Stratospheric Aerosol Retrieval From OSIRIS Limb Scattered Sunlight Spectra

A Thesis

Submitted to the Faculty of Graduate Studies and Research

in Partial Fulfillment of the Requirements

for the Degree of

Doctor of Philosophy

in the Department of Physics and Engineering Physics by

Adam E. Bourassa

Saskatoon, Saskatchewan

April, 2007

In presenting this thesis in partial fulfillment of the requirements for a Postgraduate degree from the University of Saskatchewan, the author agrees that the Libraries of this University may make it freely available for inspection. The author further agrees that permission for copying of this thesis in any manner, in whole or in part, for scholarly purposes may be granted by the professor who supervised this thesis work or, in his absence, by the Head of the Department or the Dean of the College in which this thesis work was done. It is understood that any copying or publication or use of this thesis or parts thereof for financial gain shall not be allowed without written approval from the author. It is also understood that due recognition shall be given to the author and to the University of Saskatchewan in any scholarly use which may be made of any material in this thesis.

Requests for permission to copy or to make other use of material in this thesis in whole or in part should be addressed to:

Head of the Department of Physics and Engineering Physics
116 Science Place
University of Saskatchewan
Saskatoon, Saskatchewan
Canada
S7N 5E2

To Amy

A.M.D.G.

Abstract

The recent development of satellite observations of limb scattered sunlight at optical wavelengths has afforded a new opportunity to measure the vertical structure of atmospheric composition from the upper troposphere to the mesosphere, on a global scale. The determination of profiles of atmospheric composition from observed limb radiance profiles requires two elements, a forward radiative transfer model and a species specific inversion algorithm. In this work, the development of a new, fully spherical, successive orders, radiative transfer model, SASKTRAN, for the analysis of limb scattered sunlight is presented. The model is incorporated into a novel relaxation algorithm that employs spectral ratios to retrieve profiles of stratospheric aerosols from limb radiance measurements collected by the Canadian OSIRIS instrument on the Odin satellite.

The SASKTRAN forward model results compare favorably with both OSIRIS observations as well as with other radiative transfer model calculations while remaining computationally practical for the operational inversion of large satellite data sets.

The spectral ratio relaxation algorithm is able to retrieve aerosol number density profiles at stratospheric altitudes from limb radiance profiles assuming the height profile of the aerosol particle size distribution is known. The equivalent aerosol extinction derived from the OSIRIS measurements at visible wavelengths agrees with coincident occultation measurements from other satellite instrumentation to within 15% when a size distribution appropriate for background aerosol conditions is used. Finally, it is demonstrated that the incorporation of simultaneous infra-red observations at 1530 nm into the inversion yields a useful proxy for the aerosol size distribution parameters.

Acknowledgements

This work was made possible by the Canadian Space Agency and the National Sciences and Engineering Research Council. It was also financially supported by the University of Saskatchewan and the Institute of Space and Atmospheric Studies. I am personally very grateful for the dedication, support, generosity and encouragement of my supervisors Drs. Ted Llewellyn and Doug Degenstein. They gave me a chance to learn both independently and together with them, and provided much opportunity as part of a truly international collaboration. I would also like to express my sincere appreciation to Drs. Nick Lloyd and Dick Gattinger as they were always available for discussion and advice. Finally, I am most grateful for the unwavering support of my family, especially my wife, Amy.

Contents

List of Tables

List of Figures

List of Abbreviations

API	Application Programming Interface
CCD	Charge Coupled Device
CDI	Combined Differential-Integral radiative transfer model
CDI-PI	Combined Differential-Integral Picard Iterative radiative transfer model
ECMWF	European Center for Medium range Weather Forecasting
GSLS	Gauss-Seidel Limb Scatter radiative transfer model
GSS	Gauss-Seidel Spherical
HALOE	Halogen Occultation Experiment
IRI	Infra-Red Imaging module of OSIRIS
LIMBTRAN	Limb scatter radiative transfer model
LORE	Limb Ozone Retrieval Experiment
MART	Multiplicative Algebraic Reconstruction Technique
MCC++	Monte Carlo C++ radiative transfer model
NOAA	National Oceanic and Atmospheric Administration
NPOESS	National Polar-Orbiting Operational Satellite System
OIRA	Oxygen Infra-Red Atmospheric band
OMPS	Ozone Monitoring and Profiling Suite
OS	Optical Spectrograph module of OSIRIS
OSIRIS	Optical Spectrograph and Infra-Red Imaging System
POAM	Polar Ozone and Aerosol Measurement instrument
PSC	Polar Stratospheric Cloud
RAM	Random Access Memory
SAGE	Stratospheric Aerosol and Gas Experiment
SAM	Stratospheric Aerosol Measurements
SASKTRAN	University of Saskatchewan Radiative Transfer Model
SBUV	Solar Backscatter Ultraviolet instrument
SCIAMACHY	Scanning Imaging Absorption Spectrometer for Atmospheric Chartography
Siro	Reference Monte Carlo radiative transfer model
SME	Solar Mesospheric Explorer
SOLSE	Shuttle Ozone Limb Sounding Experiment
TOMS	Total Ozone Mapping Spectrometer
UV	Ultra-Violet
VER	Volume Emission Rate
WMO	World Meteorological Organization

Chapter 1

Introduction

Satellite based remote sensing techniques are able to probe the atmosphere globally, frequently and repetitively. The data that these techniques can provide are invaluable to the understanding and monitoring of atmospheric processes. The observation of electromagnetic waves created by atmospheric processes or modified by interaction with the atmosphere forms the basis of the measurement. Optical instrumentation, measuring the ultraviolet, visible, and near infrared regions of the spectrum, has been used extensively in the past with much success in measuring processes from the upper troposphere to the mesosphere and beyond.

Almost all optical remote sensing of upper troposphere and the stratosphere, falls into two categories of measurement: solar occultation for high vertical resolution, and nadir sounding for good global coverage. More recently, a new generation of satellite instruments has employed an optical technique known as limb scatter. This measurement technique is a "side-view" observation of the radiance of sunlight scattered from the atmosphere. The spectrum of light from the sun, in the process of traveling through the atmosphere and scattering off molecules and particles, possibly several times, is imprinted with the signatures of the atmospheric composition. The observation of the limb, or side-view, radiance of the atmosphere from a satellite provides the benefits of both an occultation and a nadir sounding experiment: vertical

resolution combined with global coverage of the sunlit atmosphere.

The Optical Spectrograph and InfraRed Imaging System (OSIRIS), an optical limb scatter instrument designed and built in Canada is currently deployed on the Swedish Odin satellite. The objective of the mission is to provide new information on the extent to which humans are changing the atmospheric environment, specifically through the study of stratospheric and mesospheric ozone, and the coupling and energy budget of these regions. Stratospheric aerosols, while not explicit in the mission statement, are an important component of the radiative and chemical balances of the atmosphere and have direct effects on ozone and climate. This thin, fine mist of particles exists naturally, but is highly variable as volcanic eruptions can strongly perturb the concentration of stratospheric aerosol on a global level. Anthropogenic pollution also plays an important role in the creation and maintenance of the stratospheric aerosol layer. As such, it is valuable to use limb scatter observations of the atmosphere to measure and monitor stratospheric aerosol properties.

The limb scattered radiance is a complicated non-linear function of the atmospheric state. Therefore, in order to infer a particular aspect of the atmospheric state from the measurements there are two requirements. The first is a set of calculations, often referred to as a model, that predicts the measurements from the atmospheric state through a description of the physical processes. The second requirement is a method to work backwards from a set of measurements, using the model, to determine the atmospheric state at the time of measurement. The work in this thesis falls distinctly into these two parts: the development of a radiative transfer code appropriate for modelling limb scatter effects of stratospheric aerosol, and the inversion of the OSIRIS measurements to yield stratospheric aerosol properties.

The necessary background on stratospheric aerosol properties and an explanation of the limb scatter measurement technique, with details of the OSIRIS instrument are presented in Chapter 2. The inverse problem is formally defined in Chapter 3 and some common methods of solution presented. Specifically, an adaptation of the

Multiplicative Algebraic Reconstruction Technique (MART), previously used successfully in medical imaging and atmospheric tomography, is introduced as a method for retrieving the vertical profile of an atmospheric species from limb scatter. An application of a standard formal error analysis to the MART adaptation is also presented.

Chapter 4 includes the description of the development of a radiative transfer model, named SASKTRAN, that is suitable for the analysis of limb scatter measurements. As limb scatter is a relatively new technique, a standard model does not yet exist. Currently existing models broadly fall into two categories: approximate solutions, for example in a plane-parallel atmosphere, that are suitably efficient for use in an inversion but are known to have systematic error, and fully spherical, often statistical, models that consume large amounts of computational time. The goal in the development of SASKTRAN was to create a model with spherical geometry, suitable for the limb scatter measurements, with an efficient solution that could be used operationally for the inversion of a large satellite data set. A secondary, but important, requirement of the model is a robust and physically meaningful treatment of stratospheric aerosols.

Chapters 5 and 6 present the development of a retrieval technique for stratospheric aerosol number density and particle size distribution from limb scattered sunlight using the SASKTRAN model and the MART inversion. The OSIRIS measurements are used to demonstrate the applicability and feasibility of the technique. A significant portion of the work in this section is dedicated to the analysis of error terms in the inversion. The important findings of the thesis work and the future for the developed techniques are summarized in Chapter 7.

Chapter 2

Background

2.1 Stratospheric Sulphate Aerosols

2.1.1 Volcanoes and the Junge Layer

In 1883, the massive eruption of the island volcano Krakatoa produced brilliant red sky glow persisting long after sunset. These sunsets occurred for years afterward and were observed at locations around the world. Scientists of the day realized the sunset glow phenomena was the evidence of long lifetime, high altitude particles that were circulated in the upper atmosphere on a global scale (*Symons*, 1888). Many years later, the existence of particles at stratospheric altitudes, not directly associated with a volcanic event, was hypothesized by *Gruner and Kleinert* (1927) based on the long lasting purple colour of twilight.

The first direct confirmation of aerosols up to 30 km altitude was by in-situ measurements performed on a high altitude balloon experiment in 1959 (*Junge et al.*, 1961). In this experiment, particle densities were measured in two size ranges: 0.01 to 0.10 μm, and 0.1 to 1.0 μm. The smaller particles were found to decrease in concentration with altitude down to an undetectable amount around 20 km altitude. A surprising result of the experiment was that the concentration of the larger size particles peaked in a ubiquitous layer near 20 km. Analysis of the larger particles found

high concentrations of sulphur. The sulphur content was confirmed 10 years later by *Rosen* (1971) who, by testing the boiling point of stratospheric aerosol particles, found the composition to be consistent with a solution of 75% sulphuric acid, H_2SO_4.

Because there had not been a significant volcanic event in almost three decades (*Stothers*, 1996), Junge postulated correctly that the layer of stratospheric aerosol is formed within the stratosphere by oxidization of trace gases that originate in the troposphere and are transported to stratospheric altitudes. While the discovery of this background stratospheric aerosol layer, often termed the Junge Layer, was landmark, the experiment was limited in scope and further quantification and confirmation of the source was required.

Volcanic eruptions, with the potential to inject massive amounts of sulphur and ash directly into the stratosphere, sometimes to altitudes of 30 km, have an immense impact on the formation, maintenance, and distribution of stratospheric aerosol. Large eruptions, such as El Chichón in 1982 (12 Tg) and Mount Pinatubo in 1991 (30 Tg) caused very significant perturbations of the particles in the stratosphere. In fact, the aerosol load in the years since the discovery of the background layer has been dominated by volcanic activity. Eight volcanic events with significant impact on the stratosphere have occurred in the past 120 years, four occurring between 1880 and 1910, and four occurring between 1963 and 1991 (*Stothers*, 1996; *Deshler et al.*, 2006). The completeness of the relaxation of the aerosol layer to background levels during volcanically quiescent periods is somewhat contentious; however, aside from the time period of the original experiments in the early 1960's, *Deshler et al.* (2006) identified four time periods in the modern age of remote sensing with minimal volcanic influence: 1974, 1978-1980, 1988-1991, and 1997–present.

Part of the reason that the background aerosol level is somewhat disputed is that the lifetime of a stratospheric sulfate aerosol particle is relatively long. For example, following the first year after the Pinatubo eruption during which the aerosol was transported in the stratospheric circulation around the globe, the volcanic aerosol

mass in the stratosphere decreased with an exponential decay rate of approximately 1 year (*McCormick et al.*, 1995). The lack of a significant event in over two decades since the eruption means that in recent years the aerosol layer has reached equilibrium levels not observed since its discovery. Trend analysis of several long term data sets, though admittedly difficult, has revealed that the *background* level of the aerosol has not experienced significant change since the 1970's (*Deshler et al.*, 2003, 2006).

2.1.2 Climate and Chemistry

Aerosols in the stratosphere affect both the radiative balance and the chemistry of the atmosphere. Radiatively, the *direct* effect of aerosols in the stratosphere is to increase the planetary albedo by the scattering of visible solar radiation, which results in a cooling of the Earth. Conversely, but with an efficiency that depends on the size distribution, aerosols can also scatter the thermally emitted infrared radiation from the surface and contribute to global warming through the greenhouse effect. For typical stratospheric sulphate aerosol sizes, even volcanic, the cooling effect dominates over the warming (*Lacis et al.*, 1992; *Valero and Pilewskie*, 1992). Additionally, absorption of thermal radiation can result in local warming of the stratosphere. Aerosols also have an *indirect* radiative effect in that they may act as cloud condensation nuclei and modify cloud formation (for example *Le Treut et al.*, 1998).

These radiative effects are quite small for background levels of stratospheric aerosol; however, after a large volcanic event they are very significant (*McCormick et al.*, 1995). The effect of the Mount Pinatubo eruption in June of 1991 has been the topic of much discussion and many measurements and modelling studies were undertaken to assess the impact on the stratosphere and on the Earth's climate. For example, the monthly mean clear sky solar irradiance measured at Mauna Loa, Hawaii, a location quite remote from the Philippines, decreased by as much as 5% in the months following the eruption, and measurements of the global average lower tropospheric temperature were decreased from the average by 0.5°C in September of

that year (*Dutton and Christy*, 1992). Aerosols from the Pinatubo eruption also triggered a stratospheric warming with temperatures rising over 3 standard deviations above the 26-year mean during the following four months (*Labitzke and McCormick*, 1992).

Within the polar vortex, sulphate aerosols play a role in the formation of polar stratospheric clouds (PSC's) as cloud condensation nuclei and, therefore, contribute to the heterogeneous chemistry that leads to polar springtime chlorine activation and ozone destruction (*Solomon*, 1999). Cold conditions and elevated levels of aerosol loading outside the vortex can also contribute to ozone destruction through similar surface chemistry on the cold aerosol particle without the actual formation of a stratospheric cloud (*Erle et al.*, 1998). Other trace gas chemical reactions are significantly affected by sulphate aerosols, even at background levels. For example, the partitioning of NO_x/NO_y, which is important in the ozone cycle, is sensitive to low levels of aerosol surface area (*Crutzen*, 1970).

The radiative effects of stratospheric aerosol are of increasing importance in the study of the Earth's climate. The global cooling caused by the Pinatubo eruption exceeded the warming due to anthropogenic greenhouse gases for 18 months following the eruption, and remained significant for about three years. Although there is a reasonably large degree of uncertainty in the calculation, *Hansen et al.* (2002) estimate that the global positive radiative forcing of greenhouse gases in 1991 was $1.5\,\mathrm{W/m^2}$, and that the negative forcing of the volcanic aerosol was as high as $3.0\,\mathrm{W/m^2}$. By the end of the decade, the radiative effects of the aerosol had decayed to negligible levels while the greenhouse gas forcing continued to increase.

Climate change is most literally a hot topic. The Intergovernmental Panel on Climate Change (IPCC, 2001) document, estimates a global warming due to anthropogenic emission of 1.4–5.8 °C by the year 2100 at a rate of warming greater than any period in the last 10 000 years. Most climate predictions assume a volcanic activity level in the upcoming century identical to that of the past century; however, there

has been some recent suggestion of "geoengineering" climate stability by the deliberate creation of stratospheric aerosol. *Wigley* (2006) suggests a periodic injection of aerosol precursor gases into the stratosphere at levels similar to the Mt. Pinatubo eruption to provide cooling to offset the effect of greenhouse gas emissions. This scenario is proposed in order to provide some time to allow the economic and technological development to support a significant decrease in greenhouse gas emissions.

2.1.3 Source Gases and Microphysics

Stratospheric aerosol particles are small liquid droplets of sulphuric acid and water created from nucleation and condensation of molecules of H_2SO_4 that are formed from sulphur bearing gases that originate in the troposphere. The most important of these source gases are carbonyl sulphide (OCS) and sulphur dioxide (SO_2). OCS is the most abundant form of naturally occurring sulphur in the troposphere. It is created by marine processes and is absorbed by oceans, plants and soil (*Kettle et al.*, 2002). Because it has a relatively long atmospheric lifetime and low solubility, significant amounts of OCS can reach the stratosphere where photolysis and oxidation reactions lead to the creation of H_2SO_4 (*Crutzen*, 1976). SO_2, which also oxidizes to H_2SO_4 in the atmosphere, is a common pollutant from industrial processes used for more than a century. Because it is produced anthropogenically, its distribution in the troposphere is highly variable and tends to be concentrated over highly populated areas. As it is a pollutant, it is under emission control; the concentration over North America and Europe reached a maximum in the 1980's and has since been slowly decreasing (*Hicks et al.*, 2002). However, SO_2 remains the largest source of sulfur in the atmosphere. Because it is highly volatile and tends to react before it can reach the stratosphere, it is estimated that OCS and SO_2 contribute approximately equally to the background sulphate aerosol layer (*Thomason*, 2006). Large volcanic eruptions have the potential to inject large amounts of SO_2 directly into the stratosphere where it can cause a direct and significant increase in the sulphate aerosol load. Even smaller

more frequently occurring eruptions that do not reach stratospheric altitudes have an important contribution to the levels of SO_2 in the upper troposphere and lower stratosphere through mixing and transport (*Thornton et al.*, 1999).

The formation and maintenance of the background stratospheric layer has been described in detail by *Hamill et al.* (1997). The life cycle of a stratospheric aerosol particle contains three main stages: particle formation, development of a size distribution through transport and mixing processes, and removal from the stratosphere. Aircraft measurements of the particle size distributions suggest that the majority of stratospheric aerosol particles are formed very near the tropical tropopause in rising air caused by deep convective activity (*Goodman et al.*, 1982). Past work has theorized that the particles form on existing nuclei of meteoric origin (*Hunten et al.*, 1980) or from the troposphere (*Turco et al.*, 1979); more recent observations have shown that the particles form directly from from binary homogenous nucleation of H_2SO_4 at the tropical tropopause (*Brock et al.*, 1995). Once in the stratosphere, newly formed particles further condense and coagulate with larger existing aerosol particles. This multiplicative process produces a size distribution of particles that is approximately single mode log-normal. Several studies have shown this distribution appropriately fits observations of the background layer (*Bingen et al.*, 2004; *Bauman et al.*, 2003); however, volcanic aerosols are better characterized by multi-modal log-normal distributions (*Deshler et al.*, 2003).

Above the tropical tropopause, aerosols are rapidly transported zonally by the stratospheric winds but as meridional circulation is much slower, the particles tend to remain within tropical latitudes, a region referred to as the "tropical stratospheric reservoir" (*Trepte and Hitchman*, 1992). Poleward transport and descent into the "lowermost stratosphere" eventually occurs within the Brewer-Dobson circulation where quasi-isentropic transport of air carries the particles into the troposphere through tropopause folding events (*Holton et al.*, 1995). Sulphate aerosols are quickly lost in the troposphere where they influence cirrus cloud formation by acting as cloud

condensation nuclei.

Several determinations have been made of the aerosol particle size distribution through in situ measurements with balloon instruments (example *Deshler et al.*, 2003) and aircraft instruments (example *Brock et al.*, 1995), and remotely through inversion processes with satellite measurements (example *Bingen et al.*, 2004; *Bauman et al.*, 2003; *Thomason et al.*, 1997) and lidar measurements (example *Stein et al.*, 1994). The log-normal size distribution,

$$n(r) = \frac{1}{\sigma_g\sqrt{2\pi}}\frac{1}{r}\exp\left[-\frac{(\ln r - \ln r_g)^2}{2\sigma_g^2}\right] \tag{2.1}$$

is characterized by two parameters, r_g, and σ_g (*Hansen and Travis*, 1974). It is simply the distribution of a variable whose logarithm has a random distribution. That is if $\log(x)$ is normally distributed, the distribution of x is log-normal and is effectively skewed to larger values of x. The mode radius, r_g, is the geometric mean radius, or the mean value of the logarithm of the radii. The standard deviation of the logarithm of the radii, σ_g, is conventionally given in terms of mode width, s_g=e^{σ_g}. Typically, the background aerosol size distribution has a mode radius near 0.08 μm and a mode width of approximately 1.6; a second mode develops in volcanic aerosol with a much larger mode radius of approximately 0.4 μm and a narrower mode width near 1.1 (*Deshler et al.*, 2003).

2.1.4 Satellite Measurements

Since the advent of satellite-borne sensing of the atmosphere in the 1970's, the global extinction of solar radiation by aerosol particles in the stratosphere has been monitored almost continuously using a series of solar occultation instruments. Because of this long term data record, the extinction, generally measured at optical wavelengths, has become the standard parameter for developing a stratospheric aerosol climatology and for monitoring long term trends in the global distribution. The solar

occultation technique entails measuring the transmission of solar radiation through the atmosphere. From a satellite platform the sun is observed to rise through the atmosphere once per orbit and to set through the atmosphere once per orbit.

Measurements are made at successive tangent altitudes during each observation creating a height profile of the atmospheric transmission. Generally, a measurement of the direct solar radiation above the atmosphere is also captured during each occultation event to provide an inherent calibration for long term stability. Satellite based solar occultation is especially suited for observing the stratosphere as the optical depth is small and structures generally vary slowly along the long path length of the horizonal line of sight.

Because only two profiles are measured per orbit, at sunrise and at sunset, the latitudinal coverage of an occultation experiment is somewhat limited. For example, over the course of one day, all of the sunrise measurements occur at basically the same latitude. Depending on the inclination of the satellite orbit, the sampling latitude can remain fixed within a region, polar latitudes for example, or can be made to sweep across the hemisphere over the course of approximately one month.

The "gold-standard" set of occultation instruments is the SAM/SAGE series of NASA instruments. The first satellite based occultation instrument was the Stratospheric Aerosol Measurements (SAM) II (*McCormick et al.*, 1979). It was launched on the Nimbus 7 satellite into sun-synchronous orbit in 1978 and continued operation until 1993 measuring single wavelength (1000 nm) solar transmission at tangent altitudes from 10 to 30 km. The Stratospheric Aerosol and Gas Experiment (SAGE) was launched in 1979 and operated for only two years; however, the extinction was measured simultaneously in four photometer channels centered at 385, 450, 600 and 1000 nm (*McCormick et al.*, 1979). Similarly, SAGE II measured extinction at seven wavelengths over the same range and was operational for an astounding 21 years, from 1984 to 2005 (*Russell and McCormick*, 1989). SAGE III, launched in 2002, was a combination CCD/photometer and measured the aerosol extinction at nine

wavelengths from 385 to 1545 nm (*Thomason and Taha*, 2003). Very recently, the SAGE III mission was terminated.

Other satellite based occultation experiments include the Halogen Occultation Experiment (HALOE), which measured infrared aerosol extinction (2.45 to 5.26 μm) from 1991 to 2005 (*Russell et al.*, 1993), and the Polar Ozone and Aerosol Measurement (POAM) II and III instruments (*Glaccum et al.*, 1996; *Lucke et al.*, 1999), which were similar to SAGE II and operated from 1993-1996 and 1998-2005, respectively.

In combination, these satellite occultation instruments have collected an impressive and important long term data record for the study of stratospheric aerosol. Much has been learned about the dynamics of the stratosphere and the aerosol layer, in both the volcanic and background states. The effect of two dominant volcanoes, El Chichon (1982) and Mt. Pinatubo (1991) and the relaxation of the aerosol layer to the current background state has been monitored globally by these missions. A key recommendation of the WMO World Climate Research Program "Assessment of Stratospheric Aerosols" document (*Thomason*, 2006) states: "The importance of stratospheric aerosol in climate and atmospheric chemistry strongly supports a commitment to continuing both space-based and in-situ observations of aerosols into the foreseeable future." However, as of 2006, there are no operational satellite solar occultation experiments, and none are planned for the future. The continuation and extension of this long term global data record has become the responsibility of newer remote sensing techniques such as stellar occultation (*Vanhellemont et al.*, 2005) or limb scatter measurements (see Section 2.2). The present work focuses on the retrieval of stratospheric aerosol properties from measurements of limb scattered sunlight and the required radiative transfer. The following section details the limb scatter technique and specifically introduces the measurements made with the OSIRIS instrument on the Odin satellite that are used to demonstrate the utility of the aerosol inversion.

2.2 Limb Scatter Measurements

2.2.1 Technique

Measurements of sunlight scattered from the atmosphere have been used successfully for decades as a technique for remote sensing of the atmospheric state from satellite. The heritage of the limb scatter technique is from early nadir backscatter instruments, such as the Total Ozone Mapping Spectrometer (TOMS) and the Solar Backscatter Ultraviolet (SBUV) instrument that made spectral measurements of UV scattered sunlight in a downward looking (nadir) viewing geometry (*Heath et al.*, 1975). A global map of the total ozone vertical column can be produced from these measurements with relatively high horizontal resolution achieved by mechanically scanning the instrument field of view across the satellite track. It is possible to achieve some measure of the vertical distribution of ozone by combining measurements at wavelengths that are successively more strongly absorbed; however, the vertical resolution is very limited, especially at altitudes below the ozone peak near 30 km (*Bhartia et al.*, 1996).

The limb scatter technique used by OSIRIS is a relatively new technique for remote sensing the atmospheric state from the spectrum of scattered sunlight. The geometry of the measurement is such that the line of sight of the satellite instrument is not pointed at the Earth's surface as with nadir sounding, but is directed at an altitude within the atmosphere above the Earth's surface, i.e. at a tangent altitude in the atmospheric limb. A vertical profile of the limb radiance is obtained by scanning the line of sight of the instrument through successive tangent altitudes or by imaging along multiple lines of sight simultaneously. This observing geometry is similar to that of an occultation but is inherently a very different measurement in that the line of sight is not directed at the sun.

The benefits of the limb scatter technique are twofold: high vertical resolution similar to that of the occultation experiment, and good global coverage similar to

that of backscatter nadir sounding. There are, however, difficulties and limitations because of the long path length of the line of sight through the atmosphere. Basic inversion techniques require horizontal homogeneity along the line of sight. Tomographic techniques are indeed possible and advantageous but require fast sampling both vertically and along the satellite track to resolve horizonal structures (*Degenstein et al.*, 2003; *Degenstein et al.*, 2004). The other main difficulty with the limb scatter technique is that the presence of significant multiple scattering and reflection from the Earth's surface at optical wavelengths requires detailed radiative transfer for the retrieval of vertically resolved information on the atmospheric state.

Figure 2.1 conceptualizes the limb scatter measurement geometry. The tangent altitude of the line of sight is defined as the smallest distance between the Earth's surface and the line of sight. In an optically thin atmosphere, the sensitivity of the limb measurement to the atmospheric state is maximum at the altitude corresponding to the tangent altitude simply due to the viewing geometry. The path length of the intersection of the line of sight through infinitesimally thin atmospheric layers is maximum near the tangent point. For this reason, the tangent point is often used to define the geometry. The latitude and longitude of the measurement are taken as the point on the Earth's surface directly below the tangent point. The solar geometry is also usually defined at the tangent point. Three commonly defined angles are shown in Figure 2.1. The tangent point solar zenith angle, θ_T, is measured from the zenith direction at the tangent point to the solar direction. The tangent point solar azimuth angle, ϕ_T, is measured from the horizontal projection of the solar direction to the line of sight. Note that by definition of the tangent point the line of sight is contained within the horizontal plane. These two angles and the location on the Earth's surface directly below the tangent point are sufficient to define fully the geometry in a one dimensional sense where the atmosphere is assumed to be vertically stratified. If the fully three dimensional variability of the atmosphere is to be considered, another parameter that defines the line of sight in a geographic coordinate system is required.

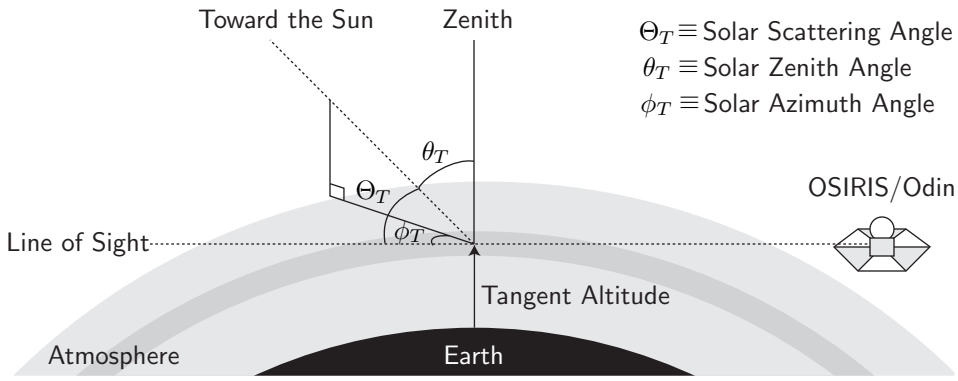

Figure 2.1: The limb scatter measurement geometry and the definition of the solar angles at the tangent point.

The solar scattering angle, Θ_T, measured directly from the solar direction to the line of sight, is a useful and common parameter for an intuitive description of the measurement geometry; however, it is not strictly required as it is simply derivable from θ_T and ϕ_T.

The Solar Mesospheric Explorer (SME) was one of the first satellites to measure profiles of limb radiance at optical wavelengths (*Barth et al.*, 1983). These were successfully used to infer mesospheric ozone (*Thomas et al.*, 1984) from the 1.27 μm emission of O_2 and to study polar mesospheric clouds (*Thomas*, 1984) from scattered sunlight.

Until quite recently, the limb scatter technique had not been used to probe stratospheric and tropospheric altitudes because of the requirement for spherical radiative transfer calculations that can accurately model the effects of multiple scattering and surface reflection, and also because of the strict instrument design requirements for high signal to noise over orders of magnitude, stringent stray light rejection and accurate pointing information. The first investigation of the possibility of using limb scatter observations at these altitudes was by the Shuttle Ozone Limb Sounding Experiment (SOLSE) combined with the Limb Ozone Retrieval Experiment (LORE) flown as a proof of concept on shuttle flight STS-87 in 1997. Using a combination of UV and visible wavelengths, a retrieval of the ozone profile from 15 to 40 km with

approximately 3 km vertical resolution was achieved (*McPeters et al.*, 2000).

In the following years to the current date, three satellite missions have performed extensive limb measurements. The Scanning Imaging Absorption Spectrometer for Atmospheric Chartography (SCIAMACHY) instrument launched on the ENVISAT satellite in March, 2002, and still currently operational, makes a large set of measurements from the ultraviolet (220 nm) to the near infrared (2380 nm) with varying spectral resolution and a vertical resolution of approximately 3 km (*Bovensmann et al.*, 1999). Limb scanning is only one operational mode of SCIAMACHY as it also operates in a nadir scanning mode. SAGE III, specially designed for for solar occultation and discussed above, also measured the limb scattered sunlight radiance between occultation events. *Rault* (2005) has presented a retrieval of ozone profiles from the SAGE III limb scatter measurements; however, as the system was designed for occultation and not limb scatter, poor stray light rejection and a lack of line of sight pointing information makes the inversion of these measurements a difficult task. OSIRIS completes the list of limb scatter satellite instruments and is the only satellite instrument of the three that is dedicated to limb scatter measurements. Details of the OSIRIS mission and instrument are covered in the following section.

A number of techniques have been developed for the retrieval of trace gases from limb scattered spectra including spectral analysis techniques (*Flittner et al.*, 2000; *von Savigny et al.*, 2003; *Rault*, 2005) and the application of Differential Optical Absorbtion Spectroscopy (*Haley et al.*, 2004; *Rault*, 2005). The retrieval of stratospheric aerosol properties from satellite limb scatter measurements forms a new field of work. It is of interest to use the limb scatter technique for the retrieval of stratospheric aerosol parameters because of the relatively high sampling resolution, both vertically and horizontally, and the need to continue monitoring of aerosol in the absence of an operational occultation experiment. The reliable retrieval of aerosol from limb scatter is also necessary because the retrievals of trace gases such as O_3 and NO_2 are sensitive to any error in the assumed aerosol profile (*von Savigny et al.*, 2005).

Loughman et al. (2005) show that even for a background aerosol loading, uncertainty in the aerosol profile is the second most significant source of error in the limb scatter retrieval of ozone.

2.2.2 OSIRIS on Odin

The Canadian Optical Spectrograph and InfraRed Imaging System (OSIRIS) is one of a new generation of satellite instruments designed to measure the atmospheric limb radiance of scattered sunlight (*Llewellyn et al.*, 2004). The instrument is onboard the Swedish satellite Odin (*Murtagh et al.*, 2002), which was launched on February 20, 2001, and continues full operation to date. The Odin orbit is circular and sun-synchronous at an altitude near 600 km with a period of 96 minutes. The orbital inclination of 98° from the equator provides near-global coverage as the corresponding sampled latitude range for nominal on-track instrument pointing is from 82° S to 82° N . The local time of the ascending node, *i.e.* the northward equatorial crossing, is 1800h. The satellite remains very close to local dusk on the entire ascending track and close to local dawn on the descending track, sweeping quickly through noon at high northern latitudes and through midnight at high southern latitudes. A plot of the satellite latitude as a function of local time is shown in Figure 2.2.

A consequence of the Odin dusk/dawn orbit is that the winter hemisphere is not illuminated by the sun at the sampled local times. Figure 2.3 is a plot of contours of the solar zenith angle at the OSIRIS tangent point as a function of latitude during 2002. In the gray shaded regions the sun is below the horizon at the tangent point ($\theta_T > 90°$) and trace gas and aerosol inversion are not possible. During these orbital segments, measurements of the nighttime airglow are made. For two time periods in each year, in approximately February and October, the Odin orbit track is closely aligned with the solar terminator and so provides a solar zenith angle at the tangent point of the line of sight that is very close to 90° at all latitudes.

A second consequence of the dusk/dawn orbit is that the variation in solar zenith

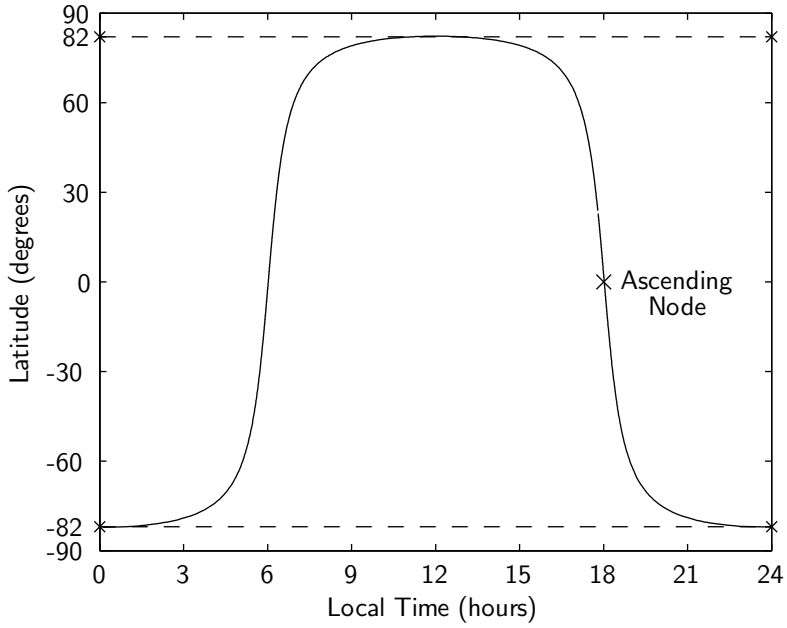

Figure 2.2: The Odin/OSIRIS latitude as a function of the local time. The sun-synchronous orbit has an ascending node local time of 1800h and descending node local time of 0600h.

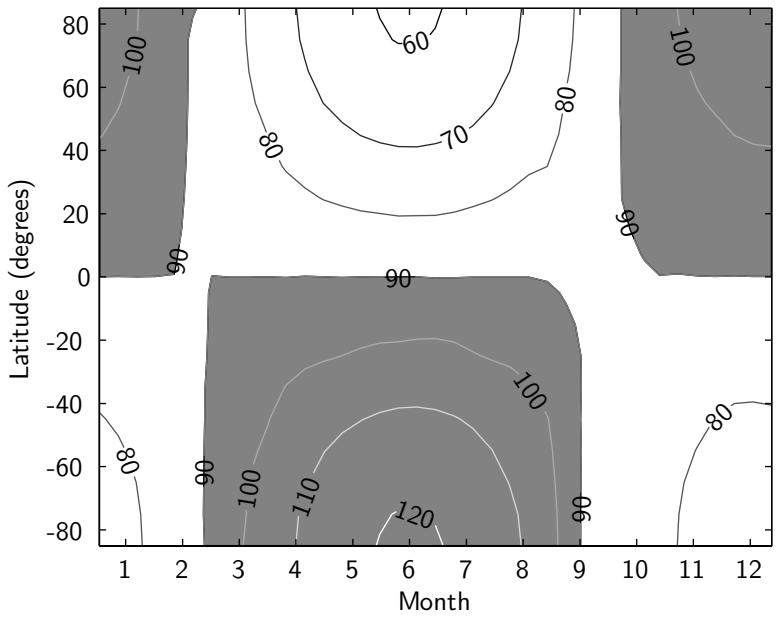

Figure 2.3: Contours of the OSIRIS tangent point solar zenith angle (degrees) as function of latitude during 2002. The shaded regions ($\theta_T > 90°$) are not illuminated by the sun.

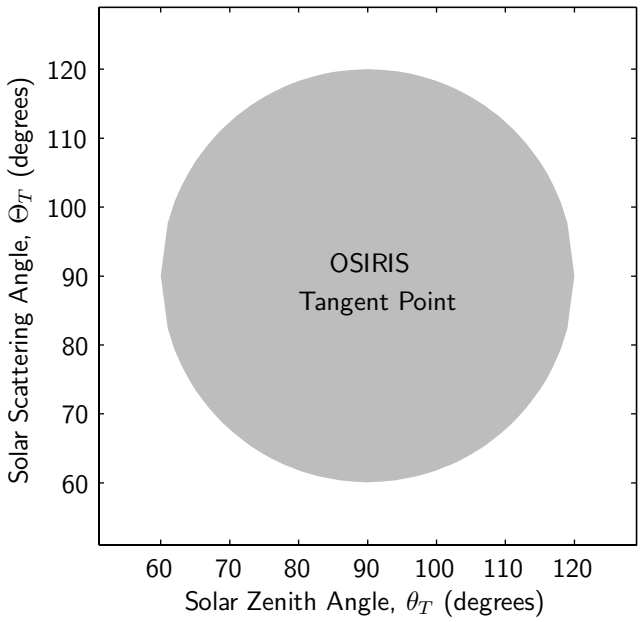

Figure 2.4: The variation in the OSIRIS solar scattering angle as function of the solar zenith angle at the tangent point. All OSIRIS measurements fall within the gray area on the plot.

angle is somewhat limited. A dusk or dawn local time implies that the sun is very close to the horizon. It can be seen from Figure 2.3 that the range of solar zenith angles over a year is between approximately 60° and 120°, with most of the variation occurring at the poles where the local time quickly sweeps from dusk to dawn or vice versa. The near polar inclination of the orbit also limits the variation in solar scattering angle. It also ranges from approximately 60° to 120° and is most often near 90°. The relationship between the solar scattering angle and the solar zenith angle is shown in Figure 2.4. All OSIRIS measurements fall within the gray area on the plot. It can be seen from the figure that when the solar zenith angle is near 90°, the full range of solar scattering angles are possible; however, as the sun rises higher in the sky, the solar scattering angle becomes closer to 90°.

The OSIRIS instrument is essentially two optical subsystems, suggested by its name: an optical spectrograph (OS) module and an infrared imager (IRI) module. The OS is grating spectrometer with a CCD detector, and measures spectra of the

limb radiance from 280-810 nm with a spectral resolution of approximately 1 nm. The field of view of the spectrograph when mapped on to the atmospheric limb at the tangent point is approximately 1 km vertically and 40 km horizontally. Vertical profiles of the limb radiance are obtained by taking OS exposures while performing a repetitive vertical scan of the single line of sight through selected tangent altitude ranges, nominally from 10 to 100 km. Successive exposures are spaced by approximately 2 km in tangent altitude. The time required for a single altitude scan is near 1.5 minutes and so allows for nearly 60 scans per orbit. To satisfy the accurate three axis astronomical pointing requirements of the other instrument on Odin, the Sub-Millimetre Radiometer, the attitude control system of the satellite equipped with two star trackers, a sun tracker, magnetometers and gyros, and is activated by magnetic torquers and momentum wheels. The reconstructed knowledge of the limb pointing is approximately ± 15 arcseconds, or ± 200 m in the vertical at the tangent point.

Figure 2.5 is a plot of typical OSIRIS limb radiance spectra at selected tangent altitudes during limb scan 06432012. A transition in optical material is used to filter higher order light diffracted from the grating. This transition region, referred to as the order sorter, contaminates the measurements at wavelengths from 475 to 535 nm. Figure 2.6 is a log-scale plot of the limb radiance profile at four wavelengths across the OS spectrum during the same limb scan. In each panel, the short wavelength is within an ozone absorption band, Hartley-Huggins in the left, Chappuis in the right, and causes a "knee" in the profile at the tangent altitude where the atmosphere becomes optically thick. The knee in the 350 nm radiance profile, near 25 km, is simply from neutral density Rayleigh scattering.

The IRI is composed of three vertical near infrared co-aligned linear array imagers that capture one dimensional images of the limb radiance at 1260, 1270, and 1530 nm with a tangent altitude resolution of approximately 1 km. The imagers have parallel bore-sights; each one consists of an identical baffling system, lens, narrow-band interference filter, and a one-dimensional linear array of 128 thermo-electrically cooled

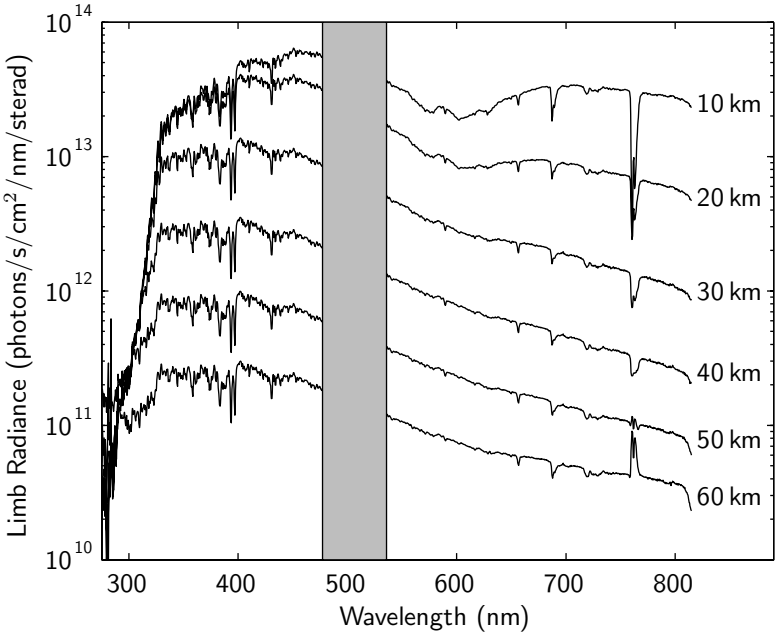

Figure 2.5: Sample OSIRIS limb radiance spectra at selected tangent altitudes during a typical limb scan, 06432012. The spectral order sorter contaminates the measurements at wavelengths between 475 and 535 nm.

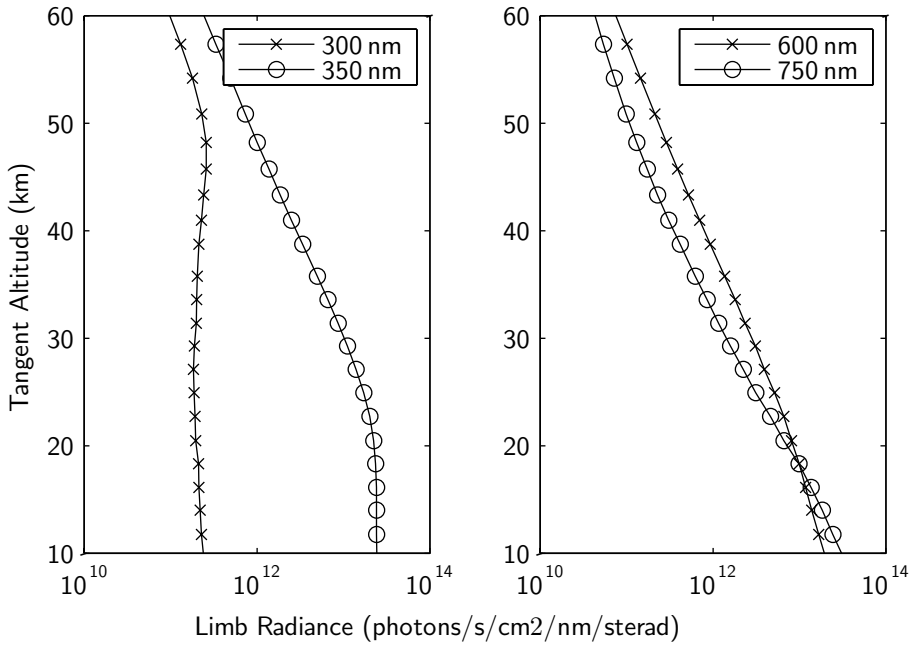

Figure 2.6: Sample OSIRIS limb radiance profiles at selected wavelengths during a typical limb scan, 06432012.

InGaAs photodetectors placed in the focal plane of the lens. Approximately 30 photodetector elements at one end of each array are covered with a mask in order to provide a continuous measure of the dark signal in the array. Therefore there are approximately 100 lines of sight from each imager channel that measure simultaneously over 100 vertical kilometers in tangent altitude. The tomographic inversion of the 1260 and 1270 nm Oxygen InfraRed Atmospheric (OIRA) band volume emission rate is the heritage of the algorithm used here for the one dimensional species inversion. The 1530 nm channel is designed to measure a Meinel band hydroxyl vibrational-rotational emission in the mesosphere. As this emission is extremely weak during the daytime, this channel also provides a measure of the limb scattered sunlight from the neutral atmosphere.

2.2.3 Upcoming Missions: OMPS

The potential of limb scatter measurements from satellites is vast. The National Polar-Orbiting Operational Satellite System (NPOESS), which is the next generation of polar orbiting operational environmental satellites operated by the National Oceanic and Atmospheric Administration (NOAA) of the United States, has selected the Ozone Mapping and Profiling Suite (OMPS) instrument (*Dittman et al.*, 2002) for the monitoring of ozone over the next 20 years in fulfilment of the Montreal Protocol. OMPS will monitor both the total ozone column through nadir backscatter measurements and the ozone vertical profile through limb scatter measurements in the optical region from the tropopause to 60 km altitude. The specified accuracy of the ozone vertical profile is 10% with a precision of 3% and a vertical resolution of 2 km.

It is clear that limb scatter measurements will form an important part of the future of global atmospheric measurements. Without an operational occultation experiment in orbit, it is essential to be able to monitor the stratospheric aerosol load with limb scatter in order to maintain the long term data record and to understand

better the effect of aerosol on climate and atmospheric chemistry. As well, a proper characterization of the effect of aerosol scattering on limb spectra is required in order to accurately monitor ozone, especially in the long term, from limb scatter. It is the fundamental objective of this thesis work to demonstrate a technique, with the required radiative transfer modelling, for the retrieval of stratospheric aerosol properties from limb scattered sunlight measurements.

Chapter 3

Inverse Theory

3.1 The Inverse Problem

Scientific theory attempts to explain our observations of nature. In physics, the theory is formed by a set of rules, composed in mathematics, that relate a natural state or phenomenon to a set of observations. The rules are most often developed in a forward sense where they are used to predict the measurement of a natural state. Practically, an approximation of the complete theory may be all that is required to predict the measurement to the desired accuracy. The theory is useful in an inverse sense where a measurement is used to work backwards to deduce the natural state. This is the formulation of the inverse problem that occurs in many experimental branches of science, especially in remote sensing in the Earth sciences.

There are two important difficulties that arise in most remote sensing applications. The first is that the measurements are a non-linear function of the natural state making the inversion, or backward application of the theory to infer the natural state from the measurements, not a simple matter. The second difficulty is that the measurements are inherently discrete, whether in time or space, and finite in number, while the desired natural state is continuous. Therefore, the formulation of the inverse problem is to determine the best estimate of the natural state from the

available measurements and to gauge the quality of the estimate.

For this work, and for many applications of inverse theory to atmospheric remote sensing, the desired natural state is the height profile of an atmospheric constituent species, $x(z)$, where x represents a species parameter, for example, the concentration of a gas or the composition of an aerosol, and z is the vertical coordinate. Unless the functional form of this continuous state is known a finite number of measurements cannot be used to determine the continuous profile. At best, it is possible to recover only samples of the natural state. Mathematically, these samples are represented as a vector, appropriately called the *state vector*,

$$\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}. \tag{3.1}$$

The n elements of $\mathbf{x}$ each correspond to the value of the species parameter at a specific altitude level.

The measurements are inherently discrete and form the *measurement vector*,

$$\mathbf{y} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{pmatrix}. \tag{3.2}$$

There are m elements of $\mathbf{y}$, each related in some way to an altitude. It should be noted that m is not necessarily equal to n.

The mathematical formulation of the measurement is

$$\mathbf{y} = f(\mathbf{x}, b) + \boldsymbol{\epsilon}\,, \tag{3.3}$$

where f describes the physical process that relates the state of the atmosphere to the measurement. This is a function of the desired state parameter, $\mathbf{x}$, and is also a

function of other atmospheric parameters denoted simply by b. Real measurements always have noise, $\boldsymbol{\epsilon}$, associated with them so that the observation is the sum of the true response of the atmosphere plus the measurement error.

The essence of the inverse problem is to use the known physics to work backwards from the measurements to determine the desired atmospheric species profile. Implementation of the physics together with the best estimate of the atmospheric parameters is used to form the *forward model*,

$$\mathbf{F}(\mathbf{x}, \tilde{b}) \cong f(\mathbf{x}, b)\,. \tag{3.4}$$

Except in trivial cases, the model only approximates the true physical process. This may be due to several reasons. It can be that the description of the theory is incomplete, or that the model may be designed in an approximate sense because an implementation of the complete theory is resource prohibitive. In addition, a perfect knowledge of the atmospheric state, b, is impossible. The best guess of the atmospheric state is denoted by $\tilde{b}$. Thus the forward model predicts the measurements for a given atmospheric state,

$$\mathbf{y} = \mathbf{F}(\mathbf{x}, \tilde{b}) + \boldsymbol{\epsilon}\ . \tag{3.5}$$

The objective of the inversion is to find the state vector that provides the best match between the model prediction and the measurements. A common statistical way to quantify the best estimate is through the method of least squares where the solution is the state vector that minimizes,

$$R^2 = \sum_{i=1}^{m} \Big(\mathrm{y}_i - \mathrm{F}_i(\mathbf{x}, \tilde{b})\Big)^2 = \Big|\mathbf{y} - \mathbf{F}(\mathbf{x}, \tilde{b})\Big|^2\ . \tag{3.6}$$

An easy interpretation of the inverse problem is demonstrated with the simple case where the model is linear and the measurements are noise-free. The measurements and the state vector are then simply related. For a typical state vector, $\mathbf{x}_a$, known

as the *a priori* state,

$$\mathbf{y} = \frac{\partial \mathbf{F}(\mathbf{x}, \tilde{b})}{\partial \mathbf{x}}(\mathbf{x} - \mathbf{x}_a) + \mathbf{F}(\mathbf{x}_a, \tilde{b})\,. \tag{3.7}$$

The partial derivative of the model with respect to the state vector is the Jacobian matrix of $\mathbf{F}(\mathbf{x}, \tilde{b})$ and in the context of an inversion is often referred to as the kernel matrix, $\mathbf{K}$, or the weighting function matrix. In general it is an $m \times n$ matrix, where each column vector represents the derivative of $\mathbf{F}(\mathbf{x}, \tilde{b})$ with respect to a single element of the state vector,

$$\mathbf{K} = \begin{pmatrix} \frac{\partial \mathrm{F}_1}{\partial \mathrm{x}_1} & \cdots & \frac{\partial \mathrm{F}_1}{\partial \mathrm{x}_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial \mathrm{F}_m}{\partial \mathrm{x}_1} & \cdots & \frac{\partial \mathrm{F}_m}{\partial \mathrm{x}_n} \end{pmatrix}. \tag{3.8}$$

If there is the same number of measurements as elements in the state vector profile then $\mathbf{K}$ is a square matrix and Equation 3.7 can be solved directly to yield the desired, or retrieved state, $\hat{\mathbf{x}}$,

$$\hat{\mathbf{x}} = \mathbf{x}_a + \mathbf{K}^{-1}\left(\mathbf{y} - \mathbf{F}(\mathbf{x}_a, \tilde{b})\right)\,. \tag{3.9}$$

3.2 Application to Aerosol Retrieval from OSIRIS

This thesis work addresses the retrieval of stratospheric aerosol profiles from OSIRIS measurements of limb scattered sunlight spectra. The state vector, $\mathbf{x}$, is composed of a discretized aerosol number density profile. Elements of the measurement vector, $\mathbf{y}$, are constructed from OSIRIS spectra, each corresponding to a different measured tangent altitude. The physics that relates the aerosol profile to the OSIRIS spectra, $\mathbf{F}(\mathbf{x}, \tilde{b})$, is a model of the radiative transfer that predicts the scattering of sunlight from the atmosphere into the OSIRIS instrument line of sight. Other atmospheric parameters required by the radiative transfer model, $\tilde{b}$, include the reflectivity, or albedo, of the Earth, the neutral density profile and the ozone profile.

The significant contribution of multiple scattering of sunlight in the atmosphere

and from the Earth's surface results in a non-linear relation between the aerosol profile and the measured spectra. The OSIRIS spectra, like all real measurements contain error, both random and systematic. The remaining sections of this chapter contain a brief survey of inverse methods commonly used in non-linear inverse problems and a specification for a formal error analysis in the retrieval. An adaptation of the MART inversion, described in Section 3.6.3, is used here for the retrieval of the aerosol number density from the OSIRIS measurements.

3.3 Direct Solution

If the problem is non-linear the solution can be found by trial and error where the forward model is evaluated for a very large set of values of the atmospheric state. The solution is chosen from the set as the vector that minimizes the difference between the measurements and model prediction. This obviously involve a large number of evaluations of the forward model, especially if there is a significant number of points in the measurement vector and state vector. If the model calculation is resource intensive this type of solution is prohibitive.

Using Gauss-Newton iteration, the number of evaluations of the forward model can be minimized by assuming the model is linear at each step. Keeping only the linear terms in the Taylor expansion of the forward model about an initial guess, or *a priori*, state yields,

$$\mathbf{F}(\mathbf{x}, \tilde{b}) = \mathbf{F}(\mathbf{x}_a, \tilde{b}) + \mathbf{K}(\mathbf{x} - \mathbf{x}_a)\,. \tag{3.10}$$

A first estimate of the retrieved state, $\hat{\mathbf{x}}^{(1)}$ is made by solving this linear equation. The kernel is then recalculated about the retrieved state and the resulting linear equation is solved again to yield a second, and better, estimate of the retrieved state. The $(n+1)^{th}$ iteration is,

$$\hat{\mathbf{x}}^{(n+1)} = \hat{\mathbf{x}}^{(n)} + (\mathbf{K}^{(n)})^{-1}\left(\mathbf{y} - \mathbf{F}(\mathbf{x}^{(n)}, \tilde{b})\right)\,, \tag{3.11}$$

where $\mathbf{K}^{(n)}$ is the kernel matrix calculated with respect to the n^{th} iteration of the state vector,

$$\mathbf{K}^{(n)} = \frac{\partial \mathbf{F}(\mathbf{x}^{(n)}, \tilde{b})}{\partial \mathbf{x}^{(n)}} \,. \tag{3.12}$$

In this formulation, the *a priori* state, $\mathbf{x}_a$, is identical to $\mathbf{x}^{(0)}$. It should be noted that this method cannot be applied in all cases because it requires the inverse kernel. In addition, it is often computationally expensive to determine the kernel matrix, and this method requires both re-evaluation of the kernel and calculation of its inverse at every iteration.

3.4 Regularization

In the least squares solution, noise in the measurements are reflected in the retrieved state vector. An additional constraint is added in the least squares formulation by Tikhonov regularization (*Tikhonov*, 1963) where the solution is the minimum of,

$$R^2 = \left| \mathbf{y} - \mathbf{F}(\mathbf{x}, \tilde{b}) \right|^2 + \gamma^2 \left| \mathbf{x} \right|^2 \,. \tag{3.13}$$

This additional term sets up a balance between the fit of the forward model to the measurements and the variation in the retrieved state. When there is random noise in the measurements, the minimum value of $|\mathbf{x}|^2$ is the smoothest retrieved profile. The ad-hoc factor, γ^2, is called the regularization parameter. It determines the importance of the smoothness in the retrieval and an optimal value is often difficult to determine. A consequence of the regularization is a loss of small scale structure in the retrieval as it is not distinguished from noise. There are several modifications of regularization that allow the smoothness to be defined through an operator matrix and by the difference between the retrieved state and the *a priori* state.

3.5 Optimal Estimation

The optimal estimation method, attributed to *Rodgers* (1976, 2000), is widely used in atmospheric science. It is another regularization method and is based on a statistical knowledge of the desired state parameter. It is developed in the Bayesian view that the probability density function of a variable, characterized by its mean and variance, describes knowledge about the variable rather than a distribution of various measurement trials.

Given a set of measurements, $\mathbf{y}$, with the associated covariance, $\mathbf{S}_\varepsilon$, and a linear forward model, $\mathbf{K}$, Bayesian statistics are used to update the *a priori* estimate of the probability density function of the state parameter to the *a posteriori* estimate by including information from the measurements,

$$P(\mathbf{x}|\mathbf{y}) = \frac{P(\mathbf{x})\,P(\mathbf{y}|\mathbf{x})}{P(\mathbf{y})}\,. \tag{3.14}$$

An important assumption in the optimal estimation formulation is that the measurement error is random and that probability density function of the *a priori* state is a Gaussian distribution with covariance, $\mathbf{S_a}$. The solution, $\mathbf{x} = \hat{\mathbf{x}}$, is found by maximizing the value of $P(\mathbf{x}|\mathbf{y})$ and results in a weighted mean of the *a priori* state, $\mathbf{x}_a$, and the forward model for the measurements,

$$\hat{\mathbf{x}} = \left(\mathbf{K}^T{\mathbf{S}_\varepsilon}^{-1}\mathbf{K} + \mathbf{S_a^{-1}}\right)^{-1}\left(\mathbf{K}^T{\mathbf{S}_\varepsilon}^{-1}\mathbf{y} + \mathbf{S_a^{-1}}\mathbf{x}_a\right)\,. \tag{3.15}$$

The covariance matrix of the retrieval is determined directly from the kernel matrix, the covariances of the measurements, and the *a priori*,

$$\mathbf{S_x} = \left(\mathbf{K}^T{\mathbf{S}_\varepsilon}^{-1}\mathbf{K} + \mathbf{S_a^{-1}}\right)^{-1}\,. \tag{3.16}$$

The probability density function of the *a priori* state is often unknown so that a climatological estimate is frequently used. The retrieved state is certainly not ran-

domly distributed about the climatology or there would be no value in making the measurement. However, the covariance of the retrieved state can be interpreted as a gauge of the amount of information added to the probability density function of the retrieved state by the measurements. It should be noted that in Equation 3.16, $\mathbf{S_x}$ can never be greater than the covariance of the *a priori*, $\mathbf{S_a}$. If the effect of the measurements is small compared to the measurement error, the retrieved state is very close to the *a priori* state and $\mathbf{S_x}$ is approximately equal to $\mathbf{S_a}$. However, if the measurement error is small compared to the effect of the measurement, the solution is constrained by the measurement and $\mathbf{S_x}$ is decreased.

If the model is non-linear, a Gauss-Newton iteration is performed using the linearized model obtained by Taylor expansion (Equation 3.10). Therefore the iterative solution is,

$$\hat{\mathbf{x}}^{(n+1)} = \hat{\mathbf{x}}^{(n)} + \left(\mathbf{S_x}^{(n)}\right)^{-1} \left(\mathbf{K}^{(n)T}\mathbf{S_\varepsilon}^{-1}\left(\mathbf{y} - \mathbf{F}(\mathbf{x}^{(n)}, \tilde{b})\right) + \mathbf{S_a^{-1}}(\mathbf{x}_a - \mathbf{x}^{(n)})\right) . \quad (3.17)$$

As with the direct solution method, this iterative method requires re-calculation of the kernel at every iteration. A feature of the optimal estimation solution is that it is generally applicable and does not require inversion of the kernel matrix. The inclusion of the *a priori* information is said to guard the solution from amplification of the measurement error. However, the effect of the *a priori* on the solution is determined through its covariance, which for a climatological estimate is not Gaussian and rarely well known.

3.6 Relaxation Methods

3.6.1 Chahine Relaxation

An ad-hoc relaxation method is often used in atmospheric inversion. It was developed by *Chahine* (1970, 1972) and takes the simple iterative form,

$$\hat{x}_i^{(n+1)} = \hat{x}_i^{(n)} \frac{y_i}{F_i(\mathbf{x}^{(n)}, \tilde{b})}\,, \tag{3.18}$$

where the update of the estimate of the state parameter is simply the ratio of the measurement to the forward model prediction evaluated at the current best guess of the state. In this formulation, the measurement vector must be constructed so that if the measurement vector element is greater than the evaluation of the current state of the forward model, increasing the value of the state parameter causes the model result to move closer to the observations. The iteration process is repeated for each element until the residuals, R_i, converge. This occurs when the evaluation of the forward model at the current state matches the measurements,

$$R_i = y_i - F_i(\mathbf{x}^{(n)}, \tilde{b}) \rightarrow \epsilon, \tag{3.19}$$

where ϵ has a magnitude that reflects the measurement error.

In the limb geometry, the sensitivity of observation y_j to a change in state x_i is strongly peaked in altitude, as shown in Figures 5.1 and 5.4. The maximum sensitivity occurs where the measurement is tangent through the element that is changed. The Chahine non-linear relaxation technique relies on this condition where the state vector elements (points on the altitude profile of the species of the interest) correspond in a one-to-one fashion with those of the observation vector (the measured tangent altitudes). This requires the length of the measurement vector and the state parameter vector to be equal and is often done by some type of simple interpolation of the smooth atmospheric state profile. For limb geometry, the altitude of each of the

state vector elements must be the tangent altitude of the corresponding measurement vector element, which requires $j = i$ as in Equation 3.18.

This method is advantageous in its simplicity and its remarkable stability. It can be extremely fast computationally as it does not require evaluation of the kernel matrix nor its inverse. The iterations can be stopped after an appropriate number so that the retrieval does not fit small structures to noise in the measurements. An often cited disadvantage is the difficulty in gauging the error in the retrieval. In addition, this method does not allow the incorporation of information from multiple measurements at the same altitude within the retrieval, nor account for how a change in one element of the state parameter affects the other elements of the state parameter except through the evaluation of the forward model as the iterations proceed.

3.6.2 Twomey-Chahine Extension

A method that allows for the retrieval of the state vector at points between measurement locations was developed by *Twomey* (1975) and *Twomey et al.* (1977) as an extension to the Chahine relaxation method. The kernel matrix, $\mathbf{K}$, is used in order to include the sensitivity in all elements of the state parameter to a change in the value of a certain element. The Chahine iteration, Equation 3.18 where i is no longer necessarily equal to j, is modified to

$$\hat{x}_i^{(n+1)} = \hat{x}_i^{(n)} \frac{y_j}{F_j(\mathbf{x}^{(n)}, \tilde{b})} K_{ij} + \hat{x}_i^{(n)} (1 - K_{ij}) \,, \tag{3.20}$$

where the kernel matrix is normalized such that the maximum value of $K_{ij} = 1$.

The iteration is applied to all measurements, y_j, for each element of the state parameter, $\hat{x}_i$, and repeated until the convergence criteria, Equation 3.19, are met. The way that the kernel is used in the iteration is somewhat arbitrary; however, it has the effect of updating the state parameter according to the ratio of a measurement at a given altitude to the model with a weight that is based on how much a change at

that altitude affects the parameter. If there is no change in the state parameter due to a change at the measurement altitude, the kernel function $K_{ij} = 0$ and the state parameter is unchanged, $\hat{x}_i^{(n+1)} = \hat{x}_i^{(n)}$. As this method requires re-evaluation of the kernel matrix at each iteration, the calculation efficiency of the Chahine technique is lost.

It should be noted that if the kernel functions are such that a change in state parameter only has an effect at the altitude of the state parameter, then $K_{ij} = \delta_{ij}$ and the Twomey iteration simplifies to the Chahine relaxation.

3.6.3 Multiplicative Algebraic Reconstruction

A two-dimensional tomographic retrieval technique implemented by *Degenstein et al.* (2003) has been used to retrieve mesospheric volume emission rates of the Oxygen InfraRed Atmospheric band measured by the InfraRed Imaging (IRI) subsection of the OSIRIS instrument. The Multiplicative Algebraic Reconstruction Technique (MART) was first developed as a statistical method to deblur Fabry-Perot images (*Lloyd and Llewellyn*, 1989) and has many similarities to the Maximum Likelihood Expectation Maximization technique used in Positron Emission Tomography (*Shepp and Vardi*, 1982) as well as the Chahine (Section 3.6.1) and Twomey-Chahine (Section 3.6.2) relaxation inversion techniques. In the present work, an adaptation of the tomographic technique is used to solve the one dimensional inverse problem for the aerosol profile from measurements of limb scattered sunlight.

The Chahine relaxation technique assumes that one measurement is made as a function of altitude. This single measurement can in fact be a combination of several measured radiances, for example the depth of an absorption band. However, if two or more independent measurements are made as a function of altitude, each depending in some way on the value of the state parameter, the Chahine method does not allow the measurements to be combined except through a retrieval of the state parameter for each measurement set and then taking some type of average value of the solutions.

In the MART formulation, each independent measurement set, denoted by k for both the measurements, $\mathbf{y_k}$, and the forward model, $\mathbf{F_k}(\mathbf{x}, \tilde{b})$, is included at each iterative update of the state parameter estimate. In addition an individual measurement is allowed to contribute to the estimate for a combination of altitudes. The iteration is,

$$\hat{x}_i^{(n+1)} = \hat{x}_i^{(n)} \sum_k \sum_j \frac{y_{jk}}{F_{jk}(\mathbf{x}^{(n)}, \tilde{b})} W_{ijk} \,. \tag{3.21}$$

The terms W_{ijk} form a weighting filter function that determines the importance of the j^{th} element of the k^{th} measurement set to the value of the i^{th} element of the state parameter. The weighting filter function can be broken down into two sets of weights: for each i^{th} element of the state parameter, a set of weights for elements of the measurement vector, W_{ij}, and a second set of weights that spans the measurements sets, W_{ik}. In an atmospheric sense, W_{ij} defines the contribution of measurements at locations, j, for the retrieval at a given altitude, i. W_{ik} specifies the contribution of each set of independent measurements, k, for example the radiance at different wavelengths, to the retrieval at a given altitude. Both sets of weights are normalized so that the sum of all elements is unity.

$$\sum_j W_{ij} = \sum_k W_{ik} = 1 \tag{3.22}$$

This implies that the normalization of the entire weighting filter function for any state parameter element, i, is also unity.

$$\sum_k \sum_j W_{ijk} = 1 \tag{3.23}$$

In a similar fashion to the Twomey extension retrieval, the simplest choice of weighting filter used with a single measurement set, $W_{ij} = \delta_{ij}$, reduces this method to the Chahine relaxation.

The MART inversion is the technique of choice for analysis of OSIRIS spectra at

the University of Saskatchewan. It has been successfully implemented for the two-dimensional recovery of the OIRA band, and one dimensional constituent species retrieval for ozone (*Roth et al.*, 2007) and NO_2 (Degenstein, private communication). It is used in this work for the retrieval of the stratospheric aerosol profile. The overall benefits of the MART inversion are threefold:

- inclusion of independent measurement sets at each iterative update of the state parameter estimate,
- inclusion of multiple elements of a measurement set for the estimate of a single element of the state parameter, and
- computational efficiency as neither the kernel nor its inverse are required.

3.7 Error Analysis

3.7.1 Theory

The error analysis of atmospheric retrievals has been formalized by *Rodgers* (1990) (see also *Rodgers*, 2000). The true error associated with a solution obtained by iterative inversion of a non-linear problem is certainly difficult to quantify satisfactorily. However, the method presented by Rodgers, while most applicable to the optimal estimation solution, can be used generally and so provides a way to estimate the contribution from various sources of error, the resolution of the retrieval process, and the effect of the *a priori* that is inherent in the solution. It relies on two linearizations, the first of the forward model about the best estimate of the atmospheric state, and the second of the inversion process itself about the measurements.

Briefly, following *Rodgers* (2000), the inversion process can be viewed as a function, R, that results in the retrieved solution for a given set of measurements, $\mathbf{y}$, all of the atmospheric inputs required by the forward model prediction, $\tilde{b}$, and the *a priori*

knowledge of the desired state, $\mathbf{x}_a$,

$$\hat{\mathbf{x}} = R\left(\mathbf{y}, \tilde{b}, \mathbf{x}_a\right). \tag{3.24}$$

The forward model, $\mathbf{F}(\mathbf{x}, \tilde{b})$ is used in the inversion to predict the measurements as a function of the desired state parameter,

$$\hat{\mathbf{x}} = R\left(\mathbf{F}(\mathbf{x}, \tilde{b}) + \Delta\mathbf{f}(\mathbf{x}, b) + \boldsymbol{\epsilon}, \tilde{b}, \mathbf{x}_a\right), \tag{3.25}$$

where ϵ represents the noise in the measurements and $\Delta\mathbf{f}(\mathbf{x}, b)$ accounts for the approximate nature of the forward model (Equation 3.4).

The model is assumed to operate linearly for values of the state parameter near the *a priori*, $\mathbf{x}_a$ and for all other atmospheric inputs near the assumed state, $\tilde{b}$. Keeping the linear terms in the Taylor series expansion of the forward model with respect to $\mathbf{x}$ and b yields,

$$\hat{\mathbf{x}} = R\left(\mathbf{F}(\mathbf{x}_a, \tilde{b}) + \frac{\partial\mathbf{F}(\mathbf{x}_a, \tilde{b})}{\partial\mathbf{x}}(\mathbf{x} - \mathbf{x}_a) + \frac{\partial\mathbf{F}(\mathbf{x}, \tilde{b})}{\partial b}(b - \tilde{b}) + \Delta\mathbf{f}(\mathbf{x}, b) + \boldsymbol{\epsilon}, \tilde{b}, \mathbf{x}_a\right). \tag{3.26}$$

The kernel matrix, or Jacobian, which quantifies changes in the forward model with respect to the assumed atmospheric inputs, is denoted by,

$$\mathbf{K}_b = \frac{\partial\mathbf{F}(\mathbf{x}, \tilde{b})}{\partial b}. \tag{3.27}$$

Using Equation 3.27 and the definition of the kernel matrix with respect to the state parameter (see Equation 3.8), Equation 3.26 becomes,

$$\hat{\mathbf{x}} = R\left(\mathbf{F}(\mathbf{x}_a, \tilde{b}) + \mathbf{K}(\mathbf{x} - \mathbf{x}_a) + \mathbf{K}_b(b - \tilde{b}) + \Delta\mathbf{f}(\mathbf{x}, b) + \boldsymbol{\epsilon}, \tilde{b}, \mathbf{x}_a\right). \tag{3.28}$$

The second linearization is of the retrieval process with respect to the measure-

ments, $\mathbf{y}$,

$$\hat{\mathbf{x}} = R\left((\mathbf{F}(\mathbf{x}_a, \tilde{b}), \tilde{b}, \mathbf{x}_a)\right) + \frac{\partial R}{\partial \mathbf{y}}\left\{\mathbf{K}(\mathbf{x} - \mathbf{x}_a) + \mathbf{K}_b(b - \tilde{b}) + \Delta\mathbf{f}(\mathbf{x}, b) + \boldsymbol{\epsilon}\right\} . \quad (3.29)$$

From this result, two definitions are formed. The sensitivity of the retrieval process with respect to the measurements, $\mathbf{y}$, is called the *contribution function* matrix,

$$\mathbf{D} = \frac{\partial R}{\partial \mathbf{y}} = \frac{\partial \hat{\mathbf{x}}}{\partial \mathbf{y}} . \quad (3.30)$$

The *averaging kernel* matrix is the sensitivity of the retrieval to the true state and is the product of the contribution function and the forward model kernel,

$$\mathbf{A} = \frac{\partial R}{\partial \mathbf{y}}\mathbf{K} = \mathbf{DK} = \frac{\partial \hat{\mathbf{x}}}{\partial \mathbf{y}}\frac{\partial \mathbf{y}}{\partial \mathbf{x}} = \frac{\partial \hat{\mathbf{x}}}{\partial \mathbf{x}} . \quad (3.31)$$

The first term in Equation 3.29 is called the bias error and is an inversion of simulated noise free measurements that are created with the forward model using the *a priori* state. For any well behaved retrieval method, this must yield the *a priori*. Therefore, the error in the solution, *i.e.* the difference between the retrieved state and the true state, can be found by rearranging Equation 3.29,

$$\hat{\mathbf{x}} - \mathbf{x} = (\mathbf{A} - \mathbf{I})(\mathbf{x} - \mathbf{x_a}) + \mathbf{D}\,\boldsymbol{\epsilon} + \mathbf{D}\left(\Delta\mathbf{f}(\mathbf{x}, b) + \mathbf{K}_b(b - \tilde{b})\right) . \quad (3.32)$$

The three terms that contribute to the error in the solution, in the same order as they appear in Equation 3.32, are referred to as the smoothing error, the measurement error, and the forward model error. The forward model error is composed of two terms: the error that arises from the approximate nature of the forward model itself and the uncertainty in the inputs that are required to describe the atmospheric state. This uncertainty in the model inputs is sometimes called the forward model parameter error. Each of these terms is addressed and explained in detail in Section 5.5.

3.7.2 Application to MART

In the optimal estimation formulation, where the solution is the peak of the Bayesian probability $P(\mathbf{x}|\mathbf{y})$ and the probability density functions are assumed to be Gaussian, the contribution function matrix and the averaging kernel can be derived analytically from the kernel and the covariance matrices of the measurements and the *a priori*. While the theory is generally applicable to other inverse methods, algebraic derivations of the contribution function and the averaging kernel are not applicable. However, *Puliafito et al.* (1995) have shown that the Rodgers formal error analysis can be applied to several inverse methods by directly calculating the required partial derivatives through numerical perturbations. The perturbations must be sufficiently small that the response is linear, yet not so small as to introduce error due to floating point precision. In this way, the Rodgers formal error analysis can be applied to the MART inversion of OSIRIS data.

Chapter 4

The SASKTRAN Radiative Transfer Model

When people thought the Earth was flat, they were wrong. When people thought the Earth was spherical, they were wrong. But if you think that thinking the Earth is spherical is just as wrong as thinking the earth is flat, then your view is wronger than both of them put together.

Isaac Asimov, 1988

4.1 Background

Radiative transfer models have always been an important tool for the atmospheric community. Such models provide the link between observation and the atmospheric state and are frequently used within data inversion processes to estimate atmospheric parameters. Thus not surprisingly, the accuracy of retrieved parameters depends upon the accuracy of the radiative transfer model; for this reason, there has been a large research effort into radiative transfer models particularly since satellites started observing the Earth and the atmosphere.

The propagation and scattering of light within a planetary atmosphere is controlled by the equation of radiative transfer, originally developed by *Chandrasekhar* (1960). In its most general form the equation is five dimensional: three dimensions

for positional dependence and two for the angular distribution about a point. Most solutions of the radiative transfer equation strive to reduce the number of dimensions to be solved by considering geometries with special symmetric properties. For example, the equation has been solved semi-analytically for a three dimensional, plane-parallel geometry (*Anderson et al.*, 1999) that has had extensive use in the satellite community for the analysis of nadir observations.

The development of satellite-based observations of scattered sunlight in the atmospheric limb, (SOLSE-LORE (*McPeters et al.*, 2000), OSIRIS (*Llewellyn et al.*, 2004), SCIAMACHY (*Bovensmann et al.*, 1999), SAGE III (*Rault*, 2005)), has created a need for operational radiative transfer models that accurately and efficiently model the Earth's limb at optical wavelengths. Figure 4.1 illustrates the limb scatter measurement and shows the different scattering paths of the solar beam into the instrument line of sight. It is well known that while plane-parallel radiative transfer models are suitable for nadir observations, they are not appropriate for limb geometry (*Collins et al.*, 1972). Consequently, there have been several attempts to numerically model radiative transfer for limb-viewing geometries (*Herman et al.*, 1994; *Oikarinen et al.*, 1999; *Griffioen and Oikarinen*, 2000; *Rozanov et al.*, 2001; *McLinden et al.*, 2002; *Kaiser and Burrows*, 2003; *Postylyakov*, 2004b). The inherent geometry and the presence of multiple scattering (*Oikarinen et al.*, 1999) significantly complicate the numerical solution and, invariably, numerical techniques can consume considerable computer time and memory.

In this chapter, the development of a new radiative transfer model, called SASKTRAN, is presented. The name is in keeping with previously developed 'TRAN's, *i.e.* LIMBTRAN, MODTRAN, HITRAN, and highlights the affiliation with the University of Saskatchewan. The basic characteristics of the model are:

- a spherical shell atmosphere of homogeneous cells with variable thickness,
- scattering by molecules and aerosols with an altitude dependent cross section and phase function; absorption by an extensible list of temperature dependent

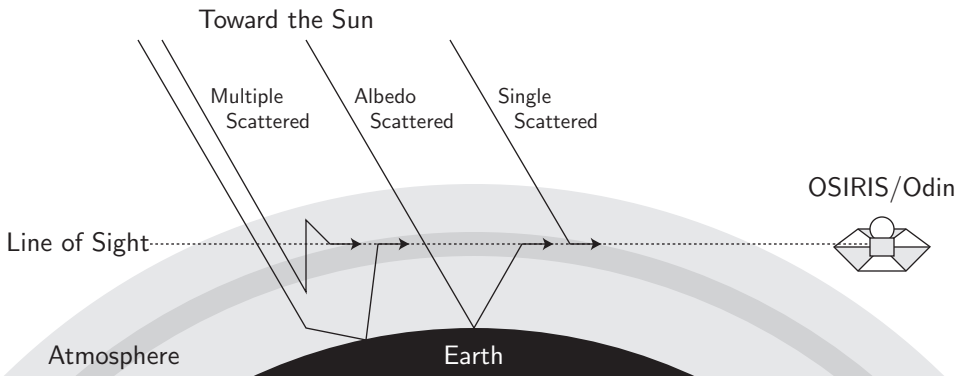

Figure 4.1: The limb scatter measurement technique. The contribution from light that has scattered multiple times in the atmosphere and/or from the ground is significant.

species,

- estimation of n-order scattering by the method of successive orders along rays traced within the spherical geometry, where
  - contribution from the first two scattering events of the solar beam, including the non-uniform illumination of the spherical Earth surface, is calculated directly,
  - contribution of third and higher scattering events to the source term at a specific point is efficiently estimated by integrating the previous order source term calculated at the local zenith,
- specialized numerical integration over the unit sphere that exploits the characteristics of the intensity field in order to optimize the ray tracing.

The code was specifically developed to provide a rapid and accurate solution and so utilizes modern programming techniques such as caching and multi-threading that make full use of multi-processor desktop computers.

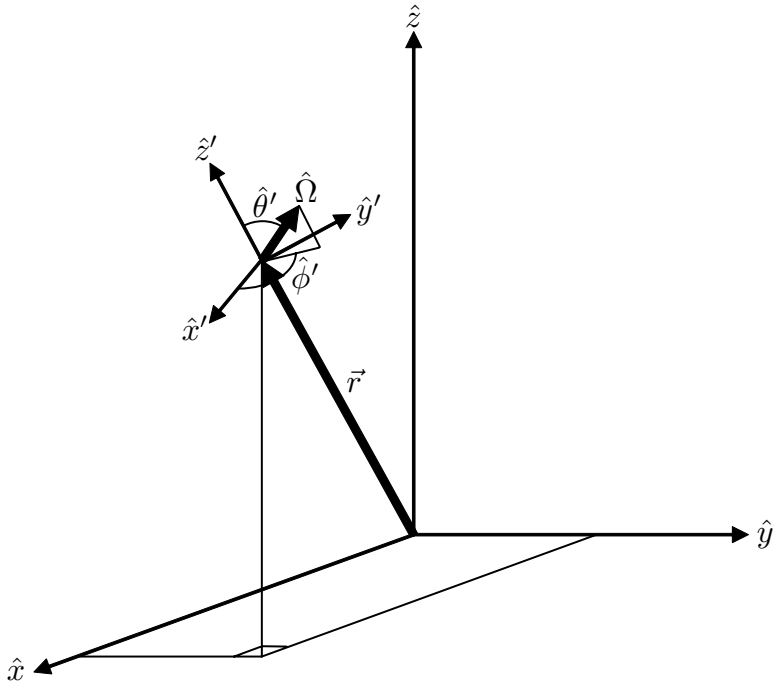

Figure 4.2: Specification of the radiation field, $I(\vec{r}, \hat{\Omega})$, requires five coordinates,: three for the position in space, $\vec{r}$, and two for the directional distribution, $\hat{\Omega}$ (based on Figure 1.7 from *Walter* (2006)).

4.2 Radiative Transfer Theory

4.2.1 Description of the Radiation Field

A monochromatic, randomly polarized radiation field, $I(\vec{r}, \hat{\Omega})$, is fully described by the magnitude and five spatial coordinates. At each position, $\vec{r}$, defined by three Cartesian coordinates, a second "primed" coordinate system is defined by a rotation of the $\hat{z}$-axis so that $\hat{z}'$ is aligned with $\vec{r}$; see Figure 4.2. In the primed coordinates, the direction of the radiation is specified by the two angular coordinates, θ' and ϕ', that compose the unit vector of the radiation propagation, $\hat{\Omega}$.

The basic quantity of the radiation field at position $\vec{r}$ is the *radiant flux density*, $F(\vec{r}, \hat{n})$. It is the energy, $\mathrm{d}E$, per unit time, $\mathrm{d}t$, and per unit area, $\mathrm{d}A$, that is incident on, or emitted from, a surface perpendicular to the direction $\hat{n}$ in the wavelength range

$(\lambda, \lambda+\mathrm{d}\lambda)$,

$$F(\vec{r}, \hat{n}) = \frac{\mathrm{d}E(\vec{r})}{\mathrm{d}t\ \mathrm{d}A\ \mathrm{d}\lambda}\,. \tag{4.1}$$

Therefore the units of radiant flux density are

$$[F(\vec{r}, \hat{n})] = \frac{\text{photons}}{\text{s cm}^2\ \text{nm}} \tag{4.2}$$

where the commonly followed convention of specifying the energy in terms of the number of photons, N, for wavelength λ requires,

$$E = N\frac{hc}{\lambda}\,. \tag{4.3}$$

The *radiance*, $I(\vec{r}, \hat{\Omega})$ is the power measured by an optical system that views an emitting surface area in the direction $\hat{\Omega}$ over a small solid angle $\mathrm{d}\Omega$. When viewed at an angle θ to the surface normal, $\hat{n}$, the differential area on the emitting surface is decreased by the factor $\cos(\theta)$. Therefore, the radiance in wavelength range (λ, $\lambda+\mathrm{d}\lambda$) is

$$I(\vec{r}, \hat{\Omega}) = \frac{\mathrm{d}E(\vec{r})}{\mathrm{d}t\ \mathrm{d}A\cos(\theta)\ \mathrm{d}\lambda\ \mathrm{d}\Omega}\,, \tag{4.4}$$

and has units

$$[I(\vec{r}, \hat{\Omega})] = \frac{\text{photons}}{\text{s cm}^2\ \text{nm sterad}}. \tag{4.5}$$

The flux density can be defined in terms of the radiance as an integral of the projection of the radiance along $\hat{n}$ over the solid angle of the hemisphere,

$$F(\vec{r}, \hat{n}) = \int_{2\pi} I(\vec{r}, \hat{\Omega}')\ \cos(\theta')\,\mathrm{d}\Omega'\,, \tag{4.6}$$

where θ' is the angle between $\hat{\Omega}'$ and $\hat{n}$.

A *Lambertian surface* is defined such that the energy emitted, or reflected, from a surface area element varies with the cosine of the angle from the surface normal. The effect is that the radiance is the same for all viewing directions. Therefore, for

a Lambertian surface, the radiance in Equation 4.6 is no longer a function of the direction $\hat{\Omega}'$ and can be taken outside of the integral,

$$\begin{aligned} F(\vec{r}, \hat{n}) &= I(\vec{r}) \int_{2\pi} \cos(\theta') \, \mathrm{d}\Omega' \\ &= I(\vec{r}) \int_0^{2\pi} \int_0^{\frac{\pi}{2}} \cos(\theta') \sin(\theta') \, \mathrm{d}\theta' \, \mathrm{d}\phi' \\ &= \pi \, I(\vec{r}). \end{aligned} \tag{4.7}$$

The *albedo*, a, of a surface is the ratio of the upward propagating reflected flux, F_{up}, to the downward propagating flux, F_{down},

$$a = \frac{F_{\text{up}}}{F_{\text{down}}} \, . \tag{4.8}$$

Therefore, for an incident down-welling flux, $F_{\text{down}}(\vec{r}, \hat{n})$ on a Lambertian surface with an albedo, a, the observed radiance is

$$I(\vec{r}) = \frac{1}{\pi} \, a \, F_{\text{down}}(\vec{r}, \hat{n}) \tag{4.9}$$

It should be noted that the albedo is generally a function of wavelength.

The formulation presented here specifies the radiation field at a point in three dimensional space, $\vec{r}$, and in the direction $\hat{\Omega}$. However, it is important to consider not just the radiation field at a specific point, but also the transport of radiation from one point to another. Radiation in the atmosphere is assumed to travel in straight lines; while this is not generally true because of refraction, it is true in a differential sense over a small distance, and often refractive effects are negligibly small. Therefore, it is useful to employ a parametrization of the transport path using a geometric path coordinate, s. Point $\vec{r}$ is written in terms of the propagation direction, $\hat{\Omega}$, and a reference point, $\vec{r}_0$,

$$\vec{r} = \vec{r}_0 + \hat{\Omega} \, s \, , \tag{4.10}$$

such that s increases in the direction $\hat{\Omega}$. The reference point can be considered to be the observation position. Therefore, the radiation traveling in the propagation direction ends at the reference point where the path coordinate, s, is defined to be zero. This is illustrated in the vector diagram shown in Figure 4.3. For clarity in further discussion, the radiation field will be written in terms of the path coordinate only. Thus the position reference vector in three dimensional space and the propagation direction are implied by the path. That is,

$$I(s) \equiv I(\vec{r}, \hat{\Omega}). \tag{4.11}$$

4.2.2 The Radiative Transfer Equation

Beer-Lambert Law

The transmission of monochromatic radiation through an infinitesimally thin region of thickness ds, along the propagation direction, $\hat{\Omega}$, is simply described by the Beer-Lambert Law,

$$\frac{\mathrm{d}I(s)}{\mathrm{d}s} = -I(s) \sum_i \sigma_i(s) n_i(s)\,, \tag{4.12}$$

where dI is the change in radiance, I, due to an interaction with atmospheric species, i, that attenuates the incident radiation through absorption and/or scattering processes. The species cross section, $\sigma_i(s)$, which quantifies the probability of interaction in units of area, and the species volume number density, $n_i(s)$, depend upon the location along the propagation path, s, in the atmosphere. The result of the summation in Equation 4.12 is the extinction,

$$k(s) = \sum_i \sigma_i(s) n_i(s), \tag{4.13}$$

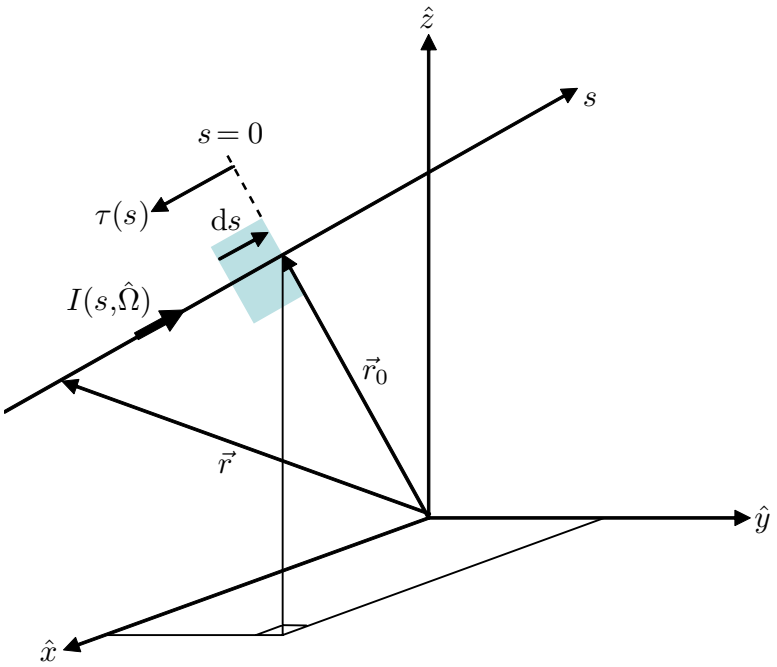

Figure 4.3: The radiation path length coordinate, s, is defined to be zero at the observation point, $\vec{r}_0$, and increases in the direction of radiative propagation, $\hat{\Omega}$.

and can be separated into components that arise from scattering, $k_{\mathrm{scat}}(s)$ and absorption, $k_{\mathrm{abs}}(s)$,

$$k(s) = k_{\mathrm{scat}}(s) + k_{\mathrm{abs}}(s). \tag{4.14}$$

It should be noted that the units of extinction are per unit length.

Figure 4.3 shows a small section of the atmosphere of thickness $\mathrm{d}s$ along the propagation direction. The observer is located at $\vec{r}_0$ corresponding to path coordinate s=0. For the assumption of a path length dependent extinction, $k(s)$, Equation 4.12 can be rewritten as

$$\frac{\mathrm{d}I(s)}{\mathrm{d}s} = -k(s)I(s)\,. \tag{4.15}$$

The radiance at any point along the path is determined by rearranging and integrating this result from s to the reference point,

$$\int_{I(s)}^{I(0)} \frac{\mathrm{d}I(s)}{I(s)} = -\int_{s}^{0} k(s')\,\mathrm{d}s'\,. \tag{4.16}$$

The *optical depth*, τ of a homogeneous layer of thickness ℓ is a dimensionless quantity that describes the fraction of the radiation attenuated by the layer,

$$\tau = k\,\ell. \tag{4.17}$$

Thus at any point along a path with varying extinction, a change in path length is related to a change in optical depth as,

$$\mathrm{d}\tau(s) = -k(s)\,\mathrm{d}s. \tag{4.18}$$

The negative sign arises because a positive change in the path coordinate means that the radiation moves closer to the observer because the observer is at the end point of propagation; therefore, the total optical depth from the observer to s is decreased at $s + \mathrm{d}s$. With reference to Figure 4.3, the optical depth is measured positive in the direction opposite to s.

Substituting Equation 4.18 into Equation 4.16 yields,

$$\int_{I(s)}^{I(0)} \frac{\mathrm{d}I(s)}{I(s)} = \int_{\tau(s)}^{\tau(0)} \mathrm{d}\tau(s), \tag{4.19}$$

with solution,

$$\ln\left(\frac{I(0)}{I(s)}\right) = \tau(0) - \tau(s)\,. \tag{4.20}$$

Since the optical depth at the reference point zero, the radiance observed at the reference point can be written in terms of the radiance at any point along the path, s, and the optical depth between that point and the reference, $\tau(s)$, as

$$I(0) = I(s)e^{-\tau(s)}. \tag{4.21}$$

Addition of the Source Term

The Beer-Lambert Law describes the loss of radiation by extinction of the incident beam, either through absorption or scattering out of the propagation direction. However, radiation can be scattered from other directions into the propagation direction and so increase the radiance. The full form of the radiative transfer equation considers the total change in the radiation along the propagation direction, both positive and negative. Including a source term, $J(s)$ in the Beer-Lambert Law, Equation 4.15, results in the familiar integro-differential form of the radiative transfer equation,

$$\frac{\mathrm{d}I(s)}{\mathrm{d}s} = k(s)(-I(s) + J(s))\,. \tag{4.22}$$

This is essentially a detailed balance equation where the radiation sources and sinks are considered at each point along the propagation direction (*Walter*, 2006). It is also a non-local equation as the radiance at point s depends on the radiance at all other points. The exact solution of Equation 4.22 at any point requires solving for the entire atmosphere. This is the essence of the modelling problem.

In the atmosphere, there are generally three processes that contribute to the radiance: the scattering of radiation into the propagation direction, thermal emission of blackbody radiation, and photochemical emission. At optical wavelengths, i.e. the near ultraviolet, visible, and near infrared ($\lambda < 2\,\mu$m), the magnitude of the scattered sunlight far exceeds that from thermal emission. At those wavelengths where there is no photochemical emission, the source term is an integral of the diffuse radiance over all directions scattered into the propagation direction,

$$J(s, \hat{\Omega}) = \frac{k_{\mathrm{scat}}(s)}{k(s)} \int_{4\pi} I(\vec{r}, \hat{\Omega}')\, \bar{\mathrm{p}}(s, \Theta)\, \mathrm{d}\Omega'\,. \tag{4.23}$$

$I(\vec{r}, \hat{\Omega}')$ is again written explicitly as a function of $\hat{\Omega}'$ to show that the diffuse radiance over all directions is scattered into the propagation direction. The effective

phase function, $\bar{\mathrm{p}}(s, \Theta)$ defines the probability of scattering per unit solid angle from direction $\hat{\Omega}'$ into the propagation direction $\hat{\Omega}$ at point s and is a weighted average of the phase functions of all scattering particles at that point. The scattering angle Θ is the angle between the diffuse radiance and the propagation direction, *i.e.*

$$\cos(\Theta) = \hat{\Omega}' \cdot \hat{\Omega}. \tag{4.24}$$

The fraction $\frac{k_{\mathrm{scat}}(s)}{k(s)}$ is required as only those species that scatter radiation, and not those that absorb radiation, contribute to the source term.

The integro-differential radiative transfer equation can be transformed into a useful integral form. If the path coordinate, s, is converted to optical depth, τ, then the radiative transfer equation (Equation 4.22) becomes,

$$\frac{\mathrm{d}I(\tau)}{\mathrm{d}\tau} = I(\tau) - J(\tau)\,. \tag{4.25}$$

The derivative of the term $I(\tau)e^{-\tau}$, is,

$$\frac{\mathrm{d}}{\mathrm{d}\tau}\left(I(\tau)e^{-\tau}\right) = e^{-\tau}\left(\frac{\mathrm{d}I(\tau)}{\mathrm{d}\tau} - I(\tau)\right). \tag{4.26}$$

Substitution of this result into Equation 4.25 yields an integrable form of the radiative transfer equation,

$$\frac{\mathrm{d}}{\mathrm{d}\tau}\left(I(\tau)e^{-\tau}\right) = -J(\tau)\,e^{-\tau}\,. \tag{4.27}$$

The integral from a reference point, τ_0, to any point along the path, τ, is,

$$I(\tau)e^{-\tau} - I(\tau_0)e^{-\tau_0} = -\int_{\tau_0}^{\tau} J(\tau')\,e^{-\tau'}\,\mathrm{d}\tau'\,. \tag{4.28}$$

By choosing the reference point as $\tau_0 = 0$, the radiance at an arbitrary point, τ, seen by the observer is,

$$I(0) = I(\tau)e^{-\tau} + \int_{0}^{\tau} J(\tau')\,e^{-\tau'}\,\mathrm{d}\tau'\,. \tag{4.29}$$

If Equation 4.29 is transformed back to path length coordinates, the commonly stated integral form of the radiative transfer equation is obtained,

$$I(0) = I(s)e^{-\tau(s)} + \int_s^0 k(s')J(s')\, e^{-\tau(s')}\, \mathrm{d}s' \,. \tag{4.30}$$

The units of the term $k(s)J(s)$, which appears in the integration along the line of sight, are

$$[k(s)J(s)] = \frac{\text{photons}}{\text{s cm}^3 \text{ nm sterad}} \,. \tag{4.31}$$

Because of this, $k(s)J(s)$ is sometimes referred to as the volume emission rate (VER); however, this term is traditionally used to describe photochemical emission. To avoid confusion with the process that arises from scattering only, the present work will follow the convention that the radiance $J(s)$ is simply called the "source term".

4.2.3 Radiative Processes

Rayleigh Scattering

The scattering of sunlight by the molecular neutral density, quantified by Lord Rayleigh in the 19$^{\text{th}}$ century, is the dominant scattering process in the atmosphere and is invoked to explain the blue color of the sky. It is an elastic scattering process so that the wavelength of the scattered light and the incident light are identical. The scattering molecule does not absorb any of the incident radiation nor does the interaction stimulate emission. Inelastic Raman scattering processes, which result in changes in wavelength due to changes in the rotational energy of the molecule, contribute only a small effect and can be neglected in the present work.

Rayleigh scattering is the small particle limit of scattering theory. Atmospheric molecules are considered as homogeneous spherical particles that are much smaller than the wavelength of the incident light. The electric field, $\vec{E}_0$, of the incident radiation induces a dipole moment, $\vec{p}$, in the molecule with a strength that depends

on the polarizability, α,

$$\vec{p} = \alpha \vec{E}_0. \tag{4.32}$$

The dipole oscillates at the same frequency as the incident electric field and generates a scattered electric field proportional to the acceleration of the dipole. For an oscillating periodic dipole of wavenumber, $k = 2\pi/\lambda$, the scattered field is proportional to k^2 and results in an intensity proportional to k^4.

The scattering plane is formed by the incident propagation vector and the scattered propagation vector. For an electric field that oscillates purely perpendicular to the scattering plane, the scattered electric field is independent of the scattering angle. For an electric field that oscillates parallel to the scattering plane, the scattered field depends on the cosine of the scattering angle. Since sunlight is randomly polarized, there is no phase relation between orthogonal components of the incident field and the total intensity is the sum of the intensity due to each component. Therefore, the directional dependence of the scattering, *i.e.* the phase function, $\mathrm{p}(\Theta)$, is proportional to $1 + \cos^2(\Theta)$. The phase function is normalized over solid angle such that,

$$\frac{1}{4\pi} \int_0^{2\pi} \int_0^{\pi} \mathrm{p}(\Theta) \, \sin(\Theta) \, \mathrm{d}\Theta \, \mathrm{d}\phi = 1. \tag{4.33}$$

Therefore, the phase function is,

$$\mathrm{p}(\Theta) = \frac{3}{4}(1 + \cos^2(\Theta)). \tag{4.34}$$

The Rayleigh scattering cross section, σ_r, defines the interaction of a single molecule with a beam of light. If the scattered intensity is integrated over solid angle, the cross section can be determined as the ratio of the scattered flux to the incident flux density,

$$\sigma_r = \frac{128\pi^5}{3} \frac{\alpha^2}{\lambda^4}. \tag{4.35}$$

The diatomic nature of the molecules that comprise air, mainly O_2 and N_2, causes

a small deviation from this result, which assumes that the scattering particles are perfect homogeneous spheres. Because of the orientation of the diatomic molecule, the induced dipole moment may not be perfectly aligned with the incident electric field. In this case, the polarizability is a tensor and results in a small wavelength dependent change in the cross section. The species dependent King correction factor (*King*, 1923) accounts for the depolarization; a weighted average over all species in dry air yields the wavelength dependent cross section. The atmospheric standard is given in *Bates* (1984). An efficient estimate of the cross section can be analytically determined from an empirical fit as a function of wavelength (*Nicolet et al.*, 1982),

$$\sigma_r = \frac{4 \times 10^{-28}}{\lambda^{3.916 + 0.074\lambda + 0.05\lambda^{-1}}} \,\mathrm{cm}^2 \,, \tag{4.36}$$

where λ is in μm.

Gas Absorption

The absorption of solar radiation by atoms and molecules in the atmosphere occurs through quantum mechanical transitions of the electronic, vibrational, and rotational energies of the particle. The ability of an atomic or molecular species to absorb light is measured by the absorption cross section, σ_i, for species i. This process is wavelength dependent and cross sections can also be a function of temperature and pressure. Cross sections are typically measured in the laboratory; Figure 4.4 shows the absorption cross sections for two molecules that are important to the OSIRIS mission, O_3 and NO_2, as measured by *Burrows et al.* (1999) and *Burrows et al.* (1998) respectively. The absorption in the atmosphere generally results in a photo-dissociation, ionization, fluorescence, or heating primarily through collisional quenching.

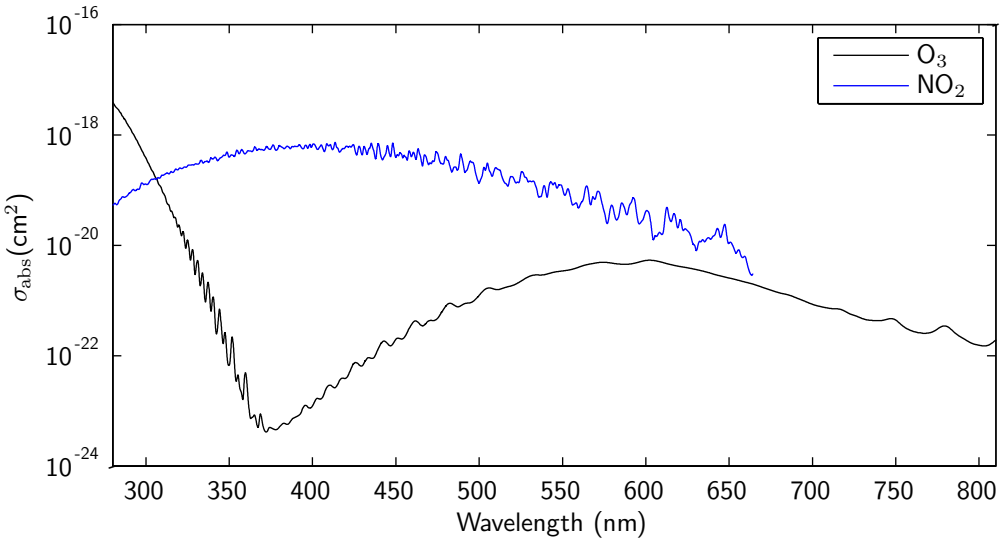

Figure 4.4: The absorption cross section of O_3 (203 K) and NO_2 (220 K) at OSIRIS wavelengths.

Aerosol Scattering

The size of stratospheric aerosol particles is of the same order as the wavelength of visible light. Under these conditions the assumptions used to describe Rayleigh scattering are not valid and more rigorous treatment of the scattering from a spherical particle is required. *Mie* (1908) was the first to publish the formal solution. It is essentially a boundary value problem for which the total transverse components of the electromagnetic field must be continuous at the surface of the dielectric sphere. Therefore, the solution is in three parts: the solution of the vector wave equation in spherical coordinates, the expansion of an incident plane wave in spherical harmonics, and boundary matching to determine the internal and external scattered waves. The geometry is illustrated qualitatively in Figure 4.5. An important difference between this scattering problem and the Rayleigh limit is that the index of refraction of the dielectric sphere is not necessarily purely real so that the solution accounts for both scattering and absorption of the incident wave.

The vector wave equation arises from the Maxwell equations in charge and current free regions. The solution can be constructed by forming the electric and magnetic

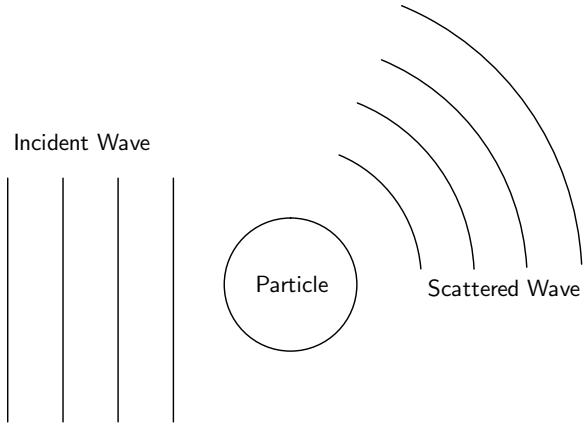

Figure 4.5: The incident plane wave and a spherical wave scattered from a dielectric sphere that is approximately the same size as the wavelength of the incident wave.

field vectors as orthogonal combinations of solutions to the scalar wave equation that can be found by separation of the variables in spherical coordinates. The solution takes the form of a combination of trigonometric functions, the Legendre polynomials (spherical harmonics of the first kind), and Hankel functions. The expansion of the plane wave is found by considering an infinite series of Legendre polynomials. At the surface of the sphere the sum of the tangential component of the incident and scattered electromagnetic field (the external field) must equal the internal field for continuity at the boundary. The final results for observations at a distance much greater than the particle radius, r, are the efficiencies of extinction and scattering, $Q(r)$ and $Q_{\text{scat}}(r)$ respectively, and the scattering phase function, $\text{p}(\Theta)$, also a function of the particle radius. For wavenumber $k = 2\pi/\lambda$, the efficiencies are,

$$Q(r) = \frac{2}{(kr)^2} \sum_{n=1}^{\infty} (2n+1)\text{Re}(A_n + B_n)\,, \tag{4.37}$$

$$Q_{\text{scat}}(r) = \frac{2}{(kr)^2} \sum_{n=1}^{\infty} (2n+1)(|A_n|^2 + |B_n|^2)\,, \tag{4.38}$$

where the coefficients A_n and B_n are given in terms of a normalized half integral order spherical Bessel functions, $\psi_n(\alpha)$, and Hankel functions, $\xi_n(\alpha)$, and the index

of refraction, m,

$$A_n = \frac{\psi_n'(mkr)\psi_n(kr) - m\psi_n(mkr)\psi_n'(kr)}{\psi_n'(mkr)\xi_n(kr) - m\psi_n(mkr)\xi_n'(kr)}\,, \tag{4.39}$$

$$B_n = \frac{m\psi_n'(mkr)\psi_n(kr) - \psi_n(mkr)\psi_n'(kr)}{m\psi_n'(mkr)\xi_n(kr) - \psi_n(mkr)\xi_n'(kr)}\,. \tag{4.40}$$

The analytic form of the scattering phase function is derived in a similar manner.

The efficiencies are related to the cross sections by the cross sectional area of the particle,

$$\sigma(r) = \pi r^2\, Q(r)\,, \tag{4.41}$$

$$\sigma_{\text{scat}}(r) = \pi r^2\, Q_{\text{scat}}(r)\,. \tag{4.42}$$

The absorption cross section is fixed by these results as,

$$\sigma_{\text{abs}}(r) = \sigma(r) - \sigma_{\text{scat}}(r)\,. \tag{4.43}$$

The solution is an infinite series that converges slowly. The calculations are quite long and can use significant computational resource. An efficient and accurate numerical evaluation of the series was the topic of much work in previous decades and resulted in the development of a some widely used standard codes. The code developed by *Wiscombe* (1980) is used in the present work.

As discussed in Section 2.1, the mixing and coagulation processes in the aerosol life cycle result in an altitude dependent particle size distribution, often modelled as log-normal (Equation 2.1). Thus the effective scattering cross section for a distribution of particle sizes requires an integration of the cross section per particle over the distribution (*Hansen and Travis*, 1974),

$$\sigma = \int_0^\infty \pi r^2\, Q(r)\, n(r)\, \mathrm{d}r\,, \tag{4.44}$$

where $n(r)$ is the distribution of radii that specifies the number of particles of size between r and $r+\mathrm{d}r$ and is normalized such that

$$\int_0^\infty n(r)\,\mathrm{d}r = 1\,. \tag{4.45}$$

The phase function is treated similarly,

$$\mathrm{p}(\Theta) = \int_0^\infty \mathrm{p}(\Theta, r)\, n(r)\,\mathrm{d}r\,. \tag{4.46}$$

Figure 4.6 shows the scattering cross section of a log-normal distribution of spherical H_2SO_4 particles with a mode radius, r_g, of 0.3 μm and mode width, s_g, values of 1.1, 1.4, 1.7 and 2.0. The integral in Equation 4.44 is evaluated by Gaussian quadrature of order 1000 from $r = r_g/(5s_g)$ to $r = r_g(5s_g)$. The Mie cross section of a single particle typically exhibits a large amount of wavelength structure and reaches a peak in the region where the wavelength is approximately equal to the particle size. The effect of increasing the width of the distribution is to smooth the structure and shift the peak to longer wavelengths that correspond to the larger particle sizes in the skewed tail of the log-normal distribution as shown in Figure 4.6.

A widely used parametrization of the spectral dependence of the aerosol cross section in the optical region is the Ångström exponent relation (*Ångström*, 1964),

$$\sigma(\lambda) = \sigma_0 \left(\frac{\lambda}{\lambda_0}\right)^{-x}, \tag{4.47}$$

where x typically ranges from -0.5 to 2.0. The largest values of the exponent x correspond to the smallest particle sizes. This approximation demonstrates an important characteristic of aerosol scattering; the wavelength dependence of the cross section is weaker than for the Rayleigh cross section, which corresponds to an exponent of $x \simeq 4$. The clear sky, dominated by Rayleigh scattered light, is blue in colour, whereas Mie scattered sunlight, commonly observed from clouds, is white in color

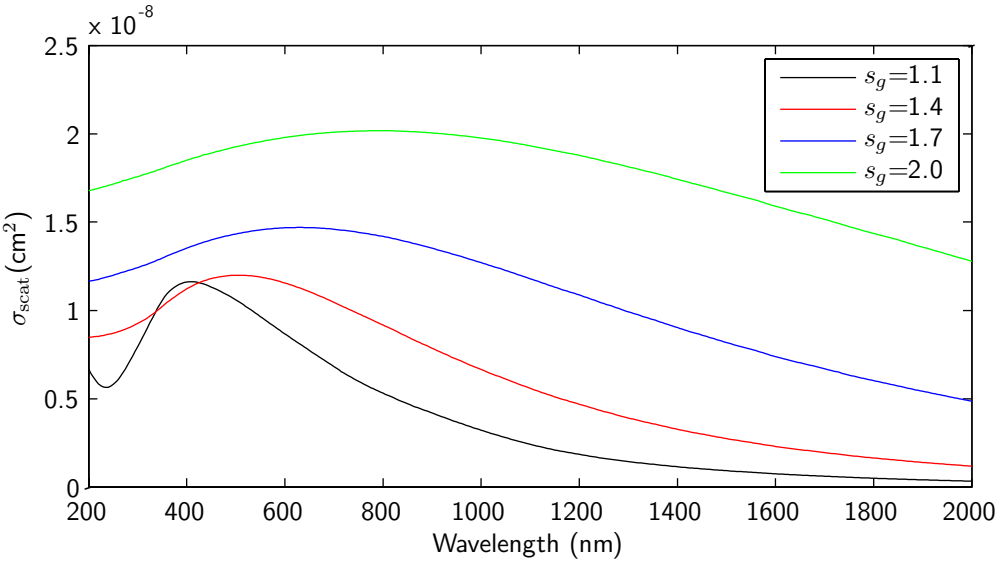

Figure 4.6: The effective scattering cross section of a log-normal distribution of spherical H_2SO_4 particles with a mode radius, r_g, of 0.3 μm and a range of mode widths.

due to the weaker wavelength dependence.

Figure 4.7 is a plot of aerosol scattering cross sections calculated from the Mie solution and from the Ångström relation. The Mie cross sections are calculated for log-normally distributed particles with mode radii of 0.1 μm and 0.5 μm and a mode width of 1.5 in both cases. These represent approximate maximum and minimum stratospheric aerosol sizes. The Ångström parameterizations shown in Figure 4.7 are for $x = -0.5$ and $x = 2.0$, both normalized to 10^{-8} cm^2 at λ=1000 nm. It is clear that the spectral dependance of the $x = 2.0$ Ångström relation is similar to that for the Mie cross sections for the smaller particle case, and that the $x = -0.5$ Ångström relation captures the weaker wavelength dependence of the larger particle Mie cross sections. However, there are important differences, and although the Ångström exponent is indicative of the particle size, it is not used in this thesis work due to the accuracy and efficiency of available Mie codes. It is included here for completeness because of its frequent occurrence in other work.

The scattering phase function of aerosol size particles is very anisotropic with a large peak in the forward direction, i.e. near $\Theta = 0$. The peak becomes more

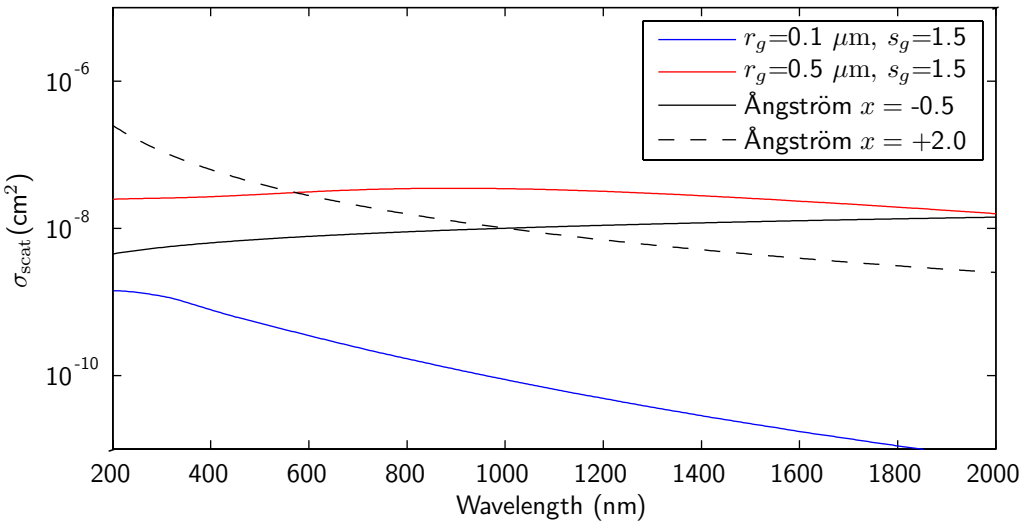

Figure 4.7: Mie scattering cross section of a log-normal distribution of particles with mode radii of 0.1 μm and 0.5 μm, and the Ångström exponent relation corresponding to maximum and minimum aerosol sizes.

pronounced for larger particles. There is also a characteristic smaller enhancement in the backward direction, near $\Theta = 180°$, called the *glory*, that is also more significant for the larger particle sizes. Figure 4.8 is plot of the scattering phase function calculated using the Mie solution for the same two sets of log-normally distributed particles shown in Figure 4.7. The ordinate axis has a logarithmic scale. The forward peak and the glory are evident in both cases, but both are more pronounced at the larger particle size. The Rayleigh phase function (Equation 4.34) is also included for reference. All phase functions are normalized to 1.0. Compared to the Rayleigh phase function, the asymmetry is clear; Rayleigh scattering appears to be almost isotropic compared to aerosol scattering.

An often used parametrization of the scattering phase function is the Henyey-Greenstein approximation (*Henyey and Greenstein*, 1941),

$$\mathrm{p}(\Theta) = \frac{1}{4\pi} \frac{1 - g^2}{(1 - 2g\cos(\Theta) + g^2)^{3/2}}\,, \tag{4.48}$$

where the factor g is called the *asymmetry factor* and is typically about 0.7 for strato-

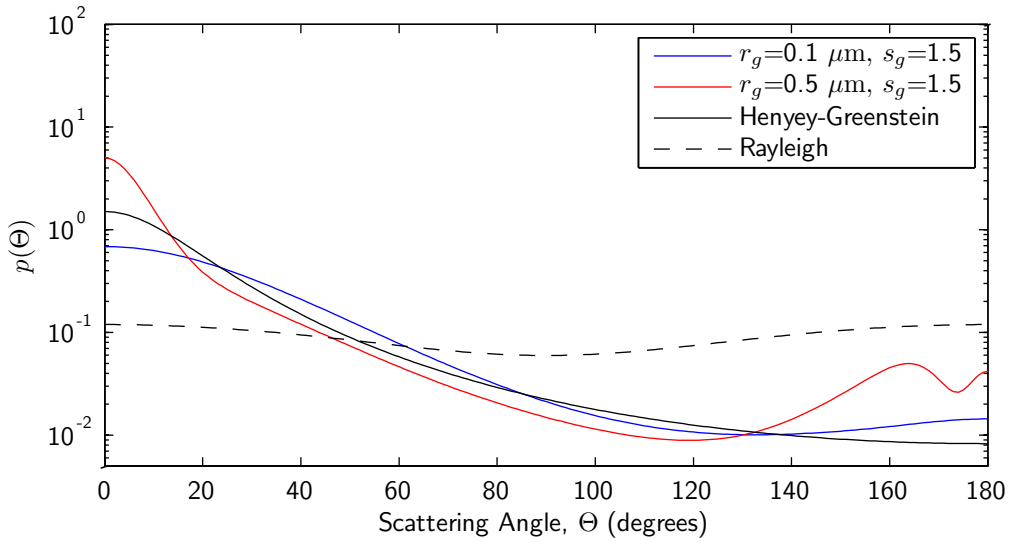

Figure 4.8: The Mie effective phase function for log-normally distributed particles with mode radius of 0.1 μm and 0.5 μm at a wavelength of 500 nm, and the Henyey-Greenstein parametrization with $g = 0.7$. Rayleigh phase function is included for reference.

spheric aerosols. This curve is also included in Figure 4.8. The Henyey-Greenstein curve is somewhat appropriate as it does capture the forward scattering feature and the approximate angular dependence. Like the Ångström exponent, the Henyey-Greenstein approximation is often used, although it is not used for the OSIRIS retrievals because of the availability of the Mie solution. However, it is used in the present work for comparison of the SASKTRAN model with other radiative transfer models in Section 4.7.1 simply because it is used in the other models.

It is important to note that for the range of OSIRIS solar scattering angles, approximately $60^\circ < \Theta_T < 120^\circ$ (see Figure 2.4), all three aerosol phase functions are relatively smooth and have a similar angular dependence. For single scattered sunlight, the phase functions are not significantly different. However, for multiply scattered light from the atmosphere or from the ground, which comes from all directions (see Figure 4.1), the difference in the phase functions is important.

4.3 Algorithm

The algorithm for the SASKTRAN model is presented here in two parts: the theoretical development, and the implementation. In the theoretical development section the problem is defined in terms of the continuous integral equations for radiative transfer. An attempt is made to describe the stages of the solution mathematically in a way that represents the model operation. The implementation section specifically addresses how the each of these stages is computationally realized and the assumptions and approximations that are required.

4.3.1 Theoretical Development

The integral form of the equation of radiative transfer, Equation 4.30, that is included again here for easy reference,

$$I(0) = I(s)e^{-\tau(s)} + \int_s^0 k(s')J(s')\, e^{-\tau(s')}\, \mathrm{d}s' \,, \tag{4.29}$$

states mathematically that the radiance observed at a location, $s = 0$, in a direction along the straight line path, s', is the integral sum of the scattered source term along the path attenuated to the observer, plus the radiance at the end of the path, also appropriately attenuated back to the observer. It is the basic operational requirement of any radiative transfer model to estimate the source function along the observer line of sight and perform the line integral along the path.

The SASKTRAN model approximates the observed radiance by multiple applications of the integral form of the radiative transfer equation through the method of successive orders (*Hansen and Travis*, 1974) along rays traced in a spherical atmosphere. Briefly, the successive orders of scattering method is initialized by the solar irradiance incident upon the atmosphere. The first scatter of the solar beam, at all locations in the sunlit atmosphere, produces the first order source term. This source is propagated throughout the atmosphere and scattered again at all locations

to produce the second order source term. Similarly the second order is propagated and scattered to generate the third order, the third order generates the fourth, and so on. The total source function at any point along the line of sight is the sum of the source terms from each scattering order. For atmospheric scattering in the optical region, the magnitude of the terms diminishes quickly and the algorithm typically converges within 8 to 10 orders. In the spherical shell geometry, many rays, or lines of sight, in many locations and in many directions, are required to estimate the source function. The fundamental ray defined by the observation line of sight of an optical instrument is specifically referred to in the following as the instrument line of sight.

For clarity, Equation 4.30 can be rewritten in terms of the absolute position of the observer, r_0 at $s = 0$, radiative propagation direction, $\hat{\Omega}$, and path end point s_1. The path length coordinate, s, is still used for terms that are evaluated along the line of sight.

$$I(\vec{r}_0, \hat{\Omega}) = \int_{s_1}^{0} J(s, \hat{\Omega}) e^{-\tau(s,0)} k(s) \mathrm{d}s + \tilde{I}(s_1, \hat{\Omega}) e^{-\tau(s_1,0)} \,. \tag{4.49}$$

This equation applies in the most general case and can be used to find the radiance at any point, $\vec{r}_0$, inside or outside the atmosphere. The end point of the path, s_1, for atmospheric ray tracing, is either at the top of the atmosphere or at the ground. For this reason, the radiance at the end point is denoted specially as $\tilde{I}$. At the top of the atmosphere $\tilde{I}(s_1, \hat{\Omega})$ is zero; otherwise $\tilde{I}(s_1, \hat{\Omega})$ is the brightness of the ground element intersected by the line of sight. It is assumed that scattering from the ground is Lambertian (see Equation 4.9) so that there is no dependence on $\hat{\Omega}$ and the radiance at the ground can be written simply as $\tilde{I}(s_1)$.

It is important to note that the above definition of $\tilde{I}(s_1)$ excludes the case where the observation direction, $-\hat{\Omega}$, is directly towards the sun. This condition is not handled by the SASKTRAN model as the subtler effects of multiple scattering that are of concern in the present work are negligible in a solar occultation.

A key aspect of the SASKTRAN algorithm is a subdivision of the source terms, including the ground radiance, by scattering order. Equation 4.49 can be expressed

as

$$I(\vec{r}_0, \hat{\Omega}) = \int_{s_1}^{0} \left[J_1(s, \hat{\Omega}) + J_2(s, \hat{\Omega}) + \sum_{i=3}^{\infty} J_i(s, \hat{\Omega}) \right] e^{-\tau(s,0)} k(s) \mathrm{d}s + \left[\tilde{I}_1(s_1) + \tilde{I}_2(s_1) + \sum_{i=3}^{\infty} \tilde{I}_i(s_1) \right] e^{-\tau(s_1,0)} \tag{4.50}$$

where $J_1(s, \hat{\Omega})$ is the single scatter source term, $J_2(s, \hat{\Omega})$ is the source term due to light that has been scattered twice, and $J_i(s, \hat{\Omega})$ is the source term due to light that has been multiply scattered three times or more. Similarly, $\tilde{I}(s_1)$ is now subscripted by 1 to denote a single scattering of the solar beam from the ground, $\tilde{I}_2(s_1)$ is the radiance of the ground element due from the second scattering event, and so on. It should be noted that in this formulation, for all multiple scatter source terms, that is J_2 through J_∞, the last scattering event of that order occurs within the atmosphere, and for the ground radiance terms $\tilde{I}_1$ through $\tilde{I}_\infty$ the last scattering event is from the ground. Equation 4.50 explicitly separates the source term into three types: single scattered light, light scattered twice, and light scattered more than twice. This reflects a specific and important aspect of the SASKTRAN algorithm in that the approximations used to calculate each of these terms are different. Three, and higher, orders of scattering are handled in exactly the same fashion and are, therefore, expressed as a single summation term in Equation 4.50. The radiative transfer equations that apply specifically to each of these terms are developed in the following sections.

Light That is Scattered Once

In the subdivided form of the integral radiative transfer equation (Equation 4.50), there are two terms that involve light that is scattered identically once. These terms are the atmospheric source along the path, $J_1(s, \hat{\Omega})$, and the radiance at the ground, $\tilde{I}_1(s_1)$. $\tilde{I}_1(s_1)$ is determined by attenuating the solar beam, $F_0(\hat{\Omega}_0)$, to the ground

and subsequent Lambertian scattering (Equation 4.9) of the downward component,

$$\tilde{I}_1(s_1) = \frac{a}{\pi} F_0(\hat{\Omega}_0)\, e^{-\tau(\text{sun},s_1)}\, \cos(\theta_{sza})\,, \tag{4.51}$$

where a is the albedo or surface reflectance, θ_{sza} is the solar zenith angle at the ground point s_1, and $\tau(\text{sun}, s_1)$ is the optical depth from the top of the atmosphere to the ground.

The single scatter source term, $J_1(s, \hat{\Omega})$, is determined by attenuating the solar beam to the points along the path, s, and then scattering into the direction $\hat{\Omega}$. Mathematically, this is

$$J_1(s, \hat{\Omega}) = \frac{k_{\text{scat}}(s)}{k(s)}\, F_0(\hat{\Omega}_0) e^{-\tau(\text{sun},s)}\, \bar{\text{p}}(s, \hat{\Omega}, \hat{\Omega}_0), \tag{4.52}$$

where $\tau(\text{sun}, s)$ is the optical depth from the sun (in reality, the top of the atmosphere) to the scattering point, and $\bar{\text{p}}(s, \hat{\Omega}, \hat{\Omega}_0)$ is the effective phase function for scattering from the solar direction, $\hat{\Omega}_0$, into the $\hat{\Omega}$ direction. Because the solar beam is the only source of light for the first order scattering, the integral over all space that is required to evaluate the source term (Equation 4.23) simplifies to Equation 4.52 as the solar direction is a delta function in solid angle. In this definition of the source function, J_1, thermal or photochemical sources could be included in the model if required.

The contribution to the radiance at any position $\vec{r_0}$ in the direction $\hat{\Omega}$ arising from a single scattering event of the solar beam in the atmosphere, or at the surface of the Earth, can now be determined from the line integral along the direction $\hat{\Omega}$,

$$I_1(\vec{r_0}, \hat{\Omega}) = \int_{s_1}^{0} J_1(s, \hat{\Omega}) e^{-\tau(s,0)} k(s) \mathrm{d}s + \tilde{I}_1(s_1) e^{-\tau(s_1,0)}. \tag{4.53}$$

This is the single scatter evaluation of Equation 4.49.

Light That is Scattered Twice

The source term in Equation 4.50 that arises from light that has been scattered twice is logically determined from a successive scattering of light that has been scattered once. Therefore, it is the integral of the radiance, from light scattered once, over all solid angles and scattered a second time with absolute efficiency $k_{\text{scat}}(s)$ and phase function $\bar{\text{p}}(s, \hat{\Omega}, \hat{\Omega}')$. Thus,

$$J_2(s, \hat{\Omega}) = \frac{k_{\text{scat}}(s)}{k(s)} \int_{4\pi} I_1(s, \hat{\Omega}')\, \bar{\text{p}}(s, \hat{\Omega}, \hat{\Omega}')\, \text{d}\Omega' \,. \tag{4.54}$$

or

$$J_2(s, \hat{\Omega}) = \frac{k_{\text{scat}}(s)}{k(s)} \int_{4\pi} \left[\int_{s_1'}^{s} J_1(s', \hat{\Omega}') e^{-\tau(s', s)} k(s) \text{d}s' + \tilde{I}_1(s_1') e^{-\tau(s_1', s)} \right] \bar{\text{p}}(s, \hat{\Omega}, \hat{\Omega}')\, \text{d}\Omega' \,, \tag{4.55}$$

after substitution of Equation 4.53 into Equation 4.54. The recursive nature of Equations 4.54 and 4.53 should be noted; $J_2(s, \hat{\Omega})$ is the spatial integral of the scattered radiance, $I_1(s, \hat{\Omega})$, which itself is the line integral of the source term $J_1(s, \hat{\Omega})$. It should also be noted again that $\tilde{I}_1(s_1')$ is zero if the ray does not intersect the ground.

In a fashion similar to the single scatter contribution, the contribution to the radiance at any position $\vec{r_0}$ in the direction $\hat{\Omega}$ arising from a two scattering events of the solar beam, $I_2(\vec{r_0}, \hat{\Omega})$, can now be determined from the line integral along the direction $\hat{\Omega}$ of the second order source term,

$$I_2(\vec{r_0}, \hat{\Omega}) = \int_{s_1}^{0} J_2(s, \hat{\Omega}) e^{-\tau(s, 0)} k(s) \text{d}s + \tilde{I}_2(s_1) e^{-\tau(s_1, 0)} \,, \tag{4.56}$$

where the radiance at the ground is an integral over all upward directions of the single scatter radiance Lambertian scattering from the ground, *i.e.*,

$$\tilde{I}_2(s_1) = \frac{a}{\pi} \int_{2\pi} I_1(s_1, \hat{\Omega}')\, \cos(\theta')\, \text{d}\Omega' \,. \tag{4.57}$$

Light That is Scattered Three or More Times

Following the same convention, $I_i(\vec{r_0}, \hat{\Omega})$ is defined as the radiance observed at any point, $\vec{r_0}$, in the direction $\hat{\Omega}$ from light that has been scattered identically i times. Therefore,

$$I_i(\vec{r_0}, \hat{\Omega}) = \int_{s_1}^{0} J_i(s, \hat{\Omega}) e^{-\tau(s,0)} k(s) \mathrm{d}s + \tilde{I}_i(s_1) e^{-\tau(s_1,0)}, \tag{4.58}$$

where

$$J_i(s, \hat{\Omega}) = \frac{k_{\mathrm{scat}}(s)}{k(s)} \int_{4\pi} \left[I_{i-1}(s, \hat{\Omega}') \right] \bar{\mathrm{p}}(s, \hat{\Omega}, \hat{\Omega}') \, \mathrm{d}\Omega' \,, \tag{4.59}$$

and $\tilde{I}_i(s_1)$ is zero if the ray does not intersect the ground.

In a fashion similar to the second order case, the radiance at the ground for third and higher order scattering events is an integral over all upward directions of the radiance arising from the previous order Lambertian scattering at the ground,

$$\tilde{I}_i(s_1) = \frac{a}{\pi} \int_{2\pi} I_{i-1}(s_1, \hat{\Omega}') \, \cos(\theta') \, \mathrm{d}\Omega'. \tag{4.60}$$

It should be noted again that this is a recursive definition. Equation 4.59 requires the evaluation of Equation 4.58 for one less order of scattering. However, Equation 4.58 requires evaluation of Equation 4.59 at the same order. Equation 4.60 is also recursive as it requires Equation 4.58 to be evaluated from the ground point to the top of the atmosphere.

4.3.2 Implementation

The subdivision of the radiative transfer source terms, detailed in the previous sections, provides the basis for implementing the SASKTRAN estimate of the radiance observed along the instrument line of sight. The algorithm is now defined as the calculation of the each of the source terms (Equations 4.52, 4.54 and 4.59) at appropriate points for numerical integration along the instrument line of sight and, if the line of sight intersects the Earth's surface, the addition of the ground radiance

(Equations 4.51, 4.57 and 4.60).

This section provides the details of the calculation of each of these terms and the assumptions and approximations that are made in the implementation. The primary assumptions in the model are the use of an osculating sphere coordinate system to approximate the oblate spheroid geometry, the homogeneity of atmospheric properties in spherical cells, and straight line ray tracing for the transfer of radiation. All integrations are approximated numerically by discretization. These include the line integral for calculating the optical depth, the line integral of source terms along a path, and the spatial integral required to scatter incoming radiances for an estimate of the next order source term. For the third and higher order source terms, a further approximation in the scattering geometry is used to provide computational efficiency without a large impact on the final result.

The Instrument Line of Sight and the Sun

The SASKTRAN model uses an Earth centered coordinate system defined by the polar axis and the prime meridian. In this coordinate system, the $\hat{z}$ axis is directed from the centre of the Earth to the North Pole, the $\hat{x}$ axis points towards the prime meridian and the $\hat{y}$ axis is set according to a standard right-handed system. The surface of the oblate spheroid that represents the shape of the Earth is approximated by the IAU1976 reference geoid, which is within 100 m of the gravitational shape of the Earth. The position of the observing instrument, $\vec{r}_0$, the look direction, $-\hat{\Omega}$, and the solar direction, $\hat{\Omega}_0$, are user configurable. In general, the observer can be located anywhere in three-dimensional space looking in any direction. The solar direction can be set explicitly or determined from ephemeris calculations for a given universal time. Optionally for a limb geometry, the configuration can be set by defining the tangent point altitude and solar angles (see Figure 2.1) or linked directly to the OSIRIS Application Programming Interface (API) code base.

This specification of the instrument line of sight and the sun position accommo-

dates the full three-dimensional solar geometry required to model satellite measurements of limb scattered sunlight.

The Osculating Sphere Grid

SASKTRAN employs a spherical geometry to approximate the oblate sphere of the Earth and the atmosphere. The best choice of sphere depends upon the location on the Earth. SASKTRAN uses the osculating sphere, which is defined by matching the curvature of the meridional ellipse on the geoid to a sphere at a surface geographical reference point. The reference point is determined as the latitude/longitude average of points along the instrument line of sight. For limb observing geometry, this is the tangent point.

By the definition of the osculating sphere, the unit vector that points from the centre of the sphere, in the geographic coordinate system, to the reference point where the sphere is tangent to the Earth is the local zenith direction. Along this direction, a change in spherical radial distance is exactly equal to a change in altitude above the surface of the Earth.

Spherical cells of variable thickness that are concentric with the osculating sphere represent the atmosphere. A confusion sometimes arises regarding the use of the term spherical "shell" in this context as it can be unclear as to whether the reference is to the boundary of the layer or to the layer itself. In the present work, the term spherical "cell" is used exclusively to refer to the layer of finite thickness through which a line of sight has a finite path length of intersection. In the zenith direction the difference between the radial distance of the centre of a cell and radius of the osculating sphere is the altitude of the cell. For directions other than the zenith, this is not the exactly true. However, the deviation is small, less than 100 m, in the angular region around the zenith covered by a line of sight. For the purpose of the present work, the cells are 1000 m thick vertically and are defined between 0 and 100 km along the zenith direction. Within a spherical cell, all atmospheric properties are

assumed to be homogeneous regardless of the absolute position in three-dimensional space. This does not require that two points in the same cell have the same radiative source terms as the solar geometry is different at each point.

It can be seen from the integral form of the radiative transfer equation (Equations 4.23 and 4.30) that three parameters are required to describe the radiative properties of the atmosphere. For the assumption of homogeneity in a spherical cell, these parameters are a function of altitude, h, only, and are the total extinction, $k(h)$, the extinction due to scattering $k_{\text{scat}}(h)$, and the effective phase function, $\bar{\text{p}}(h, \Theta)$. Computationally, for the centre point of each spherical grid cell, one value is required for each of the two extinction parameters and an array of values represents the phase function at a configurable angular resolution.

The user interface to the model provides the specification of these three optical properties through height profiles of atmospheric temperature and species densities with corresponding cross sections and scattering phase matrices. The species that are included for a given calculation are completely user configurable. Under default conditions, SASKTRAN uses neutral density and temperature profiles derived from ECMWF analyses. The Rayleigh scattering cross-sections are an implementation of the formulae given by *Bates* (1984), or alternatively by the approximation given in Equation 4.36. Stratospheric aerosols are described by a log-normal distribution of hydrated sulphate particles with user configurable size distribution parameters as a function of altitude. Aerosol scattering phase functions are calculated from the Mie scattering code developed by *Wiscombe* (1980), appropriately integrated over the particle size distribution. Default temperature dependent O_3 and NO_2 cross-sections are taken from *Bogumil et al.* (2000). The code is structured so that it is easily extended to include other atmospheric species such as other aerosol types and absorbing species such as water vapour, molecular oxygen and carbon dioxide.

A Note on Horizontal Homogeneity

The assumption of horizontal homogeneity implies one degree of symmetry in the radiative source terms associated with a particular spherical grid cell. The variation in the source term depends only on the solar zenith angle and not on the absolute sun direction. In a spherical geometry with horizontal homogeneity, there is no variation in the atmospheric column density along the path from the sun to two points in the same cell with the same solar zenith angle. In a coordinate system where the Sun-Earth vector is placed along the $\hat{z}$-axis, the radiative solution has "longitudinal" symmetry and only depends upon "latitude", *i.e.* solar zenith angle, and height, see for example *Lenoble and Sekera* (1961). Previous work has clearly shown that the variation of the multiple scatter source term with solar zenith angle at a given altitude is smooth (*Herman et al.*, 1994; *Oikarinen et al.*, 1999; *Griffioen and Oikarinen*, 2000; *McLinden et al.*, 2002). Intuitively, this represents light diffusing through the atmosphere via scattering processes and smoothing any sharp features present in the original illumination.

SASKTRAN utilizes this feature to reduce the number of points where a calculation of the multiple scatter source terms is required. In a simple formulation, a model would calculate the multiple scatter source terms at each path length segment along a line of sight. However, this is very computationally intensive. It is far more efficient to solve for the multiple scatter source terms that sensibly span the range of solar zenith angles along the line of sight and use the "longitudinal" symmetry implicit to the formulation as well as the assumed smooth variation in solar zenith angle to interpolate the source term at the required points. This reduction in dimension and the required numerical resolution in solar zenith angle is addressed in detail in Section 4.5.1.

Path Lengths Through Cells

The numerical approximation to the line integrals for the calculation of the optical depth and for the radiance along a line of sight require a step size, *i.e.* $\mathrm{d}s \simeq \Delta s$. This is accomplished in SASKTRAN by using the path lengths of the ray through the spherical cells. Δs is always taken as the distance between the intersections of the ray with the shells bounding the cell. For the entire path length through the cell the optical properties of the atmosphere are assumed to be contstant, *i.e.* the value stored for the centre of the cell. Thus the grid cell thickness has an effect on the accuracy of the quadrature. However, decreasing the cell thickness from the 1000 m resolution results used in the present work by a factor of ten changes the computed radiance by less than two percent, depending upon wavelength and solar geometry. This is discussed further in Section 4.7.1.

The path lengths must be calculated in a true three dimensional geometry. The actual lengths could be calculated in the two dimensional plane of the line of sight and the spherical zenith, but the absolute position of each path length element in three-dimensional space is required to account for the variation in solar geometry along the ray.

Calculation of Optical Depth

The optical depth defined by Equation 4.18 is integrated numerically by,

$$\tau(s_2, s_1) = \int_{s_2}^{s_1} k(s)\,\mathrm{d}s \simeq \sum_{i=0}^{n} k_i\,\Delta s_i\,, \tag{4.61}$$

where the line of sight from s_2 to s_1 intersects n grid cells. As discussed above, the path lengths Δs_i are determined as the intersections of the line of sight with the spherical cells that define the atmospheric grid. The extinction k_i is that corresponding to the centre of the associated cell. If the line of sight intersects the same cell twice, Δs_i is the total path length through that cell.

The assumption of horizontal homogeneity is adopted to improve computational efficiency in the implementation of the model. Similar paths from cell to cell are identified and the optical depth is calculated only once. For example, the path from a point in the atmosphere to the ground at a particular zenith angle has the same optical depth for all azimuth angles. Thus the calculation of the optical depth, and more importantly in terms of computation time, the calculation of the transmission, $e^{-\tau}$, is only performed for one ray in this set.

Calculation of Radiance

All radiance calculations, for example Equation 4.58, require an integration along a line of sight, and are performed numerically in the same fashion as the calculation of optical depth. The value of the source function at each path length is assumed to be the value of the source function at the radial centre of the cell.

There is, however, an important distinction between the implementation of the radiance integral and the optical depth calculation when the ray intersects the same cell twice. For the radiance, each contribution must be handled separately as the transmission to the reference point, $e^{-\tau(s)}$, is different for the two segments. This is discussed in detail in later sections; however, it is important to note that the source terms, $J_i(s, \hat{\Omega})$, are computed for a discrete resolution in solid angle. A linear interpolation of the spatial dependence to the required direction is required to determine the source term at each point along the ray.

Single Scatter Source Terms

The source term for one scattering of the solar beam, Equation 4.52, is calculated at each required point along a line of sight. It is implemented by tracing rays from the sun to the centre of each path length segment along the line of sight. The optical depth is calculated along each of these rays to determine the attenuation of the solar flux density from the top of the atmosphere down to each scattering point where it is

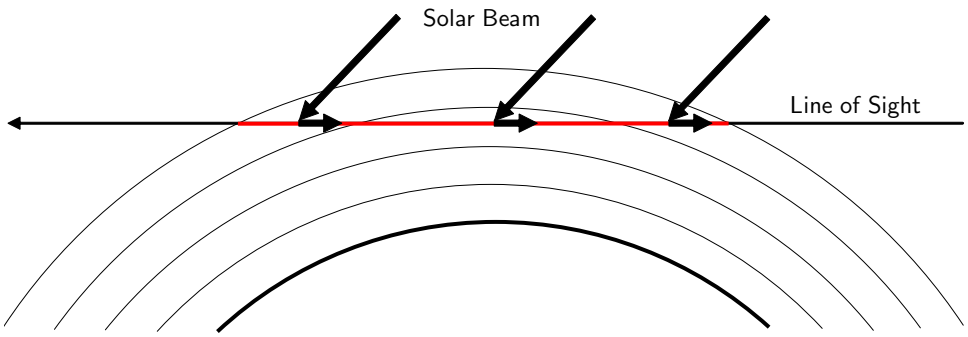

Figure 4.9: A single scattering event for the solar beam that directly radiates into the instrument line of sight at each cell. The solar beam is attenuated from the top of the atmosphere to the scattering point. The scattered radiance is attenuated along the line of sight back to the instrument.

scattered into the line of sight according to Equation 4.52 using the three dimensional solar geometry. Figure 4.9 illustrates a two-dimensional cross section of the single scattering geometry for a limb viewing instrument.

If required, the single scatter source from the ground at the end of a line of sight, Equation 4.51, is also determined from the three dimensional solar geometry without approximation. This procedure, for determining the radiance due to a single scatter of the solar beam, is identical for all lines of sight whether it is in an intermediate calculation along a ray that is required to generate the next order of scattering, or directly for the instrument line of sight.

All Multiple Scatter Source Terms

Unlike the single scatter source term, which is evaluated directly at each scattering point along every ray in the SASKTRAN model, the multiple scatter source terms are calculated for a range of solar zenith angles by exploiting the symmetry (mentioned previously) that arises from the assumption of horizontal homogeneity. The source term is evaluated at each of these points as a function of altitude. One vertical profile of the source term, $J_i(h, \hat{\Omega})$, is calculated for each order of scatter, i, greater than one; again, h represents the altitude above the osculating sphere. Similarly, the multiple scattered radiance from the Earth's surface is also determined for the zero altitude

point of each source term vertical profile. The solar zenith angle range is determined from the instrument line of sight. For example, if the solar zenith angle varies from 82 to 85° along the instrument line of sight, the vertical profile of the source term is calculated at solar zenith angles 82.5, 83.5, and 84.5°. This angular resolution is variable.

At each cell, the source term is a two-dimensional field where a set of discrete local zenith and azimuth angles are used to specify the spatial dependence over all solid angles. Each vertical profile of the multiple scatter source term is calculated independently and there is no attempt to couple the vertical profiles. This independence does not affect the solution for two scattering events, but is a limitation for the third and higher order multiple scatter source terms.

Source Term From Two Scattering Events

The source term due to two scattering events of the solar beam is calculated without approximation for the three-dimensional geometry. Consider the calculation of the vertical profile of the source term at a specific solar zenith angle. In order to determine the spatial dependence of $J_2(s, \hat{\Omega})$ at a single point, s, along this vertical direction, rays are traced from s for a distribution of zenith and azimuth angles to estimate the spatial integration (Equation 4.54). The distribution of these rays is important for the accuracy of the integration over solid angle; this is discussed in detail in Section 4.5.2. The radiance along each ray from a single scattering of the solar beam is calculated as discussed above (Equation 4.53). That is, the three-dimensional geometry is traced from s along the ray direction, and then from all scattering points along the ray, including the ground if required, to the sun (see Figure 4.9). This radiance is then scattered into a second set of discrete directions that specify the spatial dependence of the source term. It is important to recognize that there are two different angular resolutions in this calculation: the resolution used to perform the integration of the incoming radiance over solid angle, and the resolution that

specifies the angular distribution of the source term.

A two dimensional cross section of the scattering geometry with a sample distribution of rays in the zenith direction is shown in Figure 4.10. The two scatter geometry is further illustrated in Figure 4.11. The straight red line indicates the intersection of a limb viewing instrument line of sight with the bottom 100 km of atmosphere. The tangent altitude is 25 km. The black lines denote the solar beam before a scattering event, the straight green lines represent light that has been scattered once from an atmospheric constituent, and the straight blue lines follow the path of the light after it has been scattered once from a ground element. For the scattering volume element at 25 km altitude, the blue circle encompasses all rays from the element that strike the ground. The green circle is the exit point of all rays from the element that leave the top of the atmosphere. All rays that are traced to determine the second order source term at this location fall within these bounding circles. Each scattering event along each of these rays is determined in the three dimensional spherical geometry.

Source Term From Three or More Scattering Events

The calculation of the source term from three or more scattering events is similar to the calculation of the second order term in that it involves the spatial integral of the radiance from the previous order of scatter. However, an important approximation is made for these higher order terms in that the SASKTRAN model does not trace the exact light path of all scattering events through the three-dimensional grid. The approximation is made in the selection of the previous order source term at each point along the ray paths that are traced to estimate the distribution of the radiance over solid angle.

For example, consider the calculation of the third order source term at a given altitude for a particular solar zenith angle. Rays, distributed in solid angle, are traced from this location; the required radiance in every direction is the line integral of the second order source term (Equations 4.59 and 4.56). It should be noted that the

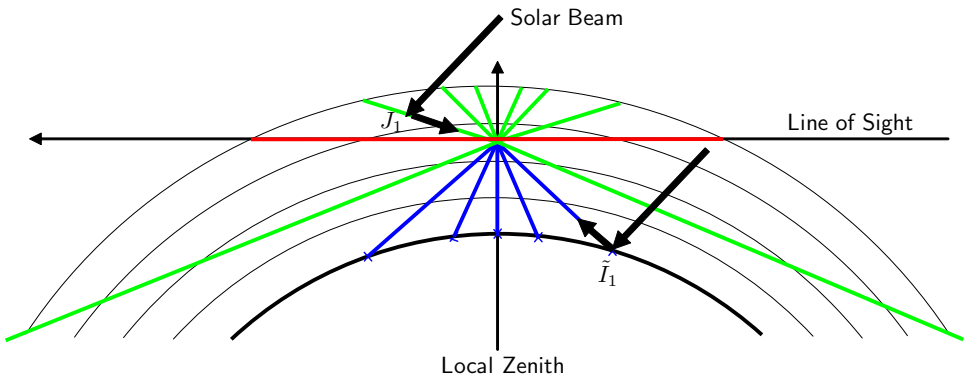

Figure 4.10: The calculation of the second order source term $J_2(s, \hat{\Omega})$. Rays are traced at an angular resolution over the unit sphere; blue rays strike the ground, green rays exit the top of the atmosphere. The incoming radiance due to a single scattering event of the solar beam, from the atmosphere, $J_1(s, \hat{\Omega})$, and from the ground, $\tilde{I}_1(s_1)$, is calculated along each ray, scattered and integrated over solid angle.

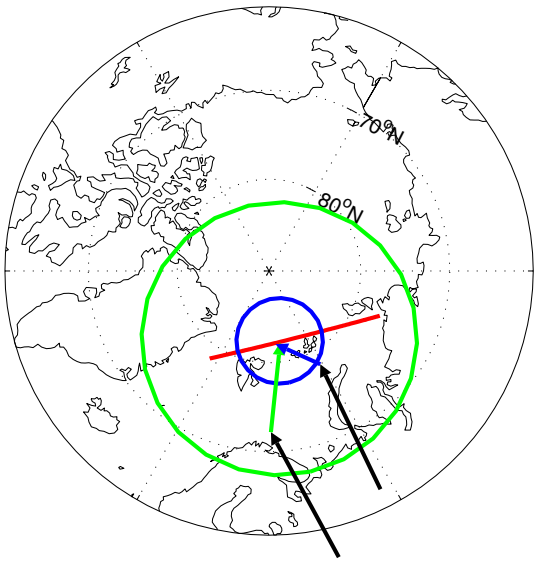

Figure 4.11: From a scattering volume element at 25 km altitude, the blue circle encompasses all rays distributed in solid angle from the element that strike the ground. The green circle is the exit point of all rays from the element that reach the top of the atmosphere. The solar geometry for all scattering within these regions is handled in a fully 3D sense. The black/green arrow path represents an atmospheric scattering of the solar beam into a ray direction; black/blue represents a Lambertian scattering of the solar beam into an ray direction that strikes the ground.

second order source term is not known in all three dimensional space, rather only at the selected solar zenith angle. Thus, the approximation in SASKTRAN is to simply use the second order source term that corresponds to the solar zenith angle at the point where the ray originated. Similarly, for the fourth order profile where third order source terms are integrated, and so on. Each time the vertical profile of the source term is required, it is geometrically transformed from the ray origin to the required location in order to maintain the proper scattering angle, altitude, and attenuation back to the origin of the ray.

To reiterate, the procedure used to calculate the vertical profile of the third, or higher order, source term at a particular solar zenith angle is:

- The vertical profile of the second order source term is determined by tracing rays through two scattering events in the full three dimensional geometry.
- The vertical profile of the third order source term profile is calculated using Equations 4.58, 4.59 and 4.60. At all points where $J_2(h, \hat{\Omega})$ is required, it is assumed that the second order source term profile evaluated at the origin of the ray applies to the current point along the ray.
- The fourth order source term vertical profile is calculated in the same fashion from the third order profiles that have been calculated.

This iterative procedure is extended in the same way to determine any arbitrary order of scattering. The ray tracing for the third order source term is shown in Figures 4.12 and 4.13. For most wavelengths and observer geometries the adopted approximation has only minimal effect. However, the impact of the approximation is addressed in Section 4.4.

The Radiance Along the Instrument Line of Sight

After all source terms for a particular set of solar zenith angles along the instrument line of sight are determined, it is simple matter to make the final numerical line integral that calculates the solution of the modelled radiance as seen by the instrument.

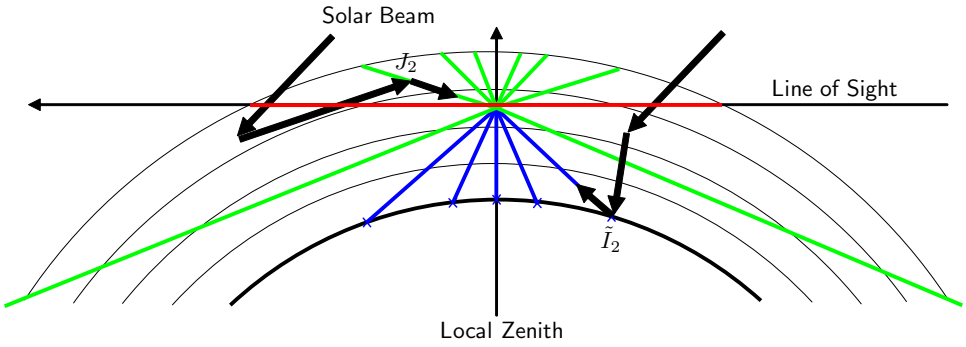

Figure 4.12: The calculation of the source term due to the second order of multiple scattering, $J_3(s, \hat{\Omega})$. Similar rays tracing is performed over the unit sphere. The incoming radiance due to two scattering events of the solar beam is calculated along each ray by a rotation of $J_2(s, \hat{\Omega})$ from the local zenith along the ray path. The reflection of the diffuse radiance field from the ground is indicated as $\tilde{I}_2(s_1)$ (see Equation 4.57).

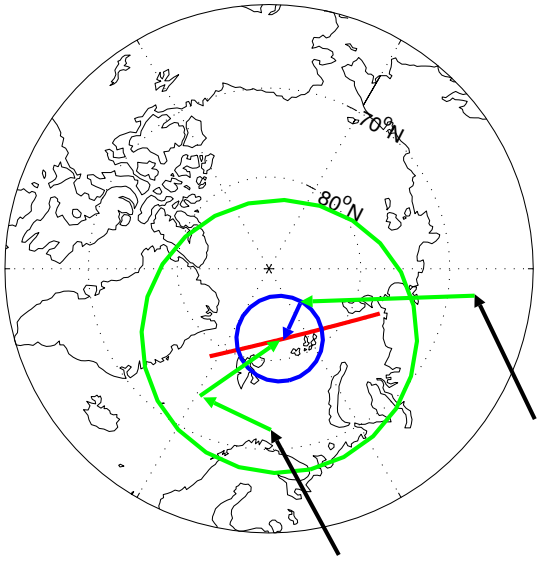

Figure 4.13: The same polar map shown in Figure 4.11. The black/green/green arrow path represents two atmospheric scattering events of the solar beam required for the calculation of $J_3(s, \hat{\Omega})$. The black/green/blue ray path represents one scatter from the atmosphere and the second from the ground into the integration ray direction. Conceptually, as shown here, the first scattering could be well outside the region used to calculate $J_2(s, \hat{\Omega})$ at the local zenith. However, $J_2(s, \hat{\Omega})$ is used as the second order source term everywhere.

The single scatter contribution to the observed radiance is determined in the same way as for any other ray. For higher orders of scattering, the solar zenith angle, the altitude, and the scattering direction along the instrument line of sight determine the source functions, $J_i(h, \hat{\Omega})$, for each $i \geq 2$. These functions are linearly interpolated in solar zenith angle and in solid angle at each cell intersection.

4.4 Importance of Multiple Scattering

The significance of multiple scattering to measurements of scattered sunlight in limb geometry is shown in Figure 4.14. The presented model results are for an OSIRIS limb viewing geometry with a 25 km tangent altitude and a scene albedo of 0.8. The limb radiance was modelled with SASKTRAN for all OSIRIS wavelengths (280-810 nm) with 10 orders of multiple scattering. Separate calculations with the model show that further orders of scattering beyond the tenth contribute less than 1% to the total radiance at all wavelengths. Each curve in the plot is the fraction of the total limb radiance that is due to the corresponding order of scattering. It is apparent from Figure 4.14 that at the shortest wavelengths, where the Hartley-Huggins O_3 band absorption is significant, the total limb signal is well represented by the single scatter calculation. At longer wavelengths the contribution from multiply scattered light is approximately half of the total signal. It is important to note the size of the contribution from the second order of scattering compared to that from successive orders. In the red end of the spectrum, the second order scatter, J_2, which occurs from either the ground or the atmosphere, is almost as important as the single scatter term; however, the contributions from higher orders are much smaller than from the first two. Thus the SASKTRAN algorithm, which approximates the third and higher orders, is appropriate for the estimation of the multiple scatter contribution. The first and second order scattering events, which are modelled without approximation in the three dimensional spherical geometry, are much more important than successive

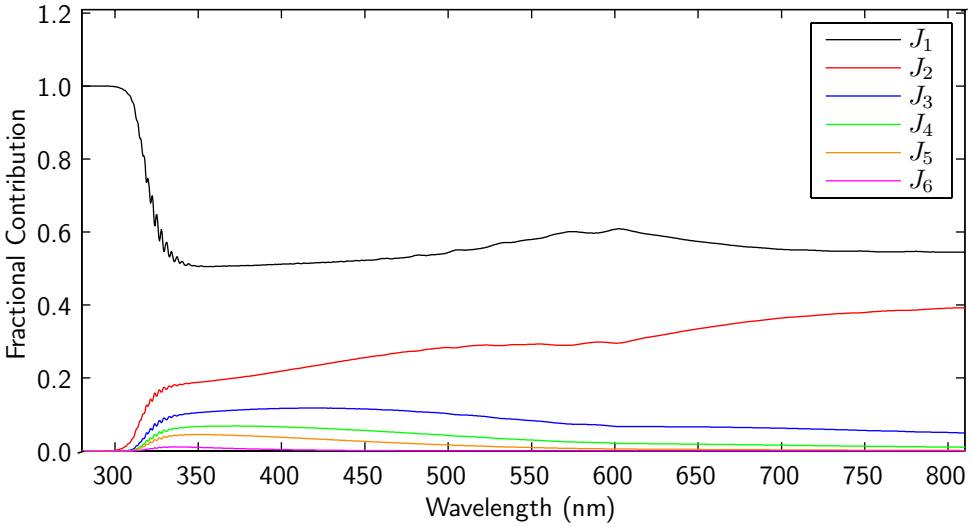

Figure 4.14: The fractional contribution of successive orders of scattering to the total limb signal at 25 km tangent altitude for all OSIRIS wavelengths. The total signal is calculated with 10 orders of multiple scattering. Each curve represents the fraction of the total signal that is contributed after n scattering events.

higher orders that are approximated in the model by integrating the previous order term at the ray origin.

The effects of this approximation can be shown by considering the condition where the OSIRIS line of sight is directed along the solar terminator. In this case there is no variation of solar zenith angle along the entire line of sight, and the Earth's surface below the tangent point is illuminated on one side of the line of sight and in shadow on the other side. In the ray tracing used to determine the vertical profile of the second order source term, J_2, the first scatter of the solar beam towards the source term element is calculated according the true solar geometry at each point. On the illuminated side of the line of sight, successive scattering orders are calculated by integrating the previous order source term from along the ray origin, *i.e.* at the terminator. The result is therefore systematically too low. The opposite is true on the shadowed side of the line of sight where the source term is systematically too high for orders beyond the second scatter. The opposite nature of these systematics introduces a partial cancelation of the error. The limb geometry dictates that this

will always be the case; if the source term is systematically high in one direction from the line of sight, it will be systematically low in the other direction.

Figure 4.14 shows that the limb radiance is dominated by first and second order scattering. Because the contribution from the third scattering order is less than approximately 15% even for high albedo, and because of this systematic compensation by under and over estimation in the higher orders, the error in the total radiance caused by this approximation is less than a few percent. A comparison of SASKTRAN with other radiative transfer models, including fully spherical radiative transfer codes, also supports this conclusion (see Section 4.7.1).

4.5 Computational Accuracy

The greatest trade-off in radiative transfer modelling of visible light in the atmosphere is between computational accuracy and speed. Accuracy in this problem is most often with regard to the integrations required for solution of the radiative transfer equation that are approximated in the model by discrete summations. Simply increasing the resolution of the discretization to a very high degree will cause the integrations to converge to the solution. However, for a successive orders ray tracing model such as SASKTRAN, increasing the resolution in any dimension incurs heavy execution time penalties. In SASKTRAN, two approximations that are specifically important for the solution of the radiative transfer equation in limb geometry are addressed in particular:

- the variation in the source term with solar geometry along the relatively long path length of the line of sight,
- the convergence of the source term integration over solid angle.

4.5.1 Variation in Solar Source Term

As noted in Section 4.3.2, the assumption of homogeneous spherical layers introduces a symmetry in the geometry of the radiative transfer problem. As shown in Figure 4.15, the $\hat{z}$-axis of a convenient geocentric coordinate system is placed along the solar direction and, because the atmospheric cells are homogeneous, the resulting radiation field is symmetric in the azimuthal coordinate direction. Thus the five dimensional solution of the radiative transfer problem is decreased to a four dimensional problem, *i.e.* only radial and zenith coordinates are required to describe the distribution of the directional radiation field.

While the single scatter component of the source term is determined at every cell intersection, the variation in the multiple scatter component of the source term is addressed in the SASKTRAN model by exploiting the symmetry of the radiation field about the solar beam. The vertical profile of the source term is determined for a set of solar zenith angles, θ_{sza}, that correspond to segments of the instrument line of sight. Figure 4.15 shows the intersection of a line of sight with the modelled atmosphere in red and a set of three locations along the line of sight where the vertical profile of the multiply scattered source term is calculated. Each of these locations corresponds to a solar zenith angle. The red circles represent a geographical area where the geometry of the scattering is solved without approximation for the first two orders. An easily understood case that demonstrates this symmetry about the solar beam is when the instrument line of sight is contained in a horizonal plane of the geocentric coordinate system shown in Figure 4.15. In this case the vertical profile of the source term does not vary along the line of sight and, therefore, only one estimate of the source term profile is required.

Thus in SASKTRAN, at least one profile of the multiple scatter source terms is always calculated. This corresponds to the zenith direction in the osculating sphere coordinates. For limb geometry this zenith direction is directed at the tangent point of the observation and is representative of the most important solar conditions along

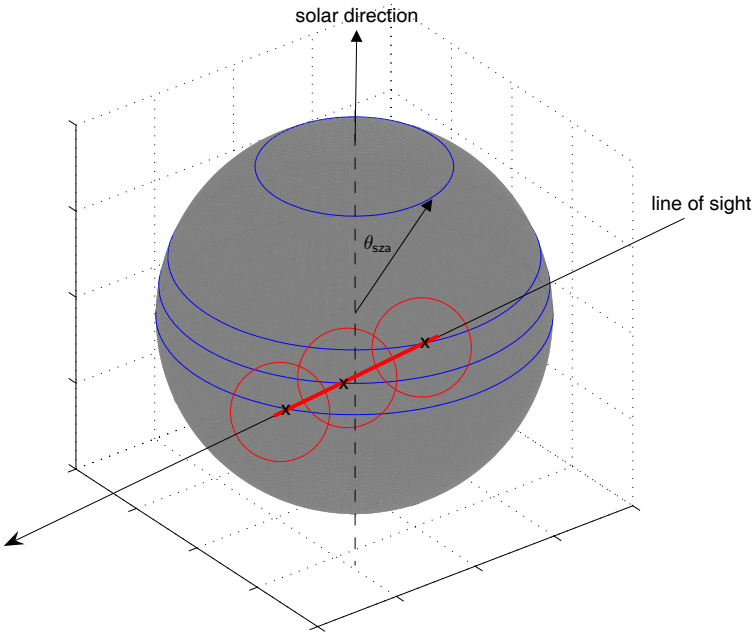

Figure 4.15: A spherical geocentric coordinate system with the global $\hat{z}$-axis placed along the solar direction. In a spherical atmosphere with homogeneous layers, the resulting radiation field is symmetric in global azimuth.

the line of sight due to the long path length at the tangent point. When the source term is calculated for additional solar zenith angles, a linear interpolation between the two source term solutions that are closest in solar zenith angle to each cell intersection point is performed during the integration of the source terms along the line of sight.

Tests with the model have shown that for tangent point solar zenith angles less than 70° the limb radiance calculated using one multiple scatter source term profile at the tangent point zenith angle is accurate to within 1%. However, when the sun is low on the horizon the instrument line of sight can span the solar terminator, i.e. cross from the sunlit part of the orbit into darkness or vice versa, and result in a variation in the multiple scatter source term that is significant. OSIRIS measurements near equatorial latitudes typically occur at solar zenith angles near 90° (see Figure 2.3). Figure 4.16 shows the result of model calculations for a typical OSIRIS equatorial

geometry. In this particular case the solar zenith angle along the line of sight is approximately 85° on the near side of the tangent point and 94° on the far side of the tangent point. In the set of limb radiance profiles shown in blue, the source term is only calculated at a solar zenith angle of 90°. In the other set, shown in red, the source term is calculated at 8 locations that correspond to solar zenith angles along the line of sight. The separation between each profile is approximately 1° in solar zenith angle. The plots compare the limb radiance profile at four wavelengths: 300, 450, 600 and 750 nm. At 300 nm, the additional source term calculations have no impact on the modelled radiance. This is because at this wavelength and altitude the single scatter radiance is the dominant term in the calculation (see Figure 4.14). At longer wavelengths the difference becomes significant at lower tangent altitudes. However, even for this extreme limb geometry, the error caused by estimating the multiple scatter source term at a single solar zenith angle is less than 10% and is limited to the lowermost tangent altitudes.

4.5.2 Solid Angle Resolution

The source term in this radiative transfer problem for all orders of scattering beyond the first scatter is the integral of the diffuse radiance field over the unit sphere given by Equation 4.23. This equation is again included here for reference,

$$J(s, \hat{\Omega}) = \frac{k_{\text{scat}}(s)}{k(s)} \int_{4\pi} I(\vec{r}, \hat{\Omega}')\, \bar{\text{p}}(s, \Theta)\, \text{d}\Omega' \,. \tag{4.22}$$

In previous work the numerical solution of this integral has been frequently approached by representing the integral over solid angle with a double integration over the spherical directional coordinates, θ' and ϕ',

$$J(s, \theta, \phi) = \frac{k_{\text{scat}}(s)}{k(s)} \int_{2\pi} \int_{\pi} I(s, \theta', \phi') \bar{\text{p}}(s, \Theta)\, \sin(\theta') \text{d}\theta' \text{d}\phi' \,. \tag{4.62}$$

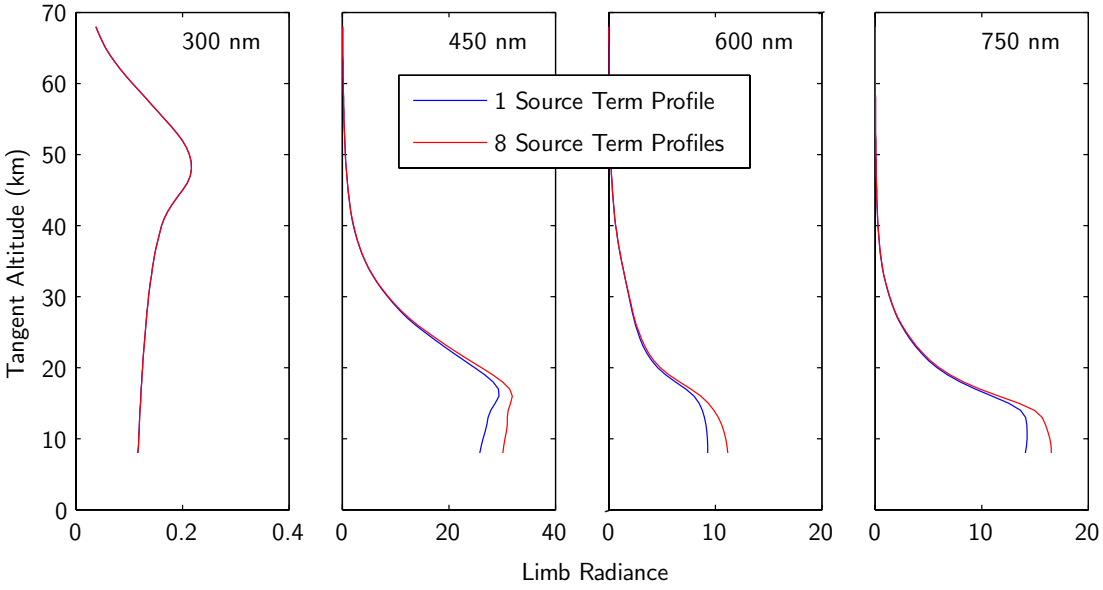

Figure 4.16: SASKTRAN limb radiance profiles (10^{12} photons/cm^2/s/nm/sterad) for an equatorial OSIRIS geometry with a large range of solar zenith angle along the instrument line of sight. The difference between the estimated multiple scatter contribution calculated only at the tangent point compared to every 1° of solar zenith angle is approximately 10% at the longer wavelengths and lowest tangent altitudes.

It is important to recognize that two physically different angular resolutions are described in Equation 4.62. The source term is calculated for the set of angles specified by θ and ϕ. For each θ and ϕ, the integration is performed over the diffuse radiance field for the resolution specified by the set of angles θ' and ϕ'. It is straightforward to apply simple numerical techniques, such as trapezoidal or Gaussian quadrature, with respect to each angular coordinate. The number of points considered in the summation is certainly important with respect to the accuracy of the solution; however, as the greatest computational cost associated with the solution of the radiative transfer equation is incurred in the estimation of the source term, it is desirable to minimize the number of points considered. For each point of the numerical integration in a ray tracing model, such as SASKTRAN, the diffuse radiance in the corresponding direction must be calculated. This involves the line integral of the source term, from the previous order of scattering, from the point of interest to the ground or the top

of the atmosphere.

In the plane parallel solution, a relatively small number of divisions in θ', typically less than 10, are used to perform the integration in the zenith direction (*Griffioen and Oikarinen*, 2000). Due to the simple geometry, it is advantageous to calculate the source term at an identical angular resolution to the quadrature resolution. For limb geometry the integration of the source function along the instrument line of the sight requires knowledge of $J(s, \theta, \phi)$ at zenith angles very near to $\theta = 90°$, typically from approximately 80° on the near side of the tangent point at the top of the atmosphere to 100° on the far side of the tangent point. Since a small number of divisions in the zenith direction provides only very few points in this range a simple interpolation of $J(\boldsymbol{s}, \theta, \phi)$ is used for the required angles in the plane parallel solution. The zenith angle interpolation is also required because, in this one dimensional problem, the solution is not defined in the horizonal direction by nature of the geometry (*McLinden et al.*, 2002). It should be noted that the azimuthal dependence in the plane parallel solution is considered in terms of a Fourier expansion of the diffuse field and the scattering phase function. The orthogonality of the components results in a simplified azimuthally-independent form of the radiative transfer equation.

In a spherical geometry, the azimuth dependence must be estimated explicitly. The appropriate placing of the divisions in the numerical integral is key to the accuracy of the solution. A uniform division of the zenith and azimuth coordinates results in an high density of rays near the poles of the unit sphere. *Steinacker et al.* (1996) have addressed this undesirable consequence. They do not consider separate divisions of the zenith and azimuth coordinates. Strictly in terms of solid angle, the numerical estimation of the source function is

$$J(s, \hat{\Omega}) = \frac{k_{\text{scat}}(s)}{k(s)} \sum_{4\pi} I(s, \boldsymbol{\Omega'}) \bar{\text{p}}(s, \Theta) \, \Delta\Omega' \,. \tag{4.63}$$

Steinacker et al. propose a set of integration points on the unit sphere in order to

construct an "optimal" distribution of rays such that the solid angle weight, $\Delta\Omega$, for each division is approximately $4\pi/N$ for N rays. This distribution of rays is used by *Rozanov et al.* (2001) in the spherical CDI-PI model with typically N=100 integration nodes on the unit sphere.

This method of uniform distribution in solid angle is generally appropriate and is a good choice for a diffuse field with an unknown spatial distribution or in an optically thick medium. However, with knowledge of the distribution of the diffuse intensity of scattered sunlight in the atmosphere, a distribution of rays can be developed that is directly suited to the situation. Figure 4.17 is a plot of the first order diffuse radiance, $I_1(s, \theta, \phi)$, at 750 nm and 25 km altitude for an albedo of 0.4. These radiances arise from a single scattering event of sunlight within the atmosphere and/or the ground, and are required for the calculation of the second order of source term, J_2 (see Equation 4.54). For this plot, the field is calculated at sub-degree angular resolution in both zenith and azimuth coordinates. The left panel of the plot shows the radiance as a two dimensional field that is a function of zenith angle and azimuth angle. The zenith coordinate is measured from the local vertical direction and the azimuth coordinate is fixed to the sun, *i.e.* $\phi = 0$ is in the plane of the solar vector and local zenith direction. The right panel of Figure 4.17 shows the same data on the surface of a unit sphere in an attempt to provide clarity for the geometry.

It is clear that there is a very significant change in the radiance field near, and just below, the local horizon, $\theta = 90^\circ$. For all directions that look upwards from this location, the radiance is small and approximately constant. This shows that the contribution to multiple scattering from upper altitudes is small. For downward looking directions, the radiance is much larger, but it is quite constant with respect to the variation at the horizon. The large radiances are due to the contribution from scattering from the thick atmosphere below and from the Lambertian reflection of sunlight from the ground. It is for those directions that are within the angular range between the horizon and the tangent direction to the Earth, typically about

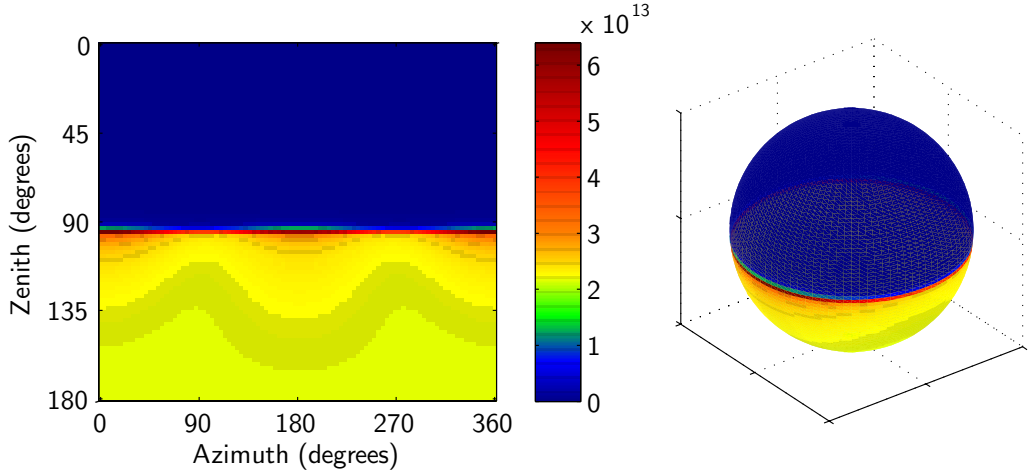

Figure 4.17: The diffuse radiance field at 750 nm (photons/s/cm^2/nm/sterad) observed in all directions from 25 km altitude due to the first scattering event of the solar beam.

$\theta = 90 - 95°$, where the large change occurs. The azimuth dependence of the radiance at every zenith angle is smooth and reflects the fact that the Rayleigh scattering phase function is maximum in the forward and backward direction (see Equation 4.34). Since $\phi = 0$ is in the plane of the solar vector the forward scattering direction is also in this plane.

The source term, J_2, that arises from the scattering of the diffuse radiance field, I_1, shown in Figure 4.17, calculated with a discretization of Equation 4.62, is shown in Figure 4.18. This calculation is made with a very high angular resolution for the integration. For each direction of $J_2(s, \theta, \phi)$, the integration of the diffuse radiance, $I_1(s, \theta', \phi')$, was performed for 720 equally spaced rays in zenith angle (θ') and 24 equally spaced rays in azimuth angle (ϕ'). $J_2(s, \theta, \phi)$ is determined for 24 equally spaced directions in zenith angle (θ) and azimuth angle (ϕ). While these integration nodes suffer the problem of high ray density near the poles of the unit sphere the solution is fully converged at the adopted high resolution. If the resolution is increased further, there is no appreciable change in the solution. The diffuse source term $J_2(s, \theta, \phi)$ shown in Figure 4.18 is represented as a three dimensional surface because both the shape and absolute magnitude of the result are important. It should be

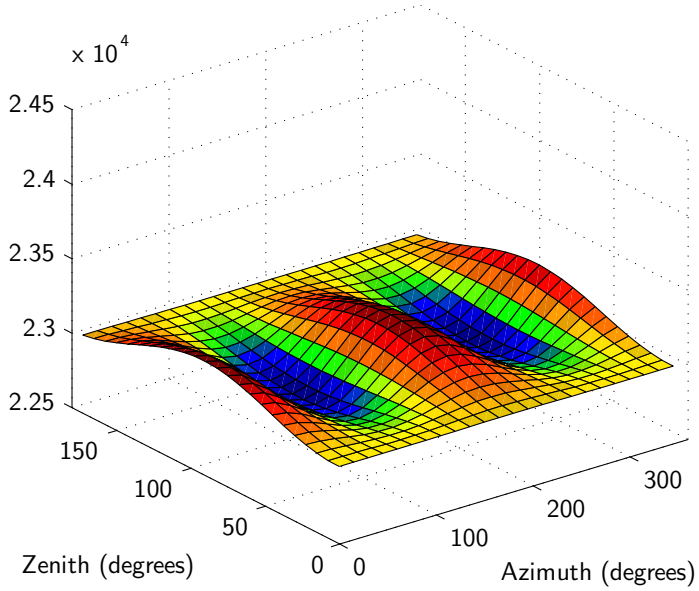

Figure 4.18: $J_2(s,\theta,\phi)k(s)$ (photons/s/cm^3/nm/sterad) at 750 nm and 25 km altitude calculated by discretization of Equation 4.62 using 17280 rays to represent the incoming diffuse radiance field: 720 rays uniformly spaced in zenith, 24 rays uniformly spaced in azimuth. $J_2(s,\theta,\phi)$ is calculated at a uniformly spaced angular resolution with 24 rays in the zenith direction and 24 rays in the azimuth direction.

noted that the source term is smooth in both directions so that simple interpolation methods give good estimates for required directions in the evaluation of successive orders that are not explicitly evaluated.

A high resolution calculation, such as that shown in Figure 4.18, requires a large computational resource and is impractical for use in a retrieval. The resolution must be decreased without compromising the accuracy of the solution. Figure 4.19 shows the result of the same calculation with only 24 equally spaced rays in both zenith and azimuth to represent the incoming diffuse radiance field. While the general shape of the angular distribution of the result is similar to that shown in Figure 4.18, the magnitude is systematically higher by approximately 5%.

To correct for this systematic error, the solid angle integration is implemented in the SASKTRAN model with a distribution of 24 rays in zenith that are concentrated near, and just below, the local horizonal in order to capture the sharp change in incoming radiance in these directions. The same number of rays are distributed

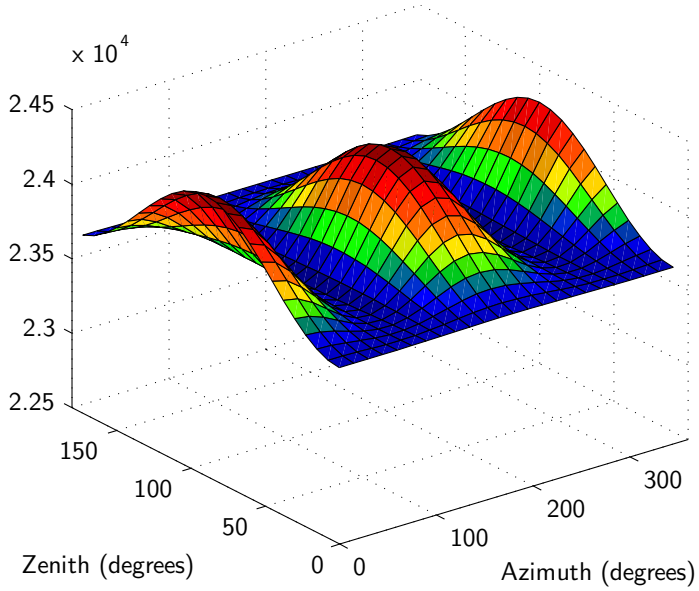

Figure 4.19: The source term distribution as shown in Figure 4.18 except calculated using only 24 rays uniformly spaced in zenith. The entire field is systematically high in comparison to the high resolution calculation presented in Figure 4.18.

equally in azimuth angle. Figure 4.20 shows a plot of the same calculation of $J_2(s, \theta, \phi)$ that is shown in Figures 4.18 and 4.19 but with the concentrated ray spacing of incoming radiance in the zenith angle coordinate. The result compares well with the high resolution solution (Figures 4.18) both in shape and in magnitude. The differences are less than 1% in all directions. These results are typical and similar results are obtained at other wavelengths and altitudes. Tests with the model have shown that with less than 24 rays distributed in zenith angle, even if carefully spaced near the horizon, the accuracy of the solution is reduced.

As previously noted, the line of sight integration of the total source term requires the evaluation of the source term at zenith angles in the range 80–100°. Like other models, a linear interpolation of the angular distribution of the source term is used to perform this line integral for those directions at cell intersections that are not explicitly calculated.

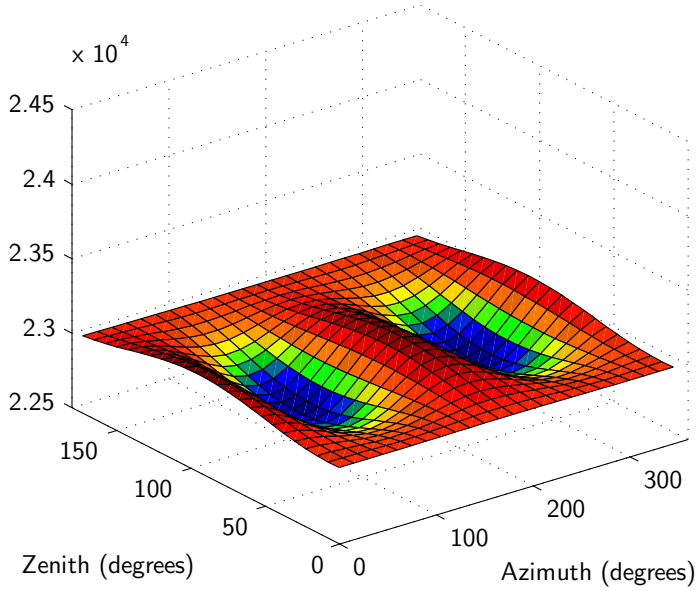

Figure 4.20: The source term distribution as shown in Figure 4.18 except calculated using only 24 rays in the zenith direction with a high density of rays just below the local horizon. The result compares very well with the high resolution calculation shown in Figure 4.18.

4.6 Model Organization and Optimization

SASKTRAN is written in C/C++ and makes extensive use of object-oriented software design. At the base level, there are two interacting, but separate code modules: the initialization and the radiative transfer engine. The engine module requires only four basic inputs:

- the observer position and look direction,
- the solar direction,
- the total wavelength dependent extinction, $k(h, \lambda)$, for each cell altitude, and
- the wavelength dependent directional scatter, $k_{\text{scat}}(h, \lambda)\bar{\text{p}}(h, \Theta, \lambda)$, for each cell altitude.

The output of the engine is the modelled limb radiance in the observer look direction for unit solar flux incident on the top of the atmosphere. The initialization module provides a user interface to the engine such that the optical properties required for the radiative transfer are specified by the atmospheric species density profiles, either

from interfaces to standard climatologies or from profiles specified by the user. The initialization also provides a specification for an arbitrary limb geometry and a direct interface to the Odin/OSIRIS attitude solution database.

To take advantage of modern desktop computing advancements, extensive use is made of RAM by pre-calculating and caching the engine inputs and storing the different orders of scattering in the source term estimation. For example, the wavelength dependent directional scatter, $k_{\text{scat}}(h,\lambda)\bar{\text{p}}(h,\Theta,\lambda)$, is a three dimensional field: wavelength, cell, and scattering angle. In the initialization stage, the array is calculated and cached for all required wavelengths, at the required cell resolution and for sub-degree (variable) scattering angle resolution. The total radiative transfer calculation for a single wavelength with n-orders of multiple scattering requires approximately 1 MB of memory. No additional memory is required for the calculation of orders of scattering beyond the second scatter. Thus the calculation of the entire OSIRIS spectrum (1352 wavelengths) to a converged result requires less than 2 GB of RAM. In order to take advantage of recent multi-core processors, the radiative transfer engine module is multi-threaded in the wavelength dimension. A converged limb radiance profile at a single wavelength, with typically more than 5 orders of multiple scattering, requires approximately 10 s of execution time on a 3.0 GHz Pentium D desktop computer. Because of the multi-threading in wavelength, the overhead required for calculation of a second wavelength is less than 0.5 s on the dual core Pentium D. Calculations with fewer orders of multiple scattering reduces execution time approximately linearly with decreasing number of orders. The currently available version of the code is compiled for Windows XP, but future plans include portability to other platforms.

4.7 Comparisons

4.7.1 Other Radiative Transfer Models

In a recent paper that has compared radiative transfer models that are suitable for limb scatter, *Loughman et al.* (2004) have identified three broad categories of model based on technique: Monte-Carlo, spherical, and approximate spherical. Six models, two of each type, were compared for a range of solar geometries, albedo conditions and aerosol load. All of the models in the study use homogenous spherical shells and calculate the single scatter contribution to the limb radiance using a spherical algorithm. In addition, all of the models attempt a full treatment of the multiple scattering component from the atmosphere and from the surface. However, it is in the multiple scatter calculation where there are important differences between the algorithms.

The Monte-Carlo models, Siro (*Oikarinen et al.*, 1999) and MCC++ (*Postylyakov*, 2004a), use a backward method of counting statistics to follow a single photon from the detector on a path through the spherical atmosphere determined by probabilistic scattering events. The first of the two spherical models, Guass Seidel Limb Scatter (GSLS) (*Loughman et al.*, 2004), is a limb scatter adaptation of the Guass Seidel Spherical (GSS) model (*Herman et al.*, 1994, 1995). The calculation of the diffuse field is performed in a so-called "pseudo-spherical" sense where the solar beam is attenuated to each spherical shell and is used as the initialization for a standard plane-parallel multiple scattering calculation. This calculation is only made for the solar zenith angle at the tangent point of the instrument line of sight and then translated along the line of sight for integration in a spherical atmosphere. The other spherical model, called CDI-PI (*Rozanov et al.*, 2001) for the Combined Differential-Integral Picard Iterative algorithm, uses a pseudo-spherical initial guess for the diffuse field and iteratively modifies the diffuse field (in a spherical atmosphere) until there is convergence of the solution to the radiative transfer equation. Finally, the two

approximate spherical models, CDI (*Rozanov et al.*, 2000) and LIMBTRAN (*Griffioen and Oikarinen*, 2000) calculate the diffuse field from a series of plane parallel multiple scattering calculations for different solar zenith angles along the instrument line of sight.

For this comparison, both the observation geometry and the atmospheric state were specified in detail in order to minimize those variables that contribute to systematic differences in the results. In this way it is possible to focus on the effects of the different algorithms. The study specified vertical profiles of the neutral density and ozone together with cross sections at each wavelength. The profile for stratospheric aerosol extinction was also specified at each wavelength and the Henyey-Greenstein phase function, with $g = 0.7$, (Equation 4.48) is assumed for the aerosol scattering phase function at all altitudes. Scattering from the Earth's surface is assumed to be Lambertian.

The general results of this comparison showed that the Monte-Carlo models, which *Loughman et al.* deem to be most correct in a general sense, agree within 1.5% in total limb radiance. The spherical models agree with the Monte-Carlo models to within approximately 2-4%, except for GSLS in those cases where the viewing geometry exaggerates the error from the assumptions used in the calculation of the multiple scatter contribution. The approximate spherical models are known to produce systematically larger total radiances than the spherical models. These differences, which increase with altitude to approximately 9% at 60 km, are directly related to the plane parallel assumptions in the calculation of the multiple scattering contribution (*Griffioen and Oikarinen*, 2000). It was also noted that the spherical models, especially the Monte-Carlo models, require a large amount of computation time, up to to several minutes on a desktop PC, to simulate a limb radiance profile at a single wavelength. This limitation is prohibitive for a retrieval, although with linearization techniques the use of Monte-Carlo models and other spherical techniques in retrievals are currently being explored (*Postylyakov*, 2004b; *Walter*, 2006).

Loughman et al. present several comparison cases for limb radiances at 325, 345, and 600 nm including single scatter, total radiance with low albedo and total radiance with high albedo, all for several solar geometries. The high albedo total radiance calculation has been chosen for the comparison with the SASKTRAN results in the present work as this case presented the largest difference between the model results. The treatment of aerosols in the SASKTRAN model had to be modified slightly in order to incorporate the direct specification of extinction rather than a particle size distribution and number density. Similarly, the phase function was modified to use the Henyey-Greenstein approximation.

Figure 4.21 is a reproduction of Figure 6(c) from *Loughman et al.* (2004) together with the SASKTRAN results. This figure shows the total radiance at 325 nm for an aerosol laden atmosphere and a scene albedo of 0.95. The Monte-Carlo model Siro is taken as the reference model so that comparisons are presented as the percent difference from the Siro result. Each subplot corresponds to a particular solar geometry defined by the solar zenith angle at the tangent point of the line of sight, θ_T, and the azimuth angle from the line of sight to the solar vector in the xy-plane, ϕ_T (see the geometrical definitions shown in Figure 2.1). It should be noted that the solar geometries that correspond to $\theta_T = 60$, 80, and 90° at solar azimuth $\phi_T = 90°$ are most representative of the OSIRIS sun-synchronous dusk/dawn orbital geometry. Figure 4.22 is the same as Figure 4.21, but for a wavelength of 600 nm, and is identical to Figure 8(c) from *Loughman et al.* (2004) with the addition of the SASKTRAN results. The comparisons at 345 nm are similar to those for the other two cases.

At 325 nm, the SASKTRAN results compare well with Siro. The overall agreement with Siro is very similar to that of the other spherical models except for GSLS, which suffers at solar geometries with high and low azimuth angles. When the sun is on the horizon, *i.e.* at $\theta_T = 90°$, there is a "knee" in the limb radiance profile (*Kaiser et al.*, 2004) due to the large optical depth along the path from the sun to the scattering point. This results in a decreased radiance at lower tangent altitudes

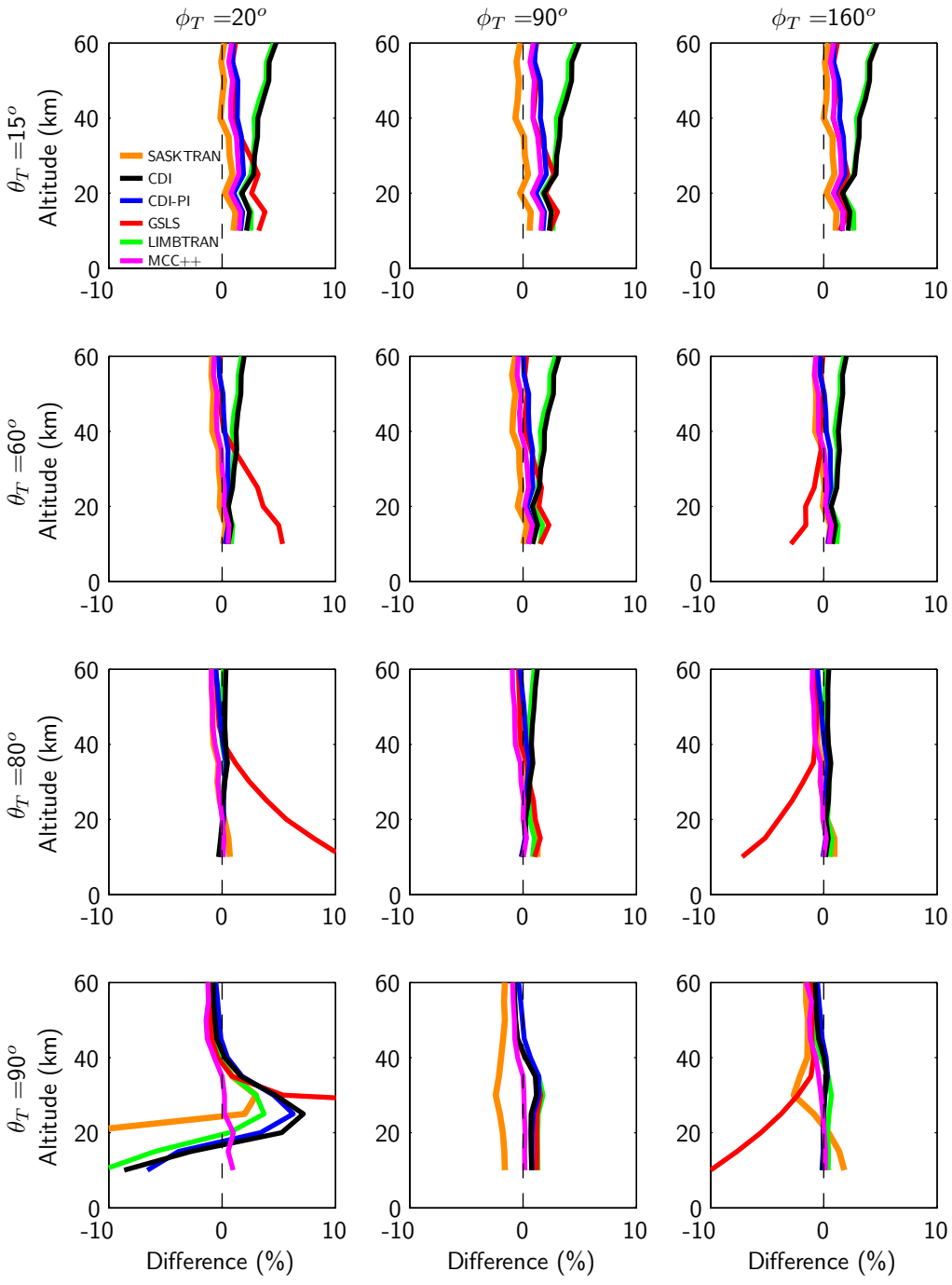

Figure 4.21: A comparison of the total limb radiance percent difference from Siro for 12 solar geometries at 325 nm with an albedo of 0.95. Identical to Figure 6(c) from *Loughman et al.* (2004) with the addition of the SASKTRAN results.

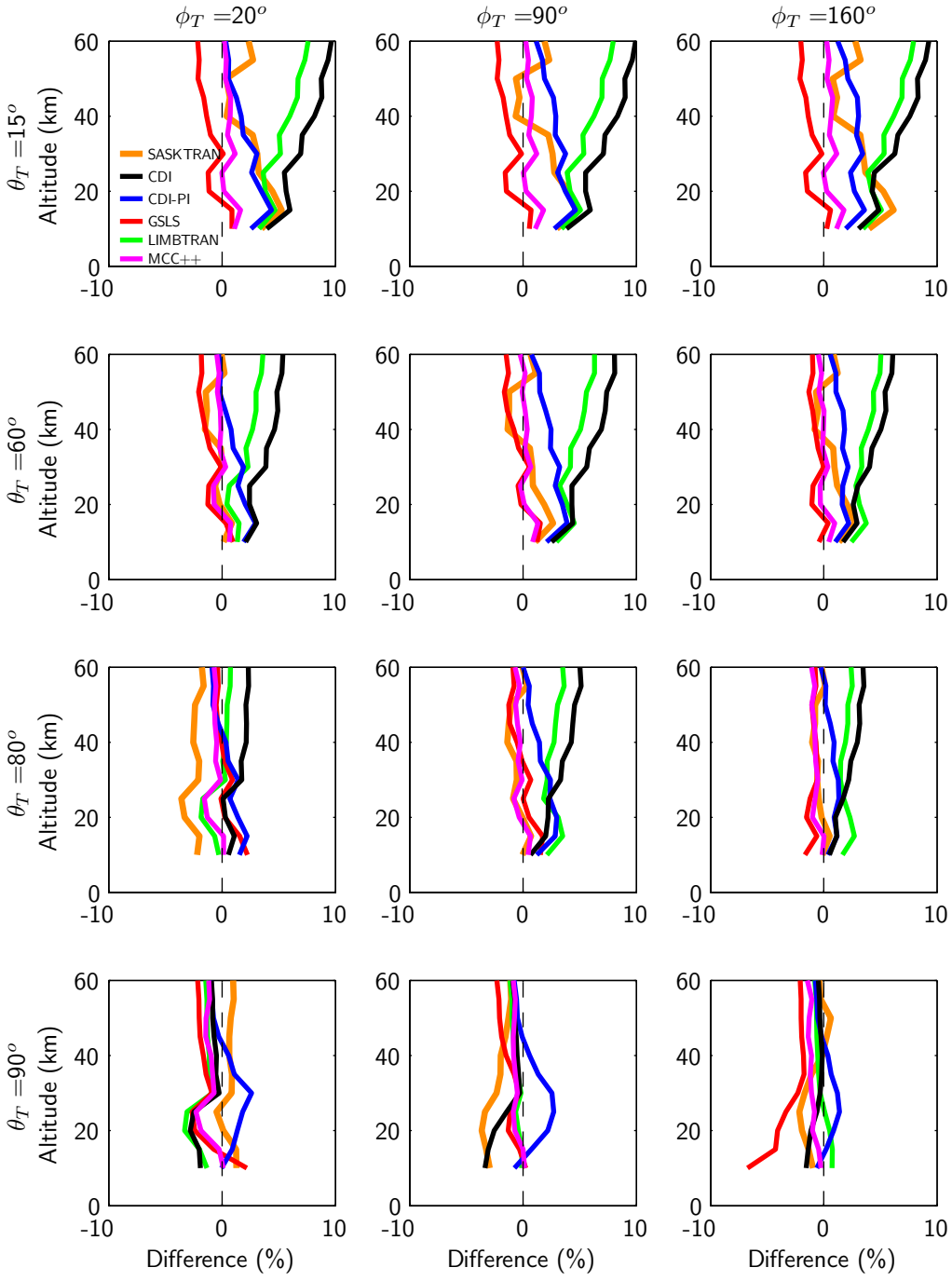

Figure 4.22: A comparison of the total limb radiance percent difference from Siro for 12 solar geometries at 600 nm with an albedo of 0.95. Identical to Figure 8(c) from *Loughman et al.* (2004) with the addition of the SASKTRAN results.

and an increased difference between the models. This geometry is difficult for the models and larger differences are to be expected. The cause of the 2–3% difference in the SASKTRAN result at $\theta_T = 90°$, $\phi_T = 90°$, is unknown, but the difference is relatively small and not of great concern when the variation in the other models with respect to Siro for all $\theta_T = 90°$ conditions is recognized. As noted by *Loughman et al.* (2004), both of the models that use plane parallel approximation of the multiple scatter source function, LIMBTRAN and CDI, produce results that are systematically high at the upper altitudes. This effect is not evident in the SASKTRAN result. Similar results are shown in Figure 4.22 for the comparisons at 600 nm, although the overall spread between the model results is greater. The systematic error in LIMBTRAN and CDI is also exaggerated. Overall the SASKTRAN agreement is good when compared with that from the other spherical models.

It can be concluded from these comparisons that the SASKTRAN model is in better agreement with the Monte-Carlo models than the approximate spherical models (LIMBTRAN and CDI) and about as well as the other spherical models (GSLS and CDI-PI); for certain cases, the SASKTRAN results agree much better than GSLS with the Siro results. The advantage of SASKTRAN is that the code efficiency makes it suitable for inversions, even when high orders of scattering are required (**?**). Of these various radiative transfer models, the only ones that are currently used for operational level processing of satellite data are LIMBTRAN, CDI, and GSLS. However, SASKTRAN provides a solution that is suitable for operational inversions and performs well in comparison to the reference models.

4.7.2 OSIRIS Measurements

The SASKTRAN results are primarily used to retrieve vertical number density profiles of atmospheric constituents from OSIRIS measurements. Therefore, it is important for this case, and in general for any modelling project, that the modelled results accurately reproduce the measurement. The plots shown in Figure 4.23 clearly

demonstrate this is the case for the SASKTRAN results and a typical OSIRIS measurement set. The observations used in this comparison were taken from an OSIRIS scan made on April 29, 2002. The mean latitude and longitude of the tangent points of the lines of sight for this scan are 82° N, 38° E. No enhanced scattering from clouds was observed at low tangent altitudes and so it is believed that this measurement set has a high uniform albedo due to snow or ice conditions. The solar zenith angle at the average tangent point of the lines of sight in this scan is 68°.

Specifically, Figure 4.23 shows the OSIRIS radiance measurements and the corresponding SASKTRAN modelled radiance at tangent altitudes of 15, 25, 35 and 40 km over the entire OSIRIS wavelength range, *i.e.* 280 to 810 nm. Although SASKTRAN can operate with a wavelength dependent albedo, a value of 0.84 was used at all wavelengths for these comparisons. This value was determined through a comparison of the measured and modelled limb radiance at 40 km tangent altitude and provides the best fit in a small wavelength range near 700 nm. This is an appropriate tangent altitude and wavelength range to use to infer the albedo because the dominant terms in the limb signal are the single scatter of the solar beam and a single reflection of the solar beam from the Earth's surface that is successively scattered into the OSIRIS line of sight. Atmospheric multiple scatter is almost insignificant in this region.

The ECMWF data for the time and location of the OSIRIS scan was used for the atmospheric temperature profile and the neutral density profile. The ozone, nitrogen dioxide and stratospheric sulphate aerosol profiles were retrieved directly from the OSIRIS measurements. Details of the gas retrievals have been presented by *Roth et al.* (2007). The aerosol retrieval is described in the following chapters. In this work the SASKTRAN model does not include other absorbing species such as molecular oxygen and water as seen by the mismatch between the model and measurements at these absorption features in the red end of the spectrum.

The SASKTRAN total limb radiance has been separated in Figure 4.23 into components that represent the contribution from a single scatter of the solar beam, and

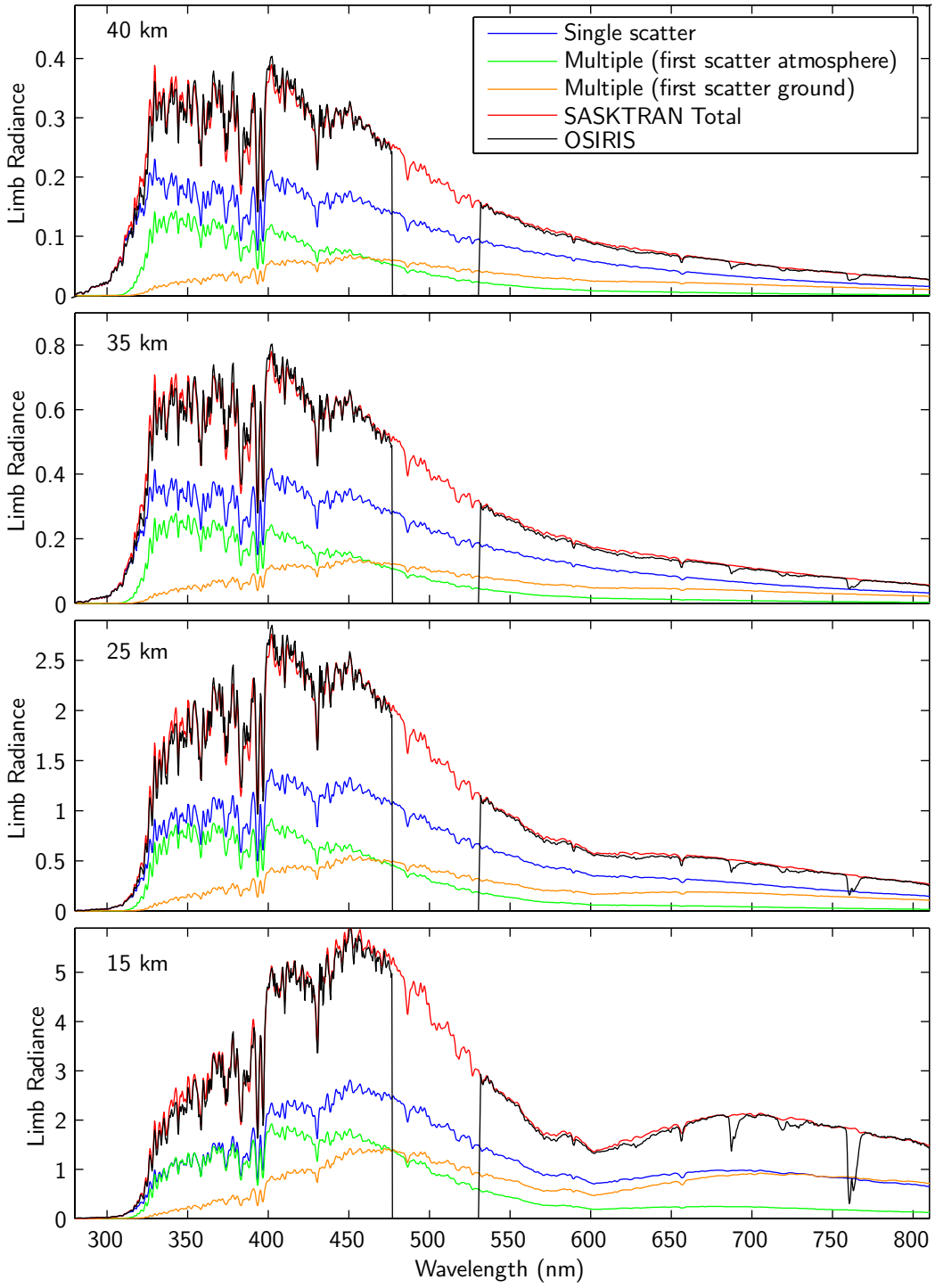

Figure 4.23: Measured (OSIRIS) and modelled (SASKTRAN) limb radiance spectra (10^{13} photons/s/cm^2/nm/sterad) at tangent altitudes of 15, 25, 35 and 40 km for scan 06432012.

from the converged multiple scatter solution. The multiple scatter contribution is further separated into components that result from a first scattering event within the atmosphere and from the Earth's surface. It is clear that at different tangent altitudes and in different wavelength ranges the contribution from each of these terms exhibits a significant variation and demonstrates the importance of an accurate radiative transfer model for analysis of limb scattered sunlight measurements.

The plots shown in Figure 4.23 are typical of the comparisons between the SASKTRAN modelled results and the OSIRIS radiance measurements. Certainly, the variability of the scene below the observations, *i.e.* spatially and spectrally dependent albedo and clouds, affects the measured result. Comparisons with the SASKTRAN model show that this level of agreement is only possible when the measurement conditions are well known. The absolute calibration of the OSIRIS instrument is also important in this comparison as the determination of the scene albedo requires matching of the absolute measured and modelled radiances. Recent astronomical measurements and internal OSIRIS documentation provide the absolute calibration used here with an uncertainty of $\pm 10\%$ (*Lloyd, personal communication,* 2006). While this uncertainty is important and does affect the comparisons presented here, the excellent match of the measured and modelled spectra across all wavelengths from 310 to 810 nm, and at all altitudes from 15 to 40 km, indicates the consistency of the measurement and the model.

4.8 Conclusions

The SASKTRAN model has been developed using a successive orders of scattering algorithm in a spherical geometry. The calculation of the multiple scatter source term is exact, with respect to the solar geometry, for the first two orders of scattering from the atmosphere and the ground. The integral over the unit sphere of the incoming diffuse radiance field that is required for the estimation of the source term is highly

variable near the local horizon. Careful spacing of the integration nodes, with a high density of rays near and just below the local horizon, provides an accurate solution of the source term to within 1% for a relatively small number of rays. When the sun is relatively high in the sky, i.e. solar zenith angle less than 70°, the estimate of the multiple scatter source term at the tangent point of the line of sight alone is sufficient; closer to the solar terminator, the multiple scatter source term must be estimated at several points along the line of sight in order to avoid a systematic error of approximately 10% at tangent altitudes below 20 km.

Comparisons at wavelengths and geometries specified by *Loughman et al.* (2004) show that the SASKTRAN results compare better with a reference Monte-Carlo model than the approximate spherical models and generally as well as the true spherical models. The SASKTRAN results compare well with the OSIRIS measurements at wavelengths from 310–810 nm and at altitudes from 15–40 km for a best fit scene albedo and the retrieved profiles of ozone, stratospheric aerosol and nitrogen dioxide. The implementation of the algorithm using caching and multi-threading computational techniques has resulted in an efficient code that is suitable for improved operational inversion of limb scatter satellite data with modern desktop computers.

Chapter 5

Aerosol Number Density Retrieval

An algorithm for the retrieval of global stratospheric aerosol profiles is presented in this chapter using the OSIRIS limb scatter measurements as an example data set. The retrieval utilizes a one dimensional version of the MART non-linear relaxation inversion suitable for limb scatter. A height profile of the particle size distribution must be assumed in order to retrieve the aerosol number density. An altitude normalized wavelength ratio measurement vector is employed to minimize effects of upwelling radiation from ground albedo and uncertainties in the neutral density profile. Using a method of numerical perturbation, a formal error analysis is performed that shows that the dominant error term is the measurement noise. Comparison of SAGE II and SAGE III coincident measurements with the OSIRIS result converted to extinction shows agreement with the limb scatter retrievals to within 15% throughout the lower stratosphere for an appropriate choice of particle size distribution. A sample set of the OSIRIS Level 2 aerosol product is presented that highlights the relatively high sampling resolution of the limb scatter technique.

5.1 Aerosol signature in limb scatter

For spherical droplets of sulphuric acid, the scattering cross section at visible and near infrared wavelengths is several orders of magnitude larger than the absorption cross section. Therefore in a basic sense, where the total atmospheric optical depth is small, stratospheric aerosols enhance the total limb radiance beyond the Rayleigh background level by a small fraction. In regions of high total atmospheric optical depth, addition of aerosols contribute to the extinction of radiation from the scattering point along the line of sight more than they contribute to additional scattering into the line of sight. This results in a reduction of the total limb radiance by a small fraction. Therefore, unlike the limb sensitivity to an absorbing gas species, which is always negative, the addition of aerosol in the atmosphere can increase or decrease the limb signal, depending on the altitude and wavelength range.

Figure 5.1 is a plot of the sensitivity, or kernel matrix (Equation 3.8), of the limb radiance at selected OSIRIS wavelengths to the aerosol number density profile for an assumed log-normal particle size distribution and under normal OSIRIS viewing conditions (solar zenith angle of 72 degrees, scattering angle of 88 degrees) calculated using the SASKTRAN forward model. Each curve in Figure 5.1 represents the sensitivity of the radiance at a single tangent altitude calculated using a typical aerosol number density profile that is successively perturbed by +1% at each altitude from 1 to 40 km. The curves tend to peak at the altitude corresponding to the tangent altitude. This is because the majority of the information in a limb spectra comes from the tangent point of the line of sight and is discussed in more detail in Section 5.3.

In general, the sensitivity to aerosol increases at longer wavelengths. At the shortest measured wavelengths (less than 300 nm), the limb radiance is essentially insensitive to stratospheric aerosols. The sensitivity at the other wavelengths, greater than 300 nm, shows an enhancement of the limb radiance due to an increase in aerosol density for upper altitudes, and an extinction of the signal at lower altitudes where the optical thickness is large. Note that the altitude where the sensitivity changes

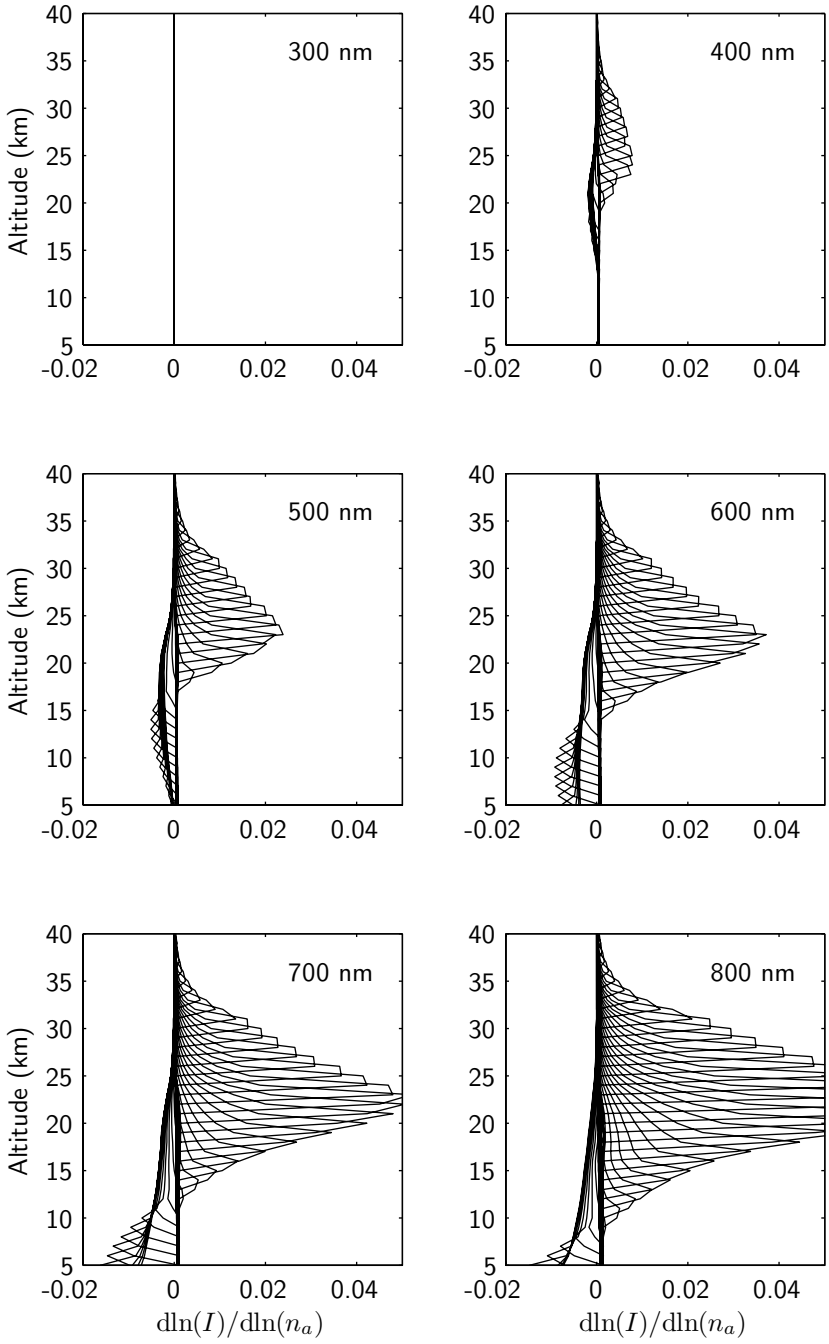

Figure 5.1: Sensitivity of the modelled limb radiance, I, at selected OSIRIS wavelengths to changes in the aerosol number density at each 1 km altitude layer.

from positive to negative is dependent on wavelength and is lower at the longest wavelengths. This point of insensitivity also moves in altitude depending on the shape and magnitude of the aerosol profile. The increased sensitivity of the limb signal to aerosol at longer wavelengths is partly due to the fact that the Mie scattering cross section of the aerosol particles does not decrease as rapidly with wavelength as the Rayleigh scattering cross section of the neutral density.

For these calculations, a height profile of a single mode, log-normal size distribution and a number density profile that are simple parameterizations consistent with SAGE II retrievals as calculated by *Bingen et al.* (2004) for post-volcanic conditions (September, 1993) are used. Mode radius varies from 0.4 to 0.3 μm from 10 to 40 km and mode width varies from 1.2 to 1.1 over the same range. This size distribution is chosen as it is significantly influenced by the larger volcanic aerosols and represents a worst case scenario for the modelling and sensitivity studies of this inversion technique. It is shown in Section 5.6 that a size distribution of smaller particles that better represents the background aerosol state of more recent years is required in order to retrieve OSIRIS extinction that agrees well with SAGE II/III measurements. However, the validity of the technique presented here, which is a retrieval of the number density for an assumed distribution, still applies.

In the forward model, two parameters are required to characterize the effect of the aerosol particles. One of these is the aerosol extinction (see Equation 4.13),

$$k(\lambda) = n_a \sigma(\lambda) \tag{5.1}$$

defined as a product of the aerosol number density, n_a and the scattering cross section, $\sigma(\lambda)$, as calculated by Mie theory based on the particle size distribution. The absorption cross section of stratospheric sulphate aerosols is negligible compared to the scattering cross section. As the extinction is a product, the effect of particle size distribution and number density are intertwined in such a way that, roughly

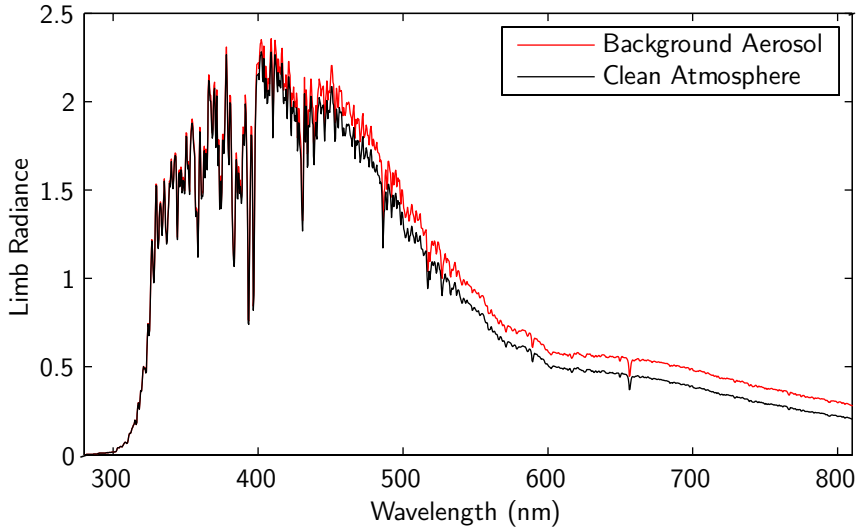

Figure 5.2: Modelled limb radiance spectra (units of 10^{13} photons/s/cm^2/sterad) at 25 km tangent altitude for clean Rayleigh/O_3 atmosphere and for the same conditions with a typical background stratospheric aerosol load.

speaking, a small number of large particles can produce the same extinction as a large number of small particles. In a retrieval sense, the extinction is not as strongly affected as the number density by incomplete knowledge of the particle size distribution; however, particle size still has an important effect through the second required parameter of the radiative transfer calculation, the scattering phase function, p(Θ). Conveniently, for stratospheric aerosol particle sizes and visible wavelengths *McLinden et al.* (1999) show that the phase function changes slowly with mode radius such that the extinction is not a strong function of particle size.

Figure 5.2 is a plot of the modelled OSIRIS spectrum at 25 km tangent altitude for a clean, Rayleigh/O_3 atmosphere and for the same conditions with a typical stratospheric aerosol load. It is clear that the addition of aerosol enhances the signal at long wavelengths; short wavelengths are relatively insensitive to aerosol loading.

5.2 The retrieval vector

The general formulation of the inverse problem as introduced in Chapter 3 is,

$$\mathbf{y} = \mathbf{F}(\mathbf{x}, \tilde{b}) + \boldsymbol{\epsilon}\,, \tag{3.5}$$

where each element of the measurement vector, $\mathbf{y}$, is associated with a measured altitude, $\mathbf{F}$ represents the physics of the measurement, *i.e.* the forward model, and is a function of the desired atmospheric parameters, $\mathbf{x}$, often referred to as the state vector. The forward model is also a function of many other variables, all denoted by $\tilde{b}$. In this case, the state vector is the unknown aerosol profile.

A spectral analysis technique, often used in limb scatter retrievals, attempts to construct each element of the measurement vector, $\mathbf{y}$, using an algebraic combination of radiance measurements at different wavelengths and altitudes designed to maximize sensitivity to the state vector parameter and minimize the effects of other unknowns.

For the retrieval of trace gases, it is also standard practice to normalize the measured and modelled limb radiance profiles to an exposure at a reference tangent altitude (*Flittner et al.*, 2000; *Haley et al.*, 2004). This technique is similar to the solar reference technique used in limb occultation and removes the need for an absolute calibration. It has been shown previously that this altitude normalization provides a degree of insensitivity to unknown ground albedo (*von Savigny et al.*, 2003) as the contribution from the upwelling radiation to the reference altitude exposure is in a similar ratio for all other tangent altitudes.

It is tempting to choose a set of individual wavelengths in the OSIRIS spectrum away from trace gas absorption, and, using a normalization technique with an upper tangent altitude, solve Equation 3.5 for the aerosol extinction profile that best makes the modelled radiance profile match the measurements at each wavelength. In this case, one must assume a reasonable scattering phase function for the aerosol (Mie or otherwise). This is certainly a possible method; however, it relies on the assumption

that the background Rayleigh atmosphere can be modelled perfectly. Any difference between the modelled and measured radiances that is due to an imperfect knowledge of the neutral density, i.e. a gravity wave, or errors in the assumed profile, is attributed to and fit by the aerosol profile. Figure 5.3 is a plot of the sensitivity of the limb radiance to a change in the neutral density for the same wavelengths and geometry used to show the aerosol sensitivity in Figure 5.1. These are also calculated with the SASKTRAN model and each curve represents the sensitivity of the measurement at a single tangent altitude to successive small (+1%) perturbations in a reference neutral density profile at each altitude (all neutral density profiles used in the following work are derived from ECMWF analysis at OSIRIS scan locations). At all wavelengths and tangent altitudes, the sensitivity to the neutral density is approximately an order of magnitude larger than the sensitivity to the aerosol profile. For a retrieval of the aerosol extinction using only information from a single wavelength, a small error in the assumed neutral density would translate to a very large systematic error in the retrieved aerosol extinction.

Therefore, a spectral analysis construct of the measurement vector is proposed. The ratio of a long wavelength, λ_l, to a shorter wavelength, λ_s, characterizes the Mie scattering wavelength dependence of the aerosol particles shown in Figure 5.2 and provides some insensitivity to uncertainties in the neutral density. Each measured spectrum $I(h, \lambda)$ at tangent altitude, h, is first normalized to a measurement at a reference altitude, h_r,

$$\bar{I}(h_j, \lambda) = \frac{I(h_j, \lambda)}{I(h_r, \lambda)} \,. \tag{5.2}$$

Then, an element of the measurement vector, y_j, is constructed for each measured tangent altitude as the logarithm of the wavelength ratio of the altitude normalized profile,

$$y_j = \log\left(\frac{\bar{I}(h_j, \lambda_l)}{\bar{I}(h_j, \lambda_s)}\right) \,. \tag{5.3}$$

The logarithm of the ratio is used as it creates a better behaved measurement vector

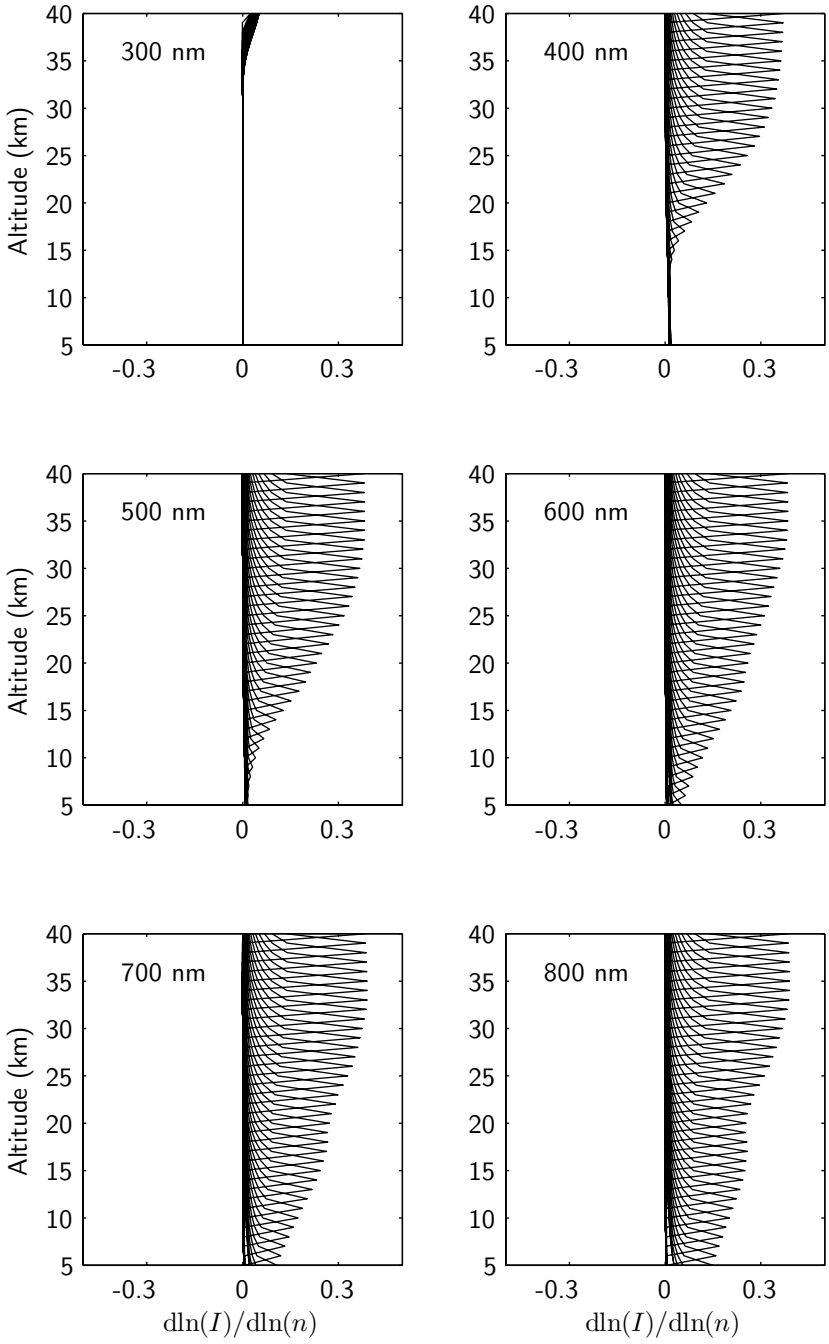

Figure 5.3: Sensitivity of the modelled limb radiance, I, at selected OSIRIS wavelengths to changes in the neutral density, n, at each 1 km altitude layer.

due to the exponential nature of the height profile.

The choice of wavelengths is of course quite critical. It is important to choose λ_l and λ_s with the largest possible separation in order to maximize the spectral characterization of the Mie scattering. However, λ_s must remain sufficiently long in order to provide sensitivity at lower stratospheric tangent altitudes as the limb optical depth quickly becomes large for blue and UV wavelengths, approximately $\lambda < 450$ nm. The OSIRIS spectral order sorter, an unavoidable characteristic of the optical design, contaminates wavelengths from 475 nm to 535 nm (*Llewellyn et al.*, 2004) and the Chappuis ozone absorption band has a significant cross section between 500 nm to 680 nm (see Figure 4.4). Therefore, $\lambda_s = 470$ nm has been chosen, essentially the longest wavelength on the short wavelength side of the order sorter. For the long wavelength $\lambda_l = 750$ nm is used to maximize separation from λ_s and avoid the O_2 A-band absorption (762 nm) and Woods anomalies from the polarization response of the grating above 780 nm (*McLinden et al.*, 2002). A reference tangent altitude of h_r =40 km was chosen as this is an upper bound for the retrieval because the aerosol number density is negligibly small above approximately 35 km. This represents a type of calibration point where the signal can be reliably modelled as it is not greatly affected by the aerosol load.

The sensitivity of this measurement vector to both the aerosol density and the neutral density calculated with SASKTRAN through successive perturbation of each species at specific altitudes is shown in Figure 5.4. The relative sensitivity to the neutral density compared to the aerosol density is indeed more favorable than for the single wavelength cases. Above 20 km altitude, the vector is more sensitive to aerosol density than it is to neutral density. For lower altitudes nearer to the tropopause, the sensitivity of the neutral density still grows beyond that of the aerosol, but down to approximately 15 km they remain on the same order of magnitude. It is worthwhile to note that this analysis only considers localized changes, or errors, in the neutral density profile. If the assumed neutral density profile was systematically too low

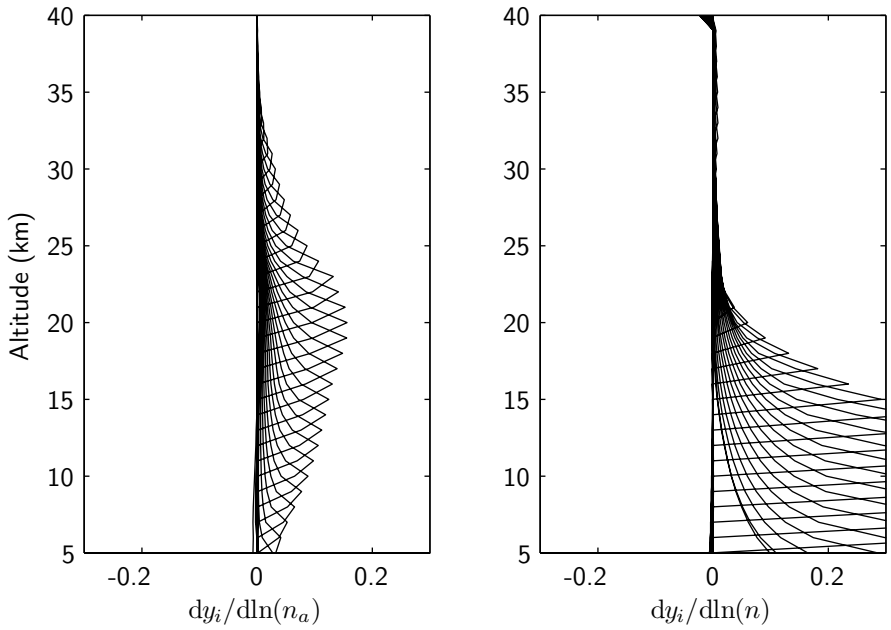

Figure 5.4: Sensitivity of the measurement vector $\mathbf{y}$ defined in Equation 5.3 to changes in aerosol density, n_a, and neutral density, n, respectively.

or too high at all altitudes, it would have almost no effect because of the altitude normalization.

5.3 Inversion

As discussed in Chapter 3, an adaptation of the Multiplicative Algebraic Reconstruction Technique (MART) has been used successfully for the two-dimensional tomographic retrieval of the mesospheric volume emission rates of the OIRA band measured by the IRI subsection of the OSIRIS instrument (*Degenstein et al.*, 2003). It was also noted that a further adaptation of the technique is used for the one dimensional species profile inversion from the OS data.

The MART relaxation iteration equation, included again here for convenience,

$$x_i^{(n+1)} = x_i^{(n)} \sum_k \sum_j \frac{y_{jk}}{F_{jk}(\mathbf{x}^{(n)}, \tilde{b})} W_{ijk} \,, \tag{3.21}$$

reduces to

$$x_i^{(n+1)} = x_i^{(n)} \sum_j \frac{y_j}{F_j(\mathbf{x}^{(n)}, \tilde{b})} W_{ij} \tag{5.4}$$

for a single spectral measurement at each tangent altitude. *Degenstein et al.* (2003) set the weighting filter terms, W_{ij}, proportional to the path length matrix based on intersections with the grid cell in the 2D atmosphere used for the tomographic inversion. For this work, a simplified, but similar variant of $\mathbf{W}$ is employed. For this one dimensional problem, the path length through a spherical shell at altitude i is maximum for the observation j that samples with a tangent altitude nearest to i. The path length for observation j falls off quickly for shells at altitudes above i. Therefore, all measurements with a tangent altitude lower than the altitude of interest, i, have a successively smaller path length of intersection through shell i. All measurements with tangent altitude above measurement j do not intersect shell i. The weighting filter, $\mathbf{W}$, allows these measurements which are lower in tangent altitude than measurement j, to contribute to the retrieval at shell i through the path length relationship. Since a large separation in tangent altitude of incorporated measurements is not desirable, a somewhat arbitrarily set maximum of three measurements are allowed to contribute to the retrieval of a given element of the state parameter.

Measurements near the bottom of the scan are handled as a special case. $\mathbf{W}$ is therefore a lower triangular square matrix where each row i corresponds to a retrieval altitude. Since the OSIRIS measurements are not uniformly spaced in tangent altitude and are not aligned with shell boundaries, $\mathbf{W}$ is different for each scan. For simplicity, it is desirable to use the same weighting filter for each scan, loosely based on the path length relationship. A typical example, for a scan of 6 exposures, where

$i = 1$ (the first row) is the lowest tangent altitude, the weighting filter is

$$\mathbf{W} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0.25 & 0.75 & 0 & 0 & 0 & 0 \\ 0.1 & 0.3 & 0.6 & 0 & 0 & 0 \\ 0 & 0.1 & 0.3 & 0.6 & 0 & 0 \\ 0 & 0 & 0.1 & 0.3 & 0.6 & 0 \\ 0 & 0 & 0 & 0.1 & 0.3 & 0.6 \end{pmatrix}. \tag{5.5}$$

With a sufficient number of iterations, any similar choice of weighting filter is equally adequate.

Equation 5.4 can be written in matrix form as

$$\boldsymbol{\alpha} = \mathbf{W}\mathbf{m} \tag{5.6}$$

where $\boldsymbol{\alpha}$ is a column vector of the factors used to update the state parameter upon each interation,

$$\alpha_i = \frac{x_i^{(n+1)}}{x_i^{(n)}}, \tag{5.7}$$

and $\mathbf{m}$ is a column vector of the ratio of the measurement vector to the forward model evaluated at the current value of the state parameter.

$$m_j = \frac{y_j}{F_j(\mathbf{x}^{(n)}, \tilde{b})} \tag{5.8}$$

Note that both $\boldsymbol{\alpha}$ and $\mathbf{m}$ have the same number of elements. For OSIRIS sampling geometry, this is equal to the number of exposures in a limb scan.

To test the retrieval, a simulation of the limb measurement using a known aerosol number density profile, labeled "control" in Figure 5.5, was calculated with SASKTRAN. Then, beginning with an initial guess profile, ten iterations of the MART inversion were performed using the measurement vector defined by Equations 5.2 and

5.3. Using the SASKTRAN model, this requires computation time of approximately two minutes on a standard desktop PC. The measurement vectors corresponding to both the initial guess profile and the retrieved profile are shown in the left panel of the figure, along with the vector directly calculated from the simulated measurements. The right panel of the figure shows the number density profile for the same three cases. The inversion is successful at adjusting the modelled measurement vector to better match the control case. This is also the case for the number density profile. The retrieval matches the control profile to within 3% except at the lowest altitudes where the difference is closer to 15%. The larger discrepancy at lower altitude is due, in part, to the limb geometry, which causes an "onion-peel" like nature of the retrieval. For a small number of iterations, the values at the highest altitudes converge the earliest. As iterations proceed, the lower altitudes then converge because the measurements that sample these altitudes with the largest sensitivity also have a smaller, but significant, sensitivity to the upper altitudes due to the path length through the upper shells. For a larger number of iterations, i.e. ≥ 30, all altitudes above approximately 12 km have converged to within a small percent difference. The difference that arises at lower altitudes is addressed further in Section 5.5.1.

5.4 Extinction Product

The retrieved solution is the number density that makes the forward model best match the measurements for the assumed aerosol particle size distribution profile. Thus the solution is really only truly meaningful in the light of this assumption. It was noted previously in Section 5.1 that the extinction (Equation 5.1) provides some measure of insensitivity to errors in the assumed particle sizes.

In order to test the variability of the retrieved extinction with respect to errors in the assumed particle size, three simulations were performed. The SASKTRAN forward model was used to construct a set of simulated observations that each used the

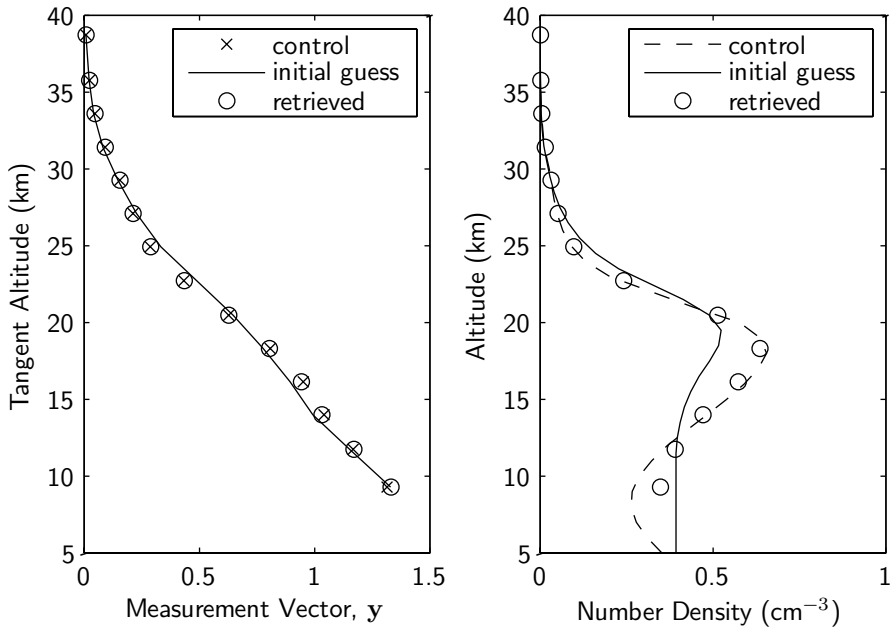

Figure 5.5: A simulation of the aerosol number density retrieval using an assumed size distribution profile. The left panel shows the measurement vector for three cases: a simulated aerosol profile labeled "control", the initial guess profile, and the retrieved profile after 10 iterations of the MART inversion. The right panel plots the number density corresponding to the same three cases.

same aerosol number density profile (labeled "control" in Figure 5.5) and a different height profile of the size distribution. Figure 5.6(a) shows the aerosol cross section at 750 nm associated with the climatological profile of the size distribution used for the simulated retrieval shown in Figure 5.5. The cross sections that result from adjusting r_g by $\pm 10\%$ are also shown. It is important to note that there is a different scattering phase function associated with each cross section. The aerosol number density was retrieved for the three sets of simulated measurements, each time assuming the same climatological size distribution profile. Effectively, the retrieval was performed for perfect knowledge of the particle distribution, and for the cases where the particles in the atmosphere are both smaller and larger than those assumed in the retrieval. These results are shown in Figure 5.6(b).

Where both the simulated measurements and the retrieval used the same cross sections, the solution is identical (to within a small percentage) to the control profile.

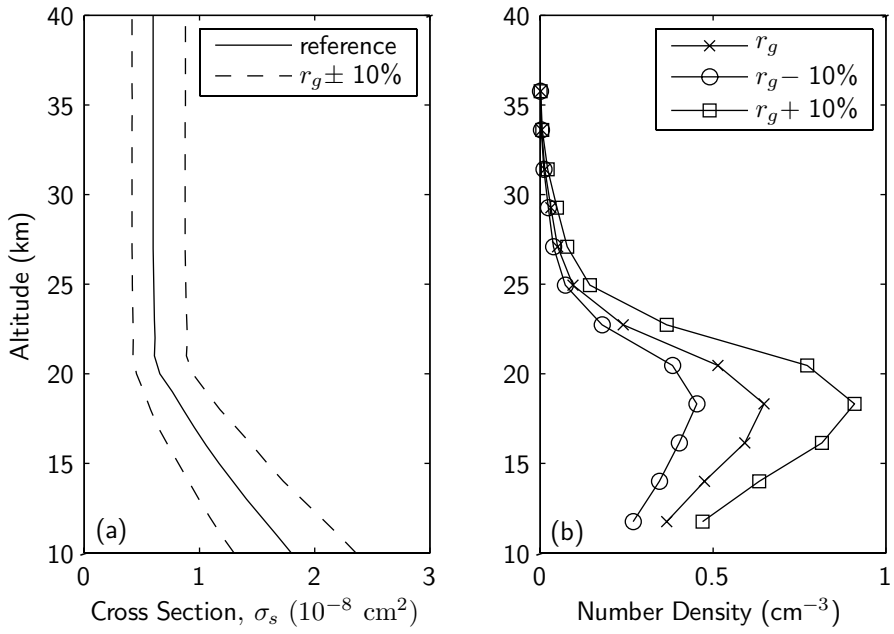

Figure 5.6: (a) Height profile of aerosol scattering cross sections at 750 nm for climatological particle size distributions and for the same with mode radius adjusted by ±10%. (b) Simulated number density retrievals for forward model cases each using the same aerosol number density and the cross sections shown in (a).

When r_g was adjusted by -10% in the simulated measurements, the retrieved number density is significantly less, by almost a factor of 2, than the control profile at all altitudes. This is essentially a compensation of the lower optical depth due to the smaller cross section. Because the retrieval still assumes the climatological cross sections, the retrieved density must be lower. The opposite is true for the case where r_g is adjusted by +10% and the retrieved density is larger than the control profile.

It is apparent that the retrieved aerosol number density is quite sensitive to errors in the assumed size distribution. However, as expected the extinction, shown in Figure 5.7 for each of the three cases, does provide some insensitivity to the size distribution. The known, or reference, extinction is calculated as a product of the control density and the three different cross section profiles. The retrieved extinction is the product of the retrieved density and the cross section for the climatological size distribution that is the same in each case. In the first case shown in Figure 5.7(a),

where the forward model and the retrieval use the same cross sections, the difference in extinction in less than 3% at all altitudes. This difference results directly from the inversion and is due to the interpolation of the profile onto the SASKTRAN altitude grid at each iteration and the convergence of the solution (30 iterations were used to improve algorithm performance at lower altitudes). Figures 5.7(b) and (c) show a comparison of the known and retrieved extinction for simulated smaller and larger mode radius, respectively. For both of these cases, the extinction profiles agree with the reference to within approximately 10%.

This larger discrepancy is due to the scattering phase function. Particles used in the forward model to simulate the measurements for cases (b) and (c) have a different phase function than the particles used for the retrieval. By adjusting the number density appropriately, the retrieval is able to compensate for the difference in optical depth caused by the different cross section. However, the difference in phase function causes an unavoidable systematic error in the retrieved extinction.

For particles, the phase function is strongly peaked in the forward and back directions and is reasonably flat for scattering near to the plane that is perpendicular to the incident light. For the OSIRIS geometry, the angle between the line of sight and the solar vector is always near 90 degrees so that the single scatter contribution to the limb radiance depends only on this relatively flat region of the aerosol phase function. However, the diffuse radiation field, which arises from atmospheric multiple scattering and upwelling radiation, requires scattering at all angles. This can explain the fact that for all three cases shown in Figure 5.7, the agreement is very good, *i.e.* $< 3\%$, at altitudes near 16-20 km. The limb radiance at λ_s=470 nm has a very large contribution from the diffuse field due to the large optical depth of the Rayleigh atmosphere, whereas for low values of ground albedo the λ_l=750 nm radiance depends largely on the single scatter component. The aerosol kernel matrix elements for wavelengths around 470 nm drop off quickly at altitudes below about 20 km (see Figure 5.1). Therefore in this region the measurement vector, which depends on

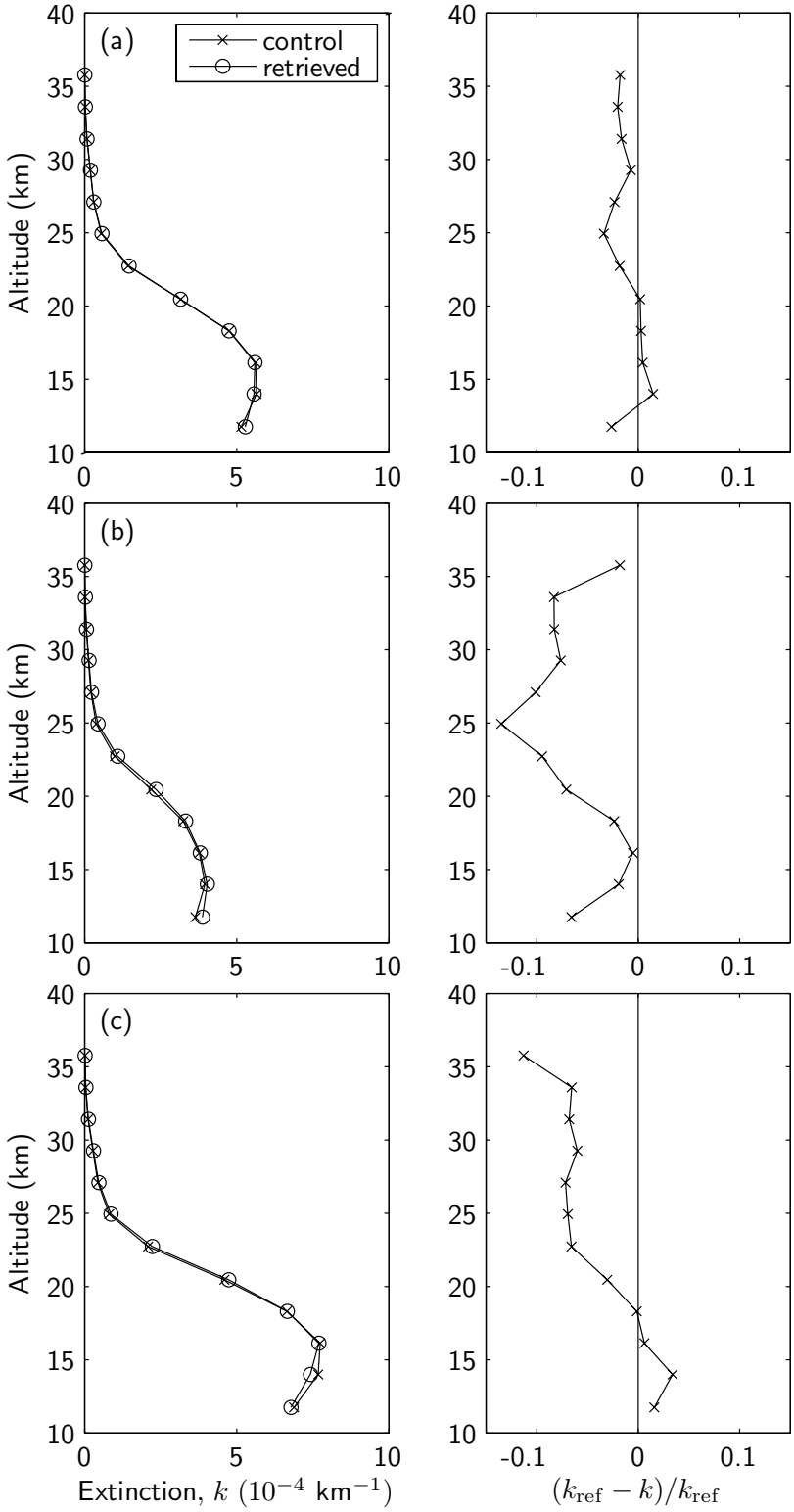

Figure 5.7: Known, labeled "control", and retrieved extinction profiles, and the relative difference for the 3 simulation cases shown in Figure 5.6. The forward model of the observations using (a) same particle size distribution as the retrieval, (b) mode radius adjusted by -10%, and (c) mode radius adjusted by +10%. Agreement is to within approximately 10% for all cases.

the ratio λ_l/λ_s, is largely unaffected by the diffuse field and does not have a large sensitivity to the phase function. Thus, the extinction agrees very well regardless of errors in phase function. Below 16 km the discrepancy in extinction increases again because even at 750 nm the optical depth is becoming large and the contribution from the diffuse field is significant.

5.5 Error

As discussed in Chapter 3, the Rodgers formal error analysis (*Rodgers*, 1976, 2000) identifies three distinct error categories that exist in the retrieval process: smoothing error from the finite resolution of the inversion, measurement error, and forward model error, including both approximations in the model and uncertainties in model parameters. Restating Equation 3.32 in slightly simplified form, the difference between the retrieved profile, $\hat{\mathbf{x}}$, and the true profile, $\mathbf{x}$, is

$$\hat{\mathbf{x}} - \mathbf{x} = (\mathbf{A} - \mathbf{I})(\mathbf{x} - \mathbf{x_a}) + \mathbf{D}\ (\mathbf{\Delta_y}) + \mathbf{D}(\mathbf{\Delta_F}) \tag{5.9}$$

where $\mathbf{x_a}$ is the *a priori* profile and is used as a linearization point to describe the dependence of the forward model on changes in the state parameter, $\mathbf{A}$ is the averaging kernel matrix, and $\mathbf{D}$, is the contribution function matrix. The first term in Equation 5.9 is related to the smoothing error where the solution is considered exact with uncertainty due to the resolution of the retrieval. The second term includes errors that come directly from the measurement process, $\mathbf{\Delta_y}$, such as detector random noise. The third term describes error that arises from the forward model, $\mathbf{\Delta_F}$, from approximations in the physics of the model, and from uncertainty in model parameters such as neutral density, albedo, etc.

As stated in Section 3.7.2, for a relaxation-type inversion, such as the MART, an algebraic derivation of these matrices in not applicable. However, following *Puliafito et al.* (1995), an attempt is made in this work to determine the matrices $\mathbf{A}$ and $\mathbf{D}$

numerically through perturbation of the true profile and the measurements using a set of simulated measurements constructed with the forward model such that the true state is known. A standard formal error analysis is then carried out; however, it should be noted that it is only strictly valid for the profile used in this numerical analysis.

5.5.1 Smoothing Error

In the absence of noise, the relaxation-type inversion always converges to the exact solution in the limit of infinite iterations (*Puliafito et al.*, 1995). This means that the averaging kernel is very close to the identity matrix, i.e. a change in the true profile is perfectly reflected in the retrieved solution.

$$\lim_{n\to\infty} \mathbf{A} = \mathbf{I} \tag{5.10}$$

To test the magnitude of the smoothing error for the aerosol inversion proposed here, columns of the averaging kernel matrix, $\mathbf{A}$, were calculated by successively performing the simulated retrieval shown in Figure 5.5, each time perturbing a single altitude of the true state, $\mathbf{x}$, by 1% and computing the differences in the retrieved profile, $\hat{\mathbf{x}}$ (see Equation 3.31). Note that $\mathbf{A}$ is dimensionless. Figure 5.8 shows a plot of the columns of $\mathbf{A}$ for altitudes between 10 and 40 km at 2 km vertical resolution. For altitudes above 20 km, the columns are very close to delta functions centered at the altitude of the perturbation. This means that inversion captured the perturbation of the true state in the retrieval very accurately. Below 20 km, the effect is blurred, or smoothed, over a range of altitudes near the perturbation resulting in a retrieval error that increases with lower altitudes. The reason for this lack of information at the lower altitudes is the very large limb optical depth of the atmosphere. Even in a perfect simulation with no measurement noise or forward model error the inversion cannot perfectly converge on the true state at large optical depths. This is simply a

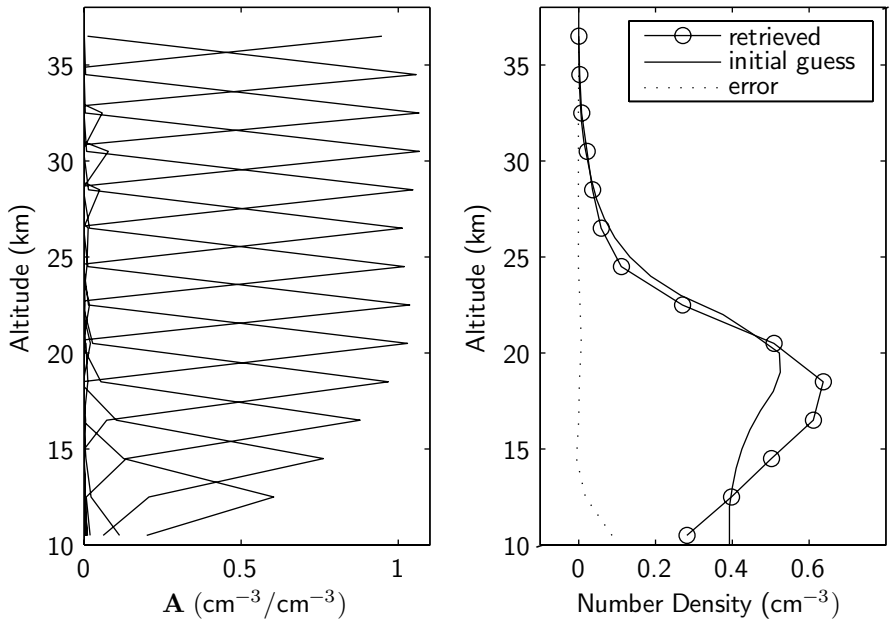

Figure 5.8: The averaging kernel matrix, **A**, and the smoothing error for a typical case.

physical limitation of the measurement and interpretation of the retrieved state must be carried out with caution at altitudes below 20 km.

The right panel of Figure 5.8 is a plot of the results of a simulated retrieval with an initial guess profile (in this context taken as the *a priori* profile) and a curve representing the smoothing error due to the inversion. It is calculated as the magnitude of the first term in Equation 5.9 and is equal to the difference between the retrieved state and true state and shows that for a typical case a smoothing error of 30% at lower altitudes is possible.

5.5.2 Measurement Error

The error in the retrieval that results directly from error in the OSIRIS measurements is determined through the contribution function matrix, **D**. Similar to the averaging kernel calcuation, columns of **D** are calculated numerically by perturbing elements of a simulated measurement vector profile and calculating the resulting difference in

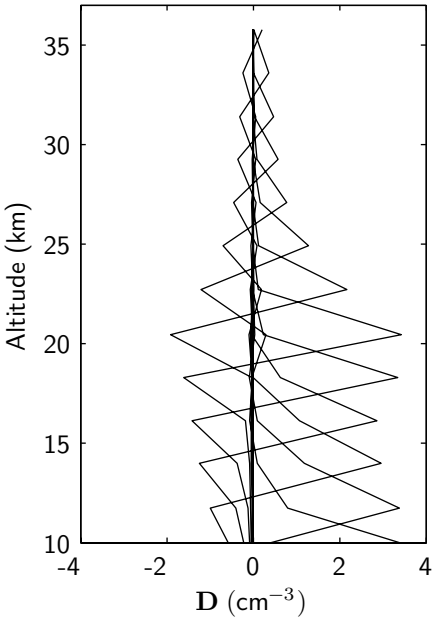

Figure 5.9: The contribution function matrix, **D**, that relates error in the measurements and the forward model to error in the retrieved profile.

the retrieved profile (see Equation 3.30). These columns are plotted in Figure 5.9. Note that **D** is in number density units as the measurement vector is dimensionless (see Equation 5.3).

The main sources of measurement error are detector noise and uncertainty in the altitude registration of each exposure. From analysis of Equation 5.3, random error in the OSIRIS radiances due to detector noise, δI, causes an error in an element of the measurement vector,

$$\delta y_j = \frac{\delta I(h_j, \lambda_l)}{I(h_j, \lambda_l)} + \frac{\delta I(h_r, \lambda_l)}{I(h_r, \lambda_l)} + \frac{\delta I(h_j, \lambda_s)}{I(h_j, \lambda_s)} + \frac{\delta I(h_r, \lambda_s)}{I(h_r, \lambda_s)} \,.$$

The covariance matrix describing the error on the measurement vector is the diagonal matrix, $\mathbf{S}_\varepsilon$, where diagonal elements correspond to the height profile of δy_j^2. From Equation 5.9, the covariance matrix for errors in the retrieval due to measurement

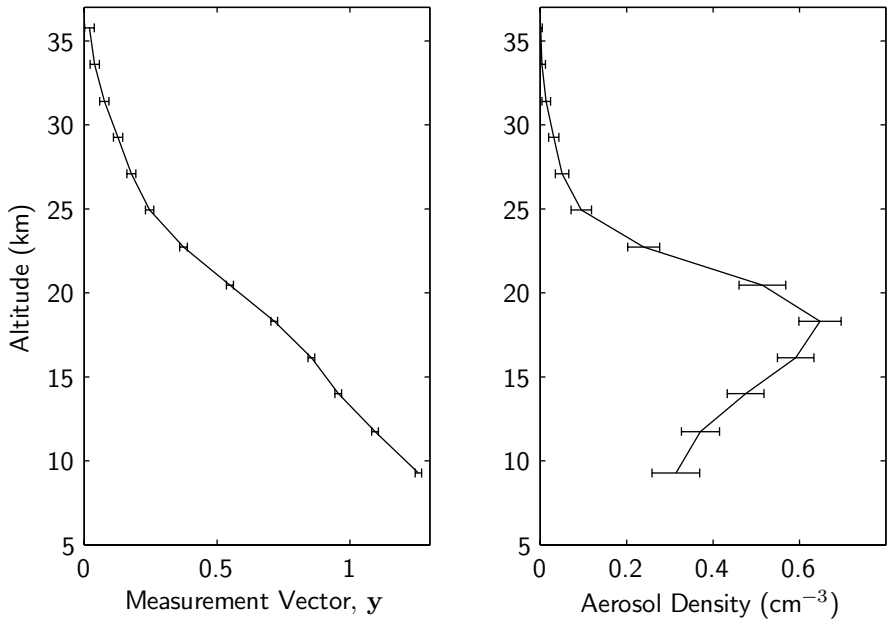

Figure 5.10: The measurement vector constructed using OSIRIS measurements for a typical scan with an error bar calculated from the detector noise. This measurement error results in the error shown here with the retrieved profile.

errors is calculated through the contribution function as

$$\mathbf{S_M} = \mathbf{D}\mathbf{S}_\varepsilon\mathbf{D}^\mathbf{T}. \tag{5.11}$$

Figure 5.10 shows, in the left panel, a plot of the measurement vector constructed from a typical OSIRIS scan with error bars calculated using Equation 5.11. The right panel of the figure shows the retrieved aerosol number density profile for these measurements and the error bar that results from the measurement error as the square root of the diagonal elements of $\mathbf{S_M}$. Near the peak of the number density profile the measurement error due to detector noise is less than 10% and at the uppermost altitudes near 35 km it increases to above 100%.

A similar analysis can be carried out to determine the error in the retrieval that arises due to uncertainty in altitude registration of the limb radiance measurement. Due to the astronomy requirements of the Odin satellite, the OSIRIS instrument

has very good pointing accuracy from on-board star trackers. The error in altitude registration is believed to be better than 200 m at the tangent point. Assuming this uncertainty, a similar analysis can be carried out by recalculating the measurement vector for simulated measurements shifted by 200 m in tangent altitude. This again results in an error in the retrieved density profile of approximately 10% for lower altitudes and larger errors of 30-40% at upper altitudes.

5.5.3 Forward Model Error

Error from the forward model can be separated into two distinct types. The first type includes all of the systematics that arise from the approximate nature of the model. These are extremely difficult to gauge as an accurate characterization requires comparison with an exact model that is, of course, impossible to attain. Often this type of error is used to describe the systematics that arise when it is possible to use the forward model in an approximate sense. While the SASKTRAN model is certainly not perfect, it is utilized to the fullest extent for the aerosol retrieval. Fully spherical geometry is always used. The multiple orders of scattering and the resolution of all spatial integrals in the calculation of the i^{th} order diffuse profile are done to a sufficiently high resolution that effectively no change in the retrieved aerosol profile is observed. No attempt is made to quantify the effect of assuming horizontal homogeneity.

The second type of error in this category, often called forward model *parameter* error, occurs due to uncertainty in the external inputs specified in the radiative transfer. These errors are most certainly systematic in nature. In this approach, the effect of each parameter is investigated independently through numerical perturbation of the simulated retrieval by an amount in the parameter that represents a realistic uncertainty. Total forward model parameter error is the quadrature sum of each term. Surface albedo and neutral density are the most important inputs to the radiative transfer calculation in terms of their uncertainty affecting the aerosol retrieval. Again,

a note about particle sizes must be made. The assumed particle size distribution is, in fact, the most important forward model parameter that affects the aerosol retrieval. However, it is such a crucial aspect of the product that it should not be considered an error, but a crucial assumption. The retrieved density is only meaningful with respect to the assumed size distribution and the two must always be considered together. The accuracy of the product as an optical depth or extinction is shown by the analysis accompanying Figure 5.7.

The reflectivity of the earth, or the albedo, represents a significant unknown in the radiative transfer modelling of the limb signal. The signature of the radiation reflected from the ground in the limb radiance is much stronger at red wavelengths than at shorter UV/blue wavelengths because shorter wavelengths suffer greater extinction along the path length to the ground and back up to the scattering point due to the Rayleigh optical depth. The reddening of the spectrum is similar to the effect of an increased aerosol load and results in potential confusion between the signal attributed to aerosol and the signal caused by an error in the assumed albedo. Even though the altitude normalization of the measurement vector tends to cancel out the majority of the effect of the albedo, it is important that the best estimate of the albedo be used in the forward model as an error in the assumed albedo does effect the solution.

Figure 5.11 is a plot of the same simulated retrieval used to investigate the effect of particle sizes but for different values of albedo. This time the forward model of the measurements is calculated using an albedo of 0.4. The retrievals are performed using assumed albedos of 0.3, 0.4, and 0.5. When the correct albedo is used, i.e. 0.4, the solution is accurate to within the same retrieval error discussed in Sections 5.4 and 5.5.1. With an assumed albedo of 0.3, the algorithm attributes some reddening of the spectrum that is actually due to the upwelling radiation to the aerosol and retrieves an aerosol number density that is systematically about 10% too high at all altitudes. The error is slightly larger at the uppermost and lowermost altitudes. The opposite is true when the assumed albedo is 0.5. That is, some reddening of the spectrum

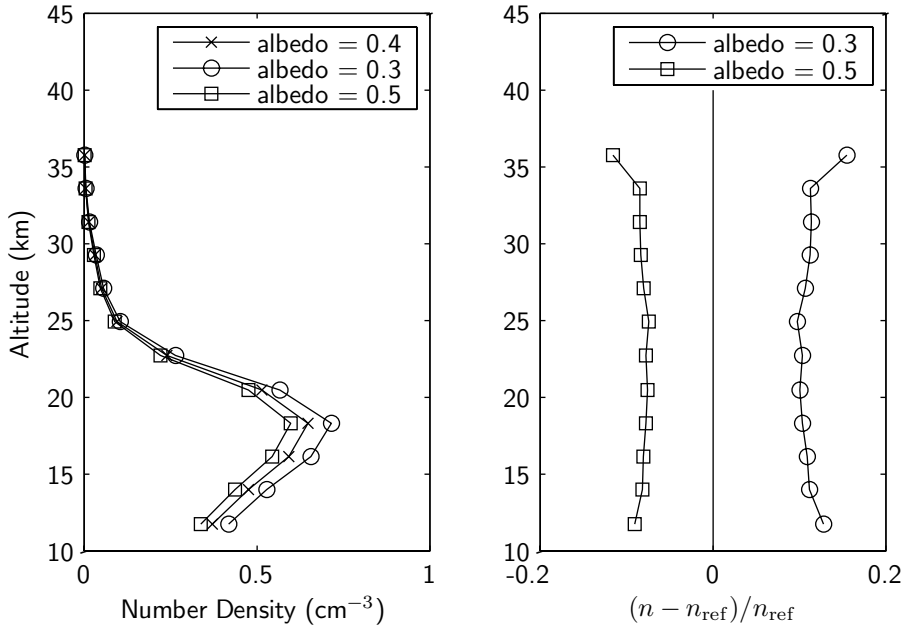

Figure 5.11: Simulations to show the variation in retrieved number density for error in the assumed albedo. One simulated measurement set with an albedo of 0.4 was used to retrieve number densities assuming an albedo 0.3, 0.4 and 0.5.

actually caused by the aerosol profile is attributed to the modelled upwelling radiation that is too large. The retrieved solution is too low, by approximately 10%, with a slightly larger error at the altitude extremes.

Because of these similar effects on the limb radiance, given a set of measurements in the region of the aerosol layer, it is difficult to estimate the albedo. However, at altitudes above the aerosol layer a much more reliable estimate of the albedo can be obtained. For OSIRIS scans, the reference altitude chosen for the measurement vector, h_{ref}=40 km, is also a suitable altitude for determining the albedo. For this work, an estimate of the albedo is obtained, as discussed in Section 4.7.2, by fitting the absolute value of the modelled limb radiance at 700 nm and 40 km altitude to the measured value by adjusting the albedo in the forward model. The reflection from the earth is assumed to be Lambertian with no variation in the horizontal direction and the same value of albedo is used at all wavelengths. Even though these sweeping assumptions must be made, it is better to attempt to estimate the albedo from the

measurements directly than to use a climatology of earth reflectivity because of the frequency of clouds that significantly modify the amount of upwelling radiation from that predicted by clear sky conditions and earth albedo.

The main problem with this technique of albedo estimation is that it relies on the absolute calibration of the instrument. Even so, the calibration is believed to be good to within less than 10% for measurements on the long wavelength side of the order sorter. This is based on extensive in-flight radiative transfer testing and the validation of other retrievals that rely on the absolute calibration (*Lloyd, personal communication*, 2006). This translates into an uncertainty in the estimated albedo of approximately 20%. In order to relate this uncertainty in the albedo to an uncertainty in the retrieved profile, the Rodgers error analysis requires the sensitivity of the forward modelled measurement vector to the albedo parameter, a,

$$\mathbf{K}_a = \frac{\partial \mathbf{F}(\mathbf{x_a}, \tilde{b})}{\partial a}. \tag{5.12}$$

Note that this is a diagonal matrix as the albedo is a single parameter that affects the entire profile of the measurement vector. It is determined numerically by computing the difference between the forward model simulation of the measurement vector due to a small change in albedo. For an albedo covariance matrix, $\mathbf{S}_a$, the forward model parameter error covariance is

$$\mathbf{S_S} = \mathbf{D}\mathbf{K}_a\mathbf{S}_a\mathbf{K}_a^{\mathbf{T}}\mathbf{D}^{\mathbf{T}}. \tag{5.13}$$

Cloud cover is handled in this work as a modification of the ground albedo. By using a tangent altitude near 40 km, the retrieved albedo is effective for the entire scene below and compensates for cloud cover modification to the upwelling radiation. A systematic error arises in the calculation of the multiple scattering component in the forward model because the light path is modified from the cloud free scenario assumed in the model; however this error is small compared to the uncertainty from

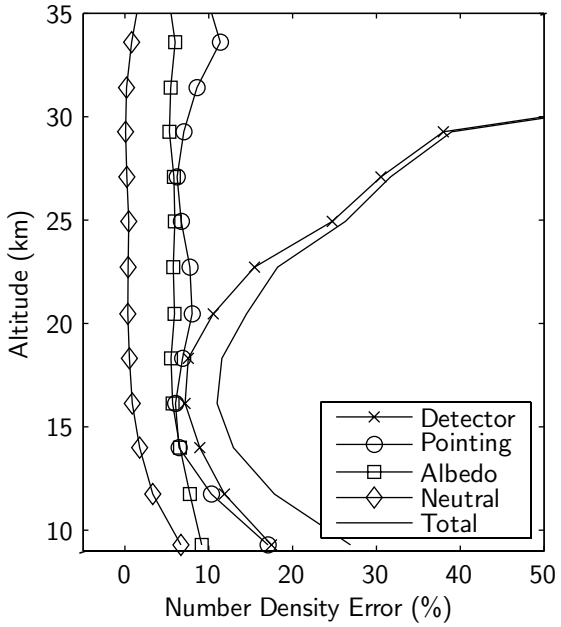

Figure 5.12: Percent error in retrieved number density from measurement error due to detector noise and attitude registration and from forward model parameter error due to uncertainty in albedo and neutral density. The total error is the quadrature sum of all terms.

the absolute calibration.

The same error analysis is applied to the sensitivity to the neutral density with an assumed uncertainty of 1% at all altitudes. Figure 5.12 is a summary plot of the error terms discussed here as a percentage of the aerosol number density. The total error is the quadrature sum of all terms. Especially at upper altitudes, the dominant uncertainty arises from detector noise. This error is very large in terms of percentage, but it is quite constant with altitude in terms of number density. The large percentage error reflects the very small aerosol load at high altitude. For this case, the total error is less than 20% between 12 and 23 km and increases rapidly above 25 km altitude.

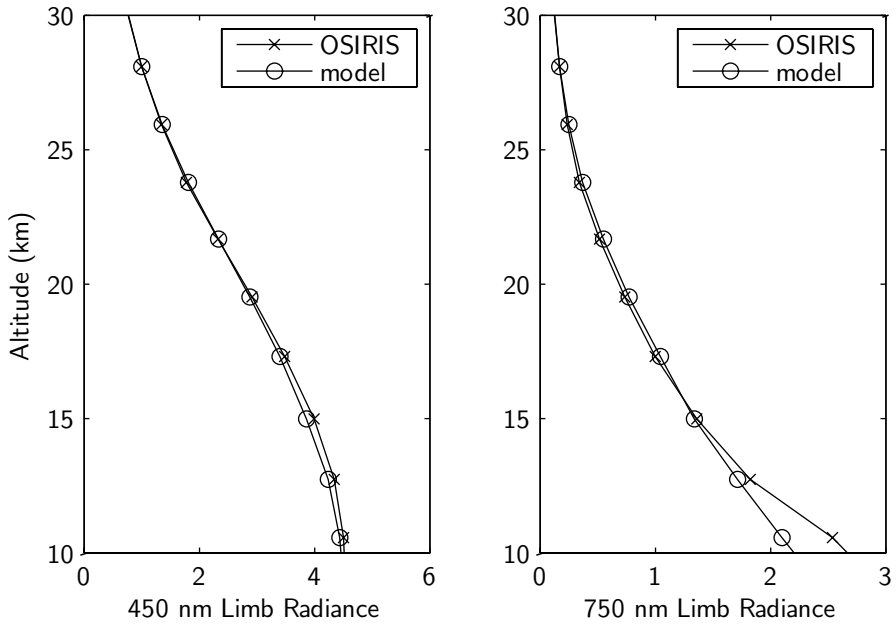

Figure 5.13: OSIRIS limb radiance profiles (units of 10^{13} photons/s/cm^2/sterad) for scan 06432019 at 470 nm and 750 nm and the forward model profiles using an initial guess aerosol density profile and retrieved albedo of 0.57.

5.6 OSIRIS Measurements

This described technique has been applied to the OSIRIS limb scatter measurements. Figure 5.13 is a plot of the limb radiance at 470 nm and 750 nm for a typical mid-latitude OSIRIS scan (06432019, 73° solar zenith angle, 104° solar scattering angle). Also shown on the plot is the limb radiance calculated using SASKTRAN with an initial guess aerosol number density and the particle size distribution used in the simulations. Radiances are shown on an absolute scale and no altitude normalization has been performed. The albedo used in the model for this scan is 0.57 and was determined using the technique described above. The largest difference between the measurements and the model is at 750 nm at altitudes below 15 km. Slight discrepancies are noticeable at both wavelengths for all altitudes up to approximately 25 km where the measurements and the model are in close agreement.

The difference between the measurement and the model is much more apparent

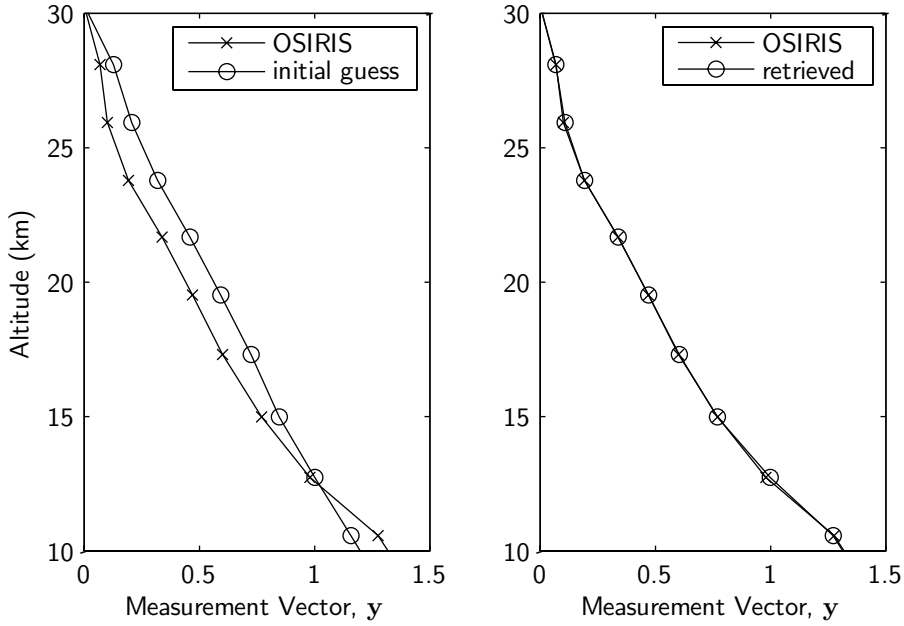

Figure 5.14: The measurement vector, $\mathbf{y}$, constructed using the OSIRIS measurements and with the forward model, before and after the retrieval of the aerosol number density.

in the measurement vector (see Equation 5.3) plotted in the left panel of Figure 5.14 for the OSIRIS measurements and for the forward model with an initial guess aerosol profile. Because the measurement vector is constructed in such a way that the kernel matrix elements are always positive, the difference between the measured and modelled vector can be simplistically interpreted as an overestimate of the initial guess aerosol load at altitudes above 13 km and an underestimate below. This is not strictly true due to the coupling between altitudes from multiple scattering.

The measurement vector calculated with SASKTRAN after 30 iterations of the MART inversion of the aerosol density is shown in the second panel of this figure. The forward model vector is now in very good agreement with the measurements. Convergence to within 2% in the measurement vector, i.e. $|\alpha_i - 1| < 0.02$, at all altitudes is obtained except at the lowest altitudes where a larger difference is still present due to the retrieval error discussed in Section 5.5.1. The retrieved aerosol number density for this OSIRIS scan is shown in Figure 5.15 as an example of a

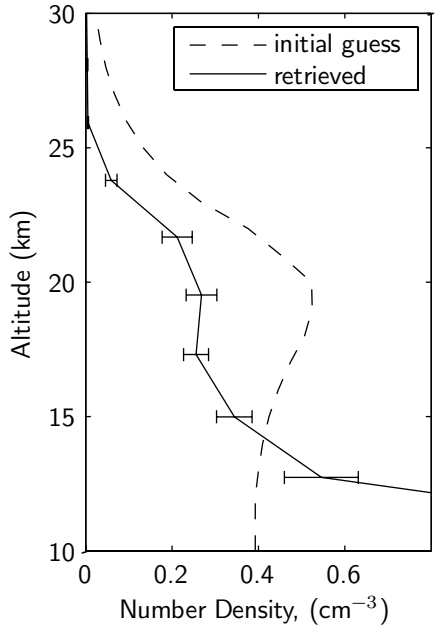

Figure 5.15: The retrieved aerosol number density for scan 06432019 assuming the particle size distribution parameters shown in Figure 5.6.

typical result.

Finally, the modelled calculation of the limb radiance at 470 nm and 750 nm was repeated using the retrieved aerosol density. The results are shown in Figure 5.16. Even though the fitting of the measurement vector does not require the agreement of the limb radiance with the measurements, at both wavelengths, close agreement is obtained at altitudes above 15 km. Below this altitude, some discrepancy remains; however, especially at 750 nm, the modelled profile is much closer to the measurements than it was before the inversion. Because the modelled radiance profiles using the retrieved profile closely match the observations, systematic error due to aerosol in the further retrieval of trace gases such as ozone is significantly reduced by this solution. The good agreement between the modelled and measured radiances produced by this aerosol retrieval is also demonstrated in the plots of the entire spectral range presented in the previous chapter, Figure 4.23.

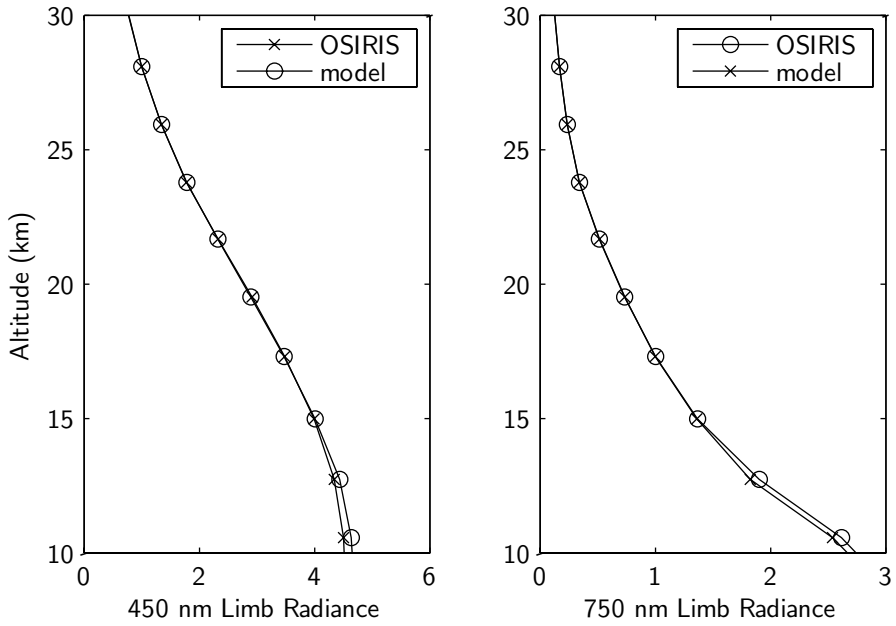

Figure 5.16: OSIRIS limb radiance profiles (units of 10^{13} photons/s/cm^2/sterad) for scan 06432019 and the forward model profiles using the retrieved aerosol number density and retrieved albedo.

5.7 SAGE II/III Comparison

While a complete comparison and validation study of the OSIRIS aerosol profiles is beyond the scope of this thesis, a brief comparison of a single profile measured by OSIRIS, SAGE II and SAGE III is shown in Figure 5.17. The coincidence of the measurement, detailed in Table 1, was chosen based on time and latitude. All three measurements occur within eight hours and less than one degree of latitude. Longitudinal separation is large; however, the relatively constant zonal nature of the stratosphere at mid-latitudes should allow for a reasonable comparison of the three results. As both SAGE II and SAGE III measure 1020 nm extinction, the Mie cross section for the assumed particle size distribution at 1020 nm is used to convert the retrieved OSIRIS aerosol number density to extinction per kilometer for this comparison. It is important to recall that the OSIRIS inversion is performed using the ratio of the radiance at 750 nm to 470 nm. Therefore, using the assumed distribution

Table 5.1: Measurement locations and time on January 5, 2004, for extinction profile comparison between OSIRIS, SAGE II, and SAGE III.

Instrument	Time (UTC)	Latitude (^{o})	Longitude (^{o})
OSIRIS	11:06:05	-37.7	113.2
SAGE II	18:45:37	-37.8	9.2
SAGE III	14:04:32	-37.8	79.1

to relate the results to 1020 nm extinction will depend much more drastically on the uncertainty in the size distribution than the 750 nm extinction where the error arises only from the scattering phase function (see Section 5.4, specifically Figure 5.7).

In the top panels of the figure, the OSIRIS retrieval assumes the *Bingen et al.* (2004) size distribution from 1993 SAGE II data inversions previously used for the simulations in this work. In the lower panels, the retrieval is performed using background layer size distribution parameters consistent with in-situ measurements by *Deshler et al.* (2003) in 2001 (mode radius of 0.08 μm, mode width of 1.6 at all altitudes). It is clear from the figure that the particle size distribution from 1993, which is skewed to larger size by volcanic aerosols, does not retrieve an extinction that is consistent with that directly measured by the occultation instruments. However, the smaller particle sizes that represent the background layer used in the lower panels results in an extinction that agrees reasonably well with SAGE II and SAGE III. The difference between the OSIRIS extinction and the SAGE II extinction between 15 and 30 km altitude is very similar to the difference between SAGE III and SAGE II. At these altitudes, all three instruments agree to within approximately 15%. Below 15 km, OSIRIS extinction is systematically high. This is likely due in part to the growing smoothing error at low altitudes because of the large optical depth.

To interpret this result, it is important to understand the nature of this limb scatter inversion. An occultation instrument measures the extinction directly. In the limb scatter case, however, the measurement is of a type of triple product of the aerosol number density, the cross section, and the scattering phase function and is coupled in altitude through the significant contribution from multiple scattering.

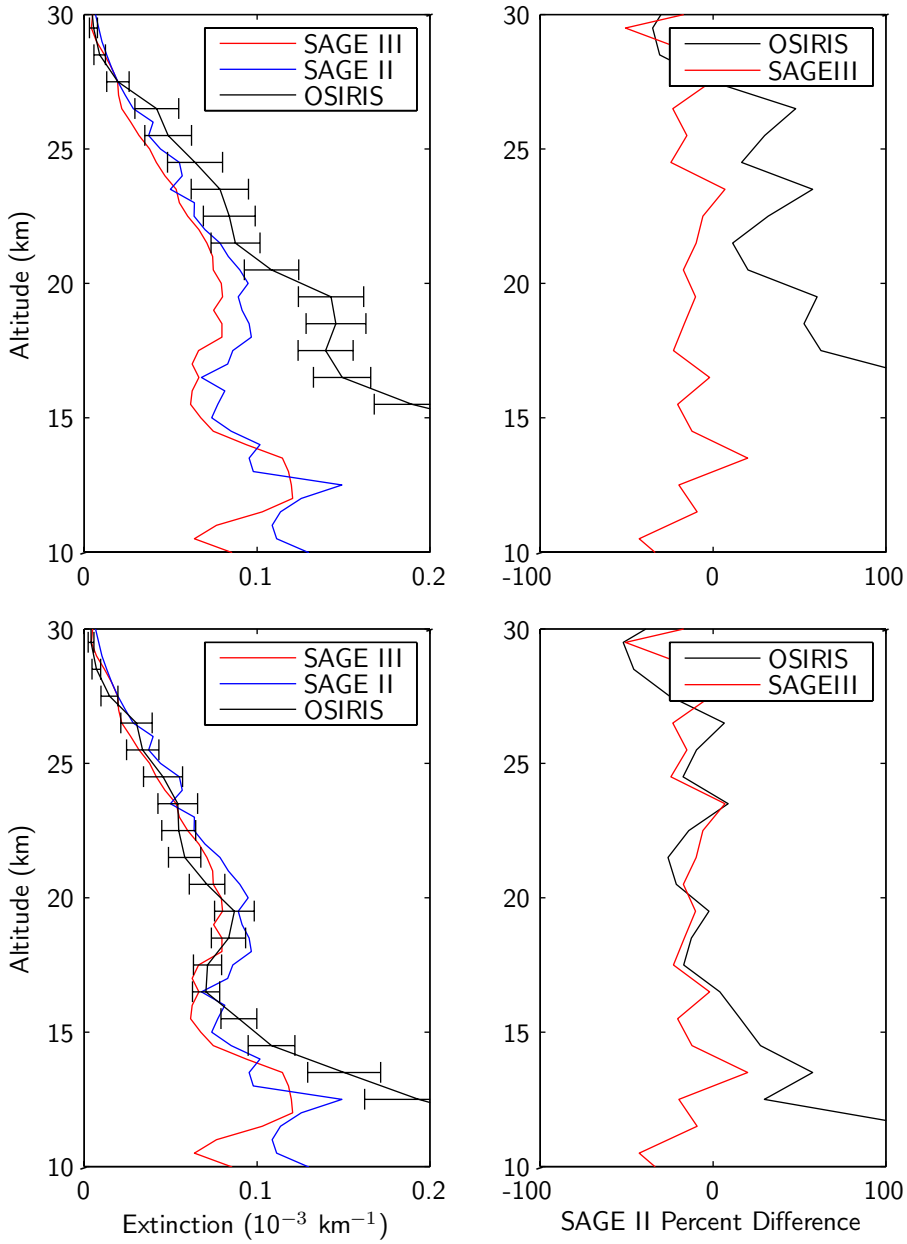

Figure 5.17: A comparison of coincident mid-latitude SAGE II, SAGE III and OSIRIS aerosol 1020 nm extinction profiles. OSIRIS number density is converted to extinction using corresponding Mie cross sections. In the top panels, the OSIRIS retrieval uses the *Bingen et al.* (2004) size distribution used for the modelling work. For the lower panels, the retrieval is performed using background layer size distribution parameters consistent with in-situ measurements by *Deshler et al.* (2003) in 2001 (mode radius of 0.08 micron, mode width of 1.6 at all altitudes). The OSIRIS profile is interpolated to the SAGE II/III 1 km altitude grid.

The specific limb scatter inversion technique presented here retrieves the number density that is most consistent with the measurement vector given the cross section and phase function from an assumed size distribution. That is, one element of this triple product is retrieved; the other two are derived by making an assumption about the particle sizes. In a geophysical sense, it is essential that the retrieved aerosol number density be considered carefully. It must be interpreted not as an absolute quantity, but as the effective amount of particles given the assumed size. Even this must be stated cautiously as it is only truly correct in terms of the aerosol effect on the radiance at the measured wavelengths.

The conversion of the number density to extinction, through multiplication by the assumed cross section, yields a more robust parameter as discussed in Section 5.4. This is especially true at the measured wavelengths, where the uncertainty due to the assumed size distribution is relatively small and arises only from error in the scattering phase function. At other wavelengths, for example the 1020 nm result presented here, systematic error arises directly from the wavelength dependence of the cross section that is calculated from the assumed size distribution. Clearly from Figure 5.17, the volcanically influenced distribution of relatively large particle sizes does not produce consistent limb scatter results across wavelengths. However, the mid-latitude background aerosol distribution, physically much more applicable during this time period, yields a very agreeable result. This distribution was not chosen in any particular way to force a match of the results, it is simply representative of typical background conditions. Therefore, with a reasonable choice of size distribution for the physical conditions, the resulting extinction, even at wavelengths that are not measured, can be reliably retrieved from limb scatter.

As noted previously in Chapter 2, the aerosol layer of this current decade is in a somewhat rare background state of equilibrium. The last volcanic eruption with significant effect on the stratosphere, Mt. Pinatubo, occurred over 25 years ago. This is, in fact, very convenient for the operational products derived from the data set as

the effect of aerosol on the limb radiance is minimized. It also allows for a simple background state size distribution to be used for the aerosol inversion with consistent results across wavelengths as shown in the SAGE comparison above.

In order for a limb scatter aerosol inversion to be useful in a more general case where volcanic influences may be significant, knowledge of the vertical and horizonal variation of the particle size distribution is required. This may be obtained from a separate experiment, for example occultation or in-situ; however, the ability to retrieve the size distribution directly from an appropriate set of limb scatter measurements is possible. This is inherently useful in a geophysical sense and greatly improves the robustness of the retrieved extinction. A technique that can be used to infer an indication of the particle size distribution with the OSIRIS measurements is presented in the following chapter.

5.8 OSIRIS Operational Aerosol Product

This retrieval algorithm and the SASKTRAN forward model were used to create an operational code for installation on the University of Saskatchewan OSIRIS Level 2 Processing Network. The entire OSIRIS mission to date comprises nearly half a million limb radiance scans that are appropriate for species inversion. The beta-version aerosol product, an implementation of the algorithm presented here, has been processed to current date. The majority of the processing occurred over a four month time period during 2006. Recent data is processed automatically within approximately 10 days of measurement. As discussed previously, conversion of the retrieved number density to extinction at the measured wavelength yields a robust result that minimizes the uncertainty due to the assumed size distribution. Therefore, the official product of the OSIRIS aerosol inversion is the 750 nm extinction.

To demonstrate the utility of the OSIRIS limb scatter data set and the beta version of the aerosol product, a global map of retrieved aerosol extinction at 750 nm

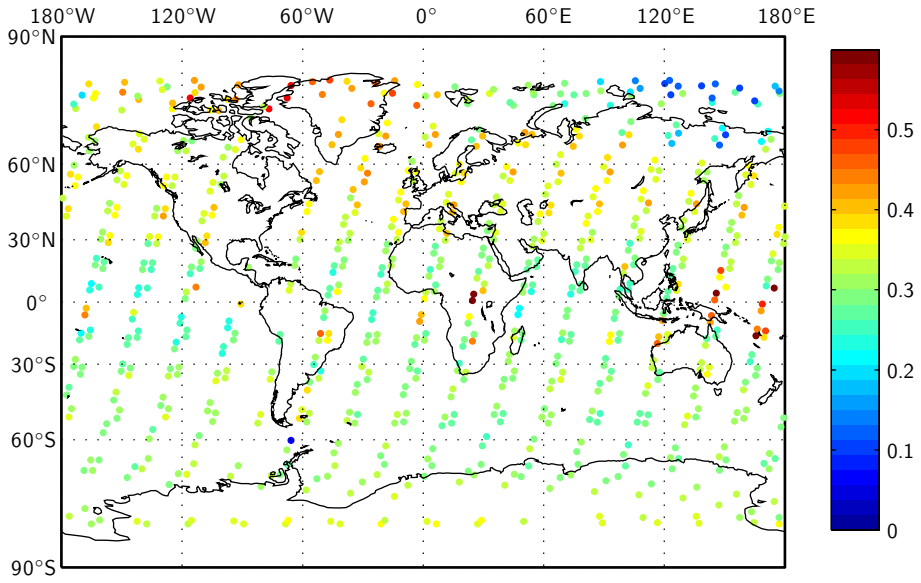

Figure 5.18: Global map of OSIRIS retrieved extinction (10^{-3} km^{-1}) at 750 nm and 20 km altitude for the time period February 27 to March 3, 2006. Measurement location indicated by marker position, extinction indicated by marker colour.

for the one week time period from February 27 to March 3, 2006, at a fixed altitude of 20 km is shown in Figure 5.18. The measurement is indicated by a circular marker, the color of which represents the extinction. During this time period, the Odin orbit plane is very close to the solar terminator affording coverage over the full latitude range in a single orbit. The following figures, 5.19 through 5.22, are a contour plot of the same at altitudes of 20, 24, 28 and 32 km. Although the interpretation of these distributions is not within the scope of this work, it is interesting to note that regions of enhanced aerosol load are detected near the Amazon, central Africa and Indonesia, all of which are often associated with source terms of aerosol through deep convection. A clean northern polar vortex over Siberia, which arises from adiabatic descent of air within the vortex, is also visible in the figure.

5.9 Conclusion

The limb scatter data set, such as that measured by OSIRIS, provides good opportunity for the study and monitoring of stratospheric aerosol. In this chapter, a

technique was presented for the retrieval of aerosol number density for an assumed size distribution at altitudes from 15 to 30 km. Forward model radiances using the retrieved aerosol agree very well with OSIRIS measured limb radiances and will result in an improvement in further retrieval of trace gases from the spectra. The absolute value of retrieved aerosol number density is meaningful only with respect to the assumed particle sizes; however, conversion of the results to extinction provides some insensitivity to the uncertainty in the size. Thus using an in-situ measured size distribution that is representative of the background layer, the OSIRIS 1020 nm extinction compares well with SAGE II and SAGE III throughout the stratosphere for a coincident case investigated here. This relatively high sampling resolution of the OSIRIS limb scatter measurements makes this data set of interest for studying the dynamics of the stratosphere, upper troposphere/lower stratosphere (UTLS) exchange, and for extending the long history of the aerosol data record from the SAGE series of occultation measurements.

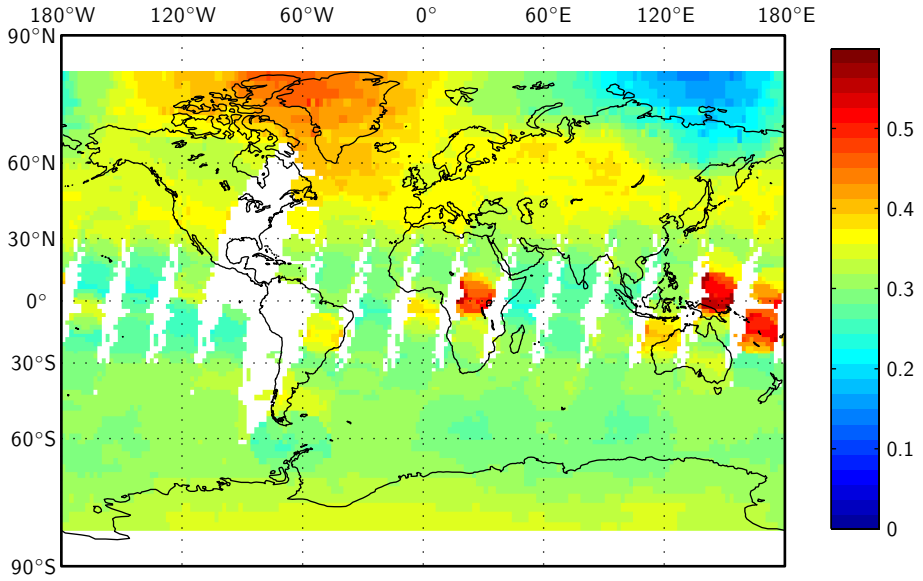

Figure 5.19: Contoured global map of OSIRIS retrieved extinction (10^{-3} km^{-1}) at 750 nm and 20 km altitude for the time period February 27 to March 3, 2006. The color white represents a region of insufficient sampling.

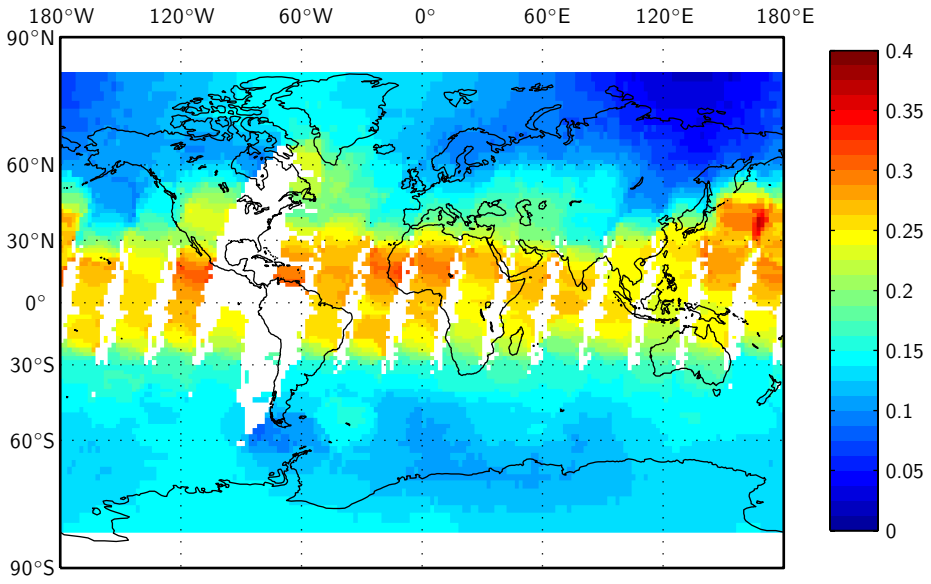

Figure 5.20: Same as Figure 5.19 except at 24 km altitude.

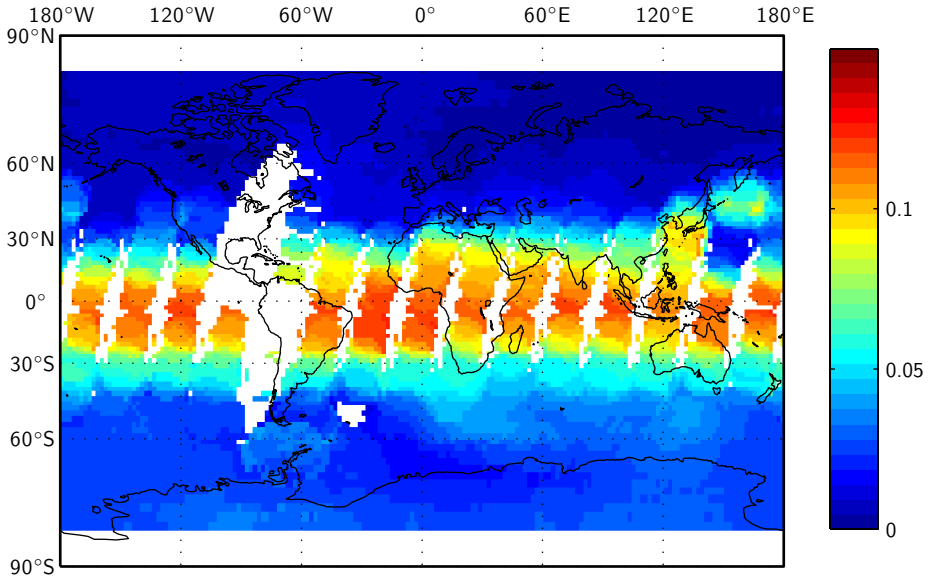

Figure 5.21: Same as Figure 5.19 except at 28 km altitude.

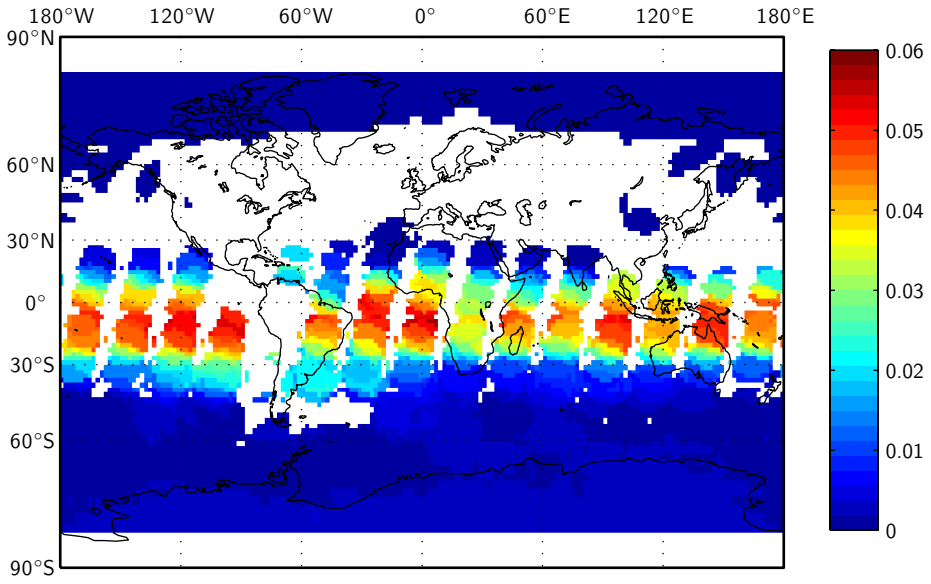

Figure 5.22: Same as Figure 5.19 except at 32 km altitude.

Chapter 6

Aerosol Particle Size Retrieval

6.1 Size Information

The comparisons of the forward modelled radiances with the OSIRIS measured spectra shown in the previous chapters, for example Figures 4.23 and 5.16, demonstrate very good agreement when the retrieved aerosol number density profile is used in the forward model. The aerosol number density corresponds to an assumed size distribution profile, and performing the retrieval for any reasonable choice of aerosol size distribution results in forward model radiances that agree with the measurements in a similar fashion.

This good agreement between the measured and modelled spectra at all wavelengths from 470 to 810 nm, as shown in Figure 4.23, suggests that there is little, or no, further information in the spectra that can be used to infer particle size. With the reduced sensitivity at shorter wavelengths (see Figure 5.1), the aerosol number density retrieval can fit to the measured spectra for any reasonable choice of size distribution. In order to achieve information about the particle size distribution, a wavelength range extending into the near infrared is required. A technique is presented in this chapter that uses measurements from the IRI subsection of the OSIRIS instrument and multiple applications of the aerosol number density retrieval to pro-

vide a measure of the particle size distribution.

The ability to infer size information is based on the difference that arises when the result obtained at the OS wavelengths is extrapolated to the near infrared using the assumed particle size distribution. This effect is shown in the comparison of the 1020 nm extinction inferred from the OSIRIS spectra, with that measured by SAGE II and SAGE III, presented in the previous chapter (see Section 5.7). For the volcanically influenced particle size distribution profile, the retrieved extinction is systematically too high compared to the SAGE II/III measurements, even though the forward model spectra are consistent with the OSIRIS measurements. When a more representative particle distribution profile corresponding to smaller particles was used, the retrieved extinction is in much better agreement with the SAGE II/III measurements. In the same fashion, differences between the forward modelled limb radiance and the OSIRIS measurements in the near infrared are an indication of the fidelity of the assumed particle size.

6.2 IRI Measurements and Model

The three channels of the IRI subsection of the OSIRIS instrument provide vertical images of the band integrated 1260, 1270 and 1530 nm atmospheric limb radiance. The 1260 and 1270 nm channels are designed to measure a molecular emission from excited state molecular oxygen that is formed through ozone photo-dissociation. This emission is quite bright compared to the Rayleigh background so making these channels useless for characterization of scattered sunlight. The 1530 nm channel is designed to measure the hydroxyl Meinel band emission that is also related to ozone photo-chemistry. However, this emission is strongest at night and so results in daytime measurements of limb scattered sunlight are not significantly contaminated by the OH emission.

The IRI measurements are inherently different than the OS measurements in

that the IRI takes a multi-pixel vertical image of the limb, at a vertical resolution of approximately 1 km, with every exposure. The spectrograph has a single line of sight corresponding to a single tangent altitude for each exposure and so requires the satellite to nod to accomplish a vertical limb scan (see Section 2.2.2). The optical axis of the OS and the IRI are co-aligned such that the pixel corresponding to the optical axis of the IRI has the same line of sight as the OS. Therefore, the OS and IRI are best used in combination by using measurements from the optical axis pixel of the IRI over the course of a vertical scan of the OS line of sight rather than an entire IR vertical image from a single exposure. However, as the exposures of the OS and the IRI are not synchronized, an interpolation of the IRI radiance profile to the OS tangent altitudes is required.

Figure 6.1 is a plot of the limb radiance measured by the OS at 750 nm and by the IRI at 1530 nm during scan 06432019. This is the same scan as that chosen to demonstrate the number density retrieval in the previous chapter. The IRI profile is taken from the optical axis pixel and interpolated to the OS tangent altitudes. The SASKTRAN forward model radiances for both wavelengths, calculated using the retrieved aerosol number density, are also shown on the plot. All profiles are normalized to the radiance at 28 km tangent altitude because the absolute calibration of the IRI is uncertain. The size distribution assumed for this retrieval is uniform in altitude and has a mode radius of 0.1 μm and a mode width of 1.5 and is approximately representative of background conditions.

The radiance profiles at 750 nm show good agreement between the OSIRIS measurements and the SASKTRAN prediction. This is essentially the same result as that presented in Figure 5.16 at the end of the previous chapter. However, there is a difference between the 1530 nm radiance predicted by the SASKTRAN model and that measured by the IRI. The modelled radiance is systematically too high for all altitudes below approximately 25 km. This difference is due to an incorrect particle size. The retrieved number density for the assumed particle size results in consis-

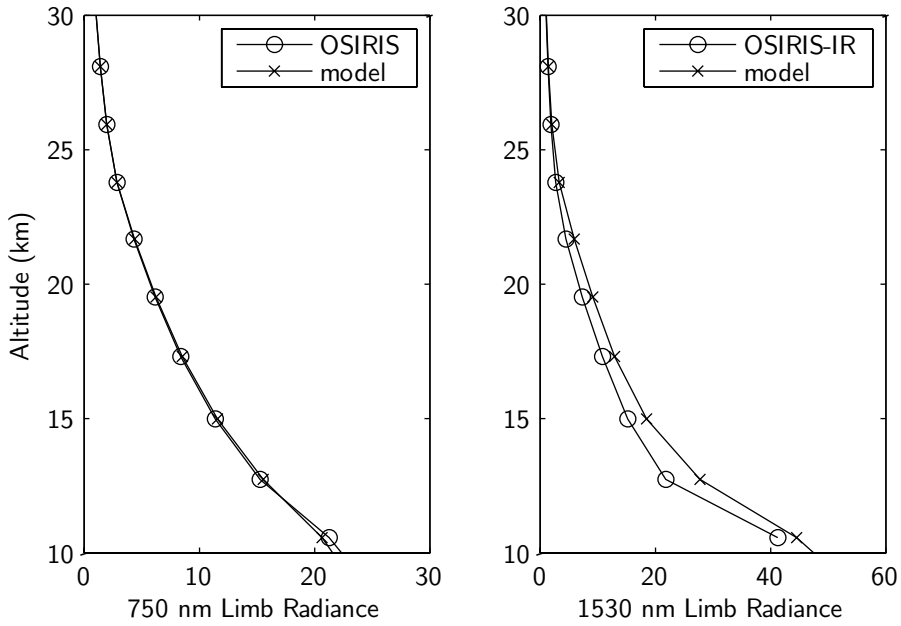

Figure 6.1: Modelled and measured limb radiance spectra (normalized to 28 km) using the retrieved aerosol number density profile assuming a uniform height profile of size distribution parameters ($r_g = 0.1\,\mu$m, $s_g = 1.5$). The IRI measurement is interpolated to OS tangent altitudes.

tent forward model radiances across the visible spectrum, but the uncertainty in the particle size results in a systematic bias in the forward model prediction at 1530 nm.

The Ångström relation introduced in Section 4.2.3 approximates the aerosol scattering cross section by a simple inverse power law, Equation 4.47, where the wavelength dependence of large particles is weaker than that for small particles. In a single scatter sense, the systematically high radiance predicted by SASKTRAN at 1530 nm in Figure 6.1, is due to an aerosol scattering cross section that is too large. Thus, using the Ångström exponent relation, this can be interpreted as a size distribution corresponding to particles that are too large; that is, in order to make the 1530 nm model prediction agree with the IRI measurements, the aerosol cross section must decrease more quickly with wavelength. Indeed, the IRI provides a measured sensitivity to the aerosol particle size distribution parameters. The following sections detail a method, developed as part of this work, to utilize the 1530 nm limb radiance

measurements to infer aerosol size information.

6.3 Methodology

The aerosol number density retrieval described in Chapter 5 characterizes the aerosol scattering with a measurement vector that is constructed using the ratio of the radiance at a long wavelength to the radiance at a shorter wavelength (see Equations 5.2 and 5.3). This technique inherently fits the model prediction to the measurement in a manner that is consistent at both the long and short wavelengths for the scattering cross section and phase function that correspond to the assumed particle size distribution. Choosing a different combination of wavelengths to form the measurement vector ratio yields a retrieved number density that is radiatively consistent with this alternate set of wavelengths. If the assumed particle size distribution accurately describes the actual size distribution, the retrieved number density does not depend on the choice of different wavelength pairs, *i.e.* any wavelength ratio with sufficient sensitivity will yield the same retrieved number density. However, if the assumed size distribution does not accurately describe the real distribution, the retrieved number density will differ so as to best match the measurement vector for the specific wavelength pair. The difference between retrieved number density profiles for different measurement vectors provides an indication of the fidelity of the assumed particle size distribution and information as to how it must be adjusted in order to provide better forward model predictions at all measured wavelengths.

For example, consider the result of performing the aerosol number density retrieval several times, where each time the long wavelength of the ratio is shifted to a longer wavelength. Because of the uncertainty in the assumed particle size distribution, the resulting retrieved number density profiles will differ. However, providing that the assumed distribution is a reasonable approximation of the actual distribution, each of the retrieved number density profiles, $n_i(h)$, can be robustly converted to

extinction at each measured wavelength using the corresponding scattering cross section, $\sigma_{\text{scat}}(\lambda, h)$. For each altitude, h, and for each wavelength, λ_i, the retrieved extinction is,

$$k(\lambda_i, h) = \sigma_{\text{scat}}(\lambda_i, h)\, n_i(h)\,. \tag{6.1}$$

Choosing one of the measured extinction profiles as a reference yields a measured extinction ratio at each of the remaining wavelengths,

$$\tilde{k}_{\text{meas}}(\lambda_i, h) = \frac{k(\lambda_i, h)}{k(\lambda_{\text{ref}}, h)}\,. \tag{6.2}$$

Theoretically, the extinction ratio is simply the cross section ratio because the actual number density, which cannot depend on wavelength, cancels in the ratio,

$$\tilde{k}_{\text{theory}}(\lambda_i, h) = \frac{\sigma_{\text{scat}}(\lambda_i, h)\, n(h)}{\sigma_{\text{scat}}(\lambda_{\text{ref}}, h)\, n(h)} = \frac{\sigma_{\text{scat}}(\lambda_i, h)}{\sigma_{\text{scat}}(\lambda_{\text{ref}}, h)} = \tilde{\sigma}_{\text{scat}}(\lambda_i, h)\,. \tag{6.3}$$

Therefore, the problem is a standard minimization. Using the set of retrieved number density profiles, one for each wavelength, a set of measured extinction ratios is determined (this will be one less than the number of wavelengths as one is used for the reference). Using Mie theory, the extinction ratio can be calculated as the cross section ratio at each measured wavelength for any size distribution. Assuming a single mode log-normal distribution reduces the problem to the determination of the mode radius and the mode width that correspond to a set of extinction ratios that best match the measured extinction ratios.

It should be noted that this method is somewhat ad-hoc in nature as the assumption is made that the retrieved number density can be accurately converted to extinction at the measured wavelength. However, as has been shown previously, variation of the scattering phase function with particle size can result in a systematic error; however, given a reasonable choice of size distribution parameters for the number density retrievals the error in the measured extinction ratios is small (see

Figure 5.7).

6.4 Implementation with OSIRIS

The key to the implementation of the algorithm presented in the previous section is the retrieval of the aerosol number density using the OS radiance at visible wavelengths and using the IRI radiance at 1530 nm. However, there are two main difficulties. The first is that the scattered sunlight limb radiance profiles measured by the IRI are of lower quality than the OS measurements. The IRI measurements contain significantly more noise at upper tangent altitudes than the OS measurements, and are contaminated with a considerable amount of stay light. The second difficulty is simply that the aerosol number density inversion is computationally expensive, and the size parameter retrieval algorithm requires multiple retrievals of the aerosol number density for a single scan, which vastly increases the required processing time. The technique using the OSIRIS data is demonstrated here as a proof of concept. With significant computational resource, it may be possible in the future to implement an operational version for improvement of the official OSIRIS aerosol product.

6.4.1 IRI Noise and Stray Light

Figure 6.2 shows the 1530 nm limb radiance measured by the optical axis pixel of the IRI during scan 06432019 and the forward model result from SASKTRAN for all tangent altitudes from 10 to 60 km. Both profiles are normalized to the radiance at 28 km. In order to demonstrate the difficulties associated with the IRI measurements, the modelled result is shown for zero aerosol load in the atmosphere. The reason for this is explained below.

The total optical depth of the atmosphere is small in the near infrared due to the rapidly decreasing Rayleigh cross section. Therefore, the limb signal is dominated by single scattered sunlight and sunlight reflected from the ground that is then,

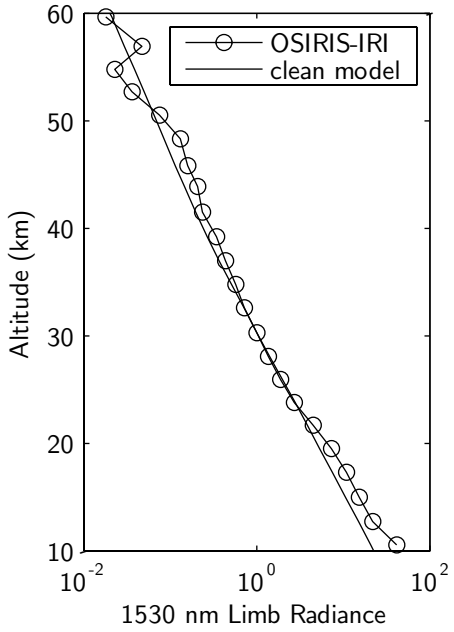

Figure 6.2: The 1530 nm limb radiance profile measured by the IRI optical axis pixel during scan 06432019, and the SASKTRAN model results for an aerosol free atmosphere. Both profiles are normalized at 28 km.

within the atmosphere, scattered directly into the line of sight. Atmospheric multiple scattering is negligible at this wavelength. For this reason, the addition of aerosol contributes to scattering, rather than extinction, along the line of sight for all tangent altitudes. This is shown in Figure 6.3 where the sensitivity, or Jacobian, of the 1530 nm limb radiance with respect to the aerosol density is positive at all altitudes. The 700 nm result is included for reference and shows considerable extinction at lower altitudes where the total optical depth is becoming large and atmospheric multiple scattering is significant. It should be noted that the sensitivity of the 1530 nm limb radiance is much greater than at 700 nm. This is due to the increasing difference between the Rayleigh cross section and the aerosol cross section with wavelength.

Because the 1530 nm aerosol Jacobian is positive at all altitudes, the modelled radiance for the clean atmosphere, shown in Figure 6.2, can only become larger with the addition of any aerosol loading at any altitude. Therefore the model result for the clean atmosphere represents a minimum value for the predicted radiance. Instrument

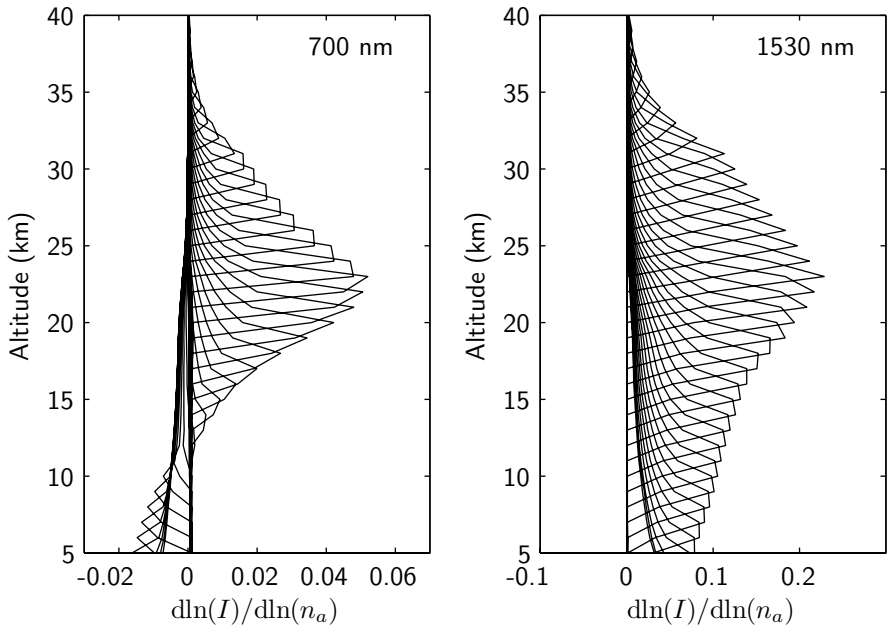

Figure 6.3: The sensitivity of the limb radiance at 700 and 1530 nm to changes in the aerosol number density at each 1 km altitude layer.

noise is clearly noticeable at tangent altitudes above 40 km. At these altitudes, the aerosol number density is insignificant and the measurement should agree with the model result for the clean atmosphere. Therefore, instrument noise should result in measured values that are randomly distributed about the modelled profile. Indeed at the upper altitudes, the measured values oscillate around the modelled profile.

The other important instrumental issue with the IRI measurement is stray light contamination. Much work has already been done to characterize and correct for the stray light in the Level 0 to Level 1 calibration of the measurement; this has resulted in a significant improvement (*Bourassa*, 2003). However, the profile shown in Figure 6.2, which is typical of the IRI Level 1 product, still shows a small, but significant, amount of stray light contamination. Between approximately 30 and 40 km, the measured radiance systematically deviates from the modelled radiance by an amount that increases with altitude. This is due to light from outside the field of view scattering from the instrument baffle system into the detector. The positive

deviation of the measured profile from the modelled profile at altitudes below 25 km is an enhancement in the limb scatter due to aerosol.

6.4.2 Normalization Tangent Altitude

Both the detector noise and the stray light contamination are relatively small components of the total signal at lower altitudes near the aerosol layer, *i.e.* below 30 km tangent altitude. However, the upper altitude normalization of the radiance profiles required for the construction of the measurement vector, Equation 5.2, is critical for the retrieval; it eliminates the need for an absolute calibration and provides some insensitivity to uncertainty in the atmospheric neutral density and the albedo. The aerosol number density retrieval algorithm used with the OS measurements presented in previous chapter adopts a reference tangent altitude near 40 km. This point was chosen as it is well above the aerosol layer and the radiance can be reliably modelled without specific knowledge of the aerosol profile. However, if this convention is applied to the IRI measurements, the stray light and detector noise at 40 km result in a non-physical scenario. This is illustrated with the two profiles shown in Figure 6.2. The modelled profile is calculated for zero aerosol density at all altitudes. If a normalization tangent altitude of 40 km is used, the modelled radiance is too low for tangent altitudes between approximately 20 and 40 km. Because the 1530 nm kernel is strongly positive at all altitudes, each iteration of the inversion then attempts to decrease the aerosol density at these altitudes in order to improve the match between the modelled and measured radiances. However, the aerosol density is already zero at all altitudes. Obviously, the stray light contamination of the IRI measurement leads to an incorrect and non-physical interpretation of the radiance for the aerosol inversion. A similar effect occurs if there is significant detector noise in the measurement at the reference altitude. Even for retrievals using the OS measurements, which have a much higher signal to noise, the detector noise is the dominant term in the error analysis as shown in Section 5.5.

It is possible to choose the reference exposure at a much lower tangent altitude, one within the aerosol layer, but this is undesirable because the sensitivity of the measurement vector to the state parameter is zero at the reference point, and is small for nearby altitudes. In this work it has been found that a measurement in the altitude region between 25 and 35 km can be used for normalization of the IRI measurements. The solution is essentially a matter of finding a point that minimizes both the impact of the instrumental effects and of the decreased sensitivity to aerosol near the normalization altitude.

The altitude where the aerosol density becomes negligibly small depends on the geographic location. For example, stratospheric circulation patterns near equatorial latitudes loft aerosol particles to high altitudes creating a significant aerosol loading that sometimes extends up to almost 40 km; descent in the Brewer-Dobson circulation often results in a very clean stratosphere above 20 km at polar latitudes. The degree of stray light contamination also varies from scan to scan as it depends of the reflectivity of the scene below the instrument line of sight as well as the solar geometry. Thus, the normalization tangent altitude must be determined on a scan by scan basis by attempting to choose the highest tangent altitude possible to maximize sensitivity to the aerosol profile, and the lowest tangent altitude possible to minimize effects from stray light and detector noise. This is accomplished by using the model result for the clean atmosphere in combination with the knowledge that the 1530 nm Jacobian is strongly positive at all altitudes.

For any given scan, the IRI 1530 nm radiance profile and the corresponding model result for zero aerosol density are successively normalized to each exposure in the tangent altitude range from 20 to 40 km. The best reference exposure is chosen as the *highest* tangent altitude that results in a normalized measured profile that is larger than the normalized modelled profile at all tangent altitudes below the reference. For most cases, this method ensures that the reference exposure is not significantly contaminated by stray light. It is possible that an aerosol enhancement at high

altitudes could be lost in the stray light signal. However, it is a fortunate coincidence of the OSIRIS sampling geometry that at equatorial latitudes, where high altitude aerosol is most probable, the stray light contamination is decreased because the solar zenith angle is always near 90 degrees; this effectively limits the upwelling radiation.

When this method is applied to the scan shown in Figure 6.2, the exposure at 28 km tangent altitude is chosen as the reference. Both the measured and modelled profiles are normalized to this point and it can be seen from the figure that 28 km is indeed a reasonable altitude for the normalization point. The enhancement in the measured radiance at the lowest altitude, which corresponds to aerosol scattering, is clearly apparent in the modelled result. Between 25 and 30 km where the aerosol density is negligibly small (see Figure 5.5) the two radiance profiles match very well. Above 30 km, the measured signal begins to deviate from the model result due to stray light and detector noise as previously discussed.

6.4.3 Inversion

Using this same scan (06432019) as representative of a typical case, the aerosol density was independently retrieved with the algorithm presented in Chapter 4 using λ_l=750 nm and using λ_l=1530 nm. In both cases, λ_s=470 nm is used as the short wavelength reference in the measurement vector ratio. The reference tangent altitude, which is the same for both retrievals, is derived from the IRI measurement at h_r=28 km as explained the previous section. A uniform height profile of particle size distribution with mode radius r_g=0.1 μm and mode width s_g=1.5 is assumed in both cases.

Detailed steps of the inversion are not presented here as the procedure is identical to that shown in Chapter 4. However, a brief summary is presented to demonstrate the applicability of the technique at both OS and IRI wavelengths. Figure 6.4 shows the measurement vector from OSIRIS for λ_l=750 nm and for λ_l=1530 nm. The measurement vectors for both retrievals constructed from the SASKTRAN forward model

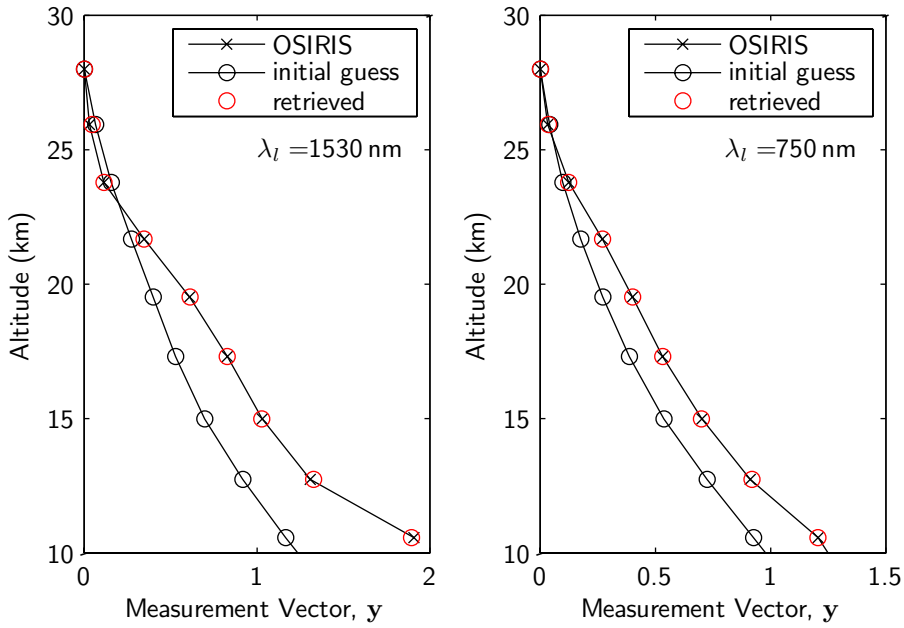

Figure 6.4: The measurement vector using the IR measurements (λ_l=1530 nm) and using the spectrograph measurements (λ_l=750 nm) for OSIRIS scan 06432019. Modelled results after the retrieval, using uniform height profile of size distribution parameters (r_g=0.1 μm, s_g=1.5), converge well in both cases.

radiances, before and after the inversion are included in this figure. The OSIRIS vectors for the two cases are significantly different in both shape and magnitude as there is a large difference in the limb scatter signal at the two wavelengths. While the forward model vector for the initial guess profile varies greatly from the OSIRIS vector, the inversion produces convergence in both cases demonstrated by the close match between elements of the retrieved result and the OSIRIS measurement.

As expected, due to the uncertainty in the assumed size distribution, each of the two inversions yields a different retrieved number density profile. These two profiles are shown in the left panel of Figure 6.5. While not drastically different, the two profiles vary by up to 50% at some altitudes and are not simply separated by a scale factor. Since the assumed size distribution profile is uniform in altitude, it is the shape of the actual size distribution profile that creates this altitude dependent difference between the two number density profiles.

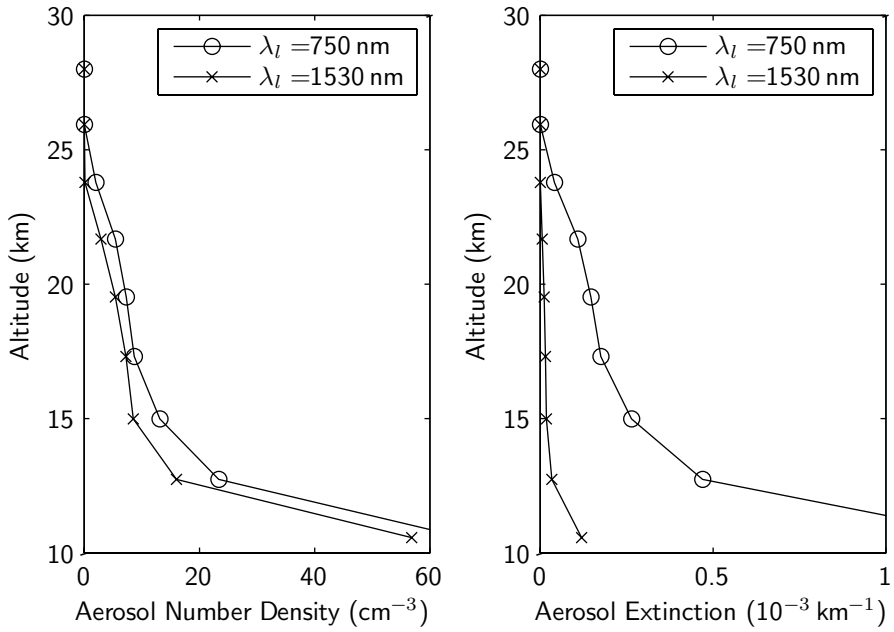

Figure 6.5: The retrieved aerosol density and extinction profiles for the assumed size distribution using λ_l=750 nm and λ_l=1530 nm for OSIRIS scan 06432019.

The determination of the particle size distribution parameters from these two retrievals requires the conversion of the result to the equivalent extinction. The scattering cross section that corresponds to the assumed size distribution is approximately an order of magnitude smaller at the IRI wavelength, i.e. 2.01×10^{-10} cm^2 at 750 nm and 2.10×10^{-11} cm^2 at 1530 nm (see Figure 4.7). Since the assumed distribution is uniform in altitude, the resulting extinction at each wavelength, shown in the right panel of Figure 6.5, is simply a scaling of the retrieved number density by this cross section. These results may be considered a valid retrieval of the extinction in an absolute sense as long as the scattering phase function for the assumed particle size distribution is not significantly in error. As shown by the simulations in Section 5.4 and Figure 5.7, the error in the retrieved extinction at 750 nm is less than the uncertainty in the mode radius. For 1530 nm, the error should be smaller yet as there is little atmospheric multiple scattering.

6.5 Particle Size Distribution Parameters

Following the method detailed in Section 6.3 the extinction ratio is calculated at each altitude. As the extinction is only measured at two wavelengths with OSIRIS in the current scenario, there is a single value of the measured extinction ratio at each altitude. The reference wavelength is chosen as 1530 nm so that Equation 6.2 becomes,

$$\tilde{k}_{\text{meas}}(750\,\text{nm}, h) = \frac{k(750\,\text{nm}, h)}{k(1530\,\text{nm}, h)}\,. \tag{6.4}$$

The theoretical extinction ratio at these wavelengths is pre-calculated using the Mie cross sections for a large set of log-normal mode radius and mode width values. This set can be efficiently searched in a least-squares sense in order to determine the combination of mode radius and mode width that corresponds to an extinction ratio that equals the measured extinction ratio.

Figure 6.6 shows the least-squares R^2 values of the measured extinction ratio at 15 km for a large set of extinction ratios that correspond to a range of mode radius and mode width values. As might be expected, the single measurement of extinction ratio does not provide a unique minimum in the least-square space. A family of solutions arise that show that for each mode width, there is a value of mode radius that corresponds to the extinction ratio that matches the measurement. The minimum, or valley, in the least squares space decreases in mode radius for increasing values of mode width. This is understandable as the distribution must widen as the radius decreases in order to maintain the same effective scattering cross section. This situation of a family of solutions is typical and a similar solution occurs at the other altitudes in this scan. Optimistically, this result provides one of the two assumed parameters. For an assumed mode width, the mode radius is fixed; conversely, for an assumed mode radius, the mode width is fixed.

Figure 6.7 shows the altitude profile of the mode radius that is determined from the least squares fit of the extinction ratio for an assumed mode width of 1.5 and

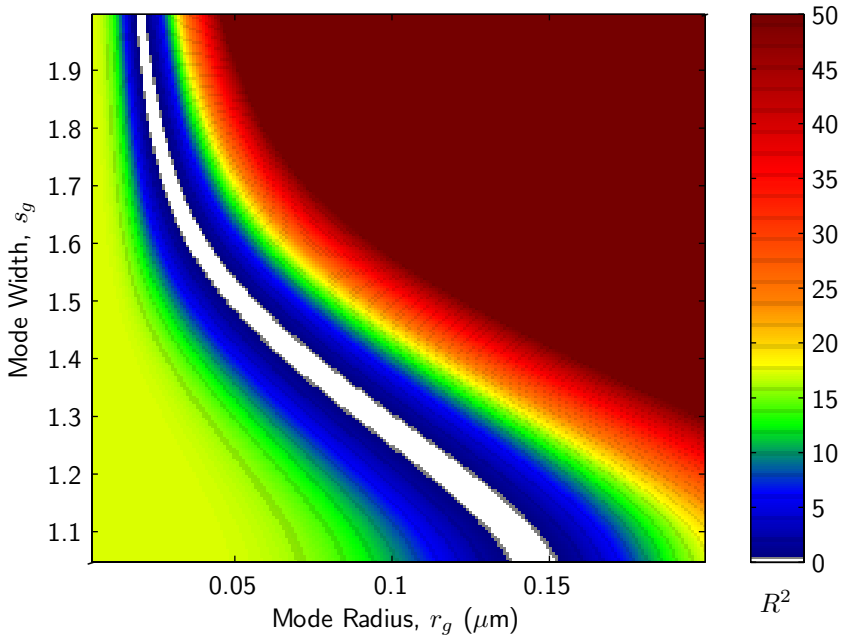

Figure 6.6: Least squares fit values of log-normal size distribution parameters for the retrieved extinction ratio at 15 km altitude for OSIRIS scan 06432019. The family of solutions is a typical result.

a large range of pre-calculated mode radii. The retrieved particles sizes are smaller at all altitudes than the initial guess as predicted from inspection of the original modelled 1530 nm radiance profile.

A non-linear least squares search, such as the Levenberg-Marquart algorithm, may be implemented here. However, because the theoretical extinction ratio can be pre-calculated with ease for a very large number of size distribution parameters and searched with a computation time that is negligible compared to the number density inversion, these more complex minimizations are not required. The value of mode width, 1.5, is somewhat arbitrary; however, it is consistent with the in-situ measurements made by *Deshler et al.* (2003). Above 24 km, the measured 1530 nm aerosol extinction is dominated by noise; therefore, the value at 24 km is extrapolated to higher altitudes.

The error bar shown on the mode radius values in Figure 6.7 is a simple propagation of the aerosol number density error bar (see Section 5.5) through the extinction

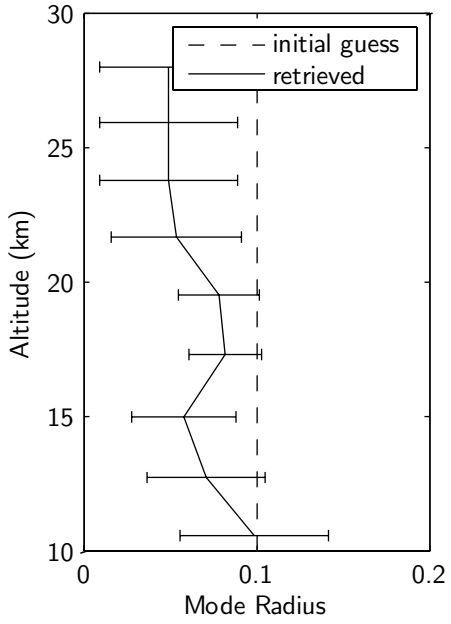

Figure 6.7: Retrieved mode radius for assumed mode width of 1.5 for OSIRIS scan 06432019.

ratio and least squares fit. It represents a worst case and so is quite large. The size of the error is partly due to the fact that the minimum in the least squares space is not narrow; for an assumed mode width, a relatively small change in the extinction ratio produces a large change in the mode radius. Thus detector noise, the dominant term in the error in the retrieved aerosol number density, creates a large error bar in mode radius.

It is generally true that the addition of the extinction ratio at more wavelengths should provide the ability to fix both size distribution parameters. As part of this work, this same method was applied to a set of retrieved extinction profiles at 700, 750, 800 and 1530 nm. Measurementes at wavelengths shorter than 700 nm are significantly contaminated by the O_3 Chappuis absorption band and cannot be used. The best fit size distribution parameters for the set can then be found simply by minimizing,

$$\chi^2(h) = \sum_{i=1}^{3} \left(\frac{\tilde{k}_{\text{meas}}(h, \lambda_i) - \tilde{k}_{\text{theory}}(h, \lambda_i)}{\delta \tilde{k}_{\text{meas}}(h, \lambda_i)} \right)^2 , \tag{6.5}$$

where 1530 nm is again used as the reference wavelength. It was found that the addition of the extinction at these additional wavelengths is unable to further determine the size distribution and almost identical results were found as shown here for the two wavelength retrieval. This is due to the fact that the Mie scattering cross section curve is quite flat over the region from 700 to 800 nm and almost no extra information is gained by incorporating these measurements. In order to fix the size distribution parameters, it is necessary to cover a much broader wavelength range, perhaps with more measurements in the near infrared up to 2 μm.

6.6 Resulting Forward Model Radiances

Using the retrieved mode radius profile shown in Figure 6.7, and a mode width of 1.5, the aerosol number density retrieval algorithm is applied a final time using OS wavelengths only. The resulting number density and size distribution profile are used to forward model the limb radiance profile at both 750 and 1530 nm. The results and the corresponding OSIRIS measurements are shown in Figure 6.8. Again, the modelled 750 nm radiance agrees well with the measurement. This is expected as the 750 nm observations are used in the number density inversion and it has been shown previously that any reasonable choice of size distribution profile results in an aerosol number density profile that provides convergence with the measurement. However, the encouraging result is the improved match in the radiance at 1530 nm. While the profile does not agree as well as that retrieved using 750 nm, the improvement over the result shown in Figure 6.1 shows the improved consistency of the retrieved mode radius with the IR measurement even though 1530 nm is not used in the final number density retrieval.

The remaining discrepancy between the measured and modelled IR radiance profiles is likely due to the uncertainty in the originally retrieved extinction ratio caused by uncertainty in the assumed scattering phase function. It is possible to iterate

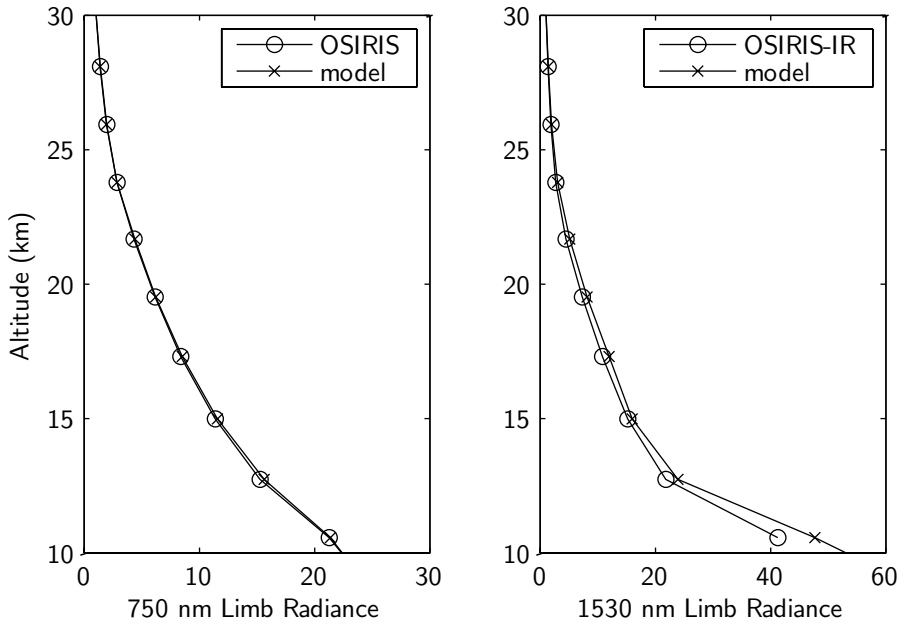

Figure 6.8: Modelled and measured limb radiance spectra (normalized to 28 km) using retrieved aerosol number density profile and retrieved mode radius profile .

this number density/particle size procedure. The current size distribution profile is used to retrieve the extinction ratio once again. This results in a second estimate of the size distribution profile. Errors due to uncertainty in the phase function should decrease with iteration. However, the iteration is very sensitive to structure in the size distribution profile and because the error bar is quite large, the noise quickly becomes unmanageable. Improvements to the technique are conceivable, such as forcing a smooth functional form for the size distribution profile or using a statistical inversion technique to create a weighted average with a smooth *a priori* state. However, this is beyond the scope of the current work and it is concluded that a first estimate of the mode radius for an assumed mode width is possible using the 750 nm/1530 nm extinction ratio.

6.7 SAGE II/III Extinction Re-comparison

To demonstrate further the effectiveness of this retrieval of the size distribution parameters, the coincident measurements of SAGE II, SAGE III and OSIRIS presented in Section 5.7 is revisited. Using the extinction ratio technique described above, the mode radius profile for this OSIRIS measurement has been determined for an assumed mode width of 1.5 at all altitudes. The result is shown in Figure 6.9.

In the same fashion as that discussed previously, this retrieved size distribution profile is used as input to the standard aerosol number density inversion. The result is converted to extinction at 1020 nm using the Mie scattering cross section corresponding to the retrieved size distribution at each altitude. The OSIRIS extinction profile is shown, together with the coincident SAGE II and III measurements, in Figure 6.10 and the result is encouraging. The OSIRIS extinction agrees within the error bar for most points between 15 and 27 km. The agreement is poor below 15 km, yet this is not surprising as it has been shown that the smoothing error begins to dominate at these low altitudes. It may be debatable whether this result is an improvement in the retrieved 1020 nm extinction over the result obtained using the in-situ measured background size distribution shown in Section 5.7. However, two distinct advantages have been gained. Primarily, the potential for the retrieval of height resolved global maps of the aerosol size distribution is, in itself, a geophysically meaningful result. Secondly, the ability of the limb scatter technique to independently infer the size distribution for the robust retrieval of aerosol extinction makes limb scatter measurements much more appealing for long term operational monitoring of the stratosphere.

6.8 Conclusions

While is it difficult to ascertain from this brief comparison the quality of the particle size information that can be obtained from the OSIRIS measurements, an important conclusion can be reached. That is, that by including the near infrared measure-

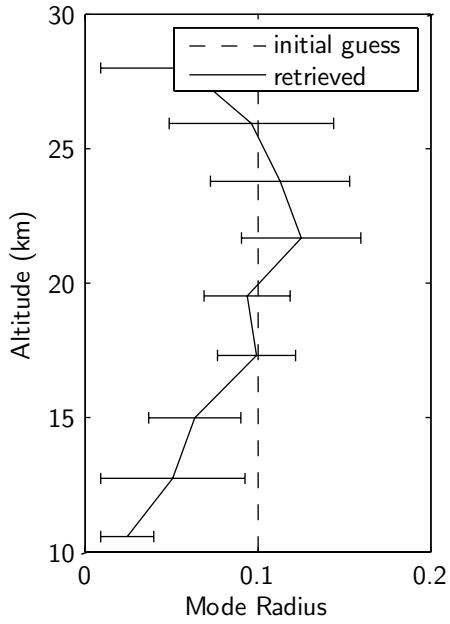

Figure 6.9: OSIRIS retrieved mode radius for assumed mode width of 1.5 for the SAGE II/III coincidence case.

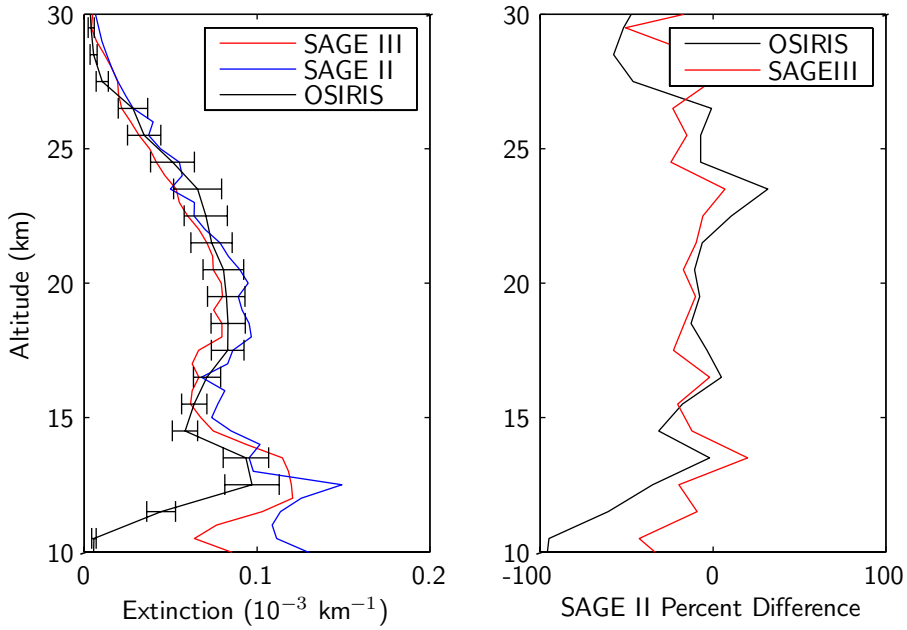

Figure 6.10: SAGE II/III and OSIRIS intercomparison of 1020 nm extinction using retrieved aerosol mode radius profile shown in Figure 6.9. The OSIRIS profile is interpolated to the SAGE II/III 1 km altitude grid.

ments an indication of the particle size distribution can be obtained purely from the consistency of the aerosol inversion in the visible and the near infrared. Indeed, the particle size indicated by the inversion of the OSIRIS measurements for the test case investigated here is, appropriately, much closer to the in-situ measurement values for 2003 than for the volcanically influenced size distribution appropriate for 1993, and provides a reasonable agreement in retrieved extinction at 1020 nm in comparison with SAGE II and III.

Chapter 7

Summary and Outlook

Satellite measurements of limb scattered sunlight spectra have vast potential for monitoring the stratospheric and upper troposphere regions of the atmosphere with global coverage and high vertical resolution. The main goal of the present work has been to demonstrate the feasibility of the retrieval of the vertical profile of stratospheric aerosol properties from these measurements. Generally the limb observations are a weak function of the aerosol load and, because of the presence of significant multiple scattering, an accurate radiative transfer model and a non-linear inversion technique are both required. The present work naturally divides into two parts: the implementation of a spherical radiative transfer code that is suitable for analysis of limb scatter measurements of an aerosol laden atmosphere, and secondly the development of a technique, which uses the model, for the inversion of the measurements to retrieve stratospheric aerosol properties from limb scatter observations, in particular for an operational inversion of the OSIRIS data set.

The radiative transfer model, named SASKTRAN, uses a successive orders of scattering ray tracing algorithm in a spherical shell atmosphere. An altitude dependent log-normal distribution of size parameters, set by the user, determines the aerosol scattering and absorption properties from the Mie particle scattering solution. The addition of other absorbing species with temperature dependent cross sections is

easily configurable. In the implementation of SASKTRAN, the contribution from the first two scattering events of the solar beam, including the non-uniform illumination of the spherical Earth surface, is calculated directly. Third and higher scattering events are efficiently approximated by integrating the previous order source term calculated at locations that correspond to a set of solar zenith angles along the instrument line of sight. At relatively small solar zenith angles (less than 70°), a single estimation of the source term at the tangent point of the line of sight is sufficient; when the sun is closer to the horizon, the multiple scatter source term must be estimated at approximately every one degree of solar zenith angle along the instrument line of sight to avoid a significant systematic error, especially at low tangent altitudes. A specialized numerical integration over the unit sphere that exploits the characteristics of the incoming radiance distribution is used to optimize the ray tracing without compromising the accuracy of the integration.

The SASKTRAN results compare very well with the OSIRIS measurements at wavelengths from 310–810 nm and at altitudes from 15–40 km for a best fit scene albedo and retrieved profiles of ozone, stratospheric aerosol and nitrogen dioxide. Comparisons of SASKTRAN results with other radiative transfer models show that the SASKTRAN results are in better agreement with a reference model than the approximate spherical models and are generally as good as the fully spherical models. In general, only the approximate spherical models are feasible for use in an inversion as the fully spherical models require too much computational resource. The implementation of the SASKTRAN algorithm using caching and multi-threading computational techniques has resulted in an efficient code that is suitable for the improved operational inversion of limb scatter satellite data with modern desktop computers.

An algorithm for the retrieval of stratospheric aerosol profiles that utilizes a one dimensional version of the MART non-linear relaxation has also been developed. It was shown that an altitude normalized spectral ratio can be used with the MART to retrieve the aerosol number density profile for an assumed height profile of the aerosol

particle size distribution. The use of the spectral ratio maximizes the sensitivity of the limb measurement to aerosol scattering and minimizes the effects of upwelling radiation from ground albedo and uncertainties in the neutral density profile. As an algebraic derivation of the error terms in the standard formal analysis is not possible for a relaxation type inversion like the MART, a method of numerical perturbation was used to show that the dominant error term in the retrieval is the measurement noise. The retrieved aerosol profile from OSIRIS observations improves the match between the measured and forward modelled limb radiance, which is important for the further retrieval of trace gases from the measured spectra.

The absolute value of the retrieved aerosol number density depends on the assumed particle size distribution. However, for an in-situ measured size distribution that is representative of the background layer for 2002, the retrieved OSIRIS extinction, extrapolated to 1020 nm, agrees to within 15% throughout the lower stratosphere with coincident SAGE II and SAGE III measurements. When the OSIRIS IRI measurements are also included in the analysis, an indication of the particle size distribution can be obtained from the consistency of the aerosol inversion in the visible and the near infrared red spectral regions. If a log-normal distribution is assumed, a family of solutions are obtained that provide a measure of one of the two distribution parameters.

The future direction of the present work focuses on upgrades to the SASKTRAN model and improvements to the aerosol inversion. An immediate improvement to SASKTRAN that may improve the current state of the OSIRIS aerosol product is the coupling of the third and higher order source terms across solar zenith angles. Also, the inclusion of a pseudo-specular reflection of sunlight from the rough ocean surface is a potentially important improvement to the model as glint from the ocean surface may be a significant term in the limb radiance (*Degenstein et al.*, 2007). In addition, if there is sufficient processing time available, an operational version of the particle size inversion with the IRI measurements is a foreseeable goal.

The identification and inversion of other aerosol types from limb scattered sunlight spectra is a next step in this field. In the present work, the focus has been on stratospheric sulphate spherical droplets. In order to use this technique to study the increasingly important processes of the upper troposphere and lower stratosphere (UTLS) region, it is critical that there be an improved understanding of the effect of cloud particles, smoke, and dust. For SASKTRAN, this will require the ability to model the scattering and absorption properties of these aerosols, and it may also require special adaptations of the multiple scatter ray tracing algorithm, especially in the presence of clouds. OSIRIS does not measure the polarization of the limb radiance directly, but for future satellite missions focused on the science of the UTLS, it is important to investigate the potential of including this capability in the SASKTRAN model. In terms of the inversion, the inclusion of other aerosol types will require an extensive study of the identification and quantification of their limb scatter spectral signature. Also with improved imaging technology that will be available in future missions it may be possible, and very advantageous, to attempt a tomographic two dimensional inversion of the aerosol properties from limb scatter sunlight spectra.

It is readily apparent that the present thesis work has shown the value of limb observations and their potential to improve our understanding of the atmospheric state.

Bibliography

Anderson, G. P., A. Berk, P. K. Acharya, M. W. Matthew, L. S. Bernstein, J. H. Chetwynd, H. Dothe, S. M. Adler-Golden, A. J. Ratkowski, G. W. Felde, J. A. Gardner, M. L. Hoke, S. C. Richtsmeier, B. Pukall, J. B. Mello, and L. S. Jeong (1999), MODTRAN4: radiative transfer modeling for remote sensing, *Proc. SPIE*, *3866*, 2–10.

Ångström, A. (1964), The parameters of atmospheric turbidity, *Tellus*, *16*, 64–75.

Barth, C. A., D. W. Rusch, R. J. Thomas, G. H. Mount, G. J. Rottman, G. E. Thomas, R. W. Sanders, and G. M. Lawrence (1983), Solar Mesosphere Explorer: Scientific objectives and results, *Geophys. Res. Lett.*, *10*, 237–240.

Bates, D. R. (1984), Rayleigh scattering by air, *Planet. Space Sci.*, *32*, 785–790, doi:10.1016/0032-0633(84)90102-8.

Bauman, J. J., P. B. Russell, M. A. Geller, and P. Hamill (2003), A stratospheric aerosol climatology from SAGE II and CLAES measurements: 1. Methodology, *J. Geophys. Res.*, *108*, 6–10, doi:10.1029/2002JD002992.

Bhartia, P. K., R. D. McPeters, C. L. Mateer, L. E. Flynn, and C. Wellemeyer (1996), Algorithm for the estimation of vertical ozone profiles from the backscattered ultraviolet technique, *J. Geophys. Res.*, *101*, 18,793–18,806, doi:10.1029/96JD01165.

Bingen, C., D. Fussen, and F. Vanhellemont (2004), A global climatology of stratospheric aerosol size distribution parameters derived from SAGE II data over the period 1984-2000: 1. Methodology and climatological observations, *J. Geophys. Res.*, *109*, 1–10, doi:10.1029/2003JD003518.

Bogumil, K., J. Orphal, and J. P. Burrows (2000), Temperature dependent absorption cross sections of O_3, NO_2, and other atmospheric trace gases measured with the SCIAMACHY spectrometer, *Proc. ERS - Symposium, Goteborg, Sweden.*

Bourassa, A. (2003), The characterization and calibration of the OSIRIS Infrared Imager, Thesis, University of Saskatchewan.

Bovensmann, H., J. P. Burrows, M. Buchwitz, J. Frerick, S. Noël, V. V. Rozanov, K. V. Chance, and A. P. H. Goede (1999), SCIAMACHY: Mission objectives and measurement modes, *J. Atmos. Sci.*, *56*, 127–150.

Brock, C. A., P. Hamill, J. C. Wilson, H. H. Jonsson, and K. R. Chan (1995), Particle formation in the upper tropical troposphere: A source of nuclei for the stratospheric aerosol, *Science*, *270*, 1650–1653.

Burrows, J. P., A. Dehn, B. Deters, S. Himmelmann, A. Richter, S. Voigt, and J. Orphal (1998), Atmospheric remote-sensing reference data from GOME: 1. temperature dependent absorption cross sections of NO_2 in the 231794 nm range, *J. Quant. Spectrosc. Radiat. Trans.*, *60*, 1025–1031.

Burrows, J. P., A. Richter, A. Dehn, B. Deters, S. Himmelmann, S. Voigt, and J. Orphal (1999), Atmospheric remote sensing reference data from GOME: 2. temperature dependent absorption cross sections of O_3 in the 234-794 nm range, *J. Quant. Spectrosc. Radiat. Trans.*, *61*, 509–519.

Chahine, M. T. (1970), Inverse problems in radiative transfer: Determination of atmospheric parameters, *J. Atmos. Sci.*, *27*, 960–967.

Chahine, M. T. (1972), A general relaxation method for inverse solution of the full radiative transfer equation, *J. Atmos. Sci.*, *29*, 741–747.

Chandrasekhar, S. (1960), *Radiative Transfer*, Dover, New York.

Collins, D. G., W. G. Blattner, M. B. Wells, and H. G. Horak (1972), Backward Monte Carlo calculations of the polarization characteristics of the radiation emerging from spherical-shell atmospheres, *Appl. Opt.*, *11*, 2684–2696.

Crutzen, P. J. (1970), The influence of nitrogen oxide on the atmospheric ozone content, *Q. J. R. Meteorol. Soc.*, *96*, 320–327.

Crutzen, P. J. (1976), The possible importance of CSO for the sulfate layer of the stratosphere, *Geophys. Res. Lett.*, *3*, 73–76.

Degenstein, D. A., E. J. Llewellyn, and N. D. Lloyd (2003), Volume emission rate tomography from a satellite platform, *Appl. Opt.*, *42*, 1441–1450.

Degenstein, D. A., E. J. Llewellyn, and N. D. Lloyd (2004), Tomographic retrieval of the oxygen infrared atmospheric band with the OSIRIS infrared imager, *Can. J. Phys.*, *82*, 501–515.

Degenstein, D. A., A. E. Bourassa, E. J. Llewellyn, and N. D. Lloyd (2007), The impact of sea-glint upon limb radiance, *Can. J. Phys.,* submitted.

Deshler, T., M. E. Hervig, D. J. Hofmann, J. M. Rosen, and J. B. Liley (2003), Thirty years of in-situ stratospheric aerosol size distribution measurements from Laramie, Wyoming (41 N), using balloon-borne instruments, *J. Geophys. Res.*, *108*, 1–13, doi:10.1029/2002JD002514.

Deshler, T., R. Anderson-Sprecher, H. Jäger, J. Barnes, D. J. Hofmann, B. Clemesha, D. Simonich, M. Osborn, R. G. Grainger, and S. Godin-Beekmann (2006), Trends in the nonvolcanic component of stratospheric aerosol over the period 1971-2004, *J. Geophys. Res.*, *111*, D01,201, doi:10.1029/2005JD006089.

Dittman, M. G., E. Ramberg, M. Chrisp, J. V. Rodriguez, A. L. Sparks, N. H. Zaun, P. Hendershot, T. Dixon, R. H. Philbrick, and D. Wasinger (2002), Nadir ultraviolet imaging spectrometer for the NPOESS Ozone Mapping and Profiler Suite (OMPS), *Earth Observing Systems VII, Proc. SPIE*, vol. 4814, 111–119.

Dutton, E. G., and J. R. Christy (1992), Solar radiative forcing at selected locations and evidence for global lower tropospheric cooling following the eruptions of El Chichon and Pinatubo, *Geophys. Res. Lett.*, *19*, 2313–2316.

Erle, F., A. Grendel, D. Perner, U. Platt, and K. Pfeilsticker (1998), Evidence of heterogeneous bromine chemistry on cold stratospheric sulphate aerosols, *Geophys. Res. Lett.*, *25*, 4329–4332, doi:10.1029/1998GL900087.

Flittner, D. E., P. K. Bhartia, and B. M. Herman (2000), O_3 profiles retrieved from limb scatter measurements: Theory, *Geophys. Res. Lett.*, *27*, 2601–2604, doi: 10.1029/1999GL011343.

Glaccum, W., R. L. Lucke, R. M. Bevilacqua, E. P. Shettle, J. S. Hornstein, D. T. Chen, J. D. Lumpe, S. S. Krigman, D. J. Debrestian, M. D. Fromm, F. Dalaudier, E. Chassefière, C. Deniel, C. E. Randall, D. W. Rusch, J. J. Olivero, C. Brogniez, J. Lenoble, and R. Kremer (1996), The Polar Ozone and Aerosol Measurement instrument, *J. Geophys. Res.*, *101*, 14,479–14,488, doi:10.1029/96JD00576.

Goodman, J., K. C. Snetsinger, G. V. Ferry, N. H. Farlow, H. Y. Lem, and D. M. Hayes (1982), Altitude variations in stratospheric aerosols of a tropical region, *Geophys. Res. Lett.*, *9*, 609–612.

Griffioen, E., and L. Oikarinen (2000), LIMBTRAN: A pseudo three-dimensional radiative transfer model for the limb-viewing imager OSIRIS on the Odin satellite, *J. Geophys. Res.*, *105*, 29,717–29,730, doi:10.1029/2000JD900566.

Gruner, P., and H. Kleinert (1927), Die dammerungerscheinen, *In iProblemeder Kosmischen Physik*, *10*, 1–113.

Haley, C. S., S. M. Brohede, C. E. Sioris, E. Griffioen, D. P. Murtagh, I. C. McDade, P. Eriksson, E. J. Llewellyn, A. Bazureau, and F. Goutail (2004), Retrieval of stratospheric O_3 and NO_2 profiles from Odin Optical Spectrograph and Infrared Imager System (OSIRIS) limb-scattered sunlight measurements, *J. Geophys. Res.*, *109*, 16,303+, doi:10.1029/2004JD004588.

Hamill, P., E. J. Jensen, P. B. Russell, and J. J. Bauman (1997), The life cycle of stratospheric aerosol particles., *B. Am. Meteorol. Soc.*, *78*, 1395–1410.

Hansen, J., M. Sato, L. Nazarenko, R. Ruedy, A. Lacis, D. Koch, I. Tegen, T. Hall, D. Shindell, B. Santer, P. Stone, T. Novakov, L. Thomason, R. Wang, Y. Wang, D. Jacob, S. Hollandsworth, L. Bishop, J. Logan, A. Thompson, R. Stolarski, J. Lean, R. Willson, S. Levitus, J. Antonov, N. Rayner, D. Parker, and J. Christy (2002), Climate forcings in Goddard Institute for Space Studies SI2000 simulations, *J. Geophys. Res.*, *107*, 4347+, doi:10.1029/2001JD001143.

Hansen, J. E., and L. D. Travis (1974), Light scattering in planetary atmospheres, *Space Sci. Rev.*, *16*, 527–610.

Heath, D. F., A. J. Krueger, H. A. Roeder, and B. D. Henderson (1975), The Solar Backscatter Ultraviolet and Total Ozone Mapping Spectrometer (SBUV/TOMS) for Nimbus G, *Opt. Eng.*, *14*, 323–331.

Henyey, L. G., and J. L. Greenstein (1941), Diffuse radiation in the Galaxy, *Astrophys. J.*, *93*, 70–83.

Herman, B. M., A. Ben-David, and K. J. Thome (1994), Numerical technique for solving the radiative transfer equation for a spherical shell atmosphere, *Appl. Opt.*, *33*, 1760–1770.

Herman, B. M., D. E. Flittner, T. R. Caudill, K. J. Thome, and A. Ben-David (1995), Comparison of the Gauss-Seidel spherical polarized radiative transfer code with other radiative transfer codes, *Appl. Opt.*, *34*, 4563+.

Hicks, B. B., R. S. Artz, T. P. Meyers, and R. P. Hosker (2002), Trends in eastern U.S. sulfur air quality from the Atmospheric Integrated Research Monitoring Network, *J. Geophys. Res.*, *107*, 4143+, doi:10.1029/2000JD000165.

Holton, J. R., P. H. Haynes, M. E. McIntyre, A. R. Douglass, R. B. Rood, and L. Pfister (1995), Stratosphere-troposphere exchange, *Rev. Geophys.*, *33*, 403–439.

Hunten, D. M., R. P. Turco, and O. B. Toon (1980), Smoke and dust particles of meteoric origin in the mesosphere and stratosphere, *J. Atmos. Sci.*, *37*, 1342–1357.

Junge, C. E., C. W. Chagnon, and J. E. Manson (1961), Stratospheric aerosols, *J. Atmos. Sci.*, *18*, 81–108.

Kaiser, J., C. von Savigny, K. Eichmann, S. Noel, H. Bovensmann, and J. Burrows (2004), Satellite-pointing retrieval from atmospheric limb-scattering of solar UV-B radiation, *Can. J. Phys.*, *82*, 1041–1052, doi:10.1139/P04-071.

Kaiser, J. W., and J. P. Burrows (2003), Fast weighting functions for retrievals from limb scattering measurements, *J. Quant. Spectrosc. Radiat. Trans.*, *77*, 273–283.

Kettle, A. J., U. Kuhn, M. von Hobe, J. Kesselmeier, and M. O. Andreae (2002), Global budget of atmospheric carbonyl sulfide: Temporal and spatial variations of the dominant sources and sinks, *J. Geophys. Res.*, *107*, 4658+, doi: 10.1029/2002JD002187.

King, L. (1923), On the complex anisotropic molecule in relation to the dispersion and scattering of light, *Proc. Roy. Soc.*, *104*, 333–357.

Labitzke, K., and M. P. McCormick (1992), Stratospheric temperature increases due to Pinatubo aerosols, *Geophys. Res. Lett.*, *19*, 207–210.

Lacis, A., J. Hansen, and M. Sato (1992), Climate forcing by stratospheric aerosols, *Geophys. Res. Lett.*, *19*, 1607–1610.

Le Treut, H., M. Forichon, O. Boucher, and Z.-X. Li (1998), Sulfate aerosol indirect effect and CO_2 greenhouse forcing: Equilibrium response of the LMD GCM and associated cloud feedbacks, *J. Clim.*, *11*, 1673–1684.

Lenoble, J., and Z. Sekera (1961), Equation of radiative transfer in a planetary spherical atmosphere, *Proc. Nat. Ac. Sci.*, *47*, 372–378.

Llewellyn, E. J., N. D. Lloyd, D. A. Degenstein, R. L. Gattinger, S. V. Petelina, A. E. Bourassa, J. T. Wiensz, E. V. Ivanov, I. C. McDade, B. H. Solheim, J. C. McConnell, C. S. Haley, C. von Savigny, C. E. Sioris, C. A. McLinden, E. Griffioen, J. Kaminski, W. F. Evans, E. Puckrin, K. Strong, V. Wehrle, R. H. Hum, D. J. W. Kendall, J. Matsushita, D. P. Murtagh, S. Brohede, J. Stegman, G. Witt, G. Barnes, W. F. Payne, L. Piché, K. Smith, G. Warshaw, D.-L. Deslauniers, P. Marchand, E. H. Richardson, R. A. King, I. Wevers, W. McCreath, E. Kyrölä, L. Oikarinen, G. W. Leppelmeier, H. Auvinen, G. Mégie, A. Hauchecorne, F. Lefèvre, J. de La Nöe, P. Ricaud, U. Frisk, F. Sjoberg, F. von Schéele, and L. Nordh (2004), The OSIRIS instrument on the Odin spacecraft, *Can. J. Phys.*, *82*, 411–422.

Lloyd, N. D., and E. J. Llewellyn (1989), Deconvolution of blurred images using photon counting statistics and maximum probability, *Can. J. Phys.*, *67*, 89+.

Loughman, R. P., E. Griffioen, L. Oikarinen, O. V. Postylyakov, A. Rozanov, D. E. Flittner, and D. F. Rault (2004), Comparison of radiative transfer models for limb-viewing scattered sunlight measurements, *J. Geophys. Res.*, *109*, 6303+, doi: 10.1029/2003JD003854.

Loughman, R. P., D. E. Flittner, B. M. Herman, P. K. Bhartia, E. Hilsenrath, and R. D. McPeters (2005), Description and sensitivity analysis of a limb scattering ozone retrieval algorithm, *J. Geophys. Res.*, *110*, 19,301+, doi: 10.1029/2004JD005429.

Lucke, R. L., D. R. Korwan, R. M. Bevilacqua, J. S. Hornstein, E. P. Shettle, D. T. Chen, M. Daehler, J. D. Lumpe, M. D. Fromm, D. Debrestian, B. Neff, M. Squire, G. König-Langlo, and J. Davies (1999), The Polar Ozone and Aerosol Measurement (POAM) III instrument and early validation results, *J. Geophys. Res.*, *104*, 18,785–18,800, doi:10.1029/1999JD900235.

McCormick, M. P., P. Hamill, T. J. Pepin, W. P. Chu, T. J. Swissler, and L. R. McMaster (1979), Satellite studies of the stratospheric aerosol., *Bull. Am. Meteorol. Soc.*, *60*, 1038–1047.

McCormick, M. P., L. W. Thomason, and C. R. Trepte (1995), Atmospheric effects of the Mt Pinatubo eruption, *Nature*, *373*, 399–404, doi:10.1038/373399a0.

McLinden, C. A., J. C. McConnell, C. T. McElroy, and E. Griffioen (1999), Observations of stratospheric aerosol using CPFM polarized limb radiances., *J. Atmos. Sci.*, *56*, 233–240.

McLinden, C. A., J. C. McConnell, E. Griffioen, and C. T. McElroy (2002), A vector radiative-transfer model for the Odin/OSIRIS project, *Can. J. Phys.*, *80*, 375–393.

McPeters, R. D., S. J. Janz, E. Hilsenrath, T. L. Brown, D. E. Flittner, and D. F. Heath (2000), The retrieval of O_3 profiles from limb scatter measurements: Results from the Shuttle Ozone Limb Sounding Experiment, *Geophys. Res. Lett.*, *27*, 2597–2600, doi:10.1029/1999GL011342.

Mie, G. (1908), Considerations on the optics of turbid media, especially colloidal metal sols, *Ann. Phys. Lpz.*, *42*, 377.

Murtagh, D., U. Frisk, F. Merino, M. Ridal, A. Jonsson, J. Stegman, G. Witt, P. Eriksson, C. Jiménez, G. Megie, J. de La Noë, P. Ricaud, P. Baron, J. R. Pardo, A. Hauchcorne, E. J. Llewellyn, D. A. Degenstein, R. L. Gattinger, N. D. Lloyd, W. F. J. Evans, I. C. McDade, C. S. Haley, C. Sioris, C. von Savigny, B. H. Solheim, J. C. McConnell, K. Strong, E. H. Richardson, G. W. Leppelmeier, E. Kyrölä, H. Auvinen, and L. Oikarinen (2002), Review: An overview of the Odin atmospheric mission, *Can. J. Phys.*, *80*, 309–319.

Nicolet, M., R. R. Meier, and D. E. Anderson (1982), Radiation field in the troposphere and stratosphere: II. Numerical analysis, *Planet. Space Sci.*, *30*, 935–983, doi:10.1016/0032-0633(82)90135-0.

Oikarinen, L., E. Sihvola, and E. Kyrölä (1999), Multiple scattering radiance in limb-viewing geometry, *J. Geophys. Res.*, *104*, 31,261–31,274, doi:10.1029/1999JD900969.

Postylyakov, O. V. (2004a), Radiative transfer model MCC++ with evaluation of weighting functions in spherical atmosphere for use in retrieval algorithms, *Adv. Space Res.*, *34*, 721–726, doi:10.1016/j.asr.2003.07.070.

Postylyakov, O. V. (2004b), Linearized vector radiative transfer model MCC++ for a spherical atmosphere, *J. Quant. Spectrosc. Radiat. Trans.*, *88*, 297–317.

Puliafito, E., R. Bevilacqua, J. Olivero, and W. Degenhardt (1995), Retrieval error comparison for several inversion techniques used in limb-scanning millimeter-wave spectroscopy, *J. Geophys. Res.*, *100*, 14,257–14,268, doi:10.1029/95JD00212.

Rault, D. F. (2005), Ozone profile retrieval from Stratospheric Aerosol and Gas Experiment (SAGE III) limb scatter measurements, *J. Geophys. Res.*, *110*, doi:10.1029/2004JD004970.

Rodgers, C. D. (1976), Retrieval of atmospheric temperature and composition from remote measurements of thermal radiation, *Rev. Geophys. Space Phys.*, *14*, 609–624.

Rodgers, C. D. (1990), Characterization and error analysis of profiles retrieved from remote sounding measurements, *J. Geophys. Res.*, *95*, 5587–5595.

Rodgers, C. D. (2000), *Inverse methods for atmospheric sounding: Theory and practice*, 167-186, World Scientific.

Rosen, J. M. (1971), The boiling point of stratospheric aerosols, *J. Appl. Met.*, *10*, 1044–1045.

Roth, C., A. Bourassa, D. Degenstein, R. Gattinger, and E. Llewellyn (2007), The retrieval of vertical profiles of the ozone number density using Chappuis band absorption information and a Multiplicative Algebraic Reconstruction Technique, *Can. J. Phys.,* submitted.

Rozanov, A., V. Rozanov, and J. P. Burrows (2001), A numerical radiative transfer model for a spherical planetary atmosphere: Combined differential-integral approach involving the Picard iterative approximation, *J. Quant. Spectrosc. Radiat. Trans.*, *69*, 491–512.

Rozanov, A. V., V. V. Rozanov, and J. P. Burrows (2000), Combined differential-integral approach for the radiation field computation in a spherical shell atmosphere: Nonlimb geometry, *J. Geophys. Res.*, *105*, 22,937–22,942, doi: 10.1029/2000JD900378.

Russell, J. M., L. L. Gordley, J. H. Park, S. R. Drayson, W. D. Hesketh, R. J. Cicerone, A. F. Tuck, J. E. Frederick, J. E. Harries, and P. J. Crutzen (1993), The Halogen Occultation Experiment, *J. Geophys. Res.*, *98*, 10,777–10,797.

Russell, P. B., and M. P. McCormick (1989), SAGE II aerosol data validation and initial data use: An introduction and overview, *J. Geophys. Res.*, *94*, 8335–8338.

Shepp, L. A., and Y. Vardi (1982), Maximum likelihood reconstruction for Positron Emission Tomography, *IEEE Trans. Med. Imag.*, *MI-1*, 113–122.

Solomon, S. (1999), Stratospheric ozone depletion: A review of concepts and history, *Rev. Geophys.*, *37*, 275–316, doi:10.1029/1999RG900008.

Stein, B., M. del Guasta, J. Kolenda, M. Morandi, P. Rairoux, L. Stefanutti, J. P. Wolf, and L. Woste (1994), Stratospheric aerosol size distributions from multispectral lidar measurements at Sodankyla during EASOE, *Geophys. Res. Lett.*, *21*, 1311–1314.

Steinacker, J., E. Thamm, and U. Maier (1996), Efficient integration of intensity functions on the unit sphere, *J. Quant. Spectrosc. Radiat. Transfer*, *56*, 97–107, doi:10.1016/0022-4073(96)00053-2.

Stothers, R. B. (1996), Major optical depth perturbations to the stratosphere from volcanic eruptions: Pyrheliometric period, 1881-1960, *J. Geophys. Res.*, *101*, 3901–3920, doi:10.1029/95JD03237.

Symons, G. (Ed.) (1888), *The Eruption of Krakatoa and Subsequent Phenomena*, Harrison and Sons.

Thomas, G. E. (1984), Solar Mesosphere Explorer measurements of polar mesospheric clouds (noctilucent clouds), *J. Atmos. Terrest. Phys.*, *46*, 819–824.

Thomas, R. J., C. A. Barth, W. Rusch, and R. W. Sanders (1984), Solar Mesosphere Explorer near-infrared spectrometer measurements of 1.27-micron radiances and the inference of mesospheric ozone, *J. Geophys. Res.*, *89*, 9569–9580.

Thomason, L. W. (2006), Assessment of Stratospheric Aerosol Processes, *World Climate Research Program*, 1–322.

Thomason, L. W., and G. Taha (2003), SAGE III aerosol extinction measurements: Initial results, *Geophys. Res. Lett.*, *30*, 33.

Thomason, L. W., L. R. Poole, and T. Deshler (1997), A global climatology of stratospheric aerosol surface area density deduced from Stratospheric Aerosol and Gas Experiment II measurements: 1984-1994, *J. Geophys. Res.*, *102*, 8967–8976, doi: 10.1029/96JD02962.

Thornton, D. C., A. R. Bandy, B. W. Blomquist, A. R. Driedger, and T. P. Wade (1999), Sulfur dioxide distribution over the Pacific Ocean 1991-1996, *J. Geophys. Res.*, *104*, 5845–5854, doi:10.1029/1998JD100048.

Tikhonov, A. (1963), On the solution of incorrectly stated problems and a method of regularization, *Dokl. Akad. Nauk SSSR*, *151*, 501–504.

Trepte, C. R., and M. H. Hitchman (1992), Tropical stratospheric circulation deduced from satellite aerosol data, *Nature*, *355*, 626–628.

Turco, R. P., P. Hamill, O. B. Toon, R. C. Whitten, and C. S. Kiang (1979), A one-dimensional model describing aerosol formation and evolution in the stratosphere: I. Physical processes and mathematical analogs, *J. Atmos. Sci.*, *36*, 699–717.

Twomey, S. (1975), Comparison of constrained linear inversion and an iterative nonlinear algorithm applied to the indirect estimate of particle size distributions, *J. Comput. Phys.*, *18*, 188–200.

Twomey, S., B. Herman, and R. Rabinoff (1977), An extension to the Chahine method of inverting the radiative transfer equation, *J. Atmos. Sci.*, *34*, 1085–1090.

Valero, F. P. J., and P. Pilewskie (1992), Latitudinal survey of spectral optical depths of the Pinatubo volcanic cloud - Derived particle sizes, columnar mass loadings, and effects on planetary albedo, *Geophys. Res. Lett.*, *19*, 163–166.

Vanhellemont, F., D. Fussen, C. Bingen, E. Kyrölä, J. Tamminen, V. Sofieva, S. Hassinen, P. Verronen, A. Seppälä, J. L. Bertaux, A. Hauchecorne, F. Dalaudier, O. Fanton D'Andon, G. Barrot, A. Mangin, B. Theodore, M. Guirlet, J. B. Renard,

R. Fraisse, P. Snoeij, R. Koopman, and L. Saavedra (2005), A 2003 stratospheric aerosol extinction and PSC climatology from GOMOS measurements on Envisat, *Atmos. Chem. Phys.*, *5*, 2413–2417.

von Savigny, C., I. C. McDade, E. Griffioen, C. S. Haley, C. E. Sioris, and E. J. Llewellyn (2005), Sensitivity studies and first validation of stratospheric ozone profile retrievals from Odin/OSIRIS observations of limb-scattered solar radiation, *Can. J. Phys.*, *83*, 957–972.

von Savigny, C., C. S. Haley, C. E. Sioris, I. C. McDade, E. J. Llewellyn, D. Degenstein, W. F. J. Evans, R. L. Gattinger, E. Griffioen, E. Kyrölä, N. D. Lloyd, J. C. McConnell, C. A. McLinden, G. Mégie, D. P. Murtagh, B. Solheim, and K. Strong (2003), Stratospheric ozone profiles retrieved from limb scattered sunlight radiance spectra measured by the OSIRIS instrument on the Odin satellite, *Geophys. Res. Lett.*, *30*, 8, doi:10.1029/2002GL016401.

Walter, H. (2006), Linearization of radiative transfer in spherical geometry, Ph.D. thesis, University of Amsterdam.

Wigley, T. M. L. (2006), A combined mitigation/geoengineering approach to climate stabilization, *Science*, *314*, 452–454, doi:10.1126/science.1131728.

Wiscombe, W. J. (1980), Improved Mie scattering algorithms, *Appl. Opt.*, *19*, 1505–1509.