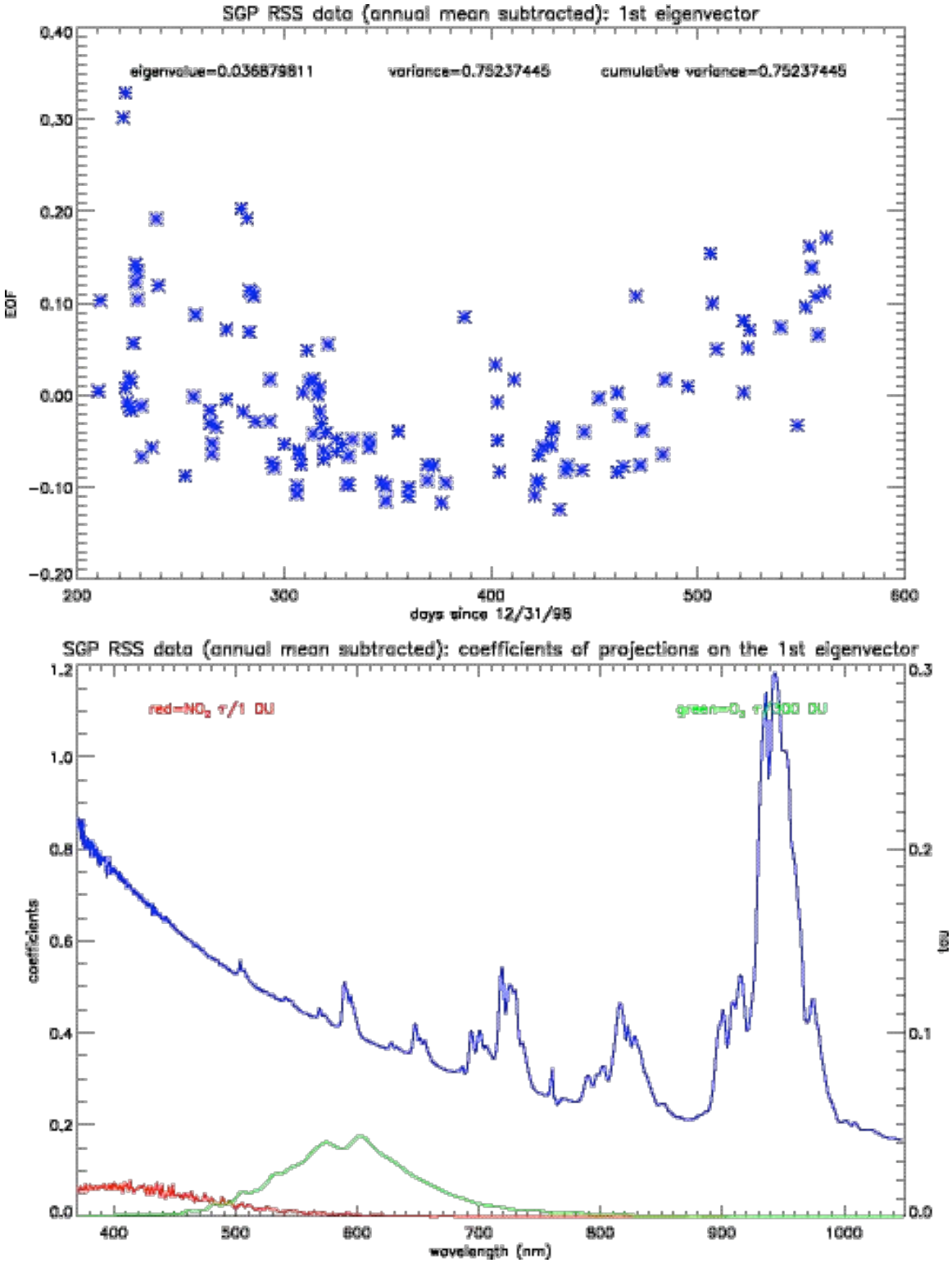

Figure 3.10: The first EOF of the RSS data, with the water vapor channel included and the means subtracted.

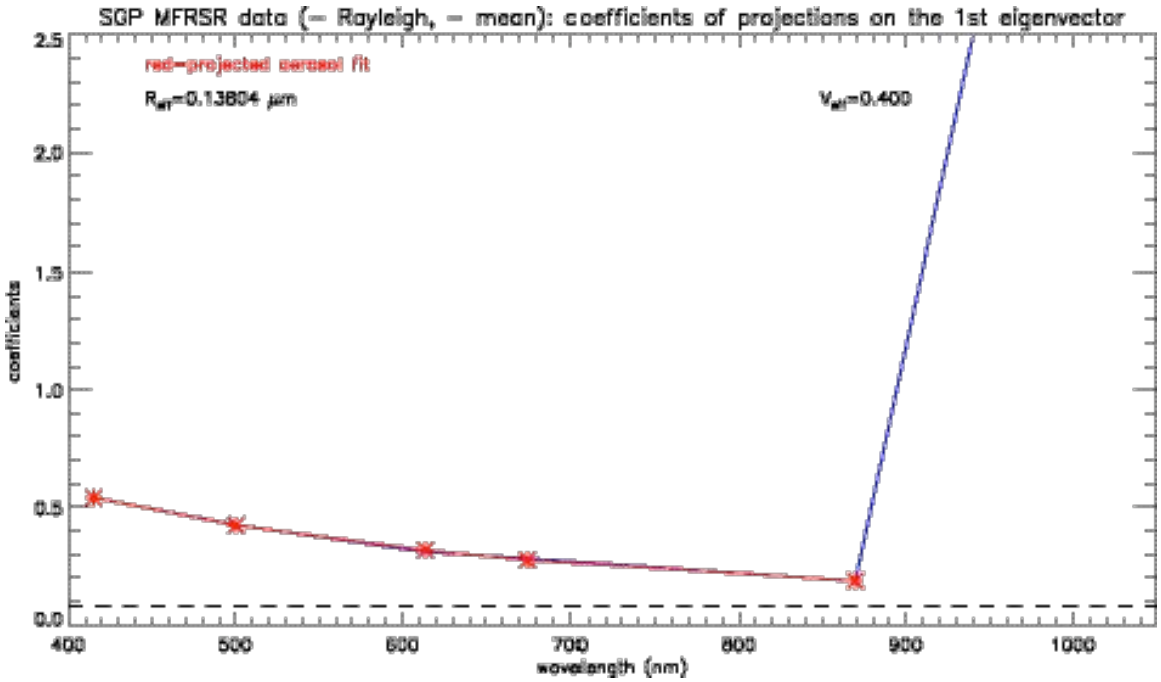

Figure 3.11: The best Mie fit to the first EOF of the MFRSR data, including the water vapor channel, with the mean subtracted.

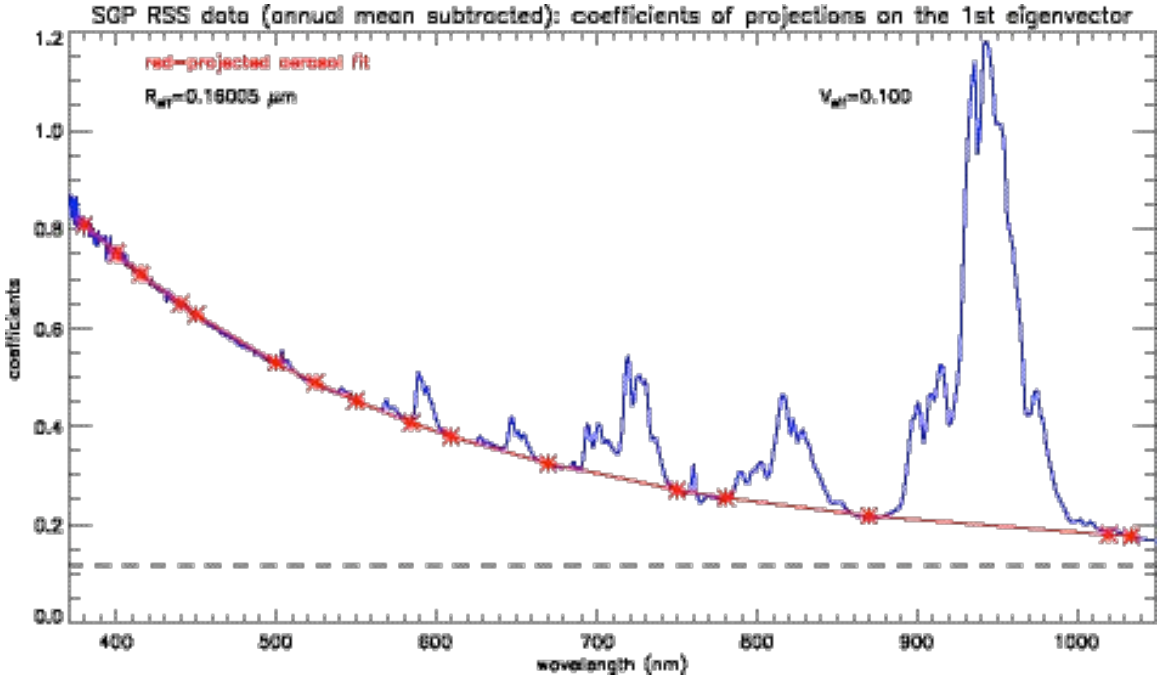

Figure 3.12: The best Mie fit to the first EOF of the RSS data, with the mean subtracted.

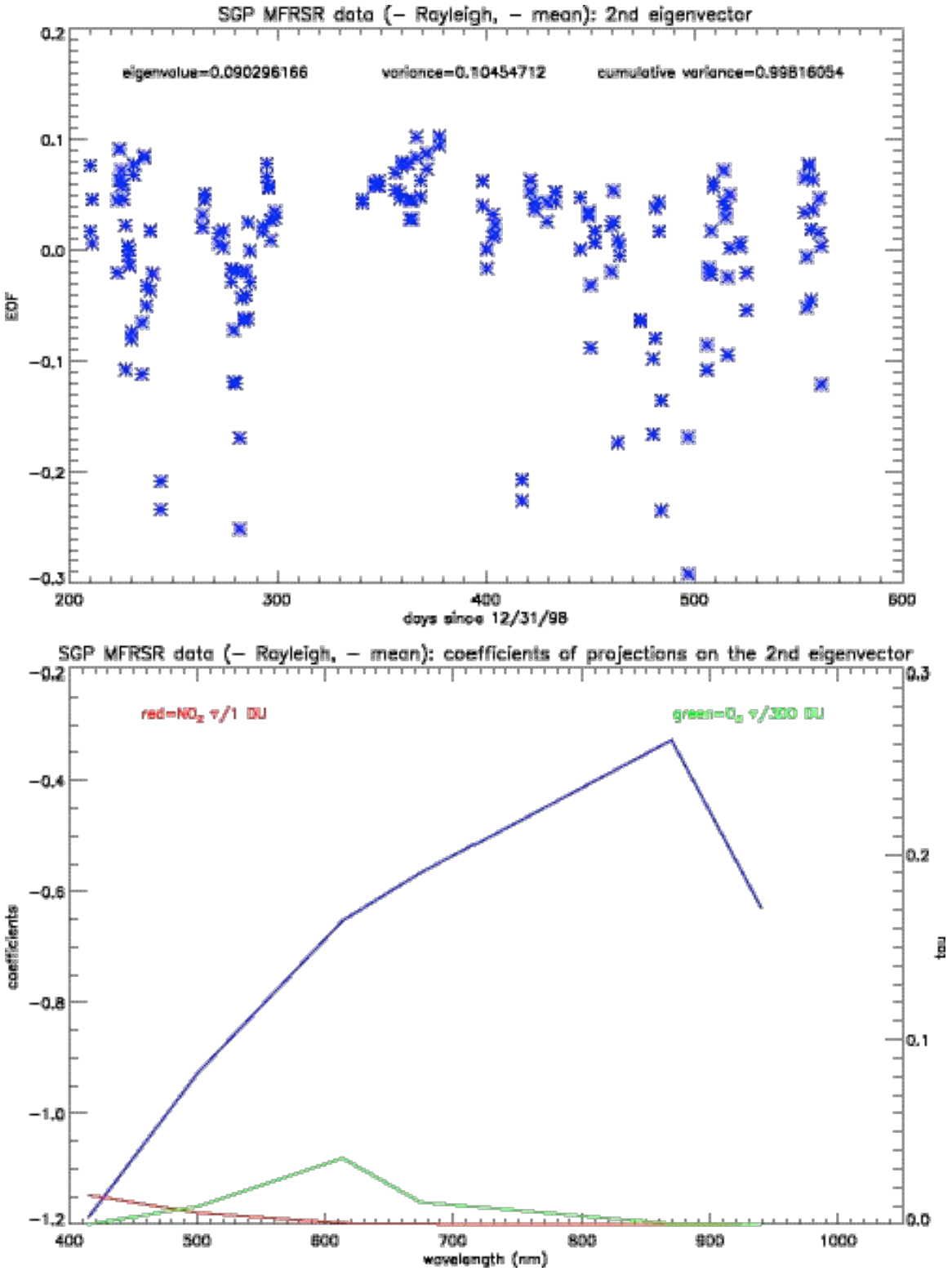

Figure 3.13: The second EOF of the MFRSR data, with the water vapor channel included and the means subtracted.

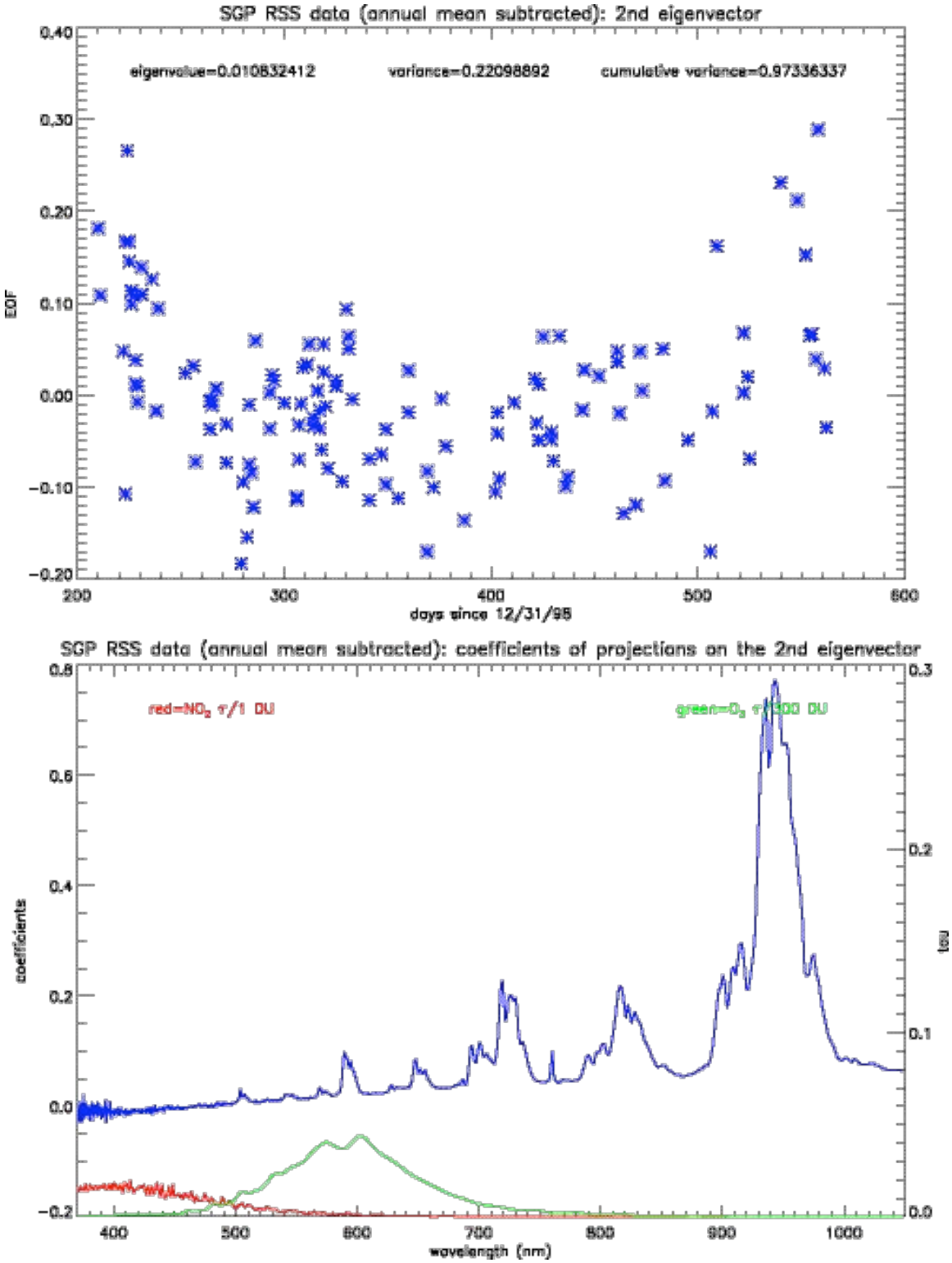

Figure 3.14: The second EOF of the RSS data, with the means subtracted. The baseline curve resembles the Mie curve of the coarse mode in a bimodal aerosol size distribution.

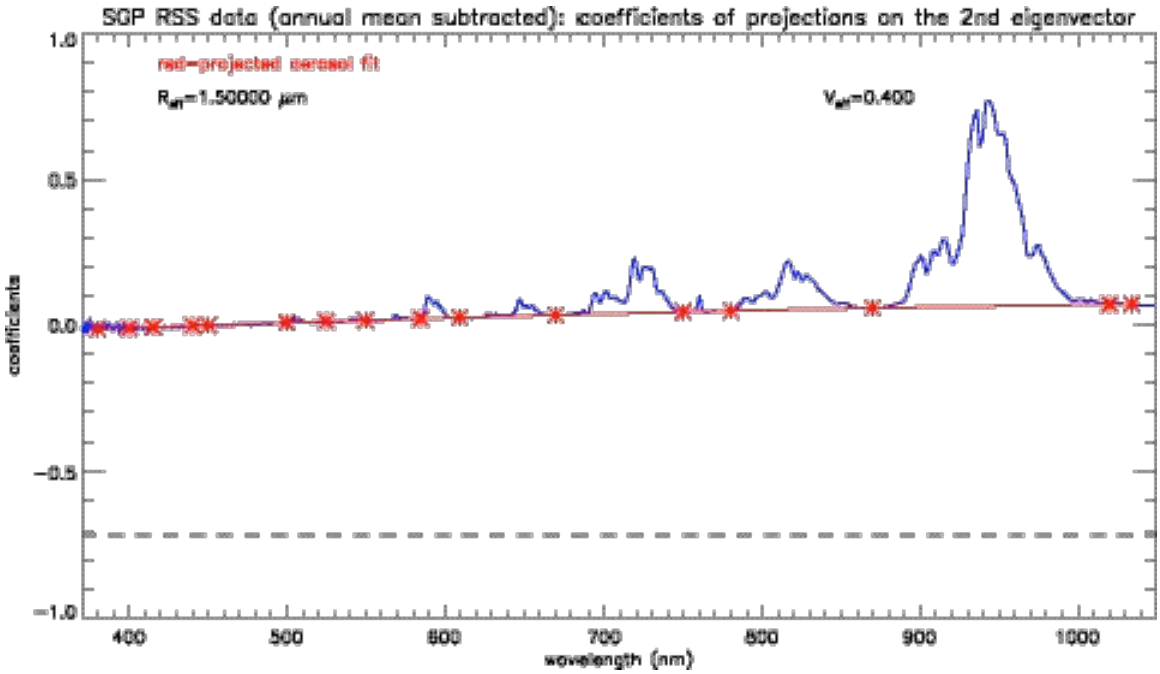

Figure 3.15: The best Mie fit to the second EOF of the RSS data, with the mean subtracted. The coarse mode fit appears reasonable, although the negative offset is significant.

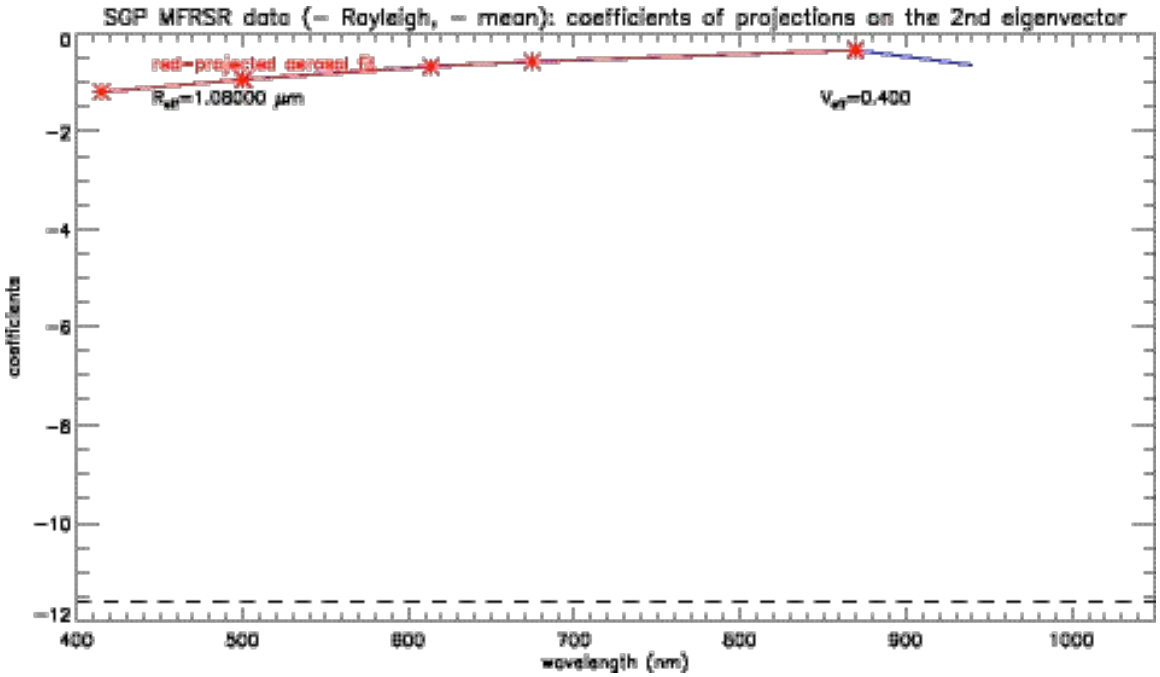

Figure 3.16: The best Mie fit to the second EOF of the MFRSR data, with the mean subtracted. The coarse mode fit produces a very large negative offset, indicating that the MFRSR data does not define a coarse aerosol mode well.

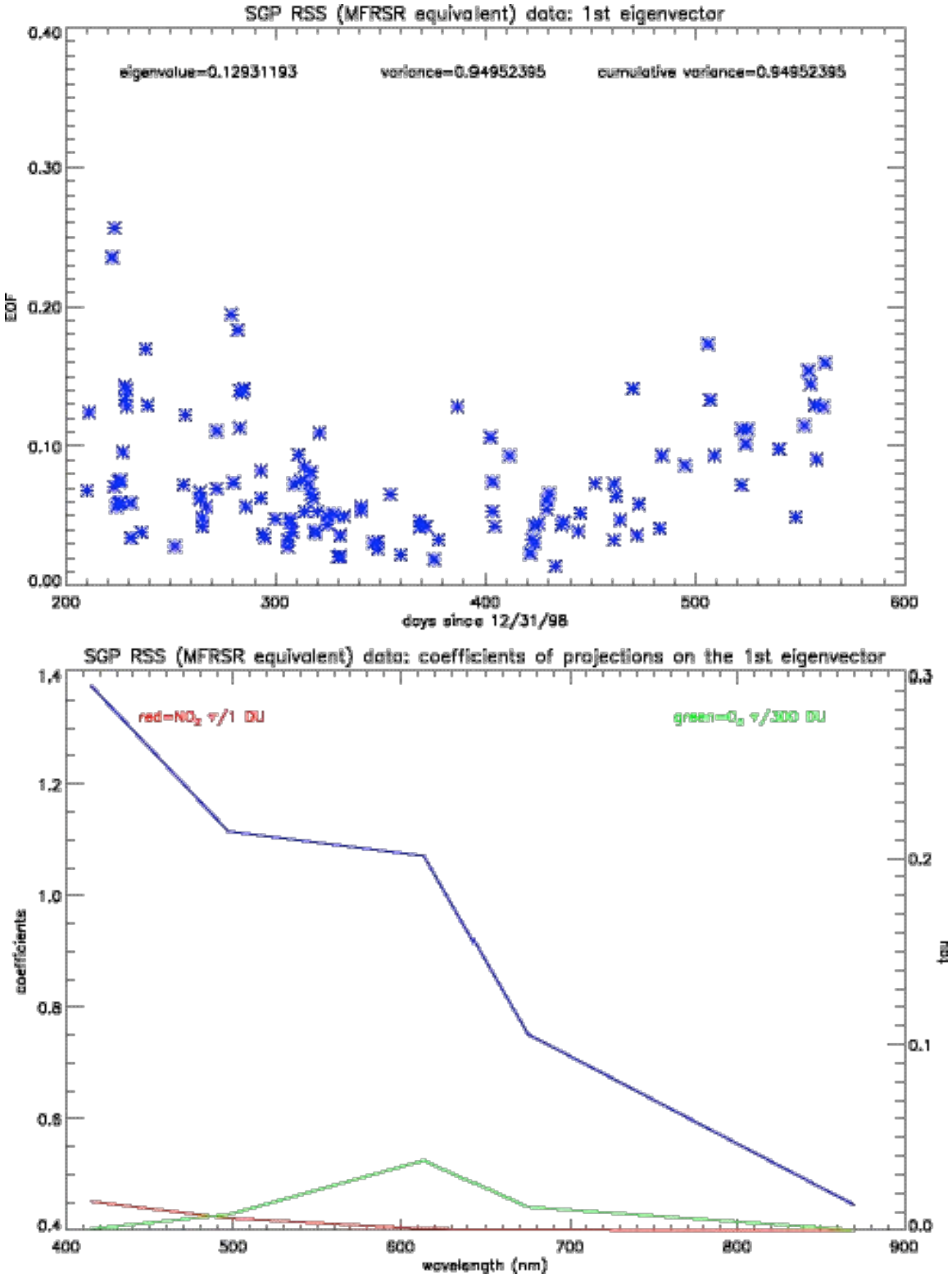

Figure 3.17: The first EOF for the "MFRSR equivalent" RSS data set, which bears strong qualitative resemblance to the EOF in Figure 3.1.

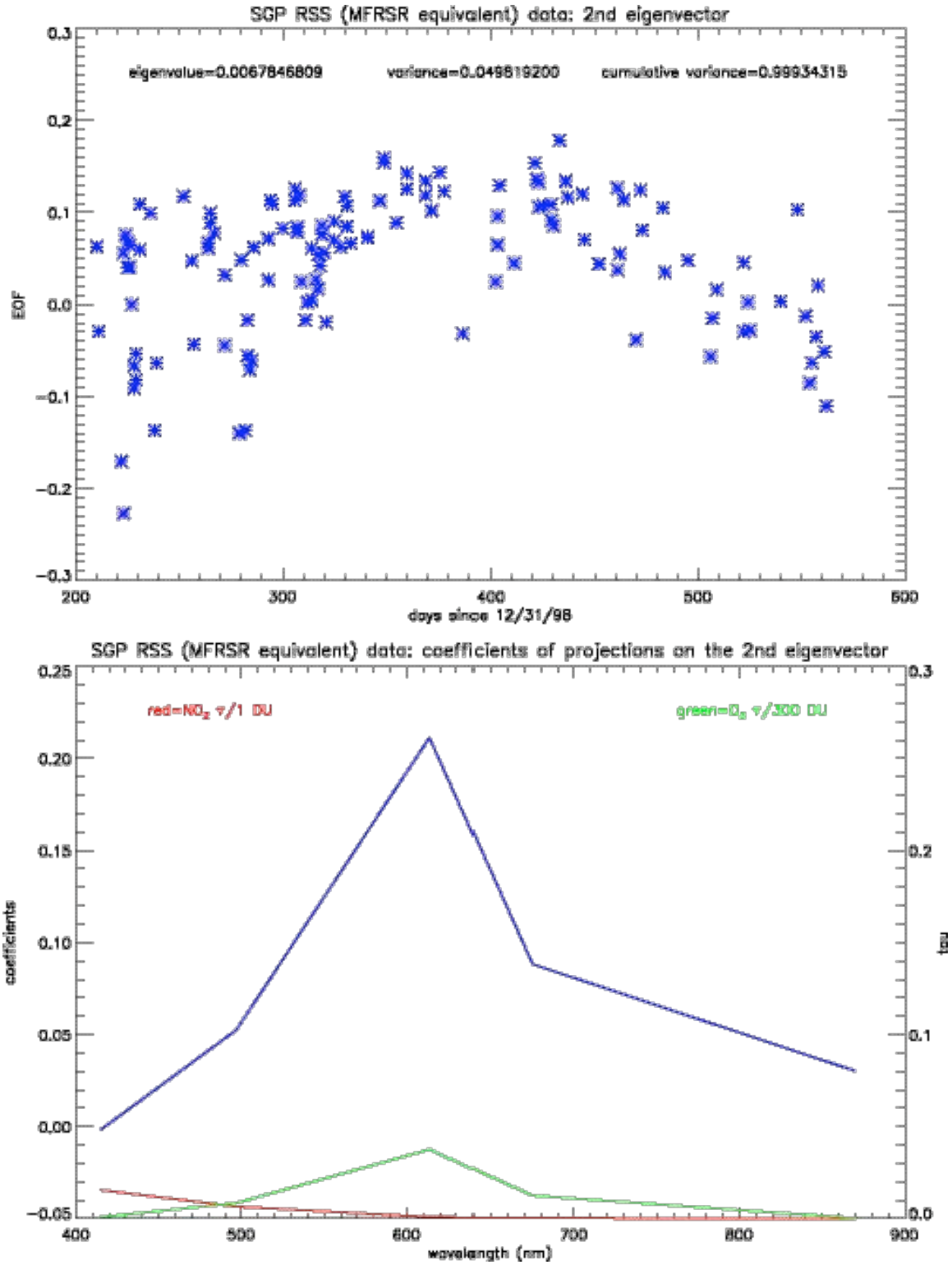

Figure 3.18: The second EOF for the "MFRSR equivalent" data set. Like Figure 3.2, this EOF is clearly dominated by ozone, but the baseline here is much smaller.

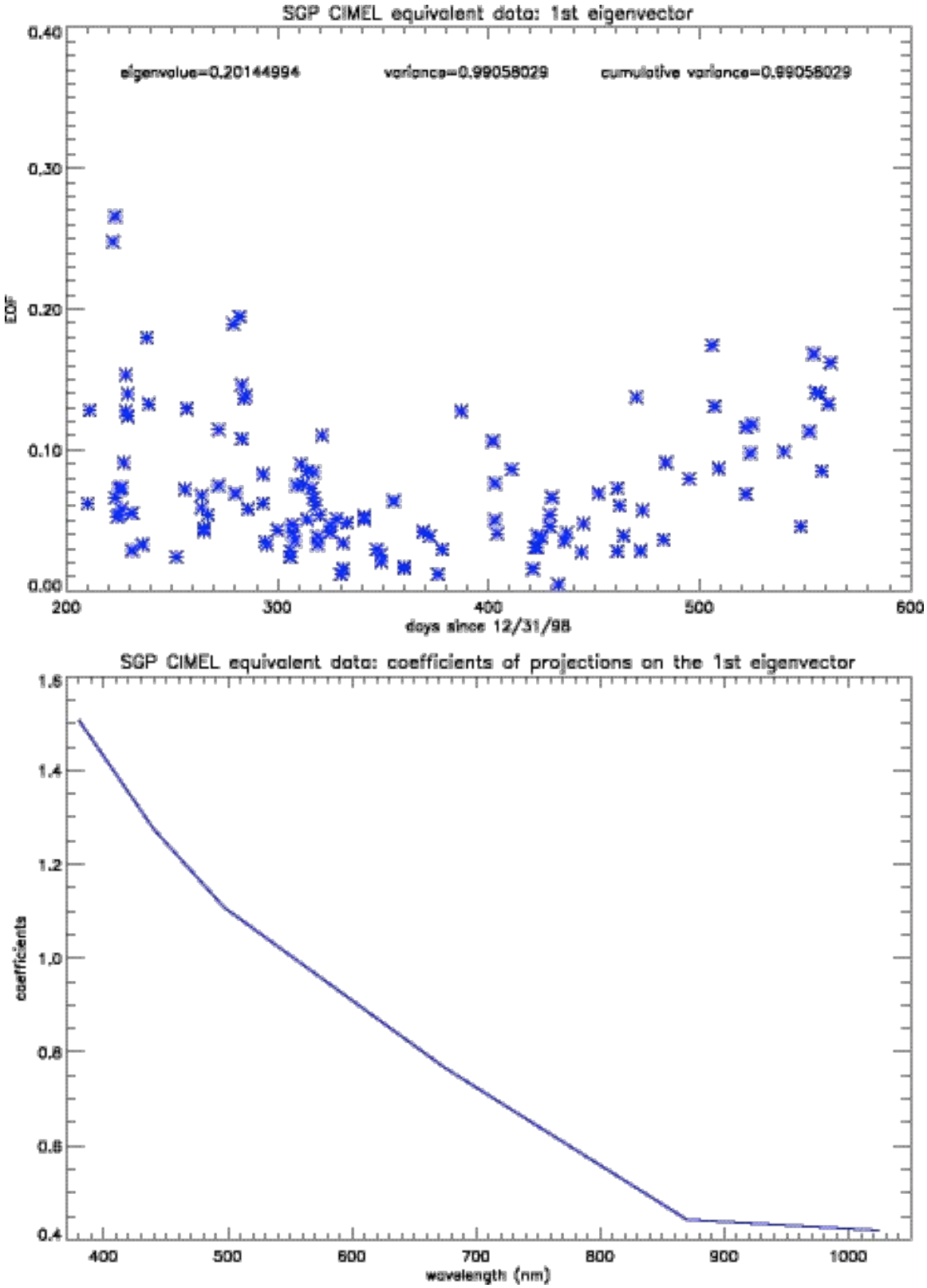

Figure 3.19: The first eigenvector of the "CIMEL equivalent" data set.

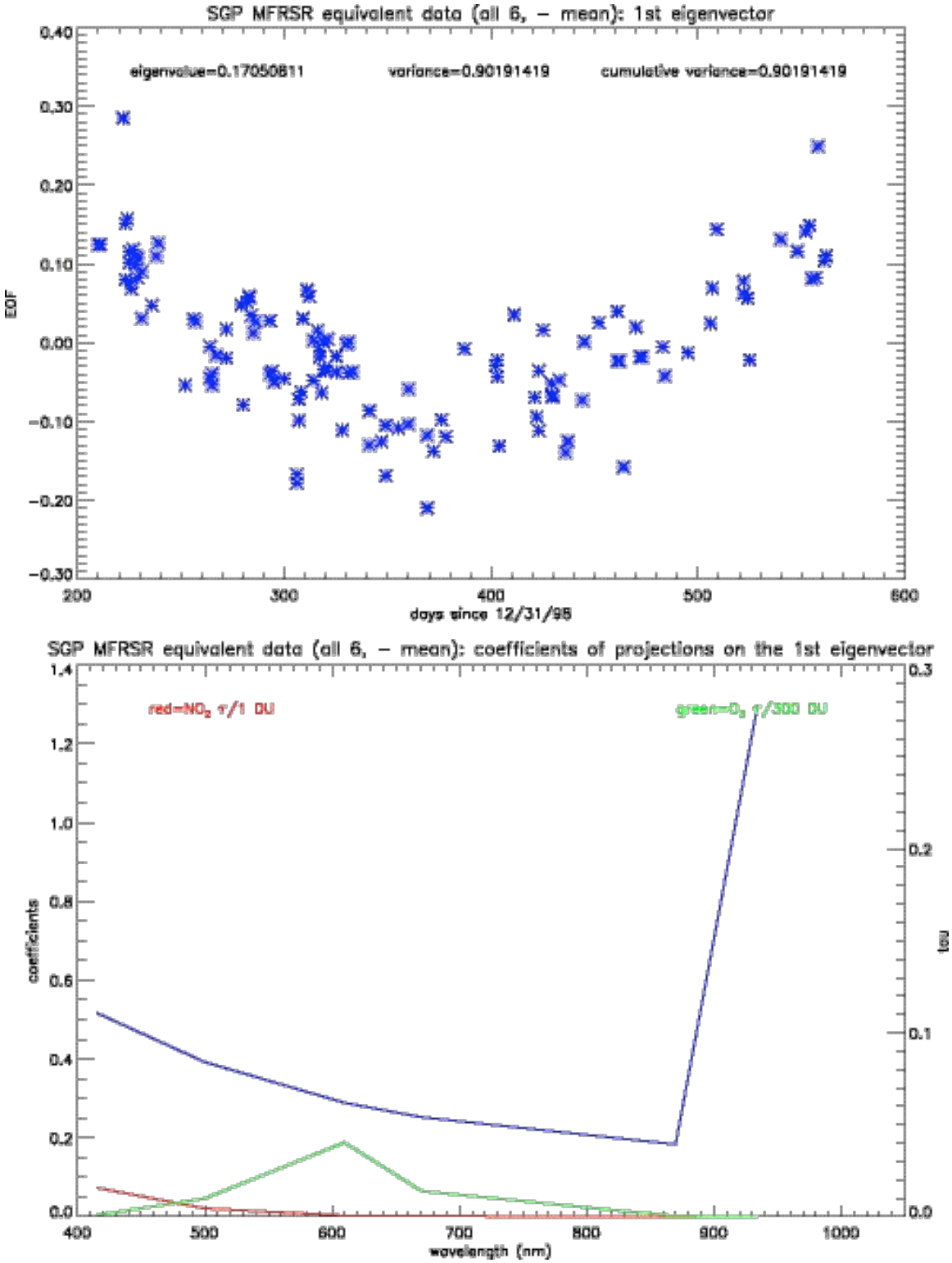

Figure 3.20: The first EOF for the "MFRSR equivalent" data set, including a water vapor absorption channel, with the annual mean subtracted.

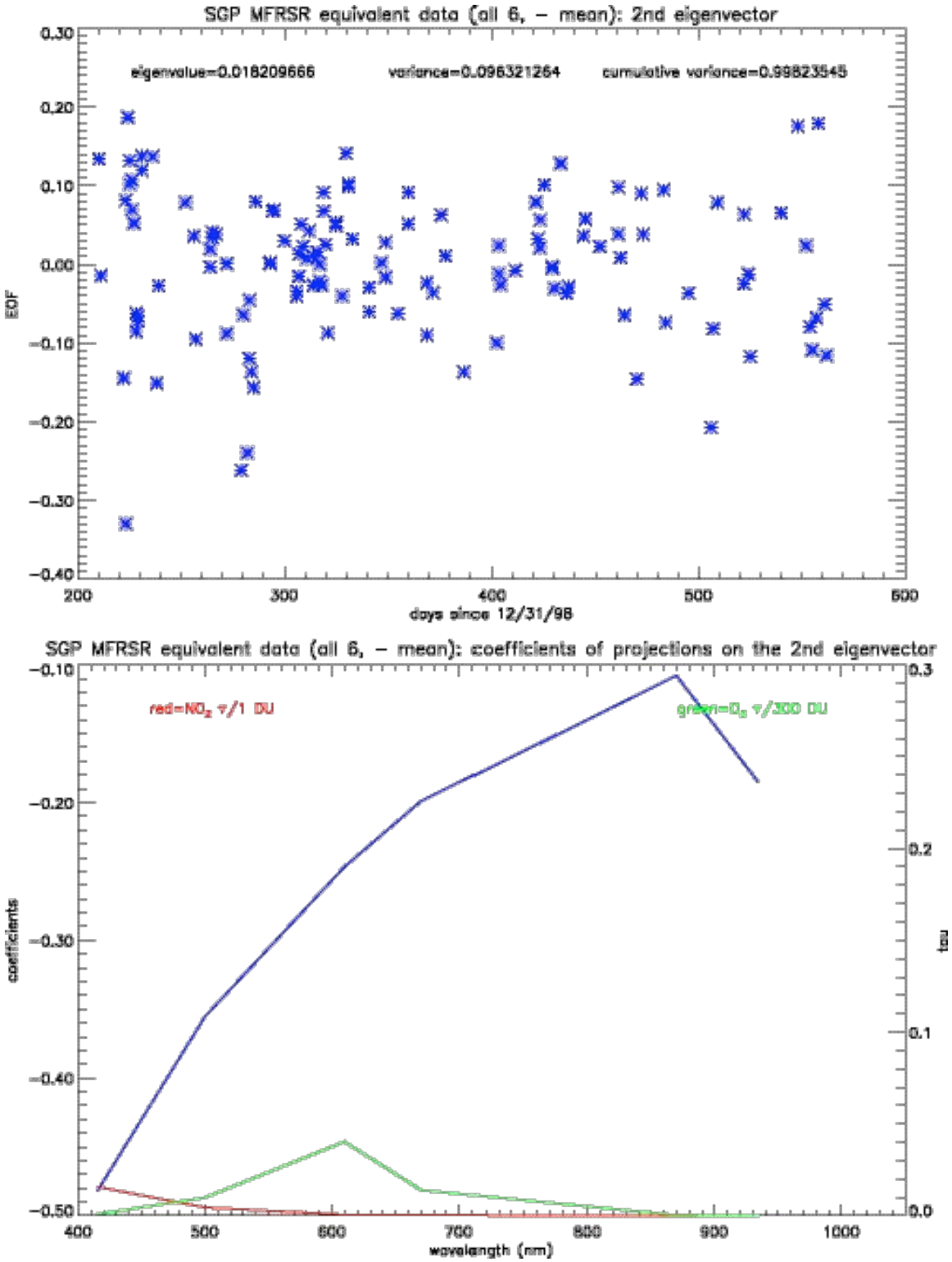

Figure 3.21: The second EOF for the "MFRSR equivalent" data set, including a water vapor absorption channel, with the annual mean subtracted.

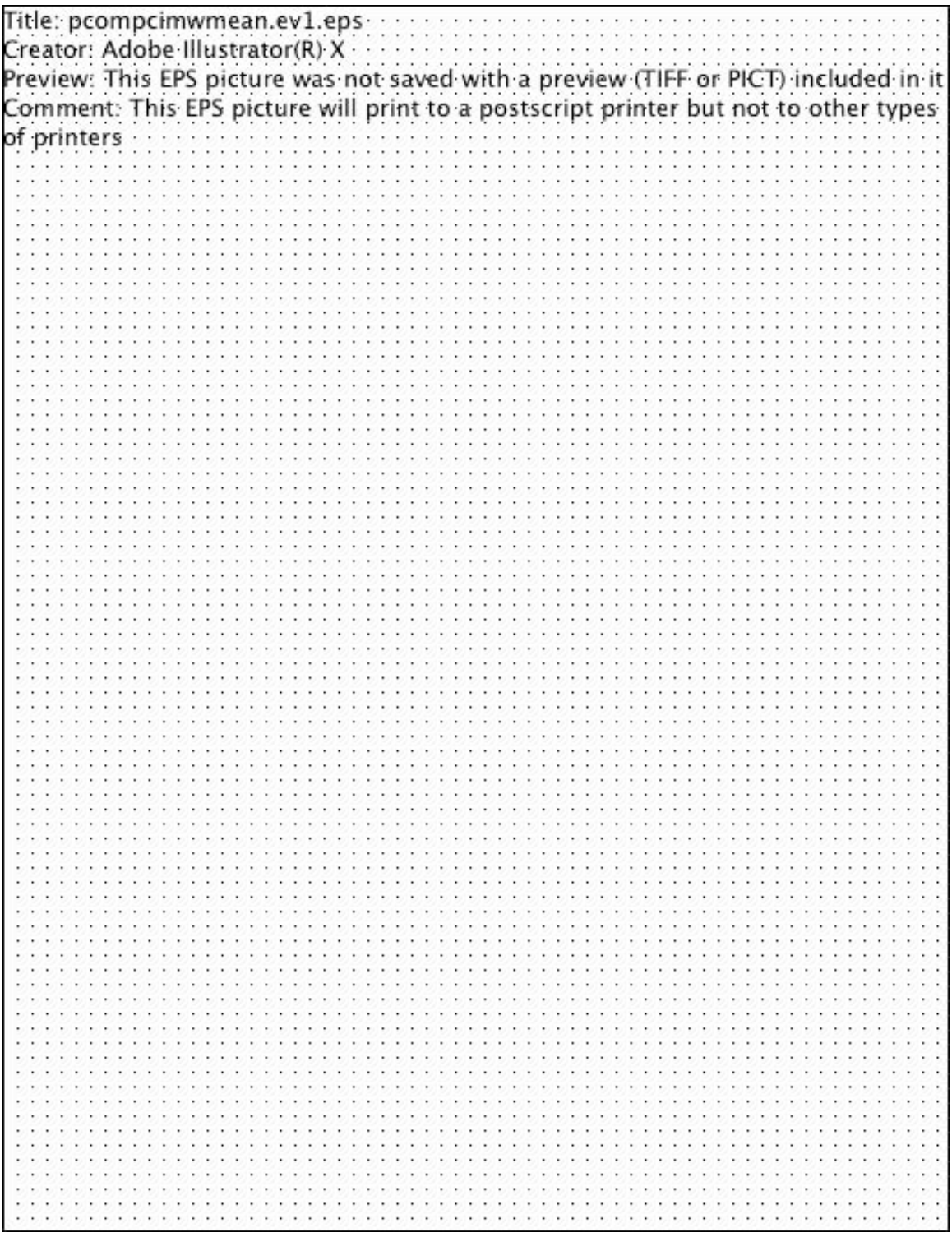

Figure 3.22: The first EOF for the "CIMEL equivalent" data set, including a water vapor absorption channel, with the annual mean subtracted.

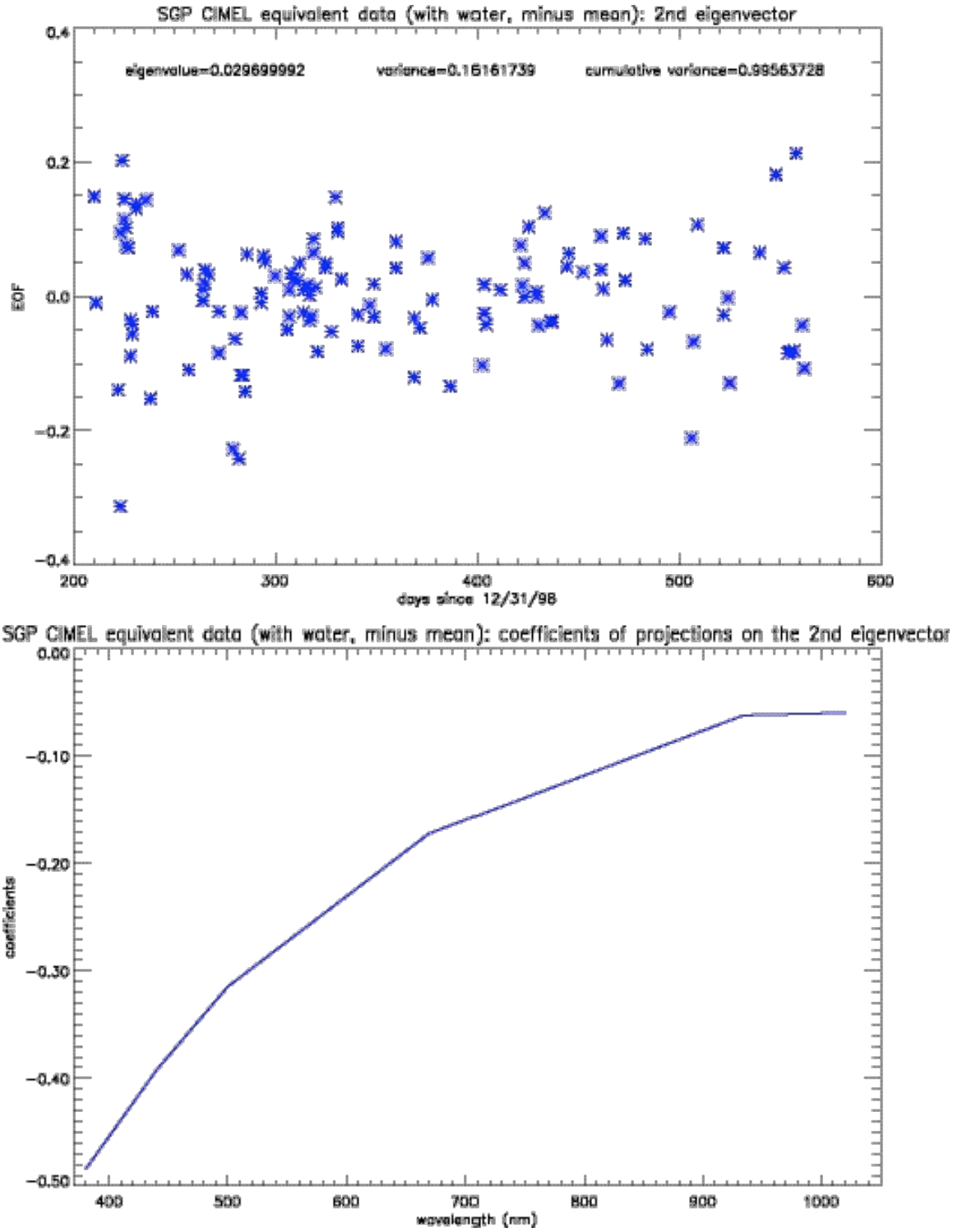

Figure 3.23: The second EOF for the "CIMEL equivalent" data set, including a water vapor absorption channel, with the annual mean subtracted.

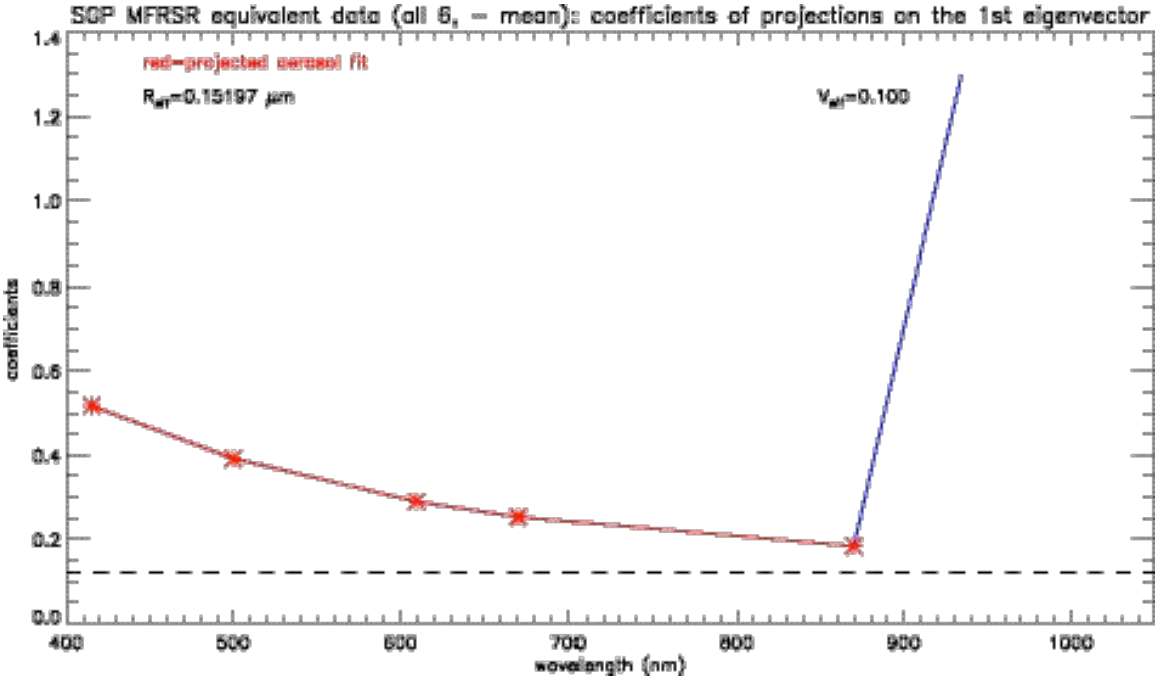

Figure 3.24: The best Mie fit to the first EOF of the "MFRSR equivalent" data, including the water vapor channel, with the mean subtracted.

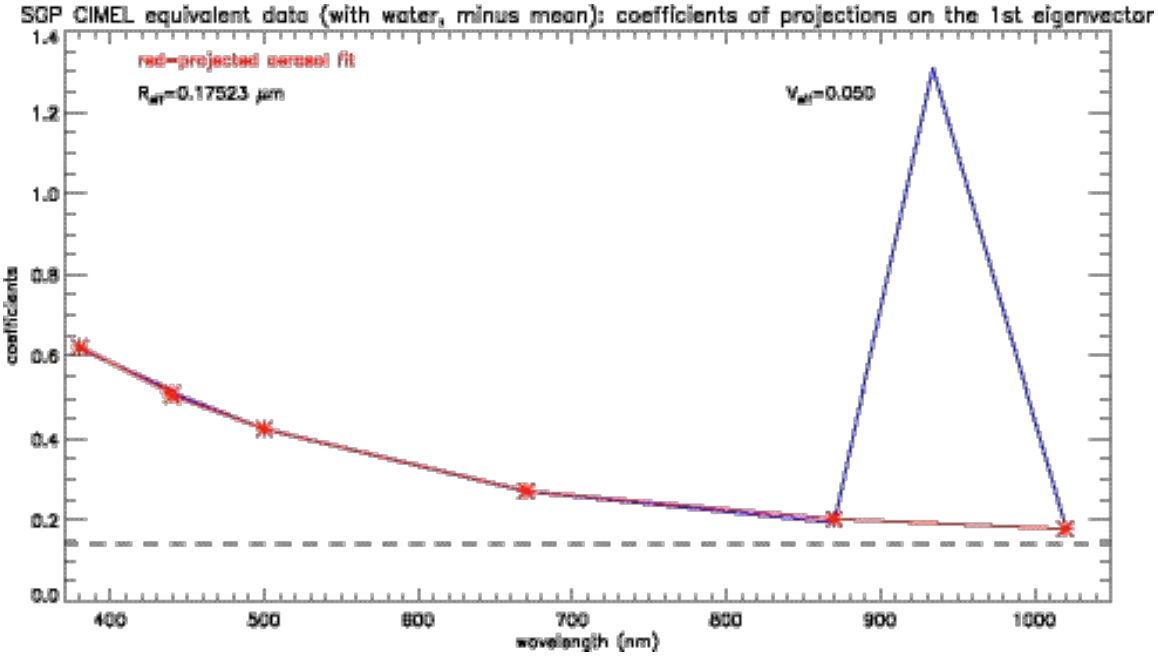

Figure 3.25: The best Mie fit to the first EOF of the "CIMEL equivalent" data, including the water vapor channel, with the mean subtracted.

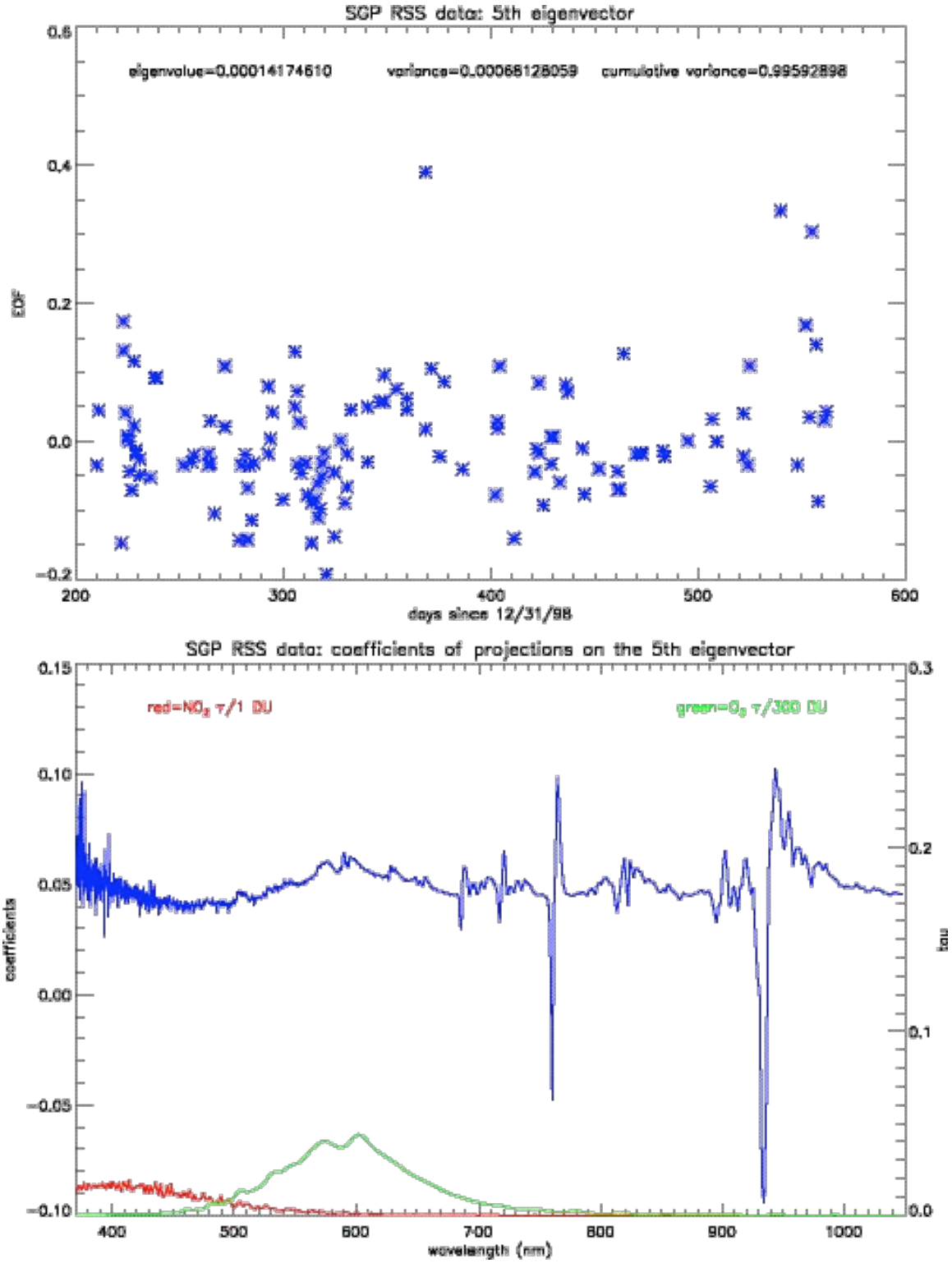

Figure 3.26: The fifth eigenvector of the full RSS data set. The pattern in the coefficients of projection, which strongly resembles the first derivatives of the H_2O and O_2 absorption coefficients, indicates that the RSS channels were affected by wavelength shifts over the duration of the data set.

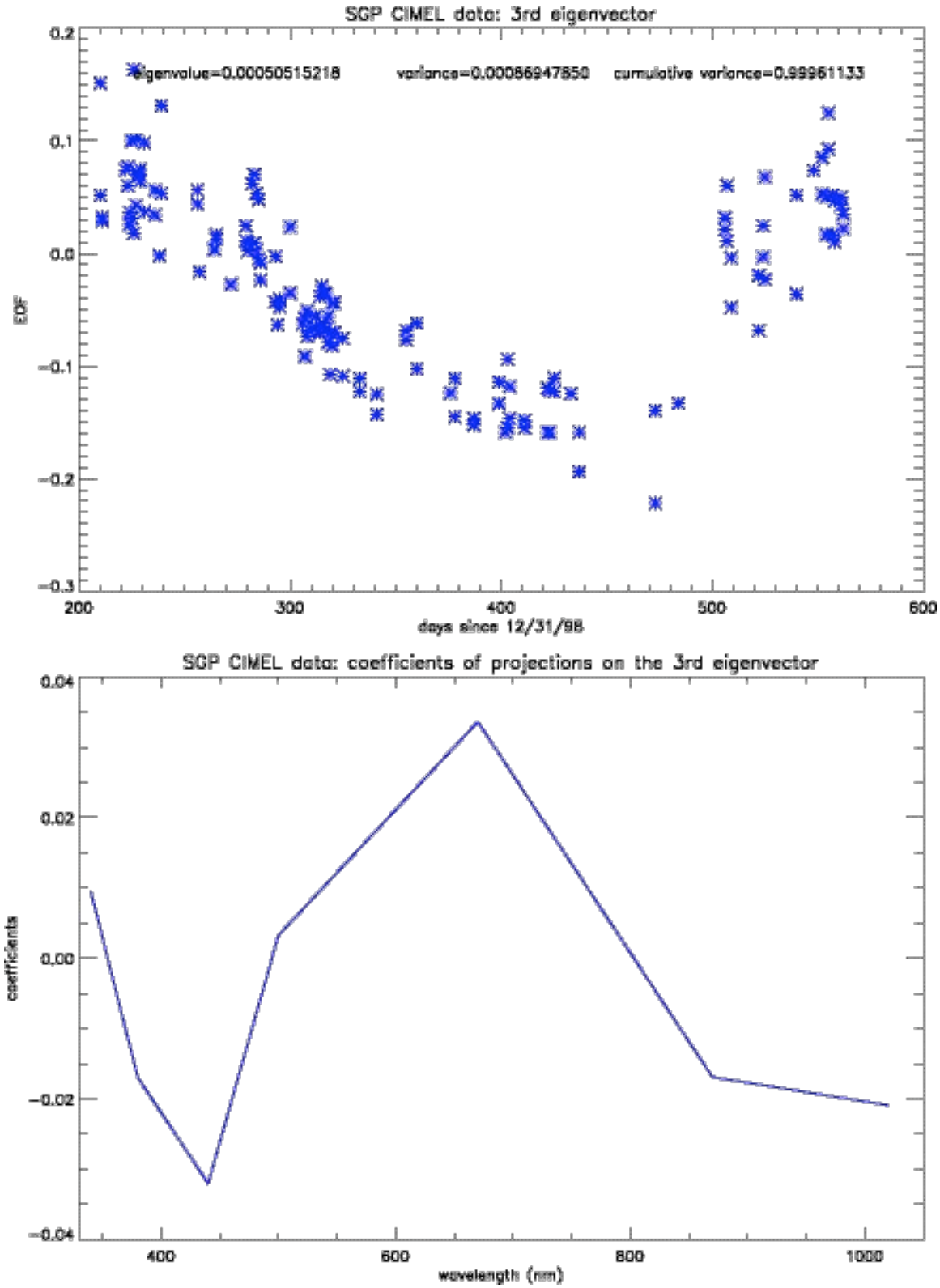

Figure 3.27: The third EOF in the CIMEL data set. A strange, but very pronounced, pattern emerges in this EOF. The coefficients suggest that it is related to the strong dips in the optical depth at 670 nm.

Chapter 3 References

Alexandrov, M. D., Lacis, A. A., Carlson, B. E., and Cairns, B., *Journal of the Atmospheric Sciences* **59**, 524-543 (2002a).

Box, G. P., Box, M. A., and Krücker, J., *Journal of Geophysical Research* **101**, 19211-19214 (1996).

Charlson, R. J., Schwartz, S. E., Hales, J. M., Cess, R. D., Coakely, J. A., Hansen, J. E., and Hofmann, D. J., *Science* **255**, 423-430 (1992).

Dubovik, O., and King, M. D., *Journal of Geophysical Research* **105**, 20673-20696 (2000).

Harrison, L., Kiedron, P., Berndt, J., and Schlemmer, J., *Journal of Geophysical Research* **108 (D14)**, 4424 (2003).

Chapter 4: The Results of the Gas and Single-Mode Aerosol Retrievals

4.1: Introduction

The results of the EOF analysis of Chapter 3 suggest that the aerosol size distribution is bimodal. If true, then the retrievals that assume a single mode aerosol size distribution will produce potentially misleading gas amounts. Any incorrect assumption about the spectral dependence of aerosol extinction will cause either aerosol extinction to be mistakenly attributed to absorption by one of the gases, or vice versa. When this happens, unnatural-looking correlations between gas and aerosol amounts, beyond what might be expected even from biomass burning and industrial pollutants, appear in the data. The higher spectral resolution of the RSS data allows for a more complete separation of gas and aerosol extinction, and provides a means to test the accuracy of the gas amounts measured using five and sixteen-channel RSS retrievals. In the case of nitrogen dioxide, any significant overestimation or underestimation of the correct amount immediately becomes obvious. Verifying the accuracy of ozone retrievals is a bit trickier, because there is less detail in the ozone absorption spectrum from which to obtain a clear and unambiguous separation of gas and aerosol extinction. Still, the EOF analysis also indicates that the large optical depth difference between the MFRSR and RSS could cause misleadingly large differences in the respective ozone retrievals, and the observation of such differences would provide further evidence that the underlying assumptions in the retrieval schemes need to be amended.

If the gases are not being retrieved accurately, this will affect the quality of the aerosol retrievals as well. The assumption of a single-mode aerosol size distribution, when the aerosols being observed actually have a bimodal distribution, could conceivably be responsible for the relatively large errors in the retrieved column amounts and the exaggerated correlation between gas and aerosol amounts. If this is indeed true, three things should be observable in the aerosol results obtained using single-mode distributions. First, the retrievals that try to solve for the effective variance will usually, if not almost always, retrieve the highest possible variance value, because broader single-mode distributions more closely approximate the spectral variation of bimodal distributions than narrower ones do. Second, the residual optical depth values will increase significantly when the gas amounts are constrained, even when the constraints are accurate, because the single-mode retrievals can no longer compensate for this incorrect aerosol size distribution assumption by wrongly attributing the aerosol extinction to the gases. Finally, the bulk of the positive residual optical depth, where the total optical depth measured exceeds the total optical depth obtained by adding together the calculated optical depths of each retrieved quantity, will be located at either end of the wavelength range, while the negative residuals will be located at the central wavelengths. This happens because the retrieved single-mode effective radius is larger than the fine mode effective radius for that day, causing an underestimation of aerosol extinction at long wavelengths, and lower than the coarse mode effective radius, causing a similar underestimation at high wavelengths. If these three characteristics are all observed in the RSS and CIMEL data, the bimodality of the aerosol size distribution would be unequivocally confirmed.

Table 4.1 shows the mean values for all the quantities measured by the various single-mode retrievals over the 54 mornings or afternoons that all three devices have in common. The highest variance used in the MFRSR retrieval algorithm is 0.4 (Alexandrov *et al*. 2002a). Therefore, for the sake of comparison, the largest variance considered for the "MFRSR equivalent" retrieval is 0.4. For the remaining RSS retrievals and the CIMEL retrievals, the effective variance can be as large as 0.5. Instead of the mean effective variance, the variance value retrieved the most often is presented, along with the percentage of times it is retrieved.

4.2: The Nitrogen Dioxide Retrieval Results

4.2.1: The Presence of the Ring Effect in the RSS Data

Before examining in detail the results of the high-resolution NO_2 retrieval using the RSS data, it is necessary to investigate how the Ring effect alters this data. Vountas *et al*. (1998) developed the rotational Raman scattering portion of the SCIATRAN radiative transfer model, with the primary purpose of correcting NO_2 measurements in DOAS-based retrievals that the Ring effect does interfere with. The Ring effect has two primary components; the first is the filling in due to the Fraunhofer absorption lines, and the second is the filling in of gaseous absorption lines. Vountas *et al*. used EOF analysis to separate these components on theoretical and experimental data, and such a procedure can be easily performed on a full day of RSS intensity data as well. The date chosen for this was March 20, 2000, on which the high-resolution NO_2 retrieval for the RSS showed

the highest column amounts. Both the total and the direct intensities are examined. The relative magnitude of the Ring effect as a function of wavelength is calculated based on the procedure used by Vountas *et al*. First, the wavelengths between 410 and 450 nm were isolated, because this is the wavelength range where NO_2 absorption is strongest, and where the Ring effect can interfere with retrievals. In addition, data points with airmass values greater than five were excluded. For each channel, the measured intensity values (I) are divided by the standard value of the incoming, top of atmosphere solar flux (I_0) at that wavelength. Without the Ring effect or NO_2 absorption, this would produce a smooth curve as a function of wavelength. To approximate this smooth curve, a third-order polynomial fit with wavelength is applied to the values of I/I_0. The Ring effect magnitude for each wavelength is then calculated as the logarithm of the ratio of the measured I/I_0 vs. the value of the fitted curve at that wavelength. This means that wavelengths where the measured intensity is greater than the fitted value, where "filling in" has taken place, will have positive values for the magnitude of the Ring effect, while those wavelengths with intensity values below the fitted curve will have negative Ring effect values. Once these values are calculated, the empirical orthogonal functions are calculated. The eigenvectors are plotted as functions of time, instead of day, while the coefficients of the projections remain functions of wavelength.

Figures 4.1, 4.2, and 4.3 show the first three eigenvectors, respectively, for the Ring effect of the total intensity for March 20. The coefficients of the first eigenvector clearly show the filling in of the Fraunhofer lines, with significant negative values at adjacent wavelengths. While the eigenvector itself shows the large degree of random noise that affected the optical depth data, it also shows a time dependence on solar zenith

angle, as Vountas *et al*. had postulated. Furthermore, the first EOF accounts for over 99% of the total variance. Unlike in the Vountas *et al*. paper, though, the second EOF doesn't reflect gaseous absorption. An abrupt shift occurs in this eigenvector, in the neighborhood of 9:30 a.m. The abruptness of the shift suggests that some sort of subtle, and probably artificial, adjustment or alteration affected the output of the device. In fact, instrument monitor checks do occur on a regular basis, and could account for the observed shift. The other full day for which this EOF analysis was performed, January 4, shows a similarly abrupt shift in the second eigenvector, occurring just before noon. However, an adjustment would not necessarily account for the downward trend in the rest of the EOF. The gas absorption emerges strongly in the coefficients of the third eigenvector. The filling in of gaseous absorption lines is very small, relative to the filling in of the Fraunhofer lines, and would be much less noticeable in the total Ring effect spectrum without the EOF analysis.

The SCIATRAN model treats all Raman-influenced light as diffuse radiation, but since the Ring effect emerged in the optical depth data, the RSS must have measured at least some of this light as part of the direct beam. Most likely, this portion of the direct beam consisted of whatever light that was shifted in wavelength, but not direction. For the sake of comparison, the Ring effect was analyzed in the direct beam for the same three days as in the total intensity. The results for first and third eigenvectors of the direct beam are shown in Figures 4.4 and 4.5, respectively. The coefficients do not differ dramatically from those for the total intensity. However, the solar angle dependence is not quite as pronounced in the direct beam as it is in the total.

Figure 4.6 is the graph in which the Ring effect was first noticed in the RSS data. It shows an odd feature that, at face value, seems to contradict the theoretical predictions. Average RSS optical depth is plotted vs. wavelength for both the thirty highest humidity days, and the thirty lowest humidity days. The high humidity days also happen to be high aerosol optical depth days, because aerosol amounts, like column water vapor amounts, peak in the summer. The modeling of Vountas *et al*. claims that the addition of aerosols will slightly dilute the Ring effect, but in the optical depth data, the Ring effect shows up much more strongly in the high aerosol days than in the low aerosol days. The reason for this lies in the fact that the magnitude of the Ring effect varies with solar zenith angle, or with airmass. The optical depth values are obtained after the data is calibrated using Langley regressions, which plot the logarithm of intensity for a given wavelength vs. airmass. As the amount of relative intensity filled in or scattered out of a particular wavelength changes with the airmass, this will affect the results of the Langley plot significantly, and produce errors in the retrieved I_0 and optical depth. Figures 4.7 and 4.8 show how these errors affect the Langley regressions for a low optical depth day and a high optical depth day, respectively, for the 431.8 nm channel. The magnitude of the Ring effect for this channel can be calculated for the entire day, and a corrected intensity can be calculated from this. In the case of the low aerosol day, December 26, the Ring effect causes a significant overestimation in the measured data, by nearly 25%, of the incoming solar flux. However, the corresponding slopes are virtually parallel, and the difference in the Langley-retrieved optical depth values is only about 0.0004, with the corrected optical depth being slightly higher. This means that the Langley-calibrated optical depth data will not be significantly altered from the actual optical depth, despite

the large error in I_0. On August 10, the greater optical depth causes a slightly smaller percentage error in I_0, but this lower I_0, in turn, lowers the calculated optical depth enough for the difference to be discernible. The alteration in the optical depth for the high aerosol day is 0.0017, large enough to produce the pattern seen in Figure 4.8. This means that in Langley regressions, the Ring effect influences the calibration significantly more than it does the calculated optical depth. An investigation of intensity measurements for a high-resolution device like the RSS will therefore reveal the Ring effect immediately, but an investigation of Langley-derived optical depth values will only reveal the Ring effect on high optical depth days, and even then only at the location of the Fraunhofer lines will its presence be readily discernible.

The alteration in the optical depth caused by the Ring effect is highest at 431.8 nm, due to the strong Fraunhofer line present in that part of the spectrum. Most of the other RSS channels between 410 and 450 nm, including the locations of the strongest NO_2 absorption lines, do not have nearly as great a discrepancy between actual and measured optical depth values. Since the Langley-related errors caused by the Ring effect wind up altering the direct-beam measurements to a much greater degree than they do the Langley-derived optical depth values, it is not clear that the high-resolution NO_2 retrievals developed for this study are compromised to a significant degree, the way that the Ring effect causes underestimations of NO_2 amounts in DOAS measurements using the direct intensity. Still, the EOF analysis indicates that the Ring effect is readily correctable in the measured intensities.

4.2.2: The High-Resolution RSS Retrieval

Because of its utilization of the full spectral resolution of the RSS, which enables the signal of nitrogen dioxide absorption to emerge distinctly from the aerosol background, the NO_2 retrieval algorithm developed specifically for this study and explained in detail in Chapter 2 is considered to produce more accurate results than any low-resolution retrieval which attempts to retrieve aerosol and NO_2 amounts simultaneously. Therefore, the results from the high-resolution retrieval are presented first, and the subsequent results of the other NO_2 are discussed in the context of these values. Figure 4.9 shows the plot of NO_2 vs. day for the RSS high-resolution retrieval. The error bars measure two standard deviations of random noise affecting the data on a particular morning or afternoon. These NO_2 results have three distinct features. First, the values are consistently small, never once exceeding 1 DU. Second, there does not appear to be any obvious seasonal cycle in the data. Third, as Figure 4.10 shows, there is no correlation between NO_2 amounts and aerosol optical depth. All three of these features contradict the findings of Alexandrov *et al*. (2002b), because the Alexandrov *et al*. algorithm does not adequately separate the extinctions due to small particle aerosols and NO_2 in the 415 nm channel. A significant portion of the aerosol extinction is mistakenly attributed to nitrogen dioxide instead. This results not only in excessive values of NO_2, but in similarly enlarged values of the aerosol effective radius. (In addition, the retrieved NO_2 values reported in Alexandrov *et al*. (2002b) follow the same seasonal pattern as the aerosols, and the correlations between NO_2 and aerosols are consequently strong.)

This result affects the expectations of what will be seen in the results from the

algorithms that retrieve aerosol and gas amounts simultaneously. Obviously, the MFRSR results will show the same erroneous patterns reported in Alexandrov *et al*. (2002b), but it is not clear if the RSS results will show these patterns as well, or to what degree this will change between the sixteen-channel retrieval and the five-channel "MFRSR equivalent" retrieval. If the RSS retrievals show the same errors, then these errors clearly do not result from anything specific to the MFRSR as a device or to the Alexandrov (2002a) algorithm. Instead, the error arises from the assumption, common to all the retrievals used, that there is enough information in the wavelengths used in the retrievals to constrain the aerosol properties completely. If the aerosols are bimodal, then this assumption simply is not true. Ultimately, the ability to separate gas and aerosol extinctions as accurately as possible, given a limit to the amount of obtainable aerosol information, depends on the accuracy of the assumptions that need to be made in order to constrain the aerosols. NO_2 is particularly sensitive to these assumptions in coarse-resolution retrievals, because the absorption is relatively weak. Ozone absorption is much stronger by contrast, so while ozone retrievals are still quite assumption-dependent even at high resolution, the percentage error in calculated concentrations is much smaller for ozone than for NO_2.

4.2.3: The MFRSR Results

Retrievals using the MFRSR data were made using effective variance values of 0.01, 0.1, 0.2, 0.3, and 0.4. It was believed that the true, single-mode effective variance fell within this range, and that the actual gas amounts would fall within the range of

values calculated in these retrievals. The large trade-off between small aerosols and nitrogen dioxide at the 415 nm channel had already been documented by Alexandrov *et al*. (2002a), and NO_2 amounts calculated using the retrieval algorithm had a very large margin of error. Because of the nature of the Alexandrov *et al*. algorithm, the calibration values of the MFRSR channels are adjusted for each variance, which eliminates any residual optical depth that might otherwise appear in the data. Therefore, no variance value is selected as producing the best fit, or given any preference over the others. Mean values of the aerosol optical depth, effective radius, column ozone, and column nitrogen dioxide are calculated for each variance during each clear morning or afternoon. As a general rule, the retrieved values for each quantity vary monotonically with variance, either steadily increasing or steadily decreasing. Therefore, this study will primarily display the results for the retrievals assuming the extreme variance values, 0.01 and 0.4.

Nitrogen dioxide produced the biggest uncertainty in the Alexandrov *et al*. (2002a) algorithm. The results from that study show NO_2 values that vary by roughly a factor of two between low variance and high variance. Similar results were obtained here, with the mean NO_2 value for the entire data set being 5.02 DU for the low variance and 2.59 DU for the high variance. There are two problems with this, however. First, even the values obtained with the high variance are substantially higher than anything obtained by the high-resolution RSS retrievals. Second, the correlation between the retrieved values of NO_2 and aerosol optical depth confirms the suspicions that much of what is being retrieved as nitrogen dioxide is, in fact, additional aerosol extinction. The plots of the retrieved NO_2 values vs. day is shown in Figure 4.11, both for the low (blue)

and high (red) variance. Both bear a strong resemblance to the plot of 870 nm optical depth vs. day in Figure 2.11.

When the NO_2 values are plotted against retrieved aerosol optical depth values, as shown in Figure 4.12, the correlation is shown to be particularly strong for the lowest variance, but it significantly decreases when the variance is increased. It is clear though, from the high-resolution retrieval, that both correlations are artificially high. A bimodal aerosol size distribution could account for the lowered correlation, because the higher variance, while still inaccurate, better represents reality than the lower variance, leading to lower NO_2 values with less correlation with aerosol optical depth.

4.2.4: The RSS Results from the Coupled Gas/Aerosol Retrievals

Figures 4.13 and 4.14 show the column NO_2 values plotted vs. day for the five-channel "MFRSR equivalent" retrieval and for the sixteen-channel full RSS retrieval assuming a single-mode aerosol size distribution. Again, the retrieved NO_2 values follow a similar pattern to the measured optical depth values. The mean values retrieved for NO_2 using RSS data are 1.17 DU for the "MFRSR equivalent" retrieval, and 1.60 DU for the sixteen-channel retrieval. Both values are much lower than the MFRSR values, although still higher than the values from the high-resolution retrieval. Likewise, the correlations between NO_2 amounts and aerosol optical depth remain fairly high, with the correlation in the five-channel retrieval being slightly higher. The differences in measured total optical depth can easily explain the lower retrieved NO_2 values. As the same error seen in the MFRSR is seen in the RSS retrievals as well, it is clear that the

error arises from a common, flawed assumption about the aerosol properties inherent in both sets of retrievals.

To confirm the overestimation of the NO_2 values obtained when the gases and aerosols are retrieved simultaneously assuming a single-mode size distribution, these retrievals are examined in the full spectral detail that the RSS provides. Figure 4.15 shows the single-mode RSS retrievals projected over the full spectrum for the morning of June 23, 2000, when the aerosol optical depth was particularly large, but the change in aerosol extinction with wavelength was relatively small. For both single-mode retrievals, but especially for the sixteen-channel one, the retrieved value of NO_2 greatly overestimates the variation in optical depth that is actually observed in the spectral region where NO_2 absorbs, establishing unambiguously that some aerosol extinction is indeed being incorrectly attributed to NO_2 in the retrievals. Furthermore, the difference in the retrieved values for ozone for this day is simply alarming. This means that at least one retrieval method, if not the both of them, fails to properly separate the extinctions due to ozone and aerosol as well.

4.3: The Ozone Retrievals

4.3.1: The MFRSR and RSS Retrieval Results

The retrieved ozone values for the low-variance and high-variance retrievals of the MFRSR data are plotted vs. day in Figures 4.16 and 4.17. The mean values (342 DU for the low variance, 329 DU for the high variance) are fairly large on the whole, just like

the retrieved values for nitrogen dioxide. The large variance, however, gives a few overly low ozone values shortly before and shortly after the gap in the data during November 1999 (days 275 to 304), while the small variance has a few anomalous values early in the data set. Possibly, there are some calibration issues for these particular days, or some other device-related problem. Despite these differences between the two graphs, a few qualitative similarities do exist. Both graphs show their lowest values on either side of November 1999, and both have higher ozone values on the whole in 2000 than in 1999.

Just like the NO_2 retrievals, the ozone retrievals show considerable differences in the results between the MFRSR and the RSS. The mean "MFRSR equivalent" retrieved ozone, 248 DU, is more than 80 DU smaller than even the mean for the high-variance MFRSR retrieval. Obviously, this discrepancy is unacceptably large, although somewhat predictable given the optical depth difference between the two devices at 610 and 670 nm. With the sixteen-channel retrieval, the additional information in the Chappuis band leads to a larger, and presumably more realistic, mean value of 265 DU. However, this mean value is still far below the values measured by the MFRSR. The mean value of the high-resolution ozone values that came with the data set is 278 DU, which only slightly lessens the discrepancy between the two devices.

The plots of the three sets of RSS-retrieved ozone values vs. day, shown in Figure 4.18, show very good qualitative agreement with each other. They all indicate that a sharp dip in the ozone values took place during November 1999 – regrettably, the very period of time when the MFRSR was inactive. They also all show a quick resurgence of ozone values in December (days 335 to 365); in two of the three cases, the highest values

in the data set occur during this month. Like the MFRSR, the ozone values on the whole are higher in 2000 than they were in 1999. Curiously, the sixteen-channel retrieval shows quite a bit less stability in its retrieved ozone values than the five-channel "MFRSR equivalent" retrieval does.

Ultimately though, the quantitative differences between the three retrievals using identical data indicate that the retrieval output, even for a fairly strong absorber like ozone, depends a lot on the specific assumptions inherent in the retrieval. In particular, if the assumptions about the spectral behavior of the aerosols are inaccurate, this will cause a trade-off of optical depth in the retrieval between aerosols and ozone at the wavelengths where ozone absorbs. The sixteen-channel retrieval may have the greatest variability in ozone amounts because the wavelength range used in the retrieval was the widest. This would cause the greatest variability in retrieved aerosol extinction at the center of the wavelength range, which also happens to be the heart of the Chappuis band. This underscores the importance of making the most accurate assumptions possible about aerosol behavior, especially if the location of a particular sun photometer does not have any independent means of measuring gas amounts.

4.3.2: Using EOF's to Objectively Compare Ozone Data

EOF's have been used in the analysis of MFRSR data before, to retrieve not only the optical depth of aerosols at the five different wavelengths, but also column ozone amounts (Taha and Box 1999). The first three eigenvectors in the MFRSR and "MFRSR equivalent" EOF's have clearly discernible contributions from ozone, and any accurate

EOF-based retrieval of ozone needs to incorporate all three EOF's. The EOF-based approach does have a couple of advantages. First, the ozone retrievals for the entire data set can be done simultaneously, and therefore very rapidly, by this method. Second, the similarities in the coefficients of the eigenvectors enable a more objective comparison of the ozone amounts measured by the two devices than their differing retrieval strategies can provide. For example, Figures 4.19 and 4.20 show the coefficients of the second eigenvectors for the MFRSR and the "MFRSR equivalent" sets, respectively. In each case, the baseline curve corresponding to the non-ozone contributors is shown as well. The values of the coefficients are markedly different from each other, being much larger for the MFRSR. The significant difference in the measured optical depth values between the two devices manifests itself in the other eigenvectors as well, but nowhere more so than here. The difference does not seem to have a whole lot to do with the ozone, however. In fact, the EOF analysis fails to show a huge difference in the ozone extinction measured by the two devices, and the very large differences obtained in the retrievals probably result from some of the discrepancy in total optical depth being mistakenly attributed to ozone, due either to calibration differences or flawed assumptions inherent in the retrieval schemes.

Unfortunately, any attempt to retrieve ozone amounts requires making assumptions about the nature of the aerosol, and it is important to remember that these assumptions are at best approximations. A polynomial fit of ln τ vs. ln λ to evaluate the aerosol extinction was applied here for the first eigenvectors of the MFRSR and "MFRSR equivalent" data sets. For the second eigenvector of the MFRSR equivalent data set, because one of the coefficients there is negative, the coefficients instead were

treated as the sum of a polynomial function of the wavelength plus a linear function of the ozone absorption coefficients. In the case of the third eigenvector, the only viable approximation was a linear fit involving the three channels with the highest wavelength. This approximation is crude at best, but the contribution of the third EOF to the total ozone amount is small.

The retrieved column ozone amounts, plotted as functions of day, are shown for the MFRSR in Figure 4.21. The EOF-retrieved ozone values for the MFRSR show some substantial differences with the values presented in Figure 4.17. Both graphs show some extreme variability in October and December of 1999, although not always on the same days. Since two different retrieval methods show the same unphysical variability over the same stretch of time for the same data, the data itself is the likely source of the problem, not either retrieval. In addition, the ozone values retrieved using the EOF's are clearly and significantly lower than those from the Alexandrov *et al*. (2002a) retrieval. Figures 4.17 and 4.21 do show some qualitative agreement, though, in that the values in the year 2000 seem to be higher in general than those for 1999. Both graphs also show a relative high value at day 506, or May 20, 2000.

The RSS ozone values in Figure 4.22 are very high on a couple of days, specifically the afternoon of August 10 and the morning of August 11, 1999, and May 20, 2000. Neither spike in Figure 4.22, however, corresponds with anything in the RSS ozone retrievals in Figure 4.18, even though the graphs otherwise show good qualitative similarity to each other. As before, the RSS values show more stability than the MFRSR values do. Like the MFRSR data, the RSS ozone values also show a significant increase

at the end of 1999 that carries well into 2000, although the RSS values are generally lower than those of the MFRSR at these days.

The mean of the EOF-retrieved MFRSR ozone values, 281 DU, is still higher than the mean of any of the RSS ozone retrievals. However, it is far lower than the mean obtained from the Alexandrov *et al*. (2002a) method, by about 40 DU for the low variance. The mean of the EOF-retrieved RSS ozone values, by contrast, is 266 DU, in very close agreement with the sixteen-channel retrieval. When only the days that the MFRSR and the RSS have in common are considered, the mean values become 286 DU for the MFRSR and 268 DU for the RSS. Both mean values are slightly higher than the means for the full set, and do not bring the agreement between the two devices any closer. When the common days in the MFRSR and RSS data sets are isolated for the high-resolution ozone retrieval, however, the mean is 285 DU, within one of the EOF-retrieved MFRSR mean. Granted, the variability in the EOF-retrieved ozone values is still much greater, and the mean is decreased by a couple of days with unrealistically low values, so the two devices are not completely reconciled. What has been established is that the retrievals in general, and those that produced the largest differences in particular, cannot completely distinguish between extinction from gas absorbers and aerosols, or even from calibration-induced differences between the total optical depths measured by the different devices.

4.3.3: Comparing Ozone Retrievals with Satellite Data

The Total Ozone Mapping Spectrometer (TOMS) and the Global Ozone Monitoring Experiment (GOME) provide global satellite coverage of column ozone amounts, along with other quantities as well. Ozone data were obtained, from both TOMS and GOME, overlooking the SGP site between July 1999 and July 2000. The days for which data was obtained did not always coincide, but both satellite retrievals produced enough data to make a comparison with the results of the MFRSR and RSS data possible. In particular, the satellite data should show, at least qualitatively, if the ground-based results adequately capture the variations in ozone amounts over the course of the data set.

Table 4.2 shows the mean and standard deviations of the column ozone amounts for the RSS high-resolution retrievals, the MFRSR using the Alexandrov *et al*. (2002a) retrieval, TOMS, and GOME. These values are shown both for the full data sets, and for the handful of days for which all four devices have concurrent data. Clearly, the MFRSR has the highest values, followed by TOMS, GOME, and the RSS. The mean values are widely spaced, with no two values agreeing particularly closely. The individual ozone values are plotted vs. day in Figure 4.23. The scatter plot shows the stratified values implied by the mean values, although a few common trends emerge in the data, most notably a sharp dip in November 1999. The MFRSR has no data for November 1999, but the drop is clear in the other retrievals. Curiously, the GOME data does not appear to drop as strongly as the others. Excepting the absence of data for November and despite being much larger on the whole, the MFRSR retrievals show some qualitative similarity

with the TOMS retrievals. Despite the large difference in mean values between the RSS and TOMS, most of the relative maxima and minima in the graph occur at the same places, or at least close to each other.

The differences between the values measured by the RSS and by the satellites are large, though, and difficult to reconcile with each other. The error might not entirely lie with the RSS, however. Bramstedt *et al*. (2002) report that TOMS ozone retrievals are systematically higher than ground-based Dobson spectrometer retrievals by one or two percent. GOME values are generally lower than TOMS values, and the mean Northern Hemisphere values retrieved by GOME are less than those retrieved by TOMS by about 7 DU. Perhaps most significantly, for both satellite retrievals, some of the largest positive discrepancies with Dobson results in the Northern Hemisphere occur at a latitude of 36°, precisely the latitude of the SGP site. Unfortunately, the SGP site does not have a Dobson device, so independent ground-based ozone data is not available for direct comparison with the RSS and MFRSR-derived results.

Despite the lack of unambiguous agreement between the MFRSR and RSS ozone retrievals, a handful of conclusions can be drawn. To begin with, the five-channel algorithms that retrieve gases and aerosols simultaneously produce gas amounts that depend too much on the aerosol amounts, not just for NO_2, but for ozone as well. The MFRSR optical depth values are very high – probably too high – and the corresponding retrieved gas values are all overestimated. By contrast, the RSS optical depth values are low – perhaps too low – and the ozone values retrieved using the "MFRSR equivalent" algorithm are underestimated, and those from the other RSS retrievals most likely are as well. The NO_2 values are also significantly lower in the "MFRSR equivalent" retrieval,

but given the strong trade-off between aerosols and NO_2, these values are still in reality too high, and too strongly correlated with aerosol optical depth. Despite following the same process as the "MFRSR equivalent" algorithm, the sixteen-channel retrieval captures more of the details of the ozone and NO_2 absorption spectra, and at least reduces the disagreement in the mean values. The EOF-based retrievals separate ozone from aerosols the same way for both the MFRSR and RSS, and produce a level of agreement between the two devices that, while not perfect, is significantly improved.

4.4: The Single-Mode Aerosol Results

4.4.1: The MFRSR Retrieval Results

As the mean retrieved values for low and high variances presented in Table 4.1 show, a change in the assumptions going into the retrieval algorithm can cause a large difference in the retrieved values of not only NO_2, but the aerosol effective radius as well. The mean effective radius for the full MFRSR data set when the small variance is assumed is 0.438 μm, fully consistent with the results for the SGP site published in Alexandrov *et al.* (2002b). Such a large value can be explained by the large NO_2 values, because if that quantity is overestimated, then the effective radius is as well. With particles this large, the aerosol extinction will start to decrease as the wavelength decreases towards the violet and ultraviolet. This is not observed in the means of the optical depth values for any of the devices (see Figure 2.12) or in the EOF analysis.

Figure 4.24 shows the aerosol optical thickness values at 550 nm retrieved for the effective variance of 0.01, plotted as a function of day. Not surprisingly, the optical depth follows the same pattern seen in Figure 2.11. For the higher variance, depicted in Figure 4.25, the daily variation qualitatively matches the lower variance, but the optical depth values are slightly larger on the whole. This occurs because less extinction is attributed to gaseous absorption when the higher variance is assumed. The effective radius values over the full data set for a variance of 4.26 are shown in Figure 5.3. The values peak in the autumn, despite a summer minimum. If NO_2 is being overestimated, though, the effective radius values will be similarly affected, and the true cycle for particle sizes, if one exists, will not emerge clearly in the data. Figure 4.27 shows the corresponding values for the large variance. Here, a handful of the radius values in the fall and winter are very large, fluctuating rather dramatically much like the ozone values for the high variance (see Figure 4.18), but the radius values on the whole are substantially lower. Some of the reasons for these strange values may be device-related – the MFRSR was out of commission for the month of November – but calibration issues may come into play as well.

4.4.2: The RSS Retrieval Results

Figure 4.28 shows the plots vs. day of the aerosol optical depth for the "MFRSR equivalent" retrieval, while Figure 4.29 shows the corresponding graph for the sixteen-channel, single-mode, coupled-gas retrievals. Both plots show the same seasonal cycle. As expected, the aerosol optical depths are lower than those of the MFRSR. As Table 4.1

shows, the mean values for aerosol optical depth in the "MFRSR equivalent" retrievals just barely agree with the MFRSR and CIMEL within the 0.02 margin of error anticipated for calibration differences. The effective radius values, presented in Figures 4.30 and 4.31, are considerably lower even than those of the high variance values for the MFRSR. Curiously, though, very high radius values similar to those in the MFRSR results show up in the sixteen-channel retrieval, but not in the "MFRSR equivalent" version.

With these retrievals, it is also possible to retrieve the effective variance, although it is not clear that the variance can be clearly separated from the other quantities. In fact, if the aerosol size distribution is bimodal, then the effective variance most likely cannot be uniquely constrained. The interrelation between optical depth and effective variance is demonstrated for the two single-mode RSS retrievals in Figures 4.32 and 4.33. In both cases, the large variance dominates the values, but the small variance oddly is the second most prominent value. The small variances occur only when the optical depth is small. With the broader wavelength range and greater spectral detail of the sixteen-channel retrieval, the number of low variance days decreases significantly – which, of course, is exactly what would be expected if the aerosol size distribution were in reality bimodal.

The plots vs. day of aerosol optical depth and effective radius, for the retrievals where the gases are retrieved separately using the high-resolution algorithms, are presented in Figures 4.34 and 4.35. The optical depth values are larger than in the coupled-gas retrievals, because less extinction is attributed to nitrogen dioxide. For the same reason, the retrieved values of the effective radius are considerably shorter. The retrieved effective radius values exceed 0.4 μm on only a handful of days in late fall and

the winter, in sharp contrast to the coupled-gas retrieval. The effective radius values are slightly larger after day 400 than before, but there does not appear to be a clearly defined seasonal cycle.

An examination of the retrieval for individual days indicates that even with the gases retrieved separately, inadequacies persist in the single-mode aerosol retrievals. Figure 4.36 depicts the results of this retrieval for the morning of June 23. The NO_2 and ozone seem to be measured adequately, but the aerosol optical depth has unacceptably large residual optical depth in excess of 0.02 at both spectral extremes and, in the opposite direction, in the middle. A bimodal aerosol distribution would create exactly this kind of pattern in the residuals if the retrieval algorithm assumes a single-mode distribution. All the criteria for establishing the inadequacy of single-mode retrievals algorithms have been met in the RSS data. The RSS retrievals, in agreement with the EOF analysis, confirm that the aerosol size distribution being observed is bimodal.

4.4.3: The CIMEL Retrieval Results

Figure 4.37 shows the values of the aerosol optical depth at 550 nm retrieved from the CIMEL data assuming a single-mode size distribution. These retrievals did not include the channel at 670 nm, due to the defect already noted, or the channel at 340 nm, for the sake of a more direct comparison with the RSS data. The maximum values do not approach those of the MFRSR, but are higher on the whole than the RSS values. This indicates that the large intercepts in Figure 2.14 do not get overcome at any point by the small slopes. Qualitatively, the CIMEL data shows the same seasonal dependence seen

in the other two devices. The size distributions show greater differences with the other two devices, although this difference can be explained by relative sharpness of the baseline curve of the CIMEL data in Figure 2.12. The effective radius values, as shown in Figure 4.38, are much smaller than either the MFRSR or the RSS values. The retrieved values do not become very large on any days in the winter either, as was the case with the other two devices. The effective variance was retrieved as 0.5 for each day in the data set, never once returning a lower value, once again consistent with the observation of bimodal aerosol distributions.

Just like the RSS, an examination of the residual optical depth values in the separate channels shows positive residuals at the ends, and negative residuals in the middle. This is illustrated in Figure 4.39, using the data from the afternoon of November 2, 1999. As this pattern repeats for every day in the set, it suggests a systematic problem rather than a random variation. Again, it is demonstrated that one Mie scattering curve alone will not adequately represent the details in the aerosol optical depth, and that introducing a second aerosol mode into the retrieval is necessary to ensure accuracy.

4.5: Conclusions

The single-mode retrievals have been shown to be flawed. Attempts to retrieve gases and aerosols simultaneously, while assuming a single-mode aerosol size distribution, produce significant errors in retrieve gas abundances, along with false correlations between gases and aerosols. While the retrieval of ozone is less sensitive to aerosol amounts than the retrieval of nitrogen dioxide, its accuracy is strongly

constrained by the accuracy of the aerosol assumptions and measurements. The large optical depth difference between the MFRSR and RSS, presumably a result of calibration errors, causes artificially large differences between the two devices in retrieved ozone values. Upon further scrutiny, these differences are mostly removed. The higher spectral resolution of the RSS data improves the measurement of the gases, which further exposes the inability of single-mode retrievals to adequately describe the aerosol extinction curves observed on a daily basis. In the following chapter, the bimodal retrieval algorithms will be tested, to see if the trade-off between aerosols and gases is eliminated or simply reduced relative to the single-mode retrievals, and to determine if the amount of retrievable aerosol information really is as limited as Box *et al*. (1996) suggested.

	mean aerosol τ, 550 nm	mean effective radius (μm)	dominant effective variance	mean NO_2 (DU)	mean ozone (DU)	mean RMS residual ε
RSS, gases retrieved separately	0.083	0.176	0.5 (94.4%)	0.39	283	0.00535
MFRSR, low variance	0.089	0.451	0.01	4.21	339	-------
MFRSR, high variance	0.097	0.348	0.4	2.32	327	-------
RSS "MFRSR equivalent"	0.078	0.240	0.4 (77.8%)	1.44	245	0.00071
RSS, gases coupled	0.073	0.268	0.5 (92.6%)	2.03	265	0.00360
CIMEL, optical depth values	0.098	0.167	0.5 (100%)	----	---	0.00757

Table 4.1: The mean values of the retrieved quantities for the single-mode retrievals, taken over the 54 common days of the three devices.

	mean O_3 (DU), full set	σ (DU), full set	mean O_3 (DU), common days	σ (DU), common days
RSS bimodal, gases retrieved separately	279	24	276	28
MFRSR	329	38	311	57
TOMS	305	26	306	29
GOME	300	25	290	19

Table 4.2: The mean retrieved ozone values from the RSS, MFRSR, TOMS, and GOME data sets.

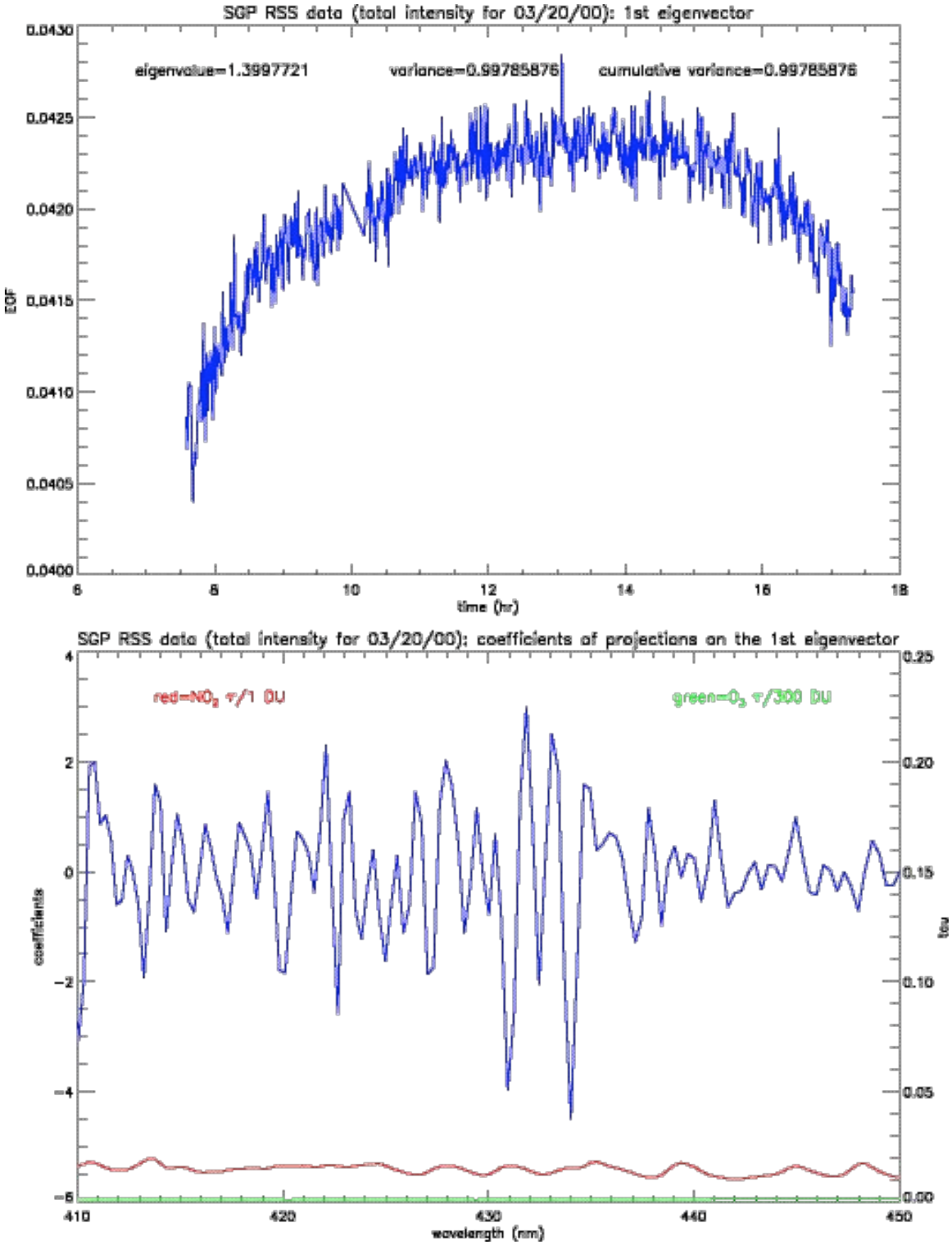

Figure 4.1: The first eigenvector of the Ring effect of the total intensity measured by the RSS on March 20, 2000.

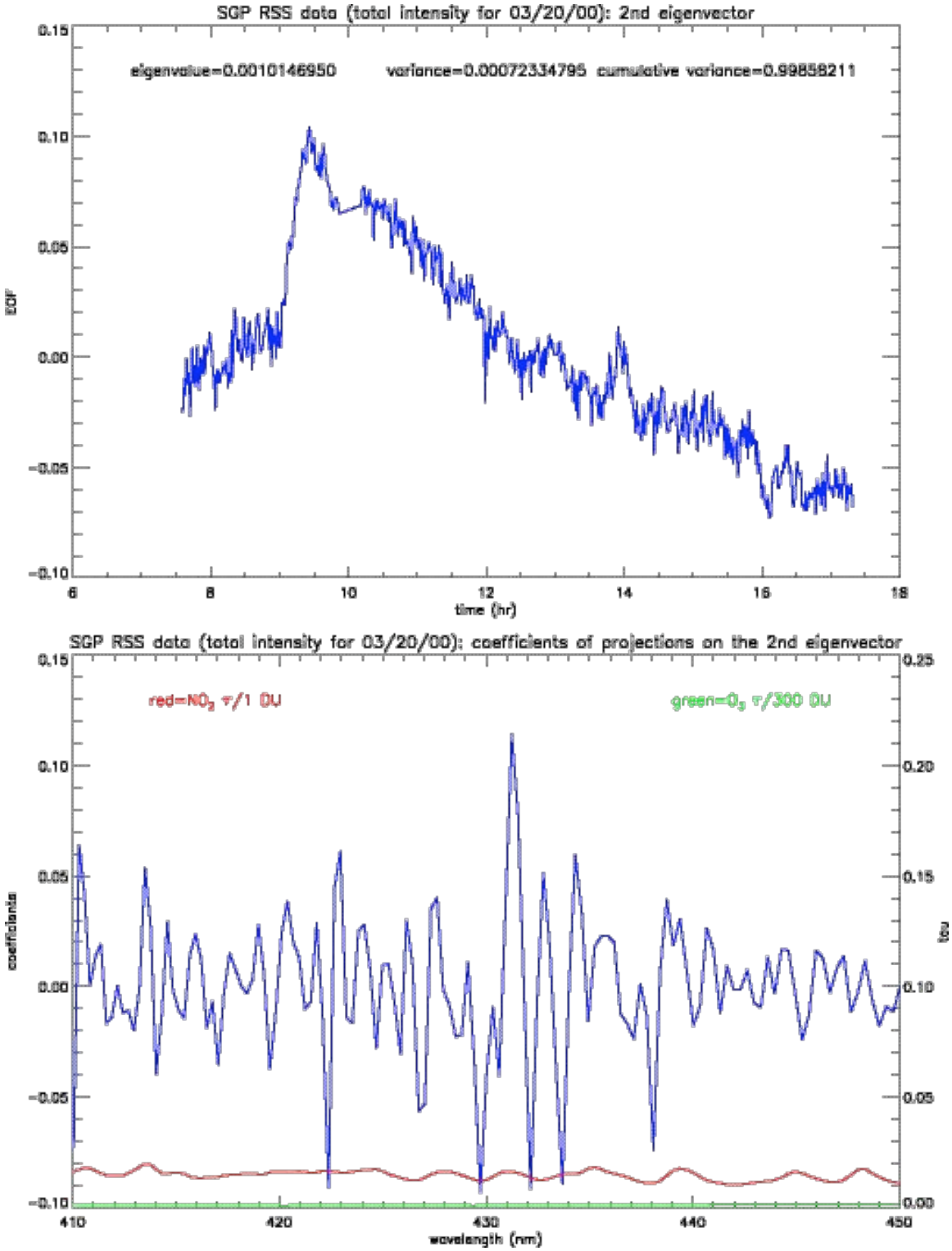

Figure 4.2: The second eigenvector of the Ring effect of the total intensity measured by the RSS on March 20, 2000. There is an abrupt shift in the EOF, taking place early in the morning.

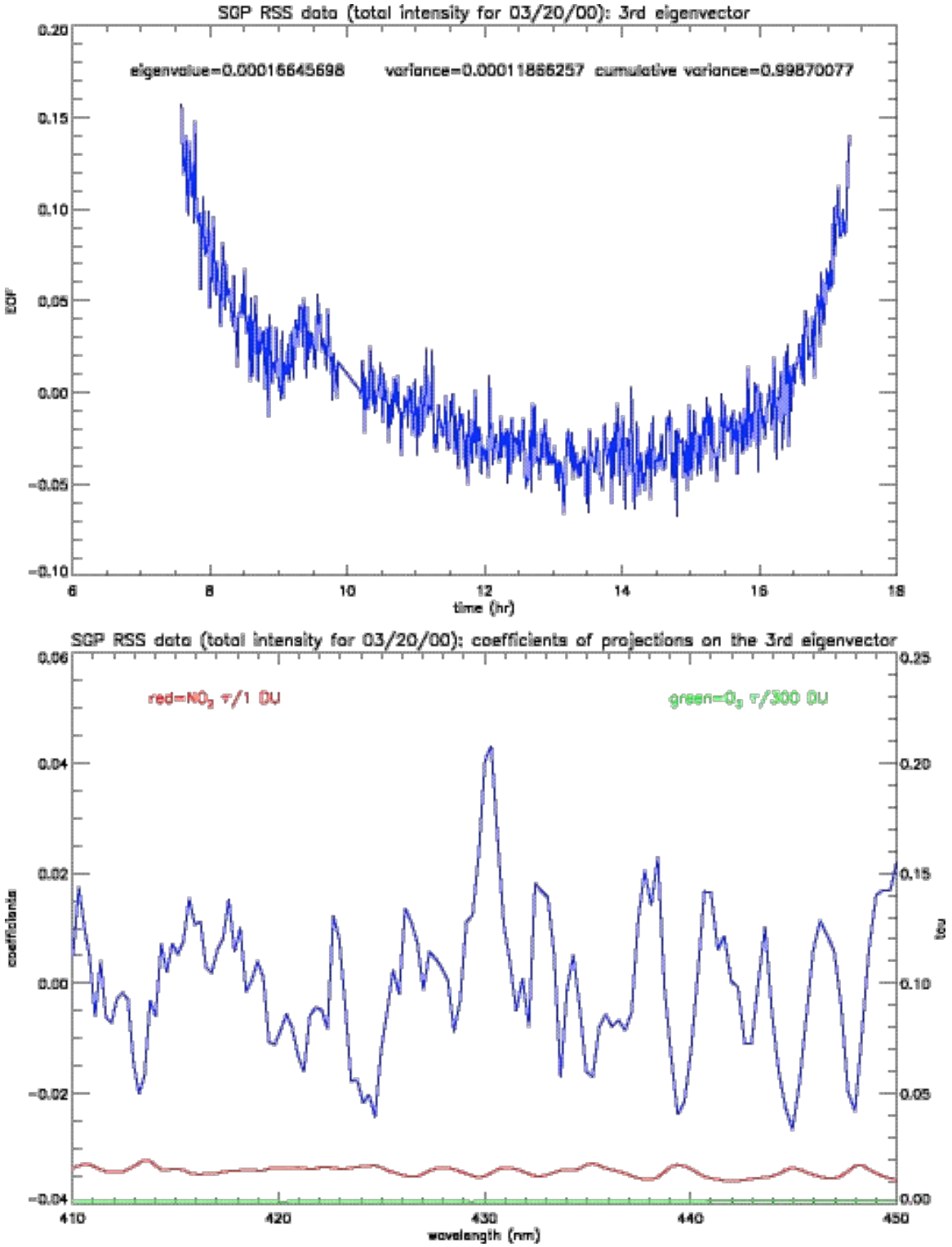

Figure 4.3: The third eigenvector of the Ring effect of the total intensity measured by the RSS on March 20, 2000. The filling of NO_2 absorption lines is particularly strong.

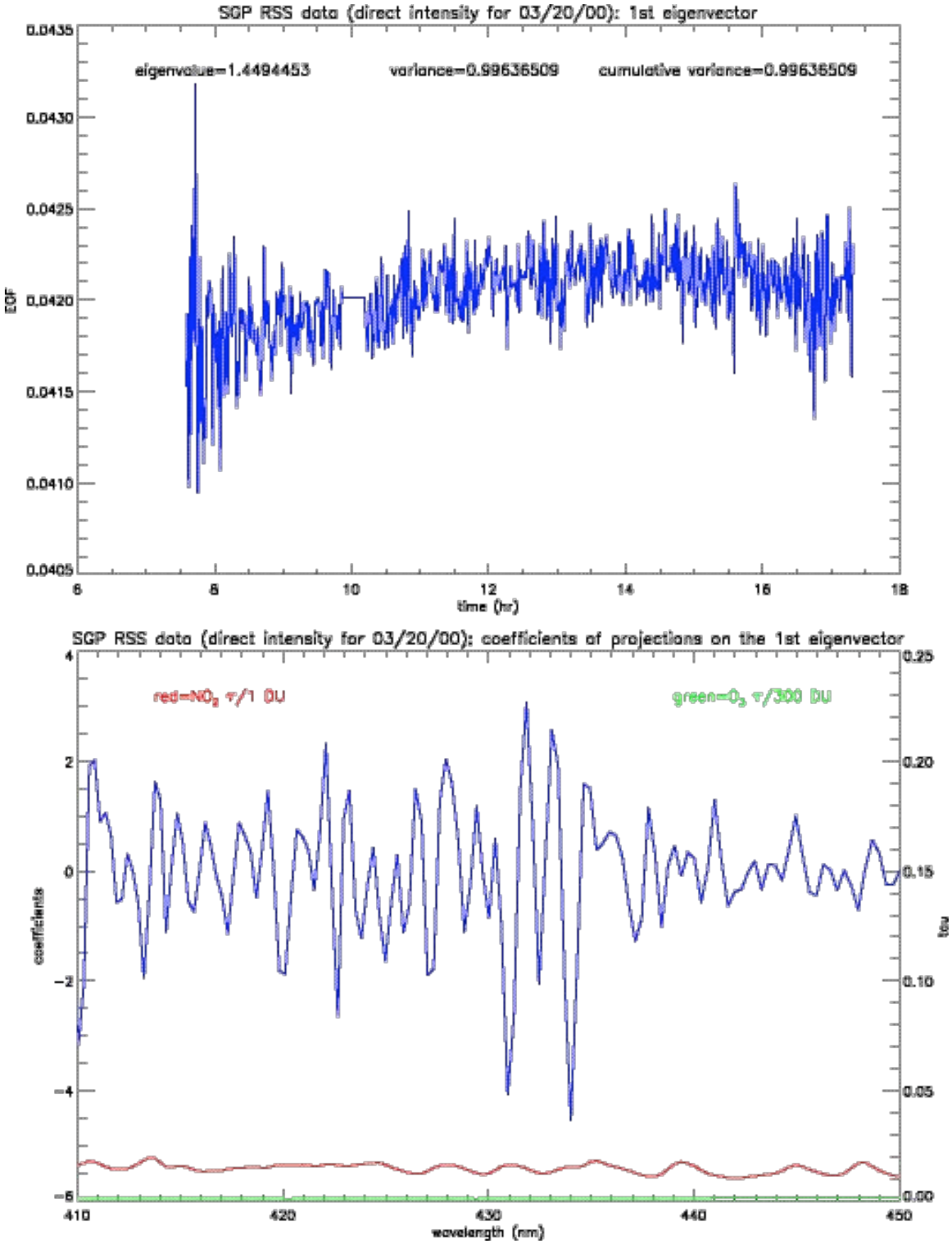

Figure 4.4: The first eigenvector of the Ring effect of the direct beam measured by the RSS on March 20, 2000.

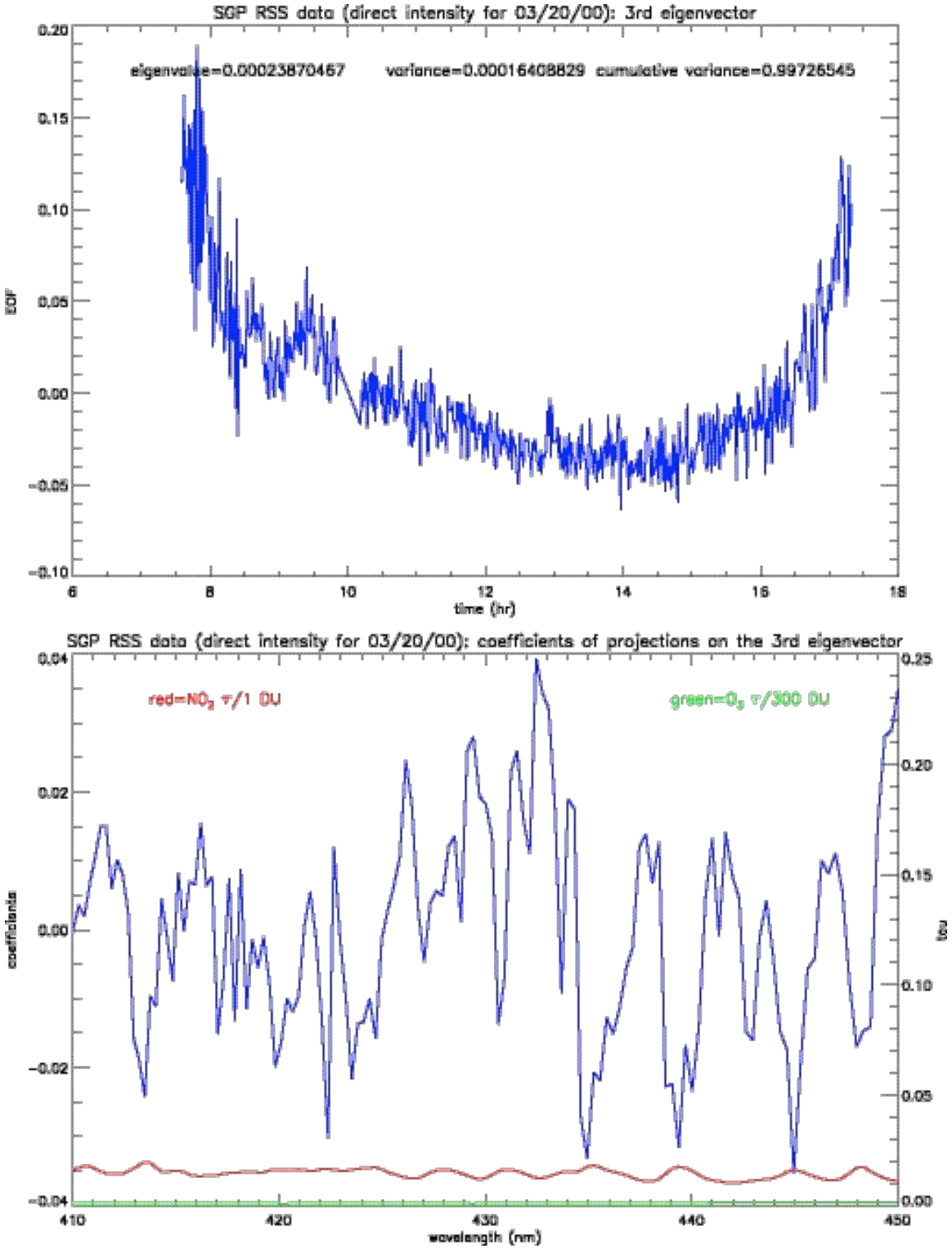

Figure 4.5: The third eigenvector of the Ring effect of the direct beam measured by the RSS on March 20, 2000.

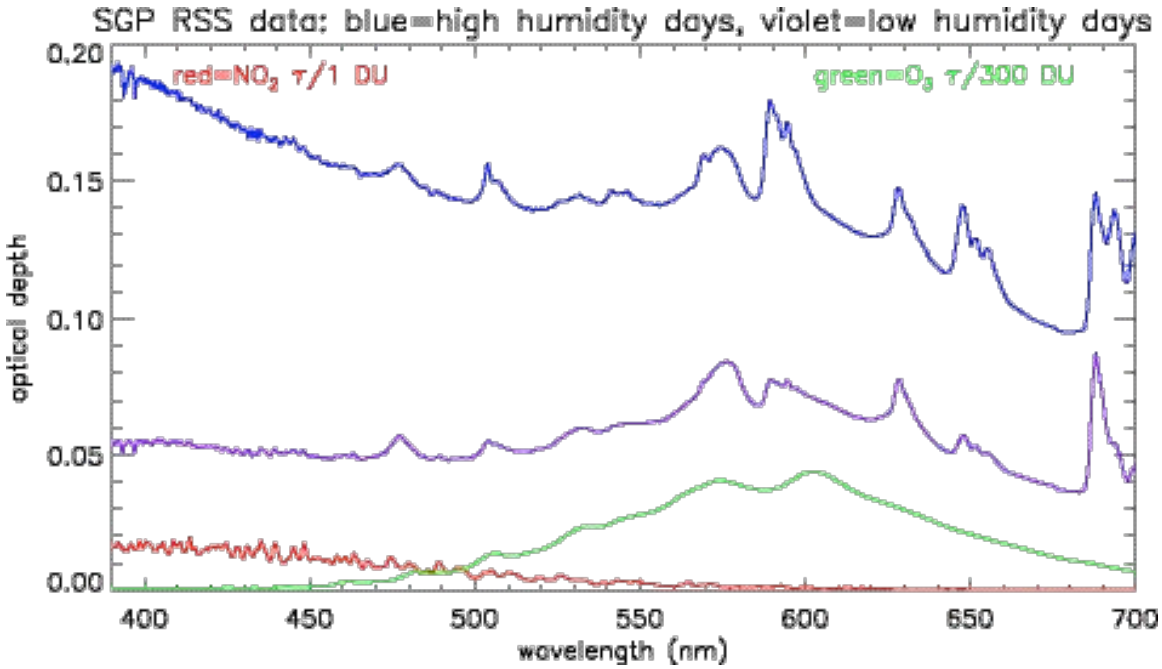

Figure 4.6: RSS mean optical depth values for the thirty highest humidity days (blue), which are also high aerosol days, and the thirty lowest humidity days (violet), which are low aerosol days. The details of the Ring effect are clear in the blue graph, but not nearly as clear in the violet graph, in apparent contradiction with theoretical predictions.

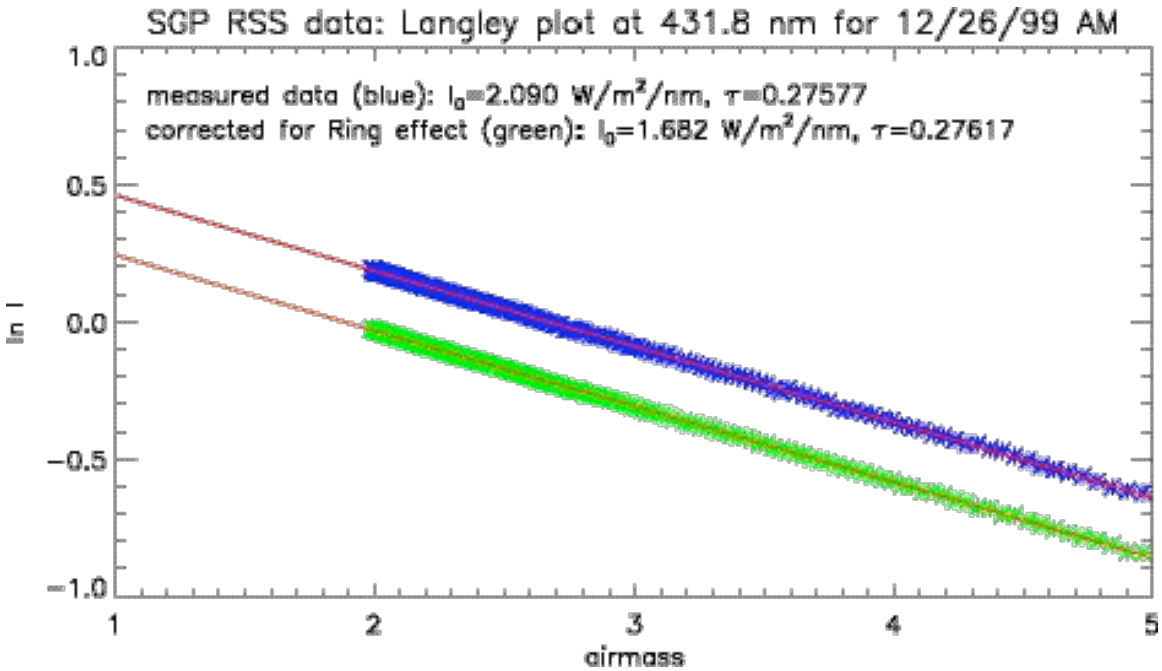

Figure 4.7: The Langley regression for the 431.8 nm channel of the RSS for the morning of December 26, 1999, a low aerosol day, both for the actual measured intensities and for the Ring-corrected intensities. The difference in the Langley optical depth is 0.00040.

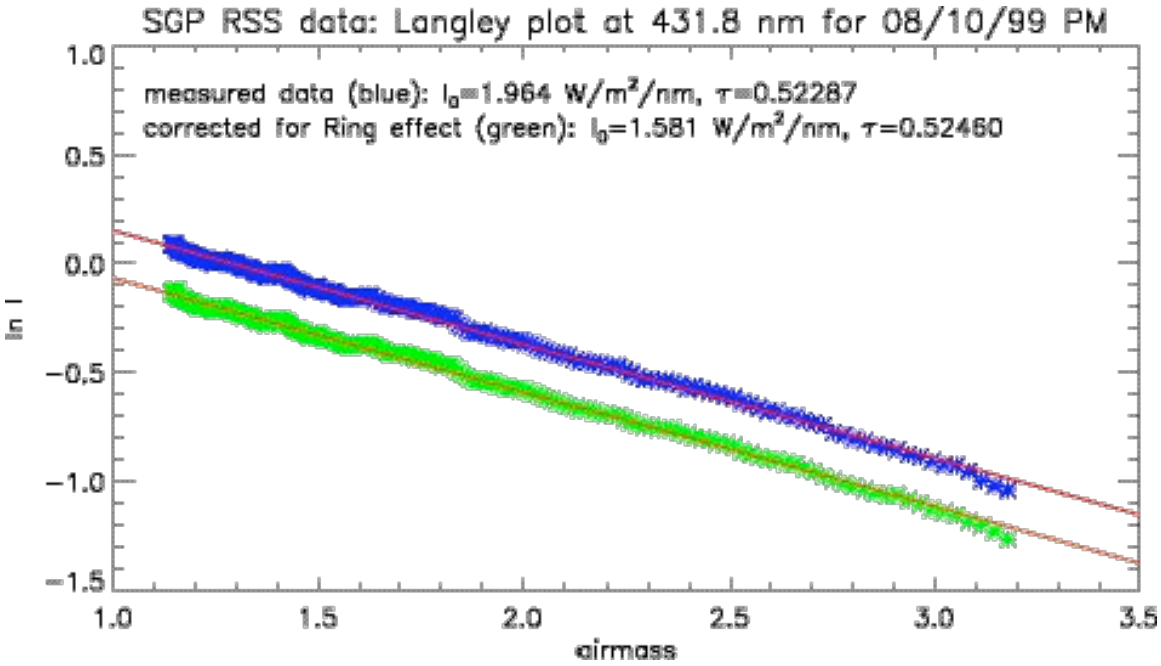

Figure 4.8: The Langley regression for the 431.8 nm channel of the RSS for the afternoon of August 10, 1999, a high aerosol day, both for the actual measured intensities and for the Ring-corrected intensities. The difference in the Langley optical depth is 0.00173, greater than the difference on December 26 by more than a factor of four.

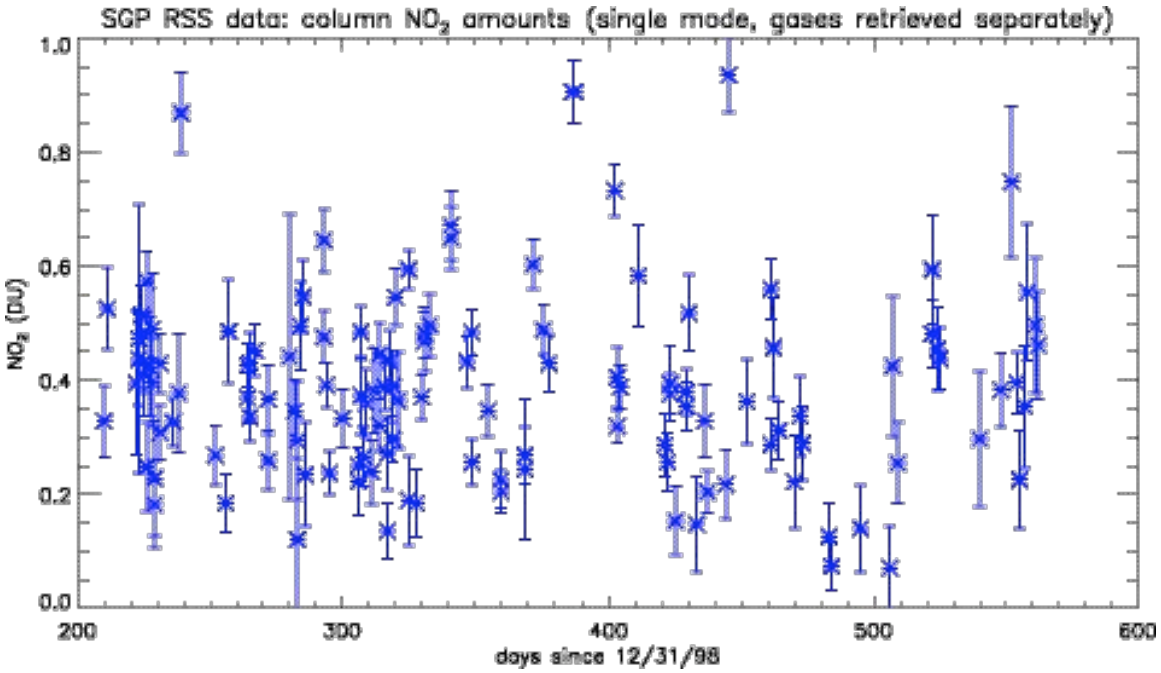

Figure 4.9: The results of the high-resolution NO_2 retrieval, with error bars denoting the level of random noise for each day.

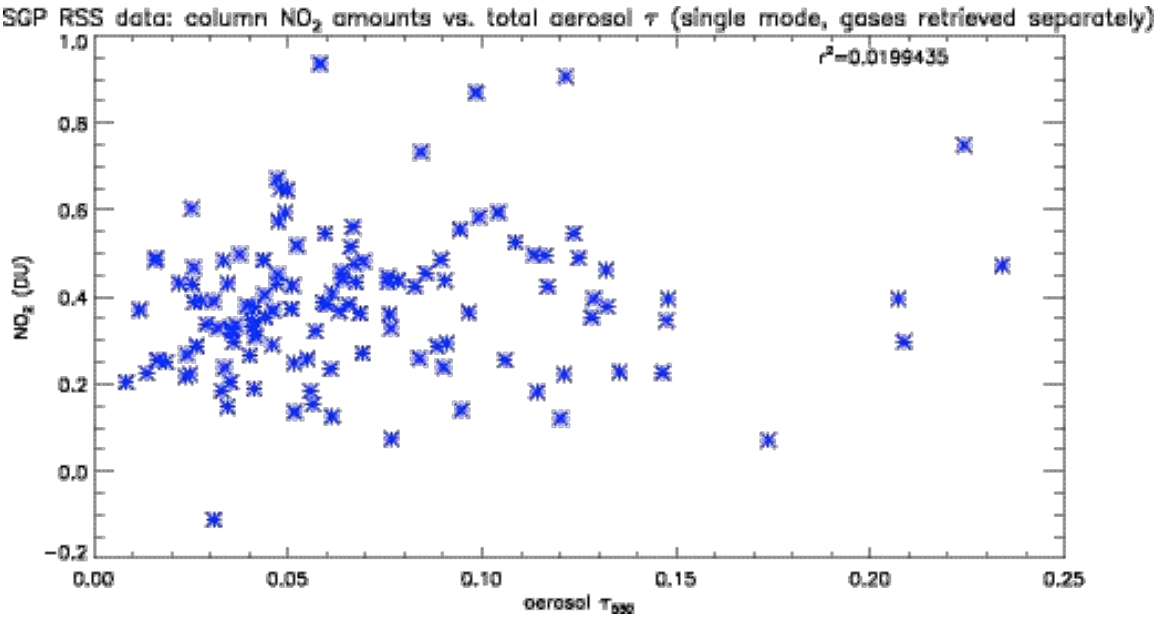

Figure 4.10: The scatter plot of NO_2 vs. aerosol optical depth at 550 nm for the single-mode RSS retrieval with the gases done separately. Note that the correlation between the two quantities vanishes almost entirely.

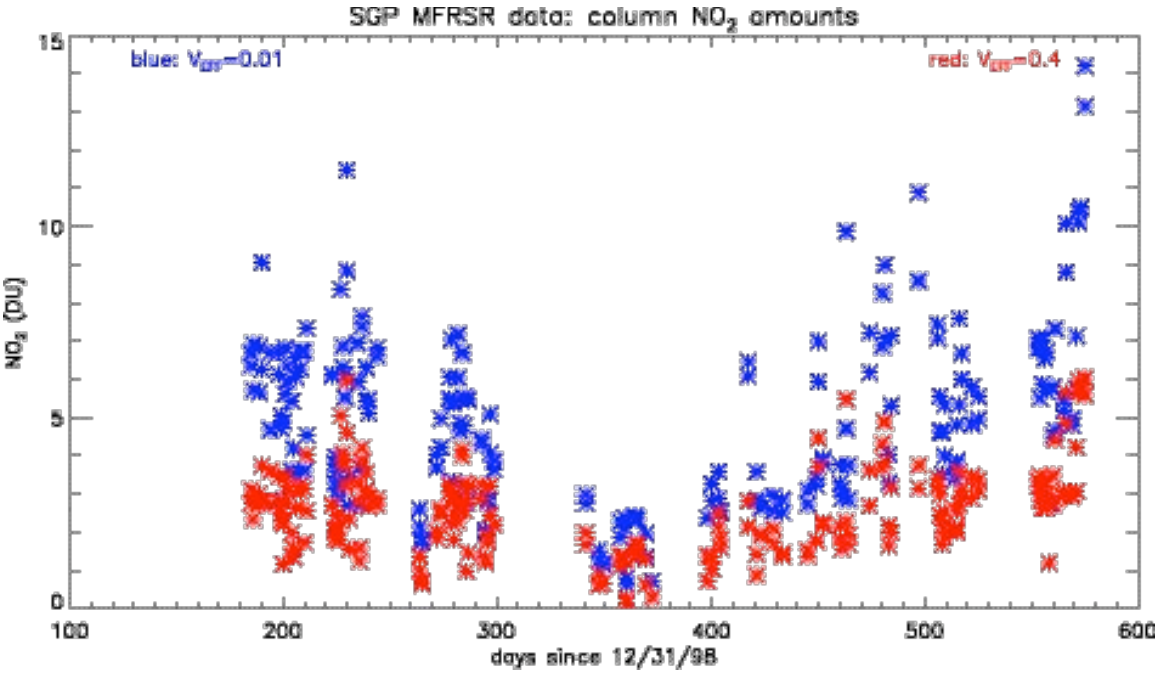

Figure 4.11: The plot of retrieved NO_2 vs. day for the low variance (blue) and the high variance (red) in the MFRSR retrieval. The patterns closely resemble the plot of optical depth vs. day in Figure 2.11.

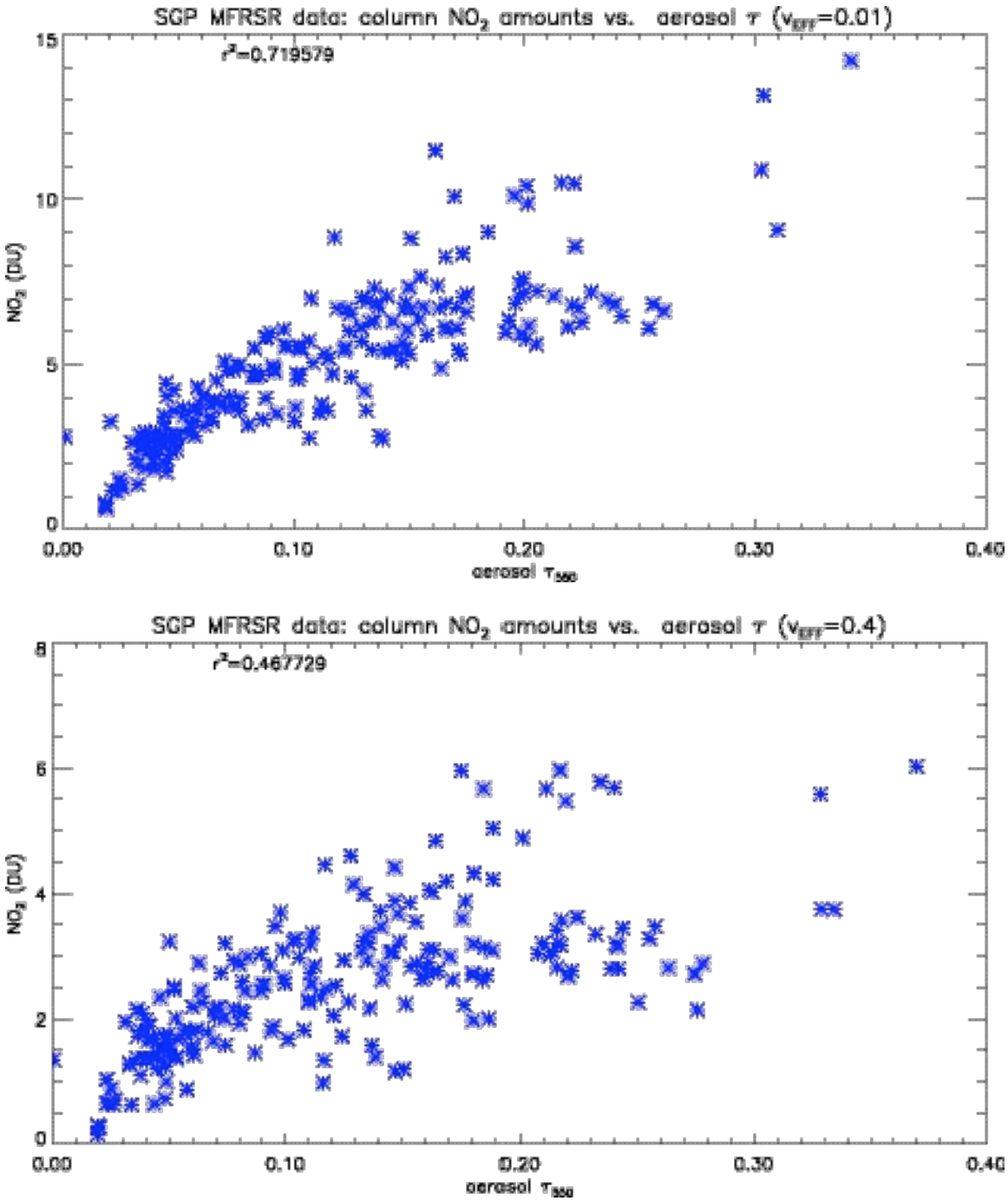

Figure 4.12: The plots of retrieved NO_2 amounts vs. retrieved aerosol optical depth for both variance values. The correlation drops sharply with increasing variance, suggesting that the large NO_2 values might be an artifact of the retrieval algorithm rather than an indication of an actual physical correlation.

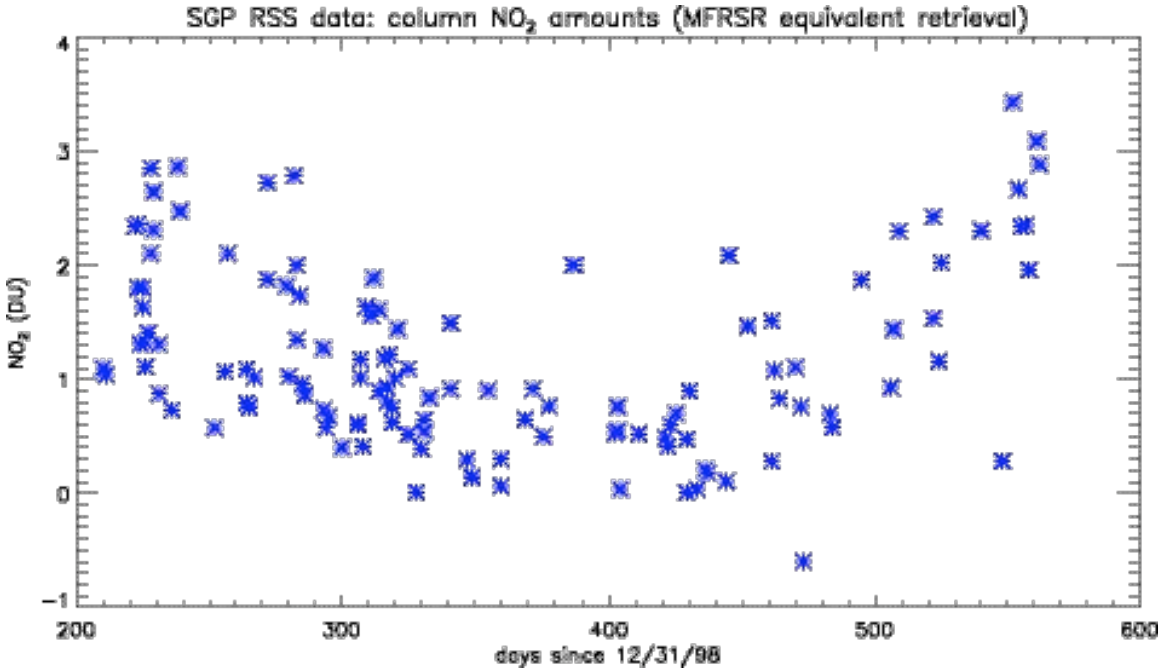

Figure 4.13: The plot of retrieved NO_2 vs. day for the "MFRSR equivalent" retrieval. These values follow the pattern of the RSS optical depth values.

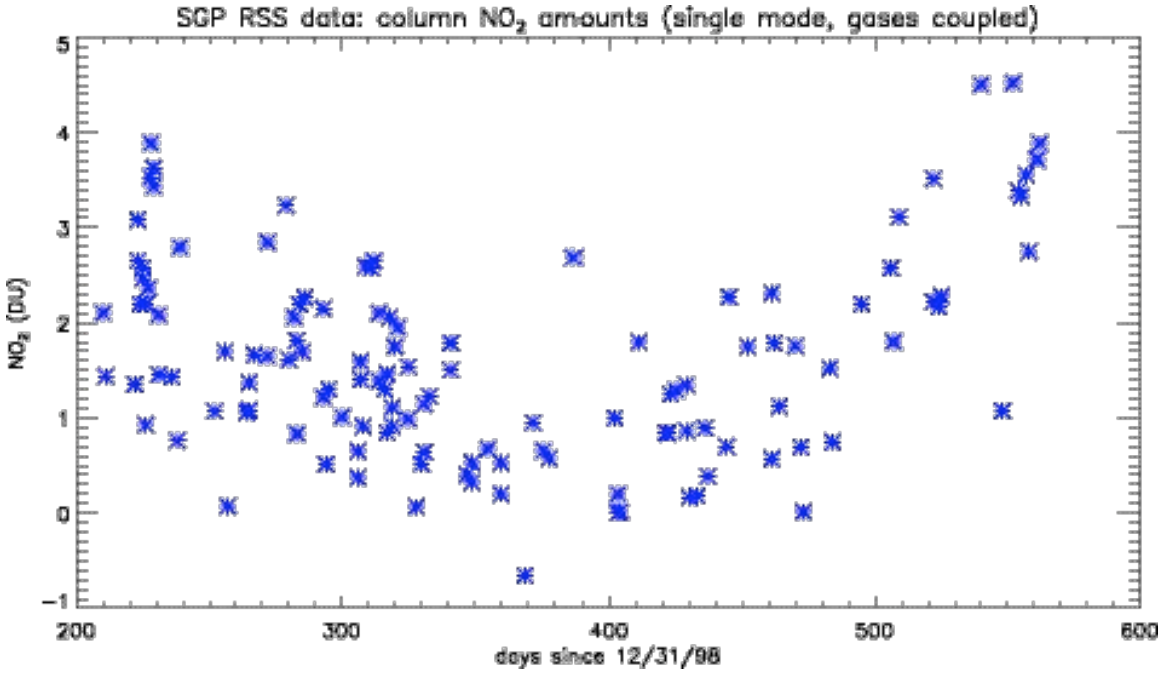

Figure 4.14: The plot of retrieved NO_2 vs. day for the sixteen-channel, full RSS single-mode retrieval. The pattern is almost identical to the one in Figure 4.13.

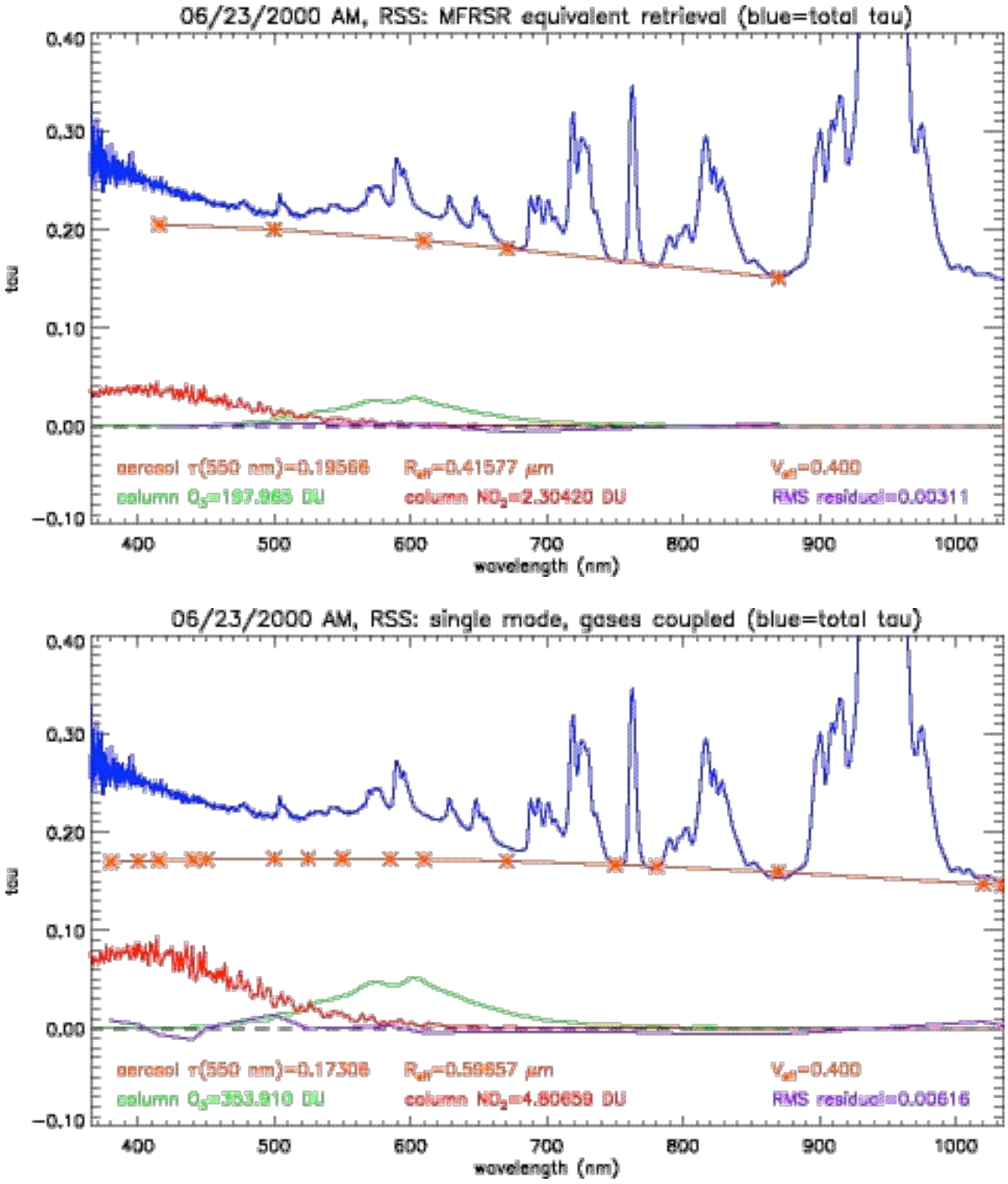

Figure 4.15: The "MFRSR equivalent" and single-mode, coupled gas retrievals using RSS data for the morning of June 23, 2000. Note the unrealistically large NO_2 values and the wildly disparate ozone values.

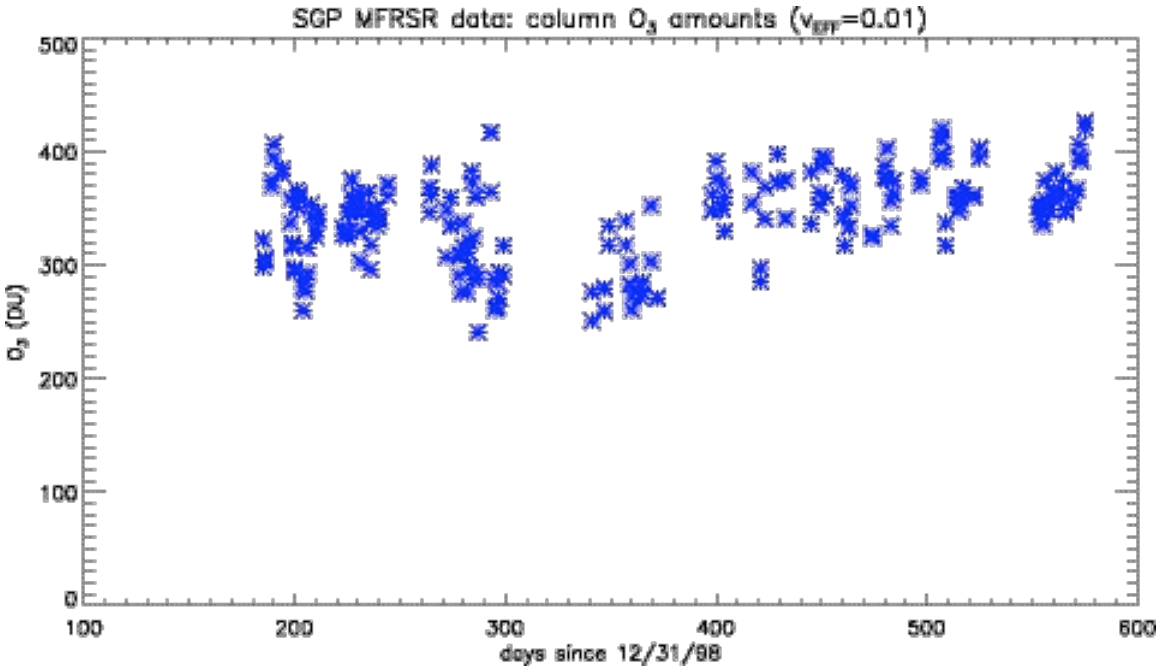

Figure 4.16: The MFRSR-retrieved values of ozone, for the low variance, plotted vs. day.

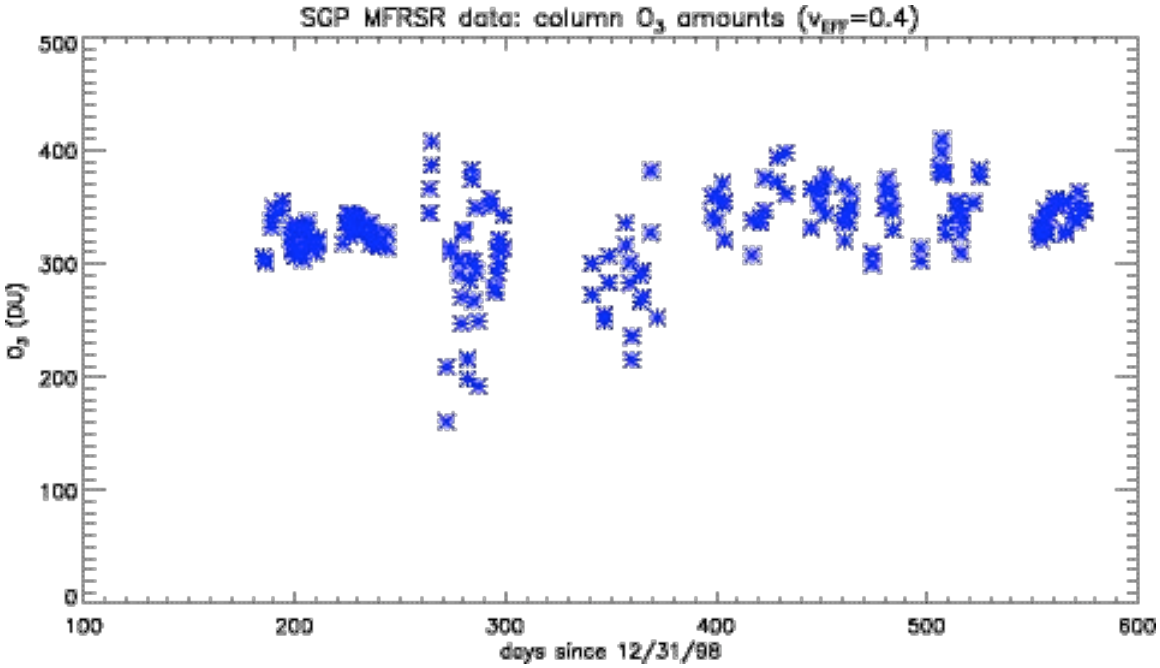

Figure 4.17: The MFRSR-retrieved values of ozone, for the high variance, plotted vs. day.

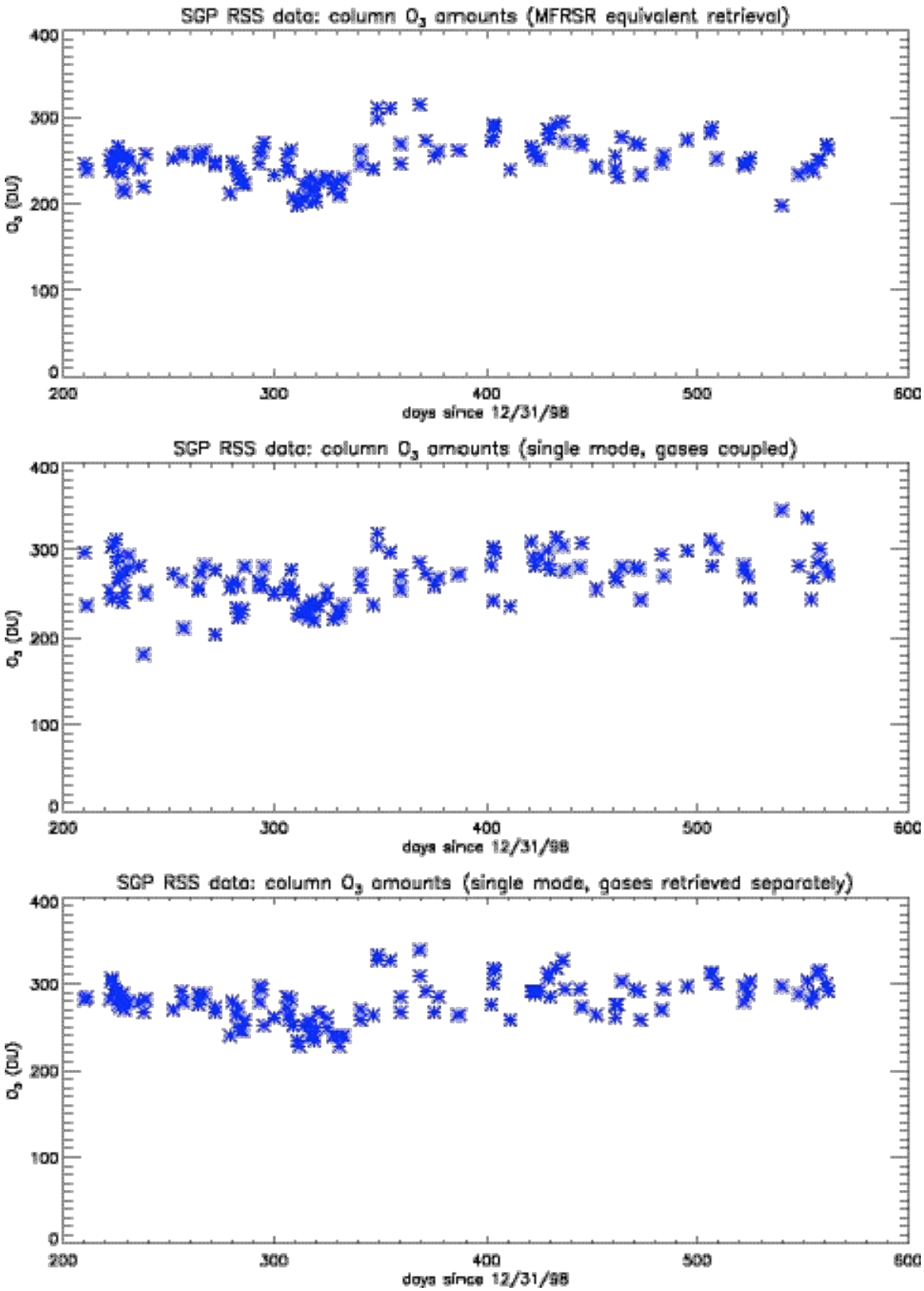

Figure 4.18: The ozone values measured using RSS data in the "MFRSR equivalent," sixteen-channel full spectrum, and high-resolution retrievals, respectively.

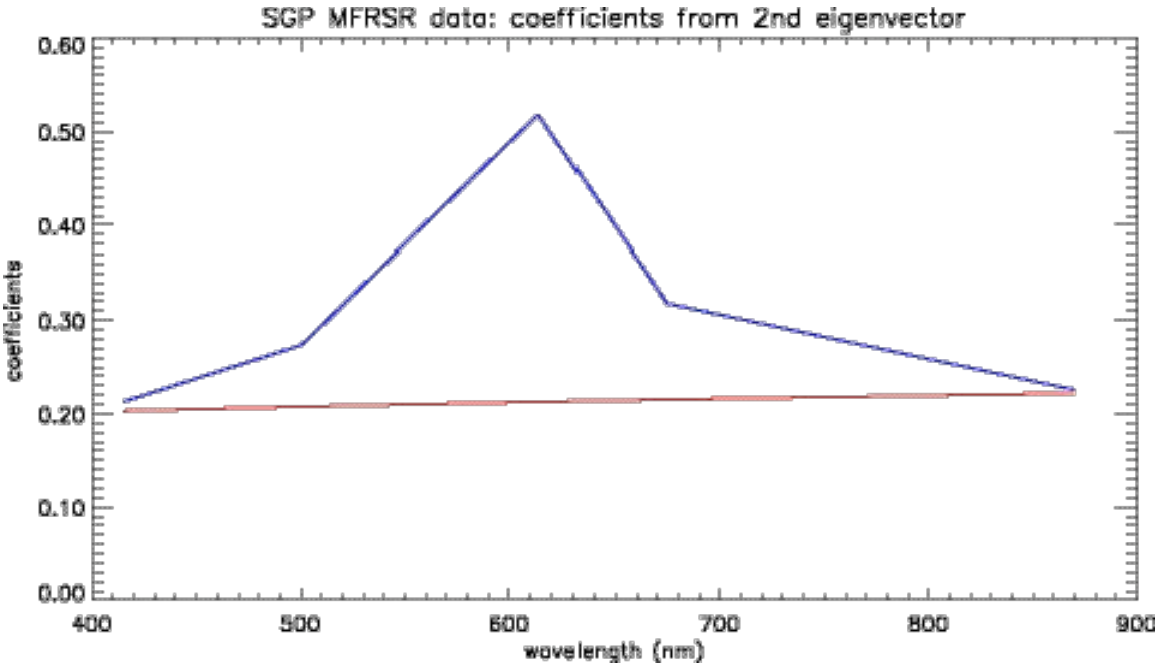

Figure 4.19: The coefficients of the second EOF of the MFRSR data. The red line indicates the baseline curve, determined after calculating the ozone contribution to the EOF.

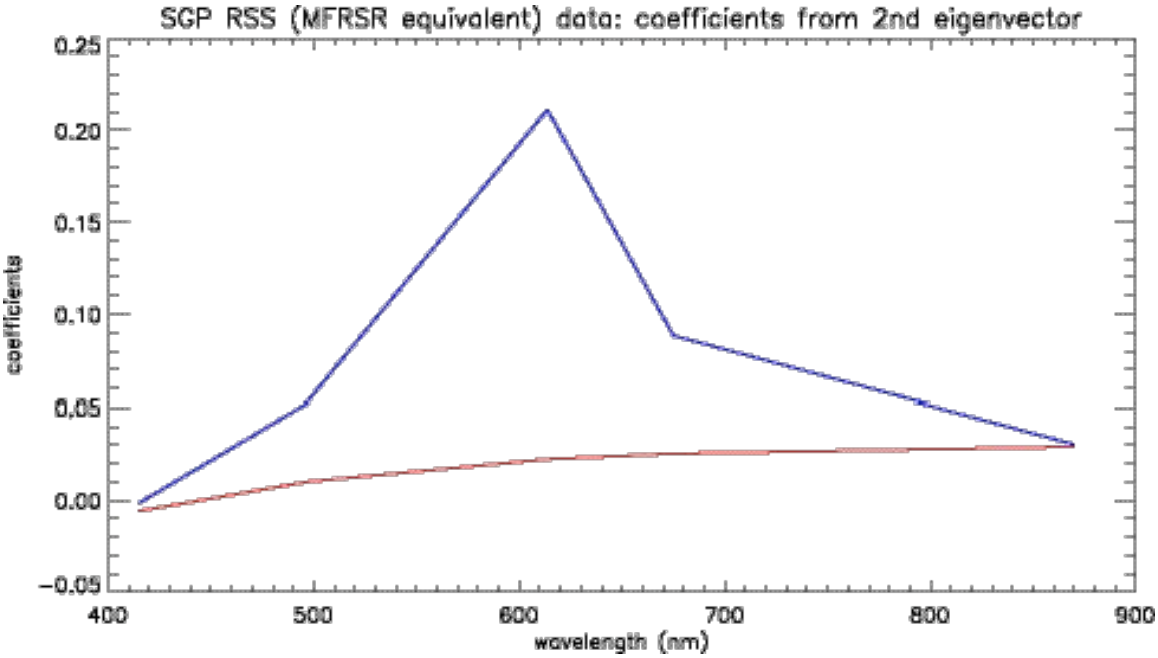

Figure 4.20: The same as Figure 4.19, except for the "MFRSR equivalent" variant of the RSS data.

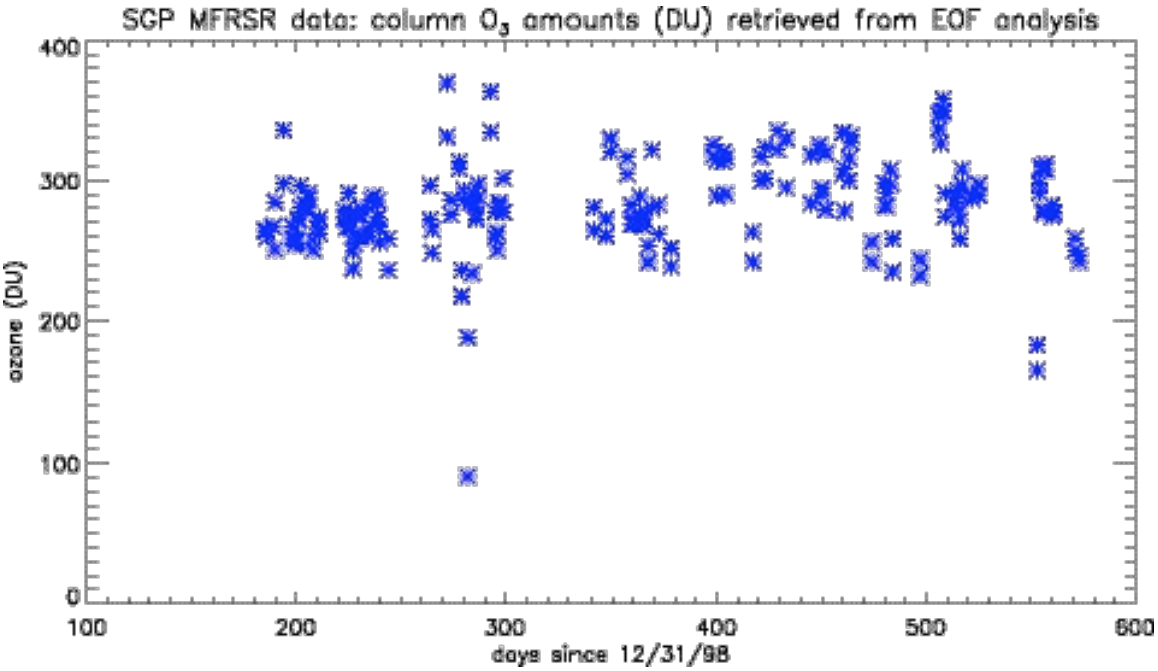

Figure 4.21: The EOF-based ozone retrievals for the MFRSR data, plotted vs. day.

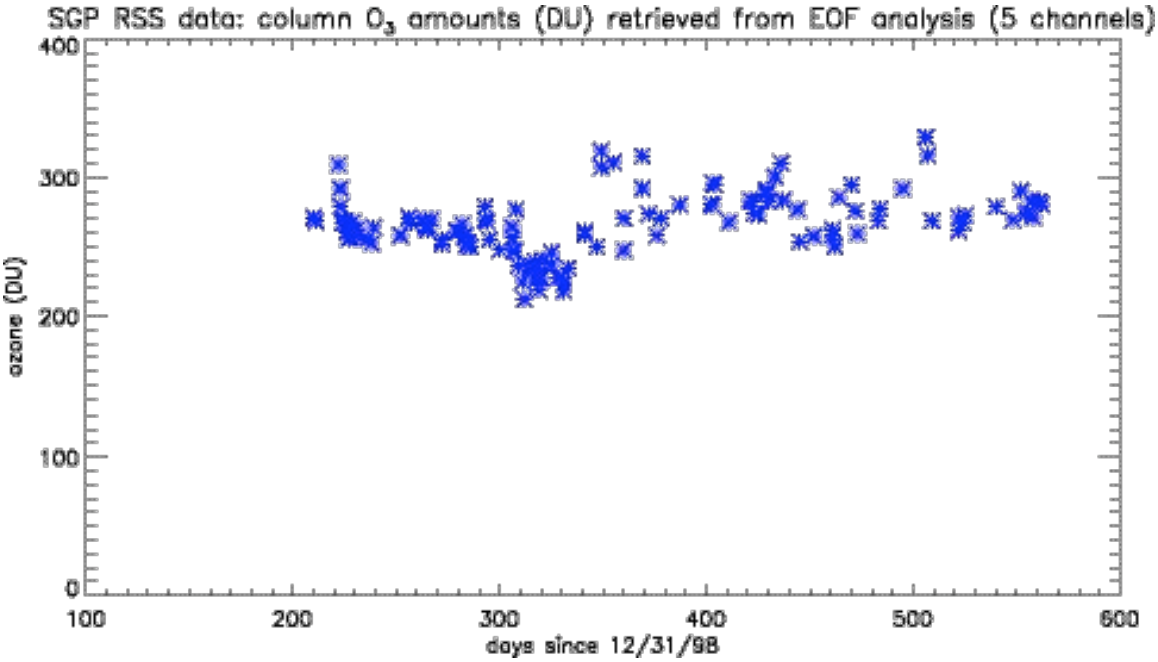

Figure 4.22: The EOF-based ozone retrievals for the RSS data, plotted vs. day.

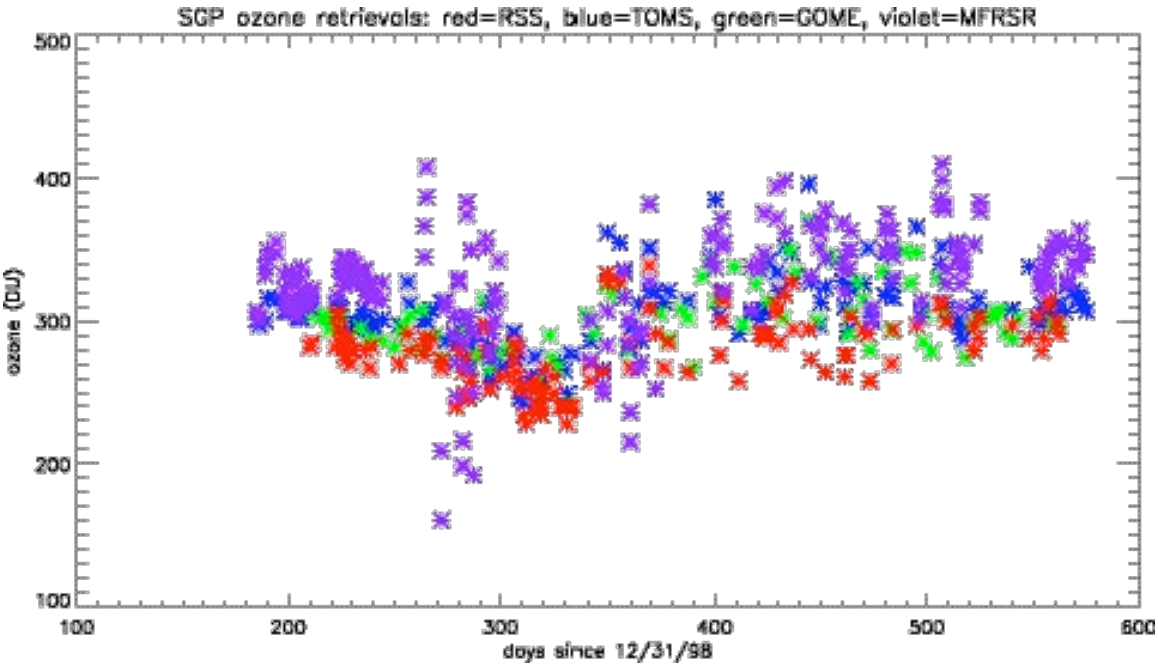

Figure 4.23: A scatter plot of the retrieved ozone values from the RSS high-resolution data (red), the MFRSR (violet), TOMS (blue), and GOME (green). The values are stratified in a quantitative sense, but some common temporal variations are apparent.

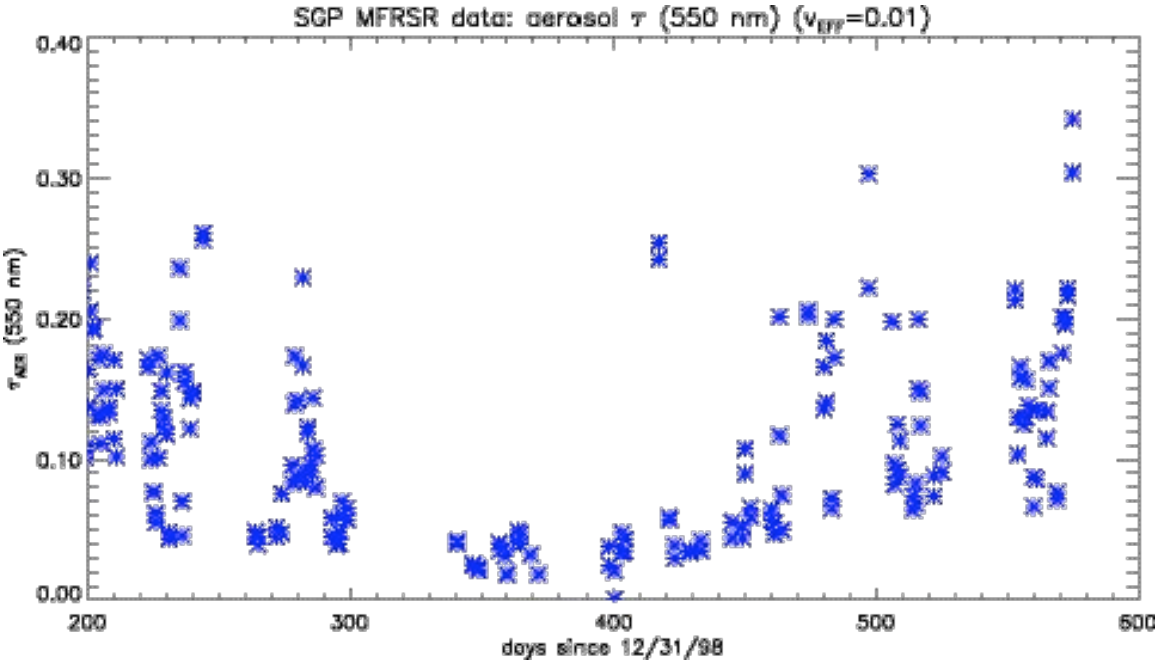

Figure 4.24: The MFRSR-retrieved aerosol optical depth values at 550 nm, from July 1999 to July 2000, for an effective variance of 0.01. Note the strong seasonal dependence.

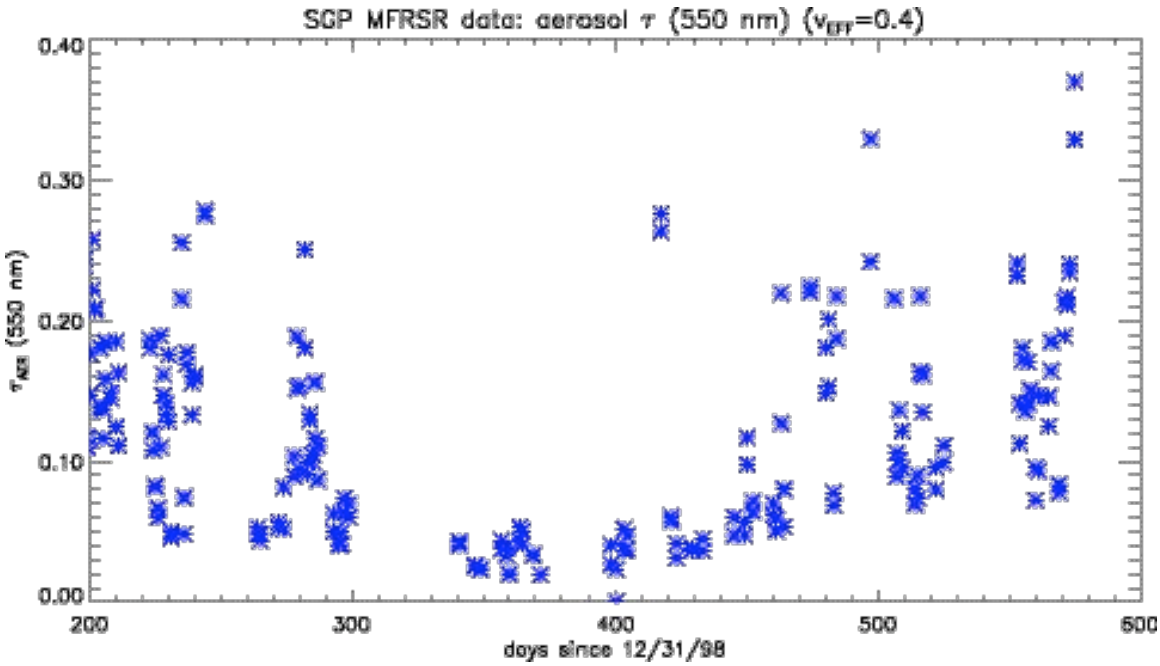

Figure 4.25: Same as Figure 4.24, except for an effective variance of 0.4. The daily variations agree qualitatively, but the mean optical depth for the higher variance is greater by 0.01.

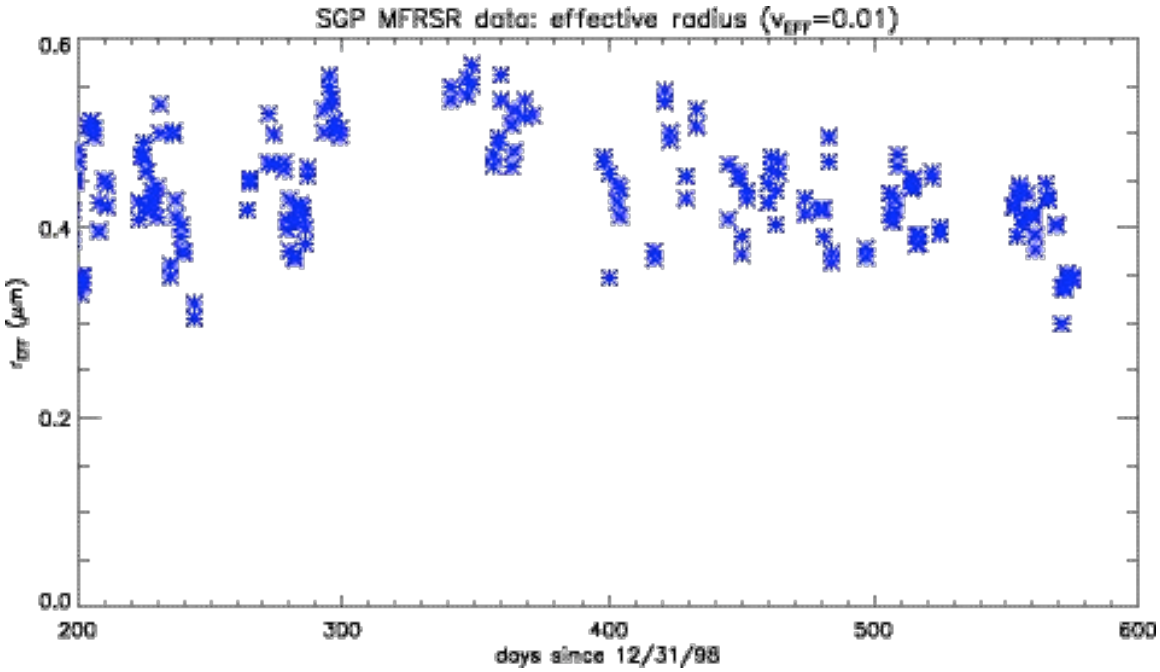

Figure 4.26: The retrieved effective radiance for a variance of 0.01. The values are lowest in the summer, but actually appear to peak in the fall.

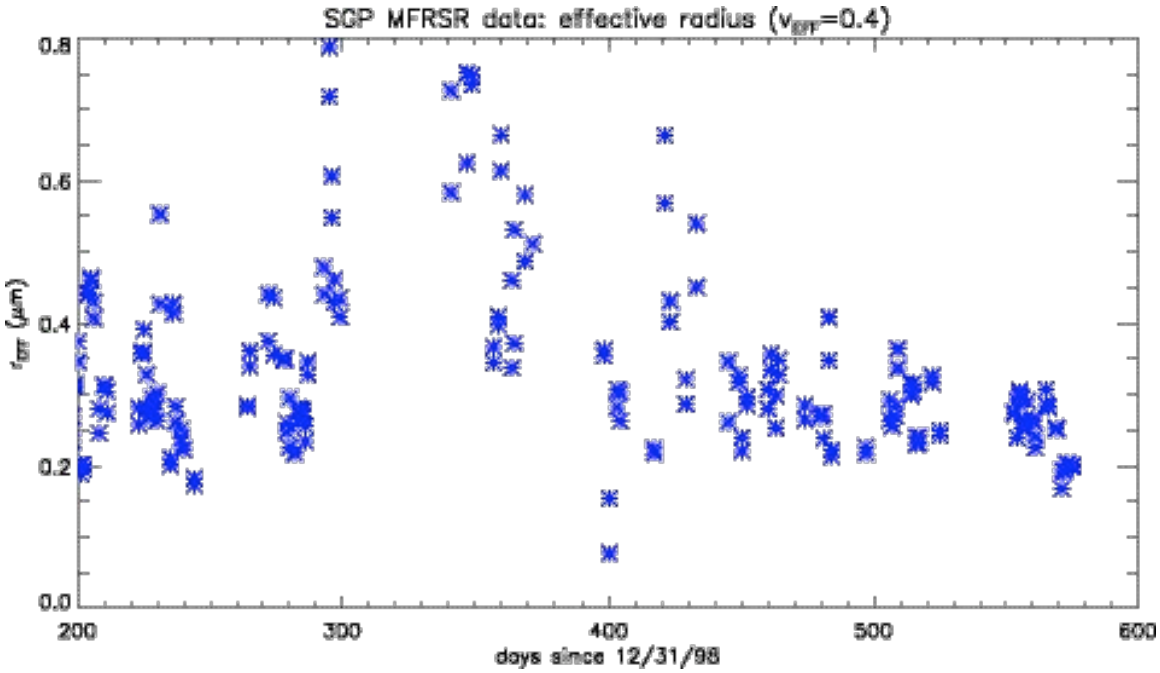

Figure 4.27: The same as Figure 4.26, except for a variance of 0.4. While lower on the whole, the radius values fluctuate significantly in the fall and the winter.

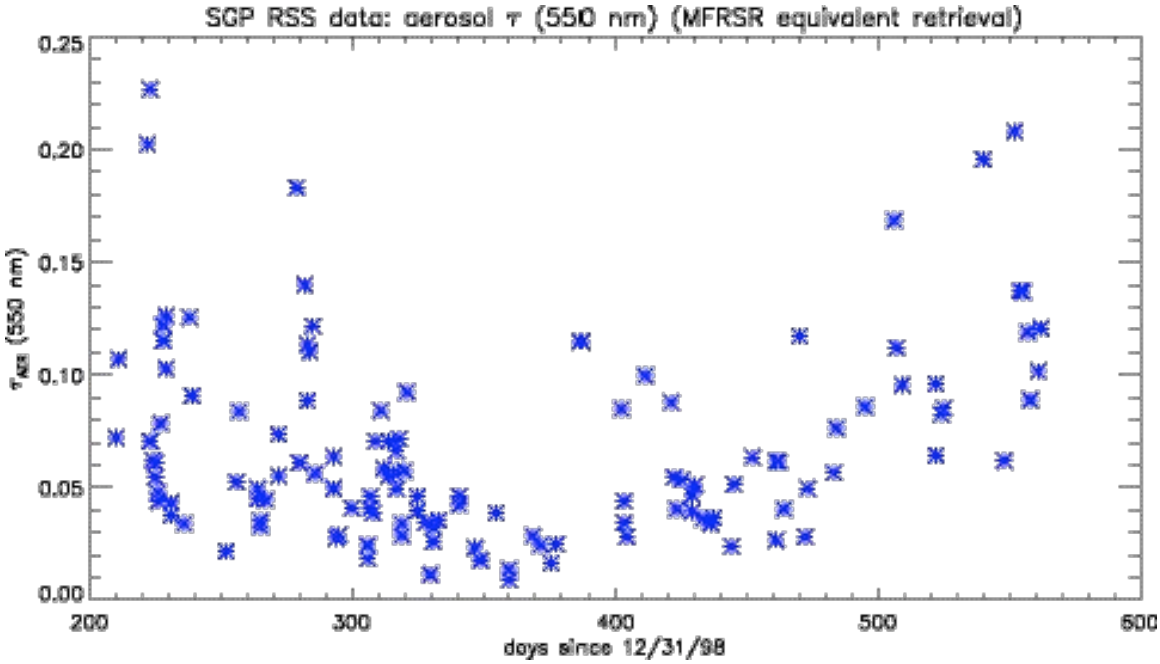

Figure 4.28: The plot of aerosol optical depth at 550 nm vs. day in the RSS "MFRSR equivalent" retrieval.

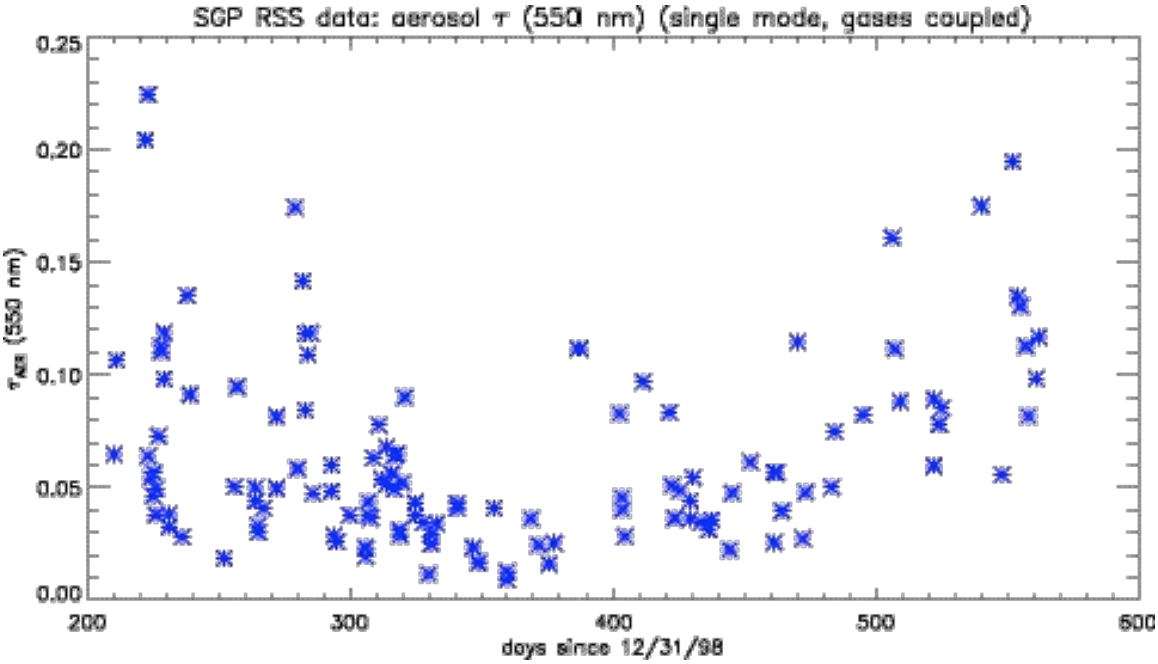

Figure 4.29: The same as Figure 4.28, but for the sixteen-channel retrieval that simultaneously retrieves aerosols and gases.

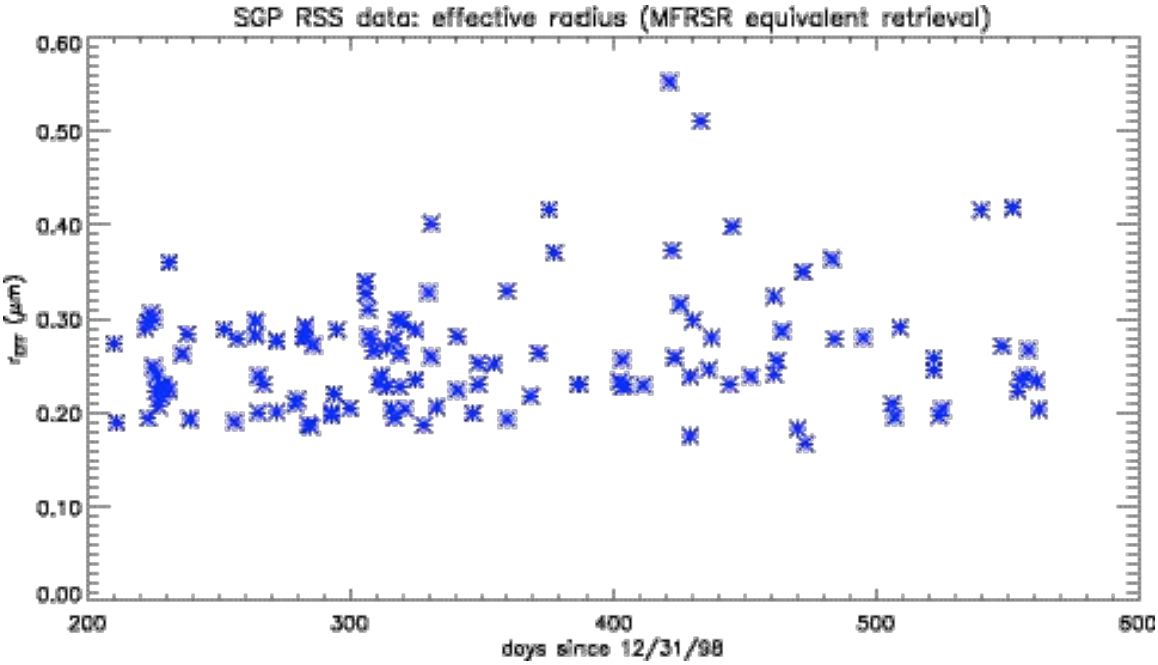

Figure 4.30: The plot of aerosol effective radius vs. day in the RSS "MFRSR equivalent" retrieval.

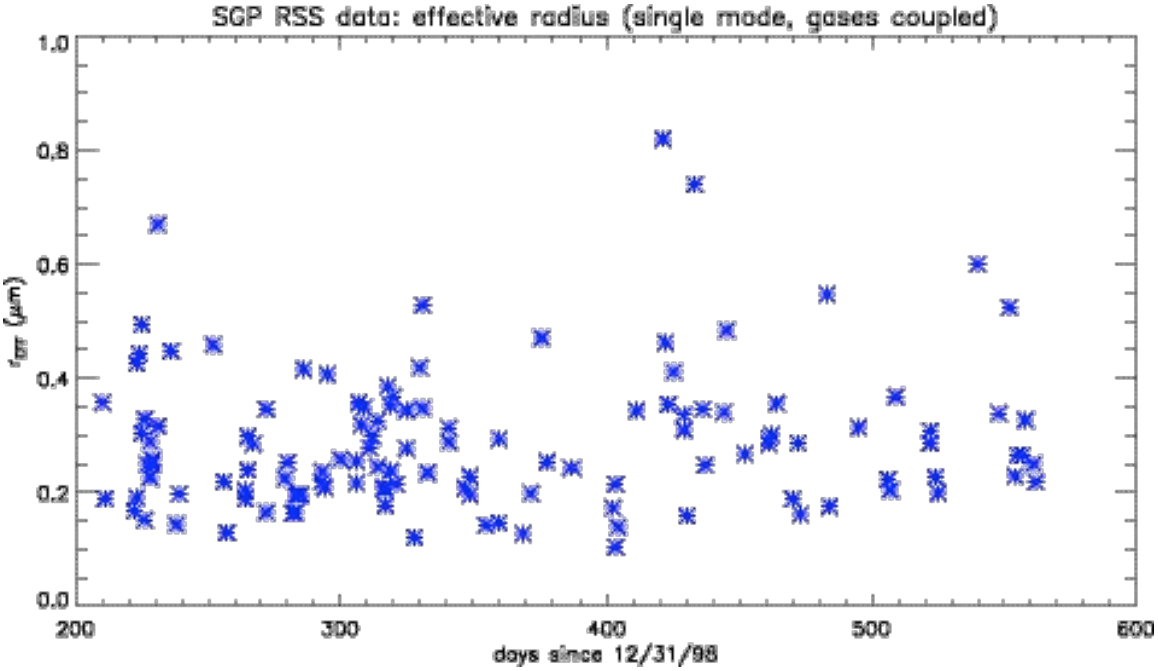

Figure 4.31: The same as Figure 4.30, but for the sixteen-channel retrieval that simultaneously retrieves aerosols and gases.

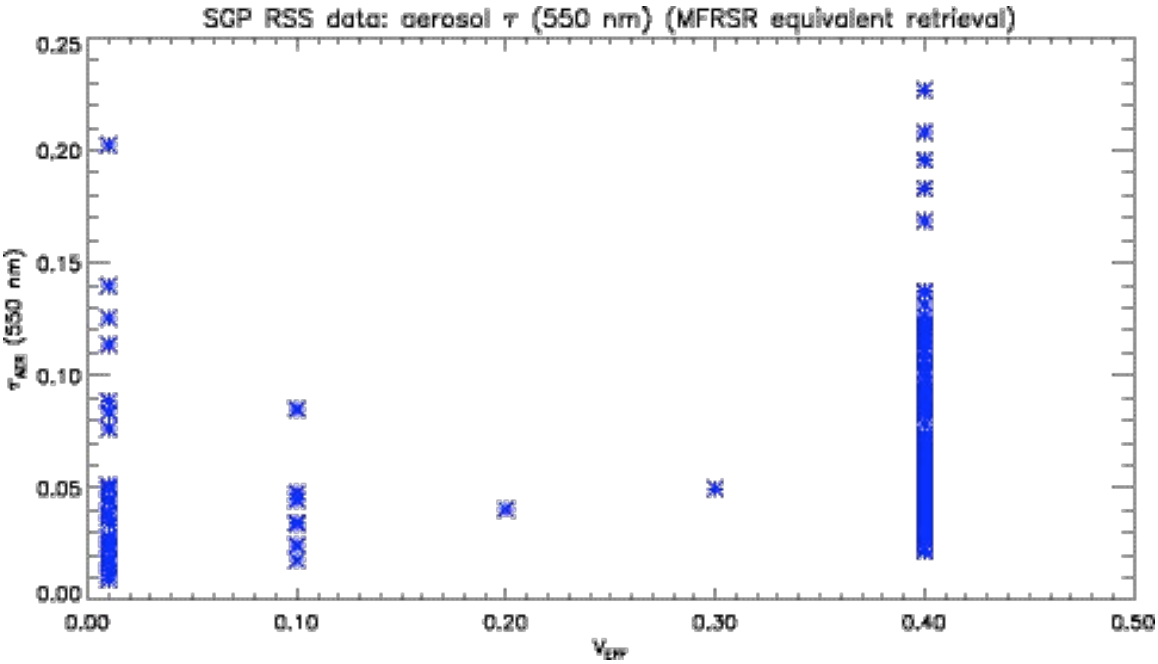

Figure 4.32: The plots of retrieved aerosol optical depth vs. retrieved effective variance for the "MFRSR equivalent" retrievals.

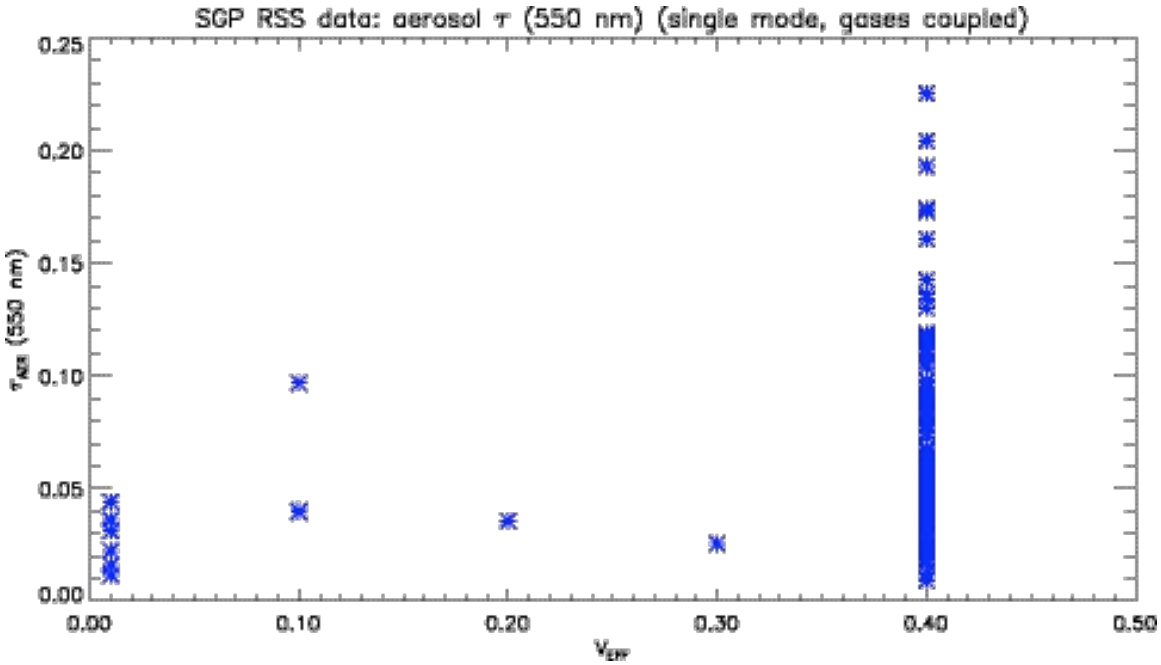

Figure 4.33: The same as Figure 4.32, except for the sixteen-channel, coupled-gas retrieval. The lower variances occur less often when the wavelength range is broadened and more spectral detail is added.

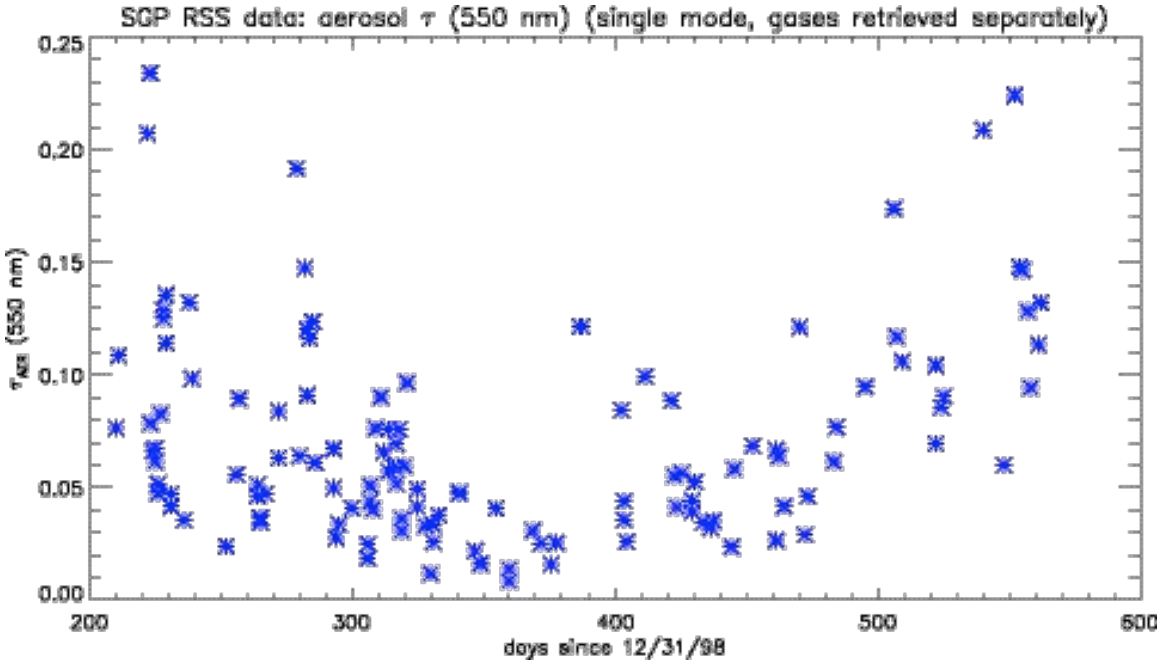

Figure 4.34: The plot of aerosol optical depth vs. day for the RSS when the gases are retrieved separately from the aerosols, using high-resolution techniques.

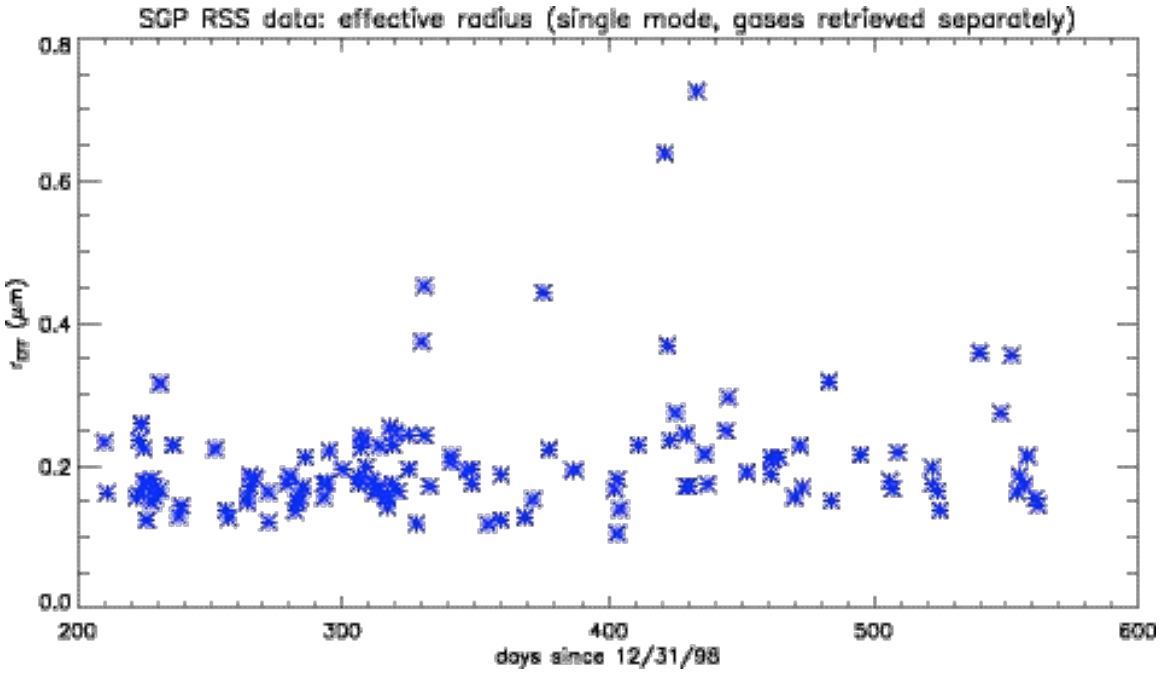

Figure 4.35: The plot of effective radius vs. day for the RSS when the gases are retrieved separately from the aerosols, using high-resolution techniques.

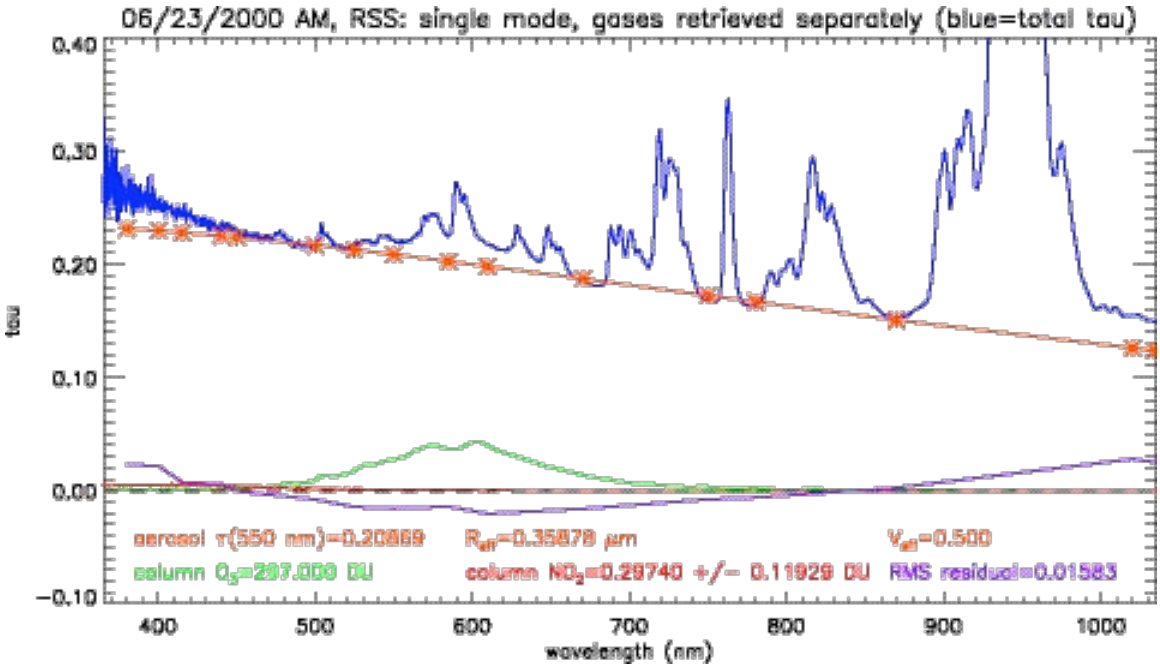

Figure 4.36: The results of the single-mode, gases retrieved separately retrieval for the morning of June 23, 2000. The combined optical depth of the retrieved quantities is too low in the extreme wavelengths, and too high in the middle wavelengths, indicating that an additional aerosol curve is necessary to optimize the fit.

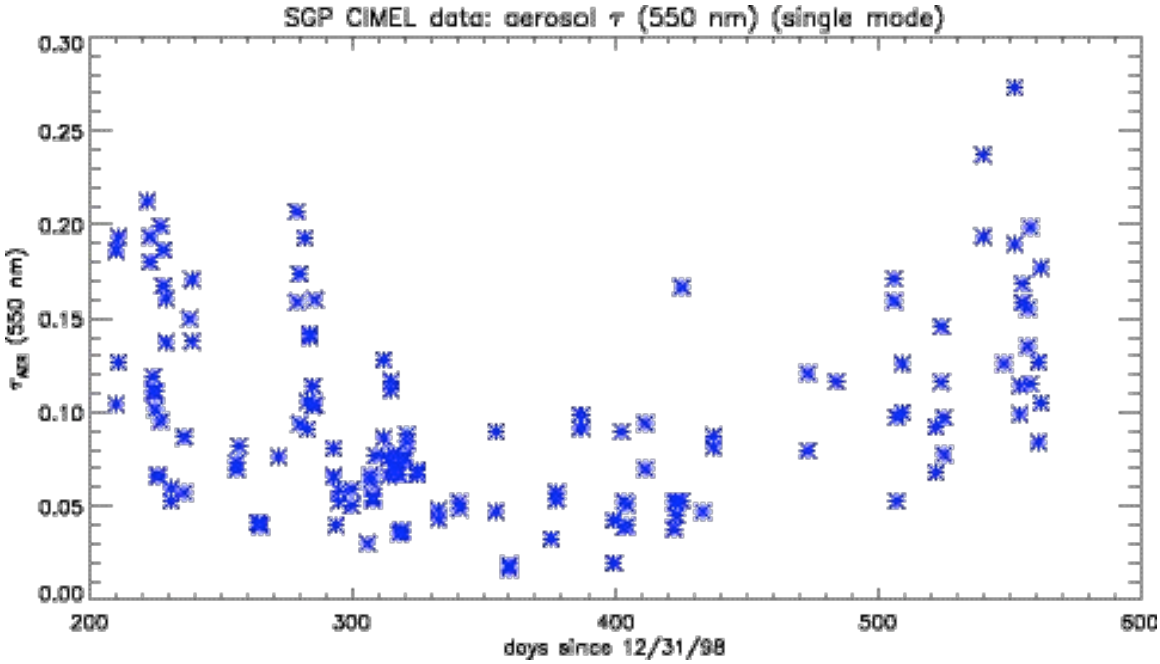

Figure 4.37: The plot vs. day of the aerosol optical depth at 550 nm, retrieved from the CIMEL data without the 340 nm channel, and assuming a single-mode size distribution.

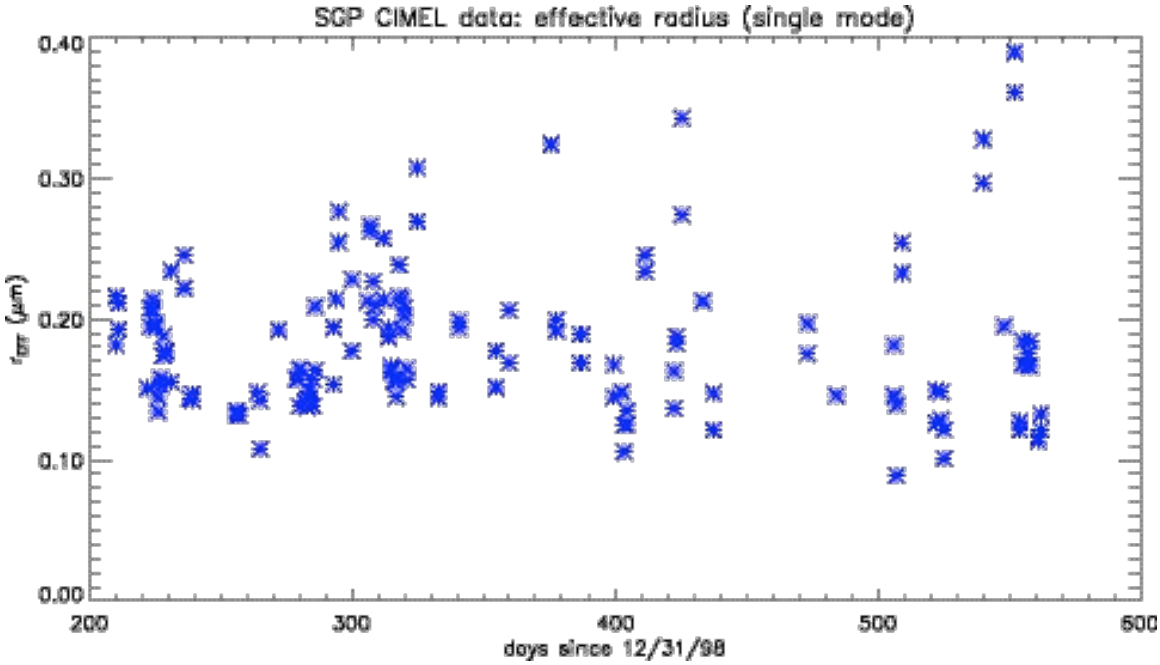

Figure 4.38: The effective radius values retrieved for the CIMEL data, excluding the 340 nm channel and assuming a single-mode size distribution.

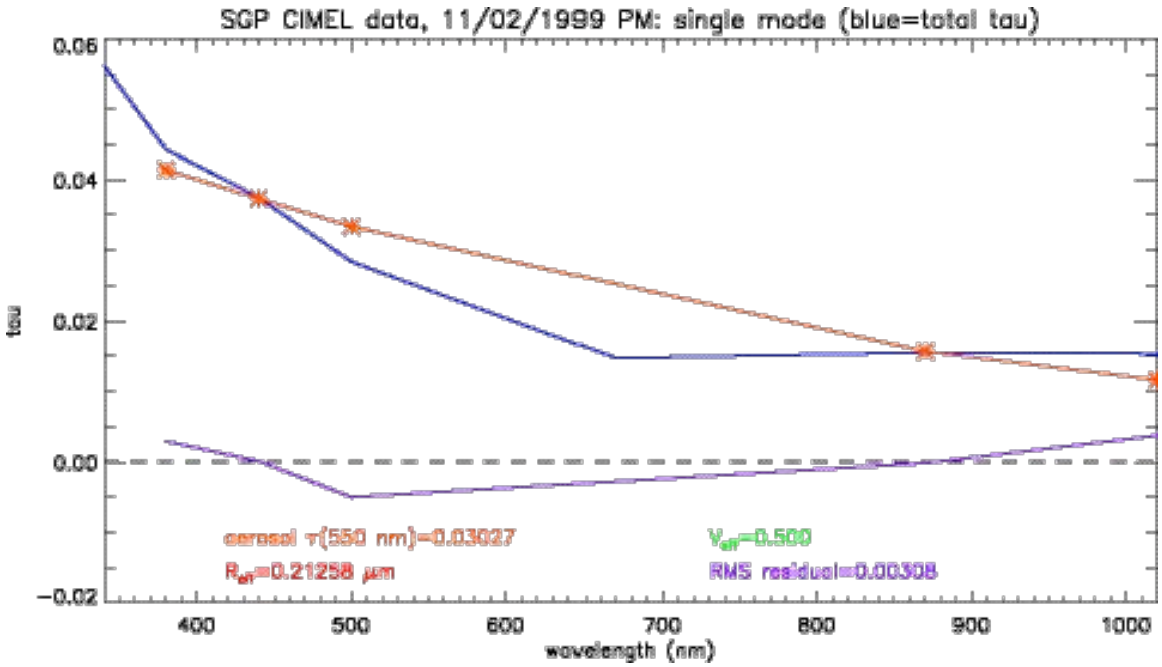

Figure 4.39: The single-mode retrieval for the afternoon of November 2, 1999. As was the case with the single-mode RSS retrievals, the pattern of positive residuals on the end and negative residuals in the middle repeats every day in the data set, suggesting that more than one aerosol mode is necessary to fully explain the extinction.

Chapter 4 References

Alexandrov, M. D., Lacis, A. A., Carlson, B. E., and Cairns, B., *Journal of the Atmospheric Sciences* **59**, 524-543 (2002a).

Box, G. P., Box, M. A., and Krücker, J., *Journal of Geophysical Research* **101**, 19211-19214 (1996).

Bramstedt, K., Gleason, J., Loyola, D., Thomas, W., Bracher, A., Weber, M., and Burrows, J. P., *Atmospheric Chemistry and Physics Discussions* **2**, 1131-1157 (2002).

Taha, G., and Box, G. P., *Geophysical Research Letters* **26**, 3085-3088 (1999).

Velders, G. J. M., Granier, C., Portmann, R. W., Pfeilsticker, K., Wenig, M., Wagner, T., Platt, U., Richter, A., and Burrows, J. P., *Journal of Geophysical Research* **106**, 12643-12660 (2001).

Vountas, M., Rozanov, V. V., and Burrows, J. P., *Journal of Quantitative Spectroscopy and Radiative Transfer* **60**, 943-961 (1998).

Chapter 5: The Bimodal Retrievals

5.1: Introduction

The results of the EOF analysis suggest that the aerosol size distribution derived from the data is bimodal in nature. The algorithms that retrieve aerosols and gases simultaneously while assuming a single-mode size distribution clearly failed to accurately retrieve the column amounts of the gases, especially nitrogen dioxide, because the algorithm mistakenly attributes some aerosol extinction to the gases. When gas amounts were retrieved as accurately as possible using the higher resolution of the RSS data, it became clear that the single-mode aerosol retrievals produced a characteristic, repeating pattern in the daily residual optical depth values that could be easily explained if the size distribution were actually bimodal, confirming the results of the EOF analysis. In this chapter, retrievals that assume a bimodal aerosol distribution are performed on the optical depth data from both the RSS and CIMEL devices. In the case of the RSS, the channels used are the same as those used in the sixteen-channel, single-mode retrieval.

From the EOF analysis, it is anticipated that the ability to retrieve a specific size distribution will be limited for both modes, but more so for the coarse mode. Values for the fine mode effective radius, at least, appear to be retrievable. Regardless of the device, all the Mie scattering fits to the first EOF's of the data sets after the mean is subtracted suggest that the value fine mode effective radius will generally fall between 0.1 and 0.2 μm. However, the results show a strong dependence on which effective variance produced the best fit. A high retrieved effective variance results in a low retrieved

effective radius, and vice versa. It is likely that this result does not reflect physical reality, but an inability of the retrieval algorithms to uniquely differentiate between the two retrieved quantities. The EOF's suggest that the coarse mode can be slightly more clearly defined in the RSS data than in the data from the other two devices, but the picture of the coarse mode obtainable even from the high-resolution RSS appears to be limited by the wavelength range. It is not clear that either the effective radius or the effective variance can be retrieved with any sort of confidence for the coarse mode.

In addition to performing bimodal retrievals on the CIMEL optical depth data, the CIMEL size distributions derived from the almucantar data (Dubovik and King 2000) are examined and compared with the RSS results. The Dubovik and King algorithm assumes a bimodal, if not multi-modal, aerosol size distribution to begin with, and on that basis has been shown to be more accurate than the MFRSR algorithm of Alexandrov *et al.* (2002a). The size distributions obtained using this algorithm can be converted into effective radius and variance values, enabling a more direct comparison with the other bimodal retrievals used in this study. A few things need to be considered regarding the almucantar-based size distribution retrievals, though. These retrievals look at both almucantar and direct sky optical depth data from the 440, 670, 870, and 1020 nm channels. Therefore, the channel that most defines the fine mode is the 440 nm channel, which is potentially sensitive to absorption not only from NO_2, but from a weak water vapor absorption band as well. If both gases are neglected, the aerosol optical depth will be overestimated in this channel, and the retrieved fine mode effective radius will correspondingly be underestimated. Also, given the same limited wavelength range as the other retrieval algorithm used in this study, it does not appear likely that the coarse

mode will be clearly defined in the almucantar algorithms, either. Finally, the EOF analysis of the CIMEL optical depth data uncovered a defect in the data that affected all of the channels and could conceivably skew the results of the retrievals.

To conclude this chapter, the bimodal retrievals are re-performed on the RSS data, this time assuming constant values for the fine mode effective variance and the coarse mode effective radius and variance, instead of trying to retrieve independent values for all of these quantities. If the amount of retrievable information is limited, then attempting to retrieve information in excess of this limit will not only produce spurious information, but could potentially mask important information contained within the data. For example, if the effective radius and variance cannot be independently retrieved for the fine mode, then a plot of the retrieved values of the fine mode effective radius vs. day will be influenced by the days on which the values of the fine mode effective variance are retrieved to be low, and on which they are high. A trend in the graph, or the absence thereof, could prove to be misleading.

5.2: Results of the RSS Bimodal Retrieval

5.2.1: Non-Uniqueness in the Coarse Mode Solutions

Figures 5.1 and 5.2 show contour plots of the residual optical depth as a function of both the coarse mode and fine mode effective radius. For the sake of simplicity, only the results for coarse and fine mode effective variance values of 0.1 are presented in this fashion. The contour plot for June 23, 2000, shown in Figure 5.1, indicates multiple

regions of convergence where a relative minimum residual is located. The fine mode effective radius for these relative minima consistently falls between 0.15 and 0.16 μm. The coarse mode effective radius has its first minimum slightly below 2 μm, but relative minima extend at least as far as 5 μm. The retrieval algorithm selects the lowest residual from among these choices, but the differences in the relative minimum residual values are too small to have confidence in the uniqueness of the final result. In addition to the series of relative minima, another individual minimum is located at a fine mode radius of 0.21 μm and a coarse mode radius of about 0.9 μm. Here, the relative minimum residual significantly exceeds the other minima. However, as Figure 5.2 shows for the afternoon of August 11, 1999, sometimes this isolated minimum produces an unambiguous best fit. It is not clear, at present, if this change in the results reflects a real change in the aerosols, or is an artifact of the algorithm. At any rate, the potentially large error in the coarse mode effective radius results in an error of ± 0.003 in the fine mode optical depth relative to the coarse mode optical depth, ignoring calibration uncertainties.

5.2.2: High-Resolution, Gases Retrieved Separately

A graphical illustration of the results of the RSS bimodal retrieval, using the high-resolution retrievals for nitrogen dioxide and ozone, is presented in Figure 5.3 for the morning of June 23, 2000. Recall that in Figure 4.15, the single-mode algorithms that retrieve aerosols and gases simultaneously produced best fits to the data that greatly overestimated the amount of NO_2 on that day. When the high-resolution gas retrievals are applied for this day, as shown in Figure 4.36, the aerosol extinction is underestimated

at both the shortest and longest wavelengths, while overestimated in the middle. Clearly, any single-mode retrieval would not produce accurate results for this particular day. The defining characteristic in the bimodal retrieval for this particular morning, almost unique in the entire data set, is that the coarse mode aerosol optical depth overwhelms that of the fine mode. The RMS residual is far lower than it was for any of the single-mode retrievals for that day, and does not show any clearly defined wavelength dependence.

Figure 5.4 shows the plot vs. day of the total aerosol optical depth – that is, the sum of the coarse mode and fine mode optical depths – measured at 550 nm by the RSS for the bimodal retrieval with the high-resolution gas amounts. The pattern shows the same temporal variations already observed, including the seasonal cycle and the series of peaks in October 1999. The fine mode optical depth, plotted vs. day in Figure 5.5, mostly follows the same pattern as the total optical depth, but conspicuous differences occur at June 23 and July 5, 2000. The large values of total optical depth on these days are not reflected by corresponding large values for the fine mode optical depth. Most notably, these differences bring the plot in Figure 5.5 into greater agreement with the first EOF's of the full RSS data, both for the base set (Figure 3.3) and the set with the mean subtracted (Figure 3.11). This supports the notion that EOF's provide information about the aerosol size distribution, and that at least some of this information deals with the two aerosol modes. In particular, it looks very likely that the coefficients of projection in Figure 3.11 really do represent a fine aerosol mode. On average, according to the retrieval results, the fine mode accounts for approximately 70% of the total aerosol optical depth at 550 nm.

Not surprisingly, the peaks that are missing from the plot of the fine mode optical depth show up in the plot of the coarse mode optical depth, presented in Figure 5.6. The coarse mode optical depth generally follows the same seasonal cycle that the fine mode optical depth does. Like the fine mode, the coarse mode has a handful of days in August and October where the optical depth rises above the background level. The coarse mode optical depth does not show the same degree of variability that the fine mode does, however. Aside from June 23 and July 5, only February 25 has a coarse mode optical depth value in excess of 0.05. A handful of days have coarse mode optical depth values that approach zero, but for no day was the retrieved value exactly zero or negative.

The fine mode effective radius is plotted vs. day in Figure 5.7. Aside from a few particularly large values, the most significant feature of the graph is that the values are significantly larger in 2000 than 1999, especially after February 14, 2000 (Day 410). On many days before February 14, the retrieved fine mode effective radius values do not exceed 0.13 μm. This value is exceeded on most of the subsequent days, however, including every single case for over 100 days after February 14. It is quite possible, but not entirely clear, that the absence of low-radius days from February to May indicates a peak in a seasonal cycle. The large variability in the results for the rest of the data set unfortunately prevents drawing a firm conclusion.

One of the conclusions drawn from the EOF analysis, based on the Mie fits to the coefficients when the annual mean was subtracted, was that the effective radius values would generally fall between 0.1 and 0.2 μm, with larger values of the effective radius corresponding to smaller values of the effective variance. The mean fine mode radius, 0.164 μm, is very similar to the value of 0.160 μm obtained from the first EOF with the

mean subtracted (see Figure 3.13). The closeness of these values provides additional support to the claim that the first EOF with the mean subtracted represents a fine aerosol mode. However, the value in Figure 3.13 corresponds to the lowest possible effective variance, 0.1, while the fine mode effective variance was allowed to vary in the bimodal retrieval. In order to determine if the effective radius and effective variance can actually be retrieved independently of each other, or if a shortage of information inherent in the data prevents the two quantities from being uniquely and independently constrained, the radius values are plotted against the variance values. Figure 5.8 clearly shows that the retrieved values of the fine mode effective radius and effective variance are interdependent. The overwhelming majority of effective radius values below 0.1 μm also have the highest variance, 0.5. In addition, all the effective radius values above 0.25 μm have the lowest variance, 0.1. This result most likely does not provide information about the actual nature of the aerosol size distribution, but rather the theoretical limits to the amount of retrievable aerosol information over a relatively narrow wavelength range, as indicated by Box *et al.* (1996). More than half of the days in the data set produce an effective variance of 0.1, with a variance of 0.5 occurring an additional 20% of the time. The predominance of the low value of the effective variance does not necessarily reflect physical reality, however.

In the plot of the coarse mode effective radius vs. day, shown in Figure 5.9, the retrieved values vary daily to a much greater degree than the retrieved fine mode radius values do. Most of the retrieved values fall between 1.1 and 1.3 μm. However, several days have values exceeding 2 μm and a handful of retrieved effective radius values fall below 1 μm as well. Curiously, the four days with values that exceed 3 μm are very

close to 5 μm, with no values at all retrieved between 2.7 and 4.7 μm. These large values lift the mean retrieved coarse mode effective radius up to nearly 1.4 μm.

Perhaps the coarse mode really is this inconsistent, or perhaps the mode is simply less clearly defined than the fine mode, due to the aforementioned limits to the completeness of the retrieval. On one hand, the relatively short lifespan in the air of large aerosol particles could cause greater variability in what is observed over a particular site, given a number of source locations for the aerosols at varying distances from the site, and changing wind speed and direction. On the other hand, it certainly does not seem likely that the aerosol mass over the SGP site could, on a small number of days, contain a coarse mode whose effective radius exceeds the largest value generally observed by a full 2 μm. The second eigenvectors in the sets of EOF's with the mean subtracted did indicate the presence of a coarse mode, but the coarse mode did not appear to be clearly defined in the EOF's, even for the RSS. As the theoretical analysis of Box *et al.* suggested that a clear analysis of the coarse aerosol mode required the inclusion of longer wavelengths in the analyzing device, it is possible that analyzing the coarse mode with the RSS necessarily yields imperfect results.

The retrieved effective variance has a value of 0.1 more than three quarters of the time. In this regard, at least, the RSS analysis of the coarse mode retains a strong degree of consistency. As Figure 5.10 shows, any interdependence between coarse mode effective radius and variance is not as apparent as it is for the fine mode. Most, but not all, of the days where the coarse mode effective radius does not exceed 1 μm have a high variance. However, many days with a large retrieved coarse mode effective radiance have a high variance as well. Like the fine mode, the retrieved coarse mode effective

variance is usually 0.5 when it is not 0.1. There is no obvious physical reason why this should be the case.

5.2.3: Gases Coupled with Aerosols

A bimodal retrieval was also performed on RSS data with the gases and aerosols retrieved simultaneously, in order to examine the degree to which improving the assumptions in the retrieval improves the retrieved values of the gas amounts, and aerosol size distributions, in low-resolution retrievals. The results of this sixteen-channel, coupled gas retrieval should not be as accurate, especially for the gas amounts, but they should establish a limit to the accuracy that can be expected of five-channel "MFRSR equivalent" retrievals that assume a bimodal distribution. The algorithms that assume a single-mode aerosol size distribution clearly need to be modified. By using the RSS, the ability of a five-channel bimodal retrieval to reproduce what can be done using sixteen channels and high resolution can be tested. The bimodal "MFRSR equivalent" retrieval can show specifically the limit to which the standard MFRSR wavelengths can reproduce accurate aerosol results. To see if a different combination of five channels would improve the results, the five-channel "MFRSR equivalent" algorithm can easily be modified to accommodate a new set of wavelengths. If the results for this new retrieval improve the agreement with the results of the sixteen-channel algorithm, then an MFRSR equipped with filters at this set of wavelengths will produce better results than the existing MFRSR can do.

Table 5.1 presents the mean values of the retrieved quantities from the sixteen-channel RSS bimodal retrievals. Similarly to Table 4.1, the dominant, or most common, values of the coarse and fine mode effective variance are shown, along with the percentage of days on which these values were obtained. For the sake of comparison, the mean values for the sixteen-channel single-mode retrievals are also presented. The mean values of the coarse and fine mode optical depths do not change significantly between the bimodal retrieval with gases coupled and with gases retrieved separately. In each retrieval, the dominant effective variance for both modes is 0.1. The low variance values occur with similar regularity in the fine mode, and exactly the same regularity in the coarse mode. Due to the improved assumptions inherent in the bimodal coupled-gas retrieval, the mean values of ozone and nitrogen dioxide more closely match those obtained using high resolution than the values obtained from the single-mode, coupled-gas retrieval do. The agreement is still not exact, however. The column NO_2 amount is still overestimated on the average, causing a corresponding overestimation in the fine mode effective radius.

In Figure 5.11, the retrieved quantities in the bimodal coupled-gas retrieval are plotted against the corresponding quantities in the separate-gas retrieval. For both aerosol modes, the optical depth values at 550 nm correlate extremely well, indicating that high-resolution retrievals are not necessary to retrieve at least these quantities accurately. Aside from the slight overestimation already noted, the fine mode effective radius values correlate acceptably well with each other. The ozone correlation likewise shows acceptable correlation, excepting one data point where the value is badly underestimated in the coupled-gas retrieval. Unfortunately, the correlation for the coarse

mode effective radius is poor. The days in which a large effective radius value is retrieved in one of the retrievals simply do not correspond to the high radius days in the other retrieval at all. This strongly suggests that even though the coarse mode optical depth appears to be measurable, the accuracy of the retrieved values of the coarse mode effective radius cannot be established. Despite close means, the values of NO_2 for the two different retrievals essentially do not correlate at all. The NO_2 values in the coupled-gas retrieval still correlate somewhat with aerosol optical depth (r^2=0.23), indicating that the trade-off between NO_2 absorption and small aerosol extinction at the short wavelengths does not entirely vanish for the coupled-gas retrieval, even when a bimodal size distribution is assumed.

5.3: The Bimodal Retrieval of the CIMEL Optical Depth Data

Figure 5.12 shows the plot vs. day of the total aerosol optical depth retrieved with the CIMEL optical depth data only, assuming a bimodal aerosol size distribution. Despite being larger on the whole than the RSS, the CIMEL optical depth values have all the same features, including the seasonal cycle, the series of large values in August and October of 1999, and the particularly large values on June 23 and July 5, 2000. Much like with the RSS, the optical depth of the fine mode measured by the CIMEL mostly, but not entirely, mirrors the pattern of the total optical depth. Once again, as Figure 5.13 shows, June 23 and July 5 do not have high fine mode optical depth values, and therefore mark the two days in the data set where the coarse mode optical depth dominated the total amount. The fact that both the RSS and the CIMEL produce this result confirms that an

aerosol mass consisting mostly of large particles passed over the SGP site on those two days. In the CIMEL data, the mean fine mode optical depth at 550 nm is 0.058, corresponding to 68.8% of the total aerosol optical depth. This percentage also shows good agreement with the RSS, despite the disparity in optical depth. The coarse mode aerosol optical depth is plotted vs. day in Figure 5.14. The CIMEL values differ from the RSS values only in that the CIMEL has usable data for both the morning and afternoon of June 23 and July 5. The RSS, by contrast, only has data for the mornings on those two days. According to the CIMEL results, the coarse mode optical depth remains significantly high throughout both days. The CIMEL and the RSS both have a high coarse mode optical depth value in late February, 2000. The high RSS value occurs on the morning of February 25 (day 421), for which the CIMEL has no corresponding data. Similarly, the high value in the CIMEL data set occurs on the afternoon of February 29, for which the RSS has no corresponding data. These differences could result from differences in the cloud-screening processes that produce the optical depth data for each device. Cloud particles tend to be much larger than 1 μm, but at the wavelength range of these devices, they do not produce a significantly different spectral signature than coarse mode aerosols do. It is therefore possible that the coarse mode aerosol retrieved on these days actually included some cloud particles. If the accuracy of retrievals of the coarse mode size distribution cannot be established, then the question of whether the aerosol data is being somewhat skewed by clouds must necessarily remain open.

The third EOF of the CIMEL data, shown in Figure 3.27, exhibits a peculiar pattern when plotted vs. day. The EOF values drop fairly steadily, until about day 500 (May 14, 2000), where a sharp discontinuity appears to occur. In the plot of the

coefficients of projections vs. wavelength for this eigenvector, the 670 nm channel gives the highest value, indicating that this eigenvector correspond to the sharp dip in optical depth values observed in this channel on many of the days. However, the pattern in the coefficients suggests that whatever affected the 670 nm channel also affected the other channels somewhat as well, and would likely interfere with the retrievals obtained from CIMEL data even if, as was done in this study, the data from the 670 nm channel is discarded. No similar pattern was observed in the single-mode CIMEL retrievals for either the aerosol optical depth (Figure 4.37), or the effective radius (Figure 4.38). However, when the bimodal retrieval is performed, the discontinuity is observed quite strongly in the retrieved values of the fine mode effective radius, as seen in Figure 5.15. The retrieved fine mode effective radius values do not show an obvious trend before day 500, but no value exceeds 0.1 μm afterwards. The discontinuity could conceivably result from a correction to the device, but the retrieved effective radius values are less plausible after the discontinuity than before. Unfortunately, this renders impossible any attempt to analyze the CIMEL data for any seasonal cycles in the retrieved value of the fine mode effective radius. As Figure 5.16 shows, the fine mode effective radius values show the same interdependence with the effective variance values seen in the RSS data.

Even more disturbingly, a distinct linear trend is observed in the plot of the coarse mode effective radius vs. day, shown in Figure 5.17. On the positive side, the CIMEL retrieval shows much less variability in the coarse mode effective radius than the RSS does. Most of the early values fall between 1.1 and 1.3 μm, just like the RSS values do. Unfortunately, the downward trend persists even after the discontinuity, and appears to be an artifact of the device. It is therefore concluded that the CIMEL optical depth

data, for the SGP site during this time period, is not reliable enough to use in evaluating the specifics of the aerosol size distribution. While it became clear from looking at a couple of the days individually that at least one channel in the CIMEL device had a defect, the EOF analysis first hinted at the full extent of the problem. Perhaps, then, a preliminary EOF analysis should be performed on all sets of sun photometer data, in order to detect any problems inherent in the data before attempting any retrievals.

5.4: Results of the CIMEL Almucantar Retrievals

The AERONET aerosol size distributions are obtained by applying the algorithm of Dubovik and King (2000) to the optical depth and almucantar values obtained from the 440, 670, 870, and 1020 nm channels of the CIMEL devices. Unfortunately, it has already been established that the CIMEL optical depth data for the SGP site during the period used for this study is skewed by at least one defect, either in the filters or in the detector, that manifests itself in altered effective radius values in the retrieval using the optical depth only. The results of the almucantar retrievals therefore have similar limits to their applicability. Still, there are reasons to think that this almucantar data can still provide enough meaningful information to justify inclusion in this comparative analysis. Unlike the MFRSR and RSS, the CIMEL measures direct and scattered radiation with different detectors, and if the second detector measures the almucantar data properly, the skew in the results will be somewhat lessened. Based on what was observed in the retrievals of just the optical depth data, the almucantar retrievals should at least show qualitative agreement with the RSS where the aerosol optical depth is concerned. Also, if

the almucantar retrievals suffer from the same shortage of obtainable information inherent in the RSS and CIMEL optical depth retrievals, then the results should show a similar interdependence between the retrieved effective radius and the effective variance for each aerosol mode. The errors caused by the defect in the optical depth data will most likely manifest themselves once again as artificial trends in the retrieved values of the coarse and fine mode effective radius, however.

As an example of the output of the almucantar-based size distribution retrieval, the graph of the size distribution measured at 2:35 PM on July 5, 2000, plotted as a function of particle size, is shown in Figure 5.18. The x-axis is the particle radius, in microns, plotted on a logarithmic scale. The y-axis is the derivative of the volume of particles at a given radius (in μm^3) in an atmospheric column, per area (in μm^2) of the column, as a function of the logarithm of the radius. While the size distribution, presented in this form, does show a dominant coarse mode in agreement with the RSS and CIMEL optical depth retrievals for that day, a direct comparison between the almucantar and optical depth retrievals cannot be made with the almucantar size distribution data in this form. However, the almucantar-retrieved size distributions can be converted to coarse and fine mode optical depth, effective radius, and effective variance values using numerical integration. First, the particle sizes below 0.5 μm are considered to be part of the fine mode, while those above 0.5 μm are grouped as the coarse mode. For a given mode, the optical depth can be defined in differential form by

$$d\tau = \pi r^2 Q_x dN(r) = \frac{3}{4r} Q_x \frac{dV}{d\ln r} d\ln r,$$

where τ is the optical depth, r is the particle radius, Q_x is the extinction efficiency factor of the aerosol determined by Mie scattering, and N equals the number of particles in an atmospheric column per cross-sectional area. The derivative dN equals $3dV/4\pi r^3$, the total differential volume of the particles divided by the volume of a particular particle. As dV/d(ln r) is a measured quantity, the optical depth can be calculated by numerically integrating over the logarithm of r. The effective radius and effective variance are calculated numerically using the equations already given in Chapter 1,

$$r_{eff} = \frac{\int_0^{\infty} r \pi r^2 n(r) dr}{\int_0^{\infty} \pi r^2 n(r) dr}, \text{ and}$$

$$v_{eff} = \frac{\int_0^{\infty} (r - r_{eff})^2 \pi r^2 n(r) dr}{r_{eff}^2 \int_0^{\infty} \pi r^2 n(r) dr}.$$

The plot vs. day of the fine mode optical depth at 550 nm is shown in Figure 5.19. This data set extends slightly beyond the range of the RSS data set, and includes some very large values on either end. In addition, there is a particularly large peak on day 417, or the afternoon of February 21. The RSS and CIMEL optical depth data do not produce anything of comparable magnitude on this day. While the almucantar-retrieved fine mode optical depth values greatly exceed those obtained solely from the optical depth measurements, the same seasonal patterns are observed, along with the series of high optical depth values in October 1999. Figure 5.20 indicates that even though the coarse mode optical depth peaks on the same day (July 5) for the almucantar as it does for the

CIMEL optical depth measurements, the peak value is about half of what was obtained previously. In general, the coarse mode optical depth values are slightly smaller for the almucantar retrievals. No almucantar retrieval was performed for June 23. The differences in the retrieved coarse and fine mode optical depth measurements become even more pronounced when the retrieved from the almucantar and optical depth values are plotted vs. the values retrieved from the optical depth values alone, for the days common to both retrievals. As Figure 5.21 shows, the fine mode optical depth values obtained from the almucantar results are consistently much larger, by more than 35%. The relationship between the two quantities is essentially linear, however, which turns out not to be the case for the coarse mode optical depth, as seen in Figure 5.22. The values retrieved by using the optical depth data alone are larger in general, and become much larger as the optical depth increases. When the total optical depth is examined, as in Figure 5.23, the almucantar data remains larger by about 20%, with an additional offset of about 0.02. The relationship between the measurements appears to be linear overall. Given the known errors affecting the optical depth values used in the almucantar size distribution retrieval, the level of agreement between the two retrievals is acceptable. Still, very little can be said conclusively in a quantitative sense about the coarse and fine optical depths measured by the CIMEL during the span of the data set.

The almucantar-retrieved values for the effective radius and variance show some interesting similarities and differences with the corresponding values retrieved from the RSS data. Figure 5.24 shows the fine mode effective radius plotted vs. day. All of the values fall between 0.1 and 0.2 μm, in contrast to the RSS results presented in Figure 5.7, where a handful of days have very large retrieved radius values. Perhaps the Dubovik

and King (2000) algorithm produces more consistent fine mode retrievals, presumably by better distinguishing the contributions of the coarse and fine aerosol modes. The mean almucantar-derived fine mode effective radius, 0.134 μm, is lower than the mean for the RSS. A number of factors could contribute to this difference, including the unaccounted-for absorption of NO_2 and H_2O in the 440 nm channel. The pattern in Figure 5.24 even suggests a seasonal cycle, with high values around day 500 (May 14, 200) and low values around day 300 (October 27, 1999). Unfortunately, whatever cycle exists in the RSS data does not emerge clearly in Figure 5.7, and the errors which affect the effective radius retrieval using only the CIMEL optical depth values (see Figure 5.15) probably affect Figure 5.24 as well. The amount of clear information available is once again too limited to draw firm conclusions about the seasonal behavior of the fine mode effective radius.

Even though the algorithm of Dubovik and King differs greatly from the other aerosol retrieval algorithms used in this study, it likely suffers from similar limits in retrievable aerosol information. In the optical depth retrievals of the RSS and CIMEL data, the information shortage manifests itself in an interdependence between the effective radius and effective variance of the fine mode. Figure 5.25 shows the plot of the fine mode effective radius vs. the fine mode effective variance in the almucantar. While the retrieved radius and variance do appear to depend on each other, the radius generally increases with the variance, in reverse to the RSS results in Figure 5.8. This indicates none of the different algorithms uniquely distinguish the effective radius and variance of the fine mode. Adding the scattered radiation in the almucantar measurements to the retrieval alters the interdependence, but does not eliminate it. Like

with the optical depth retrievals, the dominant fine mode effective variance is 0.1, occurring 62.4% of the time.

Figures 5.26 and 5.27 show the coarse mode effective radius plotted vs. day and effective variance, respectively. Just like in Figure 5.17, the corresponding plot for the CIMEL retrieval using only the optical depth values, the plot of the almucantar-retrieved coarse mode effective radius vs. day shows a fairly steady downward trend throughout most of the data set. The period after day 500 has some high values that halt this trend, even though most of the values below 2 μm occur after day 500 as well. In general, the coarse mode radius values greatly exceed those retrieved by either the RSS or the CIMEL optical depth values. The lowest values for the effective radius occur on and around July 5, where the coarse mode optical depth is highest both in the almucantar and the CIMEL direct sun data sets. A specific volume of particles at a lower radius will have a higher optical depth than the same volume of particles at a higher radius. The results of the almucantar retrieval therefore suggest that the coarse mode optical depth peaks on July 5 not because a greater volume of large particles passed over the SGP site that day, but because the aerosol mass consisted of a larger number of particles with a slightly smaller size than normal. No such connection is evident in the retrievals that use only optical depth data. Unfortunately, the defects in the optical depth data make it impossible to say if this is a reflection of the physical reality. Like the fine mode, the coarse mode shows some interdependence between the retrieved values of the effective radius and the effective variance, although the radius decreases with the variance in the coarse mode. In one case, the afternoon of February 19, the variance is very large, but the coarse mode as a whole on that day had extremely low values for dV/d(ln r).

At face value, the coarse modes retrieved from the optical depth values only and from the CIMEL optical depth and almucantar data look very different from each other. The mean effective radius retrieved with the Dubovik and King algorithm nearly doubles the mean retrieved from the RSS data, and the retrieved effective variance values also greatly exceed the RSS values. Before determining which retrieval approach best matches reality, it is necessary to examine how the different size distributions manifest themselves radiatively. This means calculating how the spectral signatures of the two distributions vary with wavelength. Size distributions having an effective radiance of 1.3 μm with a variance of 0.1 to represent a typical RSS-retrieved distribution, and a radius of 2.6 μm with a variance of 0.5 to represent the almucantar-retrieved size distribution, were input into a Mie scattering code. The resulting extinction coefficients, normalized to equal 1 at 550 nm, are plotted in Figure 5.28. The wavelengths plotted on the x-axis are the same ones used in the theoretical EOF analysis of Box *et al*. (1996). Below 800 nm, the curves are essentially identical. In other words, retrieving any details of the coarse mode size distribution using wavelengths below 870 nm is impossible. At 870 nm, the difference is at least discernible, with the value for an effective radius of 1.3 μm being slightly larger. This indicates that if bimodal fits using these distributions as the coarse mode were applied to the optical depth data of the MFRSR, or an "MFRSR equivalent" variant of the RSS data, the coarse mode optical depth at 550 nm using the 2.6 μm distribution would be larger, but the overall difference in the retrieval results would still be slight. In retrievals using MFRSR or RSS data, this difference in large particle aerosol extinction will affect the retrieved gas amounts and fine mode properties at least slightly, but not to a great enough degree to confirm that one coarse mode radius

is clearly erroneous relative to the other. The difference in the extinction coefficient increases at 1030 nm, suggesting that the RSS or the CIMEL would have a slightly better chance than the MFRSR of clearly favoring one distribution over the other, but the difference becomes much greater as larger wavelengths are added. This confirms what had been postulated in the Box *et al*. study, and strongly suggested in the EOF analysis of Chapter 3, that sun photometer devices will need wavelengths exceeding 1030 nm in order to clearly define the coarse aerosol mode. As it stands, the wavelength ranges currently used by the MFRSR, the RSS, and the CIMEL can only paint a partial picture of the aerosol extinction, and this limitation affects their ability to both precisely define the fine aerosol mode and uniquely distinguish between gas and aerosol extinction.

5.5: The RSS Retrievals with the Variances Fixed

Figure 5.8 revealed an interdependence between the retrieved values of the fine mode effective radius and effective variance in the RSS data. Given the results of the EOF analysis, it is likely that this interdependence stems not from anything physical, but from limits to the amount of information retrievable from the data. If this is indeed the case, then trying to retrieve information beyond the theoretical limit would not only introduce redundancies into the data, but could produce misleading results which actually conceal some important items of retrievable information. For example, the low effective radius values associated with the high variance values could actually mask a seasonal pattern in the plot of fine mode effective radius vs. day. The RSS bimodal retrievals were therefore re-done, this time with the coarse and fine mode effective variances fixed at 0.1.

This value was chosen because it was the dominant value of the effective variance retrieved for both aerosol modes in the RSS retrieval. This fact does not necessarily guarantee that 0.1 is the actual physical variance for either aerosol mode, however. The NO_2 and ozone values from high-resolution gas retrievals were used in this fixed-variance retrieval.

The mean values of the fixed-variance retrieval are presented in Table 5.1, along with the means from the other retrievals using RSS data. The mean coarse and fine mode optical depths do not change significantly. The mean coarse mode effective radius is slightly lowered. Due to the absence of high variance values that correspond to low radius values, the mean fine mode effective radius has increased significantly, from 0.164 μm to 0.185 μm. Most significantly, as Figure 5.29 shows, the plot of fine mode effective radius vs. day no longer shows as much daily variability as was seen in Figure 5.7, the corresponding plot when the effective variance was not held fixed. Now, excepting a small number of large values, the plot clearly does show a seasonal cycle, with a peak around day 430 (March 5, 2000) and a minimum around day 270 (September 27, 1999). This agrees roughly with the results of the Holben *et al.* (2001) aerosol climatology for the SGP site, obtained by calculating the Ångstrom coefficients for the CIMEL optical depth data between 1994 and 1999. Excluding months with less than 15 days sampled over the six-year span, the mean monthly Ångstrom coefficient had maximum values (corresponding to smaller particles) in August and October, and a minimum value (corresponding to larger particles) in April. The almucantar-retrieved fine mode effective radius values presented in Figure 5.24 do no follow exactly the same pattern, but given the known errors in the CIMEL data, the fact that there is even

marginal agreement with the RSS is remarkable. The results of Figure 5.29 ultimately show the importance of not attempting to retrieve too much information from limited data. A significant seasonal cycle would have been overlooked entirely if the retrieval that allowed both effective variances to vary was taken as the highest quality retrieval of the RSS data.

5.6: Conclusions

Clearly, limits do exist to the ability to completely retrieve aerosol properties using data with the wavelength range of the RSS or the CIMEL, even when the retrieval algorithm assumes a bimodal aerosol size distribution. For the fine mode, the effective radius and effective variance cannot both be independently retrieved. Attempting the retrieval of both quantities can mask important trends in the data, like the seasonal cycle in the fine mode effective radius. Therefore, the most appropriate retrieval algorithm for the RSS or CIMEL optical depth data keeps the effective variance values fixed. The almucantar retrievals of Dubovik and King (2000) seem to produce more consistent values for the fine mode effective radius, but the interdependence between the effective radius and the effective radius persists even here. The close agreement between mean values of the fine mode effective radius and the best Mie fits to the coefficients of projections of the first EOF's in the data sets where the mean had been subtracted lends strong credence to the idea that the EOF's capture the bimodality of the aerosol size distribution.

The coarse aerosol mode is much less clearly defined in the RSS and CIMEL data than the fine mode is. The values for the coarse mode effective radius retrieved using the almucantar data and using the optical depth alone barely agree within a factor of two. Upon examining the extinction curves produced by median coarse mode size distributions retrieved using the two types of algorithms, it becomes apparent that there is not enough information in the wavelength range of these devices to distinguish between these two size distributions, despite their obvious physical differences. Ideally, the next generation of sun photometer devices will incorporate broader ranges of wavelengths so that the coarse mode aerosol size distribution can be clearly and unambiguously measured. In the meantime, the best possible retrieval algorithms for the MFRSR and RSS data, given the confirmed bimodality of the aerosol size distribution and limit to the retrievable information, need to be designed and implemented.

	mean fine mode τ	mean fine mode r_{eff} (μm)	dominant fine mode v_{eff}	mean coarse mode τ	mean coarse mode r_{eff} (μm)	dominant coarse mode v_{eff}	mean NO_2 (DU)	mean ozone (DU)	RMS residual χ^2
RSS bimodal, gases separate	0.048	0.164	0.1 (52.0%)	0.019	1.38	0.1 (76.4%)	0.38	279	0.00163
RSS bimodal, gases coupled	0.049	0.176	0.1 (48.3%)	0.019	1.48	0.1 (76.4%)	0.51	270	0.00145
RSS single-mode, gases coupled	0.064	0.290	0.5 (87.4%)	-----	----	---	1.60	262	0.00289
RSS single-mode, gases separate	0.071	0.205	0.5 (91.3%)	-----	----	---	0.38	279	0.00441
RSS bimodal, fixed variance	0.047	0.185	0.1	0.019	1.33	0.1	0.38	279	0.00161

Table 5.1: The mean values for the retrieved quantities in the RSS bimodal retrievals.

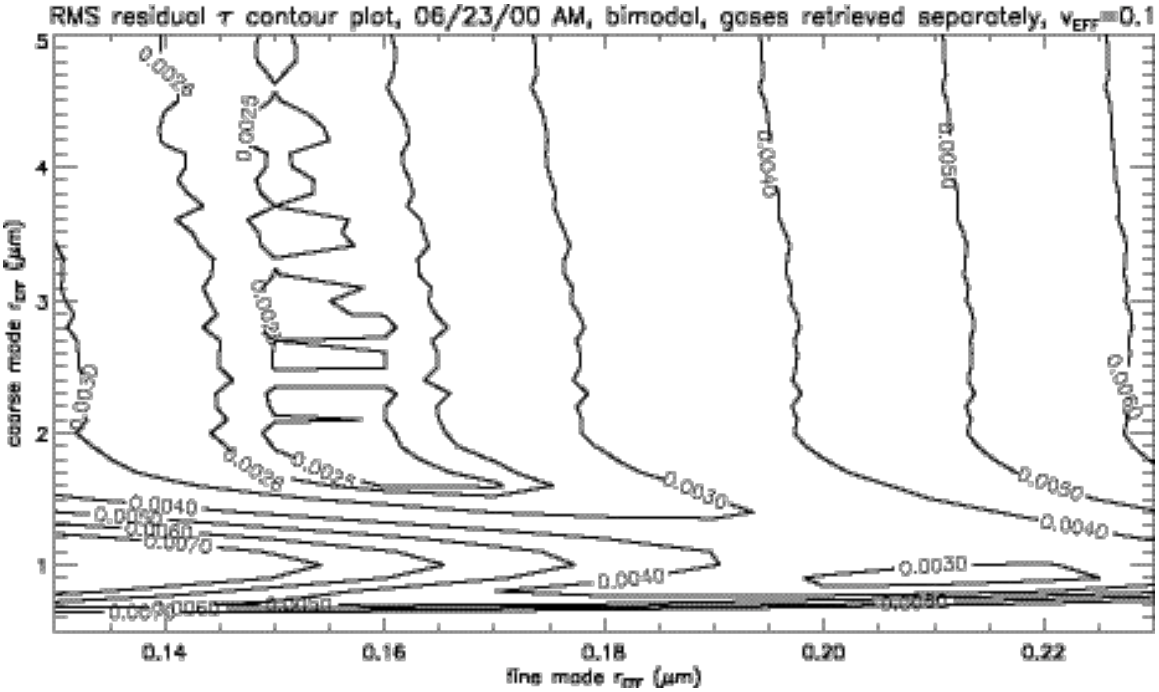

Figure 5.1: A contour plot of the residual optical depth values vs. the coarse mode (vertical) and fine mode (horizontal) effective radii, for the morning of June 23, 2000. The plot is done with the effective variance values for both modes set to 0.1.

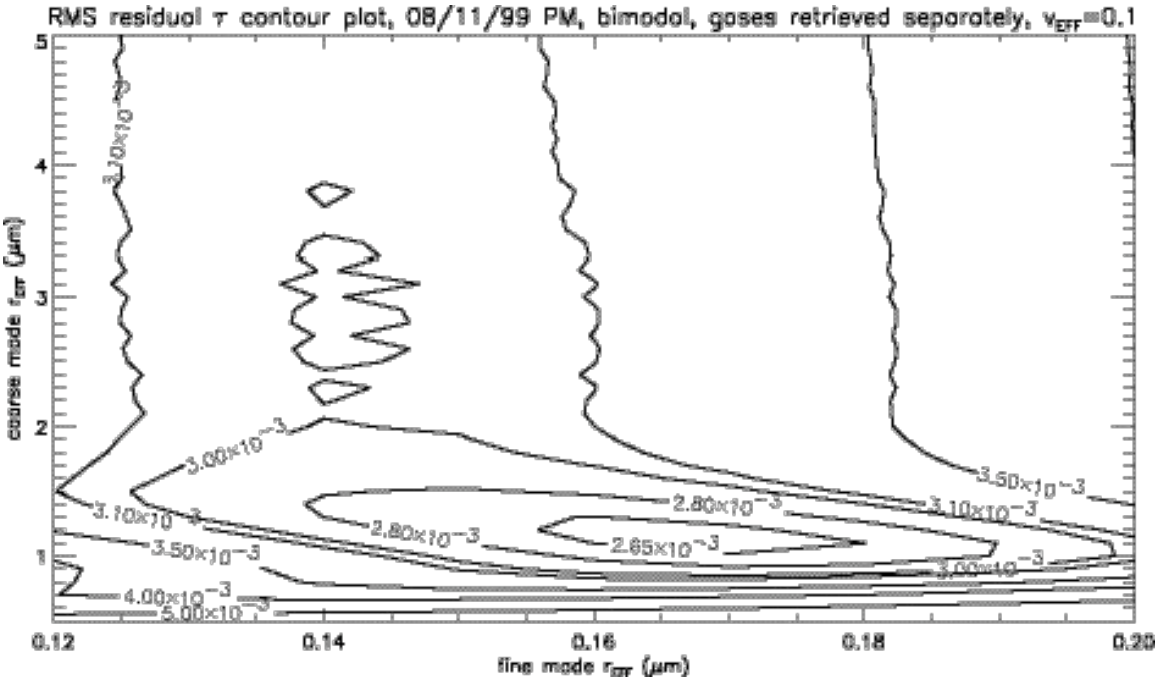

Figure 5.2: A contour plot of the residual optical depth values vs. the coarse mode (vertical) and fine mode (horizontal) effective radii, for the afternoon of August 11, 1999. The plot is done with the effective variance values for both modes set to 0.1.

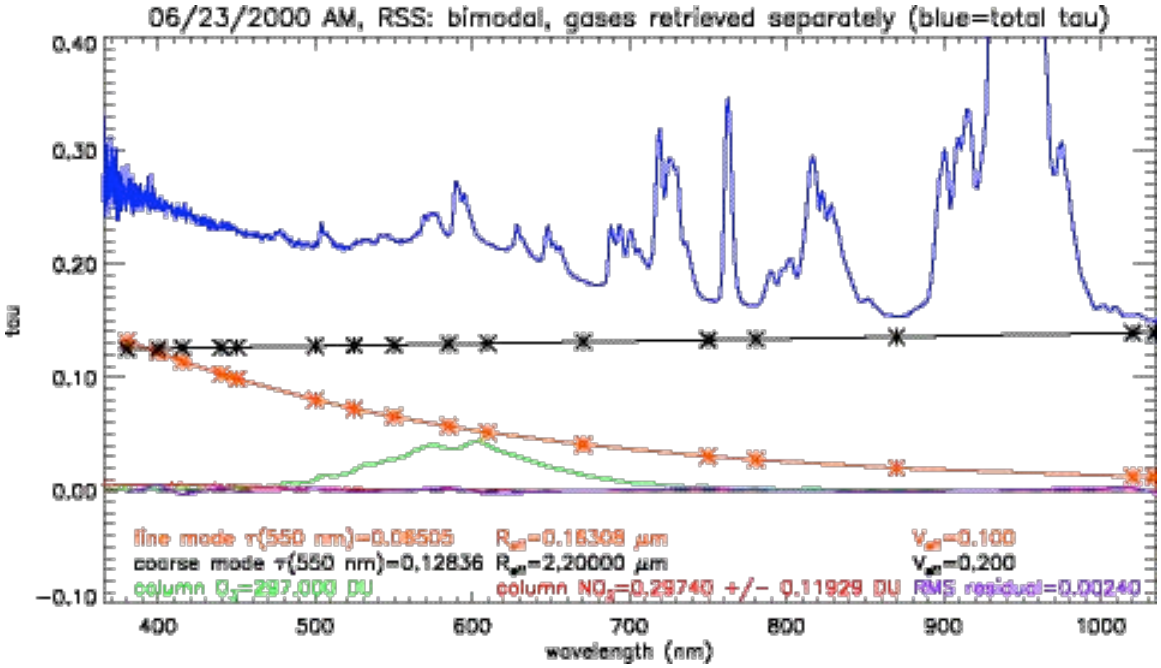

Figure 5.3: The results of the bimodal retrieval for the morning of June 23, 2000, with ozone and aerosols retrieved separately.

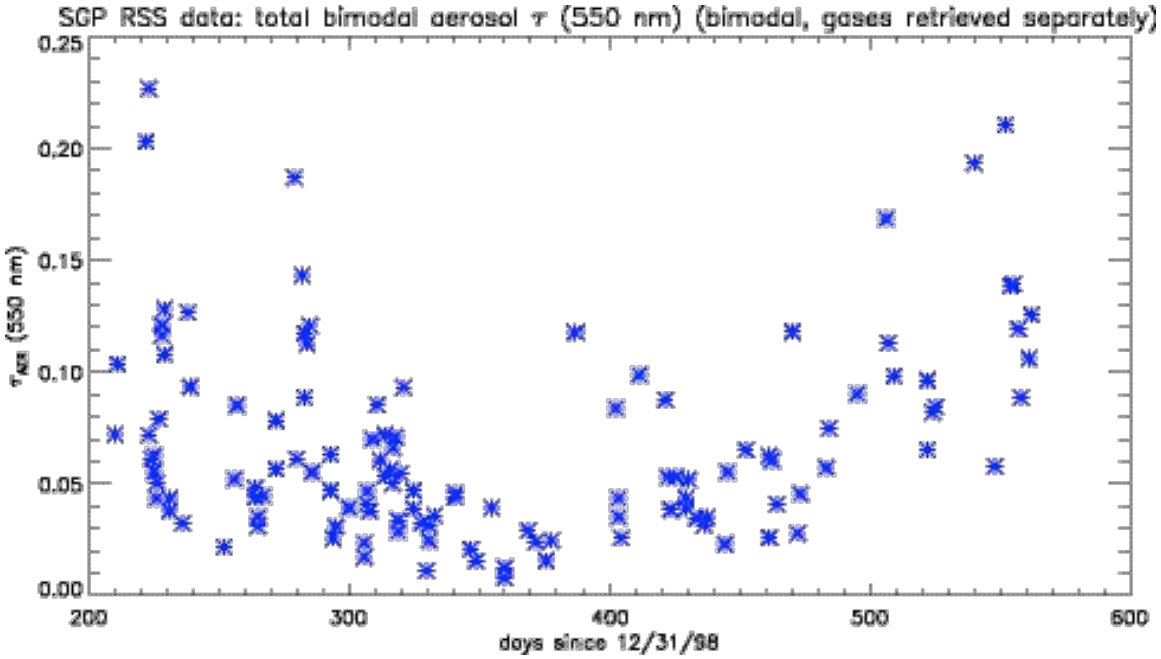

Figure 5.4: The total RSS bimodal aerosol optical depth at 550 nm, plotted vs. day.

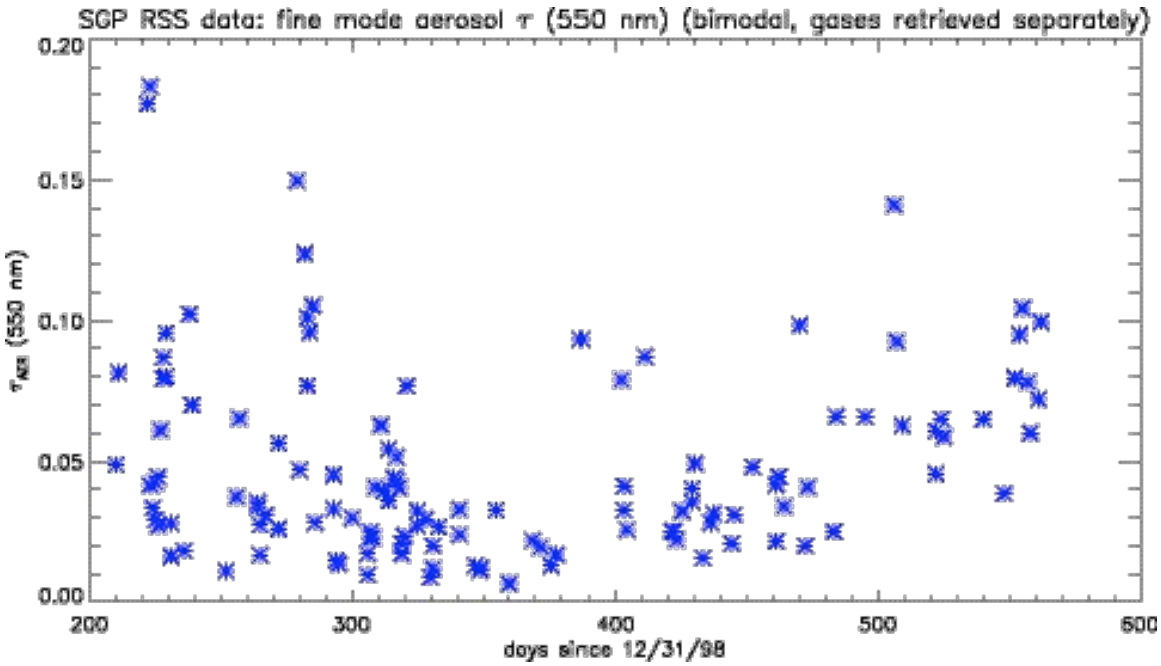

Figure 5.5: The RSS fine mode optical depth, plotted vs. day.

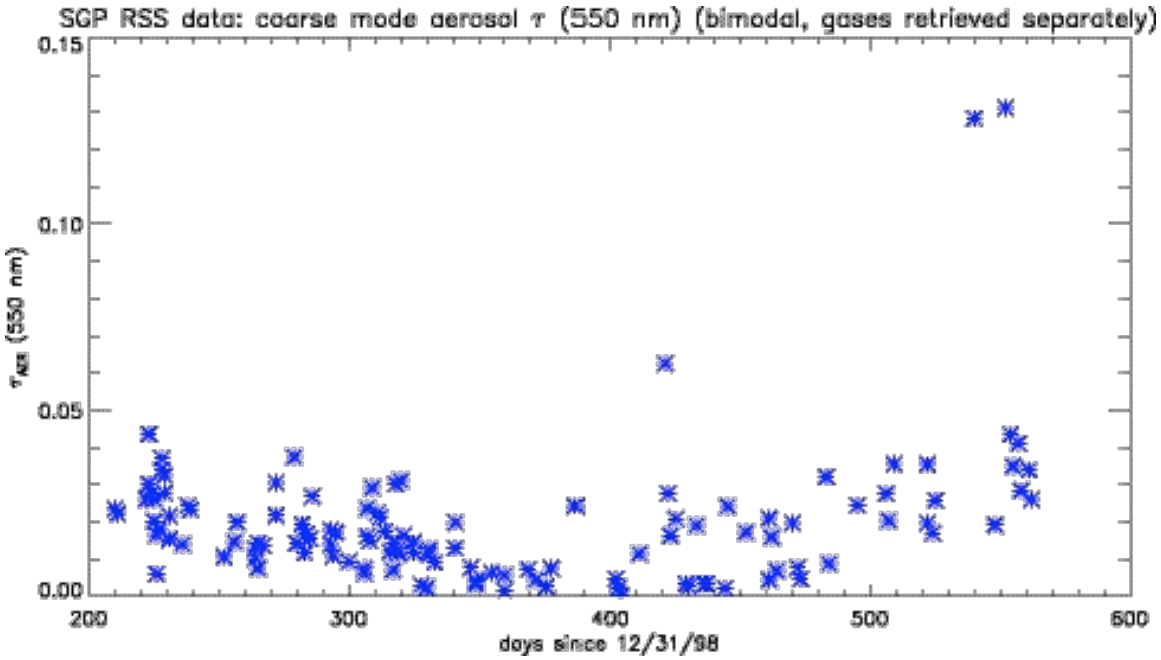

Figure 5.6: The RSS coarse mode optical depth, plotted vs. day. The three large peaks occur on July 5, June 23, and February 25.

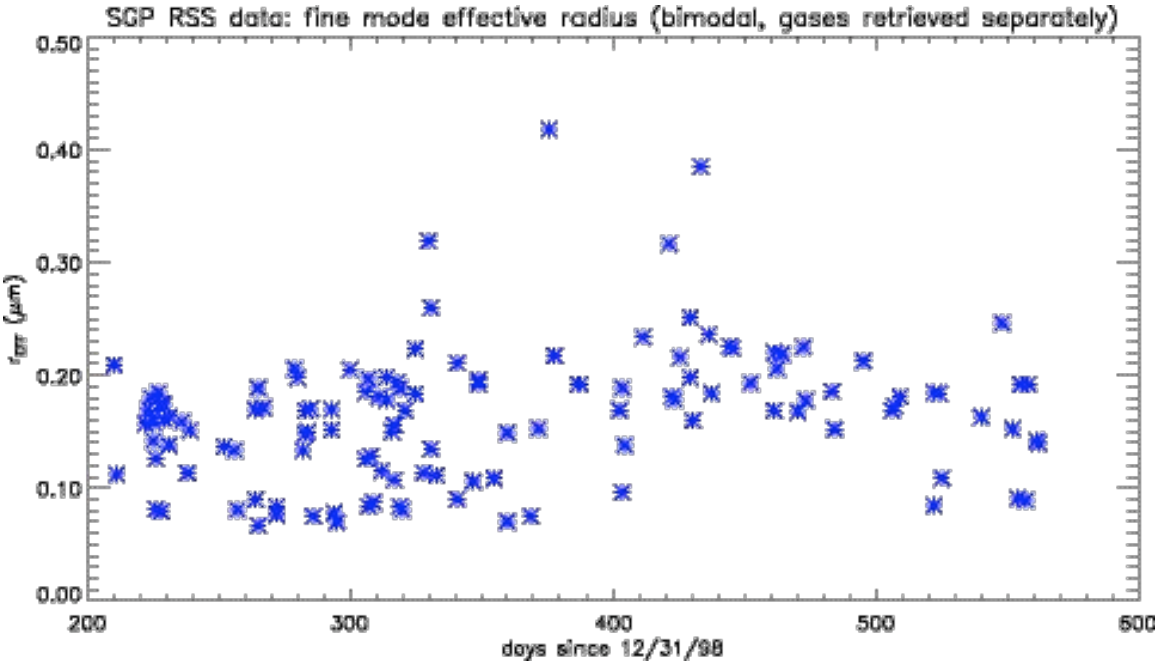

Figure 5.7: The RSS-retrieved fine mode effective radius, plotted vs. day.

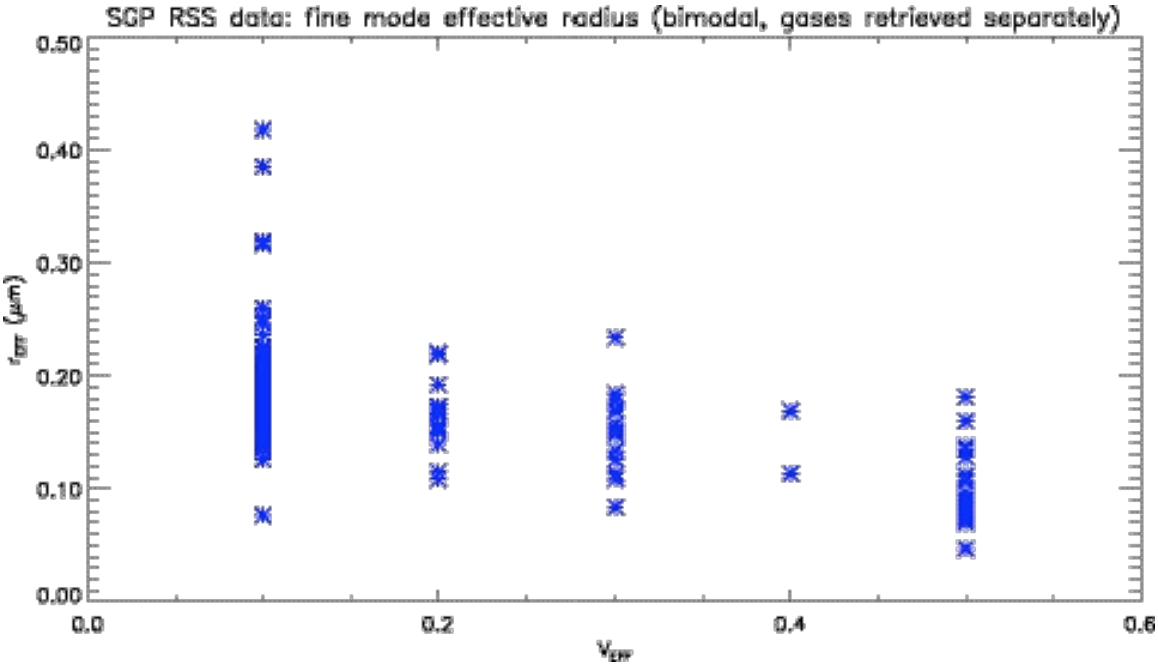

Figure 5.8: The plot of the RSS-retrieved fine mode effective radius vs. effective variance, which clearly shows that the lower radius values correspond with higher variance values.

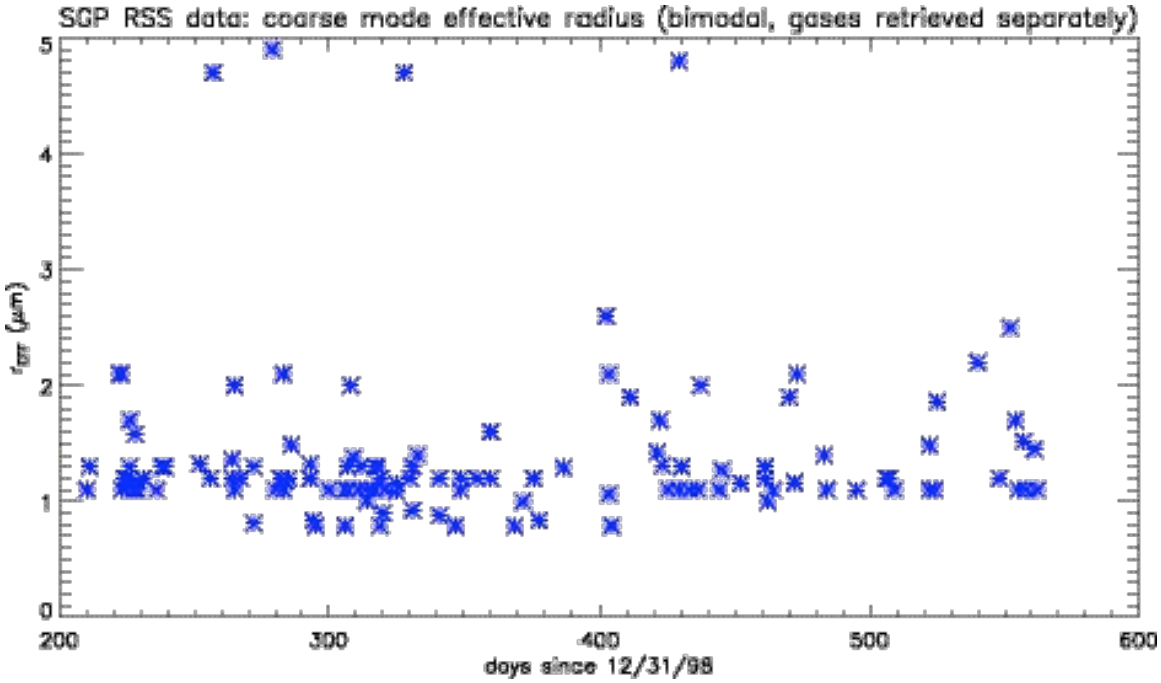

Figure 5.9: The RSS-retrieved coarse mode effective radius, plotted vs. day.

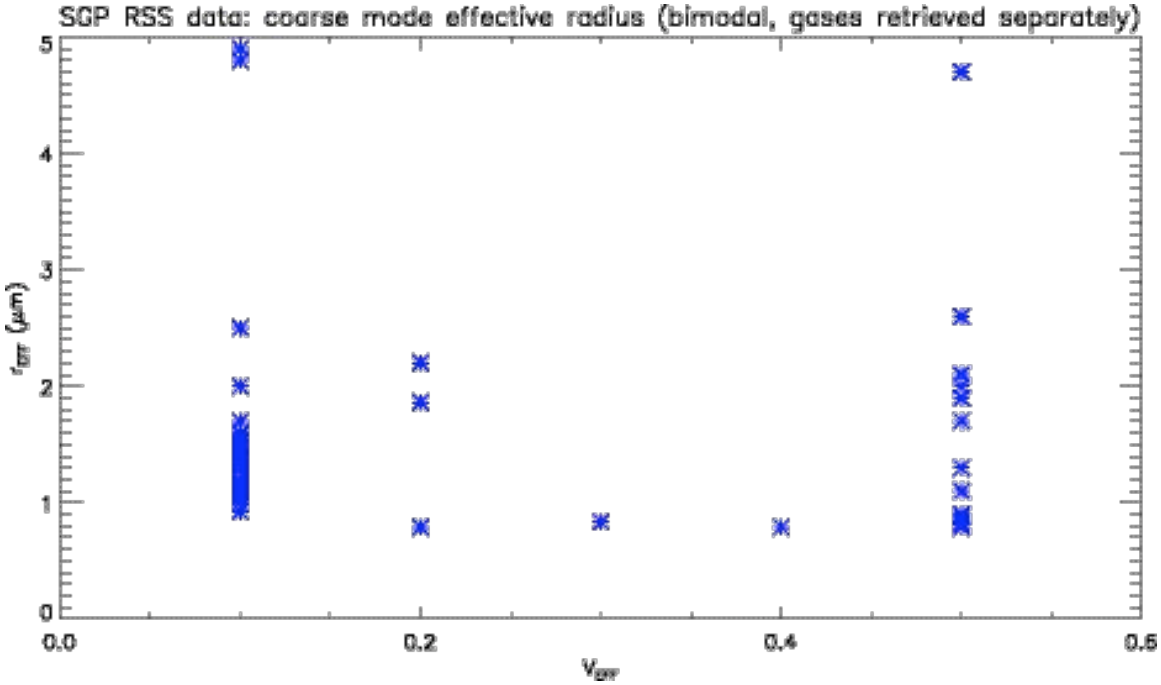

Figure 5.10: The plot of the RSS-retrieved coarse mode effective radius vs. effective variance.

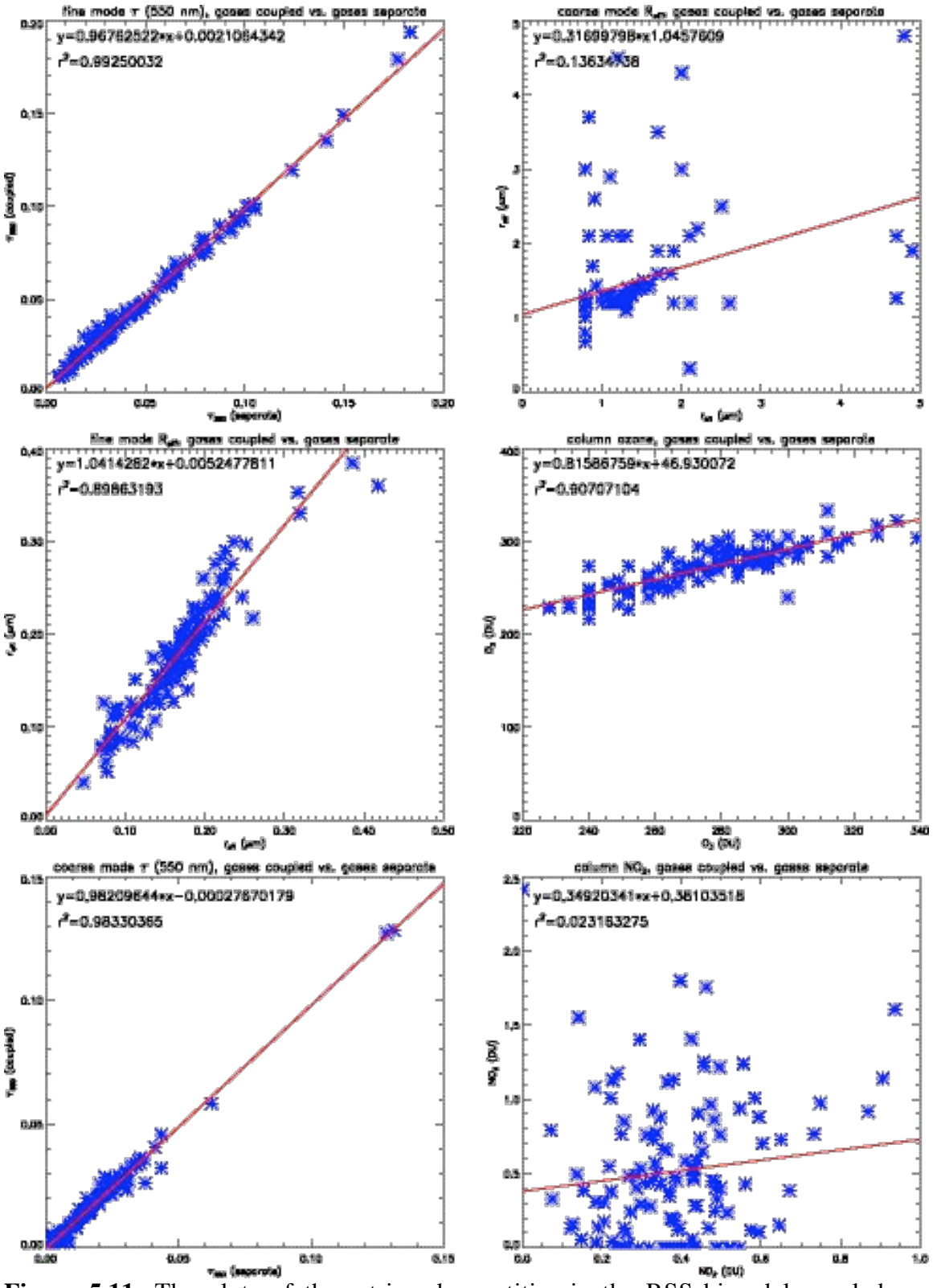

Figure 5.11: The plots of the retrieved quantities in the RSS bimodal coupled gas retrieval, vs. the same quantities when gases and aerosols are retrieved separately.

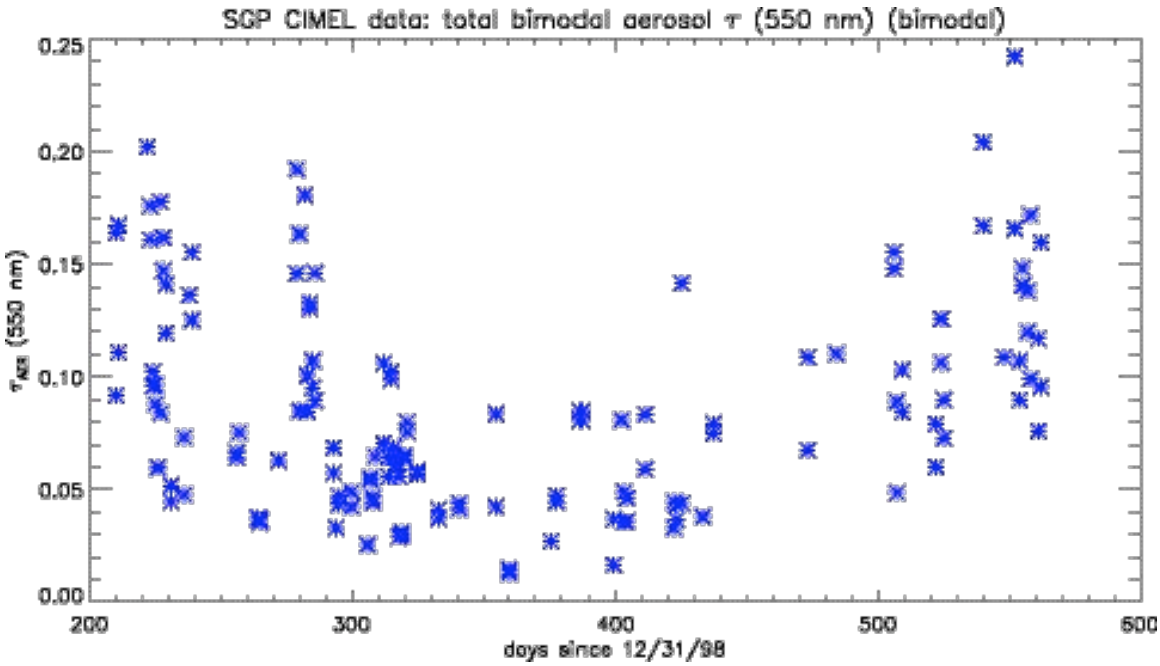

Figure 5.12: The total CIMEL aerosol optical depth at 550 nm, plotted vs. day.

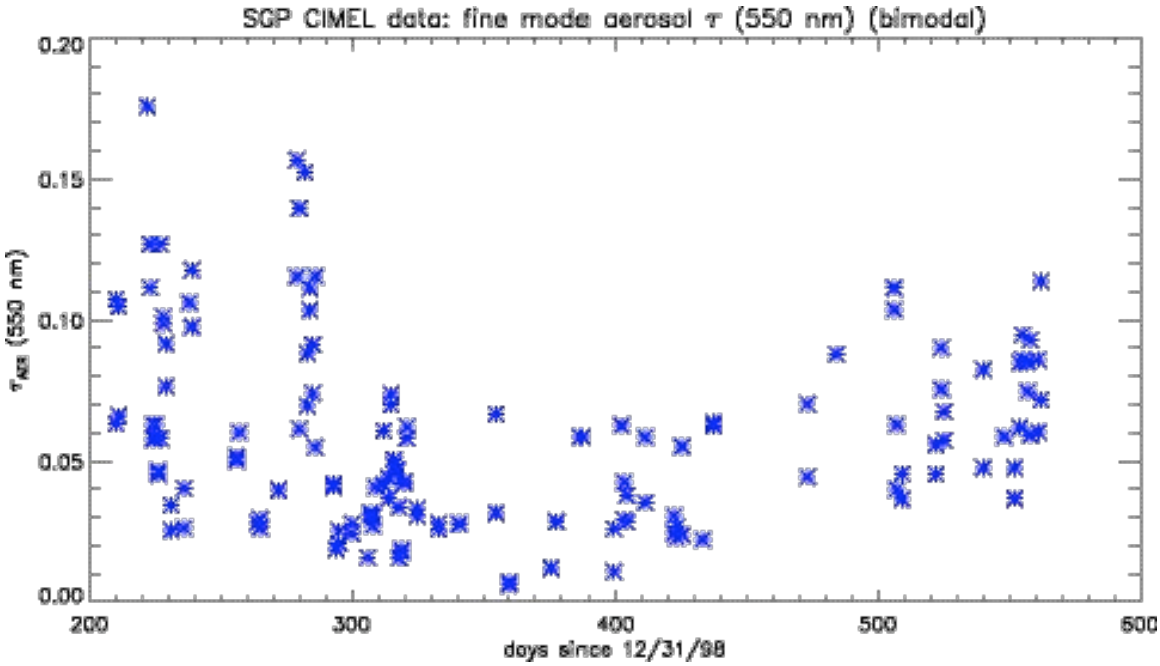

Figure 5.13: The CIMEL fine mode optical depth at 550 nm, plotted vs. day.

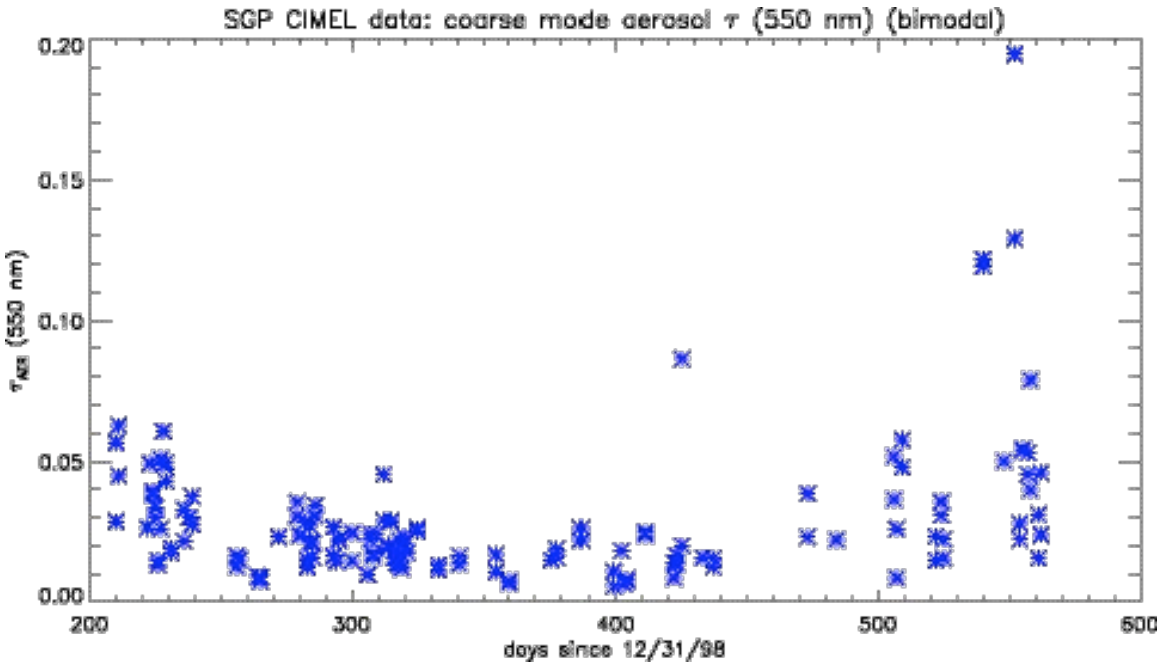

Figure 5.14: The CIMEL coarse mode optical depth at 550 nm, plotted vs. day.

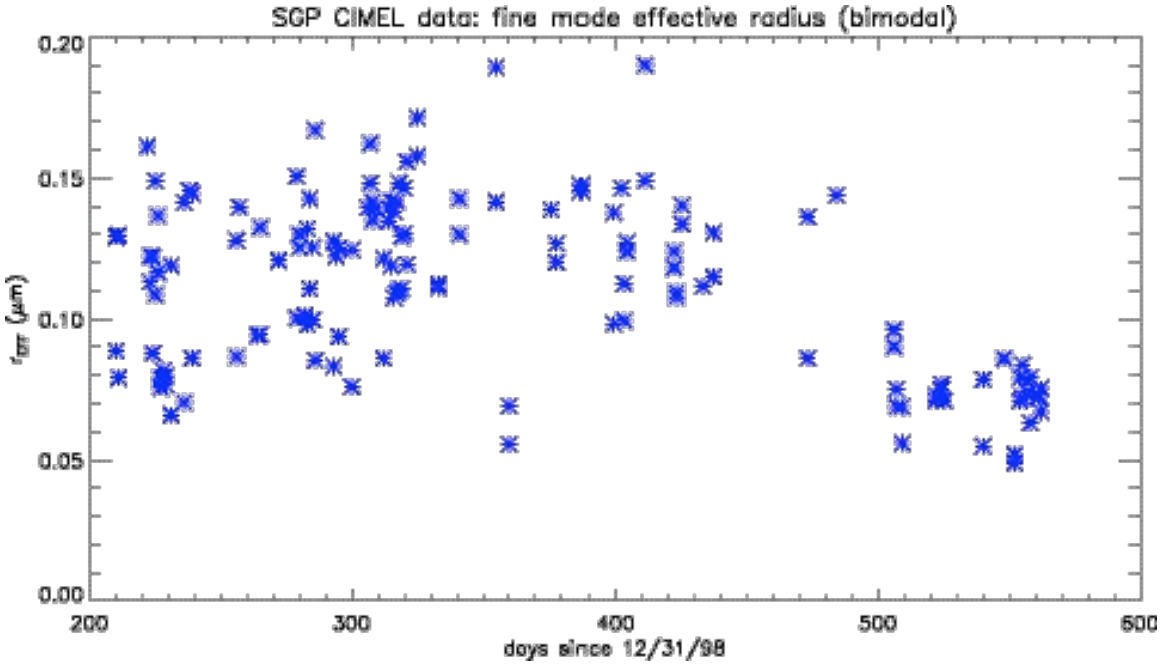

Figure 5.15: The CIMEL fine mode effective radius, plotted vs. day. A similar pattern to the one observed in the third EOF emerges in this plot.

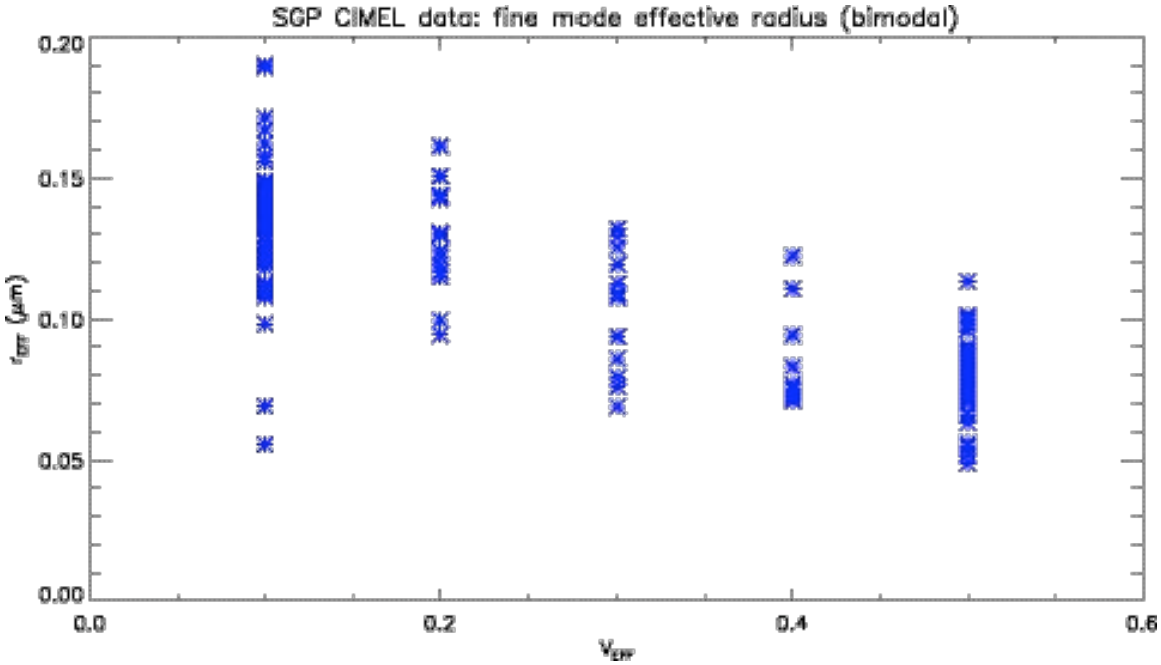

Figure 5.16: The CIMEL fine mode effective radius, plotted vs. fine mode effective variance. The same interdependence observed in the RSS data in Figure 5.6 is also observed here.

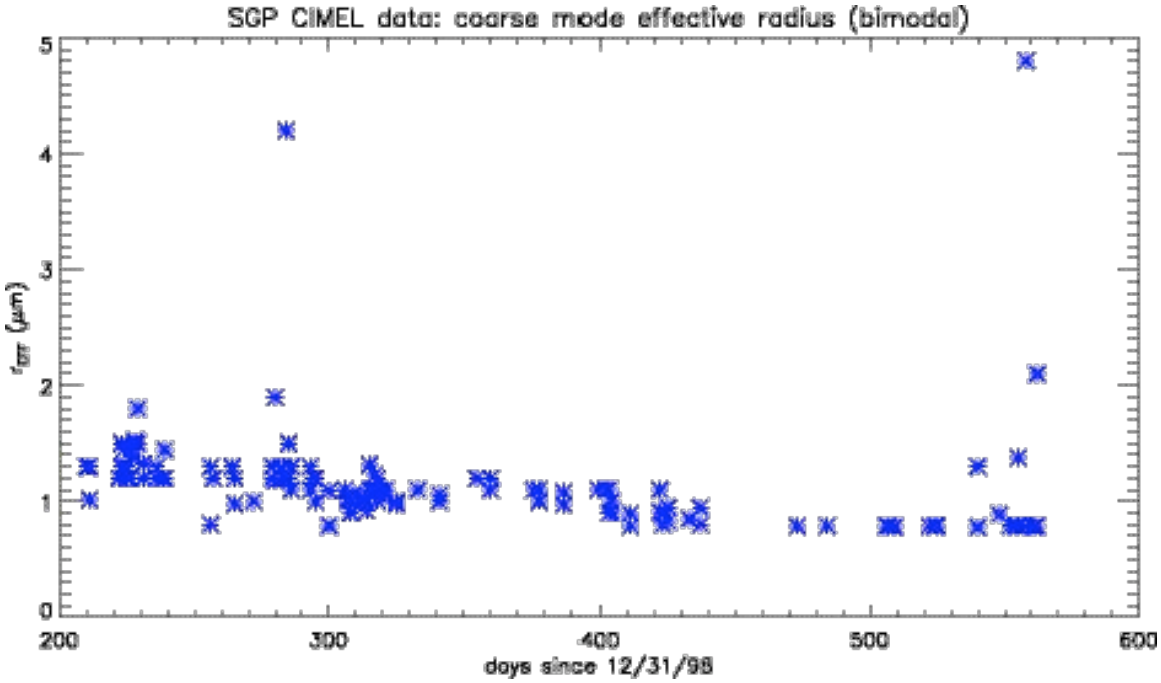

Figure 5.17: The CIMEL coarse mode effective radius, plotted vs. day. The steady downward pattern, with a few higher days near the end, also resembles the third EOF.

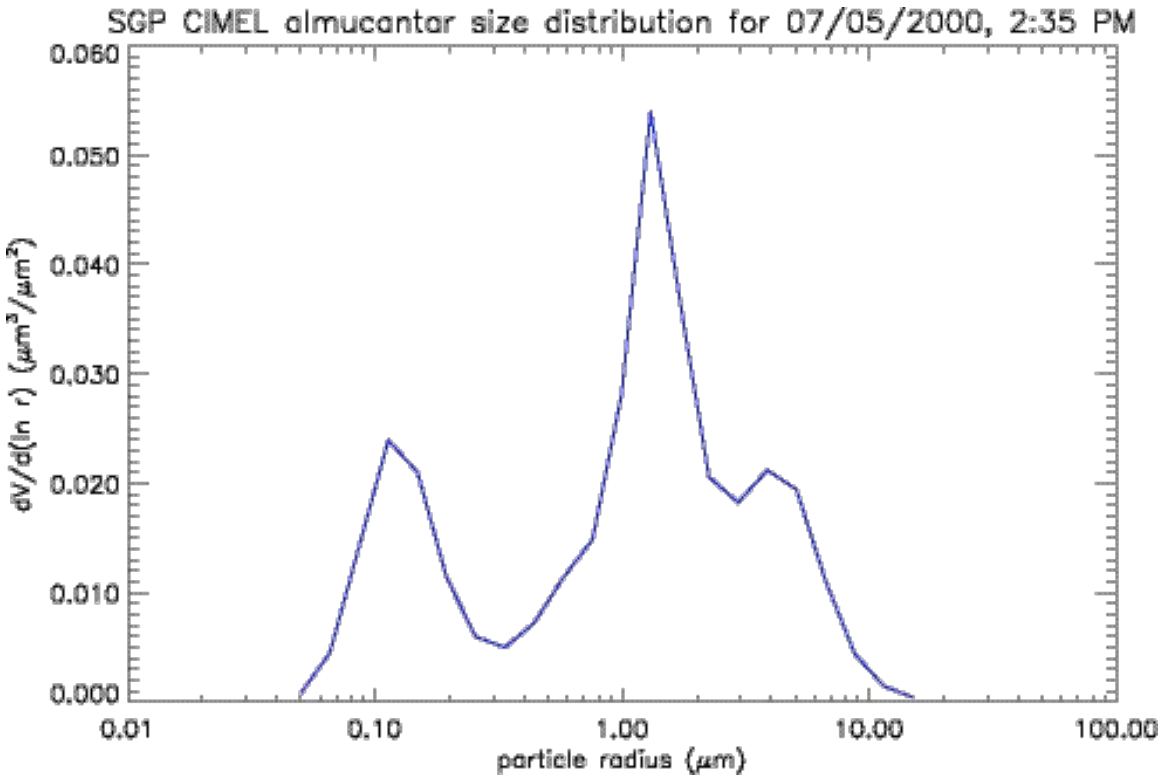

Figure 5.18: The almucantar-retrieved aerosol size distribution for 2:35 PM on July 5, 2000. The coarse mode peaks at a particularly low radius.

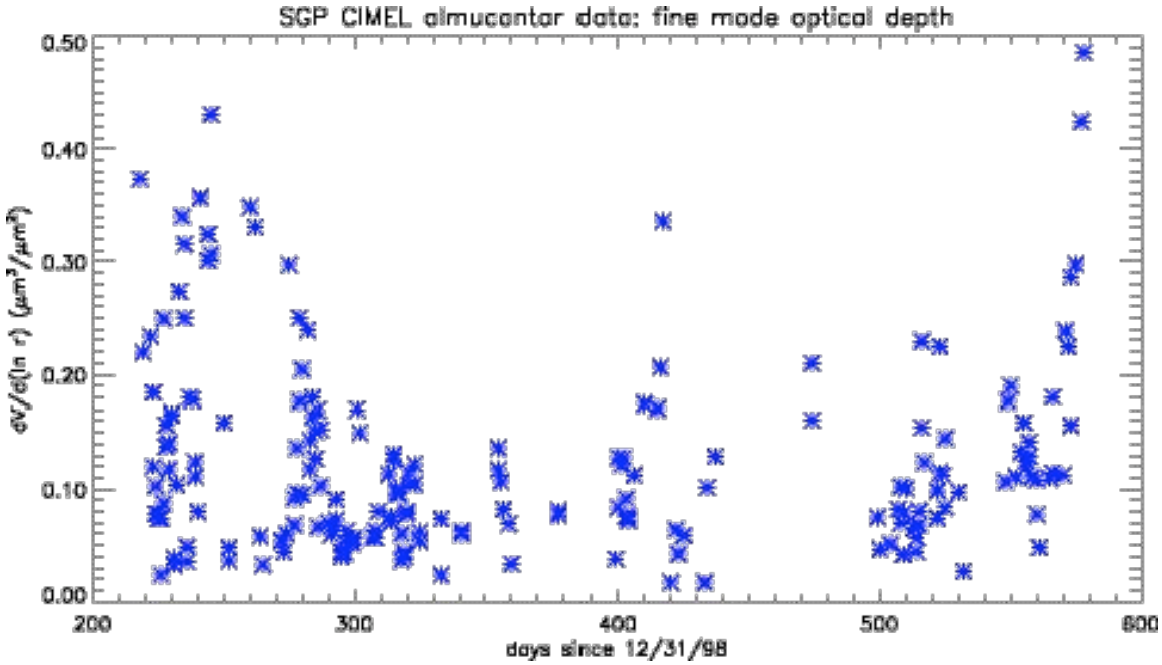

Figure 5.19: The plot of the almucantar-retrieved fine mode aerosol optical depth at 550 nm vs. day.

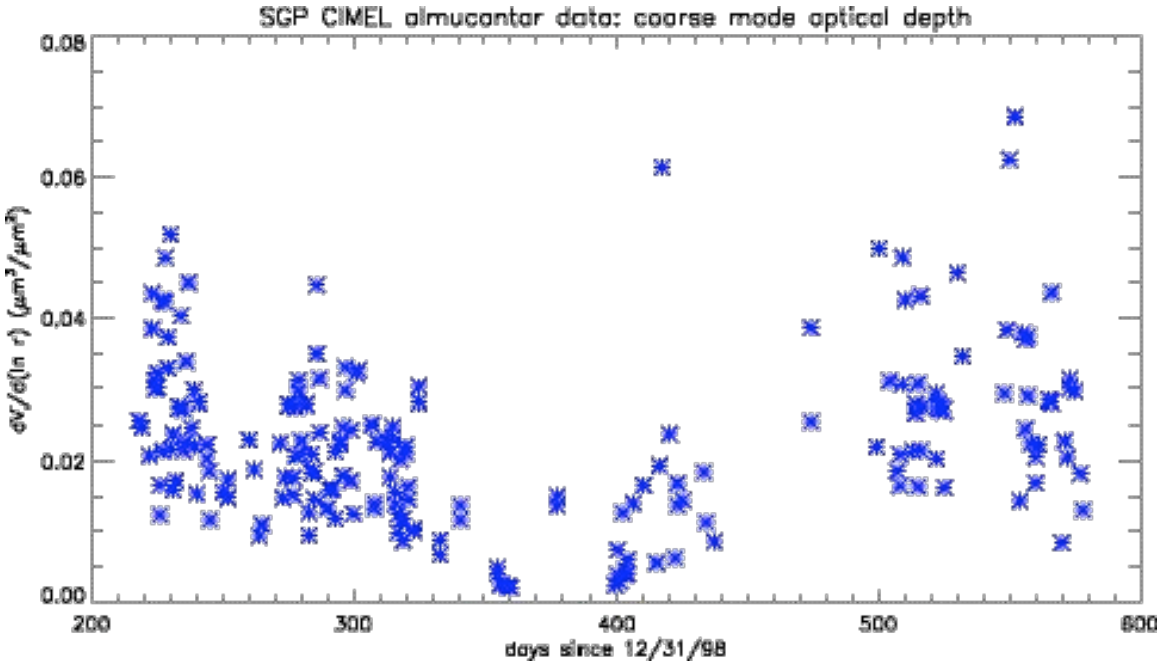

Figure 5.20: The plot of the almucantar-retrieved coarse mode aerosol optical depth at 550 nm vs. day.

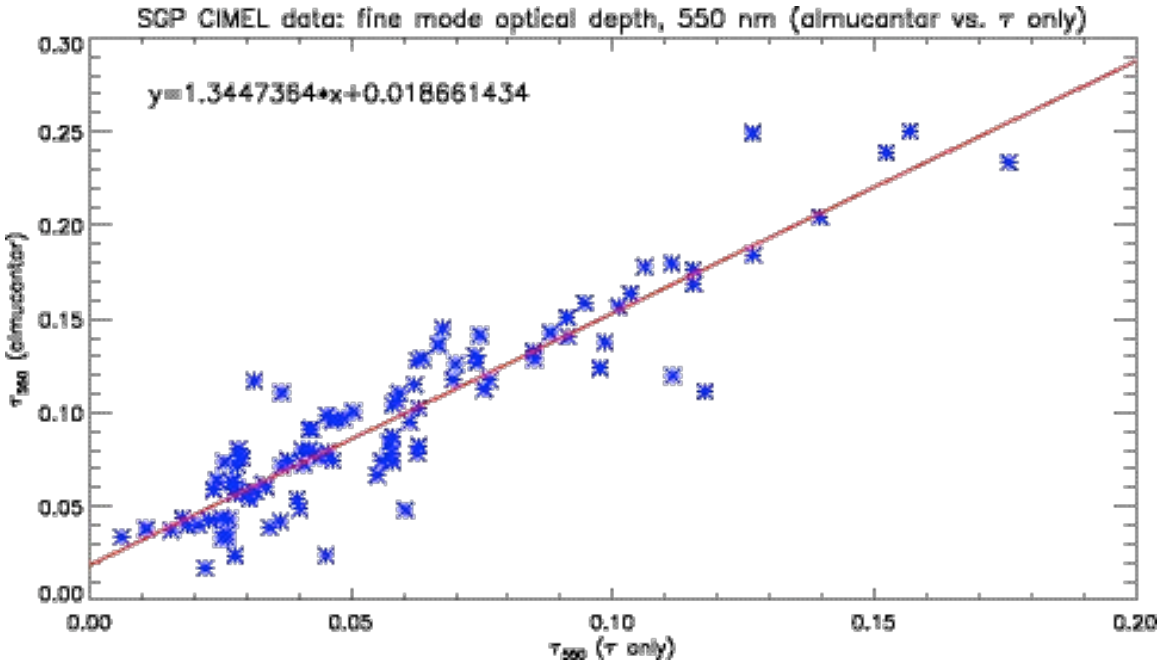

Figure 5.21: The plot of the almucantar-retrieved fine mode aerosol optical depth at 550 nm vs. the corresponding optical depth for the CIMEL retrievals using only the optical depth data.

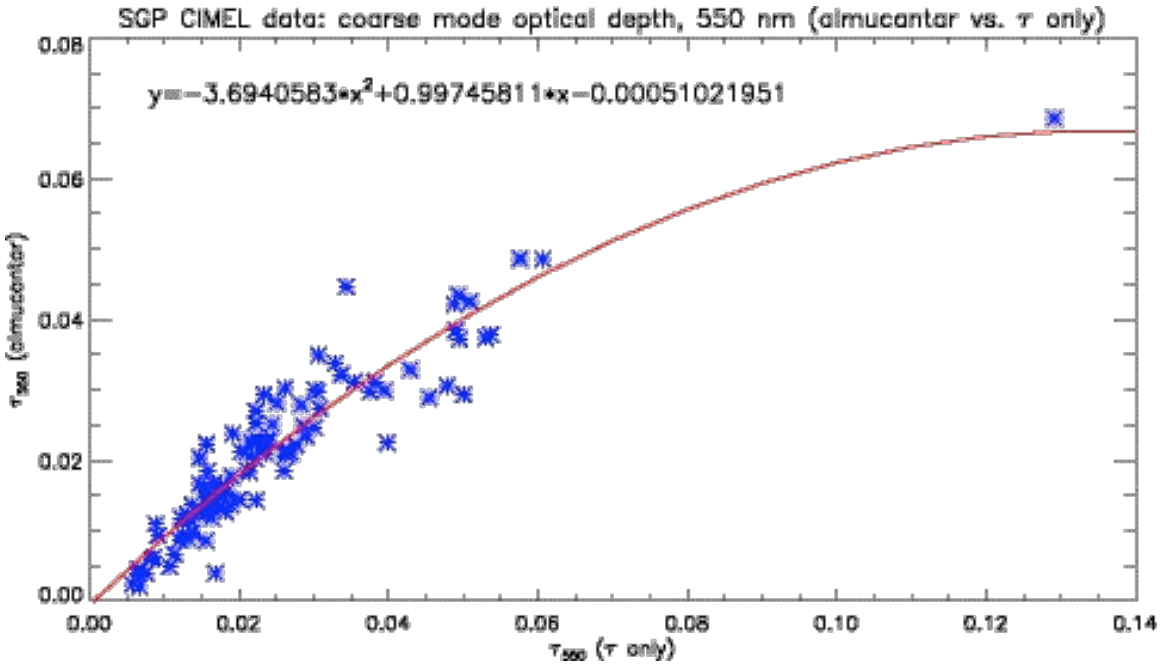

Figure 5.22: The plot of the almucantar-retrieved coarse mode aerosol optical depth at 550 nm vs. the corresponding optical depth for the CIMEL retrieval using only the optical depth data. The relationship here is clearly non-linear.

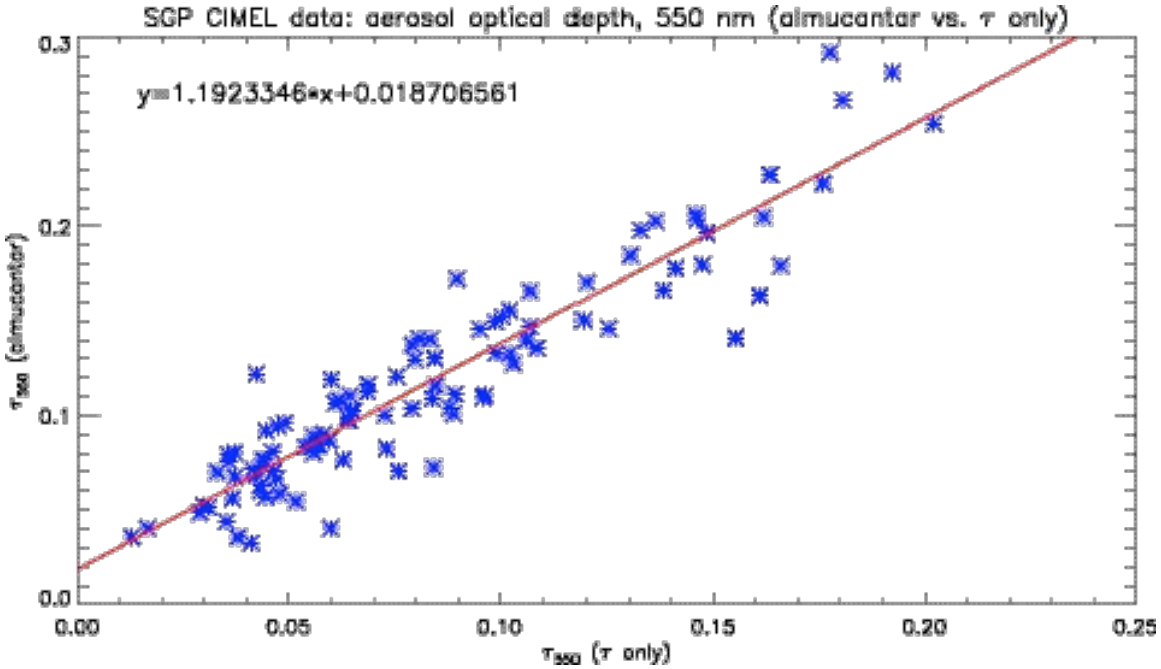

Figure 5.23: The plot of the almucantar-retrieved total mode aerosol optical depth at 550 nm vs. the corresponding optical depth for the CIMEL retrieval using only the optical depth data.

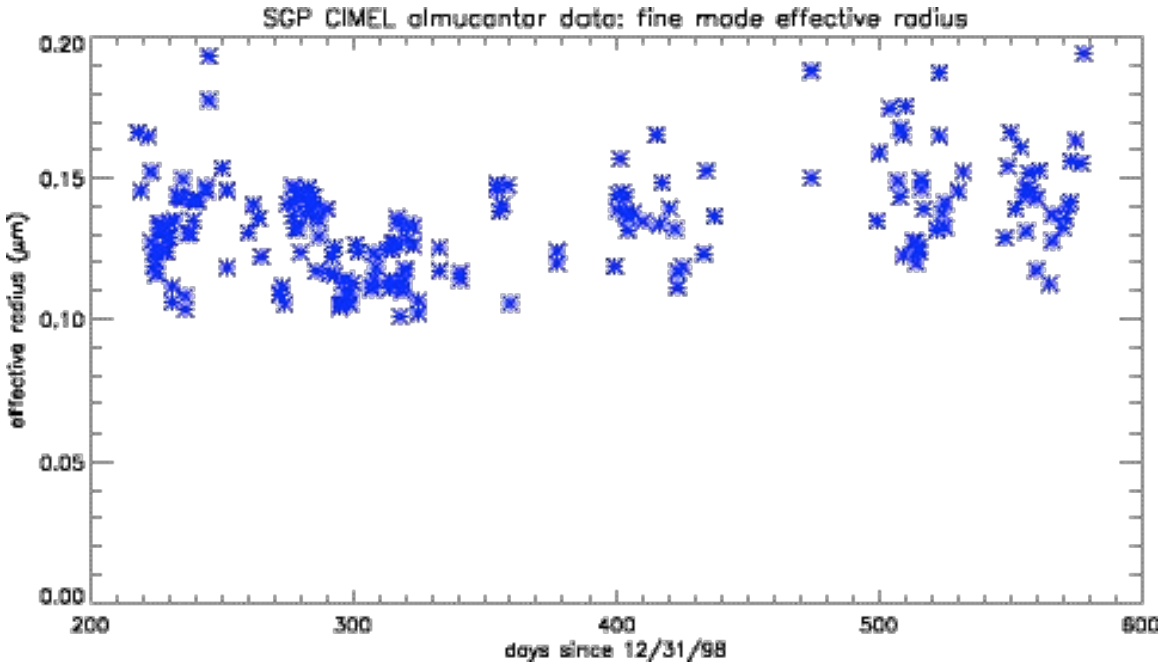

Figure 5.24: The plot of the almucantar-retrieved fine mode effective radius vs. day.

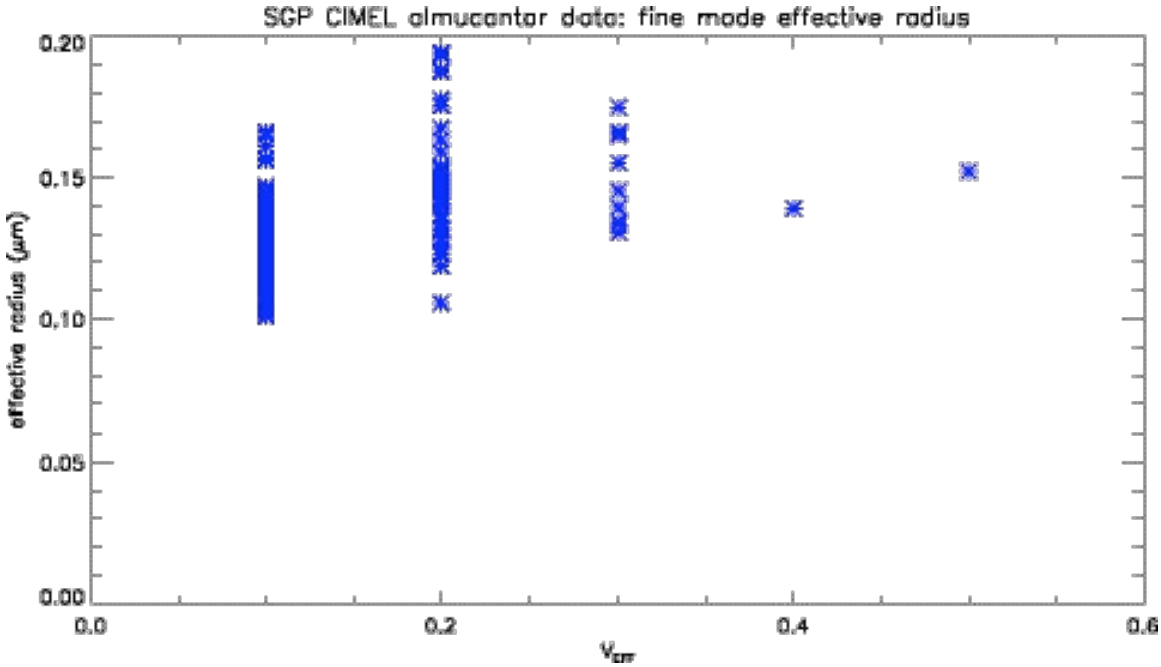

Figure 5.25: The plot of the almucantar-retrieved fine mode effective radius vs. effective variance. Unlike the RSS results in Figure 5.6, the almucantar effective radius values appear in general to increase with the variance.

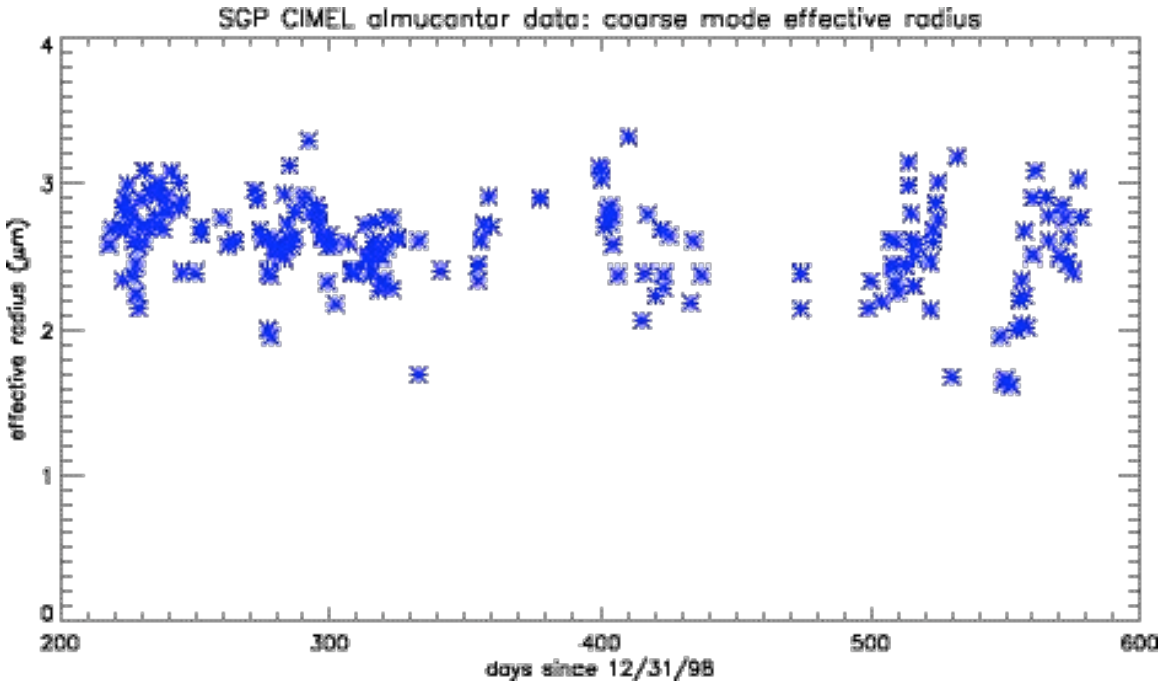

Figure 5.26: The plot of the almucantar-retrieved coarse mode effective radius vs. effective variance. The graph does not show the same abrupt changes apparent in the direct sun optical depth measurements, but there does appear to be a downward trend over the course of the data set.

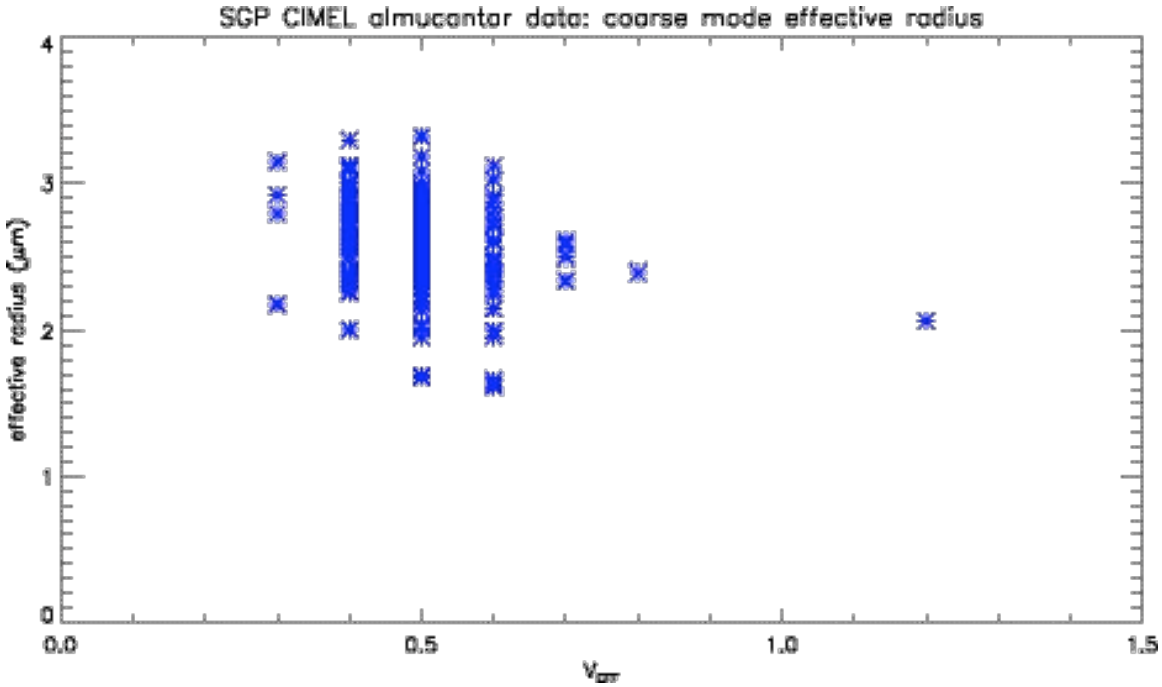

Figure 5.27: The plot of the almucantar-retrieved coarse mode effective radius vs. effective variance. The variance is extremely large in a couple of places, and the coarse mode effective radius appears to decrease in general as the variance increases.

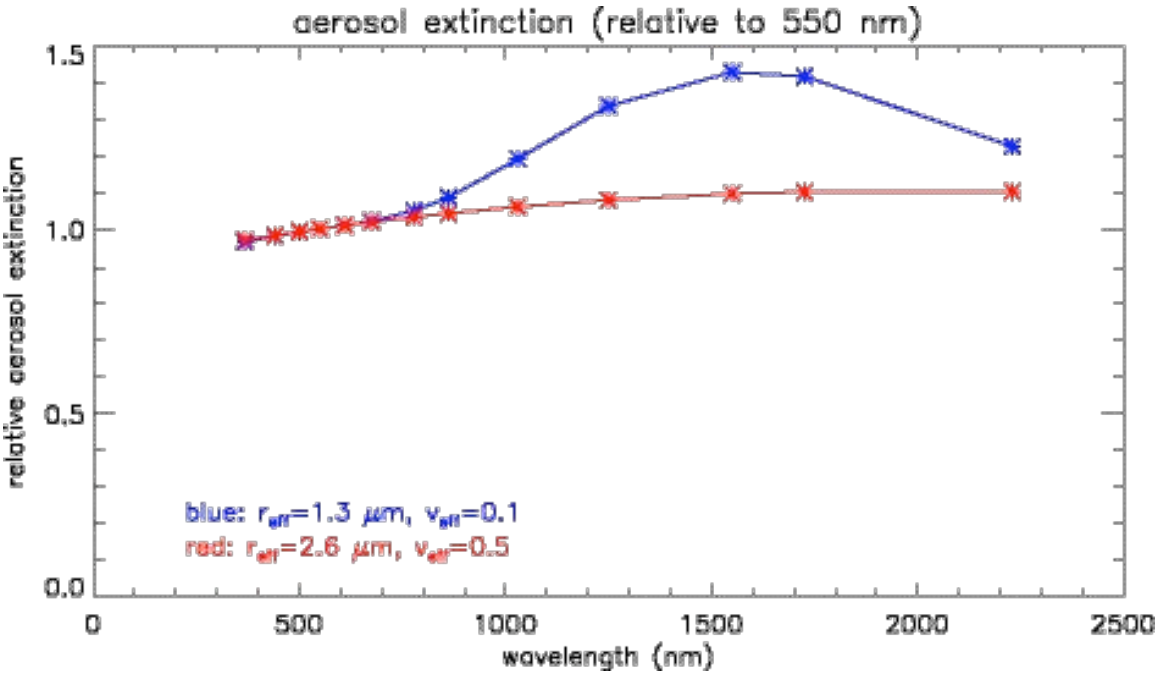

Figure 5.28: A plot of the relative aerosol extinction for the RSS-retrieved (blue) and the almucantar-retrieved (red) coarse mode size distributions.

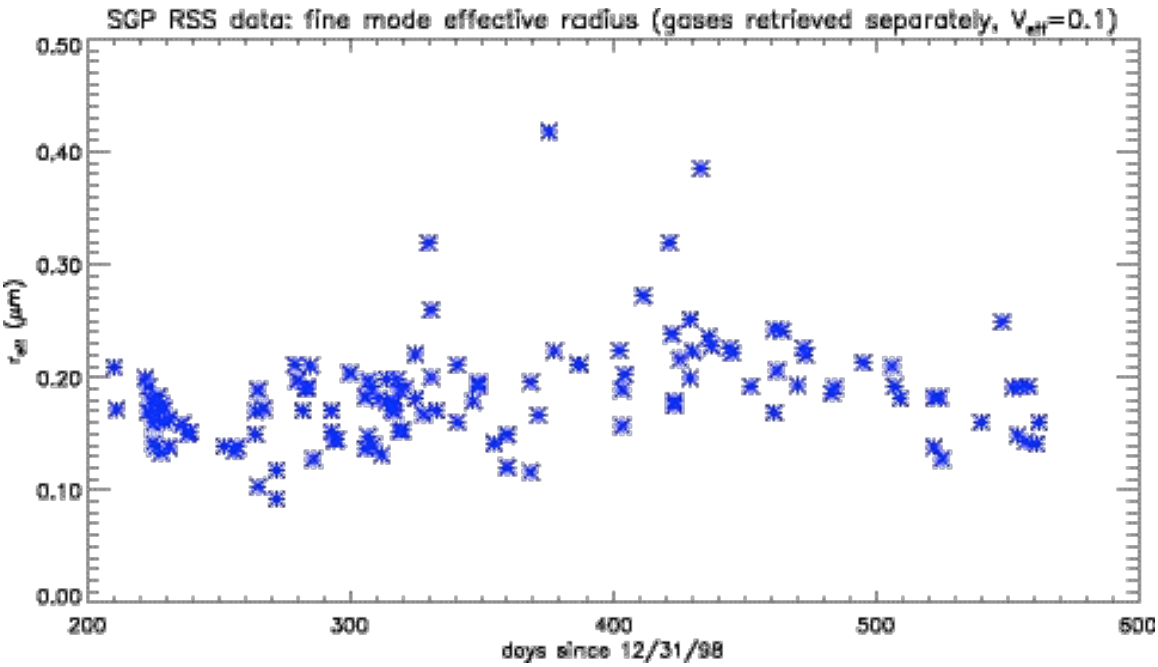

Figure 5.29: The plot of the fine mode effective radius vs. day for the RSS data, when the effective variance values are set to 0.1. When presented this way, the effective radius values appear to show a seasonal cycle, with a maximum in March and a minimum in September.

Chapter 5 References

Box, G. P., Box, M. A., and Krücker, J., *Journal of Geophysical Research* **101**, 19211-19214 (1996).

Dubovik, O., and King, M. D., *Journal of Geophysical Research* **105**, 20673-20696 (2000).

Holben, B. N., Tanré, D., Smirnov, A., Eck, T. E., Slutsker, I., Abuhassan, N., Newcomb, W. W., Schafer, J. S., Chatenet, B., Lavenu, F., Kaufman, Y. J., Vande Castle, J., Setzer, A., Markham, B., Clark, D., Frouin, R., Halthore, R., Karneli, A., O'Neill, N. T., Pietras, C., Pinker, R. T., Voss, K., and Zibordi, G., *Journal of Geophysical Research* **106**, 12067-12097 (2001).

Lacis, A. A., and Mishchenko, M. I., *Aerosol Forcing of Climate*, ed. R. J. Charlson and J. Heintzenberg, J. Wiley & Sons, pp. 11-42 (1995).

Chapter 6: Conclusions and Future Objectives

6.1: Making the Most of the Available Devices and Data

With the limits to the amount of information retrievable from the MFRSR, RSS, and CIMEL data now established, it becomes necessary to optimize the retrievals for each device within these constraints. For the RSS and CIMEL, this means not attempting to retrieve too much information, and accepting that even a 1000-channel array like the RSS cannot measure the aerosol size distribution in complete detail. The coarse mode can not be retrieved uniquely from the information obtainable with the RSS and CIMEL data, and the overall quality of the retrieval is only as good as the accuracy of the assumptions made about the coarse mode. These assumptions can be improved on, by examining additional data from other devices with broader wavelength ranges or more scattering angles.

Additionally, it remains to be determined if one of these three devices would boast any unambiguous advantages over the others as the primary device in an aerosol monitoring network. The RSS can much better constrain gas amounts than the MFRSR can, but it is not clear that the ozone values retrieved by it are more accurate than climatology. Furthermore, the combination of random noise and the Ring effect limits the precision of the NO_2 retrievals, which consequently affects the aerosol size distribution retrievals as well. As indicated by the EOF analysis, the much greater number of channels, and the added expense that goes with it, do not produce more than a slight increase in the amount of retrievable aerosol information. A significant drop in

price would probably make the RSS the best choice, but as it stands, it probably serves best as a means of evaluating less costly devices, as was done in this study.

The CIMEL, too, has its strengths and weaknesses. Because of the 340 nm channel, the CIMEL makes measurements over a broader wavelength range than the RSS does. However, it is doubtful that the 340 nm channel provides any additional constraints on the coarse mode, where the information is most needed. In addition, NO_2 absorption declines very gradually as the wavelength decreases into the ultraviolet, remaining significant in the CIMEL's two UV channels. It has been very clearly demonstrated that overestimating NO_2 amounts will increase the retrieved values of the aerosol effective radius; likewise, underestimating NO_2 amounts by setting them to zero will produce errors of similar magnitude in the opposite direction. The almucantar retrievals can help constrain the aerosol size distribution, but superior results to retrievals using only optical depth data, especially where the coarse mode is concerned, have yet to be confirmed. The CIMEL is ideally suited for measurements in relatively clean regions where NO_2 values are known to be consistently low, and ozone values do not vary significantly from climatology.

Being the least expensive of the three devices, the MFRSR has the potential to be the superior device, but not before significant changes are made to the retrieval strategy and the choice of wavelengths used. Every facet of the retrieval strategy for the MFRSR needs to be revised, beginning with the fundamental assumptions on which the retrieval is based. Even with improved assumptions, the bimodal MFRSR retrieval may still have difficulties differentiating between aerosol extinction and gas absorption, especially of NO_2. These problems could conceivably be mitigated in future MFRSR data by

replacing one of the filters in the current version of the device with a filter centered on a different wavelength, whose inclusion would broaden the overall spectral range of the device. To test this possibility, a bimodal retrieval is performed on the "MFRSR equivalent" variant of the RSS data, first with the initial set of wavelengths, and then with an altered set of wavelengths.

6.2: Using the Bimodal Results to Improve the MFRSR Retrievals

The high-resolution retrievals of nitrogen dioxide using the RSS data have demonstrated the inadequacy of low-resolution, single-mode retrievals like the MFRSR algorithm of Alexandrov *et al*. (2002a) in separating the contributions of NO_2 and aerosols to the total extinction at 415 nm. The EOF analysis has indicated, and the retrieval results have confirmed, that the aerosol size distribution is bimodal, and that any attempt to retrieve aerosol properties needs to take this bimodality into account. With that in mind, the MFRSR retrieval strategy needs to be redesigned. The "MFRSR equivalent" retrieval can be modified for bimodal distributions, and the results can then be compared to those for the full RSS, high-resolution retrievals. Furthermore, one of the channels can be replaced in the "MFRSR equivalent" retrieval, with another channel for which filters exist. If the results of this retrieval show better agreement with the full RSS than the "MFRSR equivalent" retrieval does, then the existing MFRSR devices could be improved upon simply by substituting the filter at the replaced wavelength with a filter at the new one.

The EOF analysis of Chapter 3 has shown that the MFRSR contains most of the aerosol information that the RSS does. However, the lack of spectral resolution makes the contributions of ozone and nitrogen dioxide harder to separate from the aerosols. Accurately retrieving these gases depends on the accuracy of the a priori assumptions made about the aerosol size distribution. For example, it has already been demonstrated that assuming a single-mode size distribution will result in unacceptably large errors in the retrieved values of both gases. Only three independent items of aerosol information can be retrieved, but the optical depth, effective radius, and effective variance of two separate distributions need to be considered. The predominant value of the fine mode effective variance from both the RSS and CIMEL almucantar retrievals turned out to be 0.1, so the first assumption used in an alternative "MFRSR equivalent" retrieval of the RSS data is that the fine mode variance is 0.1. The remaining assumptions concern the coarse mode size distribution. The RSS and almucantar retrievals produced results which, while significantly different physically, are only slightly distinguishable from each other in a radiative sense. At present, it is not clear whether the actual coarse mode size distribution can be better expressed by an effective radius of 1.3 μm and an effective variance of 0.1, corresponding to the mean radius and dominant variance from the RSS retrieval, or an effective radius of 2.6 μm and an effective variance of 0.5, similarly derived from the results of the almucantar retrievals. The 1.3 μm size distribution is chosen for this analysis, in order to provide the most direct comparison with the full, high-resolution RSS retrieval.

Along with a revised "MFRSR equivalent" retrieval, additional five-channel retrievals using the RSS data are performed substituting either a channel at 375 nm or a

channel at 1034 nm for the one at 670 nm. Two retrievals are performed for each wavelength. The first solves for NO_2 along with ozone and the aerosol properties, and the second imports the results of the high-resolution NO_2 retrievals while solving for the remaining quantities. Table 6.1 shows the mean values of all the retrieved quantities for each of the different retrievals performed in this chapter. In addition, for the sake of comparison, the results from the sixteen-channel RSS results assuming constant variances are also presented. The error ranges were established by an analysis of the contour plots of residual optical depth, like those in Figures 5.1 and 5.2. The coarse mode optical effective radius cannot generally be determined with much precision, as multiple values produce acceptable answers. This, in turn, causes some uncertainty in the separation of coarse and fine mode optical depth at 550 nm, even if the devices are properly calibrated. Gas amounts, when retrieved by low-resolution methods, will face similar uncertainties.

When the new retrieval algorithm is applied to the "MFRSR equivalent" data set, the mean coarse mode optical depth for the data set is 0.013. This value is lower than what was retrieved from the RSS and CIMEL almucantar data sets before, when wavelengths above 1000 nm were included in the retrieval. In addition, as Figure 6.1 shows, negative values of the coarse mode optical depth were retrieved on a couple of days. Unfortunately, the retrieved NO_2 values presented in Figure 6.2 remain poorly defined and highly variable. Many days had values well in excess of 1 DU, while several days also had retrieved values at or below zero. The mean retrieved NO_2 amount was 0.74 DU, about double the mean obtained from the retrieval algorithm developed in Chapter 2, but much less than the means obtained from the single-mode RSS retrievals. The plot of ozone vs. day, shown in Figure 6.3, qualitatively resembles the other ozone

vs. day plots retrieved for the RSS data, with the lowest values occurring in November 1999 (after day 300) and a series of peaks following afterwards. The mean value of 260 DU is significantly higher than those retrieved using the previous "MFRSR equivalent" algorithm, indicating better initial assumptions, although slightly smaller than the mean obtained using the single-mode retrieval with greater spectral resolution, and more significantly smaller than the mean obtained from the full-resolution, separate-gas retrieval. This indicates that the revised assumptions have improved the quality of the ozone results for the 5-channel retrieval, but still not enough to match the higher-resolution retrievals. The plot of fine mode effective variance vs. day, shown in Figure 6.4, reflects the same seasonal dependence seen in Figure 5.29. The large mean value of 0.241 μm results from the small retrieved coarse mode optical depth values. The extra optical depth attributed to the fine mode brings the ratio of the fine mode aerosol extinction at any wavelength to the extinction at 500 nm closer to one, consistent with a larger particle size.

6.3: Finding the Five Best Wavelengths for a New MFRSR

As indicated by the mean values presented in Table 6.1, for neither wavelength combination is the agreement for ozone or the aerosol properties especially good or especially poor, but the retrievals with the 375 nm channel produce clearly superior mean values for NO_2. The superiority becomes even more obvious in Figure 6.5, a comparative set of plots vs. day of the retrieved NO_2 values for the full RSS, the MFRSR, and the five-channel retrievals using both 375 nm and 1034 nm. The 1034 nm retrieval only has

a handful of days where the retrieved NO_2 value exceeds zero. In addition, the 1034 nm retrieval experiences a similar problem obtaining nonzero values for the coarse mode optical depth. The mean RMS residuals are also lower for the retrievals with the 375 nm. Therefore, it is concluded that substituting a filter in the neighborhood of 375 nm for the 670 nm filter will produce the most accurate bimodal MFRSR retrievals, provided that accurate values for NO_2 cannot be obtained by other means. However, if the MFRSR is located at a site where accurate values of column amounts of NO_2 are available, the best choice of a new wavelength is 1034 nm. Table 6.1 clearly shows that when the results of the high-resolution NO_2 retrieval are input into the new five-channel retrievals, the retrieval with the 1034 nm channel included produces results for the other measured quantities that much more closely agree with the sixteen-channel retrieval assuming constant variances than the retrieval with the 375 nm channel. Therefore, the ideal channel to substitute into the MFRSR depends on how much additional data for a given site is available.

Retrievals are also performed for the new "MFRSR equivalent" wavelengths with NO_2 set to 0 and ozone values set to the high-resolution RSS amounts, to parallel the CIMEL optical depth measurements. In the case with a 375 nm channel, there is a slight addition to the mean fine mode optical depth, with a corresponding subtraction from the coarse mode optical depth. The mean value for the fine mode effective radius does not change significantly. However, when the 1034 nm channel is used instead, the mean value of the fine mode effective radius drops dramatically. The coarse mode optical depth increases, naturally at the expense of the fine mode optical depth. This proves that while the wavelength combination that includes the 1034 nm will produce more accurate

retrievals if NO_2 can be measured independently, the aerosol retrievals with this wavelength combination are far more sensitive to errors in NO_2 than those with the 375 nm channel would be.

In Figure 6.6, the values for the quantities obtained by the five-channel retrieval using a 375 nm channel are plotted against the corresponding values from the full RSS retrieval with the effective variances set to 0.1. The optical depth values for both modes show a strong correlation. The y-intercept is positive for the fine mode and negative for the coarse mode, but the slope is slightly less than one in both cases. While the y-intercept values confirm that on low optical depth days, the fine mode optical depth is getting overestimated while the coarse mode is underestimated. The low slopes for both graphs indicate two additional features, however. First, the error in the fine mode optical depth decreases as the optical depth increases. Second, some coarse mode optical depth gets attributed to the gases as well. The fine mode effective radius is not as well correlated as either optical depth. The slope is still close to one, but there is a significant positive offset, primarily due to higher overall values of nitrogen dioxide and the fine mode effective variance. The plot of the retrieved ozone values is also reasonably linear, but the slope is more significantly lower than one. Unfortunately, the correlation with the NO_2 values is very limited, so while the mean values remain in fairly good agreement, it is clear that nitrogen dioxide absorption is still not being clearly separated from aerosol extinction. Figure 6.7 shows the series of comparative plots for the retrievals using the 1034 nm channel and the high-resolution NO_2 retrieval results. The correlation is superior for the fine mode effective radius, and the slopes for the other quantities are all closer to one.

6.4: The "CIMEL Equivalent" Retrieval

To complete the study, wavelengths corresponding to the six CIMEL wavelengths other than 340 nm were selected from the RSS data for a "CIMEL equivalent" retrieval. To most fully approximate CIMEL optical depth data, the high-resolution ozone values were subtracted as climatology, and nitrogen dioxide values were set to zero. As shown in Table 6.1, the mean retrieved values very closely mirror those of the five-channel retrieval using the 375 nm channel. This indicates that the additional channel beyond 1000 nm does not improve the retrieval results significantly. The limited wavelength range of the RSS prevents the CIMEL's 340 nm channel from being evaluated in this manner. Presumably, extending the wavelength range would further reduce the sensitivity of the aerosol retrieval to errors in the measured or assumed values of NO_2. Miscalculation of the optical depth due to Rayleigh scattering would likely be a greater source of error, in fact, when the 340 nm wavelength is included.

6.5: Looking Forward

The bimodal analysis of the data from the various devices resolves a number of issues, but leaves others vague. On the positive side, as the EOF analysis suggested, the two aerosol modes can be reasonably well separated, and the fine mode effective radius adequately measured, even with only five channels. When the fine mode effective variance is fixed at 0.1, a seasonal pattern emerges in the RSS effective radius data. This pattern remains, regardless of the choice of wavelengths used. Replacing the 670 nm

channel with a 375 nm channel would improve the quality of the NO_2 and fine mode effective radius retrievals in the MFRSR, without adversely affecting the other quantities. However, the ability of the MFRSR to accurately measure ozone and nitrogen dioxide on a consistent basis remains limited. No NO_2 retrieval based on a low-resolution combination of wavelengths produced good agreement with the high-resolution full RSS. Only with the 375 nm channel replacing the 670 nm did the "MFRSR equivalent" retrieval produce a mean NO_2 amount within a factor of two of the high-resolution mean, and even there, the daily values were very poorly correlated. The ability to retrieve ozone well using Chappuis band data, and to unambiguously define the coarse aerosol mode, remain limited by the wavelength range used in these sun photometer devices.

Over the short term, the existing MFRSR data can be re-evaluated using bimodal retrievals. While the quality and accuracy of the results should improve, the amount of obtainable information will necessarily remain limited. The 670 nm filters can be replaced by existing filters with central wavelengths around 375 nm, thereby further improving future data obtainable with the existing network of devices. However, the EOF analysis of Chapter 3 suggests that a sun photometer would need filters extending more deeply into the infrared in order to clarify the properties of the coarse aerosol mode. Lacking these additional wavelengths, external sources of information will be needed to fill in the gaps in information. One such source is polarimetric data, coming from devices like the Research Scanning Polarimeter (RSP). The RSP has been used to monitor aerosols from high-altitude airplanes (Chowdhary *et al.* 2001), and has two advantages that distinguish it from other ground-based detectors. First, it has the broadest spectral range, extending from 410 nm to 2250 nm. Second, the RSP measures

polarization in addition to intensity, providing further information with which to distinguish aerosol characteristics. These two features make the RSP better capable of distinguishing between the coarse and fine aerosol modes than any of the sun photometer devices used in this study. With RSP data over a site equipped with an RSS or CIMEL device, the properties of the coarse aerosol mode should be clearly determined, and the accuracy of the values of the coarse mode effective radius retrieved by both sun photometer can then be evaluated.

The other site at which an RSS is stationed, the North Slope site at Barrow, Alaska, also has data from which empirical orthogonal functions can be determined. This data can be used to show how EOF's change with location, and to what extent changes in the EOF correspond to changes in the retrieval results. Also, the conclusion from this chapter that the MFRSR could be improved by changing a filter should be verified, by means of a comparative analysis between an MFRSR with the original filters, an MFRSR with the new filters, and an RSS. These projects would expand upon the work of this present study, and hopefully also lead to further improvements in ground-based monitoring of aerosols and gases.

Still, each of the three devices examined in this study has relative strengths and weaknesses, and none has proven clearly superior to the other two. If money were no object, then the RSS would be best suited for use in a combined aerosol/gas monitoring network. On the other hand, the CIMEL works nicely when measuring ozone and nitrogen dioxide amounts is not a priority. Perhaps something higher than zero ought to be assumed for the nitrogen dioxide value in CIMEL retrievals, but the two ultraviolet channels dampen the sensitivity of the aerosol retrievals to NO_2 errors. The key issue for

the CIMEL, provided that no device defects render certain channels useless over the course of a data set, is correct representation of the Rayleigh optical depth. However, nitrogen dioxide and ozone might be worth measuring in areas affected by biomass burning or industrial pollution, and the combination of a CIMEL and a device designed specifically to measure NO_2 might not prove as cost-effective as an RSS. The MFRSR suffers from similar drawbacks. With a 375 nm filter replacing the one at 670 nm, the error in aerosol retrievals resulting from incorrect NO_2 values diminishes, but does not go away entirely. If a 1034 nm filter is used instead, the MFRSR retrievals can very closely match those of the RSS, but only if NO_2 is measured by other means. The only possible means of accomplishing this without an additional device, and consequently additional expenses, would be to use the diffuse intensity measured by the 415 nm filter. No algorithm for using the diffuse to measure NO_2 presently exists, but such an algorithm could be easily tested on RSS data once devised.

	fine mode τ	fine mode r_{eff} (μm)	coarse mode τ	coarse mode r_{eff} (μm)	ozone (DU)	NO_2 (DU)	RMS residual τ
RSS 16-channel, bimodal, fixed v_{eff}	0.047 (±0.003)	0.185 (±0.01)	0.019 (±0.003)	1.33 (±0.8)	279	0.38	0.00161
bimodal "MFRSR equivalent"	0.053	0.241	0.013	1.3	260	0.74	0.00012
with 375 nm	0.052	0.216	0.015	1.3	265	0.46	0.00013
with 1034 nm	0.048	0.169	0.019	1.3	269	0.14	0.00047
with 375 nm, using hi-res NO_2	0.052	0.209	0.016	1.3	264	0.38	0.00050
with 1034 nm, using hi-res NO_2	0.048	0.184	0.019	1.3	270	0.38	0.00061
with 375 nm, NO_2=0, hi-res O_3	0.051	0.182	0.017	1.3	279	0.00	0.00144
with 1034 nm, NO_2=0, hi-res O_3	0.044	0.138	0.022	1.3	279	0.00	0.00141
"CIMEL equivalent"	0.050	0.182	0.018	1.3	279	0.00	0.00121

Table 6.1: The mean values of the quantities obtained by the different five-channel bimodal RSS retrievals, along with the "CIMEL equivalent" retrieval. The sixteen-channel fixed variance RSS retrieval is also shown, for the sake of comparison.

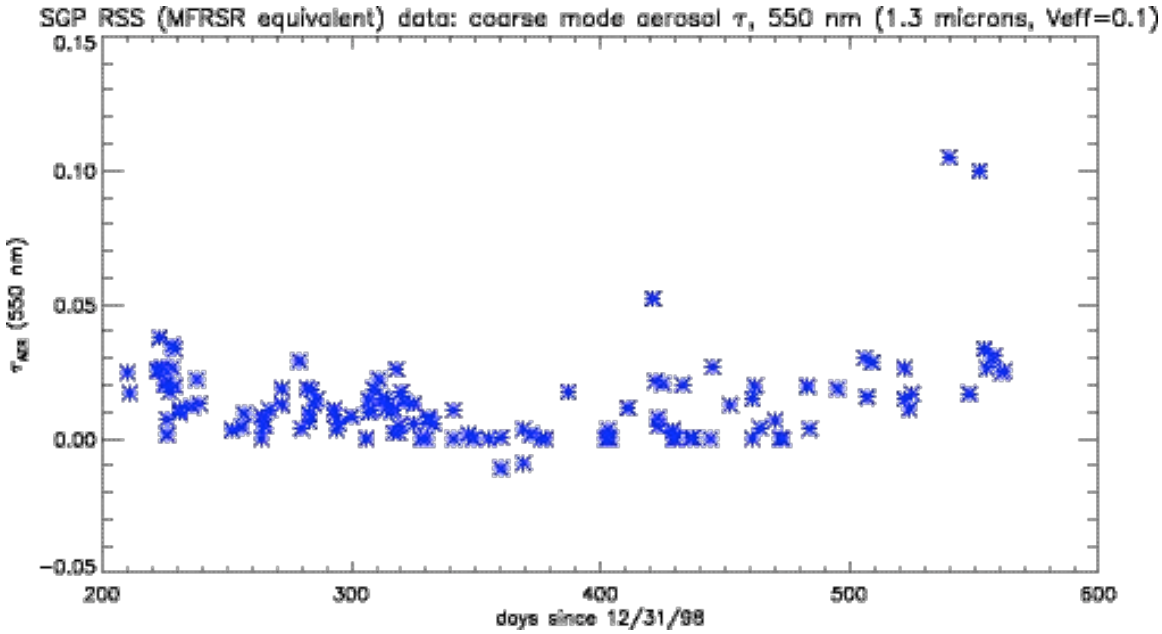

Figure 6.1: The plot of coarse mode optical depth vs. day for the bimodal "MFRSR equivalent" retrieval of the RSS data.

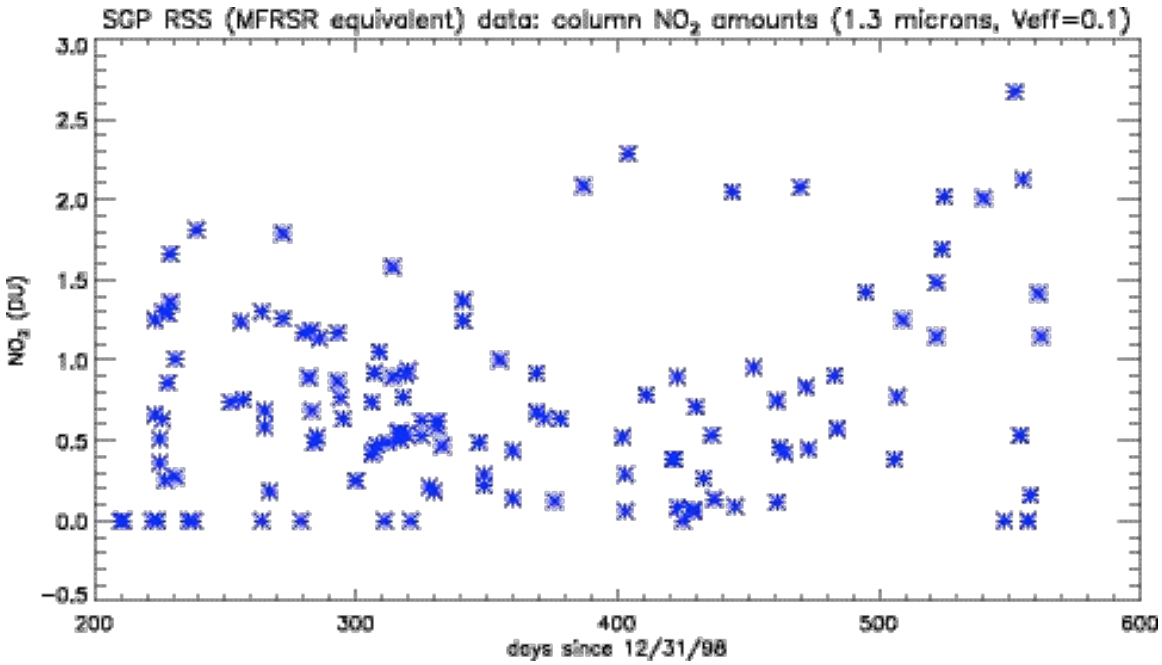

Figure 6.2: The plot of nitrogen dioxide amounts vs. day for the bimodal "MFRSR equivalent" retrieval of the RSS data. A handful of values are excessively large, and a handful are at or lower than zero.

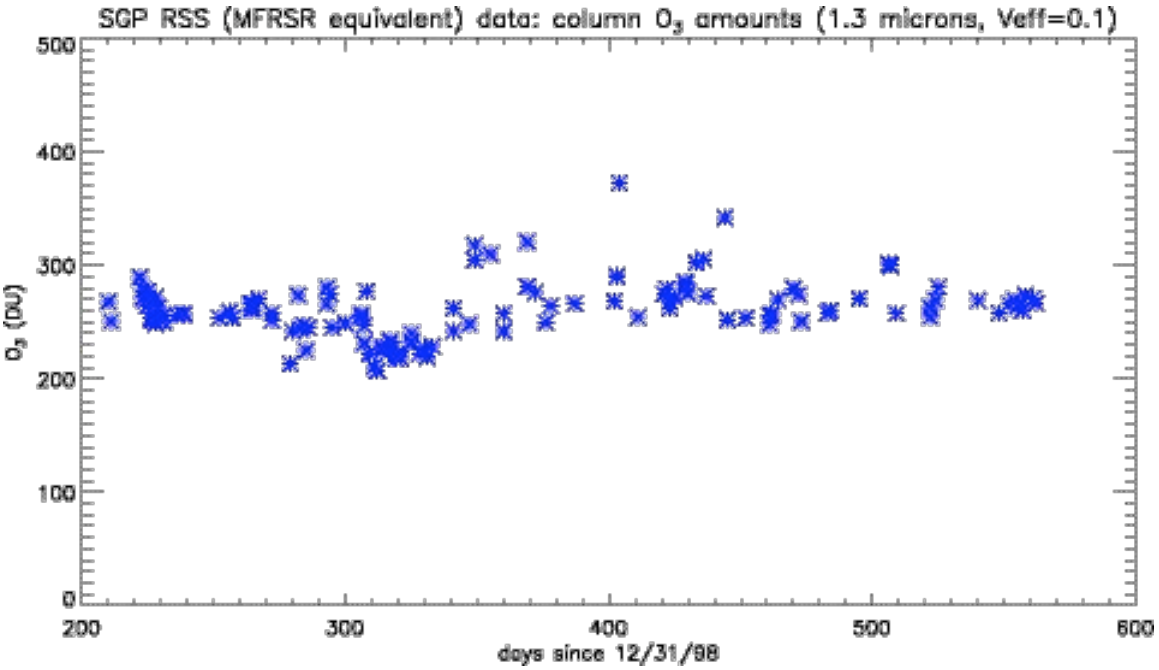

Figure 6.3: The plot of ozone vs. day for the bimodal "MFRSR equivalent" retrieval of the RSS data.

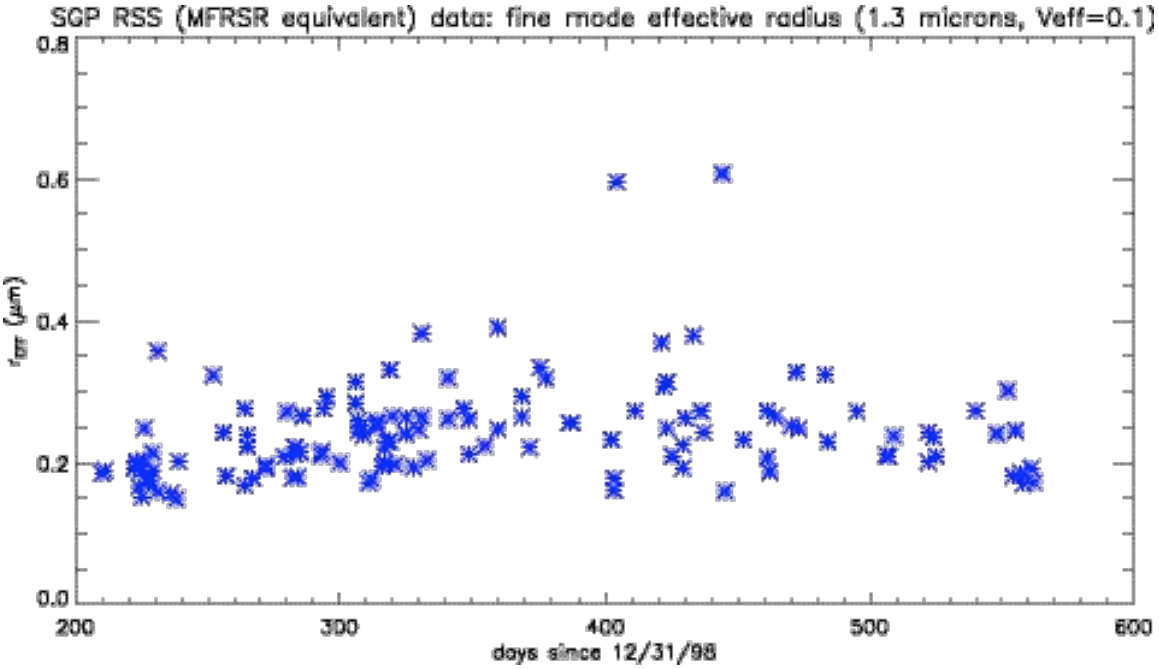

Figure 6.4: The plot of the fine mode effective radius vs. day for the bimodal "MFRSR equivalent" retrieval of the RSS data. The seasonal dependence of the effective radius persists in this graph.

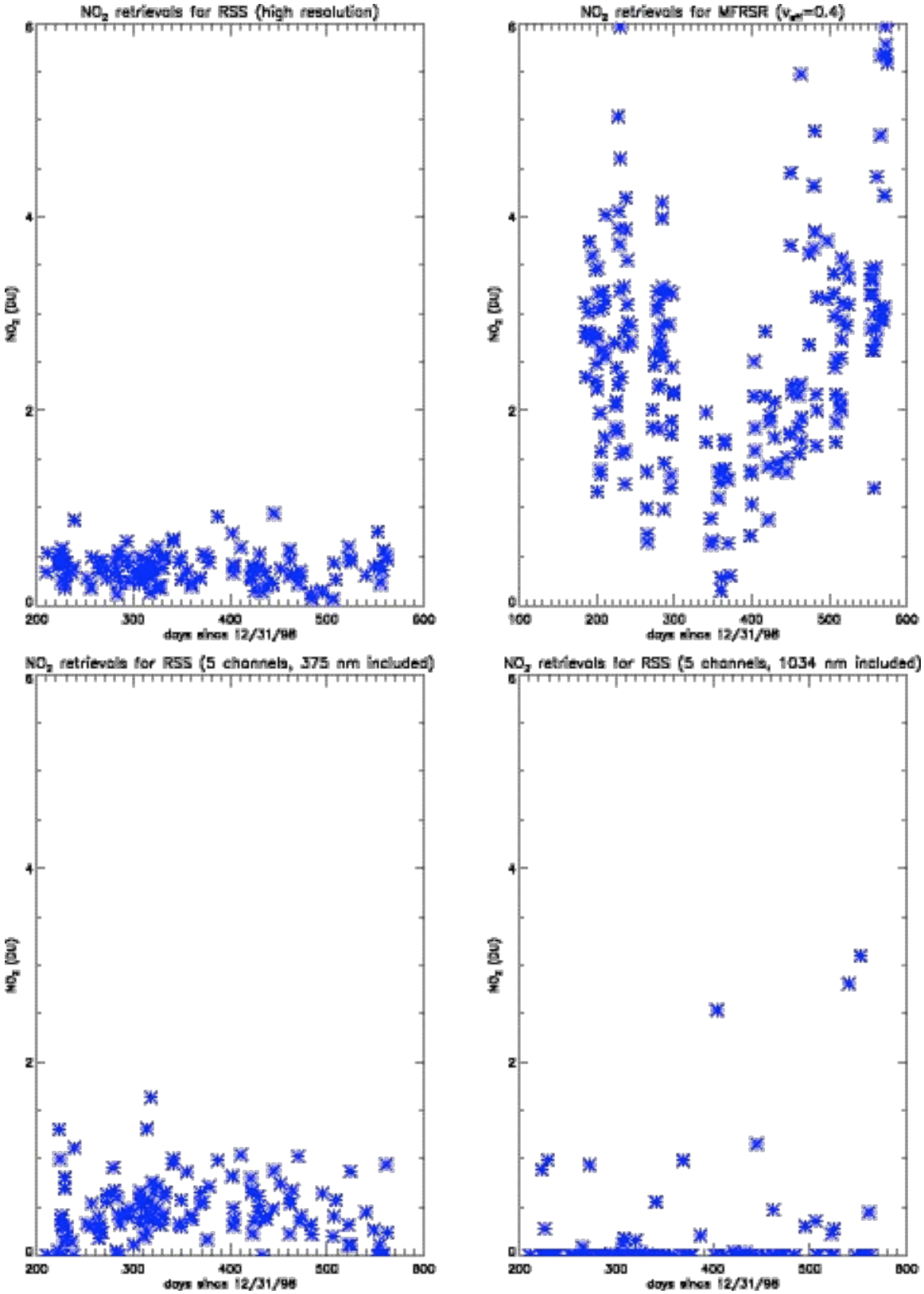

Figure 6.5: The plots of NO_2 vs. day for the full RSS, the MFRSR, and the five-channel RSS retrievals using 375 nm and 1034 nm.

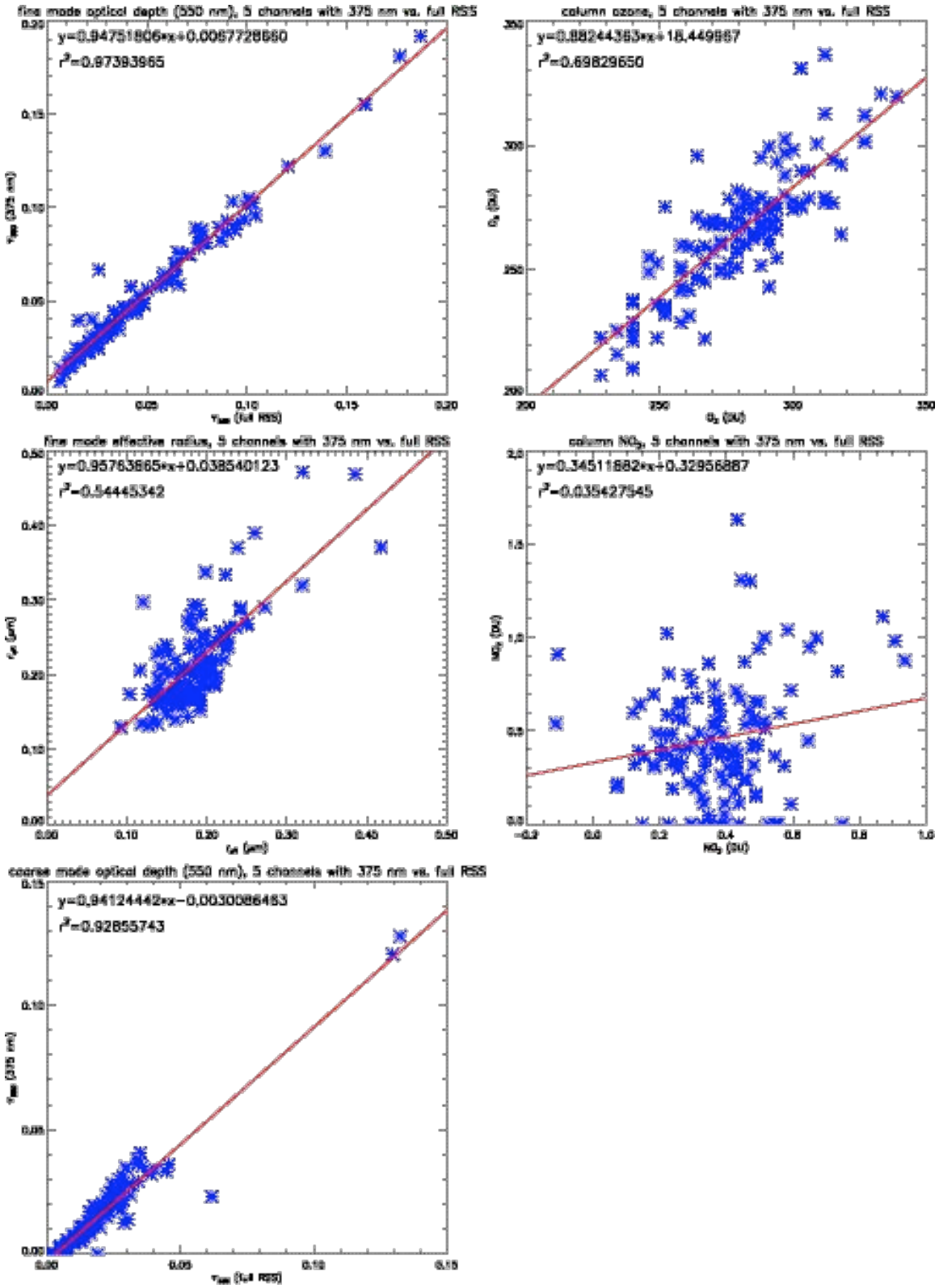

Figure 6.6: The quantities obtained from the five-channel RSS retrieval using 375 nm in place of 670 nm, plotted vs. the corresponding quantities obtained from the full RSS retrieval.

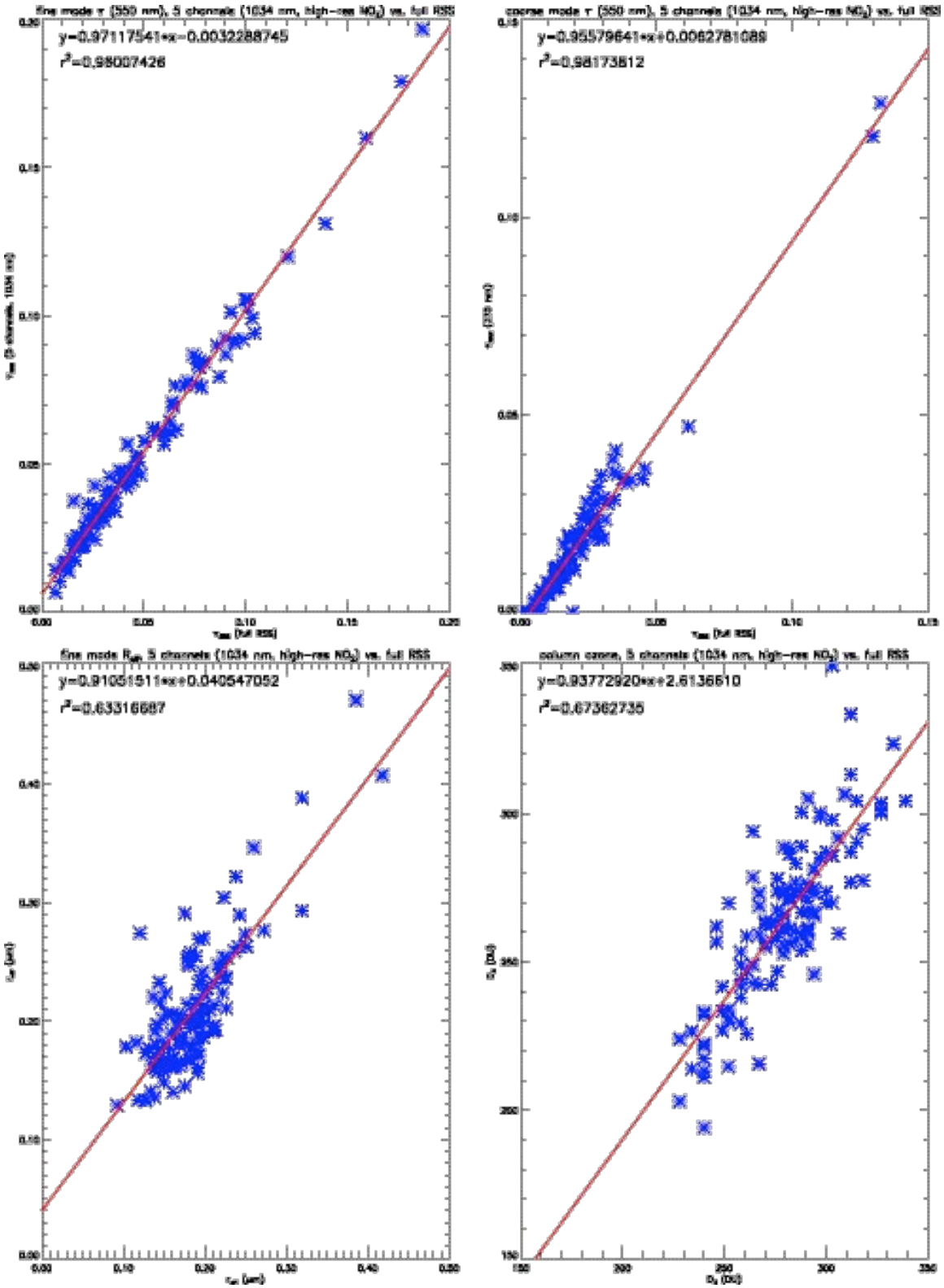

Figure 6.7: The quantities obtained from the five-channel RSS retrieval using 1034 nm in place of 670 nm and inputting the high-resolution NO_2 values, plotted vs. the corresponding quantities obtained from the sixteen-channel RSS retrieval with the effective variances held constant.

Chapter 6 References

Alexandrov, M. D., Lacis, A. A., Carlson, B. E., and Cairns, B., *Journal of the Atmospheric Sciences* **59**, 524-543 (2002a).

Chowdhary, J., Cairns, B., Mishchenko, M., and Travis, L., *Geophysical Research Letters* **28**, 243-246 (2001).

Appendix

The algorithms to retrieve aerosol properties and NO_2 column amounts using RSS data were developed on IDL Version 5.0 for UNIX, and are available on request by e-mailing me at sgianell@kdist.giss.nasa.gov.

The MFRSR retrieval algorithm was designed and developed by Mikhail Alexandrov, also using IDL. He can be contacted at malexandrov@giss.nasa.gov. The CIMEL data was obtained from the AERONET website, which can be accessed at http://aeronet.gsfc.nasa.gov/.

The SCIATRAN radiative transfer code, whose Raman scattering subroutines were used to analyze the Ring Effect, can be accessed at http://www.iup.physik.uni-bremen.de/sciatran/.