CHAPTER 2

BACKGROUND

# 2.1 Introduction

Stratospheric aerosol plays an important role in the global radiative forcing balance by scattering solar irradiation and causing an overall cooling effect that depends on the particle size distribution and the concentration (*Kiehl and Briegleb*, 1993; *Stocker et al.*, 2013). These climate effects are an important and recent focus of research due to the potential contribution of stratospheric aerosol to the so-called global warming hiatus (*Solomon et al.*, 2011; *Haywood et al.*, 2014; *Fyfe et al.*, 2013) and efforts to quantify the variability and trends in the global stratospheric aerosol load are underway with various ground-based and satellite data sets (e.g. *Rieger et al.*, 2015; *Ridley et al.*, 2014).

Since its discovery with stratospheric balloon observations (*Junge et al.*, 1961), stratospheric aerosol has been measured with various techniques, although due to the variability of physical composition and particle size, the observations are always limited to some degree and no single measurement technique can fully determine the full range of aerosol properties unambiguously. In-situ balloon observations continue to be used and have provided highly valuable data sets, including most notably the long time series of Optical Particle Counter (OPC) measurements from Laramie, WY (*Deshler et al.*, 2003; 2008; *Kovilakam et al.*, 2015). Aircraft-borne nephelometers (*Beuttell and Brewer*, 1949; *Charlson et al.*, 1969) acquire detailed in-situ measurements, providing, for example, plume composition (*Murphy et al.*, 2014), but are spatially limited to the aircraft track. Ground based lidars have been used to do detailed studies of the extent of volcanic aerosol plumes (*Chazette et al.*, 1995; *Sawamura et al.*, 2012) and provide valuable insight into long term local variability and trends in the aerosol layer. For example, lidar observations were used by *Hofmann et al.* (2009) to first report the observed increase in stratospheric aerosol over approximately the last decade. However, the global distribution, which can only really be obtained with satellite observations, provides invaluable insight into aerosol processes and variability. A good example of this is the use of satellite observations by *Vernier et* (2011b) to determine that the increased stratospheric aerosol load reported by *Hofmann et al.* (2009) was in fact due to a series of relatively minor, mostly tropical, volcanic eruptions.

Several recent studies have highlighted the requirement for continued global stratospheric aerosol observations and especially the need to resolve, both vertically and horizontally, aerosol in the lowermost stratosphere and the upper troposphere. This is the case for tracking the evolution of aerosol from volcanic eruptions, which can have a substantial effect on the aerosol optical depth in the lowermost stratosphere (*Ridley et al.*, 2014; *Andersson et al.*, 2015). Furthering the understanding of the transport of aerosol near and across the tropopause would also benefit from higher spatial and temporal resolution observations. This is evident in the case of volcanic plumes, such as that from Nabro in 2011, the transport and origin of which has been studied extensively and the conclusions are somewhat controversial (*Bourassa et al.*, 2012c; 2013; *Vernier et al.*, 2013; *Fromm et al.*, 2013; 2014; *Fairlie et al.*, 2014; *Clarisse et al.*, 2014). However, this is also the case for the formation of background-level aerosol, particularly in the region of the Asian and North American monsoons, which have been identified as a source of substantial, seasonal and highly structured aerosol formation from precursor tropospheric source gases (*Vernier et al.*, 2011a; *Neely et al.*, 2014; *Thomason and Vernier*, 2013).

Continued stratospheric aerosol observations from space are drastically needed though few, if any, planned missions with such capability are underway. In this work, we present the design and test of a prototype instrument for potential future satellite-based stratospheric aerosol observation. The Aerosol Limb Imager (ALI) concept is a relatively small, low-cost, low-power, passive instrument, suitable for microsatellite deployment with the capability to provide high spatial resolution measurements, both vertically and horizontally, of the visible/NIR aerosol extinction coefficient. The basic idea is to leverage the clear advantages of the limb scatter technique as a passive, and therefore low mass and low power, means to obtain daily global coverage, with a two dimensional hyperspectral imager for filling cross-track observation.

The ALI instrument concept is built around the use of an Acousto-Optic Tunable Filter (AOTF), which is a novel filtering technology that provides the ability to rapidly select the central wavelength of an image with no moving parts. These filters, which have recently been developed as large aperture imaging quality devices, operate very efficiently in the red and near infrared spectral range, which is a well matched spectral range for limb scatter sensitivity to aerosol and cloud (*Rieger et al.*, 2014). Additionally, the spectral bandpass of the AOTF, which is typically between 3-6 nm at these wavelengths, is very suitable for the broadband scattering characteristics of the aerosol limb signal. The two dimensional imaging nature of the design provides the capability to achieve at least sub-kilometer resolution at the tangent point, which is on the order of the scale size of the upper troposphere and lower stratosphere (UTLS) aerosol features mentioned above.

It should be noted that the basic instrument design concept of ALI is very similar to that of the Atmospheric Limb Tracker for the Investigation of the Upcoming Stratosphere (ALTIUS) (*Dekemper et al.*, 2012), which is a Belgian instrument concept from the Belgian Institute for Space Aeronomy (BIRA) and has been recently selected for small satellite deployment by the European Space Agency. ALTIUS is designed to measure limb scattered sunlight; however, it also has solar, stellar, and planetary occultation modes and is scientifically focused on trace gas measurements, particularly for ozone, whereas ALI is optimized for aerosol observation from limb scattering observations.

# 2.2 Stratospheric Aerosol

In the late 18th century, it was known that atmospheric temperature decreased with altitude and a theory had been raised that at a specific altitude the temperature must eventually go to absolute zero (*Hoinka*, 1997). This lead to a series of balloon campaigns, which were noisy and unreliable due to the technology available, to discover this mysterious altitude in the atmosphere. However, in the late 19th century the technology used in these sounding balloons had improved to a point where the atmospheric temperature could be accurately measured and it was found that at approximately 12 km an inversion point occurs where the temperature starts to increase and thus the tropopause, which separates troposphere and the stratosphere, was discovered. The stratosphere is the region of the atmosphere above this temperature inversion point where atmospheric temperature increases. The tropopause, which is the lower bound of the stratosphere, ranges in altitude from approximately 10 km to 16 km from the high latitudes to the tropics (*Andrews*, 1987), and extends up to approximately 50 km. It is a thermodynamically stable and fairly dry (*Boucher*, 2015) region of the atmosphere, and the characteristic stability of the stratosphere limits vertical transport, leading to long lifetimes, spanning from months to years, for non-volatile species (*Volk et al.,* 1997; *Brasseur and Solomon*, 2005)

The stratosphere undergoes exchange of air with the troposphere though a series of dynamical processes including tropical convection, polar vortices, and tropopause folding (*Holton et al.,* 1995). Meridional circulation within the stratosphere is dominated by the slow Brewer-Dobson circulation, although zonal circulation, *i.e*. along a constant latitude, is much faster (*Plumb and Eluszkiewicz*, 1999) and tends to cause zonal symmetry of composition. The transport of gases emitted from sources on the surface, or chemically created in the troposphere, through the tropopause and into the stratosphere is an important aspect of stratospheric composition. This is the case for stratospheric aerosol where the oxidation of sulfur-bearing compounds, transported from the troposphere to the stratosphere, forms the aerosol layer discovered by *Junge et al.* (1961) with stratospheric balloon sondes measurements. These aerosols are primarily droplets of hydrated sulfuric acid (H­2­SO­4) formed from the oxidation of sulfur-containing source gases, primary OCS and SO2 (*Brock et al.*, 1995). This stable layer of aerosol exists in the stratosphere from the altitude of the tropopause to approximately 30 km.

## 2.2.1 Aerosol Sources

The source gases that eventually form stratospheric aerosol are emitted or produced in the troposphere through both natural and anthropogenic processes. These sulfur sources enter the atmosphere in various ways and undergo a chain of chemical reactions to form sulfate aerosol within the stratosphere.

One primary source of atmospheric sulfur is OCS, which originates from natural marine processes, biomass burning, and industrial processes (*Kettle et al.*, 2002; *Notholt et al.*, 2003). OCS has a long lifetime in the troposphere and low solubility allowing for a significant portion to reach the stratosphere where some of it oxidizes and hydrates to form sulfate aerosol (H2SO4) where it contributes to the background aerosol layer (*Crutzen*, 1976).

Another important source of atmospheric sulfur is sulfur dioxide (SO2), which originates in the troposphere through the burning of fossil fuels. SO2 has a short lifetime in the troposphere, and its concentration varies regionally; however, this emitted SO2 still can enter the stratosphere through transport processes and contribute to the background aerosol layer (*Thomason and Peter*, 2006). A second source of SO2 is from volcanic eruptions; which, although highly variable in location and time, can inject a large amount of sulfur directly into the stratosphere. Volcanic eruptions can inject such large amounts of sulfur that they in fact dominate the variability of the stratospheric aerosol layer, essentially episodically perturbing the background levels. This volcanic enhancement of the aerosol layer occurred on a massive level following the volcanic eruptions of El Chichon in 1982 (12-20 Tg of SO2) (*McCormick and Swissler*, 1983; *Hofmann and Rosen*, 1983) and Mount Pinatubo in 1991 (20-30 Tg of SO2) (*McCormick and Veiga*, 1992). However, after the Mount Pinatubo eruption a volcanically quiescent period occurred where aerosol layers returned to background until the early 2000’s. Following this period, a series of relatively minor, mostly tropical, volcanic eruptions have increased the background aerosol layer in the amount of 4-7% per year from 2000 to 2009 (*Vernier et al.*, 2011b). These small eruptions have continued to the present day and include eruptions such as Kasatochi (1.2-2.2 Tg of SO2) in 2008 (*Prata et al.*, 2010), Nabro (1.0- 1.5 Tg of SO2) in 2011 (*Clarisse et al.*, 2016), Kelut (0.1-0.3 Tg of SO2) in 2014, and Calbuco (0.2 -0.5 Tg of SO2) in 2015 (*Carn et al*., 2016). The enhancement of the aerosol layer from these eruptions is well captured in the satellite record (*Rieger et al., 2015).*

## 2.2.2 Aerosol Microphysics

These sulfur source gases undergo a series of chemical reactions and are converted into molecular H2SO­4 that then nucleates and condenses to form liquid droplets of approximately 25% H2O and 75% H2SO4 (*Rossen*, 1971; Wang *et al.*, 1989). These spherical droplets coagulate into various sizes distributed between approximately 0.05 to 1.0 µm, depending on the various contributions and stages of the processes of nucleation, evaporation and condensation (*Junge et al.* 1961; *Brock et al.*, 1995; *Bingen et al.*, 2004). A log-normal distribution is often used to approximate the distribution of particle radii, *r*, in the form of

|  |  |
| --- | --- |
|  | (2.1) |

where is the aerosol number density, is the mode radius, and is the mode width (*Jäger and Hofmann,* 1991; *Hamill et al.* 1997). In this case, the particle radii are distributed normally over the logarithm of the radius. For a typical non-volcanic background aerosol, with a mode radius and mode width of 0.08 µm and 1.6 respectively (*Deshler*, 2003), the log-normal distribution is shown in **Figure 2-1**.



Figure 2-1: Sample log-normal distribution for typical non-volcanic stratospheric aerosol.

Optical Particle Counters (OPC) have been used on board stratospheric balloon flights from Laramie, Wyoming over the past 40 years to make *in situ* measurements of aerosol particle sizes in bins between 0.15 to 2.0 µm (*Deshler et al*., 2003). These measurements provide a valuable and unique long term set of size-resolved measurements of sulfate aerosol. These particle size distributions are primarily unimodal during non-volcanic periods and the log-normal distribution given in Equation 2.1 can be used to approximate background period particle sizes. But during volcanic episodes, a bimodal log-normal distribution of aerosol particles, which includes what is referred to fine and coarse modes, is more representative of the measured distributions (*Deshler et al.*, 2003; 2008; *Kovilakam et al.*, 2015). The coarse mode has larger particles than the fine mode and complicates the determination of aerosol microphysical parameters since the number of required parameters has increased to six: a number density for both the fine and coarse mode, two mode radii, and two mode widths. Figure 5 from *Deshler et al.* (2003), recreated in Figure 2-2, demonstrates two bimodal particle size distributions from balloon OPC measurements. The first distribution is from a volcanic period in 1993 after the Mount Pinatubo eruption and another from a background period in 1999. It should be noted that even though a bimodal distribution is found for the background case in Figure 2-2, the number density of the coarse mode is very small and can generally be ignored in non-volcanic periods so a unimodal approximation is typically sufficient.



**Figure 2-2**: Bimodal particle size distributions fits from OPC. (a) Distributions from a volcanic period after the Mount Pinatubo eruption recorded in 1993. (b) Distributions from a background aerosol period recorded in 1999. Both of the aerosol distribution measurement are from 20 km altitude with the solid line being the fine mode and the dashed line is the coarse mode. Figure is recreated from Figure 5 of *Deshler et al.* (2003).

## 2.2.3 Climate Effects

Stratospheric aerosol can have several effects on the climate of the planet, and particularly due to the variability of the volcanic contribution, there is a large amount of uncertainty in the overall effect (*Solomon et al.*, 2007). Through the so-called “direct effect”, aerosol particles scatter incoming visible solar radiation away from earth increasing the albedo causing a cooling of the surface of the planet (*Lacis et al.*, 1992). The albedo is the amount of incoming solar irradiance that is reflected back to space. A secondary direct effect from aerosols, which is highly dependent on aerosol particle size distribution, is a greenhouse-like effect that is caused by scattering of infrared radiation emitted from the earth’s surface (*Kiehl and Briegleb*, 1993). Aerosol also introduces a so-called “indirect effect” to the radiative balance. This is also known as the cloud albedo effect. This is caused by condensation of water onto existing aerosol particles. These become cloud condensation nuclei and stimulate cloud formation, which leads to an increase of the planetary albedo, which then also contributes to cooling the planet's surface. These types of cloud forming particles also tend to increase the overall lifetime of the cloud, increasing the overall cloud coverage and thus increasing the planetary albedo (*Charlson et al.*, 1992). Overall, the cooling effect of the aerosol particles dominates the warming effect and cools the surface of the planet (*Solomon et al.*, 2011).

During background periods without substantial volcanic contribution to the aerosol load the cooling effect from stratospheric aerosols is very small, but this greatly changes during periods of volcanic activity where the layer concentrations can be significantly enhanced. After the eruption of Mount Pinatubo in 1991 the sulfate aerosol load was increased by 5 to 10 fold causing cooling of the lower atmosphere by 0.5◦C (*McCormick et al.* 1995; *Soden et al.*, 2002) and 0.1 to 0.3◦C on the surface (*Thompson et al.*, 2009; *Canty et al.*, 2013). The surface temperatures did not return to pre-Pinatubo level until approximately three years after the eruption as the atmosphere filtered out the additional aerosol (*Hansen at al.*, 1996). More recently, a series of small to moderate volcanic eruptions have increased the background stratospheric aerosol layer (*Vernier et al.*, 2011b). This additional volcanic aerosol load has been proposed to be linked to a larger effect of decreased warming, known sometimes as the global warming hiatus (*Solomon et al.*, 2011; *Haywood et al.*, 2014; *Fyfe et al.*, 2013).

# 2.3 Aerosol Measurements

There are essentially three fundamental approaches to measuring atmospheric aerosol: ground based systems, *in-situ* measurements, and satellite based remote sensing. Each of these methods has certain strengths and the combination of data from all three approaches is essential for the future. Ground based and *in-situ* measurements typically provide detailed information about a specific localized area. Global coverage can only really be obtained with satellite measurements, but these are essentially always information-limited by the remote sensing technique. Ground-based, *in-situ* and satellite measurements have important roles in monitoring the planet’s aerosol content and each of these methods have inherent advantages and disadvantages. An overview is given here on some of the common methods for stratospheric aerosol measurements in order to place the ALI requirements and design in context.

## 2.3.1 In-situ Measurements

In-situ measurement are typically performed using balloon- or aircraft-based platforms. In-situ balloon instruments directly measure aerosol particles during the assent and can determine the height profile of the particle size distribution. The OPC is an active instrument that uses a light source internal to the device to optically count aerosol particles. This type of instrument has been launched from Laramie, Wyoming since 1971, and has successfully measured aerosol mixing ratio and particle size distributions (*Deshler et al.*, 2003; 2008; *Kovilakam et al.*, 2015) on many flights since that time. Similarly, aircrafts have been used to carry nephelometers to acquire detailed in-situ measurements (*Beuttell and Brewer*, 1949; *Charlson et al.*, 1969) including plume composition (*Murphy et al.*, 2014) but are limited spatially to the aircraft track.

## 2.3.2 Occultation

Satellite instrumentation capable of measuring stratospheric aerosol has been in use since the 1970s, beginning with limb sounding solar occultation measurements. The viewing geometry of this technique is shown in Figure 2-3. Solar occultation measurements have provided a reliable, accurate and essentially continuous long term record of vertically resolved aerosol extinction coefficient measurements, mostly from the series of Stratospheric Aerosol and Gas Experiment (SAGE) instruments including SAGE I in 1979, SAGE II in 1984, and SAGE III in 2001 (*Russell and McCormick*, 1989; *Thomason and Taha*, 2003). These SAGE measurements, which have a vertical resolution of approximately 1 km, have generally compared well with ground-based and in-situ measurements, although there are challenges associated with determining microphysical parameters and comparison between instruments can be challenging (*Russell and McCormick*, 1989; *Kovilakam et al.*, 2015). However, solar occultation is generally a robust and stable technique as it directly measures atmospheric optical depth, along with the exo-atmospheric solar spectrum with each scan, allowing for straightforward retrieval of aerosol extinction coefficients (*Damadeo et al*, 2013). The major drawback to occultation instruments is that a sunrise or sunset event is required to perform a measurement limiting is the number of scans per day to 16-48 measurements depending on the orbit.



**Figure 2-3**: An occultation instrument monitoring the atmosphere by scanning the atmosphere by looking directly at the sun.

The series of SAGE missions came to an end in 2006 with the failure of SAGE III. The occultation measurements from the currently operational MAESTRO and ACE-Imager instruments on SciSat (*McElroy et al.*, 2007; *Gilbert et al.*, 2007) have had some success producing stratospheric aerosol extinction products (*Vanhellemont et al.*, 2008; *Sioris et al.*, 2010). Furthermore, a manifestation of SAGE III is planned for deployment on the International Space Station in 2016 (*Cisewski et al.*, 2014) to continue the valuable occultation-based aerosol record.

## 2.3.3 Lidar

A method known as lidar can determine atmospheric parameters through the pulsing of a laser and the subsequent measurement of the intensity of the backscattered light at different wavelengths and polarizations. Lidar has been used at ground based facilities to measure aerosol layers dating back to the 1960s (*Fiocco and Grams*, 1964) and are still used today. More recently lidar instruments have been used on satellite missions including the Ice, Cloud, and land Elevation Satellite (ICESat) from 2002 to 2010 (*Schutz et al.*, 2005) and Cloud-Aerosol Lidar and Infrared Pathfinder Satellite Observations (CALIPSO) which launched in 2006 (*Winker et al.*, 2007). More recently Cloud Aerosol Transport System (CATS) (*Chuang et al.*, 2013) has been mounted on the international space station in 2015 with a three year planned mission. Traditionally lidar instrument observed in the nadir or zenith directions (*i.e*. straight down or up); however, some instruments are mounted slightly off-nadir. Both orientations are shown in Figure 2-4 for a space based geometry. Lidar measurements have been used to make highly spatially resolved measurements of aerosol plumes from volcanic eruptions (*Chazette et al.*, 1995; *Sawamura et al.*, 2012) as well as monitoring long term trends (*Hofmann et al.*, 2009).

One of the most notable lidar instruments for stratospheric aerosol measurements is CALIPSO, which is joint mission developed between the National Aeronautics and Space Administration (NASA) and the Centre National d'Etudes Spatiales (CNES) of the United States and France respectively. It uses a two wavelength polarized lidar system to achieve high resolution aerosol and cloud retrievals along the satellite's orbital track with global coverage from 82◦S to 82◦N (*Young and Vaughan*, 2009). CALIPSO nominally measures backscatter profiles approximately every 300 m along track with approximately 200 m vertical resolution. However, the stratospheric backscatter signal is weak and requires averaging of only the night time measurements over several days, typically yielding resolutions of 0.5 km vertically and 500 km horizontally (*Vernier et al*, 2011b). Additionally, the uncertainty in the calibration with respect to the molecular background is on the order of the stratospheric aerosol signal and leads to a potential bias in the stratospheric measurements (*Rogers et al.*, 2011). CALIPSO was launched in 2006 and although it is presently still operational, it is operating beyond its design lifetime.



Figure 2-4: Lidar instrument showing measurements in both the nadir and off-nadir lines of sight.

## 2.3.4 Limb Scatter

The limb scatter technique has an observing geometry that is similar to occultation but measures sunlight light that is scattered into the line of sight of the instrument from atmospheric interactions rather than directly observing the sun. These scattering interactions can include both single and multiple scattering events. Single scattering occurs when light from the sun interacts with a particle in the atmosphere and scatters it directly into the line of sight of the instrument. Multiple scatter is when the photon undergoes several scattering events before entering the line of sight. This can include scattering from multiple particles in the atmosphere and scattering from the ground surface. In general, these events can occur any number of times before entering the instrument. The geometry for the limb scatter technique is shown in Figure 2-5 along with the fundamental angles used to describe this condition. All angles are defined from the tangent point, which is the point where the distance between the line of sight and the surface of the earth is minimized, represented by the black dot. The Solar Zenith Angle (SZA) is the angle between the local vertical and the direction of the sun; the Solar Scattering Angle (SSA) is the defined to be the angle between the direction of the sun and the line of sight; and the Solar Azimuth Angle (SAA) is the angle between the projection of the sun direction onto the horizontal plane at the tangent point and the line of sight.



Figure 2-5: Limb scattering geometry measurement for an instrument where single and multiple scattering events occur.

The limb scatter method yields relatively good vertical resolution, comparable to occultation, and allows for measurements to be taken during any daylight period; however, it requires the use of a complex forward model to calculate the scattering events along with typically some *a priori* knowledge of the aerosol scattering cross section in order to retrieve the extinction coefficient profile. The model needs to accurately determine the effect of multiple scattering since it consists of 10-50% of the measured signal depending of the specific geometry and wavelength (*Oikarinen et al.*, 1999). Furthermore, due to the complex nature of the problem; a large amount of computational time and memory is required for an accurate calculation.

The first use of limb scatter was on the Solar Mesosphere Explorer (SME) (*Barth et al.*, 1983) to measure mesospheric ozone profiles in 1981. Much later, other limb scatter instruments were launched into low earth orbit that had the capability to determine aerosol extinction including the Optical Spectrograph and InfraRed Imaging System (OSIRIS) launched on the Odin satellite in 2001 (*Llewellyn et al.*, 2004) and the SCanning Imaging Absorption spectroMeter for Atmospheric CHartographY (SCIAMACHY) on Envisat launched in 2002 (*Bovensmann et al.*, 1999). Both of these instruments are scanning grating spectrometers with a single line-of-sight and scan the atmosphere vertically to complete a vertical profile measurement.

The OSIRIS version 5.07 data product provides 750 nm aerosol extinction profiles at approximately 2 km vertical resolution (*Bourassa et al.*, 2007) and has been shown to agree relatively well, generally within 30%, with SAGE II and SAGE III occultation measurements (*Bourassa et al.*, 2012b; *Rieger et al.*, 2015). The SCIAMACHY instrument uses a retrieval technique essentially similar to OSIRIS to retrieve aerosol profiles at 750 nm with approximately 3 km vertical resolution (*Ernst et al.*, 2012; *von Savigny et al.*, 2015). However SCIAMACHY observations ceased with the demise of Envisat in 2012 and although OSIRIS continues to operate, it is now in the sixteenth year of a mission designed for two years.

The most recently launched limb scatter instrument is the Ozone Mapping Profiler Suite Limb Profiler (OMPS-LP) on the Suomi-NPP satellite in 2011. Although similar in spectral range and vertical resolution to OSIRIS, OMPS-LP is an imaging spectrometer that vertically images the limb in a single measurement. The imaging capability of OMPS-LP provides a decrease in the time required to obtain a limb profile and so increases the along track sampling. Recent work on the feasibility of aerosol retrieval from OMPS-LP measurements show promising results (*Rault and Loughman*, 2013).

An instrument that is currently under development for a European Space Agency satellite mission is ALTIUS (*Dekemper et al.*, 2012), which is a concept from the Belgium Institute for Space Aeronomy. ALTIUS is designed to image limb scattered sunlight, both vertically and horizontally across the track through the use of Acousto-Optic Tunable Filter (AOTF) technology (see section 3.2). ALTIUS is also designed to have solar, stellar, and planetary occultation modes. ALTIUS is scientifically focused on trace gas measurements, particularly for ozone and the instrument has three hardware channels, each channel with a separate AOTF, overall measuring wavelengths from 250-2000 nm. These measurements could eventually be used for aerosol extinction retrieval.

The limb scatter technique is the one selected for the ALI instrument in this thesis work. The limb scatter technique was selected for ALI due to the potential for global coverage, high vertical resolution and the high quality of aerosol extinction retrieval as proven the OSIRIS heritage. Like ALTUIS, ALI also uses an AOTF to spectrally image the limb scattered signal. As discussed in detail in Chapter 3, the hyperspectral imaging nature of the ALI design with the AOTF allows for rapid image collection for the retrieval of high spatial resolution aerosol extinction.

# 2.4 Radiative Transfer

The use of the limb scatter technique requires a detailed understanding of radiative transfer, and the modeling of the complex scattering interactions of light within the atmosphere is somewhat involved. In this section, a necessary overview of scalar radiative transfer is provided, followed by the necessary modifications to the theory to form the polarized radiative transfer equation. A description of scattering interactions important to aerosols is also developed. Finally, an overview of the SASKTRAN radiative transfer model used within this work is provided.

## 2.4.1 Scalar Radiative Transfer

The following presents a derivation of radiative transfer equations for the atmosphere in terms of the scalar radiance, a theory which does not account for polarization. In order to accurately discuss radiative transfer, a coordinate system must first be defined. If we assume that a ray of light, , is propagating in a given direction, , and starts at a location, , with the initial position of , then the position of the ray along the path direction can be completely defined by its path length, . The basis of path length is used to define the radiative transfer equations.

The fundamental theory for radiative transfer is known as the Beer-Lambert law. The law describes the change in intensity or radiance of light, , as it interacts with a thin layer of space or atmosphere, . The thin layer has particles which affect the attenuation of the light which is dependent on the number of particles, , and the particle cross section, . If there are several different particles, the attenuation is a summation of the number densities and cross sections. The Beer-Lambert Law describes the change in radiance along the path

|  |  |
| --- | --- |
|  | (2.2) |

The extinction of the particles is a measure of the loss of light over a given distance and is defined as

|  |  |
| --- | --- |
|  | (2.3) |

Integrating Equation 2.2 forms the following result

|  |  |
| --- | --- |
|  | (2.4) |

The optical depth, , is defined as the extinction over the path length simplifying Equation 2.4 to

|  |  |
| --- | --- |
|  | (2.5) |

The above gives the radiance at point after it has gone through attenuation from .

Although this form of the Beer-Lambert’s Law is useful for describing the loss of light through scattering or absorbing from an initial source though a medium, in the atmosphere we must also account for incoming light that is scattered into the line of sight from other directions or directly emitted from particles along the path. To account for this additional source of light a source term, , is added to Equation 2.2 to yield

|  |  |
| --- | --- |
|  | (2.6) |

Using the fact that the change in optical depth is defined as

|  |  |
| --- | --- |
|  | (2.7) |

Equation 2.6 is rearrange into

|  |  |
| --- | --- |
|  | (2.8) |

Using the following derivative

|  |  |
| --- | --- |
|  | (2.9) |

and substituting it into Equation 2.8 yields

|  |  |
| --- | --- |
|  | (2.10) |

This form can now be integrated over the optical depth giving

|  |  |
| --- | --- |
|  | (2.11) |

Selecting the reference point at the observer to be and converting the equation back to path lengths yields

|  |  |
| --- | --- |
|  | (2.12) |

which gives the radiance as seen from an observer at a point, , along the line of sight.

In the atmosphere there are three sources of additional radiation that contribute to the source term: thermal emissions, photochemical reactions, and scattered light. For wavelengths from the visible to the near infrared (*i.e.* wavelengths less than 2 µm) there is negligible contribution from thermal emissions. Furthermore, as long as distinct wavelengths where photochemical reactions emit are avoided this source term can also be ignored. This leaves scattering as the only significant source of light to be added into the line of sight. The source term for scattered sunlight is given by

|  |  |
| --- | --- |
|  | (2.13) |

The diffuse radiance is given by and is the radiation scattered into the line of sight from all directions. The phase function, , described the probability that a ray of light is scattered from a direction, , into the line of sight propagation direction, . The scattering angle, , is defined as

|  |  |
| --- | --- |
|  | (2.14) |

Lastly, is the extinction only caused by scattering and not absorption. The term only allows the fraction of particles that scatter radiation, and not absorb it, to contribute to the source term.

As a final note, the calculation of the diffuse radiance is what makes this a computationally heavy problem. To completely solve for the diffuse radiance, the radiance at every point in the atmosphere must be determined. Furthermore, the light can be scattered multiple times in the atmosphere, requiring a diffuse radiance for each order of scatter. Each successive scattering adds smaller contributions to the final radiance at the observer. Through this iterative process the full multiple scatter solution to the radiative transfer equation is

|  |  |
| --- | --- |
|  | (2.15) |

The multiple scatter term is calculated for each successive order until the contribution is sufficiently small to be negligible.

## 2.4.2 Vector Radiative Transfer

The scalar radiative transfer equation works well for systems that do not measure polarized light as the effect of polarization on the total radiance is small. However, for instruments that measure polarized light, a vector radiative transfer equation is required. Before polarization can even be discussed, a method to quantify polarization must be defined which is the Stokes vectors. The Stokes vectors are given as

|  |  |
| --- | --- |
|  | (2.16) |

where is the scalar or total radiance, is the difference between horizontal polarization to vertical polarization, is the difference between +45◦ diagonal polarization to -45 polarization, and is the difference between the counter clockwise circular polarization to clockwise polarization (*Bickel and Bailey*, 1985). Using a reference frame where the local x-axis is defined to be the horizontal polarization leads to the following definition for the Stokes vector

|  |  |
| --- | --- |
|  | (2.17) |

The electric field aligned with the x and y-axis are and respectively, the star is the complex conjugate, and and are the “real part of” and “imaginary part of” respectively. The degree of polarization can be determined with the Stokes vectors. If the equality holds true then the light is fully polarized, otherwise it is only partially polarized if .

With the addition of polarization, the radiative transfer and source term equations (Equations 2.12 and 2.13) need to be rewritten with polarization state included. The polarized radiative transfer equation are

|  |  |
| --- | --- |
|  | (2.18) |
|  | (2.19) |

which are the vector radiative transfer and source term equations respectively (*Mishchenko et al.*, 2002). With polarization a scattering reference frame is defined and incoming radiance is rotated into the scattering frame multiplied by the scattering matrix, , then returned to the original propagation frame. The rotation matrix is defined as

|  |  |
| --- | --- |
|  | (2.20) |

where is the angle between the propagation and scattering reference frame. The radiance and the source terms are now Stokes vectors in 4 by 1 matrices and the scattering matrix, , is a 4 by 4 tensor that is related the probability of the incoming light to be scattered in the propagation direction with a specific polarization. As a note, the operation of is commonly referred to as the phase matrix. The polarization equation adds extra computation and memory consumption since the polarization must be computed at each scattering in the radiative transfer equation, which is nontrivial, and stored in memory, essentially four times the size of a standard scalar radiance calculation.

With the complete vector polarized radiative transfer expression the two scattering interactions that pertain to determining aerosol are described. The first interaction is Rayleigh scattering, which defines the scattering process from the molecular atmosphere, and Mie scattering, which determines how incoming light scatters from spherical particles, *i.e*. a model for stratospheric aerosol scattering.

## 2.4.3 Rayleigh Scattering

Rayleigh scatter is the scattering process by the molecular background atmosphere, *i.e*. by molecules of the air. The first calculation of molecular atmospheric scattering cross sections was by Lord Rayleigh where he assumed the molecules were dielectric spheres with radii much less than the wavelength of the light. Later, the King correction was added to the Rayleigh scattering cross section, , to yield the following expression

|  |  |
| --- | --- |
|  | (2.21) |

which is highly dependent on wavelength, . The parameters and are the volume polarizability, given per unit volume, and the depolarization ratio, which is unitless (*Sneep and Ubachs*, 2005).

The other important quantity for scattering is the scattering matrix, which is analogous to the scattering phase function in the scalar theory. For Rayleigh scattering, the vector model scattering matrix is given by the Rayleigh-Gans approximation (*Mishchenko et al.*, 2002)

|  |  |
| --- | --- |
|  | (2.22) |

Each component of the scattering matrix itself is smoothly varying and analytically determined, which allows for easy and accurate calculation of Rayleigh scattering.

## 2.4.4 Mie Scattering

For larger particles, like sulfate aerosol, Rayleigh scattering no longer holds since the size of the particles is on the order of the wavelength and Mie scattering must be used. *Mie* (1908) solved Maxwell’s equations for scattering from a dielectric sphere in a general sense with a solution using a series of spherical Bessel and Henkel functions. Only the fundamental Mie scatter equation is presented here but a full derivation of Mie scatter can be found in *van de Hulst* (1957). The scattering cross section from Mie theory is given by

|  |  |
| --- | --- |
|  | (2.23) |

where is the wavenumber, r is the particle radius and the coefficients and are given by

|  |  |
| --- | --- |
|  | (2.24) |
|  | (2.25) |

The index of refraction of the particle is given by , and and are the normalized half-integer order Bessel functions of the first kind and Henkel functions of the second kind respectively. The scattering matrix for Mie scatter for a vector solution has the following form (*Hansen and Travis*, 1974)

|  |  |
| --- | --- |
|  | (2.26) |

The terms in the scattering matrix, and , are known as the amplitude functions and are given by

|  |  |
| --- | --- |
|  | (2.27) |
|  | (2.28) |

where are the Legendre polynomials.

In the atmosphere, various particle sizes occur and a log-normal distribution (Equation 2.1) is assumed for aerosols. In order to determine effective scattering cross-section, a weighted average over the particle radius is performed

|  |  |
| --- | --- |
|  | (2.29) |

The weighted average is similarly performed to determine the effective phase matrix for a particle size distribution.

There are stark differences between Rayleigh and Mie scattering cross sections and scattering matrices. Using standard atmospheric conditions, a comparisons of the extinctions and thus the scattering cross sections as a function of wavelength are shown in Figure 2-6a. The “wavelength to the fourth” dependence is shown for Rayleigh scattering, in blue, (Equation 2.21) whereas Mie scattering, in red, has a flatter spectral dependence, which changes with particle size. Due to this difference, the contribution of Mie scattering becomes more predominant as wavelength increases. The terms of the scattering phase matrix are also radically different for the two scattering processes (Figure 2-6b). For the first term of the scattering matrix, , the Rayleigh term is smooth with even forward and back scattering and decreases at 90◦ scattering angle. The scattering term for Mie on the other hand has a strong tendency for forward scattering compared with much lower values for scattering angles past 90◦. Further, for certain wavelengths and size parameters the Mie phase matrix can have oscillations super-imposed onto the curve making the phase matrix even more complex than Rayleigh scattering.

It should be noted that although the theory is well understood, the calculation of the Mie scattering cross sections and phase matrices in practice is computationally intensive since the terms consist of infinite sums of Bessel and Henkel functions. Comprehensive work done by *Wiscombe* (1980) has allowed for accurate and efficient computation of the Mie scattering coefficients, and these are used within the modelling work presented here.

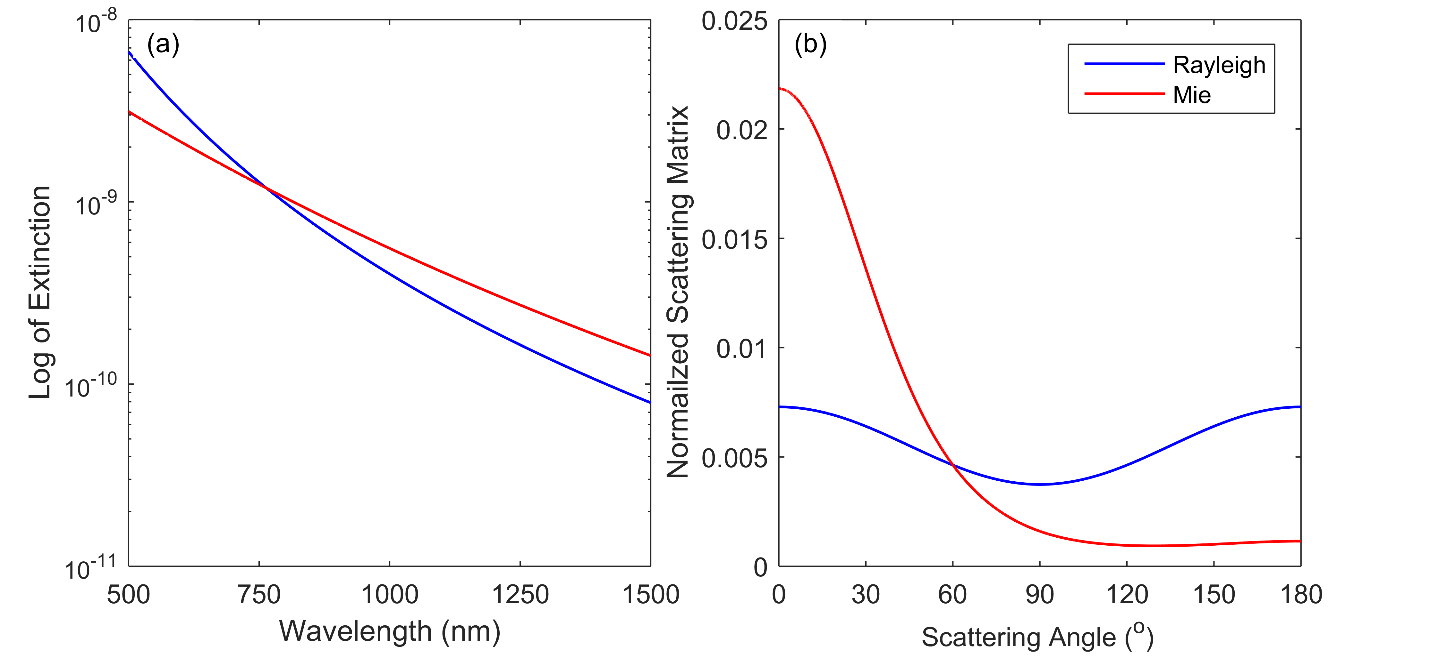


Figure 2-: (a) Change in scattering cross section for Rayleigh and Mie scattering over wavelength. The Mie scattering uses a log-normal distribution with a mode width of 1.6 and a mode radius of 0.08 µm. (b) The first term of scattering matrix, , for Rayleigh and Mie scattering across scattering angle.

## 2.4.5 SASKTRAN Radiative Transfer Model

The SASKTRAN radiative transfer was first developed to solve the scalar radiative transfer equation in a fully spherical atmosphere for both single and multiple scattering with a one dimensional atmosphere, *i.e*. considering the variation in altitude only (*Bourassa et al*., 2008). The first source term, , is the light from the sun attenuated and scattered into the instrument line of sight, and it is assumed the incoming solar irradiance encounters the earth in parallel randomly polarized rays. To include higher order terms, a successive orders method is used to simulate second, third and higher orders of scattering within the atmospheric model. Another important assumption in the SASKTRAN model is that the ground reflection is assumed to have a depolarizing Lambertian distribution, which will evenly distribute the incoming radiance evenly in all outgoing directions with the efficiency of the planetary albedo.

Recent upgrades have been performed on SASKTRAN and have led to a new solver, or “engine” known as SASKTRAN High Resolution or SASKTRAN-HR (*Zawada et al.*, 2015) which has expanded the model capabilities to be able to perform radiative transfer calculations with a fully three dimensional atmosphere. This update allows the model to vary the atmospheric concentrations in not just the vertical direction, like the original SASKTRAN, but in both of the horizontal geometries (*i.e.* latitude and longitude), allowing for more accurate modelling of the variation of atmospheric parameters along the instrument line of sight and the potential for tomographic inversions.

The most important update to the SASKTRAN-HR model for this thesis work is the capability to calculate the vector or polarized radiances (*Dueck et al.*, 2016). Solving the vector radiative transfer equation allows for SASKTRAN to compute the Stokes vectors in the reference frame of the model, which can then be rotated into any desired instrument frame of reference. The polarization algorithm implemented in SASKTRAN-HR provides polarized calculations up to a user-specified number of orders of scattering, with all successive orders approximated as randomly polarized.

# 2.5 Inversion Techniques

Remote sensing methods indirectly measure the desired atmosphere state and require a method to be able to transform the measurement into the desired physical atmospheric quantity. This process is known as a measurement “inversion”, “retrieval”, or more generally as the inverse problem. A measurement vector, , is constructed from the spectral radiance that has sensitivity to the desired physical parameter while reducing sensitivity to other physical parameters as much as possible. A forward model (for example in this case SASKTRAN-HR), is used to compute the measurement vector using an input atmospheric state, , and desired physical parameter state, . The length of the measurement vector and state vector are and respectively. It is important to note that the length of the measurement vector and state vector do not have to be the same. This leads to the following formation

|  |  |
| --- | --- |
|  | (2.30) |

where , is the measurement noise. This is the equation that needs to be inverted to yield the desired atmospheric state.

The inverse is found directly if it is assumed that there is no measurement error and the problem is linear. Using an initial guess or *a priori*, , the retrieved parameter state can be found through

|  |  |
| --- | --- |
|  | (2.31) |
|  | (2.32) |

The Jacobian is represented by , and is the partial derivative of the forward model to the state vector. The Jacobian is an matrix. However, remote sensing methods are generally non-linear and contain measurement noise; additionally numerical approximations are usually used to calculate the forward model and Jacobian in realistic time frames. These issues make the direct method generally ineffective and iterative methods are used. This section will briefly cover common methods used for atmospheric inversions.

## 2.5.1 Optimal Estimation

Very commonly in atmospheric remote sensing, a Bayesian approach is used to update the atmospheric state, and this is known as optimal estimation (*Rodgers,* 2000). This method uses statistical knowledge of the *a priori* and state parameter with measurement noise to determine the probability that the state parameter is given a measurement of **,** and is given by

|  |  |
| --- | --- |
|  | (2.33) |

A solution is found by maximizing the probability of . If it can be assumed that the covariance of the measurement vector and *a priori* have a Gaussian distribution, the above can be solved to yield

|  |  |
| --- | --- |
|  | (2.34) |

The covariance matrices of the *a priori* and measurement error are and respectively. For non-linear problems, this equation is iterated until a converged solution is found. It should be noted that if the covariance of the *a priori* is unknown or is not well modeled by a Gaussian distribution this method can provide less “optimal” results.

## 2.5.2 Levenberg-Marquardt

The Gauss-Newton iterative method has been classically used to solve for non-linear inversion problems. However, if the initial guess is far from the solution and if the solution space is not well described by a quadric the method will fail. *Levenberg* (1944) proposed another method for the non-linear least squares fit that was later modified by *Marquardt* (1963) given as

|  |  |
| --- | --- |
|  | (2.35) |

and is known as the Lebenberg-Marquardt algorithm. The damping factor, , reduces the step size in iteration to keep the problem in a linear region. If this damping factor is small the method approaches the Gauss-Newton method whereas if is large the method steps down the direction of the gradient descent. The diagonal matrix, , is a scaling matrix for the damping factor since the state vector may have different dimensions and magnitudes resulting in the benefit of larger step sizes in more linear areas of the solution space. However, the determination of the damping factor can be difficult to determine from numerical methods and usually an ad hoc method is used.

## 2.5.3 Multiplicative Algebraic Reconstruction Technique

The Multiplicative Algebraic Reconstruction Technique (MART), which is used for the operational processing of the OSIRIS data products, is a form of relaxation techniques similar to the well-known Chahine relaxation (*Chahine*, 1970). The MART algorithm has the modification of a weighting matrix, , which relates the importance of each measurement vector, , and tangent altitude, , to each retrieved state altitude, . The algorithm is given by

|  |  |
| --- | --- |
|  | (2.36) |

Unlike the previous methods, this method allows for computation of the state vector without the computation of the Jacobian or any matrix inversions allowing for a fast and efficient algorithm. In the OSIRIS aerosol algorithm, only one spectral measurement vector is used (*Bourassa et al.*, 2007; 2012b) and the measurement vector summation over *k* from Equation 2.36 can be dropped.

# 2.6 ALI Prototype Instrument and Stratospheric Balloon Flight

The work presented in this thesis is focused on the development of the Aerosol Limb Imager (ALI). The core concept of ALI is to use the limb scatter technique together with a rapid hyperspectral imaging design approach to ultimately provide global coverage of highly spatially resolved and accurate observations of stratospheric aerosol. The central design feature of ALI is the use of the novel AOTF technology, which provides the ability to rapidly select a filtered wavelength with no moving parts, an ideal feature for space application. AOTFs have recently been developed with large apertures and high quality crystals allowing for imaging capabilities. The AOTF technology operates efficiently in the red and near infrared region of the spectrum, which is well-matched for limb scatter sensitivity to aerosol and clouds (*Rieger et al.*, 2014) and has a typical spectral bandpass of 3‑6 nm, which is ideal for the broadband scattering characteristics of aerosol.

## 2.6.1 ALI Specifications

As mentioned previously, the ultimate goal of this work is the eventual realization of ALI on a future satellite mission. However, for the prototype design and test work presented here, the available flight opportunity was on stratospheric balloon. This choice of test platform has resulted in some specific design decisions for the prototype that would ultimately require revision for a satellite-based implementation.

Stratospheric balloon flights provide stable float altitude of 35 km. This results in a minimum vertical field of view for ALI of 6◦ in order to image the limb from the ground to the float altitude. For a satellite platform the required field of view would be substantially reduced to approximately 1◦ in order to cover the same range. The structural features in the aerosol distribution discussed in section 2.2 are generally less than one kilometer vertically and to resolve these features a minimum spatial resolution of 250 m is required. For this design, similar specifications are used for the horizontal (cross-track) direction.

Spectrally, aerosol scattering is a broadband phenomenon with spectrally smooth scattering cross sections such that high spectral resolution is not required; however, a large spectral range is desirable. Particle size sensitivity in limb scattering increases from the visible to the near infrared (*Rieger et al*., 2014). When combined with the single octave range capability of the AOTF, ALI is designed to measure the 600-1200 nm spectral range. Since the aerosol (Mie) scattering cross section varies slowly with wavelength, measurements every 25-50 nm within the range are required. Furthermore, this slow varying cross section also allows for low spectral resolution with a spectral bandpass requirement of less than 10 nm.

The nature of imaging limb scattering results in a situation where the lower altitudes have high signal and the upper altitudes have low signal. Therefore signal to noise levels are the most important at the high altitudes due to the exponential signal drop-off. The highest measurement altitudes (*i.e.* 30-35 km) require that the signal to noise ratio must remain above 5 to be able to use these high altitude measurements to determine aerosol extinction. Additionally, the system should be able to take a single image with an exposure time on the order of a second or less to achieve a high measurements density.

Finally, due to the limitations of the stratospheric balloon platform both mass, power consumption, and thermal rages are a concern. The balloon platform can only sustain so much mass and has a limited amount of power. As such, the total mass of ALI must remain below 50 kg and power consumption below 80 W is required. Furthermore, the thermal environment of the balloon at stratospheric altitudes is challenging with ambient temperatures that are colder than ‑40◦C. Additionally, heating from direct sunlight cause large thermal gradients in the instrument resulting in concerns of overheating. In order to mitigate these concerns ALI must be thermally stable within all equipment’s standard operating ranges.