Sensitivity Study on Canadian Air Quality Measurements from Geostationary Orbit

A Thesis Submitted to the
College of Graduate and Postdoctoral Studies
in Partial Fulfillment of the Requirements
for the degree of Master of Science
in the Department of Physics and Engineering Physics
University of Saskatchewan
Saskatoon

By
Lukas Fehr

Permission to Use

In presenting this thesis in partial fulfilment of the requirements for a Postgraduate degree from the University of Saskatchewan, I agree that the Libraries of this University may make it freely available for inspection. I further agree that permission for copying of this thesis in any manner, in whole or in part, for scholarly purposes may be granted by the professor or professors who supervised my thesis work or, in their absence, by the Head of the Department or the Dean of the College in which my 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 my written permission. It is also understood that due recognition shall be given to me and to the University of Saskatchewan in any scholarly use which may be made of any material in my thesis.

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

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

Or

Dean of the College of Graduate and Postdoctoral Studies
University of Saskatchewan
116 Thorvaldson Building, 110 Science Place
Saskatoon, Saskatchewan, S7N 5C9
Canada

Abstract

Tropospheric Emissions: Monitoring of Pollution (TEMPO) is a satellite-based remote sensing air quality instrument destined for geostationary orbit over North America beginning in 2022. TEMPO will take hourly measurements with unprecedented resolution which will greatly benefit air quality forecasting, monitoring of emission sources, and health impact studies related to air quality. The field of regard of TEMPO contains a significant portion of Canada, including regions of particular interest such as major population centers and the Alberta oil sands. However, the standard retrieval algorithms that will be used to process TEMPO data do not explicitly account for some of the challenges that exist for measurements over Canada, such as pervasive snow cover, shallow lines of sight, and limited daylight hours. With the ultimate goal of creating new or optimized algorithms that address these challenges and allow Canada to take full advantage of TEMPO, standard retrieval algorithms for nitrogen dioxide and ozone have been replicated and studied. These algorithms use differential optical absorption spectroscopy (DOAS), the technique that will be used to create the standard TEMPO products, and they will serve as a baseline for comparison with future algorithms. The SASKTRAN radiative transfer framework, developed at the University of Saskatchewan, has been utilized to calculate air mass factors, key quantities in the DOAS-style retrieval, using three complementary methods which are all in agreement with each other. End-to-end retrievals modelled after cutting-edge algorithms used by modern instruments have been implemented, and they have been used to conduct a preliminary sensitivity study that quantifies the major sources of uncertainty in DOAS retrievals using synthetic TEMPO measurements.

Acknowledgements

I would like to firstly thank my supervisors Adam Bourassa and Doug Degenstein, for providing guidance and freedom in appropriate measures, and for being excellent role models as teachers, researchers, supervisors, and people. I am also indebted to the other members of our group, who are always generous with their time and expertise.

I am grateful to Chris McLinden and Environment and Climate Change Canada for funding and for valuable collaboration opportunities.

I would like to thank my parents for always enabling and encouraging me to pursue my interests throughout my life.

Finally, I am especially thankful for my wife Kiara and her constant encouragement and support, and for the ways she continually inspires me.

Contents

List of Tables

List of Figures

List of Abbreviations

AMF	Air mass factor
BRDF	Bidirectional reflectance distribution function
CDF	Cumulative density function
CFC	Chloroflurocarbon
DALY	Disability-adjusted life-years
DO	Discrete Ordinates
DOAS	Differential optical absorption spectroscopy
DOMINO	Dutch OMI NO_2
ECCC	Environment and Climate Change Canada
FOR	Field of regard
GEMS	Geostationary Enviornment Monitoring Spectrometer
GEO-CAPE	Geostationary Coastal and Air Pollution Events
GEOS-5 ESM	NASA Goddard Earth Observation System Model version 5 Earth System Model
GOME	Global Ozone Monitoring Instrument
HR	High Resolution
HSP	Hybrid sigma-pressure
LIDORT	Linearized Discrete Ordinates Radiative Transfer
LER	Lambert Equivalent Reflectivity
MC	Monte Carlo
NASA	National Aeronautics and Space Administration
OMI	Ozone Monitoring Instrument
OMPS	Ozone Mapping and Profiler Suite
OSIRIS	Optical Spectrograph and InfraRed Imaging System
PDF	Probability density function
SCD	Slant column density
SCIAMACHY	SCanning Imaging Absorption SpectroMeter for Atmospheric CHartographY
SNR	Signal-to-noise ratio
SP2	Version 2.1 Ozone Monitoring Instrument NO_2 Standard Product
SZA	Solar zenith angle
TEMPO	Troposphierc Emission: Monitoring of POllution
TOMS	Total Ozone Mapping Spectrometer
TROPOMI	TROPOspheric Monitoring Instrument
VCD	Vertical column density
VMR	Volume mixing ratio
VZA	Viewing zenith angle

1 Introduction

The atmosphere of the Earth is made primarily of nitrogen and oxygen, which together account for over 99 % of dry air. Most of the remaining 1 % is argon, a non-reactive gas, with well under 0.05 % of the atmosphere belonging to the remaining species which are collectively referred to as trace gases. Despite their low concentrations, trace gases have a tremendous impact on life on Earth. Greenhouse gases, which include water vapour in addition to trace gases such as carbon dioxide and methane, trap heat in the atmosphere by absorbing infrared radiation that would otherwise escape into space. Aerosols, referring to larger particles such as desert dust and forest fire smoke, cool the Earth by reflecting sunlight away before it can be absorbed. Trace gases affect more than just temperature: ozone absorbs harmful solar ultraviolet rays before they reach the surface, and many trace gases that pollute air near the surface have significant negative impacts on human health. Because life as we know it is tied so intricately to the trace gases found in our atmosphere, monitoring the composition of the atmosphere is crucial.

The importance of this task is augmented by the fact that manmade emissions contribute significantly to atmospheric composition; accurately determining and predicting the impact of human activity on climate is arguably the biggest task of atmospheric study in modern times. Careful measurements are necessary to identify atmospheric conditions that pose a threat to human health and ways of life so that policy changes can be put into place which will begin to rectify such conditions. Continued measurements are necessary to evaluate the effectiveness of policy changes, and also to provide constraints for atmospheric models. These models are powerful tools which are used to understand the atmosphere and to attempt to predict the future, potentially allowing dangerous conditions to be mitigated before they occur. Additionally, measurements and models are important for everyday activities, as air quality

forecasting systems around the world warn citizens about unhealthy atmospheric conditions. Poor air quality due to manmade pollution is a significant contributor to mortality; it is estimated that globally in 2015, exposure to particles with aerodynamic diameter less than 2.5 µm ($PM_{2.5}$) accounted for 4.2 million deaths and 103.1 million disability-adjusted life-years (DALYs), and exposure to ozone accounted for 254 000 deaths and 4.1 million DALYs, where DALYs are an estimate of the years lost to ill health, disability, or premature death (Cohen et al., 2017).

The discovery of the ozone hole in 1985, followed by the Montreal protocol of 1987, highlights the destructive potential of anthropogenic emissions, the ability of global policy to mitigate this destructive potential, and the importance of measuring atmospheric composition. In 1974 it was predicted that chlorofluorocarbons (CFCs), which were being widely used as coolants and spray can propellants, could deplete stratospheric ozone. Because the molecules are extremely stable, it was proposed that they would live long enough to be transported to the stratosphere, where they would be broken apart by solar radiation, at which point the chlorine atoms would react destructively with the ozone layer (Molina & Rowland, 1974). This was verified in 1985 when dramatic losses in ozone over the Antarctic were discovered using a ground-based spectrometer (Farman, Gardiner, & Shanklin, 1985). Data from satellite instruments such as the Total Ozone Mapping Spectrometer (TOMS) played a crucial role in validating their conclusions and demonstrating the extent of the so-called ozone hole. In the Montreal protocol of 1987, many countries agreed to phase out ozone depleting substances, and decades later the ozone hole is showing signs of recovery. It is worth noting that the conclusion that the ozone hole is healing is non-trivial, and is the result of decades of meticulous, continuous measurement of ozone concentrations and careful analysis by many international scientific experts (World Meteorogical Organization, 2018).

Today many different trace gas measurement techniques are employed regularly. In situ techniques, in which the measurement is made at the same location as the instrument, are used on the surface for air quality measurements, and on aircraft or high altitude balloons to take measurements higher up in the atmosphere. In situ measurements give very accurate results, but they cannot practically be used to cover large areas. Remote sensing techniques, in

which the measurement is made at some distance from the instrument, overcome some of the limitation of in situ techniques. Remote sensing techniques utilize spectroscopy, which takes advantage of the fact that different molecules absorb light better at certain wavelengths and therefore leave a distinct signature on the light that passes through. Careful measurements of sunlight that has interacted with the atmosphere therefore contain information about the composition of the atmosphere. Spectrometers which perform such measurements can be placed on aircraft or high altitude balloons for short term measurements, but also on the ground or in orbit around the Earth for long term measurements. Ground-based remote sensing is an effective option for taking frequent and long term measurements in areas of interest such as large cities, but like in situ techniques it is not practical to use to cover large areas. Space-based remote sensing is more difficult, but it is the only option that can continuously cover large regions, even the entire globe. Having measurements that are taken on a global scale is essential for understanding the chemistry and the transport mechanisms of the atmosphere, and for answering the big questions regarding global climate change.

Atmospheric composition has been measured from space since 1970 when the American backscatter ultraviolet ozone experiment was launched. While instrument quality has certainly increased steadily over the decades, until now all space-based measurements have been made from polar orbits, in which the satellite orbits over the poles while the Earth spins underneath it. Polar orbits permit global coverage, but are fundamentally limited to one measurement per day in any given location. The newest generation of space-based atmospheric composition instruments will fly in geostationary orbits, in which the satellite orbits at the same rate that the Earth rotates and therefore appears to hover over a single location. From geostationary orbits, coverage will be continental, not global, but many measurements can be taken per day at the same location. Tropospheric Emissions: Monitoring of Pollution (TEMPO) is the name of the instrument designed by the National Aeronautics and Space Administration (NASA) which will take such measurements over North America, planned for launch in 2022.

The purpose of this thesis is to begin the process of developing algorithms for analyzing TEMPO measurements of the atmosphere over Canada. These algorithms, known as re-

trievals, are given measurements of sunlight which has passed through the atmosphere, and are tasked with extracting information about the composition of the atmosphere. While TEMPO has inherited such algorithms from similar instruments which came before it, and these official algorithms will be used to analyze all TEMPO measurements including those over Canada, they fail to address some of the key challenges that are specific to these northern measurements. The primary challenges include taking measurements over scenes with pervasive snow cover and limited daytime hours, while viewing these scenes from a shallow angle as opposed to from above. The first step in addressing these unique challenges, and also the scope of this thesis, is to establish and test existing algorithms as a baseline for comparison with future experimental algorithms. Section 2 contains background information that is pertinent to such algorithms, including information on TEMPO itself, on radiative transfer, which refers to the study of light propagating through a medium such as the atmosphere, and on differential optical absorption spectroscopy (DOAS), the primary retrieval technique used by the existing algorithms. Section 3 describes the implementation of existing DOAS algorithms, and Section 4 describes the application of these algorithms in two recent studies. Firstly, one step of the algorithm was utilized by Environment and Climate Change Canada (ECCC) for their analysis of new satellite measurements over the Alberta oil sands, and secondly a sensitivity study was performed on synthetic TEMPO measurements which highlights major sources of uncertainty.

2 Background

2.1 Solar Backscatter Instruments

Measurements of trace gases in the atmosphere date back as far as 1840, when German-Swiss chemist Christian Friedrich Schönbein made semi-quantitative measurements of ozone using chemical techniques (Platt & Stutz, 2008). Spectroscopic techniques were not applied until the 1920s, when French physicists Fabry and Buisson discovered that ozone was responsible for the absorption of ultraviolet sunlight, and correctly determined that the total vertical column is equivalent to a layer of pure ozone approximately 3 mm thick (1921). Their techniques were soon improved upon by Gordon Dobson (Dobson & Harrison, 1926), who would go on to make many important contributions to the study of atmospheric ozone including the invention of the Dobson spectrometer. These spectrometers measured ultraviolet sunlight at discrete wavelengths and determined ozone amounts from ratios of these intensities. This measurement concept was first sent to space in 1970 with the Backscatter Ultraviolet ozone experiment which flew aboard the American Nimbus-4 satellite. This experiment spawned the Solar Backscatter Ultraviolet and TOMS series of instruments, which beginning in 1978 provided multiple-decade global ozone data products (Platt & Stutz, 2008). In 1995 the Global Ozone Monitoring Experiment (GOME) started the next generation of satellite-based spectrometers, which moved away from the discrete wavelength concept and instead measured a wide spectrum at a modest spectral resolution. Different techniques such as differential optical absorption spectroscopy (DOAS) were required to analyze these measurements, which now contained information on a variety of trace gases in addition to ozone (Burrows et al., 1999).

TEMPO is built on the heritage of the satellite-based solar backscatter instruments that

precede it, notably the instruments listed in Table 2.1. The majority of these instruments are nadir, pointing at the Earth, but some, such as the Optical Spectrograph and InfraRed Imaging System (OSIRIS) developed at the University of Saskatchewan, are limb instruments, pointing just above the surface of the Earth looking sideways through the atmosphere. Nadir instruments excel at mapping total columns with high horizontal spatial resolution, and they include information that extends right down to the surface, which is necessary to address concerns regarding air quality. The primary weakness of nadir measurements is that they contain very little information about the vertical distribution of the trace gas species being measured, so they must rely on assumptions and approximations which fundamentally limit the precision of the concentration measurement, regardless of the quality of the spectroscopic measurement. Limb instruments lack horizontal spatial resolution and boundary layer information, but excel at obtaining vertical distribution information, which reduces their dependence on assumptions and approximations.

Since GOME, the basic concept for nadir sounders has not changed: measure backscattered sunlight with a downward-pointing spectrometer placed in a sun-synchronous polar orbit that permits global coverage. Technological advancements have steadily improved spatial resolution, but temporal resolution is fundamentally limited to one measurement per day in sun-synchronous orbits, meaning that a single instrument is incapable of providing data useful for understanding diurnal variations. This limitation is overcome by the upcoming generation of geostationary instruments, which are capable of taking hourly measurements across entire continents.

2.2 TEMPO

TEMPO is an ultraviolet-visible solar backscatter spectrometer, designed to measure concentrations of trace atmospheric species over North America from a geostationary orbit. With a nominal ground pixel size of 2.1 $\times$ 4.4 km near the center of its field of regard (FOR), it has unmatched spatial resolution, and along with its counterparts Sentinel-4 and the Geostationary Environment Monitoring Spectrometer (GEMS) which will fly over Europe and Asia respectively, it has unprecedented temporal resolution. TEMPO evolved from

Instrument	Orbit	View	Wavelengths (nm)	Ground Pixel (km)	Operation
GOME (Burrows et al., 1999)	polar	nadir	240-790	40 × 320	1995-2011
OSIRIS (Llewellyn et al., 2004)	polar	limb	280-800	N/A	2001-present
SCIAMACHY (Bovensmann et al., 1999)	polar	both	240-1700 1940-2040 2265-2380	30 × 240	2002-2012
OMI (Levelt et al., 2006)	polar	nadir	270-500	15 × 30	2004-present
GOME-2 (Munro et al., 2016)	polar	nadir	240-790	40 × 10	2006-present
OMPS (Flynn, Seftor, Larsen, & Xu, 2006)	polar	both	250-380	50 × 50	2011-present
TROPOMI (Veefkind et al., 2012)	polar	nadir	270-500 675-775 2305-2385	7 × 7	2017-present
GEMS (Kim et al., 2019)	geo	nadir	300-500	3.5 × 8	2020
TEMPO (Zoogman et al., 2017)	geo	nadir	290-490 540-740	2.1 × 4.4	2022
Sentinel-4 (Gulde et al., 2014)	geo	nadir	305-500 750-775	8 × 8	2023

Table 2.1: Satellite-based solar backscatter instruments.

the Geostationary Coastal and Air Pollution Events (GEO-CAPE) mission concept, which was conceived following the 2007 National Research Council Decadal Survey "Earth Science and Applications from Space". Originally planned for launch in between 2013 and 2016, GEO-CAPE aimed to address ocean coastal biophysics and atmospheric pollution chemistry, specifically the effects of anthropogenic activity in these systems. TEMPO has been designed to meet as many of the pollution-related requirements of this original mission while keeping costs down by partnering with a commercial geostationary satellite platform (Zoogman et al., 2017).

TEMPO measures light from 290-490 nm and from 540-740 nm with a spectral resolution of approximately 0.6 nm and with spectral sampling of approximately 0.2 nm. It takes images along the latitudinal direction and uses a scan mirror to step across the continent

in the longitudinal direction. The total FOR has a height of 4.82°, spanning latitudes from 20°N to 60°N, and a width of 8.38°, which encompasses the width of the continent. See Figures 2.1 and 2.2 for visualization of the FOR. TEMPO data products will include ozone (profile and total column), NO_2, SO_2, H_2CO, $C_2H_2O_2$, H_2O, BrO, IO, aerosol properties, cloud parameters, ultraviolet B radiation, and foliage properties (Zoogman et al., 2017). The instrument was completed by Ball Aerospace and Technologies Corporation in December of 2018 (Atkinson, 2018), and Maxar Technologies was secured as the host satellite provider in July of 2019, with an anticipated launch in 2022 (Northon, 2019).

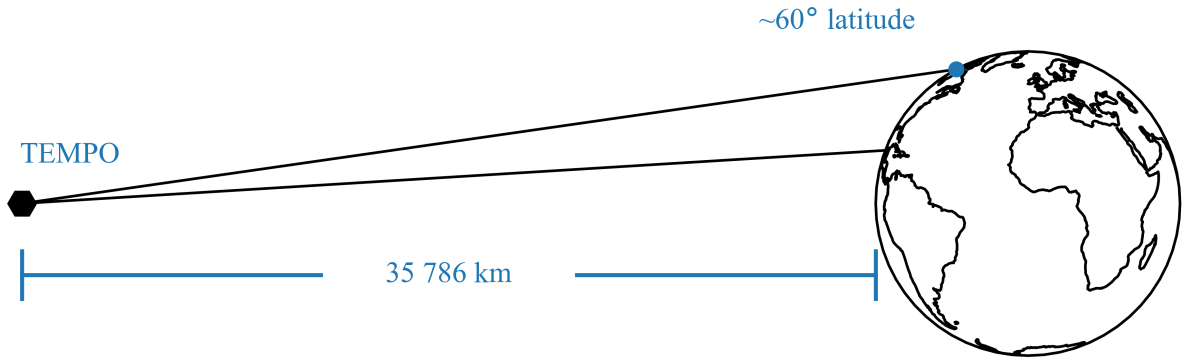


Figure 2.1: Side profile of the orbit and field of regard of TEMPO.

The motivation for TEMPO includes more accurate population exposure monitoring, more effective emission-control strategies, and improved emission inventories, databases that track pollution emissions from significant sources which are used to inform regulatory action and to drive atmospheric models. These goals will be met as the sources, sinks, and transport mechanisms of key pollutants are characterized through measurements of their concentration with sufficient precision, resolution, and coverage. The high spatial resolution of TEMPO will allow for emission sources to be resolved down to a sub-urban scale, and the high temporal resolution will allow for diurnal variation to be monitored and studied in addition to seasonal variation and long term trends. (Zoogman et al., 2017).

The motivation for analyzing TEMPO data over Canada includes the base reasons listed above, but also extends to more specific objectives such as assimilation with the current

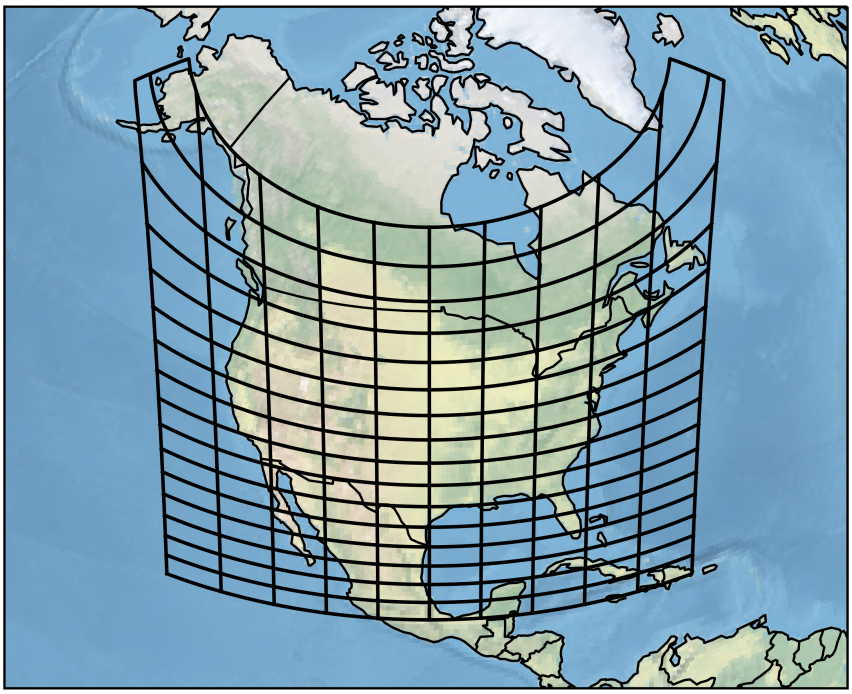


Figure 2.2: TEMPO field of regard. Each box represents 125 by 125 pixels.

air quality forecast managed by ECCC, monitoring emissions over the oil sands, conducting health impact studies, and tracking aerosols from forest fires. However, there are several challenges standing in the way of retrieving accurate trace gas quantities over Canada. The first challenge is the northern latitude of Canada and its relation to the geostationary orbit that TEMPO will fly in. As illustrated by Figure 2.1, the northern edge of the FOR skims the surface, in contrast to the southern region where the line of sight is much closer to the local vertical. With these large viewing zenith angles (VZAs) over Canada, the curvature of the Earth and the atmosphere have a larger effect on the radiative transfer. SASKTRAN, the radiative transfer model developed at the University of Saskatchewan (Bourassa, Degenstein, & Llewellyn, 2008; Dueck, Bourassa, & Degenstein, 2017; Zawada et al., 2015), is currently well equipped to investigate these effects as it is able to perform radiative transfer calculations with both a fully spherical atmosphere and a plane-parallel atmosphere. The solar zenith angle (SZA) will also be large as a result of the northern latitudes in Canada, especially during winter, which will contribute further to any spherical effects and also significantly reduce signal-to-noise ratio (SNR) as fewer photons will reach the detector. One final challenge that is unique to the Canadian portion of the FOR of TEMPO is the presence of snow. Snow is

challenging for tropospheric retrievals because it often interferes with the interpretation of clouds and aerosols, the dependence of the reflective properties on age, depth and landscape is not well understood, and snow cover data products can be unreliable. Many algorithms simply flag results over snow as unreliable, but such an approach would render most of the winter data unusable.

2.3 Radiative Transfer

The field of radiative transfer deals with the propagation of light through a medium, including processes such as absorption, emission and scattering. It originated in the beginning of the 20th century when some of the underlying concepts were developed by astrophysicists, such as the early atmospheric scattering calculations performed by Schuster (1905). Chandrasekhar published a book mid-century which would become the standard for the field, describing radiative transfer in atmospheres as a branch of mathematical physics (1950), and the application of radiative transfer to remote sensing of the atmosphere of the Earth was first suggested by Kaplan when he described the possibility of inferring temperatures and water vapour concentrations from infrared radiation viewed by an aircraft or a satellite (1959). Accurate modelling of radiative transfer is a crucial component of the retrievals used for all solar backscatter instruments.

2.3.1 Radiance

Modelling of radiative transfer is done primarily with the radiometric quantity known as spectral radiance. Consider radiant energy $dE_{\hat{\mathbf{\Omega}}}$ in the spectral window from λ to $\lambda + d\lambda$ passing through area dA over a time interval dt and propagating in direction $\hat{\mathbf{\Omega}}$ within the solid angle $d\Omega$ as shown in Figure 2.3. The angle between $\hat{\mathbf{\Omega}}$ and the normal to dA is labelled θ. Spectral radiance I is then defined as the radiant power per unit solid angle per unit projected area per unit wavelength,

$$I = \frac{dE_{\hat{\mathbf{\Omega}}}}{dt\, d\Omega\, dA\, \cos\theta\, d\lambda}\,, \tag{2.1}$$

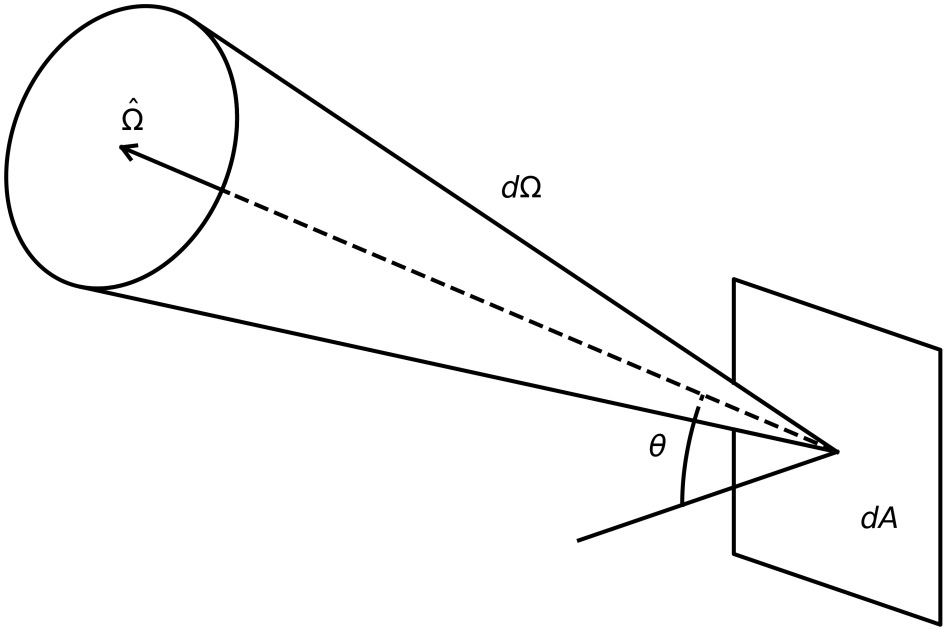


Figure 2.3: Radiance visualization.

where the projected area is the area of dA projected onto the plane perpendicular to $\hat{\mathbf{\Omega}}$. The SI unit for spectral radiance is W/sr/m^3, but alternative units which are more compatible with instrument measurements such as photons/s/sr/cm^2/nm are frequently used. Spectral radiance has the useful property that it remains constant as it propagates through a vacuum, which means that any changes to radiance are always due to interaction with a medium. Spectral radiance is also intuitively compatible with instrument measurements: by integrating over solid angle, wavelength, time, and area, all of which can be defined for a given measurement made by a given instrument, the number of photons entering the system can be calculated. These integration regions are typically small enough that the radiance can be assumed to be approximately constant, so the integration becomes multiplication which can easily be reversed. Therefore instruments which count photons can easily estimate radiance.

A second useful radiometric quantity, called spectral irradiance, exists and can be defined in a similar fashion, considering the radiant energy dE in the spectral window λ to $\lambda + d\lambda$ passing through area dA over a time interval dt from all directions. It can also be defined relative to the spectral radiance: the normal component of the spectral radiance integrated

over the entire hemisphere:

$$F = \frac{dE}{dt\, dA\, d\lambda} = \int_{2\pi} I\, cos\theta\, d\Omega\,. \tag{2.2}$$

Irradiance is used to describe approximately collimated light such as direct sunlight, as specifying the angular dependence of direct sunlight within its angular width of about half a degree is unnecessary for most radiative transfer applications. For brevity, the spectral prefix will be dropped from both radiance and irradiance for the remainder of this document, but unless is it stated otherwise the spectral dependence will remain implicit.

Assuming that emissions and terrestrial sources are negligible, which is a good assumption in ultraviolet and visible wavelengths during the daytime, all radiance in the atmosphere originates from the sun. Therefore it is often convenient to normalize the radiance by dividing it by the incident solar irradiance. This is beneficial for atmospheric remote sensing applications because most of the spectral features of the solar spectrum, which contain information on the solar atmosphere, disappear after normalization, and information from the Earth's atmospheric species remain. Inelastic processes such as Raman scattering, which is discussed in Section 2.3.4, are responsible for the faint solar signal that remains. This ratio $I(\lambda)/F(\lambda)$ has been called the reflectance spectrum, but there is some ambiguity since this same name has also referred to the quantity $\pi I(\lambda)/\mu_0 F(\lambda)$, where μ_0 is the cosine of the SZA. In more recent publications the direct ratio $I(\lambda)/F(\lambda)$ goes by different names such as the sun-normalized radiance (van Geffen, Eskes, Boersma, Maasakkers, & Veefkind, 2019) or simply the normalized spectrum (Bucsela et al., 2013), differentiating it from the quantity $\pi I(\lambda)/\mu_0 F(\lambda)$. The ratio $I(\lambda)/F(\lambda)$ will be called the normalized radiance for the remainder of this document.

2.3.2 Absorption

Absorption occurs when a photon is incident on a molecule and the molecule absorbs the energy of the photon, leaving the molecule in a higher energy state. The energy levels of a molecule are quantized, so absorption will only occur if a higher energy state exists such that

the difference in energy between the high energy state and the current state is equal to the energy of the photon. Thus the energy levels of a molecule define a set of absorption lines at specific wavelengths. After absorbing the photon, the excited molecule can dissociate into smaller molecules, or it can relax back into a lower energy state by emitting a photon or interacting with neighbouring particles.

Quantized molecular energy levels can be characterized by three different types of energy: electronic, vibrational, and rotational. Energy differences between electronic states are the largest, with purely electronic transitions often corresponding to ultraviolet and visible light. Purely vibrational transitions have lower energy and are typically in the infrared region, and purely rotational transitions have even lower energy, extending into the microwave region. Transitions of different types occur simultaneously, which results in absorptions bands containing many lines in close proximity. For example, a large electronic transition will combine with many smaller vibrational transitions to create a family of absorption lines near the wavelength corresponding to the pure electronic transition.

Ozone absorption is a crucial component of radiative transfer, as it significantly changes the intensity of radiation in the atmosphere, in contrast to many other gases which leave unique absorption signatures but have a negligible impact on the net radiant energy. The strongest absorption occurs in the Hartley bands at wavelengths in the 200 nm to 300 nm range, causing very little radiation in this window to reach the surface. Weaker absorption also occurs in the Huggins bands from 300 nm to 360 nm and the Chappuis bands from 440 nm to 1180 nm. These absorption bands correspond to electronic transitions in ozone which are immediately followed by dissociation into diatomic oxygen and oxygen atoms. This photodissociation plays a key role in atmospheric chemistry, and the reaction releases heat which drives the temperature in the upper atmosphere.

For an absorption event to occur, the photon energy must match the energy level difference to within some finite precision which is determined by several physical processes. This is called line broadening, as it creates a small window around the central wavelength in which absorption can occur. While natural line broadening, due to quantum uncertainty associated with

the lifetime of the excited state, does exist, it is negligible when compared to the following two processes. In the lower atmosphere, line broadening is dominated by collisions between the absorbing molecule and other molecules, which perturbs the energy levels and therefore broadens the line. In the upper atmosphere the Doppler effect becomes important as well, as the decrease in pressure results in fewer collisions which reduces collisional broadening. In Doppler broadening, the thermal velocity of the absorbing molecules is large enough to induce small Doppler shifts.

The absorption of a given molecule, accounting for all of these effects, is characterized by a quantity called the absorption cross section, which specifies the probability of a photon of a given wavelength being absorbed by a molecule at a given temperature and pressure. Cross sections have units of area, which lends itself to the analogy where the absorption event is likened to a random projectile striking a target: the probability of success is proportional to the area of the target. The absorption cross section will be represented by $\sigma(\lambda)$, explicitly showing the wavelength dependence but leaving implicit any temperature or pressure dependence, which is much weaker.

The cross section gives the probability of interaction with a single molecule; a new quantity is needed to characterize the probability of interaction with a medium. The absorption extinction $k(\boldsymbol{r}, \lambda)$ is the product of the cross section $\sigma(\lambda)$ and the volume number density $n(\boldsymbol{r})$,

$$k(\boldsymbol{r}, \lambda) = n(\boldsymbol{r})\sigma(\lambda) \tag{2.3}$$

where $n(\boldsymbol{r})$ is simply the number of molecules per unit volume. Although it is written here without position dependence, the cross section $\sigma(\lambda)$ has a slight dependence on position through its dependence on temperature and pressure. The extinction has units of inverse length, and represents the probability that a photon will be absorbed per unit length travelled by the photon.

Closely related to the extinction is the optical depth, a unitless quantity which represents the probability a photon will be absorbed anywhere along a given path. The optical depth

$\tau_C(\lambda)$ for path C is calculated by integrating the extinction along the path:

$$\tau_C(\lambda) = \int_C k(l, \lambda) dl \,. \tag{2.4}$$

The optical depth represents the attenuation of radiation along a path: each unit of optical depth corresponds to a reduction of the radiance by a factor of $1/e$. This property will be referred to as Beer's law, although it also goes by Bouguer's law, Lambert's law, or a combination of the three names. Mathematically this can be written as

$$I_C(\lambda) = I_0(\lambda) e^{-\tau_C(\lambda)} \,, \tag{2.5}$$

or

$$\ln I_C(\lambda) = \ln I_0(\lambda) - \tau_C(\lambda) \,, \tag{2.6}$$

where $I_0(\lambda)$ is the original radiance and $I_C(\lambda)$ is the radiance after it has been attenuated by absorption.

2.3.3 Rayleigh Scattering

Rayleigh scattering describes the interaction between a photon and a molecule which is much smaller than the wavelength of the photon. Classically, the incident oscillating electric field induces an oscillating electric dipole moment on the small molecule, which modifies the incident field, resulting in radiation being scattered in all directions. The induced dipole description is only valid when the molecule is small, so different theory is required for larger molecules. The polarizability of a molecule is represented by the parameter α, which is the constant of proportionality between the induced dipole moment and the applied electric field, with units of volume. Solving for the electric field given randomly polarized light with wavelength λ incident on a molecule with polarizability α gives the scattering cross section,

$$\sigma_s(\lambda) = \frac{\alpha^2 128 \pi^5}{3\lambda^4} \,, \tag{2.7}$$

and the scattering phase function,

$$p(\Theta) = \frac{3}{16\pi}(1 + \cos^2\Theta)\,, \tag{2.8}$$

which describes the angular distribution of the scattered radiation, where Θ is the scattering angle, as shown in Figure 2.4. The scattering phase function is normalized over the solid angle sphere, so that its integral over any solid angle region represents the probability that a scattered photon will pass through that region. The scattering cross section is inversely proportional to wavelength to the power of four, a dependence which is responsible for the blue color of the sky, as the blue end of the spectrum is more likely to be scattered than the red end of the spectrum due to its shorter wavelength.

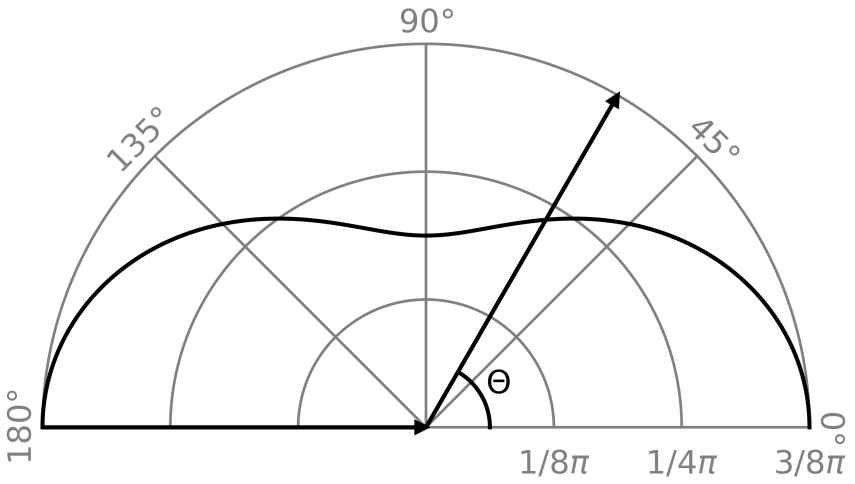


Figure 2.4: Rayleigh scattering phase function for randomly polarized light, as displayed in Equation (2.8). Scattering is most probable in the forward ($\Theta = 0°$) and backward ($\Theta = 180°$) directions. The phase function has azimuthal symmetry such that scattering into any direction within a cone of constant scattering angle is equally likely.

These two terms can be used to create a scattering source term, which describes the scattering contribution per unit length along a line of sight. Consider a scattering volume dV with thickness ds along direction $\hat{\mathbf{\Omega}}$ and cross sectional area dA perpendicular to $\hat{\mathbf{\Omega}}$ which is scattering incoming irradiance $I'(\hat{\mathbf{\Omega}}')d\Omega'$, resulting in outgoing radiance $dI_{s,\hat{\mathbf{\Omega}}'}(\hat{\mathbf{\Omega}})$ within solid angle $d\Omega$, shown in Figure 2.5. The volume contains scattering particles with number density n.

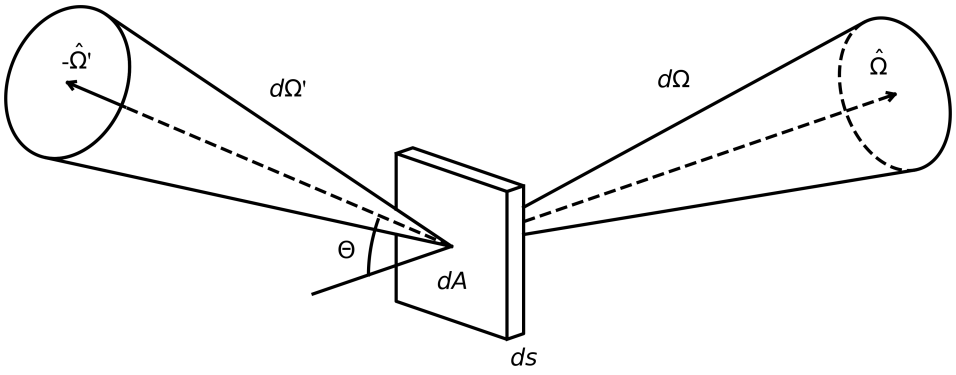


Figure 2.5: Scattering source visualization.

The incident power dP' on the volume dV is given by

$$dP' = I'(\hat{\mathbf{\Omega}}') \cos\Theta \, dA \, d\Omega' \, . \tag{2.9}$$

The scattering optical depth $d\tau_{\text{scat}}$, which is equal to the fraction of the incident power that will be scattered, is the product of the scattering extinction k_{scat} and the path length of the incident beam within the volume:

$$d\tau_{\text{scat}} = k_{\text{scat}} \, dl = n\sigma_{\text{scat}} \frac{ds}{\cos\Theta} \, . \tag{2.10}$$

The total scattered power is therefore $dP' d\tau_{\text{scat}}$, and $p(\Theta)d\Omega$ is the fraction of this scattered power that is scattered into the solid angle $d\Omega$. Therefore the scattered power in solid angle $d\Omega$, denoted dP, is given by

$$dP = dP' \, d\tau_{\text{scat}} \, p(\Theta) d\Omega = k_{\text{scat}} I'(\hat{\mathbf{\Omega}}') p(\Theta) \, dA \, d\Omega \, d\Omega' \, ds \, . \tag{2.11}$$

The corresponding outgoing radiance $dI_{s,\hat{\mathbf{\Omega}}'}$ is then given by

$$dI_{s,\hat{\mathbf{\Omega}}'}(\hat{\mathbf{\Omega}}) = \frac{dP}{dA \, d\Omega} = k_{\text{scat}} I'(\hat{\mathbf{\Omega}}') p(\Theta) d\Omega' \, ds \, . \tag{2.12}$$

The outgoing radiance due to scattering within segment ds from all incoming directions is

found by integrating the previous expression over the solid angle sphere:

$$dI_s(\hat{\mathbf{\Omega}}) = k\left(\omega_0 \int_{4\pi} I'(\hat{\mathbf{\Omega}}')p(\Theta)d\Omega'\right)ds \equiv kJ(\hat{\mathbf{\Omega}})ds\,, \tag{2.13}$$

where ω_0 is the single scatter albedo, defined as the ratio between the scattering extinction k_{scat} and the total extinction $k = k_{\text{scat}} + k_{\text{abs}}$ which accounts for both scattering and absorption, and where the defined term $J(\hat{\mathbf{\Omega}})$ is called the scattering source term.

2.3.4 Rotational Raman Scattering

Rotational Raman scattering is inelastic scattering where energy is exchanged between the photon and the molecule, resulting in a change in the quantum rotational state of the molecule and a change in the wavelength of the photon. Rotational Raman scattering by diatomic nitrogen (N_2) and oxygen (O_2) is of interest for remote sensing in ultraviolet and visible wavelengths as these processes introduce non-negligible features to measured spectra.

Rotational Raman transitions will only occur when a certain set of selection rules are met. The gross selection rule, which specifies the general physical properties that must exist in order for the phenomena to occur, is that the polarizability of the molecule must be anisotropic. Linear molecules such as N_2 and O_2 meet this criteria. The specific selection rule, which governs the allowed quantum number transitions, for linear molecules requires that $\Delta J = 0, \pm 2$, where J is the quantum number for the total angular momentum of the molecule. Note that the $J = 0$ transitions correspond to Rayleigh scattering and are not of interest here. The following theory is based on the work of Chance and Spurr (1997).

The rotational Raman transitions that are relevant for atmospheric radiative transfer occur when the N_2 and O_2 molecules are in the $\Lambda = 0$ electronic state (denoted Σ), where Λ is the quantum number representing the projection of the electronic angular momentum on the internuclear axis. They are described using Hund's case b, a limiting case in which the spin-orbit coupling is assumed to be negligible and only the spin-rotation coupling is accounted for. Under these assumptions, the relevant good quantum numbers (in addition to Λ) are J, N, and S, where N describes the total angular momentum excluding contributions from

spin, and S describes the angular momentum from spin. The spin quantum number is $S = 0$ for N_2 and $S = 1$ for O_2, and the remaining allowed quantum numbers are $N = \Lambda, \Lambda + 1, \ldots$ and $J = |N - S|, |N - S| + 1, \ldots, N + S$, with the further constraint $J \geq 0$.

For a Hund's case b molecule in a Σ state, cross sections for a scattering event following a transition from a state described by N and J to a state described by N' and J' are given by

$$\sigma_{N,J\to N',J'}(\lambda) = \frac{256\pi^5}{27\lambda^4}\gamma^2 f(N,J)b(N,N',J,J')\,, \tag{2.14}$$

where the change in energy of the scattered photon ΔE is equal to the difference between the energy of the initial and final states: $\Delta E = E(N,J) - E(N',J')$. Here, γ represents the polarizability anisotropy of the scattering molecule; Chance and Spurr have provided empirical functions for N_2,

$$\gamma_{N_2} \times 10^{25} = -6.01466 + \frac{2385.57}{186.099 - \sigma^2}\,, \tag{2.15}$$

and for O_2,

$$\gamma_{O_2} \times 10^{24} = 0.07149 + \frac{45.9364}{48.2716 - \sigma^2}\,, \tag{2.16}$$

where σ is the wavenumber in units of μm^{-1}.

The fractional population of the initial state $f(N,J)$ is given by

$$f(N,J) = \frac{g(N)(2J+1)\exp(-E(N,J)/kT)}{Z}\,, \tag{2.17}$$

where $g(N)$ is the nuclear spin statistical weight factor, k is the Boltzmann constant, T is the temperature, and Z is the state sum which is the sum of the numerator in Equation (2.17) over all states. Finally, $b(N,N',J,J')$ is called a Placzek-Teller coefficient, which for a Hund's case b molecule is given by

$$b(N,N',J,J') = (2N+1)(2N'+1)(2J'+1)\begin{pmatrix} N & 2 & N' \\ 0 & 0 & 0 \end{pmatrix}^2 \begin{Bmatrix} N & 2 & N' \\ J' & S & J \end{Bmatrix}^2\,, \tag{2.18}$$

where the last two terms follow the standard definitions for 3-j and 6-j Wigner coefficients respectively.

For N_2, the case b formulation is sufficient, but for O_2 the spin-rotation coupling causes deviations from this limiting case. To find the necessary corrections to the state energies and the Placzek-Teller coefficients, perturbation theory must be applied while adding the energy of the spin-rotation interaction to the Hamiltonian. This calculation is outside the scope of this thesis.

2.3.5 Ground Reflectance

Ground reflectance is described by a function called the bidirectional reflectance distribution function (BRDF). The BRDF $B(\hat{\mathbf{\Omega}}, \hat{\mathbf{\Omega}}')$ captures the full angular dependence of a surface's reflective properties, defining the reflectivity for every combination of incoming directions $\hat{\mathbf{\Omega}}'$ and outgoing directions $\hat{\mathbf{\Omega}}$ over the hemisphere. It is defined as the ratio of the outgoing radiance to the incoming power density:

$$B(\hat{\mathbf{\Omega}}, \hat{\mathbf{\Omega}}') = \frac{dI(\hat{\mathbf{\Omega}})}{I(\hat{\mathbf{\Omega}}')\mu(\hat{\mathbf{\Omega}}')d\Omega'} \,. \tag{2.19}$$

where $\mu(\hat{\mathbf{\Omega}}')$ is the cosine of the zenith angle made by direction $\hat{\mathbf{\Omega}}'$. The differential terms $dI(\hat{\mathbf{\Omega}})$ and $d\Omega'$ account for the fact that the radiance $I(\hat{\mathbf{\Omega}})$ in any given direction is made up of outgoing contributions $dI(\hat{\mathbf{\Omega}})$ resulting from infinitesimal solid angles $d\Omega'$ from all incoming directions $\hat{\mathbf{\Omega}}'$. Therefore the total $I(\hat{\mathbf{\Omega}})$ can be found by integrating over the hemisphere:

$$I(\hat{\mathbf{\Omega}}) = \int_{2\pi} \frac{dI(\hat{\mathbf{\Omega}})}{d\Omega'} d\Omega' \,. \tag{2.20}$$

Substituting Equation (2.19) into Equation (2.20) gives an expression for the outgoing radiance in terms of the BRDF and the incoming radiance:

$$I(\hat{\mathbf{\Omega}}) = \int_{2\pi} B(\hat{\mathbf{\Omega}}, \hat{\mathbf{\Omega}}') I(\hat{\mathbf{\Omega}}') \mu(\hat{\mathbf{\Omega}}') d\Omega' \,. \tag{2.21}$$

If a surface is illuminated by collimated light, the outgoing radiance $I(\hat{\mathbf{\Omega}})$ can be calculated without an integral,

$$I(\hat{\mathbf{\Omega}}) = B(\hat{\mathbf{\Omega}}, \hat{\mathbf{\Omega}}')\mu(\hat{\mathbf{\Omega}}')F(\hat{\mathbf{\Omega}}')\,, \tag{2.22}$$

where $\hat{\mathbf{\Omega}}'$ is the direction of the incident light and $F(\hat{\mathbf{\Omega}}')$ is the incoming collimated irradiance.

Since the BRDF describes reflection only, the outgoing energy should never exceed the incoming energy. Consider a differential surface with area dA illuminated by collimated light $F(\hat{\mathbf{\Omega}}')$, and define the albedo $a(\hat{\mathbf{\Omega}}')$ as the ratio of outgoing to incoming power:

$$a(\hat{\mathbf{\Omega}}') \equiv \frac{\displaystyle\int_{2\pi} \Big(\mu(\hat{\mathbf{\Omega}})I(\hat{\mathbf{\Omega}})dA\Big)\, d\hat{\mathbf{\Omega}}}{\mu(\hat{\mathbf{\Omega}}')F(\hat{\mathbf{\Omega}}')dA}\,. \tag{2.23}$$

To conserve energy this ratio must not exceed one. Substituting for $I(\hat{\mathbf{\Omega}})$ using Equation (2.22) and simplifying produces the constraint,

$$a(\hat{\mathbf{\Omega}}') = \int_{2\pi} \mu(\hat{\mathbf{\Omega}})B(\hat{\mathbf{\Omega}}, \hat{\mathbf{\Omega}}')d\hat{\mathbf{\Omega}} \leq 1\,, \tag{2.24}$$

which must be met by every physical BRDF for every incoming direction.

The BRDF is bound by Helmholtz reciprocity, which requires that incident and reflected rays of light behave identically if the direction is reversed. As a result, swapping incoming and outgoing directions leaves the BRDF unchanged: $B(\hat{\mathbf{\Omega}}, \hat{\mathbf{\Omega}}') = B(\hat{\mathbf{\Omega}}', \hat{\mathbf{\Omega}})$. A consequence of this is that the constraint shown in Equation (2.24) also applies when the integral is taken over all incoming directions.

2.3.6 Equation of Radiative Transfer

The differential form of the equation of radiative transfer is

$$\frac{dI}{ds} = -k(s)I(s) + J(s)\,, \tag{2.25}$$

where the first term represents losses in radiance due to extinction $k(s)$, which accounts for radiation being both absorbed and scattered, and the second term represents gains in radiance due to the source term $J(s)$. In general the source term consists of elastic scattering, emissions, and inelastic scattering, but for the present work only elastic scattering is considered. Note that the inelastic scattering background in Section 2.3.4 is not included for the purpose of explicit forward modelling, as SASKTRAN is not currently capable of modelling inelastic scattering, but rather because it is used in DOAS retrievals (see Section 3.2). A diffuse radiance field has five dimensions, three spatial and two directional, and must satisfy Equation (2.25) equation in every direction at every point in space.

The integral form of this equation, now with the spatial and directional coordinates shown explicitly, takes the following form, based on the theoretical development presented by Bourassa et al. (2008):

$$I(\boldsymbol{r}, \hat{\boldsymbol{\Omega}}) = \int_{s_1}^{0} k(\boldsymbol{r}_s) J(\boldsymbol{r}_s, \hat{\boldsymbol{\Omega}}) e^{-\int_s^0 k(\boldsymbol{r}_t) dt} ds + \tilde{I}(\boldsymbol{r}_{s_1}, \hat{\boldsymbol{\Omega}}) e^{-\int_{s_1}^0 k(\boldsymbol{r}_s) ds} \tag{2.26a}$$

$$J(\boldsymbol{r}, \hat{\boldsymbol{\Omega}}) = \omega_0(\boldsymbol{r}) \int_{4\pi} I(\boldsymbol{r}, \hat{\boldsymbol{\Omega}}') p(\boldsymbol{r}, \hat{\boldsymbol{\Omega}}, \hat{\boldsymbol{\Omega}}') d\Omega' \tag{2.26b}$$

$$\tilde{I}(\boldsymbol{r}, \hat{\boldsymbol{\Omega}}) = \int_{2\pi} B(\boldsymbol{r}, \hat{\boldsymbol{\Omega}}, \hat{\boldsymbol{\Omega}}') I(\boldsymbol{r}, \hat{\boldsymbol{\Omega}}') \mu(\hat{\boldsymbol{\Omega}}') d\Omega' \,. \tag{2.26c}$$

The radiance $I(\boldsymbol{r}, \hat{\boldsymbol{\Omega}})$ at point $\boldsymbol{r}$ in direction $\hat{\boldsymbol{\Omega}}$ is the sum of the line integral of all source terms $J(\boldsymbol{r}_s, \hat{\boldsymbol{\Omega}})$ along the line of sight, which starts at $\boldsymbol{r}$ and points opposite to the direction $\hat{\boldsymbol{\Omega}}$, and the ground contribution $\tilde{I}(\boldsymbol{r}_{s_1}, \hat{\boldsymbol{\Omega}})$, which is zero if the line of sight does not terminate at the ground. Note that $\boldsymbol{r}_s \equiv \boldsymbol{r} + s\hat{\boldsymbol{\Omega}}$ is a point along the line of sight where s is negative, and the point $\boldsymbol{r}_{s_1}$ represents the termination of the line of sight either at the ground or at the defined top of atmosphere.

The source term $J(\boldsymbol{r}, \hat{\boldsymbol{\Omega}})$ consists of scattered light, as defined in Section 2.3.3, and the ground term $\tilde{I}(\boldsymbol{r}, \hat{\boldsymbol{\Omega}})$ consists of reflected light, as defined in Section 2.3.5. All contributions are attenuated back to the observer according to Beer's law, as defined in Section 2.3.2, using the total extinction $k(\boldsymbol{r})$, which is the sum of the scattering extinction $k_{\text{scat}}(\boldsymbol{r})$ and the absorption extinction $k_{\text{abs}}(\boldsymbol{r})$. The equations are heavily coupled: to find the radiance

at a given point in a given direction, the source functions at all points along the line of sight must be known, and to know the source function at a given point, the radiance in all directions at that point must be known. Solving these equations in a realistic atmosphere is no simple task, calling for carefully selected approximations and advanced numerical techniques. Three different methods for solving this equation, all of which are currently implemented in SASKTRAN, will be described in the upcoming sections.

2.3.7 Successive Orders Method

The successive orders method uncouples the integral equations by calculating source functions and radiances one order of scattering at a time. It is a Neumann series solution where each term in the series has a physical interpretation: a radiance field representing all sunlight that has been scattered a specific number of times. The primary advantage of this method over the discrete ordinate method is that it is not constrained to a plane-parallel atmosphere; it is compatible with a spherical atmosphere, or any atmosphere in theory, which better represents the true atmosphere of the Earth.

Direct sunlight is represented by the initial radiance field $I_0(\boldsymbol{r}, \hat{\boldsymbol{\Omega}})$:

$$I_0(\boldsymbol{r}, \hat{\boldsymbol{\Omega}}) = F_0(\hat{\boldsymbol{\Omega}}_0) e^{-\int_{s_2}^{0} k(\boldsymbol{r}+t\hat{\boldsymbol{\Omega}}_0)dt} \delta(\hat{\boldsymbol{\Omega}} - \hat{\boldsymbol{\Omega}}_0)\,, \tag{2.27}$$

where $F_0(\hat{\boldsymbol{\Omega}}_0)$ is the incident solar irradiance at the defined top of atmosphere, $\hat{\boldsymbol{\Omega}}_0$ is the direction of this irradiance, s_2 is a negative number such that $\boldsymbol{r} + s_2\hat{\boldsymbol{\Omega}}_0$ is the point where the irradiance that reaches point $\boldsymbol{r}$ enters the atmosphere, and $\delta(\hat{\boldsymbol{\Omega}} - \hat{\boldsymbol{\Omega}}_0)$ is a delta function that constrains the field to a single direction, ignoring the small angular dependence of the incident sunlight. The delta function has the following fundamental property:

$$\int_{\Omega} f(\hat{\boldsymbol{\Omega}})\delta(\hat{\boldsymbol{\Omega}} - \hat{\boldsymbol{\Omega}}_0)d\Omega = \begin{cases} f(\hat{\boldsymbol{\Omega}}_0) & \hat{\boldsymbol{\Omega}}_0 \in \Omega \\ 0 & \hat{\boldsymbol{\Omega}}_0 \notin \Omega \end{cases}, \tag{2.28}$$

where $f(\hat{\boldsymbol{\Omega}})$ is an arbitrary function and Ω is an arbitrary solid angle region.

Source terms $J_n(\boldsymbol{r}, \hat{\boldsymbol{\Omega}})$ and ground terms $\tilde{I}_n(\boldsymbol{r}, \hat{\boldsymbol{\Omega}})$ are computed using the radiance field of the previous order, and radiance fields $I_n(\boldsymbol{r}, \hat{\boldsymbol{\Omega}})$ for $n \geq 1$ are computed using the source terms and ground terms of the same order:

$$I_n(\boldsymbol{r}, \hat{\boldsymbol{\Omega}}) = \int_{s_1}^{0} J_n(\boldsymbol{r}_s, \hat{\boldsymbol{\Omega}}) k(\boldsymbol{r}_s) e^{-\int_s^0 k(\boldsymbol{r}_t) dt} ds + \tilde{I}_n(\boldsymbol{r}_{s_1}, \hat{\boldsymbol{\Omega}}) e^{-\int_{s_1}^0 k(\boldsymbol{r}_s) ds} \tag{2.29a}$$

$$J_n(\boldsymbol{r}, \hat{\boldsymbol{\Omega}}) = \omega_0(\boldsymbol{r}) \int_{4\pi} I_{n-1}(\boldsymbol{r}, \hat{\boldsymbol{\Omega}}') \bar{p}(\boldsymbol{r}, \hat{\boldsymbol{\Omega}}, \hat{\boldsymbol{\Omega}}') d\Omega' \tag{2.29b}$$

$$\tilde{I}_n(\boldsymbol{r}, \hat{\boldsymbol{\Omega}}) = \int_{2\pi} B(\boldsymbol{r}, \hat{\boldsymbol{\Omega}}, \hat{\boldsymbol{\Omega}}') I_{n-1}(\boldsymbol{r}, \hat{\boldsymbol{\Omega}}') \mu(\hat{\boldsymbol{\Omega}}') d\Omega' \,. \tag{2.29c}$$

To solve these equations numerically, a number of discretizations need to be made. The atmosphere must be divided into suitable cells to facilitate the line integrals required for Equation (2.29a). Grids must be chosen for the source terms, specifying the positions, incoming directions, and outgoing directions at which to calculate the terms for each order. The radiance field can be calculated at a given position and incoming direction by tracing a line of sight through the atmosphere and dividing it into segments according to the atmospheric cells, and then interpolating in order to get the source terms and atmospheric properties that are necessary to integrate each segment. The source terms can be calculated at each position by integrating over all incoming directions while scattering them into all outgoing directions.

It would be computationally expensive to fully specify a three dimensional atmosphere and a three dimensional grid for the source terms, and to follow this full process for each order of scattering, so approximations will be necessary for most applications. Approximations such as a horizontally homogeneous atmosphere or carefully selected lookup tables in the place of explicit calculations can speed up calculations significantly without seriously compromising accuracy.

2.3.8 Monte Carlo Method

The Monte Carlo method is a non-deterministic method that can be used to solve the equation of radiative transfer in a spherical atmosphere to an arbitrary precision. It traces photons along random paths through the atmosphere, guided by appropriate probability distributions, in such a way that the desired radiance is equal to the expectation value of the radiance from

each path. The primary advantage of the Monte Carlo method compared to the successive orders method is that its accuracy is not sensitive to spatial and angular grids, as these grids are not necessary for this technique. The primary disadvantage is that many photon paths must be traced in order to reduce the uncertainty in the final radiance, which typically results in much longer computation times. The following theory is based on the work of Zawada et al. (2015), which describes the Monte Carlo implementation in SASKTRAN.

Monte Carlo integration, described in Appendix B, can be used to evaluate the integrals in Equation (2.29) which define the radiance and source terms of a particular scattering order. It is convenient to first transform Equation (2.29a) under the change of variables from path length s to transmission $T(s) = e^{-\int_s^0 k(\boldsymbol{r}_s)ds}$,

$$I_n(\boldsymbol{r}, \hat{\boldsymbol{\Omega}}) = \int_{T(s_1)}^{1} J_n(\boldsymbol{r}_{s(T')}, \hat{\boldsymbol{\Omega}}) dT' + \tilde{I}_n(\boldsymbol{r}_{s_1}, \hat{\boldsymbol{\Omega}}) T(s_1)\,, \tag{2.30}$$

where the inverse function $s(T)$ describes path length as a function of transmission, which cannot be determined analytically. Since $\tilde{I}_n(\boldsymbol{r}_{s_1}, \hat{\boldsymbol{\Omega}})$ is independent of T, this can be rewritten as two integrals,

$$I_n(\boldsymbol{r}, \hat{\boldsymbol{\Omega}}) = \int_{T(s_1)}^{1} J_n(\boldsymbol{r}_{s(T')}, \hat{\boldsymbol{\Omega}}') dT' + \int_{0}^{T(s_1)} \tilde{I}_n(\boldsymbol{r}_{s_1}, \hat{\boldsymbol{\Omega}}') dT'\,, \tag{2.31}$$

which can then be combined into a single integral,

$$I_n(\boldsymbol{r}, \hat{\boldsymbol{\Omega}}) = \int_{0}^{1} J_n^*(\boldsymbol{r}, \hat{\boldsymbol{\Omega}}, T') dT'\,, \tag{2.32}$$

using the piecewise source function,

$$J_n^*(\boldsymbol{r}, \hat{\boldsymbol{\Omega}}, T) = \begin{cases} \tilde{I}_n(\boldsymbol{r}_{s_1}, \hat{\boldsymbol{\Omega}}) & T \leq T(s_1) \\ J_n(\boldsymbol{r}_{s(T)}, \hat{\boldsymbol{\Omega}}) & T > T(s_1) \end{cases}. \tag{2.33}$$

It is also helpful to split the BRDF $B(\boldsymbol{r}, \hat{\boldsymbol{\Omega}}, \hat{\boldsymbol{\Omega}}')$ found in Equation (2.29c) into two parts:

$a_0(\boldsymbol{r}, \hat{\boldsymbol{\Omega}})$, the albedo of the BRDF for outgoing direction $\hat{\boldsymbol{\Omega}}$,

$$a_0(\boldsymbol{r}, \hat{\boldsymbol{\Omega}}) = \int_{2\pi} \mu(\Omega') B(\Omega, \Omega') d\Omega' \,, \tag{2.34}$$

and $b(\boldsymbol{r}, \hat{\boldsymbol{\Omega}}, \hat{\boldsymbol{\Omega}}')$, a probability density function (PDF) which is normalized over the solid angle hemisphere:

$$b(\boldsymbol{r}, \hat{\boldsymbol{\Omega}}, \hat{\boldsymbol{\Omega}}') = \frac{\mu(\hat{\boldsymbol{\Omega}}') B(\boldsymbol{r}, \hat{\boldsymbol{\Omega}}, \hat{\boldsymbol{\Omega}}')}{a_0(\boldsymbol{r}, \hat{\boldsymbol{\Omega}})} \,. \tag{2.35}$$

No such rearrangement is necessary for the phase function $p(\mathbf{r}, \Theta)$, which is already normalized over the solid angle sphere. With all of these changes, Equation (2.29) takes the following form:

$$I_n(\boldsymbol{r}, \hat{\boldsymbol{\Omega}}) = \int_0^1 J_n^*(\boldsymbol{r}, \hat{\boldsymbol{\Omega}}, T) dT \tag{2.36a}$$

$$J_n(\boldsymbol{r}, \hat{\boldsymbol{\Omega}}) = \omega_0(\boldsymbol{r}) \int_{4\pi} I_{n-1}(\boldsymbol{r}, \hat{\boldsymbol{\Omega}}') \bar{p}(\boldsymbol{r}, \hat{\boldsymbol{\Omega}}, \hat{\boldsymbol{\Omega}}') d\Omega' \tag{2.36b}$$

$$\tilde{I}_n(\boldsymbol{r}, \hat{\boldsymbol{\Omega}}) = a_0(\boldsymbol{r}, \hat{\boldsymbol{\Omega}}) \int_{2\pi} I_{n-1}(\boldsymbol{r}, \hat{\boldsymbol{\Omega}}') b(\boldsymbol{r}, \hat{\boldsymbol{\Omega}}, \hat{\boldsymbol{\Omega}}') d\Omega' \tag{2.36c}$$

$$J_n^*(\boldsymbol{r}, \hat{\boldsymbol{\Omega}}, T) = \begin{cases} \tilde{I}_n(\boldsymbol{r}_{s_1}, \hat{\boldsymbol{\Omega}}) & T \leq T(s_1) \\ J_n(\boldsymbol{r}_{s(T)}, \hat{\boldsymbol{\Omega}}) & T > T(s_1) \end{cases} . \tag{2.36d}$$

These integrals can be evaluated using Monte Carlo integration:

$$I_n(\boldsymbol{r}, \hat{\boldsymbol{\Omega}}) = \left\langle J_n^*(\boldsymbol{r}, \hat{\boldsymbol{\Omega}}, T) \right\rangle_{T \sim \text{uni}(0,1)} \tag{2.37a}$$

$$J_n(\boldsymbol{r}, \hat{\boldsymbol{\Omega}}) = 4\pi \omega_0(\boldsymbol{r}) \left\langle I_{n-1}(\boldsymbol{r}, \hat{\boldsymbol{\Omega}}') \right\rangle_{\hat{\boldsymbol{\Omega}}' \sim p(\mathbf{r}, \hat{\boldsymbol{\Omega}}, \hat{\boldsymbol{\Omega}}')} \tag{2.37b}$$

$$\tilde{I}_n(\boldsymbol{r}, \hat{\boldsymbol{\Omega}}) = 2\pi a_0(\boldsymbol{r}, \hat{\boldsymbol{\Omega}}) \left\langle I_{n-1}(\boldsymbol{r}, \hat{\boldsymbol{\Omega}}') \right\rangle_{\hat{\boldsymbol{\Omega}}' \sim b(\mathbf{r}, \hat{\boldsymbol{\Omega}}, \hat{\boldsymbol{\Omega}}')} . \tag{2.37c}$$

Note that the incident radiation $I_0(\boldsymbol{r}, \hat{\boldsymbol{\Omega}})$ requires no Monte Carlo evaluation, as it can be determined without an integral in the same way as the successive orders method, as done in Equation (2.27). In order to fill the sample spaces to estimate the expectation values in Equation (2.37), rays are traced backwards out of the instrument. This is physically valid, as scattering and reflection events are equally probable when the directions are reversed, and it

ensures that only rays that terminate at the instrument are traced. To begin, a scatter point or a ground reflectance point is selected along the line of sight by generating T according to Equation (2.37a). Solar irradiance is scattered through this point into the observer, with appropriate reductions for absorption and scattering, and the resulting radiance is taken as a sample of $I_1(\mathbf{r}, \hat{\mathbf{\Omega}})$. From the scatter point, a new propagation direction is selected by sampling $\hat{\mathbf{\Omega}}'$ according to Equation (2.37b) if the point is in the air, or Equation (2.37c) if the point is on the ground. Choosing a propagation direction from the ground can be done following the process for sampling a two-dimensional PDF described in Appendix B. A new scatter or reflection point is selected from this new line of sight in the same manner, the solar irradiance is scattered through both points to the observer, and the resulting radiance is taken as a sample of $I_2(\mathbf{r}, \hat{\mathbf{\Omega}})$. This process is continued up until a maximum order of scatter has been reached, or until the radiance contribution becomes negligible. This process is illustrated in Figure 2.6. By reusing the ray history from a sample of I_n to sample I_{n+1}, the computation time is reduced by a factor of n. A consequence of this is that the samples are no longer independent, but this can be accounted for when the variance of the total radiance is estimated. This entire process is repeated, creating new sets of samples for each scatter order which are independent from the previous sets, until a suitable sample space has been created for $I_n(\mathbf{r}, \hat{\mathbf{\Omega}})$ for each order n.

2.3.9 Discrete Ordinates Method

In the discrete ordinates method, an approximate analytic solution is formed for the equation of radiative transfer in a plane-parallel atmosphere. This method is typically much faster than the successive orders or Monte Carlo methods, but suffers some accuracy loss due to the plane-parallel approximation. These accuracy losses are reduced in modern implementations, but these improved algorithms are not available in SASKTRAN and are outside the scope of this thesis. The method described here is based on the linearized discrete ordinates radiative transfer (LIDORT) package (Spurr, Kurosu, & Chance, 2001).

The plane-parallel approximation reduces the equation from five dimensions to three, as horizontal homogeneity is assumed in the atmospheric properties and therefore in the radiance

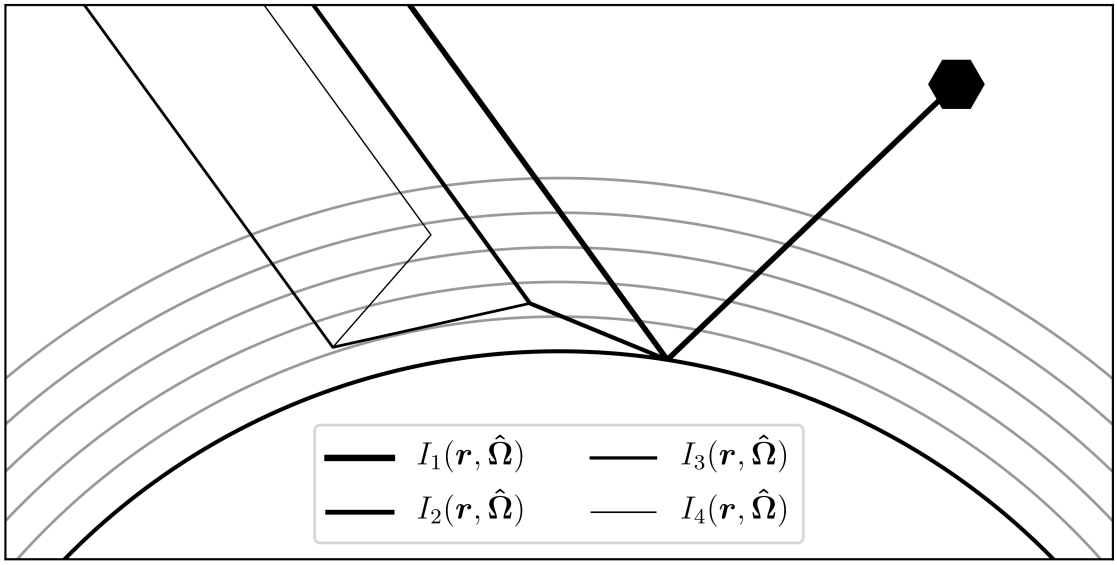


Figure 2.6: Monte Carlo ray tracing visualization, showing the shared ray history of one set of radiance samples.

field, leaving one vertical coordinate and two directional coordinates. The vertical coordinate is chosen to be optical depth τ such that the top of the atmosphere is $\tau = 0$. Taking μ and ϕ, the zenith cosine and azimuth angle, as the directional coordinates, the equation of radiative transfer takes the form,

$$\mu \frac{dI(\tau, \mu, \phi)}{d\tau} = I(\tau, \mu, \phi) - J(\tau, \mu, \phi)\,, \tag{2.38}$$

where the source term $J(\tau, \mu, \phi)$ is the sum of scattered diffuse light,

$$J_{\text{diffuse}}(\tau, \mu, \phi) = \omega(\tau) \int_0^{2\pi} \int_0^1 p(\tau, \mu, \phi; \mu', \phi') I(\tau, \mu', \phi') d\mu' d\phi'\,, \tag{2.39}$$

and scattered sunlight,

$$J_{\text{beam}}(\tau, \mu, \phi) = F_\odot \omega(\tau) p(\tau, \mu, \phi; -\mu_0, \phi_0) e^{-\tau/\mu_0}\,, \tag{2.40}$$

where μ_0 and ϕ_0 represent the direction of the sun.

The azimuth dependence is removed by expanding the radiance and the BRDF in a cosine

Fourier series,

$$I(\tau,\mu,\phi) = \sum_{m=0}^{2N-1} I^m(\tau,\mu)\cos m(\phi_0-\phi) \tag{2.41}$$

$$B(\mu,\phi,\mu',\phi') = \sum_{m=0}^{2N-1} B_m(\tau,\mu)\cos m(\phi-\phi')\,, \tag{2.42}$$

and expanding the phase function in a Legendre polynomial series,

$$p(\tau,\mu,\phi;\mu',\phi') = p(\tau,\cos\Theta) = \sum_{l=0}^{2N-1} \beta_l(\tau)P_l(\cos\Theta)\,, \tag{2.43}$$

where the scattering angle Θ is defined by

$$\cos\Theta = \mu\mu' + \sqrt{(1-\mu^2)(1-\mu'^2)}\cos(\phi-\phi')\,. \tag{2.44}$$

This generates $2N$ equations,

$$\mu\frac{dI^m(\tau,\mu)}{d\tau} = I^m(\tau,\mu) - \int_{-1}^{1} D^m(\tau,\mu,\mu')I^m(\tau,\mu')d\mu' - \frac{F_\odot}{2\pi}e^{-\tau/\mu_0}(2-\delta_{m0})D^m(\tau,\mu,\mu_0)\,, \tag{2.45}$$

one for each Fourier component, where the auxiliary quantities $D^m(\tau,\mu,\mu')$ are defined by

$$D^m(\tau,\mu,\mu') \equiv 2\pi\omega_0(\tau)\sum_{l=m}^{2N-1}\beta_l(\tau)P_l^m(\mu)P_l^m(\mu')\,. \tag{2.46}$$

The addition theorem for Legendre polynomials was used to express the Legendre polynomials $P_l(\cos\Theta)$ in terms of the associated Legendre polynomials $P_l^m(\mu)$, which have been normalized in such a way that the sum in Equation (2.46) has no coefficients. The integral in Equation (2.45) is evaluated by a quadrature sum with $2N$ Gauss-Legendre quadrature abscissae μ_i and weights a_i, where i and j take the values $\pm1,\ldots,\pm N$, and where the positive and negative indices correspond to upward and downward radiance respectively. This creates a system of $2N$ linear equations to be solved for each of the Fourier components,

$$\mu_i\frac{dI_i^m(\tau)}{d\tau} = I_i^m(\tau) - \sum_j a_j D^m(\tau,\mu_i,\mu_j)I_j^m(\tau) - \frac{F_\odot}{2\pi}e^{-\frac{\tau}{\mu_0}}(2-\delta_{m0})D^m(\tau,\mu_i,\mu_0)\,, \tag{2.47}$$

where $I_i^m(\tau) = I^m(\tau, \mu_i)$.

The problem can be solved by splitting up the atmosphere into K homogeneous layers, finding the general solutions of the $2N$ systems of $2N$ equations within each layer, and applying appropriate boundary conditions. The boundary conditions are shown in Equation (2.48), where the layer with subscript p is bound by optical depths τ_{p-1} and τ_p, and where $\tau_0 = 0$ is the top of the atmosphere.

$$I_{i,1}^m(\tau_0) = 0 \qquad i < 0 \tag{2.48a}$$

$$I_{i,p}^m(\tau_p) = I_{i,p+1}^m(\tau_p) \qquad p = 1, \ldots, K-1 \tag{2.48b}$$

$$I_{i,K}^m(\tau_K) = \pi(1+\delta_{m0}) \sum_{j=1}^{N} a_j \mu_j B_m(\mu_i, -\mu_j) I_{-j,K}^m(\tau_K) + F_\odot \mu_0 e^{-\frac{\tau_K}{\mu_0}} B_m(\mu_i, -\mu_0) \qquad i > 0 \tag{2.48c}$$

The three boundary conditions state that the downward diffuse radiance is 0 at the top of the atmosphere, that the radiance is continuous at each layer boundary, and that the upward diffuse radiance at the surface consists of the reflection of the downward diffuse radiance and the reflection of the direct solar beam, both according to the BRDF.

2.3.10 Weighting Functions

Continuous Weighting Functions

It is often helpful to use radiative transfer models to calculate not only the radiance, but also derivatives of the radiance, which are known as weighting functions. The weighting function $w_p(z)$ is defined here as the functional derivative of the radiance $I[p(z)]$ with respect to $p(z)$,

$$w_p(z) = \frac{\partial I}{\partial p}(z)\,, \tag{2.49}$$

where $p(z)$ could be any altitude-dependent atmospheric property that affects the calculation of $I[p(z)]$. Note that this formulation assumes that the property $p(z)$ is horizontally homogeneous. Functional derivatives are formally defined by the integral of their product with an

arbitrary function, denoted $\phi(z)$ here:

$$\int \frac{\partial I}{\partial p}(z)\phi(z)dz \equiv \lim_{\epsilon \to 0} \frac{I[p(z) + \epsilon\phi(z)] - I[p(z)]}{\epsilon}\,, \tag{2.50}$$

or equivalently:

$$\int \frac{\partial I}{\partial p}(z)\phi(z)dz = \left.\frac{d}{d\epsilon} I[p(z) + \epsilon\phi(z)]\right|_{\epsilon=0}\,. \tag{2.51}$$

The weighting function can be isolated by choosing the delta function $\delta(z'-z)$ as the arbitrary function:

$$w_p(z) = \frac{\partial I}{\partial p}(z) = \left.\frac{d}{d\epsilon} I[p(z') + \epsilon\delta(z - z')]\right|_{\epsilon=0}\,. \tag{2.52}$$

The term weighting function can also refer to functional derivatives of the logarithm of the radiance, rather than the radiance itself. Weighting functions of this type will be denoted $\omega_p(z)$, and differ from the previous weighting functions by a factor of $I[p(z)]$:

$$\omega_p(z) = \left.\frac{d}{d\epsilon} \ln I[p(z) + \epsilon\delta(z - z')]\right|_{\epsilon=0} = \frac{1}{I[p(z)]}\left.\frac{d}{d\epsilon} I[p(z) + \epsilon\delta(z - z')]\right|_{\epsilon=0} = \frac{w_p(z)}{I[p(z)]}\,. \tag{2.53}$$

Discrete Weighting Functions

Due to the complexity of the radiative transfer calculation, it is not feasible to calculate analytical weighting functions from an analytical profile properties; the properties are typically specified in a discretized form, and the resulting weighting functions are also discretized. If the property $p(z)$ is represented by the vector $\boldsymbol{p}$ containing discrete points p^i, then the radiance can be written as $I(\boldsymbol{p}) \equiv I[p(z)]$, and the Jacobian of $I(\boldsymbol{p})$ becomes a sensible way to define a set of discrete weighting functions:

$$w_p^i \equiv \frac{\partial I(\boldsymbol{p})}{\partial p^i}\,. \tag{2.54}$$

Consider the case where the property $p(z)$ is represented by constant values p^i within homogeneous layers defined by the altitude boundaries z_{i-1} and z_i. The derivative of the radiance with respect to p^i now represents the response of the radiance to a constant perturbation to

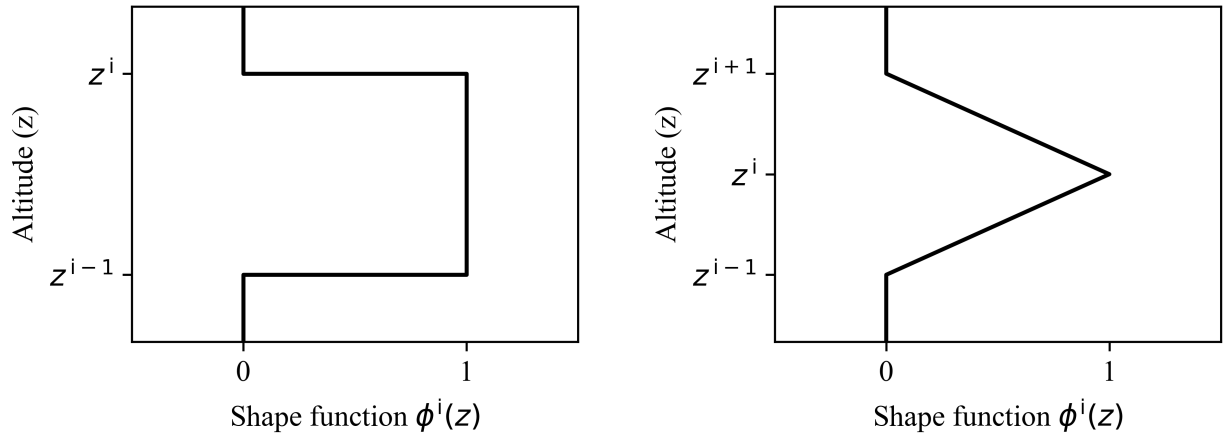
$p(z)$ within layer i, which can be written as

$$w_p^i = \frac{\partial I(\boldsymbol{p})}{\partial p^i} = \lim_{\Delta p \to 0} \frac{I[p(z) + \Delta p \phi^i(z)] - I[p(z)]}{\Delta p}, \tag{2.55}$$

where $\phi^i(z)$ is equal to one within layer i and zero elsewhere, as depicted in Figure 2.7a. Combining Equation (2.55) with the functional derivative definition in Equation (2.50) gives the following expression relating the discrete and continuous weighting functions:

$$w_p^i = \int w_p(z) \phi^i(z) dz \tag{2.56}$$

Similarly, if the property $p(z)$ is represented by linear interpolation between points p^i defined at altitudes z^i, the derivative of the radiance with respect to p^i represents the response of the radiance to a triangular perturbation to $p(z)$. This can also be expressed using Equations (2.55) and (2.56), but with $\phi^i(z)$ defined as the triangular function depicted in Figure 2.7b.



(a) Homogeneous layers. **(b)** Linear interpolation.

Figure 2.7: Shape functions for discretized weighting functions.

These definitions of the discretized weighting function are useful for retrievals based on non-linear minimization, in which the values p^i are being adjusted to find the best match between the modelled radiance $I(\boldsymbol{p})$ and the measured radiance. However, it is important to note that these derivatives are not true samples of the continuous weighting function, due to the shape function $\phi^i(z)$ being unnormalized. The relationship between discrete weighting functions

and samples of the continuous weighting function is given by

$$w_p^i = \int w_p(z)\phi^i(z)dz \approx w_p(z^{i*}) \int \phi^i(z)dz \equiv w_p(z^{i*})\Delta z^i \tag{2.57}$$

where Δz^i is the integral of $\phi^i(z)$ and z^{i*} is an altitude representing $\phi^i(z)$.

2.4 Differential Optical Absorption Spectroscopy

Differential optical absorption spectroscopy (DOAS) is a technique used to extract trace gas concentrations from measurements of light. The technique is defined by the separation of broadband spectral features from the narrow features which are of interest. DOAS takes on many different forms: it can be active or passive, done in the lab or in the atmosphere, done with direct light or scattered light, and ground-based or space-based. All methods utilize something similar to the spectral fitting described in Section 2.4.2, but only the more complex arrangements that measure scattered light require the air mass factor (AMF) correction described in Section 2.4.4. The focus of this document is on DOAS with space-based measurements of scattered light, but much of the theory is applicable to other arrangements.

2.4.1 Slant Column Densities

The slant column density (SCD) is a quantity that is fundamental to the DOAS method, as the first step is always to extract the SCD from the measured spectra. Before the SCD is defined, first consider the definition of the closely related vertical column density (VCD), denoted V:

$$V = \int_0^H n(z)dz\,. \tag{2.58}$$

The VCD is the integral of the number density $n(z)$ along a vertical path, representing the amount of a given species present in a vertical column per unit horizontal area. Note that the VCD is a function only of the location and the atmospheric state; it is completely independent of any measurement being made, in contrast to the SCD which depends on both the atmospheric state and the measurement parameters. The SCD is the integral of the number density along what is called the slant path, which is the weighted average of all light

paths that contributed to a given measurement. Note that in the ultraviolet-visible spectrum these light paths can be varied and complex: most paths consist of a single scattering or reflection event, but a non-negligible portion of the signal will consist of paths with multiple scattering or reflection events, as shown in Figure 2.8. One helpful way to visualize the SCD is to consider individual photons: if you could trace the path of each individual measured photon and integrate the number density along these paths, the average of all these integrals would be equal to the SCD. If L represents the slant path with differential segments $dl(\lambda)$, then the SCD $S(\lambda)$ can be represented by the integral

$$S(\lambda) = \int_L n(l) dl(\lambda) \,. \tag{2.59}$$

Note that this integral is not a strict line integral, as it captures both the integration and the averaging described above, but it is helpful notation nonetheless. The SCD is wavelength dependent, as it is a function of the slant path which is shaped by many scattering, absorption and reflection events, all of which are wavelength dependent.

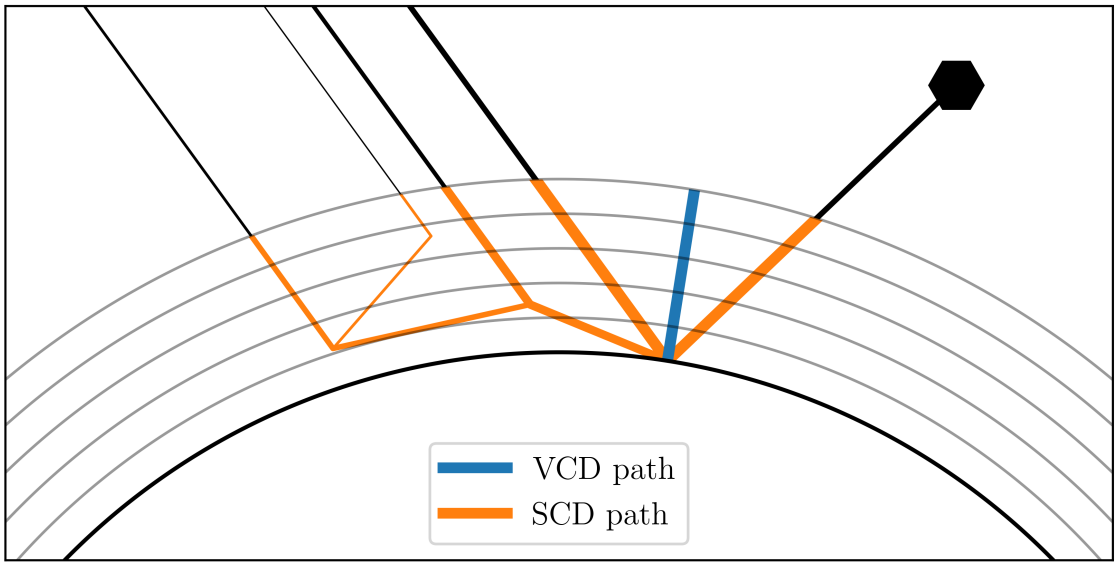


Figure 2.8: Illustration of the integration paths for the SCD and the VCD. Note that the true integration path of the SCD is an average over many more paths than are shown here.

2.4.2 Spectral Fitting

The SCD is useful because can be readily computed from radiance measurements with little computational effort. As shown in Figure 2.9, a trace gas leaves a unique spectral absorption signature that is visible in the normalized radiance spectrum. For each trace gas of interest, a wavelength window is chosen in which the absorption is strong enough to show above measurement noise and in which the spectral features vary rapidly. Because a single SCD will be derived from the spectrum within this window, the wavelength dependence of the SCD must be ignored. The window must be small enough that the SCD wavelength dependence is negligible, but large enough to acquire sufficient signal for the following analysis. Because the absorption features vary much more rapidly than other wavelength dependent processes, such as Rayleigh scattering which is responsible for the gentle downward slope of the radiance in Figure 2.9, they can be easily distinguished from each other. Note that this step is completely decoupled from any radiative transfer analysis: all of the information regarding the radiative transfer is contained in the AMF correction that is applied later in the algorithm.

To calculate the SCD, a simplified radiance model is fit to the measurement. This model accurately describes absorption processes using Beer's law, but represents the slow changing features empirically with a low order polynomial:

$$I(\lambda) = \left(\sum_i p_i \lambda^i \right) e^{-\tau(\lambda)} \,. \tag{2.60}$$

The optical depth $\tau(\lambda)$ in this model accounts for all of the absorbing species seen by the light entering the instrument; in other words it is the optical depth along the slant path. This optical depth is the product of the SCD S and the cross section $\bar{\sigma}(\lambda)$,

$$\tau(\lambda) = \int_L n(l)\sigma(\lambda, l)dl = S\bar{\sigma}(\lambda) \,, \tag{2.61}$$

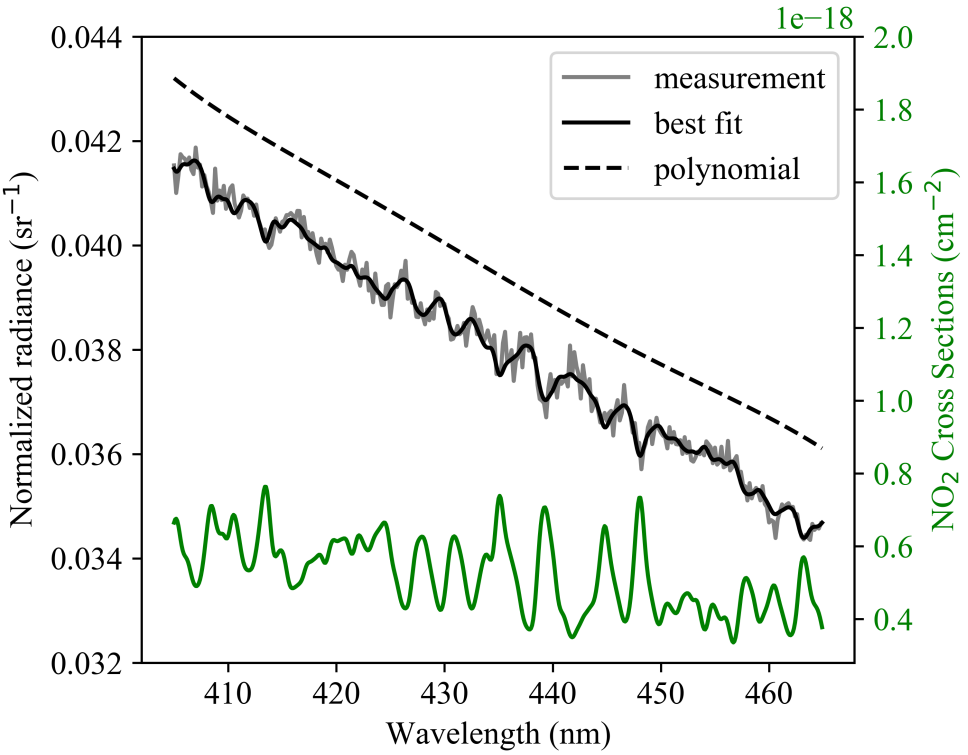


Figure 2.9: Spectral fitting example, showing a noisy measurement, the best fit using Equation (2.63), as well as the low order polynomial term from Equation (2.63). The NO_2 cross sections, which are responsible for most of the features in this wavelength window, are included for reference.

where $\bar{\sigma}(\lambda)$ is the average cross section weighted by the number density along the slant path:

$$\bar{\sigma}(\lambda) = \frac{1}{S}\int_L n(l)\sigma(\lambda, l)dl = \frac{\int_L n(l)\sigma(\lambda, l)dl}{\int_L n(l)dl}\,. \tag{2.62}$$

The position dependence of the cross section is due to its dependence on temperature and pressure, and is generally quite weak. In practice this average cross section is approximated since the exact slant path is not known. Combining Equations (2.60) through (2.62), and generalizing the equations to include multiple absorbing species, results in the following radiance model, often referred to as the DOAS equation:

$$I(\lambda) = \left(\sum_{i=0}^{N} p_i \lambda^i\right) \exp\left(-\sum_{j=1}^{M} S_j \sigma_j(\lambda)\right)\,. \tag{2.63}$$

For any given spectral window, there are typically multiple absorbing species that have a significant effect on the radiance. All of these must be a part of the spectral fit, even when the goal is to retrieve the SCD for only one species. For example, the Version 2.1 Ozone Monitoring Instrument (OMI) NO_2 Standard Product (SP2) algorithm fits NO_2, water vapour, and ozone cross sections in the 405-465 nm window in order to retrieve NO_2 SCDs (Bucsela et al., 2013).

Optimal coefficients p_i and S_j can be found using a non-linear least squares optimization algorithm such as the Levenberg-Marquardt algorithm, minimizing the following cost function:

$$\Phi(\boldsymbol{\beta}) = \sum_{k=1}^{N} \left(\frac{y_k - I(\lambda_k, \boldsymbol{\beta})}{\sigma_k} \right)^2 , \tag{2.64}$$

where $\boldsymbol{\beta}$ is a vector containing coefficients p_i and SCDs S_j, and where y_k are the sun-normalized radiance measurements taken at wavelengths λ_k with uncertainties σ_k, which are not to be confused with the absorption cross sections $\sigma_j(\lambda)$. A covariance matrix corresponding to $\boldsymbol{\beta}$ can also be found during the leasts squares optimization (see Appendix C), and the square roots of its diagonal entries corresponding to the SCDs can be taken as the uncertainty estimates of the SCDs.

A common alternative to the non-linear DOAS equation in Equation (2.63) is the linear variant, which fits to the logarithm of the radiance rather than the radiance itself:

$$\ln I(\lambda) = \sum_{i=0}^{N} p_i \lambda^i - \sum_{j=1}^{M} S_j \sigma_j(\lambda) \, . \tag{2.65}$$

The spectral fitting is now a linear least squares problem, which is much more straightforward than its non-linear counterpart. However, the non-linear DOAS is required under certain conditions, such as when the Ring Effect technique described in Section 2.4.3 is applied (Bucsela et al., 2006).

2.4.3 Ring Effect

The Ring Effect refers to the filling in of the solar Fraunhofer lines when sunlight is observed through the atmosphere, an effect which was first recorded in the 1960s by by Grainger and Ring (1962). The primary cause was not immediately established, but it has since been shown that the Ring effect is primarily caused by inelastic molecular scattering, including rotational Raman scattering. The Ring Effect is important for NO_2 retrievals, as the Ring signal is an order of magnitude larger than the NO_2 absorption signal (Marchenko et al., 2015).

One consequence of the Ring Effect is that the sun-normalized radiance is no longer completely free of solar spectrum information, as solar features leak out to nearby wavelengths and are therefore not completely eliminated when the radiance is divided by the solar irradiance. This extra information can be accounted for without significantly changing the DOAS method by modelling the Ring Effect as a pseudo-absorber, as has been done in the OMI NO_2 algorithms (Bucsela et al., 2006; Bucsela et al., 2013). An effective cross section $\sigma_R(\lambda)$, also known as the differential Ring spectrum, is calculated and its amplitude c_R is included as another parameter to fit in the non-linear least squares minimization problem:

$$I(\lambda) = \left(\sum_{i=0}^{N} p_i \lambda^i\right) \exp\left(-\sum_{j=1}^{M} S_j \sigma_j(\lambda)\right) (1 + c_R \sigma_R(\lambda)) \tag{2.66}$$

The calculation of the effective cross section $\sigma_R(\lambda)$ is described by Chance and Spurr (1997). The rotational Raman cross sections, representing the probability of transitions between quantum rotational states, are computed for diatomic nitrogen and diatomic oxygen. An incident solar irradiance spectrum is prepared which matches the resolution of the measuring instrument. The Raman cross sections are convolved with this solar spectrum to produce the Ring source spectrum. This spectrum is divided by the same solar spectrum, and then a low order polynomial is fit and subtracted to produce the differential Ring spectrum. The differential Ring spectrum contains the necessary information about the rotational Raman transitions as well as the incident solar spectrum.

2.4.4 Air Mass Factors

The SCD is not a pure measurement of the state of the atmosphere, as it depends on external factors such as the position of the observer, the position of the sun, and the ground reflectance. In order to decouple the measurement from these factors, the SCD must be converted into the VCD, which is a function of atmosphere alone. This is the purpose of the AMF, a quantity which is defined as the ratio between the SCD and the VCD:

$$A(\lambda) = \frac{S(\lambda)}{V}\,. \tag{2.67}$$

The AMF inherits the same weak wavelength dependence as the SCD, which is usually ignored by calculating the AMF at a single wavelength near the center of the wavelength window used for the spectral fit. The wavelength dependence will be dropped for the remainder of this section. The overall strategy for DOAS in a scattering atmosphere is to find the SCD through spectral fitting, calculate the AMF using the process described below, and divide them to calculate the VCD. The calculation of the AMF is complicated, requiring complex radiative transfer calculations and underinformed assumptions about the state of the atmosphere, so the AMF is often the largest source of error in DOAS retrievals.

AMFs are calculated from quantities known as box-AMFs or scattering weights. The box-AMF a^i is defined as the ratio of the portions of the SCD and the VCD that lie within a single layer:

$$a^i = \frac{s^i}{v^i} = \frac{\int_{L^i} n(l)dl}{\int_{z^{i-1}}^{z^i} n(z)dz}\,, \tag{2.68}$$

where L^i is the portion of the slant path L that lies within the discrete layer i between altitudes z^{i-1} and z^i, as illustrated in Figure 2.10.

By definition the total SCD S and the total VCD V are the sum of the partial SCDs s^i and

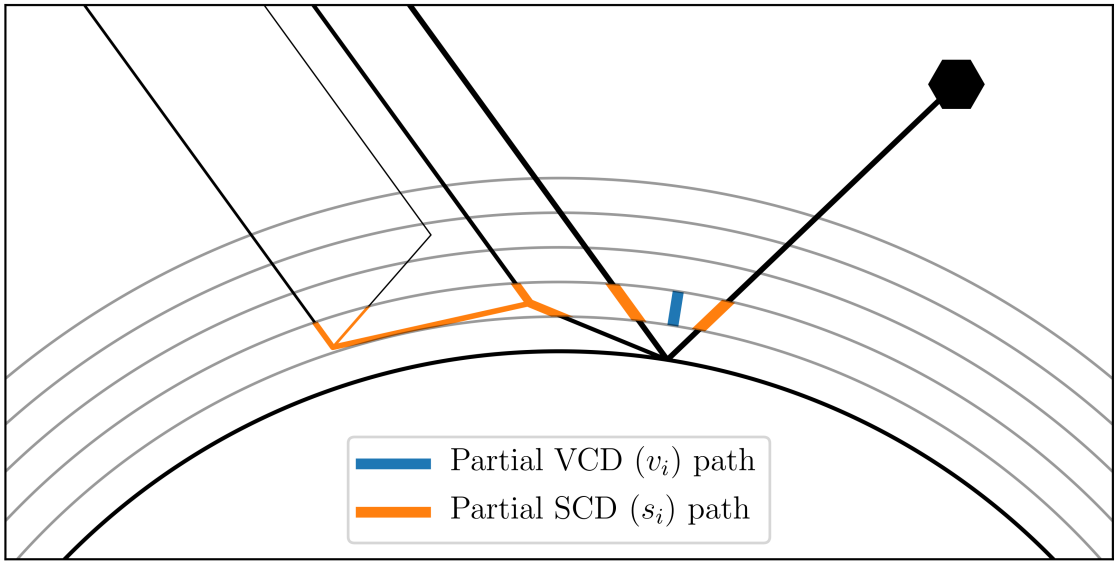


Figure 2.10: Illustration of the integration paths for the partial SCD s_i and the partial VCD v_i. Note that the true integration path of the partial SCD is an average over many more paths than are shown here.

partial VCDs v^i over all layers,

$$S = \sum_i s^i \qquad V = \sum_i v^i \,, \tag{2.69}$$

which makes the total AMF A an average of all box-AMFs weighted by the partial VCDs:

$$A = \frac{1}{V} \sum_i v^i a^i \tag{2.70}$$

If the absorber number density is assumed to be approximately constant within each layer, the box-AMFs become mostly independent of the number density profile:

$$a^i = \frac{s^i}{v^i} = \frac{\int_{L^i} n(l) dl}{\int_{z^{i-1}}^{z^i} n(z) dz} \approx \frac{\int_{L^i} n^i dl}{\int_{z^{i-1}}^{z^i} n^i dz} = \frac{\int_{L^i} dl}{\int_{z^{i-1}}^{z^i} dz} \,. \tag{2.71}$$

This is advantageous because the same box-AMFs can be used to process many scenes with different absorber profiles. Typically, an AMF lookup table is used, which contains box-AMF profiles which can be combined with any absorber profile, and which span an appropriate

range of relevant variables such as SZA, VZA, relative azimuth angle, surface albedo, and surface pressure. The lookup table is necessary because calculating the box-AMFs directly requires a radiative transfer calculation which would be too time consuming for most operational algorithms. With this framework, calculating the AMF for a given scene is a three step process: first take the parameters of the scene and linearly interpolate on the AMF table to acquire the box-AMFs a^i, then choose an a priori absorber density profile to acquire the partial VCDs v^i, then finally use Equation (2.70) to calculate the AMF.

Equivalence of Air Mass Factors and Weighting Functions

Box-AMFs can be calculated with most radiative transfer models because they are closely related to weighting functions, and are equivalent under certain circumstances. A qualitative argument can be made for this equivalency. Weighting functions are a measure of the sensitivity of the radiance to the addition of extra absorbers in a layer, and box-AMFs are a measure of the amount of measured light that passed through the layer. Therefore, if more light has passed through a layer, the addition of extra absorbers in that layer should result in a larger change in absorption and therefore a larger change in radiance.

To show this equivalence mathematically, the following argument is presented, based on the work of Rozanov and Rozanov (2010). First, we define the continuous AMF as a limiting case of the box-AMF:

$$A(z) = \lim_{\Delta z^i \to 0} a^i = \lim_{\Delta z^i \to 0} \frac{\int_{L^i} n(l)dl}{\int_{z^{i-1}}^{z^i} n(z)dz} = \frac{n(l)dl}{n(z)dz} = \frac{dl}{dz}\,, \tag{2.72}$$

where $\Delta z^i = z^i - z^{i-1}$ and where interval (z^{i-1}, z^i) contains z even as Δz^i goes to zero. Note that in this limiting case, the number density terms cancel and the continuous AMF $A(z)$ is completely independent of the number density profile $n(z)$.

To go in the opposite direction and calculate a finite box-AMF given the continuous AMF, this result $dl = A(z)dz$ can be used as a change of variables in Equation (2.68) to show that a box-AMF is a continuous weighted average of the continuous AMF within a layer:

$$a^i = \frac{\int_{L^i} n(l)dl}{\int_{z^{i-1}}^{z^i} n(z)dz} = \frac{\int_{z^{i-1}}^{z^i} n(z)A(z)dz}{\int_{z^{i-1}}^{z^i} n(z)dz} \,. \tag{2.73}$$

Consider the radiance $I[k(z)]$ resulting from the extinction profile $k(z)$ at some single wavelength λ which will be dropped from the notation for this argument. The weighting functions $\omega_{k_0}(z)$ can be used to Taylor expand the logarithm of the radiance $\ln I[k(z)]$ around $\ln I[k_0(z)]$,

$$\ln I[k(z)] = \ln I[k_0(z)] + \int_0^H \omega_{k_0}(z)\Delta k(z)dz + O(\Delta k^2) \,, \tag{2.74}$$

where $k_0(z)$ is the base extinction profile and $\Delta k(z)$ is a small perturbation such that $k(z) = k_0(z) + \Delta k(x)$. Note that the weighting functions $\omega_{k_0}(z)$ are derivatives of the logarithm of the radiance, not the radiance itself.

A similar equation can be constructed by applying Beer's law. The integral of the extinction along the slant path computes the optical depth for the entire measurement, so the relationship between the radiance without absorbers $I[k = 0]$ and the true radiance with absorbers $I[k(z)]$ can be written as

$$I[k(z)] = I[k = 0] - \int_L k(l)dl \,. \tag{2.75}$$

Forming the same equation for the base profile $k_0(z)$ and subtracting,

$$\ln I[k(z)] = \ln I[k_0(z)] - \int_L \Delta k(l)dl \,, \tag{2.76}$$

and performing a change of variables using the continuous AMF definition found in Equation (2.72),

$$\ln I(k) = \ln I(k_0) - \int_0^H \Delta k(z) A(z) dz \,, \tag{2.77}$$

we now have an equation of the same form as the Taylor expansion in Equation (2.74). Comparing the two equations, it is evident that if the perturbation $\Delta k(z)$ is made sufficiently small, the AMFs are equal to the weighting functions up to a sign:

$$A(z) = -\omega_{k_0}(z) \,. \tag{2.78}$$

Substituting this result into Equation (2.73), the box-AMF a^i is equal to the discretized weighting function $\omega^i_{k_0}$,

$$a^i = -\frac{\int_{z^{i-1}}^{z^i} n(z)\omega_{k_0}(z)dz}{\int_{z^{i-1}}^{z^i} n(z)dz} = -\int \omega_{k_0}(z)\phi^i(z)dz = -\omega^i_{k_0} \,, \tag{2.79}$$

where the normalized shape function,

$$\phi^i(z) = \begin{cases} n(z)/v^i & z^{i-1} < z < z^i \\ 0 & \text{else} \end{cases} \,, \tag{2.80}$$

is zero outside layer i and has the shape of the number density profile $n(z)$ within. While this precise shape function is required for mathematical equivalence, in practice the discretization is typically insensitive to the chosen shape function, and the use of a simple triangular or rectangular shape function as discussed in Section 2.3.10 will result in negligible change. The sensitivity to the chosen shape function increases as the gradient of the weighting function $\omega_{k_0}(z)$ increases, which may become a problem if the gradient of the number density $n(z)$ is also high, as this would make the true shape function $\phi^i(z)$ highly asymmetrical and a symmetric shape function would sample $\omega_{k_0}(z)$ at the wrong altitudes within the layer.

These conditions can occur near the ground over polluted scenes with low surface albedo, as the low albedo causes the weighting function to rapidly decrease and the pollution causes the number density to rapidly increase. Under these circumstances it may be easier to decrease the width of the layers rather than alter the discretization shape function.

Air Mass Factors and Clouds

The presence of clouds significantly affects the calculation of AMFs. Below the clouds the box-AMFs decrease and above the clouds they increase as the cloud scatters sunlight away from the Earth. To account for clouds, a second AMF lookup table is created which assumes complete cloud cover. Over partially cloudy scenes, an interpolation is performed between the clear-sky table and the cloudy-sky table as first suggested by Martin et al. (2002),

$$a^i = w a^i_{\text{cloudy}} + (1 - w) a^i_{\text{clear}} \,, \tag{2.81}$$

where w is the fraction of the measured radiance that was scattered by clouds, a value that is typically larger than the geometric cloud fraction due to the high reflectivity of clouds. To calculate box-AMFs for cloudy scenes two additional parameters are required: the cloud radiance fraction and the cloud pressure, both of which can typically be estimated by performing a separate retrieval on the same measurement. The algorithms used to derive these parameters are outside the scope of this thesis.

It is important that the calculation of the cloudy-sky AMF table uses the same set of assumptions as the cloud parameter retrievals. For example, the SP2 algorithm assumes that clouds are Lambertian reflectors with an albedo of 0.8, which is the same assumption used by the OMI cloud algorithm (Bucsela et al., 2013). This is particularly important when aerosol is not accounted for explicitly because weakly absorbing aerosols have an effect on the measured radiance similar to that of clouds, so in reality the cloud radiance fraction and the cloud pressure are effective parameters that account for clouds and aerosols together.

Air Mass Factor Uncertainty

Estimating the uncertainty in the AMF is an inexact science which requires many assumptions and approximations, and which has been done with a wide variety of approaches. The method described here is based on the error analysis for the SP2 algorithm described by Bucsela et al. (2013), but it is very similar to the more modern NO_2 data product from the Tropospheric Monitoring Instrument (TROPOMI) described by van Geffen et al. (2019). The concept is straightforward: estimate an uncertainty for each AMF input that the AMF is sensitive to, and propagate it through the AMF equation by taking its derivative with respect to that parameter to find the corresponding uncertainty in the AMF. Combining Equations (2.70) and (2.81) and changing to a vector notation gives the following expression for the AMF:

$$A = \frac{\boldsymbol{v}^{\boldsymbol{T}}(w\boldsymbol{a}_{\text{cloudy}}(p_c) + (1-w)\boldsymbol{a}_{\text{clear}}(a_s))}{V}\,, \tag{2.82}$$

where $\boldsymbol{v}$ is a vector containing the partial VCDs v^i, V is the sum of all partial VCDs v^i, $\boldsymbol{a}_{\text{cloudy}}$ and $\boldsymbol{a}_{\text{clear}}$ are vectors containing box-AMFs a^i_{cloudy} and a^i_{clear}, and where the dependence on cloud pressure p_c and surface albedo a_s is shown explicitly. The four inputs that will be considered for uncertainty analysis are $\boldsymbol{v}$, w, p_c, and a_s, as other inputs have an insignificant effect. An uncertainty must be estimated for each of the four inputs, a single value σ_X for each of the scalar inputs $X = w, p_c, a_s$ representing the standard deviation of the assumed normal distribution, and a covariance matrix Σ_v for the vector input $\boldsymbol{v}$ representing the assumed multivariate normal distribution. These uncertainties are linearly propagated through Equation (2.82) as discussed in Appendix A.2:

$$\sigma_{A_w}^2 = \left(\frac{\partial A}{\partial w}\right)^2 \sigma_w^2 \qquad \frac{\partial A}{\partial w} = \frac{\boldsymbol{v}^T\left(\boldsymbol{a}_{\text{cloudy}}(p_c) - \boldsymbol{a}_{\text{clear}}(a_s)\right)}{V} \tag{2.83a}$$

$$\sigma_{A_{p_c}}^2 = \left(\frac{\partial A}{\partial p_c}\right)^2 \sigma_{p_c}^2 \qquad \frac{\partial A}{\partial p_c} = \frac{w\boldsymbol{v}^T}{V}\frac{\partial \boldsymbol{a}_{\text{cloudy}}}{\partial p_c} \tag{2.83b}$$

$$\sigma_{A_{a_s}}^2 = \left(\frac{\partial A}{\partial a_s}\right)^2 \sigma_{a_s}^2 \qquad \frac{\partial A}{\partial a_s} = \frac{(1-w)\boldsymbol{v}^T}{V}\frac{\partial \boldsymbol{a}_{\text{clear}}}{\partial a_s} \tag{2.83c}$$

$$\sigma_{A_V}^2 = (\boldsymbol{\nabla}_v A)^T \Sigma_v (\boldsymbol{\nabla}_v A) \qquad \boldsymbol{\nabla}_v A = \frac{w\boldsymbol{a}_{\text{cloudy}}(p_c) + (1-w)\boldsymbol{a}_{\text{clear}}(a_s) - A\boldsymbol{u}}{V}\,, \tag{2.83d}$$

where $\boldsymbol{u}$ is a vector the same size as $\boldsymbol{v}$ containing ones. Note that the partial derivatives of the box-AMFs, $\partial \boldsymbol{a}_{\text{cloudy}}/\partial p_c$ and $\partial \boldsymbol{a}_{\text{clear}}/\partial a_s$, must be estimated numerically. To compute the total AMF uncertainty σ_A, the uncertainties are assumed to be independent so that the variances can be summed,

$$\sigma_A^2 = \sigma_{A_w}^2 + \sigma_{A_{p_c}}^2 + \sigma_{A_{a_s}}^2 + \sigma_{A_V}^2\,, \tag{2.84}$$

although this assumption is certainly an approximation, as the retrieval of cloud parameters is sensitive to the assumed surface albedo.

2.4.5 Stratosphere-Troposphere Separation

The troposphere and the stratosphere are the two lowest temperature-defined layers of the Earth's atmosphere. The troposphere is the lowest layer, and is primarily heated by the surface so that temperature decreases as altitude increases. The top of the troposphere is defined by the tropopause, the altitude where this cooling stops and the temperature is relatively constant with altitude, which ranges from about 9 km at the poles to about 17 km at the equator. The stratosphere begins at the tropopause and extends to about 50 km or 55 km - this region is primarily heated by the absorption of ultraviolet sunlight by the ozone layer, which causes the temperature to increase with altitude. The stratosphere is very stable, lacking the turbulence that produces such rapidly changing weather in the troposphere. The

distinction between stratosphere and troposphere is important for remote sensing because the sources and the behaviour of trace gases can change significantly from one region to the other. Ozone and NO_2, for example, occur naturally in the stratosphere and do not typically have any rapidly changing horizontal structure. In the troposphere, however, they are primarily produced by human activity, and the combination of sporadic sources and turbulence results in significant horizontal structure.

If a stratosphere-troposphere distinction is not necessary, the total VCD V can be calculated by dividing the SCD S by the AMF A, and its uncertainty σ_V is calculated by propagating the uncertainties of the SCD σ_S and the AMF σ_A:

$$V = \frac{S}{A} \qquad \sigma_V^2 = \left(\frac{\sigma_S}{A}\right)^2 + \left(\frac{S\sigma_A}{A^2}\right)^2 . \tag{2.85}$$

However, if the distinction is necessary, which is the case for NO_2 where the polluting tropospheric component is of interest due to human health concerns, it is necessary to estimate the stratospheric component of the VCD V_s, as well as to compute separate AMFs for the troposphere A_t and the stratosphere A_s. The tropospheric VCD V_t and its uncertainty σ_{V_t} is then given by

$$V_t = \frac{S - V_s A_s}{A_t} \qquad \sigma_{V_t}^2 = \left(\frac{\sigma_S}{A_t}\right)^2 + \left(\frac{A_s\sigma_{V_s}}{A_t}\right)^2 + \left(\frac{V_s\sigma_{A_s}}{A_t}\right)^2 + \left(\frac{V_t\sigma_{A_t}}{A_t}\right)^2 . \tag{2.86}$$

Computing the tropospheric and stratospheric AMFs is done by picking a tropopause altitude and then applying Equation (2.70) to the layers above and below this altitude separately. The estimation of of the stratospheric column has been done using several different techniques.

In early algorithms such as the NO_2 algorithm for GOME, the stratospheric component was assumed to be constant longitudinally, and was estimated using measurements over the Pacific where the troposphere was assumed to contain minimal NO_2 pollution (Martin et al., 2002). A similar strategy was adopted for the OMI NO_2 Standard Products, but the measurements were taken from low-pollution locations all over the globe, and more complex

processes such as planetary wave analysis (Bucsela et al., 2006) or image processing (Bucsela et al., 2013) replaced the assumption of longitudinal invariance for the purpose of extending the stratospheric field over polluted regions. The low pollution measurements can also be assimilated into a chemical transport model (CTM) to extend the stratospheric field over polluted regions, as has been done for the Dutch OMI NO_2 (DOMINO) product (Boersma, Eskes, & Brinksma, 2004) and the TROPOMI NO_2 product (van Geffen et al., 2019). The most robust method is to use collocated limb measurements to measure the stratospheric column directly, either with a single instrument with nadir and limb modes as has been done with the Scanning Imaging Absorption Spectrometer for Atmospheric Cartography (SCIAMACHY) (Beirle, Kuhl, Pukite, & Wagner, 2010), or with two separate instruments as has been done with with OMI and OSIRIS (Adams et al., 2016). However, this is not always possible as collocated limb measurements are not always available.

3 DOAS Algorithm Development

3.1 Slant Column Densities

Two SCD fitting algorithms have been implemented, following the method outlined in Section 2.4.2. The current implementations are based on OMI algorithms for NO_2 (Bucsela et al., 2013) and total column ozone (Veefkind, de Haan, Brinksma, Kroon, & Levelt, 2006), but they are easily extendible to other similar algorithms. The least squares minimization for both implementations is done using the Levenberg-Marquardt algorithm via the *curve_fit* function from the *optimize* module of the Python package *SciPy*, which is a wrapper around the *lmder* algorithm from the FORTRAN library *MINPACK*, based on the work of Jorge Moré (1978). Analytic Jacobians of Equations (3.1) and (3.2) with respect to the fitting parameters are supplied to *curve_fit* to speed up the algorithm by replacing the numerical Jacobian estimation, since expressions for the derivatives exist and their evaluation does not require significant additional computational effort.

NO_2 Slant Column Densities

The NO_2 algorithm uses the radiance model described by Equation (2.66) in the spectral window from 405 nm to 465 nm with a fifth order polynomial, three species (NO_2, O_3, and H_2O), and an optional Ring spectrum. This results in ten fitted parameters: polynomial coefficients p_0 through p_5, SCDs S_{NO_2}, S_{O_3}, and S_{H_2O}, and Ring amplitude c_R. All cross sections are taken at a temperature of 220 K; this is not accounted for in the SCD stage of the DOAS algorithm, but a temperature correction is applied to the AMF. The model is

given by

$$I_{NO_2}(\lambda) = \left(\sum_{i=0}^{5} p_i \lambda^i\right) \exp\left(-S_{NO_2}\sigma_{NO_2}(\lambda) - S_{O_3}\sigma_{O_3}(\lambda) - S_{H_2O}\sigma_{H_2O}(\lambda)\right)\left(1 + c_R\sigma_R(\lambda)\right) . \tag{3.1}$$

Ozone Slant Column Densities

The ozone algorithm utilizes the spectral window from 331.6 nm to 336.6 nm, and adds an effective temperature to the SCD fitting. The Ring correction described in Veefkind et al. (2006) has not been implemented. The effective temperature is equal to the weighted average of the temperature along a vertical profile with the ozone concentration as the weights, which is typically quite close to the temperature at the peak of the ozone layer in the stratosphere. There are five fitted parameters: polynomial coefficients p_0 through p_2, ozone SCD S_{O_3}, and effective temperature T_{eff}. The model is given by

$$I_{O_3}(\lambda) = \left(\sum_{i=0}^{2} p_i \lambda^i\right) \exp\left(-S_{O_3}\sigma_{O_3}(\lambda, T_{\text{eff}})\right) , \tag{3.2}$$

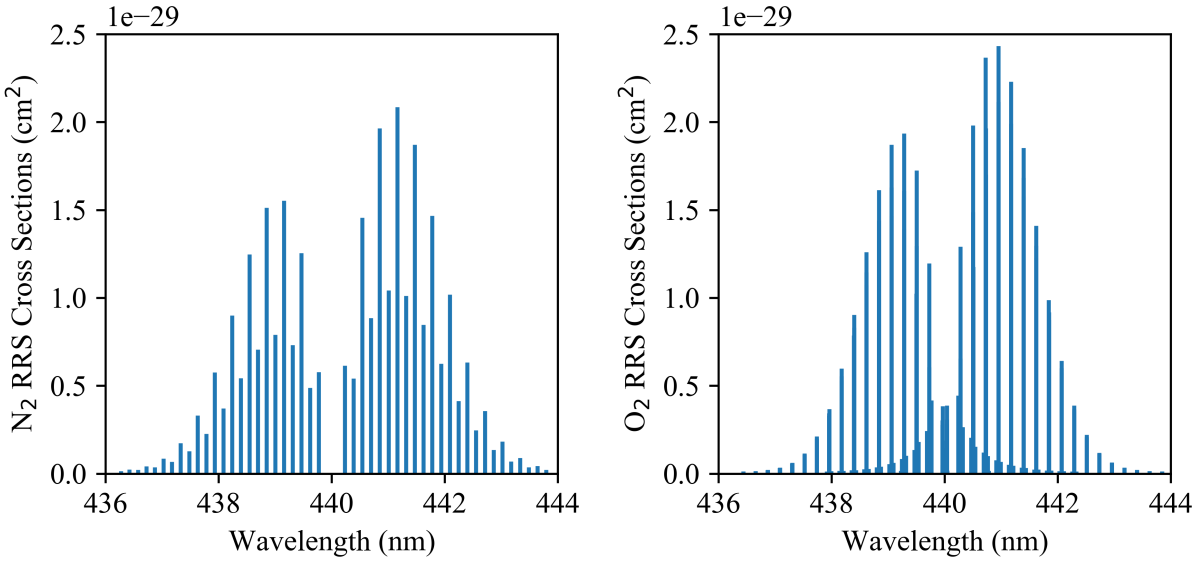
where the temperature dependence of the cross section $\sigma_{O_3}(\lambda, T_{\text{eff}})$ has been linearized to decrease computation time. It is implemented as a linear interpolation or extrapolation between ozone cross sections at two reference temperatures, $T_1 = 228\,\text{K}$ and $T_2 = 243\,\text{K}$:

$$\sigma_{O_3}(\lambda, T_{\text{eff}}) = \left(\frac{T_2 - T_{\text{eff}}}{T_2 - T_1}\right)\sigma_{O_3}(\lambda, T_1) + \left(\frac{T_{\text{eff}} - T_1}{T_2 - T_1}\right)\sigma_{O_3}(\lambda, T_2) . \tag{3.3}$$

3.2 Ring Effect

The implemented Ring effect spectrum follows Chance and Spurr (1997). The first step, described in Section 2.3.4, consists of calculating the rotational Raman cross sections for N_2 and O_2. Tabulated values which accompany *Spectroscopy and Radiative Transfer of Planetary Atmospheres* by Chance and Martin (2017) are used. The tabulated values include the initial state energy, the transition energy, and the Placzek-Teller coefficient for every rotational

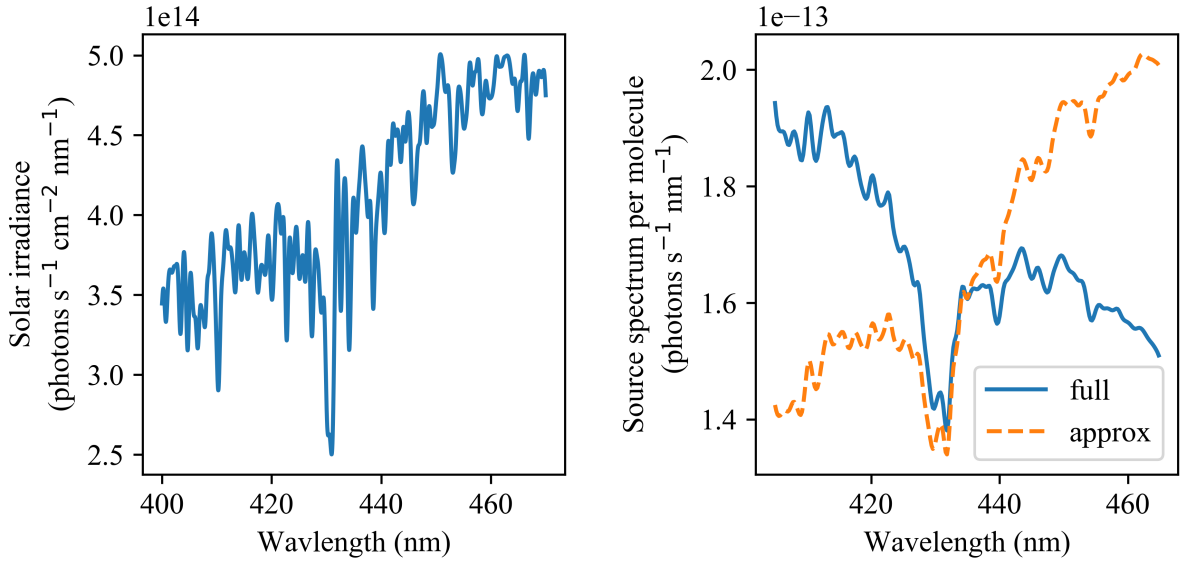
Raman line with an intensity greater than 0.1% of the strongest line, and they include the necessary corrections for O_2. The resulting cross sections are displayed in Figure 3.1. The line broadening described by Chance and Spurr is currently neglected.



(a) N_2 **(b)** O_2

Figure 3.1: Rotational Raman cross sections with an excitation wavelength of 440 nm at a temperature of 250 K. The O_2 lines account for deviation from the Hund's case b basis.

The second step is to use these cross sections to produce the differential Ring spectrum, as described in Section 2.4.3. For the convolution of the cross sections and the solar spectrum, Han et al. calculated the cross sections at a single center wavelength (2011), as opposed to calculating it at many wavelengths across the spectral region of interest. This significantly affects the low-frequency behaviour of the source spectrum, as seen in Figure 3.2. However, when a fitted polynomial is removed to produce the differential Ring spectrum as shown in Figure 3.3, this approximation has a reasonably small effect. Both the simplified convolution and the full convolution are available in the current implementation, but the full convolution is preferred, since the claim by Han et al. that the simplified convolution has negligible effects on the retrieval has not been verified, and rapid calculation of Ring spectra is not currently a priority.



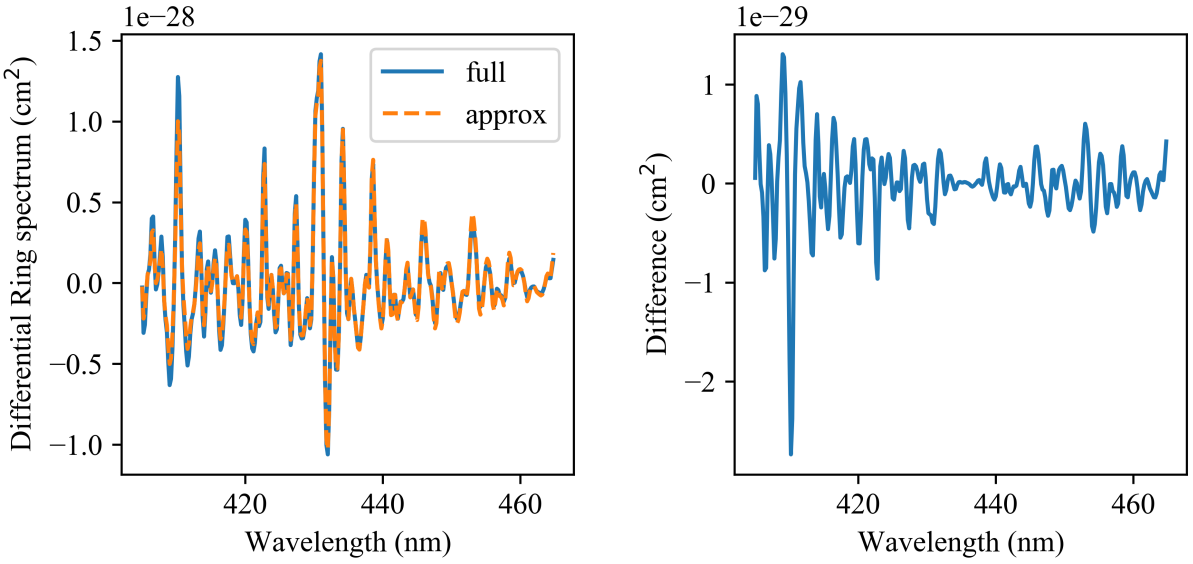
(a) Solar spectrum.

(b) Rotational Raman source spectrum.

Figure 3.2: An example solar spectrum and its corresponding rotational Raman source spectrum, calculated by convolving the solar spectrum and the rotational Raman cross sections for air molecules. The full convolution calculates the cross sections across the window, whereas for the approximation they are calculated at a single wavelength.

3.3 Air Mass Factors

Tools for calculating AMFs were developed in Python, using three separate engines available within SASKTRAN: the high resolution (HR) engine which uses the successive orders method, the Monte Carlo (MC) engine, and the discrete ordinates (DO) engine. SASKTRAN is built in C++, but it is typically accessed via a python interface; the HR and DO AMFs were implemented in python using the standard interface, but the MC AMFs had to be implemented in the core C++ engine. These tools can be used both for direct calculation of AMFs for a specific scene, as well as for creating AMF lookup tables. Currently, lookup tables are implemented using the Python package *xarray* for its built-in multidimensional linear interpolation and its ability to easily export tables as netcdf files.



(a) Differential Ring spectra.

(b) Differential Ring spectra difference.

Figure 3.3: An example solar spectrum and its corresponding rotational Raman source spectrum, calculated by convolving the solar spectrum and the rotational Raman cross sections for air molecules. The full convolution calculates the cross sections across the window, whereas for the approximation they are calculated at a single wavelength.

3.3.1 Box Air Mass Factors Via Weighting Functions

The HR and DO box-AMFs are calculated by first calculating weighting functions, and then transforming them into box-AMFs. Both engines have built-in weighting functions, described by Zawada et al. for the HR engine (2015), and derived for the discrete ordinates method by Spurr et al. (2001). These weighting functions w_n^i are the response of the radiance $I[n(z)]$ to a triangular perturbation in the number density $n(z)$, which can be written in terms of the continuous weighting function, following the notation used in Section 2.3.10:

$$w_n^i = \int \frac{\partial I}{\partial n}(z)\phi^i(z)dz\,, \tag{3.4}$$

where the shape function $\phi^i(z)$ is a symmetric triangle centered on z^i such that $\phi^i(z^i) = 1$ and $\int \phi^i(z)dz = \Delta z^i$. The altitudes z^i and the corresponding widths Δz^i are user-specified. As discussed in Section 2.3.10, this is approximately equal to the value of the continuous

weighting function evaluated at z^i multiplied by the width:

$$w_n^i \approx \Delta z^i \frac{\partial I}{\partial n}(z^i)\,. \tag{3.5}$$

The continuous AMF is equal to the negative of the functional derivative of the logarithm of the radiance with respect to extinction (see Section 2.4.4), which can be expressed as

$$A(z) = -\frac{\partial \ln I}{\partial k}(z) = -\frac{1}{I[n(z)]}\frac{1}{\sigma(z)}\frac{\partial I}{\partial n}(z)\,, \tag{3.6}$$

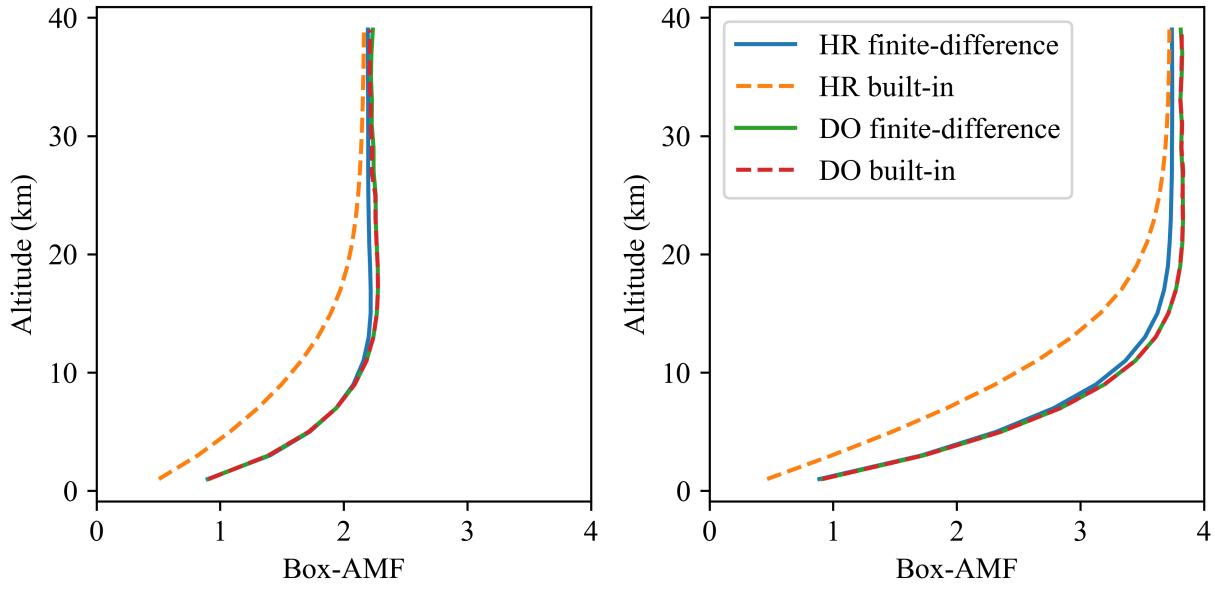
where the factor of $I[n(z)]$ accounts for the logarithm and the cross section $\sigma(z)$ accounts for the change from extinction $k(z)$ to number density $n(z)$. If the layer is sufficiently thin, and if the gradients of the continuous AMF $A(z)$ and the number density $n(z)$ are sufficiently small, then the sample $A(z^i)$ is a good approximation of the box-AMF a^i for the layer containing z^i. Combining this with Equations (3.5) and (3.6) gives the following expression, which is the basis for the HR and DO box-AMFs:

$$a^i \approx -\frac{w_n^i}{I[n(z)]\sigma(z^i)\Delta z^i}\,. \tag{3.7}$$

The weighting functions w_n^i are calculated at the midpoint of each AMF layer, with the weighting function width Δz^i equal to half the width of the layer. These parameters are chosen to keep the triangular perturbation entirely within the layer, which prevents issues in the bottom layer where a wider perturbation would be clipped by the ground.

The built-in weighting functions were tested against a finite-difference scheme, in which a triangular perturbation was added to the number density profile to compute a finite-difference derivative. The built-in HR weighting functions have only been verified for limb viewing geometries (Zawada et al., 2015), and it was found that the they were not sufficiently accurate for computing AMFs with a nadir viewing geometry, therefore finite-difference weighting functions must be used. The built-in weighting functions for the DO engine were found to agree well with their finite-difference counterparts, which means the built-in option is preferred, as it is faster and does not depend on the chosen finite difference parameters.

Example box-AMF profiles for all four methods are shown in Figure 3.4. The figure also illustrates how DO and HR box-AMFs agree for small SZAs and VZAs, but begin to deviate as these values grow due to the plane-parallel assumption made by the DO engine. The DO box-AMFs can be computed much more quickly, motivating their use for first order work and for scenes with small SZAs and VZAs, but the use of the HR engine is necessary for accurate work with large zenith angles.



(a) Mexico City (SZA 26.5°, VZA 22.8°) **(b)** Fort McMurray (SZA 44.7°, VZA 65.3°)

Figure 3.4: Box-AMF comparison between the HR and DO engines, as well as between built-in and finite-difference weighting functions, showing typical examples at the northern and southern extents of TEMPO's field of regard on a summer afternoon.

3.3.2 Monte Carlo Air Mass Factors

The MC engine was originally motivated as a means of verifying the radiances produced by the HR engine, which are sensitive in both accuracy and computation time to certain engine parameters, in particular those related to diffuse profiles, which are vertical profiles of scattering source terms. Specifically, HR calculations are sensitive to the incoming and outgoing angular resolutions chosen for each diffuse point within the diffuse profiles, as well as the density and placement of the profiles (Zawada et al., 2015). The MC engine is independent of these parameters, and can be used to determine optimal settings for HR that achieve

reasonable accuracy and computation time. The motivation for calculating box-AMFs with the MC engine is the same, as it allows HR engine parameters to be chosen with confidence. It also provides verification for the assumptions described in Section 3.3.1 regarding the shape of the weighting function perturbation and the approximate equality of the discrete and continuous AMFs, as the MC method does not rely on these assumptions.

Total Air Mass Factors

The basis of the MC AMF algorithm is the calculation of the SCD alongside the radiance, after which it is a simple matter to divide the SCD by the VCD to get the AMF. The SCD can be calculated alongside the radiance by integrating the number density along each photon path that is traced, and combining these integrals in the appropriate way.

Consider the radiance $I(\boldsymbol{r}, \hat{\boldsymbol{\Omega}})$ in an arbitrary direction $\hat{\boldsymbol{\Omega}}$ at an arbitrary point in the atmosphere $\boldsymbol{r}$. There is a SCD that is associated with this radiance, which will be denoted $S^I(\boldsymbol{r}, \hat{\boldsymbol{\Omega}})$; note that this is the same SCD that is discussed in Section 2.4.1, except that it has now been extended to an arbitrary location in the atmosphere and an arbitrary direction, not just the location and direction where the measurement is being taken. This SCD is equal to the average of the number density integrals along all contributing photon paths. Note that it is impractical to trace individual photon paths and take a strict average, but if all contributing light paths can be parameterized in some way, the SCD can be computed by taking a weighted average, with weights equal to the radiance contribution associated with each path.

For the following work, however, it will be necessary to work instead with the product of the SCD and the corresponding radiance, $W^I(\boldsymbol{r}, \hat{\boldsymbol{\Omega}}) = I(\boldsymbol{r}, \hat{\boldsymbol{\Omega}}) S^I(\boldsymbol{r}, \hat{\boldsymbol{\Omega}})$, which will be referred to as the weighted SCD. The weighted SCD can also be computed by considering all contributing paths, but now they must be combined with a simple sum, rather than a weighted average. The radiance $I(\boldsymbol{r}, \hat{\boldsymbol{\Omega}})$ can be divided into many contribution, each of which can be multiplied by the number density integral along its corresponding light path; the weighted SCD is the sum of this product over all contributions. Order-specific radiances $I_n(\boldsymbol{r}, \hat{\boldsymbol{\Omega}})$ also have a corresponding weighted SCD, $W_n^I(\boldsymbol{r}, \hat{\boldsymbol{\Omega}})$, which shares the definition of $W^I(\boldsymbol{r}, \hat{\boldsymbol{\Omega}})$ but with

the added condition that every contributing light path has exactly n scattering or reflection events. Additionally, the weighted SCD concept can be extended to source terms, defining $W_n^J(\boldsymbol{r}, \hat{\boldsymbol{\Omega}})$ and $W_n^{\tilde{I}}(\boldsymbol{r}, \hat{\boldsymbol{\Omega}})$ corresponding to $J_n(\boldsymbol{r}, \hat{\boldsymbol{\Omega}})$ and $\tilde{I}(\boldsymbol{r}, \hat{\boldsymbol{\Omega}})$ respectively, as these source terms are also made up of many radiance contributions for which the same calculation can be made.

The weighted SCD can be calculated following the same strategy used to calculate the radiance, as described in Section 2.3.8, where radiance terms of one order are used to calculated source terms of the next order, which are subsequently used to calculate radiance terms of that same order. Firstly, the weighted SCD $W_{I_0}(\boldsymbol{r}, \hat{\boldsymbol{\Omega}})$ corresponding to direct sunlight $I_0(\boldsymbol{r}, \hat{\boldsymbol{\Omega}})$ is the product of that radiance with a single line integral:

$$W_0^I(\boldsymbol{r}, \hat{\boldsymbol{\Omega}}) = I_0(\boldsymbol{r}, \hat{\boldsymbol{\Omega}}) S(\boldsymbol{r}, \hat{\boldsymbol{\Omega}}_0, s_2)\,, \tag{3.8}$$

where $S(\boldsymbol{r}, \hat{\boldsymbol{\Omega}}, s)$ represents the following integrated number density:

$$S(\boldsymbol{r}, \hat{\boldsymbol{\Omega}}, s) \equiv \int_s^0 n(\boldsymbol{r} + t\hat{\boldsymbol{\Omega}}) dt\,. \tag{3.9}$$

Secondly, weighted SCDs associated with source terms are calculated by integrating the weighted SCDs associated with all incoming contributions:

$$W_n^J(\boldsymbol{r}, \hat{\boldsymbol{\Omega}}) = \omega_0(\boldsymbol{r}) \int_{4\pi} W_{n-1}^I(\boldsymbol{r}, \hat{\boldsymbol{\Omega}}') \bar{p}(\boldsymbol{r}, \hat{\boldsymbol{\Omega}}, \hat{\boldsymbol{\Omega}}') d\Omega' \tag{3.10}$$

$$W_n^{\tilde{I}}(\boldsymbol{r}, \hat{\boldsymbol{\Omega}}) = a_0(\boldsymbol{r}, \hat{\boldsymbol{\Omega}}) \int_{2\pi} W_{n-1}^I(\boldsymbol{r}, \hat{\boldsymbol{\Omega}}') b(\boldsymbol{r}, \hat{\boldsymbol{\Omega}}, \hat{\boldsymbol{\Omega}}') d\Omega'\,. \tag{3.11}$$

Note that these equations match Equations (2.36b) and (2.36c), where the source terms themselves are calculated by integrating the incoming contributions. The piecewise function can be defined in the same way, matching Equation (2.36d):

$$W_n^{J^*}(\boldsymbol{r}, \hat{\boldsymbol{\Omega}}, T) = \begin{cases} W_n^{\tilde{I}}(\boldsymbol{r} + s(T)\hat{\boldsymbol{\Omega}}, \hat{\boldsymbol{\Omega}}) & T \leq T(s_1) \\ W_n^J(\boldsymbol{r} + s(T)\hat{\boldsymbol{\Omega}}, \hat{\boldsymbol{\Omega}}) & T > T(s_1) \end{cases}. \tag{3.12}$$

Similarly, Equation (2.36a) is mirrored here to calculate the weighted SCD $W_n^I(\boldsymbol{r}, \hat{\boldsymbol{\Omega}})$ corresponding to higher order radiance terms $I_n(\boldsymbol{r}, \hat{\boldsymbol{\Omega}})$, but now one additional term is required. To understand this additional term, first consider Equation (2.36a), which states that the radiance $I_n(\boldsymbol{r}, \hat{\boldsymbol{\Omega}})$ consists of infinitesimal contributions $J_n^*(\boldsymbol{r}, \hat{\boldsymbol{\Omega}}, T)dT$. Each contribution $J_n^*(\boldsymbol{r}, \hat{\boldsymbol{\Omega}}, T)dT$ represents the light that is scattered into the line of sight at the point $\boldsymbol{r}+s(T)\hat{\boldsymbol{\Omega}}$. As illustrated in Figure 3.5, the corresponding weighted SCD, $W_n^{J^*}(\boldsymbol{r}, \hat{\boldsymbol{\Omega}}, T)$, accounts for all of the incoming light paths that lead to the point $\boldsymbol{r} + s(T)\hat{\boldsymbol{\Omega}}$, but it does not account for the final segment that the light must traverse from the final scatter point $\boldsymbol{r} + s(T)\hat{\boldsymbol{\Omega}}$ to $\boldsymbol{r}$. To account for this final segment, the product of its number density integral $S(\boldsymbol{r}, \hat{\boldsymbol{\Omega}}, s(T))$ and its corresponding radiance contribution $J_n^*(\boldsymbol{r}, \hat{\boldsymbol{\Omega}}, T)dT$ must be included:

$$W_n^I(\boldsymbol{r}, \hat{\boldsymbol{\Omega}}) = \int_0^1 \left(W_n^{J^*}(\boldsymbol{r}, \hat{\boldsymbol{\Omega}}, T) + S(\boldsymbol{r}, \hat{\boldsymbol{\Omega}}, s(T)) J_n^*(\boldsymbol{r}, \hat{\boldsymbol{\Omega}}, T) \right) dT \,. \tag{3.13}$$

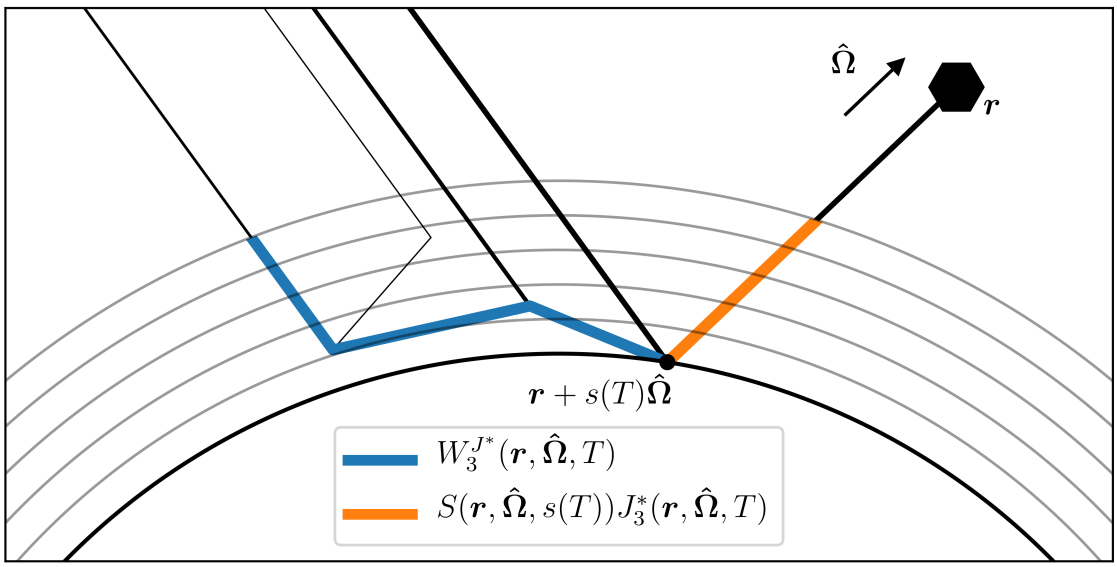


Figure 3.5: Illustration of the paths that are accounted for by each term in the weighted SCD calculation in Equation (3.13). Note that $W_3^{J^*}(\boldsymbol{r}, \hat{\boldsymbol{\Omega}}, T)$ accounts for many paths leading to the point $\boldsymbol{r} + s(T)\hat{\boldsymbol{\Omega}}$, not just the one shown here.

Equations (3.10), (3.11), and (3.13) can be evaluated with Monte Carlo integration, producing

the following equations, which are analogous to Equation 2.37:

$$W_n^I(\boldsymbol{r},\hat{\boldsymbol{\Omega}}) = \left\langle W_n^{J^*}(\boldsymbol{r},\hat{\boldsymbol{\Omega}},T) + S(\boldsymbol{r},\hat{\boldsymbol{\Omega}},s(T))J_n^*(\boldsymbol{r},\hat{\boldsymbol{\Omega}},T)\right\rangle_{T\sim\text{uni}(0,1)} \tag{3.14a}$$

$$W_n^J(\boldsymbol{r},\hat{\boldsymbol{\Omega}}) = 4\pi\omega_0(\boldsymbol{r})\left\langle W_{n-1}^I(\boldsymbol{r},\hat{\boldsymbol{\Omega}}')\right\rangle_{\hat{\boldsymbol{\Omega}}'\sim p(\boldsymbol{r},\hat{\boldsymbol{\Omega}},\hat{\boldsymbol{\Omega}}')} \tag{3.14b}$$

$$W_{\tilde{I}_n}(\boldsymbol{r},\hat{\boldsymbol{\Omega}}) = 2\pi a_0(\boldsymbol{r},\hat{\boldsymbol{\Omega}})\left\langle W_{n-1}^I(\boldsymbol{r},\hat{\boldsymbol{\Omega}}')\right\rangle_{\hat{\boldsymbol{\Omega}}'\sim b(\boldsymbol{r},\hat{\boldsymbol{\Omega}},\hat{\boldsymbol{\Omega}}')}\,. \tag{3.14c}$$

The sample spaces in Equations (2.37) and (3.14) can be filled simultaneously using common ray tracing, which introduces correlation between terms, but greatly reduces computation time. Let N ray traces be performed following the technique described in Section 2.3.8. Each trace, indexed here by j, produces radiance samples $I_{n,j}$, as before, but also weighted SCD samples $W_{n,j}$ for each order of scattering n. To reduce complexity in the current implementation, all orders are summed at the time the sample is taken, resulting in two samples, I_j and W_j, for every ray trace. This means that order-specific SCDs and their variances cannot be computed, but it eliminates the need to keep extensive correlation information between orders. The means, the variances, and the covariance of $I(\boldsymbol{r},\hat{\boldsymbol{\Omega}})$ and $W(\boldsymbol{r},\hat{\boldsymbol{\Omega}})$ are computed with the following running sums, following the results from Appendix B.3:

$$W = \frac{1}{n}\sum_{j=1}^{n} W_j \qquad \text{Var}(W) = \frac{1}{n-1}\left(\frac{1}{n}\sum_{j=1}^{n} W_j^2 - W^2\right) \tag{3.15}$$

$$I = \frac{1}{n}\sum_{j=1}^{n} I_j \qquad \text{Var}(W) = \frac{1}{n-1}\left(\frac{1}{n}\sum_{j=1}^{n} I_j^2 - I^2\right) \tag{3.16}$$

$$\text{Cov}(I,W) = \frac{1}{n-1}\left(\frac{1}{n}\sum_{j=1}^{n} I_j W_j - IW\right)\,. \tag{3.17}$$

The final SCD is then the ratio of W and I,

$$S^I = \frac{W}{I}\,, \tag{3.18}$$

with variance given by the following linear approximation derived in Section A.3:

$$\mathrm{Var}(S^I) = \frac{W^2}{I^2}\left(\frac{\mathrm{Var}(W)}{W^2} - 2\frac{\mathrm{Cov}(W,I)}{WI} + \frac{\mathrm{Var}(I)}{I^2}\right). \tag{3.19}$$

As a final step, the SCD $S_I(\boldsymbol{r}, \hat{\boldsymbol{\Omega}})$ and its variance are divided by the VCD to calculate the AMF and its variance.

Box Air Mass Factors

A small change to the process described in Equations (3.8) through (3.19) allows for box-AMFs to be computed in the place of the total AMF. Replacing Equation (3.9) with an integral that is constrained to a single layer,

$$S^i(\boldsymbol{r}, \hat{\boldsymbol{\Omega}}, s) \equiv \int_s^0 n(\boldsymbol{r} + t\hat{\boldsymbol{\Omega}})\phi^i(\boldsymbol{r} + t\hat{\boldsymbol{\Omega}})dt\,, \tag{3.20}$$

where $\phi^i(\boldsymbol{r})$ is equal to one when $\boldsymbol{r}$ is within layer i and zero otherwise, results in a layer-specific SCD. The box-AMF of this layer is then equal to the layer-specific SCD divided by the VCD within the layer. This can be done concurrently for many layers. If the layers are small enough that the number density does not change significantly within the layer, box-AMFs can be approximated by replacing the SCD with the path length,

$$S^i(\boldsymbol{r}, \hat{\boldsymbol{\Omega}}, s) \equiv \int_s^0 \phi^i(\boldsymbol{r} + t\hat{\boldsymbol{\Omega}})dt\,, \tag{3.21}$$

and by dividing by the vertical thickness of the layer rather than the VCD.

Example box-AMF calculations are shown in Figure 3.6, alongside HR box-AMFs for comparison. The figure illustrates how at one million photon paths traced, the MC AMFs appear to have converged to a reasonably smooth solution. It also demonstrates the motivation for the MC box-AMFs, which is to verify the HR box-AMFs and to help choose reasonable HR parameters: here it can be seen that with a single diffuse profile, the HR box-AMFs are underestimated by about 1%, which is corrected by introducing 4 additional diffuse profiles.

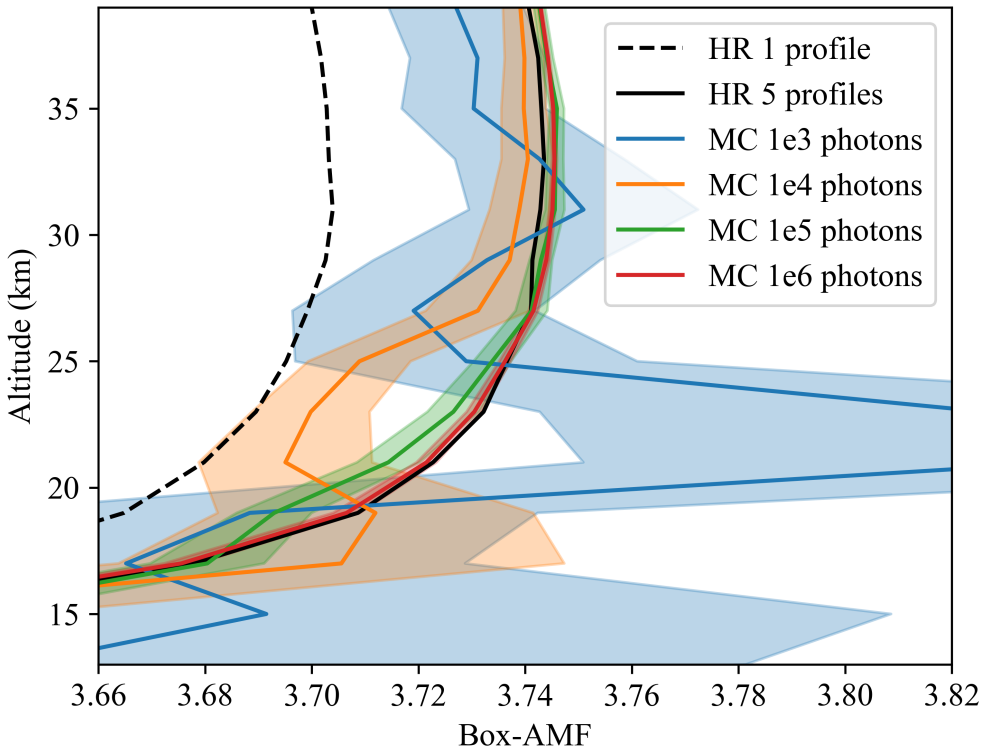


Figure 3.6: Comparison between HR and MC AMFs. The highlighted regions show the estimated standard deviation of the MC AMFs. The conditions match those in Figure 3.4b.

3.4 Retrievals

End-to-end retrieval algorithms have been implemented, utilizing all of the tools previously described in this section. Two algorithms are currently implemented, corresponding to the SP2 algorithm (Bucsela et al., 2013) and the OMI total ozone algorithm (Veefkind et al., 2006) referenced in Section 3.1. The implementations are object oriented, designed to allow for any individual piece of the algorithm to be changed with little effort. For example, by default box-AMF profiles are provided to the algorithm via lookup tables, but it is relatively simple to replace this piece of the algorithm with direct scene-by-scene box-AMF calculations, which would be prohibitively slow for large datasets but which may be useful for verification on a smaller scale. Retrievals over cloudy scenes have not yet been implemented, but the structure is all in place for when this becomes necessary.

Ozone Total Column Retrieval

The ozone total column retrieval utilizes a cloud-free version of Equation (2.82) to calculate the total AMF,

$$A = \frac{\boldsymbol{v}^T \boldsymbol{a}_{\text{clear}}(a_s)}{V}\,, \tag{3.22}$$

and Equation (2.85), repeated here, to calculate the VCD and its uncertainty:

$$V = \frac{S}{A} \qquad \sigma_V^2 = \left(\frac{\sigma_S}{A}\right)^2 + \left(\frac{S\sigma_A}{A}\right)^2\,. \tag{3.23}$$

A thorough AMF uncertainty analysis following Equations (2.83) and (2.84) has not been implemented for ozone, although it will a simple matter to copy this over from the NO_2 retrieval when it is required. In summary, the implemented retrieval consists of the following steps:

1. Perform a spectral fit to retrieve the ozone SCD and its uncertainty (see Section 3.1).
2. Calculate the box-AMF profile (see Section 3.3).
3. Acquire an a priori ozone profile.
4. Calculate the total AMF and assume a constant percent uncertainty using Equation (3.22).
5. Calculate the VCD and its uncertainty using Equation (3.23).

NO_2 Tropospheric Column Retrieval

The NO_2 tropospheric column retrieval introduces a temperature correction to the AMF calculation,

$$A = \frac{\boldsymbol{v}^T \alpha(\boldsymbol{T}) \boldsymbol{a}_{\text{clear}}(a_s)}{V}\,, \tag{3.24}$$

where $\alpha(\boldsymbol{T})$ is the temperature correction factor, a diagonal matrix with elements $\alpha^i = 1 - 0.003(T^i - T_0)$, and where T^i is the temperature in layer i and T_0 is the temperature of the cross sections used for the spectral fitting, in this case 220 K. The temperature correction factor accounts for the effects of this constant temperature cross section, and the value

0.003 was determined empirically by analyzing synthetic radiance data (Bucsela et al., 2013). Equation (3.24) is applied twice, once for the tropospheric AMF and once for the stratospheric AMF, with each of them only using atmospheric layers from their respective regions.

In the following uncertainty analysis, several constant uncertainties are assumed based on suggestions from Bucsela et al. (2013). For the tropospheric AMF A_t, the uncertainty σ_{A_t} is calculated using cloud-free variants of Equations (2.83) and (2.84) that include the temperature correction factor:

$$\sigma_{A_{a_s}}^2 = \left(\frac{\partial A}{\partial a_s}\right)^2 \sigma_{a_s}^2 \qquad \frac{\partial A}{\partial a_s} = \frac{\boldsymbol{v}^T \alpha(\boldsymbol{T})}{V} \frac{\partial \boldsymbol{a}_{\text{clear}}}{\partial a_s} \tag{3.25a}$$

$$\sigma_{A_V}^2 = (\boldsymbol{\nabla}_v A)^T \Sigma_v (\boldsymbol{\nabla}_v A) \qquad \boldsymbol{\nabla}_v A = \frac{\alpha(\boldsymbol{T})\boldsymbol{a}_{\text{clear}}(a_s) - A\boldsymbol{u}}{V} \tag{3.25b}$$

$$\sigma_{A_t}^2 = \sigma_{A_{a_s}}^2 + \sigma_{A_V}^2 \,, \tag{3.26}$$

where the uncertainty σ_{a_s} in the surface albedo a_s is assumed to be a constant 0.015.

The stratospheric AMF A_s has a smaller uncertainty, as it is much less sensitive to uncertainties in ground reflectance, and the stratospheric NO_2 profile has much less variability then its tropospheric counterpart. The uncertainty σ_{A_s} is assumed to be a constant 2 %. The retrieved tropospheric column is then given by Equation (2.86), repeated here:

$$V_t = \frac{S - V_s A_s}{A_t} \qquad \sigma_{V_t}^2 = \left(\frac{\sigma_S}{A_t}\right)^2 + \left(\frac{A_s \sigma_{V_s}}{A_t}\right)^2 + \left(\frac{V_s \sigma_{A_s}}{A_t}\right)^2 + \left(\frac{V_t \sigma_{A_t}}{A_t}\right)^2 , \tag{3.27}$$

where the uncertainty σ_{V_s} of the stratospheric column V_s is assumed to be a constant $2 \times 10^{14}\,\text{cm}^{-2}$. In summary, the implemented retrieval consists of the following steps:

1. Perform a spectral fit to retrieve the NO_2 SCD and its uncertainty (see Section 3.1).
2. Calculate the box-AMF profile and its derivative with respect to albedo (see Section 3.3).
3. Acquire an a priori NO_2 profile and its covariance matrix.
4. Calculate the tropospheric AMF and its uncertainty, assuming an uncertainty of 0.015 for the albedo using Equations (3.24) and (3.25).

5. Calculate the stratospheric AMF and assume a 2 % uncertainty using Equation (3.24).
6. Calculate the VCD and its uncertainty, assuming an uncertainty of $2 \times 10^{14}\,\mathrm{cm}^{-2}$ for the stratospheric column using Equation (3.27).

4 DOAS Applications

4.1 TROPOMI Validation Campaign

The tools described in Section 3.3 were used to develop an AMF lookup table that was used by collaborators at Environment and Climate Change Canada (ECCC) for their validation campaign for TROPOMI over the oil sands in Alberta. The ECCC validation campaign consisted of comparing TROPOMI NO_2 data with ground based remote measurements and in situ aircraft measurements, and the results were published in late 2018 (Griffin et al., 2018). These AMFs were requested by ECCC to facilitate future improvements, such as accounting for aerosols and non-Lambertian ground reflection features, especially over snow.

Preliminary results showed minor improvements with the new AMFs before implementing any of these new features. These results are shown in Figure 4.1, which is an early variant of Figure 2 from Griffin et al. (2018), received from Griffin via personal communication during the preparation of the publication. The figure illustrates a low bias in the TROPOMI NO_2 data product over the snow-covered oil sands which is improved by using new AMFs in the place of the TROPOMI AMFs. The left plots compare the original TROPOMI VCDs with the in situ aircraft VCDs, the middle plots compare the recalculated VCDs with the aircraft VCDs, and the right plots compare AMFs from all three methods for every scene. In the right plot, it can be seen that the aircraft AMFs, which are considered the most accurate measurement in this scenario, compare better with the new AMFs than with the original TROPOMI AMFs. The TROPOMI AMFs are consistently biased high, resulting in the low bias of the TROPOMI VCDs. The top row in Figure 4.1 was made using AMFs calculated by ECCC, and the bottom row was made using AMFs calculated with SASKTRAN as part of the present work. The SASKTRAN AMFs compare slightly better with the aircraft AMFs than

the ECCC AMFs, although the improvement is small enough that it should be interpreted as verification of the SASKTRAN AMFs, rather than as a significant upgrade.

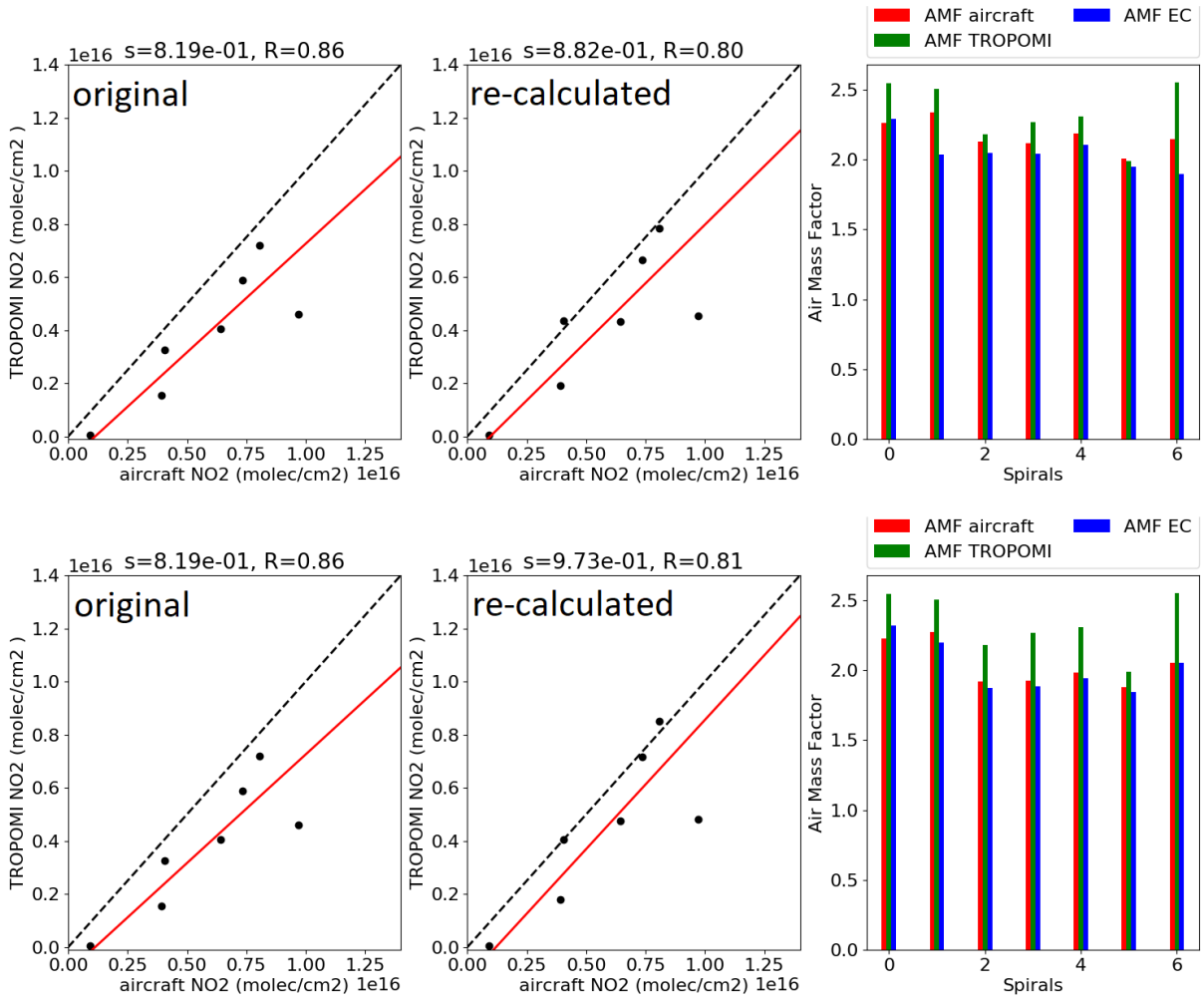


Figure 4.1: Comparison between TROPOMI data and aircraft measurements, using TROPOMI AMFs, ECCC AMFs, and SASKTRAN AMFs. Early iteration of Figure 2 of Griffin et al. (2018), received from Griffin via personal communication.

The AMF tables sent to ECCC were calculated using the DO engine, with 500 m spacing up to an altitude of 12 km. One table was calculated for clear conditions, and one for cloudy conditions, with entries corresponding to every combination of the applicable parameters defined in Table 4.1. Since this work was published, in continued collaboration with ECCC, the table has been recalculated using the HR engine, which is more computationally expensive but more accurate, especially at high zenith angles, and extended to an altitude of 50 km. Aerosol properties are currently being incorporated into the table to facilitate its use in the

analysis of measurements taken over forest fires.

Parameter	Values
Solar Zenith Angle (°)	0, 30, 50, 60, 65, 70, 73, 76, 78, 80
Viewing Zenith Angle (°)	0, 30, 50, 60, 65, 70, 72
Solar-Viewing Azimuth Angle Difference (°)	0, 30, 60, 90, 120, 150, 180
Surface Pressure (kPa)	60, 80, 90, 100
Surface Albedo (Clear-sky only)	0, .03, .06, .09, .12, .2, .3, .5, .75, 1
Cloud Pressure (kPa) (Cloudy-sky only)	20, 40, 60, 80, 90
Cloud Albedo (Cloudy-sky only)	.8

Table 4.1: AMF table parameter definition.

4.2 Simulated TEMPO Sensitivity Study

The purpose of the following study is to begin to quantify and understand the distribution of the nominal uncertainties in TEMPO retrievals over Canada. Four sources of uncertainty have been considered: SCD uncertainty, stratospheric uncertainty, albedo uncertainty, and NO_2 profile shape uncertainty. The SCD uncertainty is a result of measurement noise, and is driven primarily by the SNR. In general it is possible to retrieve the albedo from the measurements themselves, but the NO_2 profile shape cannot be retrieved, and must therefore be taken from an external source such as a climatology. In both cases, uncertainty in the input data is propagated through to uncertainty the final retrieved quantities. Clouds were not included in this study, so the uncertainty due to cloud parameters, which is significant for scenes with significant cloud cover, was not considered. Another challenging source of uncertainty that was excluded from this study is the identification of snow, particularly during periods of snowing and melting.

The general method used for this study was to simulate TEMPO measurements (Sections 4.2.1 and 4.2.2), and then to perform a full retrieval with uncertainty analysis on the simulated measurements (Section 4.2.3). Simulations were performed across the entire FOR of TEMPO, but at a spatial resolution lower then TEMPO itself, in order to observe continent-wide trends while keeping computation time reasonable. Four days were analyzed, near each of the four seasonal solstices and equinoxes, in order to observe the extremes of seasonal effects, and simulations were performed on two hour intervals in order to observe diurnal variations.

4.2.1 Synthetic Radiance

Synthetic Radiance Inputs

Synthetic radiances were calculated for this study using the SASKTRAN-HR engine, on a 1° latitude by 1.25° longitude grid, in the 405 nm to 465 nm window, with a spectral sampling of 0.2 nm, and a spectral resolution of 0.6 nm. The spatial resolution of this configuration is much lower than that of TEMPO, but the spectral sampling and resolution are a good approximation. The atmosphere was defined for each ground location by altitude profiles in pressure, temperature, ozone concentration, and NO_2 concentration, as well as albedos representing Lambertian ground reflectance. Default SASKTRAN values were used for the pressure and temperature and ozone profiles, based on the MSIS-90 model by Hedin (1991) and the Labow climatology by McPeters, Labow, and Logan (2007) respectively. More detailed and realistic data products were used for NO_2 profiles and albedos, as the retrieval is much more sensitive to these parameters. The NO_2 profiles were taken from a global simulation of tropospheric chemistry performed using NASA Goddard Earth Observing System Model version 5 Earth System Model (GEOS-5 ESM) with GEOS-Chem as a chemical model (G5NR-Chem) at a resolution of 12.5 km (Hu et al., 2018). The albedos were taken from a climatology of Lambert-equivalent reflectivity (LER) based on over five years of GOME observations (Koelemeijer, de Haan, & Stammes, 2003). Example tropospheric NO_2 columns across the entire TEMPO FOR are shown in Figure 4.2, example stratospheric columns are shown in Figure 4.3, and example albedos are shown in Figure 4.4. Figures 4.2 and 4.3, as well as Figures 4.7 through 4.18 include lines of constant SZA for reference.

Synthetic Radiance Approximations

Explicitly calculating each spectrum was found to be prohibitively computationally expensive, so a series of approximations were used to reduce computation time. Firstly, one pressure profile and one temperature profile were chosen to represent the entire FOR, which permitted the construction of a lookup table containing radiances calculated in the absence of any absorbing species. This calculation includes the computationally expensive multiple scatter analysis, and it can be done at much fewer wavelengths while introducing limited interpola-

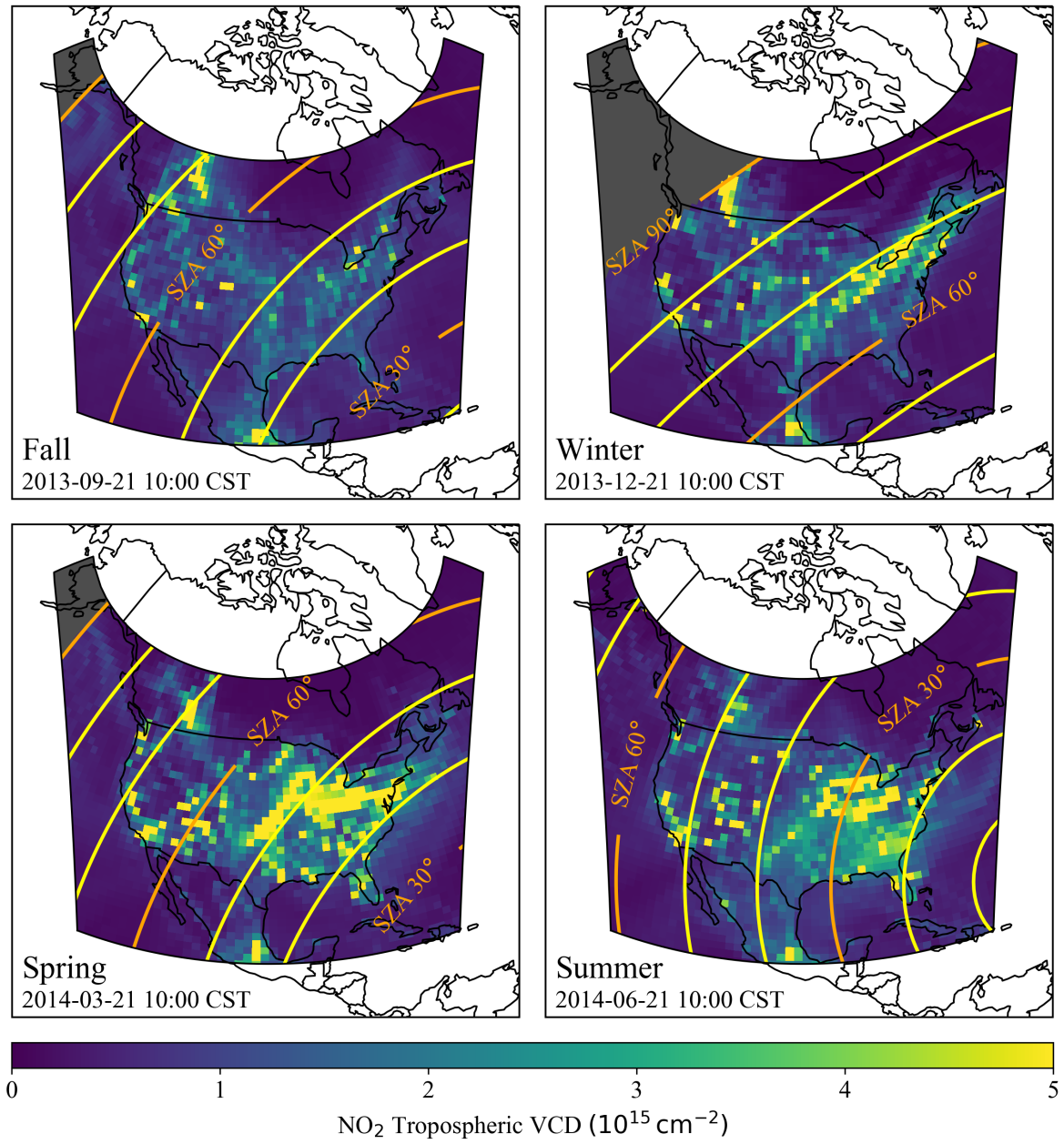


Figure 4.2: NO_2 Tropospheric Columns. Hotspots exceeding $5 \times 10^{15}\,\mathrm{cm}^{-2}$ are seen over urban locations and over the Alberta oil sands, and increased pollution is seen in spring and summer.

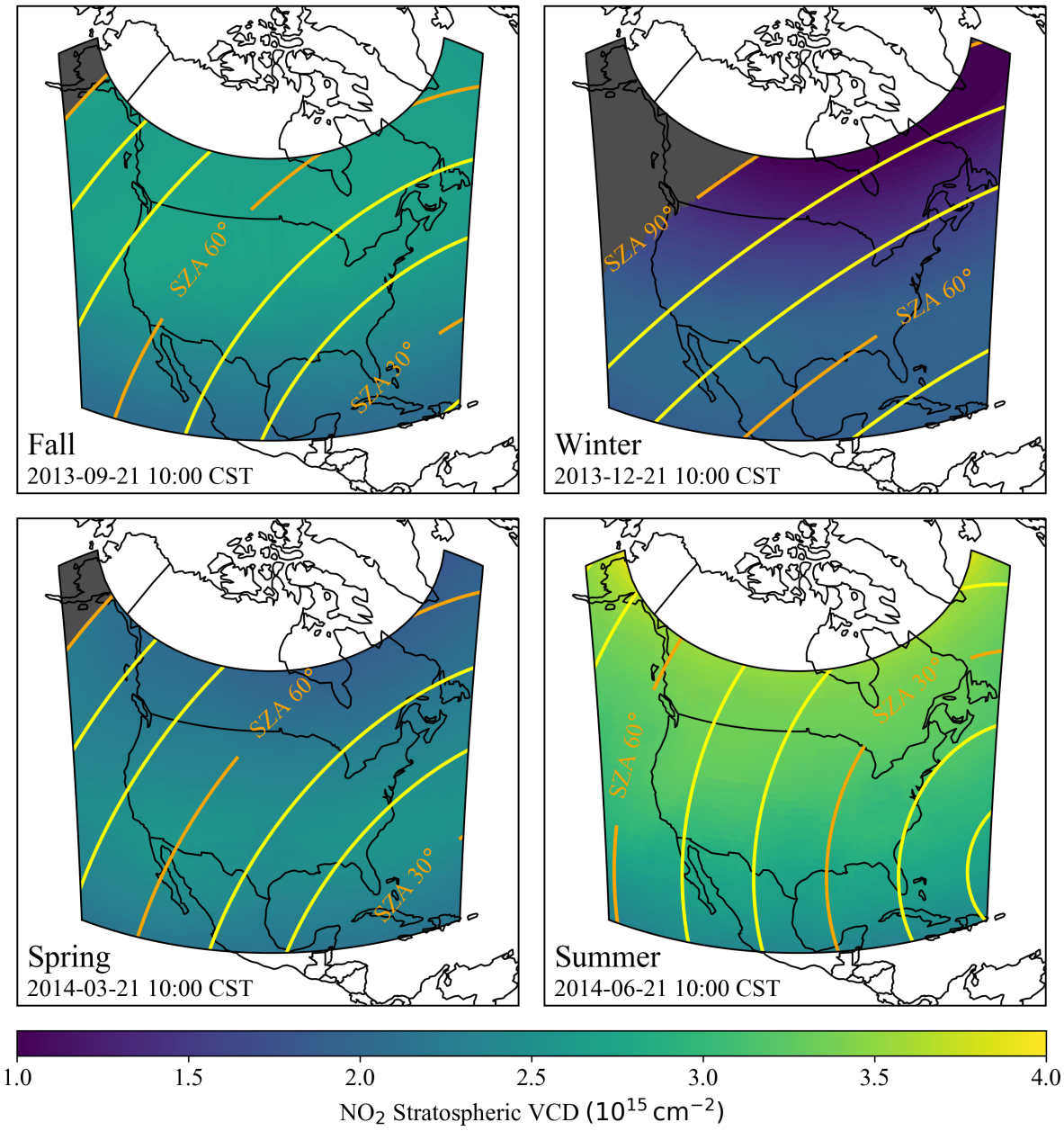


Figure 4.3: NO_2 Stratospheric Columns. The stratospheric column varies slowly and smoothly with SZA and latitude.

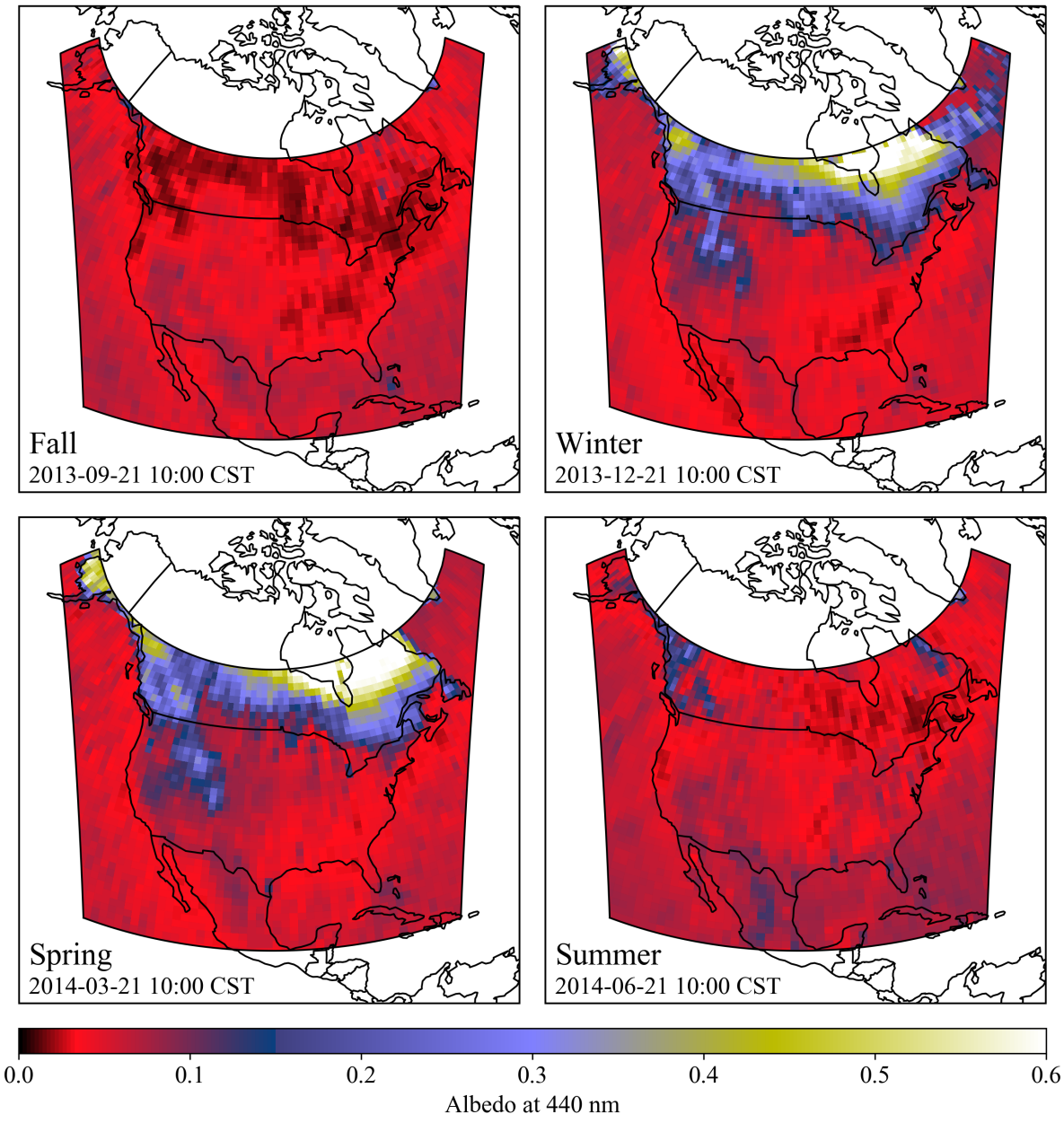


Figure 4.4: Lambertian equivalent reflectivity climatology from GOME. Typical scenes have albedos near 0.04, with much higher values corresponding to snow cover.

tion errors due to the slowly changing wavelength dependence of Rayleigh scattering. The lookup table had 24 SZAs and 24 VZAs spanning 0° to 90°, 13 azimuth difference angles spanning 0° to 180°, 4 albedo values (0.0, 0.1, 0.6, and 1.0), and 4 wavelengths (405 nm, 425 nm, 445 nm, and 465 nm). Linear interpolation was used for all of the angles, and a two dimensional cubic interpolation was used for the albedos and the wavelengths. This configuration minimized computation time, which was driven primarily by the number of required albedos and wavelengths, and kept interpolation errors under 1 % for zenith angles under 89°.

Secondly, absorption features were added to the absorption-free radiance spectra $I_0(\lambda)$ using Beer's law (see Section 2.3.2),

$$I(\lambda) = I_0(\lambda) e^{-\tau(\lambda)} \,, \tag{4.1}$$

where the optical depth $\tau(\lambda)$ is the total absorption due to NO_2 and ozone,

$$\tau(\lambda) = \tau_{NO_2}(\lambda) + \tau_{O_3}(\lambda) \,, \tag{4.2}$$

which can be calculated by integrating their respective extinctions along the slant path,

$$\tau_{NO_2}(\lambda) = \int_L k_{NO_2}(\lambda, l) dl \qquad \tau_{O_3}(\lambda) = \int_L k_{O_3}(\lambda, l) dl \,. \tag{4.3}$$

For ozone, the position dependence of the cross section was ignored, which decouples it from the slant path integral and allows the optical depth to be calculated as the product of the cross section and the SCD.

$$\tau_{O_3}(\lambda) = \sigma_{O_3}(\lambda) \int_L n_{O_3}(\lambda) dl = S_{O_3} \sigma_{O_3}(\lambda) \tag{4.4}$$

To efficiently approximate the ozone SCD, box-AMFs were combined with the vertical ozone profile, in the same way that the total AMF is calculated in Equation (2.70), but without dividing by the total VCD:

$$S_{O_3} = \sum_i s^i_{O_3} = \sum_i a^i v^i_{O_3} \,. \tag{4.5}$$

The box-AMFs were taken from the table described in Section 4.1, but with the altitude extended to 50 km and the SZA extended to 90°.

For NO_2, the position dependence of the cross section was not ignored, as this would compromise the retrieval, which explicitly accounts for the temperature dependence of the NO_2 cross section. The temperature dependence of the NO_2 cross section was linearized, in the same way the ozone cross section was linearized in Section 3.1, and the cross section was incorporated into the sum over layers:

$$\tau_{NO_2}(\lambda) = \sum_i a^i v^i_{NO_2} \sigma_{NO_2}(\lambda, T^i)\,, \tag{4.6}$$

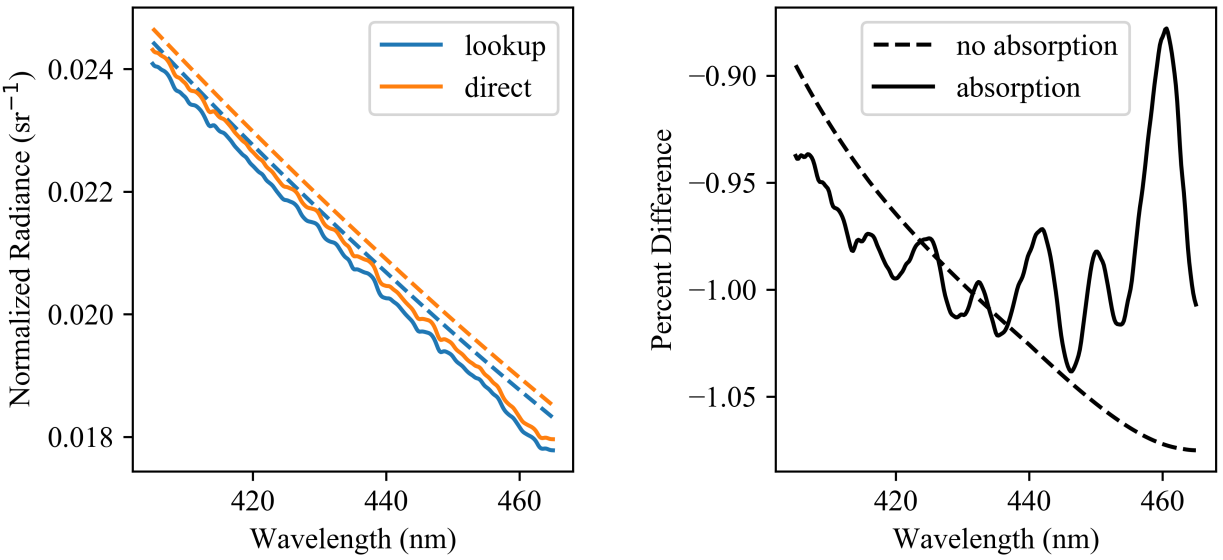
where T^i is the average temperature in layer i.

The error introduced by all of the assumptions described above is illustrated in Figure 4.5. For the nominal scene shown in the figure, the combination of interpolation errors and neglecting temperature and pressure profile variations results in an error in the absorption-free radiance on the order of 1 %, and the remaining assumptions made while applying the absorption features does not introduce significant error. Note that most of the additional error introduced in the second step is due to the difference in ozone profile, as the signature in the percent difference matches the signature of the ozone cross section. This does not impact the retrieval significantly because the retrieval extracts ozone alongside the NO_2.

After the synthetic radiances were calculated, noise was added. Wavelength-dependent SNRs were estimated for every scene, following the method described in Section 4.2.2. Then, for every simulated wavelength, Gaussian noise was added to the radiance values:

$$I_{\text{noise}}(\lambda) = I(\lambda)\left(1 + \frac{z}{\text{SNR}(\lambda)}\right)\,, \tag{4.7}$$

where z is a sample of a standard normal random variable.



(a) Synthetic radiances.

(b) Percent difference.

Figure 4.5: Comparison of synthetic radiances: lookup method versus direct calculation.

4.2.2 Signal-to-Noise Ratio

Instrument Noise

Estimating measurement noise is a crucial step in the uncertainty analysis process, as noise estimates are used as uncertainties of radiance measurements, which are propagated through the spectral fitting process to produce uncertainties in the SCDs, which are one of the primary contributions to uncertainties in the retrieved VCDs. There are three primary sources of measurement noise: stray light, dark current, and shot noise. Stray light noise is a consequence of the optical design in which light from outside the intended FOR for a given pixel is detected by that pixel. Dark current is all electronic noise that exists in the absence of all incident light. Shot noise, also called Poisson noise, is a consequence of the discrete nature of light. The emission of a single photon (and therefore the detection of a single photon) is a random event which is independent of other emission events. As a result, photons arriving at a detector follow a Poisson process, so that the photons counted by the detector during its integration time have a variability equal to the square root of the number of photons counted. This variability would be observed even with a perfectly consistent source and a

perfectly efficient detector; the only method for reducing shot noise is to increase the number of counted photons, for example by increasing the integration time so that the relative variability is reduced.

Stray light and dark current characterization are both important for fine tuning and for accurately analyzing scenarios with low signal levels, but they are not essential components of the current study. Therefore for the current study, it is assumed that all noise is shot noise. This assumption creates a simple relationship between the SNR and the number of photons counted by the detector. Letting the number of detected photons be n, the SNR becomes the ratio of the signal n and the noise $\sqrt{n}$:

$$\text{SNR} = \frac{n}{\sqrt{n}} = \sqrt{n}\,. \tag{4.8}$$

The number of photons detected by any given pixel can theoretically be calculated, given the incident radiance, if sufficient information about the instrument is known, as discussed in Section 2.3.1. To a good approximation, this can be done using the following formula,

$$n = \Delta t\,\Delta\Omega\,\Delta A \int \eta(\lambda)\text{ILS}(\lambda)I(\lambda)d\lambda\,, \tag{4.9}$$

where Δt is the integration time, $\Delta\Omega$ is the solid angle corresponding to the FOR of the pixel, ΔA is the area of the pixel projected onto the instrument aperture, $\eta(\lambda)$ is a fraction representing the end-to-end efficiency of the detector, $\text{ILS}(\lambda)$ is the instrument line shape of the pixel, and $I(\lambda)$ is the incident radiance.

Estimating SNR

SNRs averaged across selected wavelength windows were reported by Zoogman et al. (2017), and a more detailed breakdown of SNR requirements and margins can be found in presentations from TEMPO Science Team Meetings (Nicks, 2014; 2016). The reported values for 2016 appear to be in reasonable agreement across the different sources, as shown in Figure 4.6. The 2016 Science Team Meeting values were used for the following analysis, as they were the most detailed and up to date values found (Nicks, 2016).

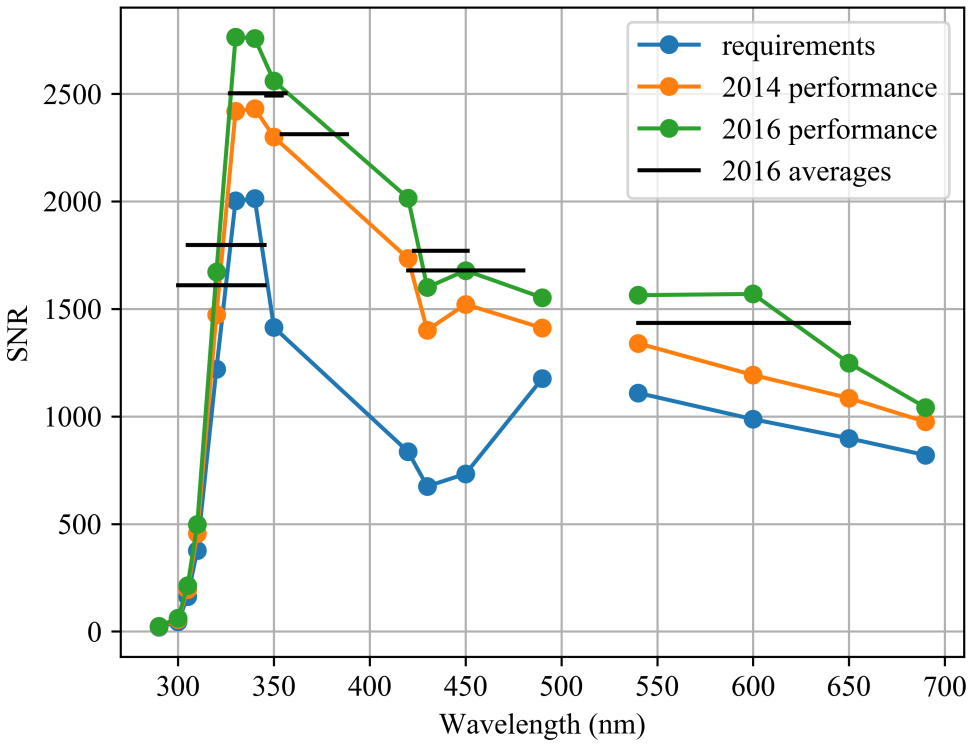


Figure 4.6: TEMPO SNR requirements (Nicks, 2014), performance in 2014 (Nicks, 2014), performance in 2016 (Nicks, 2016), and window-averaged performance in 2016 (Zoogman et al., 2017).

SNR does not depend only on the instrument itself, but also on the scene that the instrument is viewing, and the exact details of the assumed scene associated with the SNR values were not reported by any of the above sources. A nominal scene was assumed, pointing at the center of the TEMPO FOR at 36.5°N latitude and 100°W longitude, at midday near the summer solstice, with a SZA of 25°, a VZA of 42°, and with albedo values taken from the GOME LER climatology. Note that this is similar to the scene that is explicitly described for a nominal ozone profile retrieval test case in Figure 8 of Zoogman et al. (2017).

By comparing radiances from this assumed nominal scene to the reported nominal SNRs, the instrument can be characterized well enough to estimate SNRs for arbitrary scenes. Let the radiance of the assumed nominal scene be given by $I_0(\lambda)$, and let n be the number of corresponding photons detected by a given pixel. Then Equation (4.9) can be approximated as follows,

$$n \approx \alpha(\lambda_n) I_0(\lambda_n)\,, \tag{4.10}$$

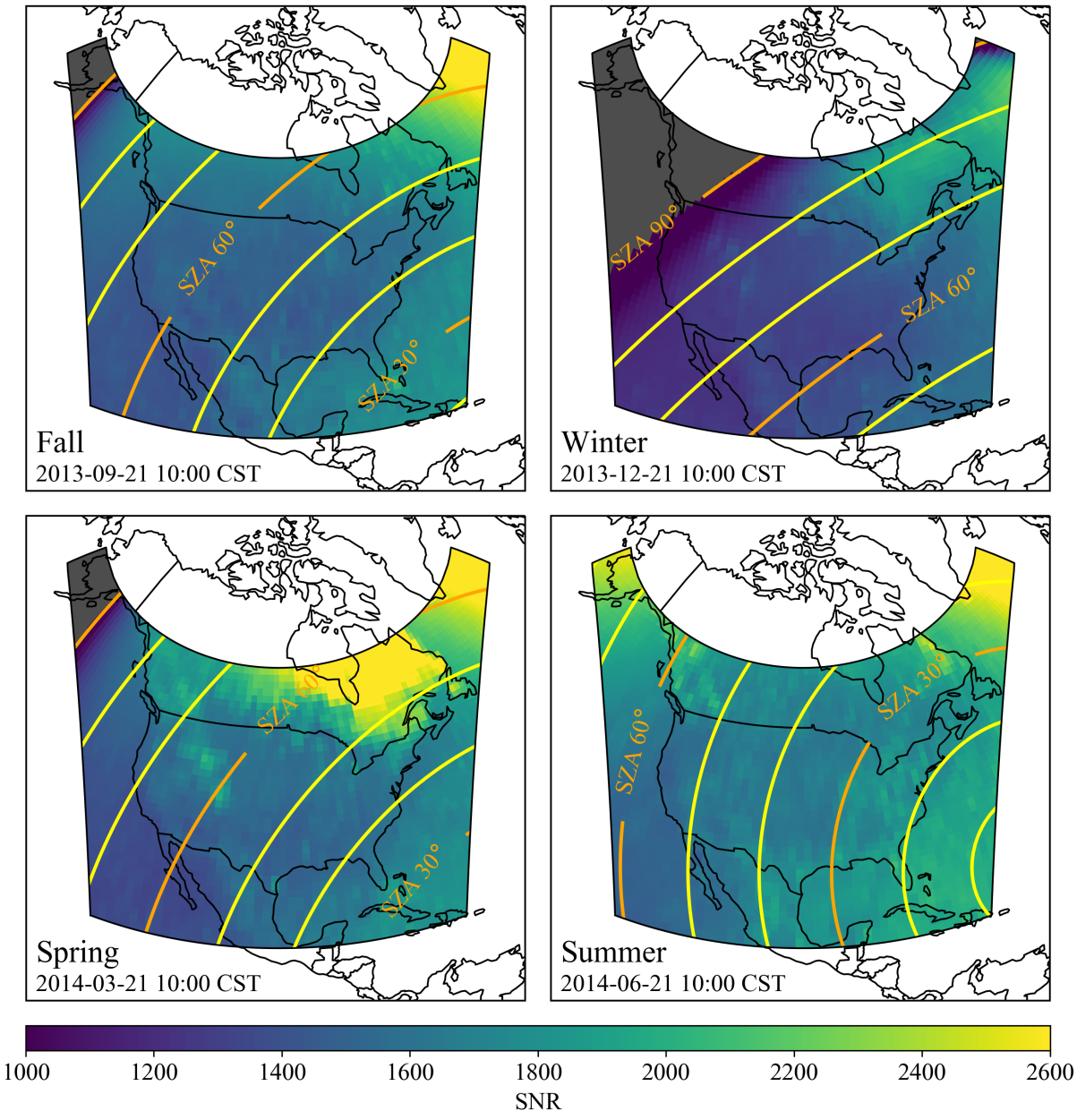


Figure 4.7: Average SNR from 405 nm to 465 nm. SNR decreases as SZA decreases due to increased attenuation. Attenuation also increases with VZA, but the SNR tends to increase with VZA due to increased scattering into the longer line of sight. SNR is sensitive to albedo, increasing significantly over snow-covered regions.

where λ_n is the central wavelength of the pixel and $\alpha(\lambda)$ is a constant of proportionality containing all of the instrument information described in Equation (4.9). Using Equation (4.8), the SNR associate with this scene and this pixel is then given by

$$\mathrm{SNR}_0(\lambda_n) = \sqrt{\alpha(\lambda_n) I_0(\lambda_n)}\,, \tag{4.11}$$

which can be used to solve for $\alpha(\lambda_n)$,

$$\alpha(\lambda_n) = I_0(\lambda_n)\mathrm{SNR}_0(\lambda_n)^2\,. \tag{4.12}$$

This was done for every wavelength where the SNR was specified (see Figure 4.6) to determine the function $\alpha(\lambda)$, which was then used to estimate the SNR for arbitrary scenes:

$$SNR(\lambda) = \sqrt{\alpha(\lambda) I(\lambda)}\,. \tag{4.13}$$

Figure 4.7 shows an example of SNR calculated over the TEMPO FOR using this strategy.

4.2.3 Retrieval and Uncertainty Analysis

The NO_2 retrieval algorithm with uncertainty analysis described in Section 3.4 was applied to the synthetic radiances described in Section 4.2.1 with the estimated SNRs described in Section 4.2.2.

NO_2 Profile Variability

To determine the NO_2 profile variability required for the tropospheric AMF uncertainty analysis, represented by Σ_V in Equation (3.25b), a monthly climatology was constructed from the same GEOS-5 ESM with GEOS-Chem simulation that was used to calculate the synthetic radiances in Section 4.2.1. A monthly mean $\boldsymbol{\mu}_V$ and a corresponding covariance matrix Σ_V were calculated as follows:

$$\boldsymbol{\mu}_V = \frac{1}{N}\sum_{i=1}^{N} \boldsymbol{V}_i \qquad \Sigma_V = \frac{1}{N}\sum_{i=1}^{N} \boldsymbol{V}_i \boldsymbol{V}_i^T - \boldsymbol{\mu}_V \boldsymbol{\mu}_V^T\,, \tag{4.14}$$

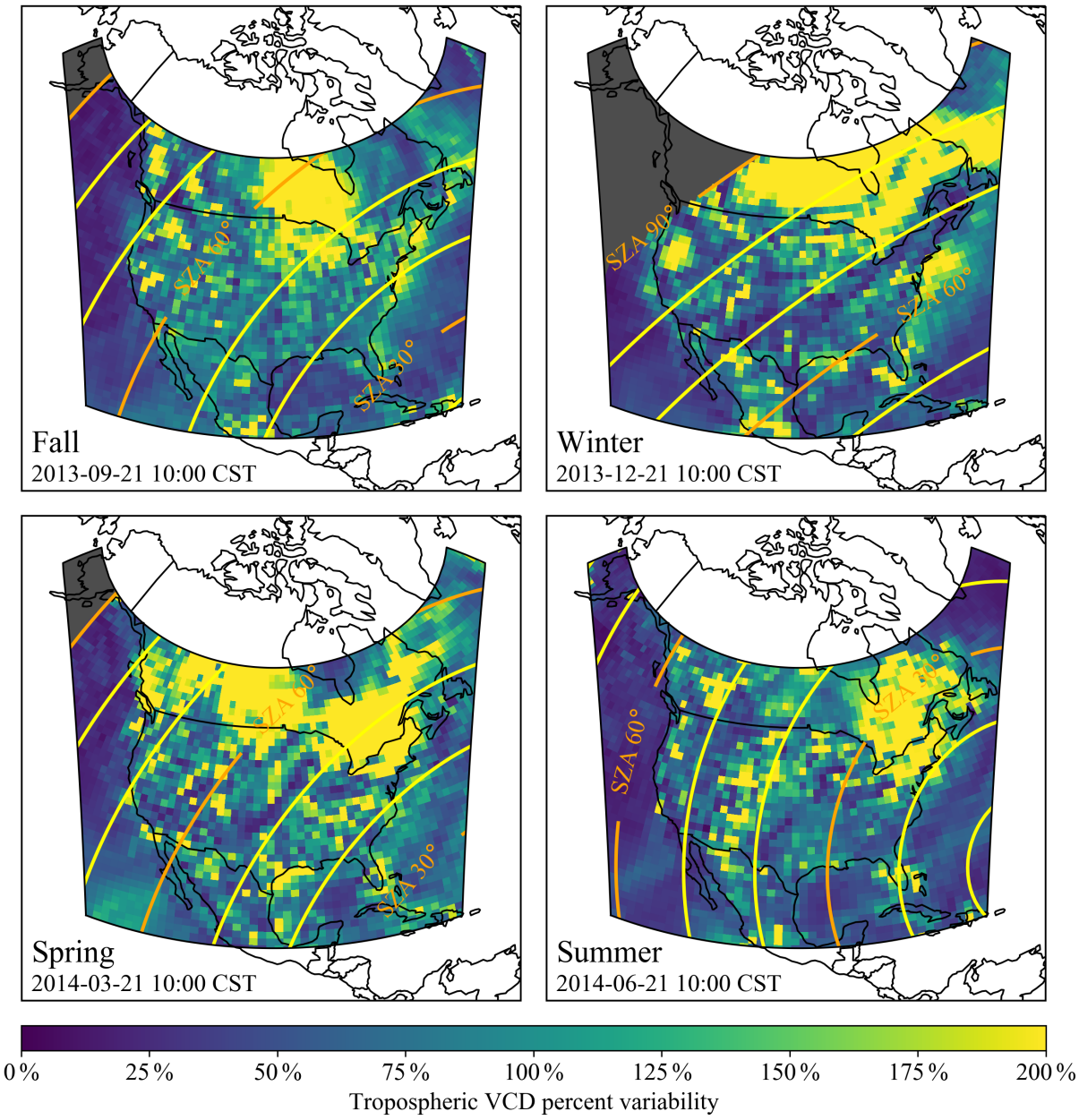


Figure 4.8: Relative uncertainty in the average tropospheric NO_2 VCD in monthly climatology constructed from the GEOS-5 ESM with GEOS-Chem simulation.

where N is equal to two times the number of days in the month, and the corresponding summation i is over two profiles per day taken from the early afternoon, approximately 13:00 and 14:00 local time. The vectors $\boldsymbol{V}_i$ contain NO_2 volume mixing ratios (VMRs) defined at the 24 lowermost hybrid sigma-pressure (HSP) levels, discussed in Appendix D.1, which was found to be sufficient to cover the troposphere. Scene-specific pressure profiles were taken from the MSIS-90 model to integrate the VMR-HSP data within the altitude-defined layers on which the box-AMFs are defined, as discussed in Appendix D.3. The variability of this climatology is illustrated in Figure 4.8. More precisely, what is shown in this figure is the relative uncertainty in the tropospheric VCD σ_V/V, where V is the average tropospheric VCD, equal to the sum of the elements of μ_V, and where σ_V^2 is the variance of V, equal to the sum of the elements of Σ_V as discussed in Appendix B.3. The high relative uncertainty is found in regions of infrequent pollution - for example a location that has low pollution throughout an entire month except for one day, when a plume from a nearby source passed through, will have a low average but a relatively high variability. Lower relative uncertainty is found in high-pollution regions such as the Alberta oil sands and central USA during spring.

Intermediate Results

Snapshots of the retrieved SCDs, the tropospheric AMFs, and the stratospheric AMFs are shown in Figures 4.9 through 4.11. The retrieved VCDs are not displayed here, as they would look nearly identical to the input VCDs that are already shown in Figure 4.2. It is worth noting that the purpose of this study is not to show that the retrieved VCDs match the VCDs used to create the synthetic radiances. This result is essentially trivial, as the exact parameters used to create the synthetic radiances are known, and were used as a priori information in the retrievals. The purpose, rather, is to examine the relative weight of all the uncertainty terms. Snapshots of SCD uncertainty and AMF uncertainty are shown in Figures 4.12 through 4.14.

VCD Uncertainty

The contributions of the intermediate value uncertainties illustrated in Figures 4.12 through 4.14 to the uncertainty in the final retrieved VCD is shown in Figures 4.15 through 4.18.

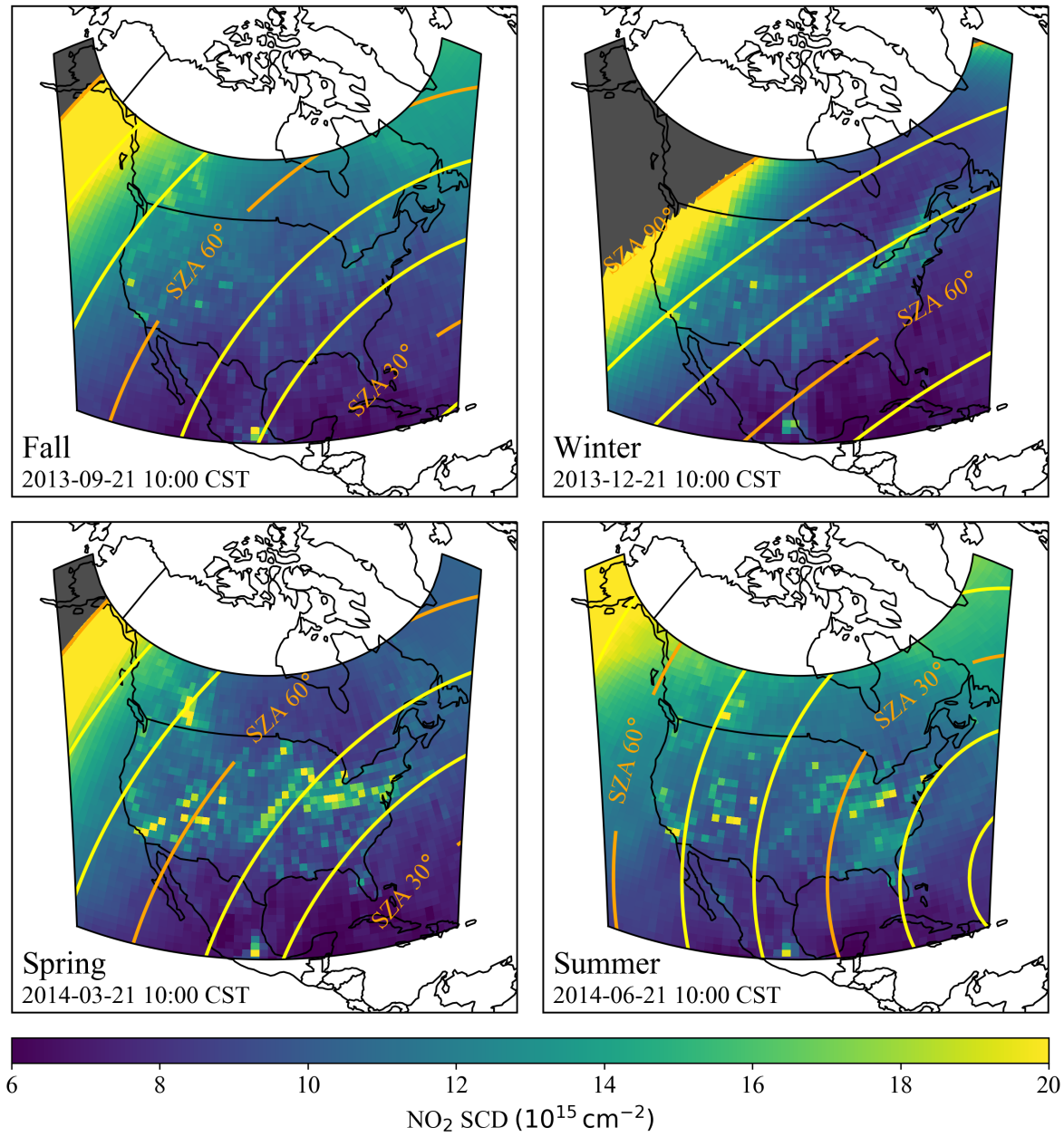


Figure 4.9: NO_2 SCDs. The SCD is large over polluted scenes such as major cities and the Alberta oil sands due to increased NO_2 concentration, and also grows as the SZA and VZA increases due to the longer slant path.

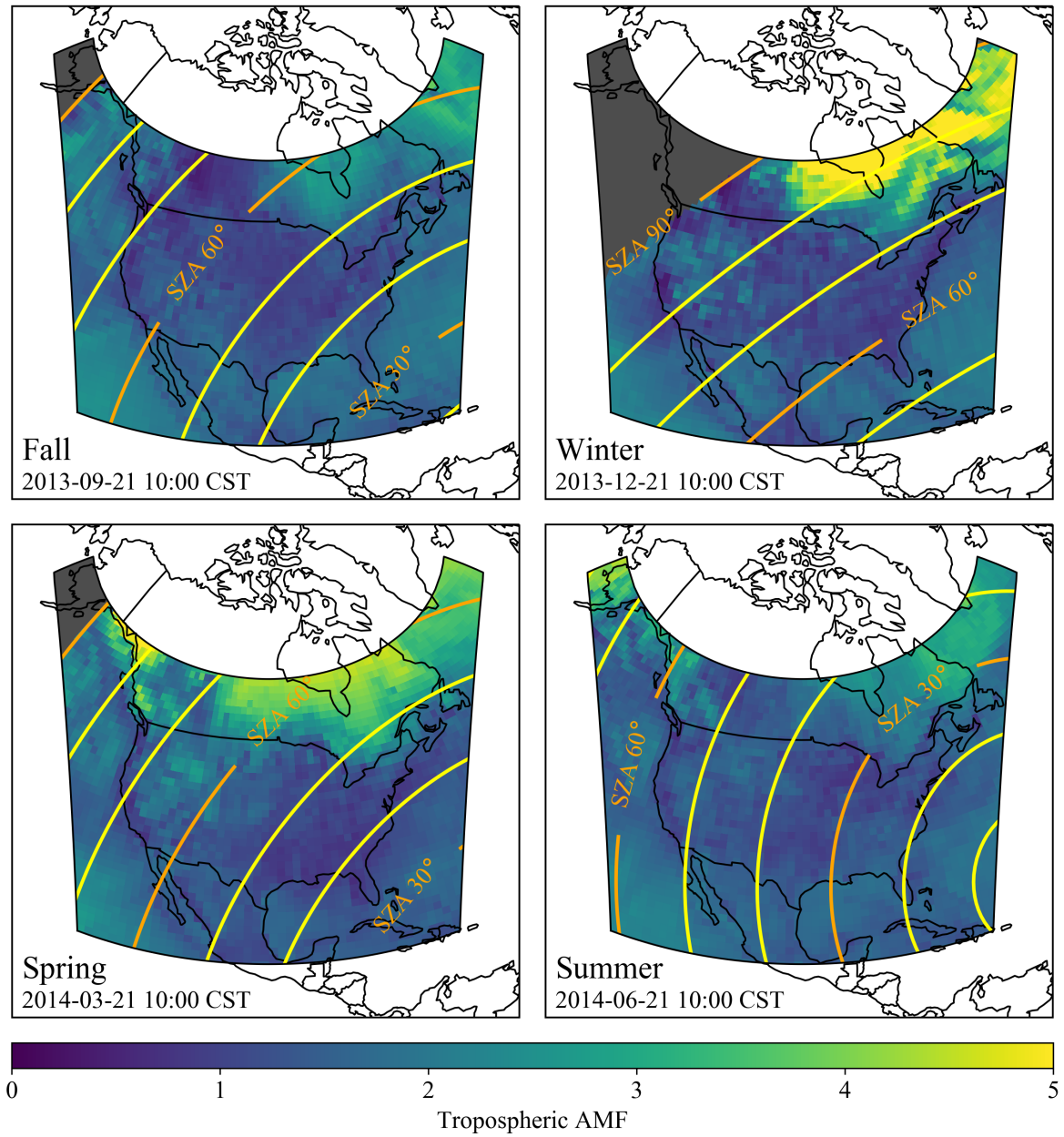


Figure 4.10: Tropospheric AMFs. The tropospheric AMF appears to be primarily sensitive to surface albedo, as the high reflectivity of snowy scenes increases measurement sensitivity near the surface.

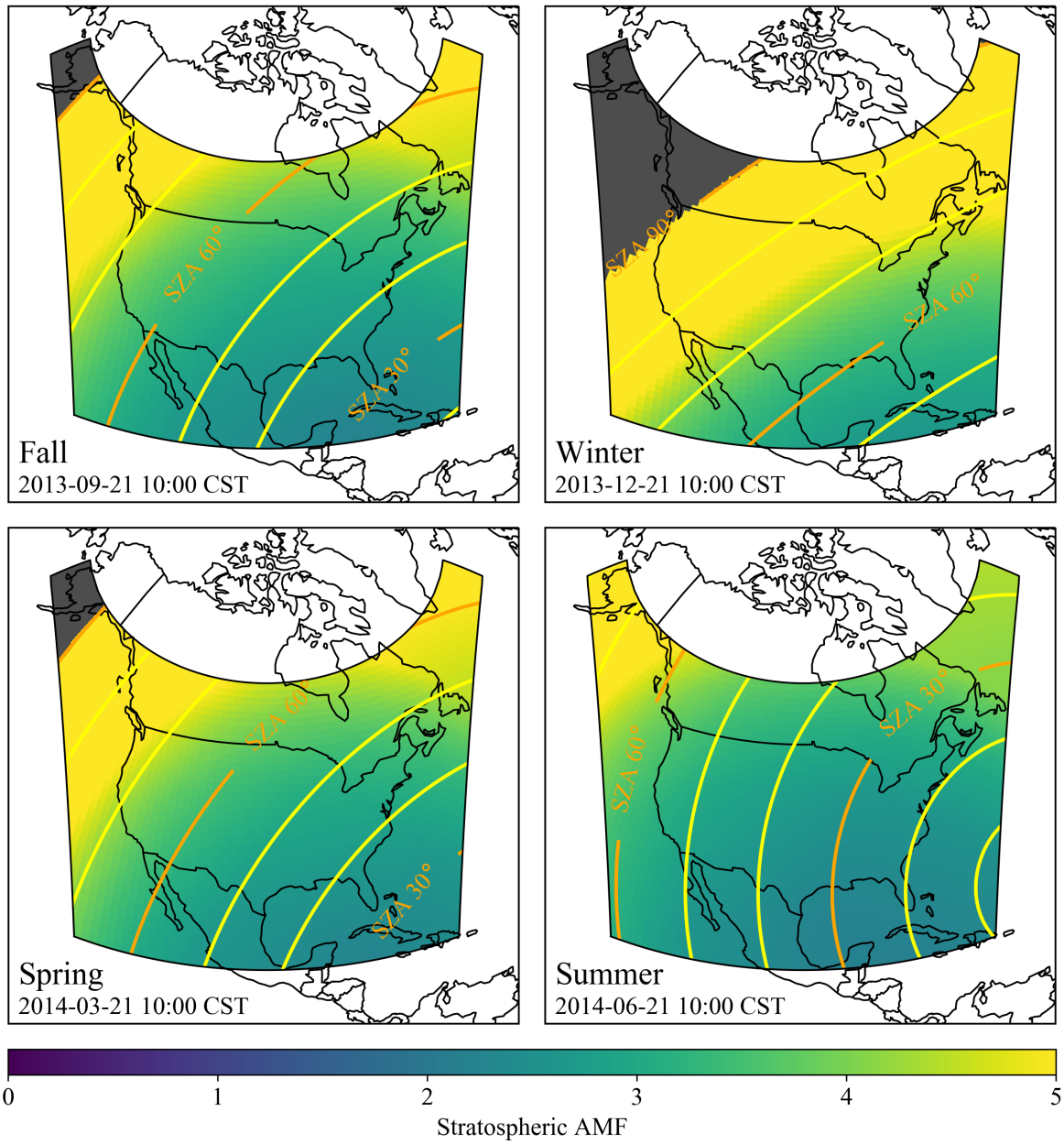


Figure 4.11: Stratospheric AMFs. In contrast to the tropospheric AMF, the stratospheric AMF is slowly varying, is insensitive to surface albedo, and has a strong dependence on SZA and VZA.

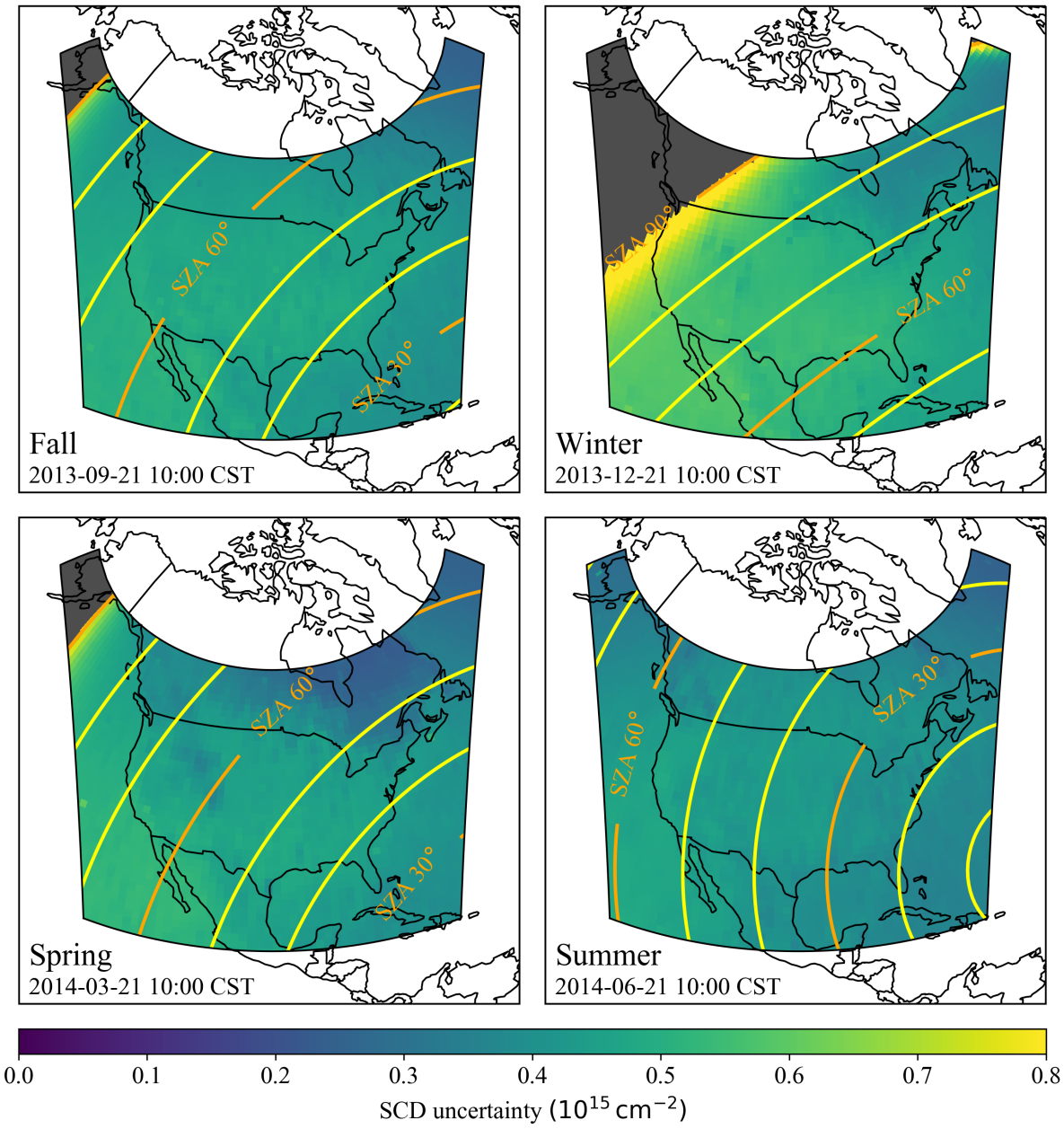


Figure 4.12: NO_2 SCD uncertainty. As expected, the SCD uncertainty has a clear anti-correlation with the SNR in Figure 4.7.

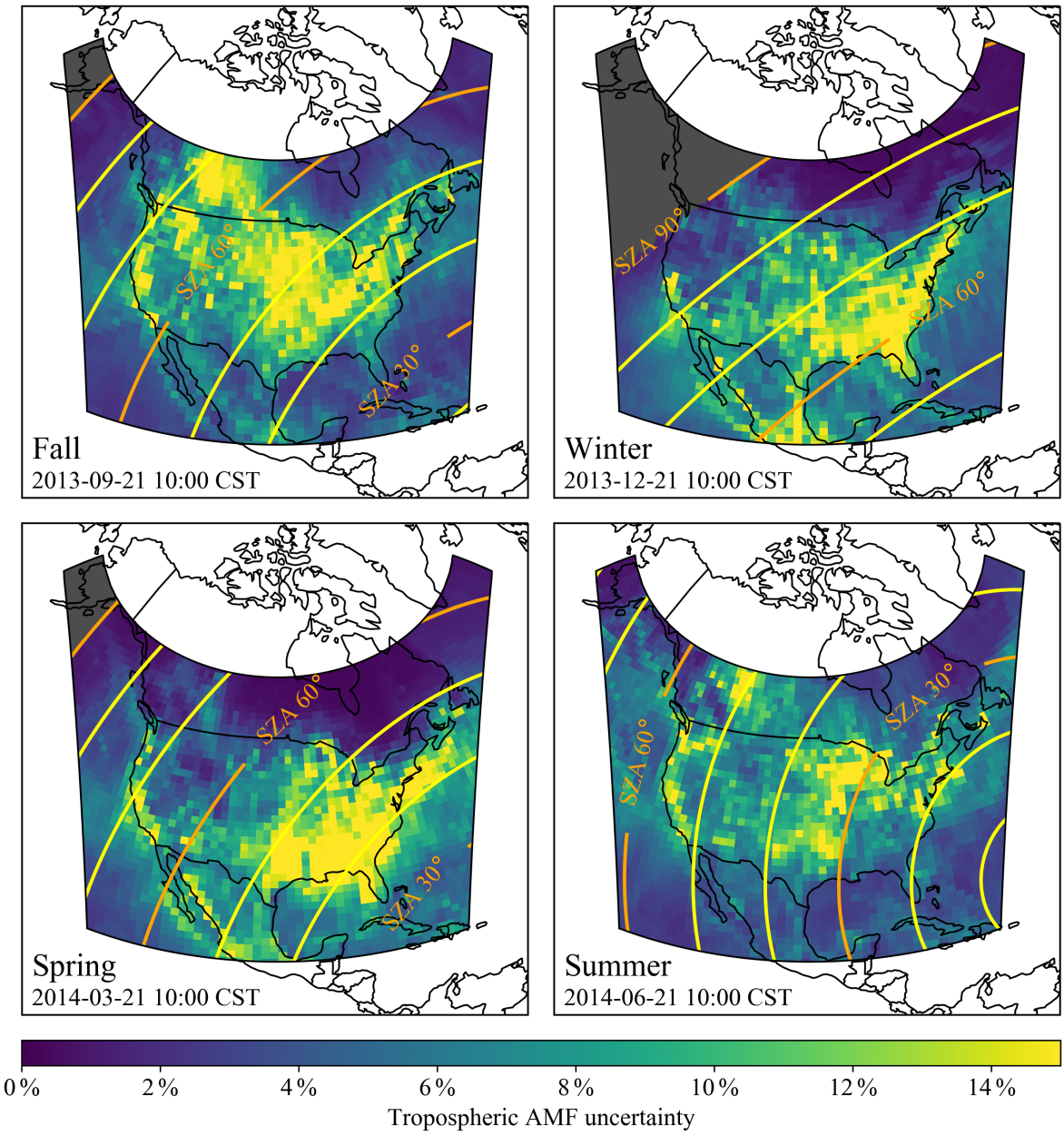


Figure 4.13: Tropospheric AMF uncertainty due to an albedo uncertainty of 0.015. This uncertainty is at a maximum over polluted regions with low albedo.

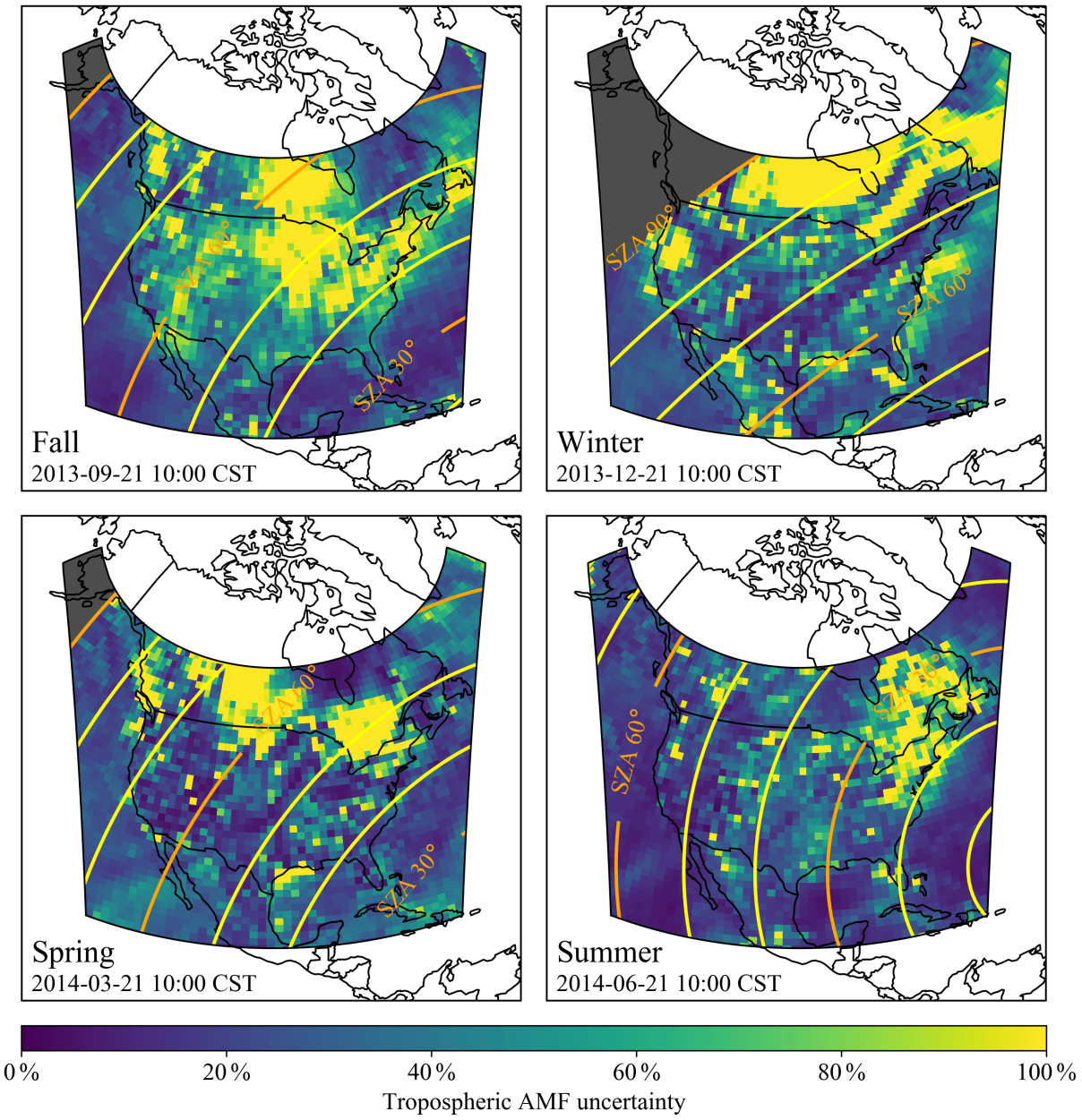


Figure 4.14: Tropospheric AMF uncertainty due to NO_2 profile variability. This uncertainty correlates well with the relative VCD variability in Figure 4.8.

The four uncertainties displayed in these figures are the square roots of the variances found in Equation (4.15), which is a combination of Equations (3.26) and (3.27) describing the tropospheric VCD uncertainty σ_{V_t}. The contribution of the SCD uncertainty σ_S to the tropospheric VCD uncertainty σ_{V_t} is found in the top left of Figures 4.15 through 4.18. The contributions of the stratospheric column uncertainty σ_{V_s} and the stratospheric AMF uncertainty σ_{A_s} are combined and displayed in the top right. The contribution of $\sigma_{A_{a_s}}$, the tropospheric AMF uncertainty due to albedo uncertainty, is found in the bottom left, and the contribution of σ_{A_V}, the tropospheric AMF uncertainty due to NO_2 profile uncertainty, is found on the bottom right.

$$\sigma_{V_t}^2 = \left(\frac{\sigma_S}{A_t}\right)^2 + \left(\frac{A_s\sigma_{V_s}}{A_t}\right)^2 + \left(\frac{V_s\sigma_{A_s}}{A_t}\right)^2 + \left(\frac{V_t\sigma_{A_{a_s}}}{A_t}\right)^2 + \left(\frac{V_t\sigma_{A_V}}{A_t}\right)^2 \tag{4.15}$$

The uncertainty in the tropospheric VCD σ_{V_t} is sensitive to the tropospheric AMF A_t, as the tropospheric AMF represents the amount of measured light that has interacted with the troposphere, and therefore the amount of information about tropospheric constituents that is contained within the measurement. This can be seen in Equation (4.15), where A_t shows up in the denominator of every term. Figure 4.10 shows that the highest values of A_t occur over northern Ontario and Quebec, especially during winter and spring, where snow cover increases surface reflectivity and causes more of the light that reaches the troposphere to be reflected up towards the instrument. The effects are seen in Figures 4.15 through 4.18, where all four uncertainty contributions are reduced over northern Ontario and Quebec.

The SCD contribution shown in Figures 4.15 through 4.18 is relatively constant in comparison to the other contributions, remaining around $3 \times 10^{14}\,\mathrm{cm}^{-2}$ for most scenes. The contribution drops dramatically over northern Ontario and Quebec during spring and winter where there is significant snow cover. This is due to the increased albedo; partially through the increase in tropospheric AMF as discussed above, but also through the decrease in SCD uncertainty (Figure 4.12) that results from the increase in SNR (Figure 4.7). Similarly, in Figure 4.16 it is evident that the SCD contribution becomes more significant when the sun is low on the horizon, which can be explained by the decrease in SNR. In Figure 4.15 it can be seen that the SCD contribution spikes over the heavily polluted oil sands, which at first seems

counterintuitive, as heavy pollution should result in a larger SCD and therefore less relative uncertainty for a given SNR. The cause is low albedo: note in Figure 4.4 that the albedo at the north end of the FOR, including the oil sands, is particularly low. Also note that the corresponding albedo in spring is much higher, and in Figure 4.17 there is no spike in the SCD contribution over the oil sands in spring. Low albedo effectively shields the pollution that resides in the lowermost layers, as most of the light that reaches these layers is absorbed by the ground and therefore does not return. The result is reduced SCDs, or equivalently reduced AMFs, and therefore increased relative SCD uncertainty.

The stratospheric contribution is slightly larger than the SCD contribution for nominal scenes, and increases sharply at high SZAs. Most of the contribution is due to the assumed stratospheric VCD uncertainty $\sigma_{V_s} = 2 \times 10^{14}\,\mathrm{cm}^{-2}$, with very little impact from the assumed stratospheric AMF uncertainty $\sigma_{A_s} = 0.02 A_s$. The constant stratospheric VCD uncertainty has a larger impact on the final uncertainty as the SZA increases, as the measurement becomes more sensitive to the stratosphere due to the long light paths travelling nearly horizontally through the stratosphere. This can be seen in Figure 4.11 where the stratospheric AMF increases sharply at large SZAs: here the information contained in the measurement becomes dominated by the interactions with stratospheric species, reducing sensitivity to the tropospheric species. Mathematically, this dependence can be seen in the second term in Equation (4.15), where the assumed stratospheric VCD uncertainty is scaled by the ratio of the stratospheric and tropospheric AMFs.

The assumed constant albedo uncertainty $\sigma_{a_s} = 0.015$ results in a maximum tropospheric AMF uncertainty on the order of 15 %, as shown in Figure 4.13. However, the resulting contribution to tropospheric VCD uncertainty is only significant over polluted hotspots with low albedo, as seen in Figures 4.15 through 4.18. Tropospheric pollution, especially in the boundary layer near the surface where it is commonly found over emission sources such as large cities, enhances the sensitivity of the retrieval to albedo. As shown in Equation (3.24), the tropospheric AMF is effectively an average of all the tropospheric box-AMFs, weighted by the amount of NO_2 in each respective layer. Therefore when the lowest layers, whose box-AMFs are the most sensitive to albedo, contain high concentrations of NO_2, the

tropospheric AMF becomes more sensitive to the albedo. Although this phenomena only affects a small percentage of the total scenes, it is significant because these heavily polluted scenes are typically the most important scenes to monitor. It is also worth noting that the current analysis shows that the albedo contribution is negligible over unpolluted snowy scenes, but the assumption of a constant albedo uncertainty does not capture the risk of snow misidentification, which will need to be addressed properly in future iterations.

The NO_2 profile contribution is the most significant, due primarily to the large NO_2 profile variability found in the monthly climatology constructed from the GEOS-5 simulation (see Figure 4.8). As seen in Figure 4.14, the tropospheric AMF uncertainty due to NO_2 profile uncertainty exceeds 100 % in some regions. The contribution is augmented by high tropospheric pollution and low tropospheric AMFs, as seen in the fifth term of Equation (4.15). The magnitude of these uncertainties highlights the weakness of the climatology approach to acquiring a priori NO_2 profiles and their uncertainties. The NO_2 profile becomes even more difficult to predict as the spatial resolution of the instrument improves, as pointed out in the documentation for the TROPOMI NO_2 retrieval, where this was stated as the reason for increasing the tropospheric AMF uncertainty from 10 % to 20 % in some of their preliminary work van Geffen et al., 2019. In the present study, the NO_2 climatology was on a 1° latitude by 1.25° longitude grid, which near the center of the FOR of TEMPO equates to an area on the order of 100 km by 100 km, significantly coarser than the 2.1 km by 4.4 km resolution of TEMPO.

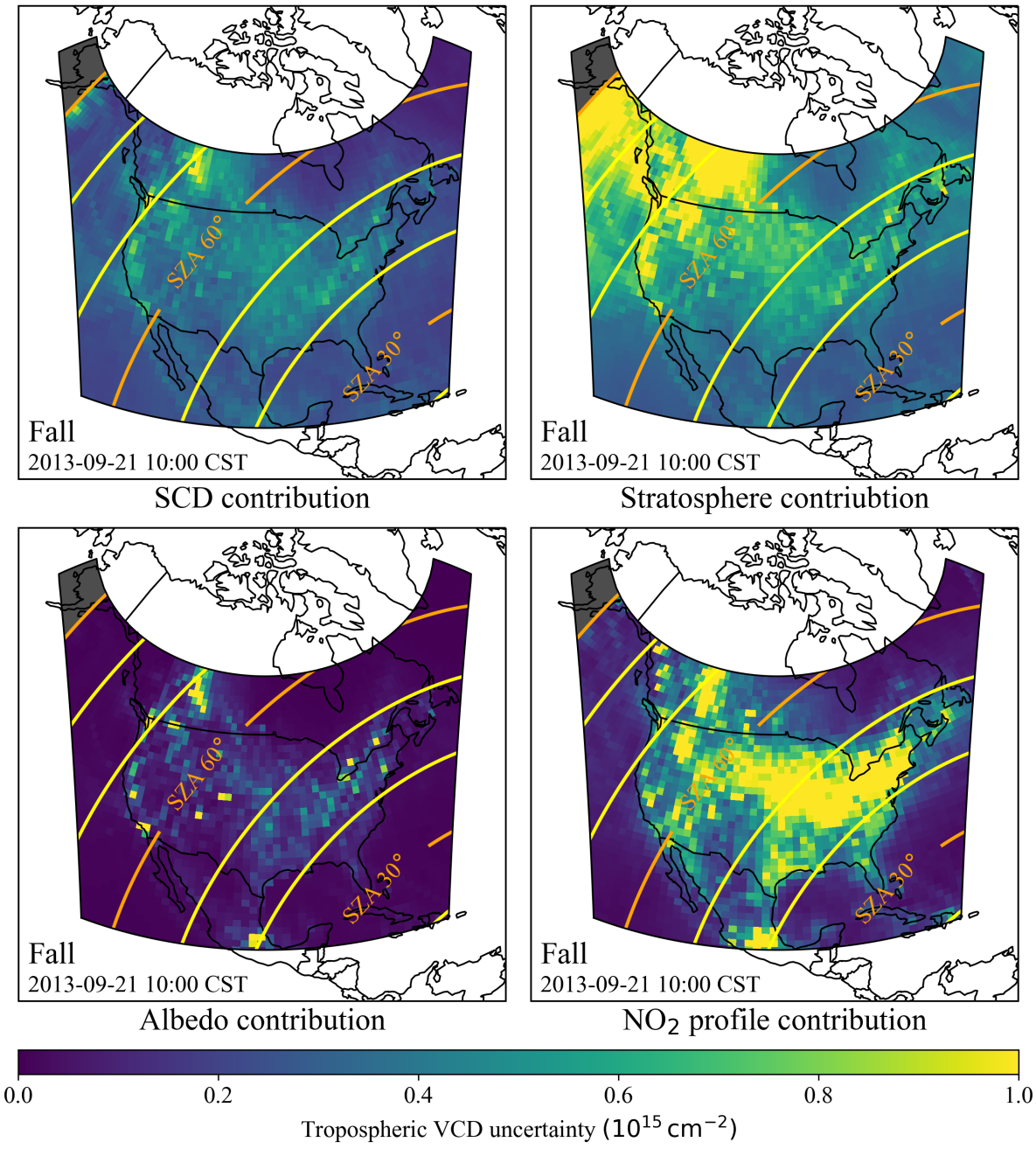


Figure 4.15: Tropospheric VCD retrieval uncertainty summary in fall.

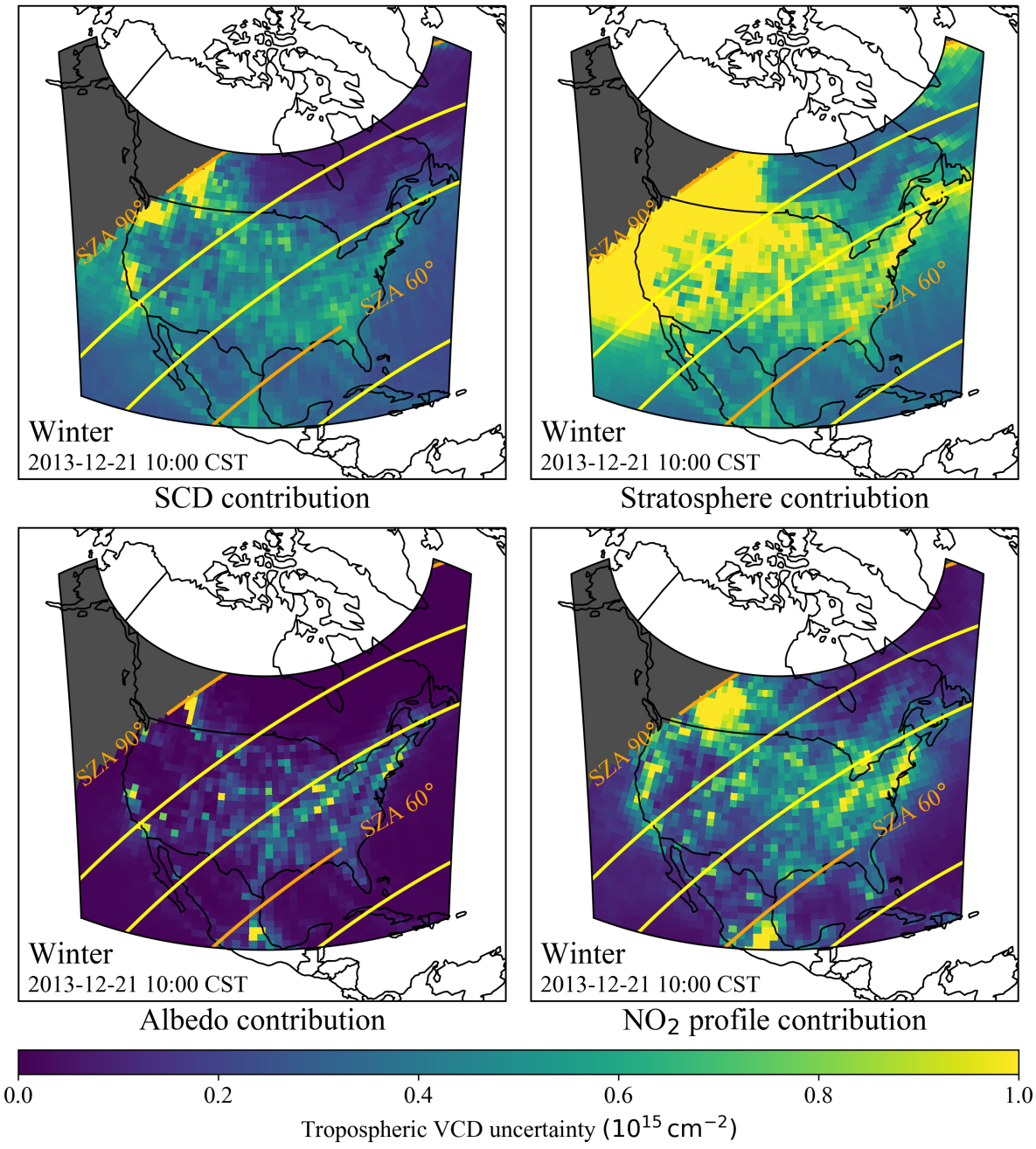


Figure 4.16: Tropospheric VCD retrieval uncertainty summary in winter.

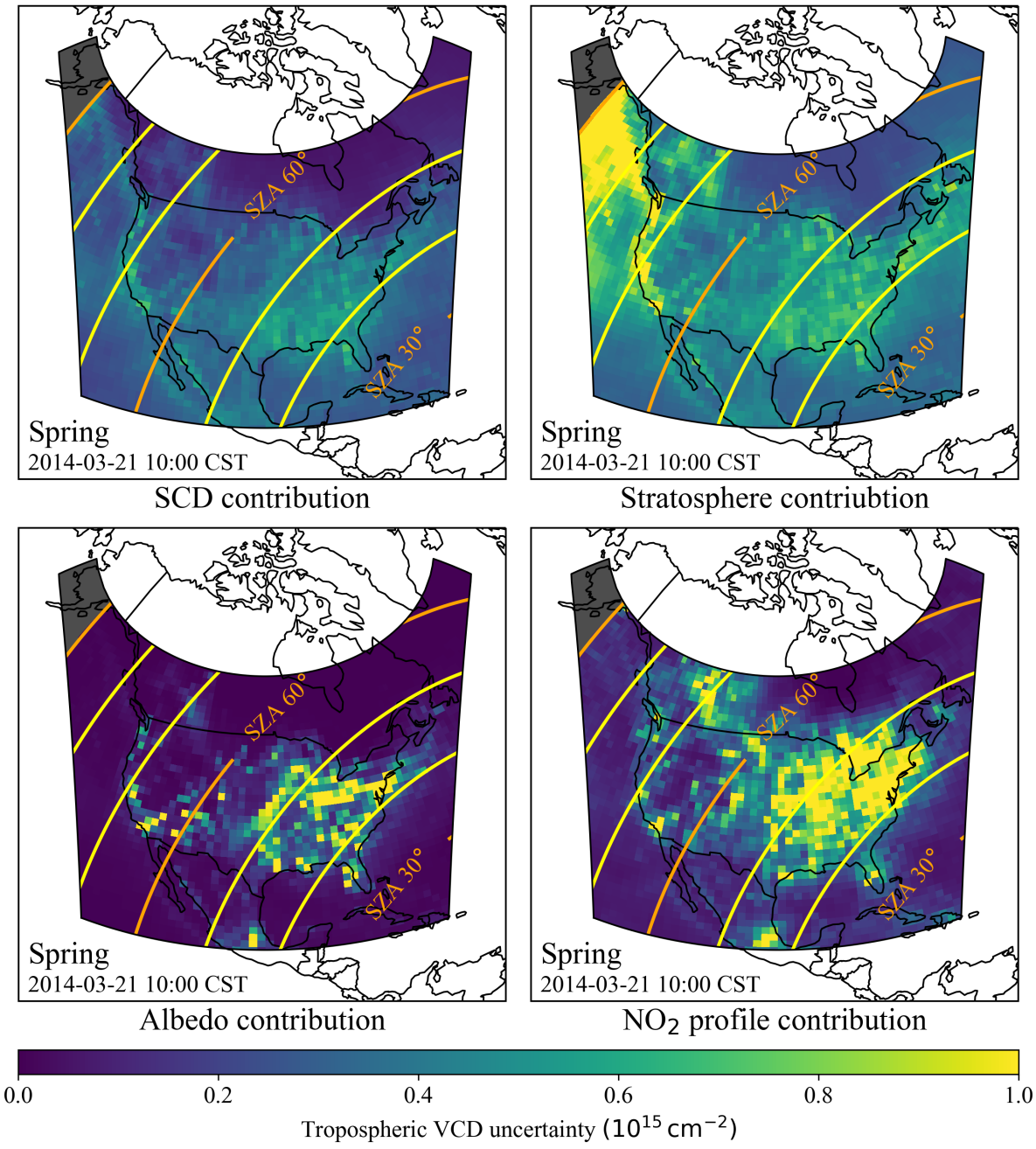


Figure 4.17: Tropospheric VCD retrieval uncertainty summary in spring.

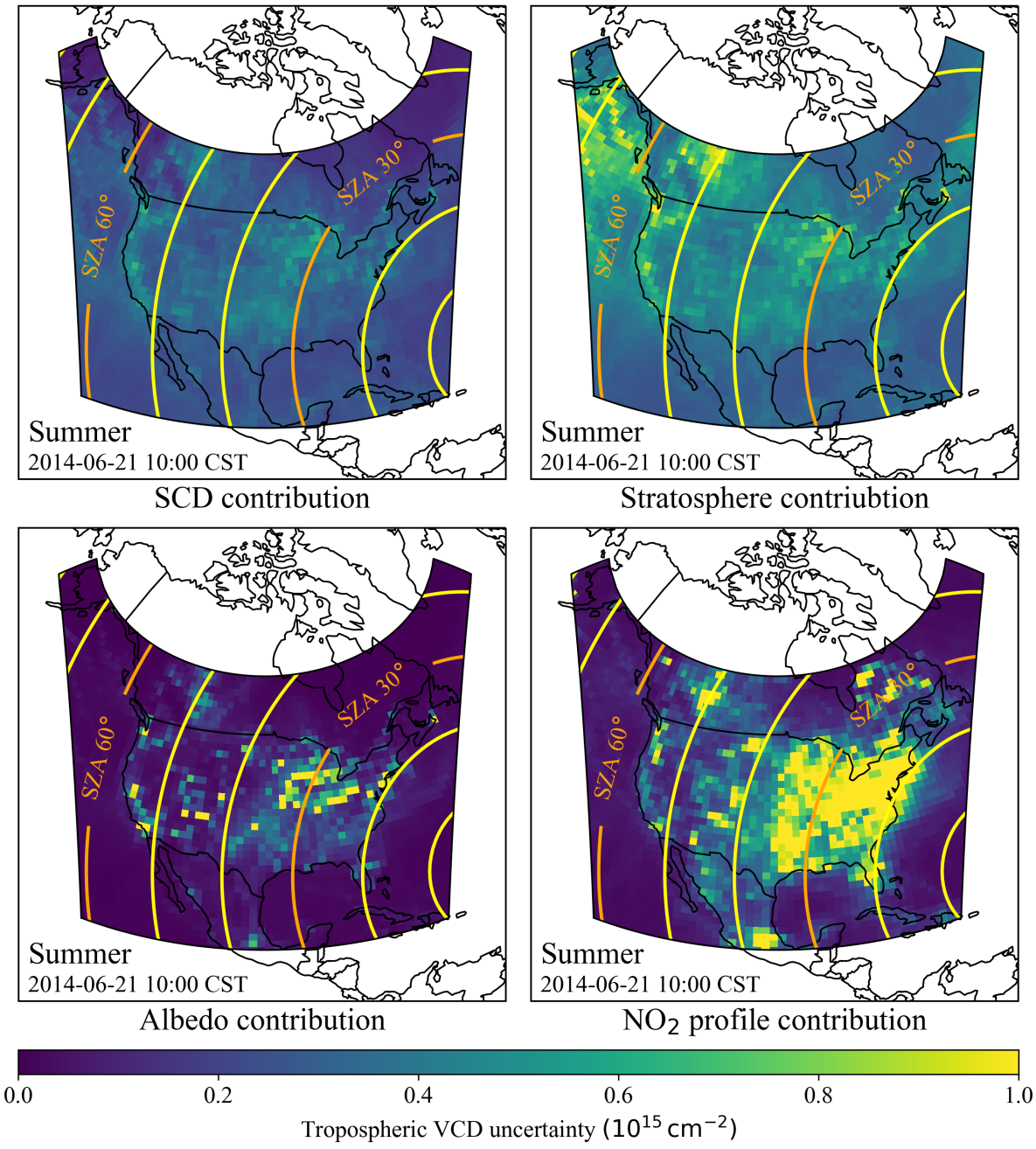


Figure 4.18: Tropospheric VCD retrieval uncertainty summary in summer.

5 Conclusion

The ultimate goal of the work done for this thesis is the creation of retrieval algorithms for TEMPO measurements that are optimized for use over Canada. To this end, fundamental tools for DOAS-style retrievals have been researched, implemented, and tested within the SASKTRAN radiative transfer framework. These tools are based on cutting edge algorithms used by current missions, and will serve as a baseline for comparison with optimizations and alternative techniques that will be experimented with in future work. The currently implemented tools can calculate box-AMF profiles using three different techniques which are all in agreement, and can perform retrievals that are based on OMI algorithms for NO_2 (Bucsela et al., 2013) and ozone (Veefkind et al., 2006). The box-AMF tools have been used to create AMF tables in collaboration with ECCC, and the retrieval tools have been used in a sensitivity study that explores the sources of uncertainty in NO_2 retrievals.

Box-AMF profiles can now be calculated using all three engines within SASKTRAN: the HR engine using the successive orders method, the DO engine using the discrete ordinates method, and the MC engine using the Monte Carlo method. The HR box-AMFs are calculated via weighting functions through a finite-difference scheme in which the absorber profile is perturbed within each layer individually, so that a full radiance calculation is required for every box-AMF layer. The DO box-AMFs are also calculated via weighting functions, but the existing built-in weighting function calculation was found to be sufficiently accurate so that a finite-difference scheme was not required. The MC box-AMFs required a completely different approach: rather than calculating finite-difference weighting functions, which would be very inefficient when combined with a non-deterministic approach such as Monte Carlo, box-AMFs are calculated directly by keeping track of the slant path within each layer alongside the usual radiance calculation. These additions had to be made within the core C++

engine, in contrast to the other implementations which were done in python via the standard interface. All three methods were found to be in good agreement: for moderate zenith angles, where the plane-parallel assumption made by the DO engine has limited effect, HR and DO agree, and they begin to diverge as the zenith angles grow, as expected. Agreement between HR and MC was observed as long as the various discretizations within HR were set with sufficient resolution. Each method has a role: the fastest method, DO, can be used for first order work or for scenes with small zenith angles, HR should be used for the remaining cases, and the MC method, which is the slowest, should be used to ensure that the HR settings are optimized. The box-AMF tools had their first significant usage and validation when a box-AMF table was provided to ECCC for use in their TROPOMI validation campaign over the Alberta oil sands (Griffin et al., 2018).

The purpose of the sensitivity study was to begin to quantify the typical distribution of retrieval uncertainty during realistic conditions. Cloud-free simulations and retrievals were performed at a reduced spatial resolution across the entire FOR of TEMPO, on four days near the four seasonal solstices and equinoxes, and on two hour intervals throughout each day. Four sources of uncertainty were propagated through the retrieval in order to understand their effect on the final retrieved tropospheric NO_2 VCD: uncertainty in the retrieved SCD, in the assumed stratospheric properties, in the assumed surface albedo, and in the assumed NO_2 profile. It was found that SCD uncertainty contributes only moderately to the final uncertainty, illustrating that the instrument design and the quality of the spectrographic measurements is not the bottleneck of the current retrieval strategy. The contribution of the stratospheric uncertainty was also found to be moderate for most scenes, but it grew significantly as the sun became low on the horizon, highlighting the decreased tropospheric sensitivity as the sun sets but also perhaps highlighting the inadequacy of the assumption of a flat stratospheric column uncertainty under all conditions. Another key observation is that the combination of low albedo and high tropospheric pollution in certain regions, such as the oil sands or large cities, results in low tropospheric AMFs, or in other words it greatly reduces the sensitivity of the retrieval to the troposphere, which amplifies the effects of all sources of uncertainty. The most significant source of uncertainty in this analysis came from the

NO_2 profile variability, which was estimated by calculating monthly means and covariances from a GEOS-5 tropospheric chemistry simulation. The large uncertainties indicate that this climatological approach to acquiring a priori NO_2 profiles is potentially insufficient, a problem that is only amplified by the high spatial resolution of TEMPO. This motivates the need for some information about the vertical distribution of NO_2, sparse as it may be in a nadir configuration, to be incorporated into the retrieval in order to reduce dependency on a priori assumptions.

Future work will continue towards the same objective of developing optimized retrieval algorithms for TEMPO measurements over Canada. Some potential directions include expansion to other retrievals such as SO_2 columns and vertical ozone profiles, investigation of the reflectivity of snow and how explicit snow information can be incorporated into retrievals, and experimentation with simultaneous two-dimensional retrievals that leverage the unique configuration of measurements over Canada. Accurate representation of snow reflectivity will be one of the largest challenges to address. While SASKTRAN is already equipped with a BRDF for pure snow developed by Kokhanovsky and Breon (2012), the real challenges are related to accurate snow identification, especially during melting periods, and how snow reflectivity interacts with physical geography. Two dimensional retrievals have the potential to be more robust than the traditional one dimensional approach, as they could extract some vertical information from the atmosphere and reduce dependency on external information as discussed above. Over Canada, they have a higher chance of success, as lines of sight and incident solar rays that travel through the atmosphere closer to horizontal have the potential to contain more vertical information, and the northern latitudes of Canada induce such conditions. The feasibility of incorporating explicit snow information or performing two dimensional retrievals is not yet clear, but both approaches have the potential to help Canada extract as much useful information as possible out of TEMPO measurements.

References

Adams, C., Normand, E. N., McLinden, C. A., Bourassa, A. E., Lloyd, N. D., Degenstein, D. A., Krotkov, N. A., Rivas, M. B., Boersma, K. F., & Eskes, H. (2016). Limb–nadir matching using non-coincident NO_2 observations: Proof of concept and the OMI-minus-OSIRIS prototype product. *Atmospheric Measurement Techniques*, *9*, 4103–4122. doi:10.5194/amt-9-4103-2016

Atkinson, J. (2018). NASA's tempo air pollution space sensor completed. Retrieved October 4, 2019, from https://www.nasa.gov/press-release/langley/nasa-s-tempo-air-pollution-space-sensor-completed

Beirle, S., Kuhl, S., Pukite, J., & Wagner, T. (2010). Retrieval of tropospheric column densitites of NO_2 from combined SCIAMACHY nadir/limb measurements. *Atmospheric Measurement Techniques*, *3*, 283–299. doi:10.5194/amt-3-283-2010

Boersma, K. F., Eskes, H. J., & Brinksma, E. J. (2004). Error analysis for tropospheric NO_2 retrieval from space. *Journal of Geophysical Research*, *109*. doi:10.1029/2003JD003962

Bourassa, A. E., Degenstein, D. A., & Llewellyn, E. J. (2008). SASKTRAN: A spherical geometry radiative transfer code for efficient estimation of limb scattered sunlight. *Journal of Quantitative Spectroscopy & Radiative Transfer*, *109*, 52–73. doi:10.1016/j.jqsrt.2007.07.007

Bovensmann, H., Burrows, J. P., Buchwitz, M., Frerick, J., Noël, S., & Rozanov, V. V. (1999). SCIAMACHY: Mission objectives and measurement modes. *Journal of the Atmospheric Sciences*, *56*(2), 127–150. doi:10.1175/1520-0469(1999)056⟨0127:SMOAMM⟩2.0.CO;2

Bucsela, E. J., Celarier, E. A., Wenig, M. O., Gleason, J. F., Veefkind, J. P., Boersma, K. F., & Brinksma, E. J. (2006). Algorithm for NO_2 vertical column retrieval from the Ozone Monitoring Instrument. *IEEE Transactions on Geoscience and Remote Sensing*, *44*(5), 1245–1258. doi:10.1109/TGRS.2005.863715

Bucsela, E. J., Krotkov, N. A., Celarier, E. A., Lamsal, L. N., Swartz, W. H., Bhartia, P. K., Boersma, K. F., Veefkind, J. P., Gleason, J. F., & Pickering, K. E. (2013). A new stratospheric and tropospheric NO_2 retrieval algorithm for nadir-viewing satellite instruments: Applications to OMI. *Atmospheric Measurement Techniques*, *6*, 2607–2626. doi:10.5194/amt-6-2607-2013

Burrows, J. P., Weber, M., Buchwitz, M., Rozanov, V., Ladstatter-Weibenmayer, A., Richter, A., DeBeek, R., Hoogen, R., Bramstedt, K., Eichmann, K.-U., & Eisinger, M. (1999). The Global Ozone Monitoring Experiment (GOME): Mission concept and first scientific results. *Journal of the Atmospheric Sciences*, *56*(2), 151–175. doi:10.1175/1520-0469(1999)056⟨0151:TGOMEG⟩2.0.CO;2

Chance, K., & Martin, R. V. (2017). *Spectroscopy and radiaitve transfer of planetary atmospheres*. Oxford, United Kingdom: Oxford University Press.

Chance, K., & Spurr, R. J. D. (1997). Ring effect studies: Rayleigh scattering, including molecular parameters for rotational Raman scattering, and the Fraunhofer spectrum. *Applied Optics*, *36*(21), 5224–5230. doi:10.1364/AO.36.005224

Chandrasekhar, S. (1950). *Radiative transfer*. Oxford, United Kingdom: Oxford University Press.

Cohen, A. J., Brauer, M., Burnett, R., Anderson, H. R., Frostad, J., Estep, K., Balakrishnan, K., Brunekreef, B., Dandona, L., Dandona, R., Feigin, V., Freedman, G., Hubbell, B., Jobling, A., Kan, H., Knibbs, L., Liu, Y., Martin, R., Morawska, L., III, C. A. P., Shin, H., Straif, K., Shaddick, G., Thomas, M., van Dingenen, R., van Donkelaar, A., Vos, T., Murray, C. J. L., & Forouzanfar, M. H. (2017). Estimates and 25-year trends of the global burden of disease attributable to ambient air pollution: An analysis of data from the Global Burden of Diseases Study 2015. *The Lancet*, *389*, 1907–1918. doi:10.1016/S0140-6736(17)30505-6

Dobson, G. M. B., & Harrison, D. N. (1926). Measurements of the amount of ozone in the earth's atmosphere and its relation to other geophysical conditions. *Proceedings of the Royal Society A*, *110*(756), 660–693. doi:10.1098/rspa.1926.0040

Dueck, S. R. F., Bourassa, A. E., & Degenstein, D. A. (2017). An efficient algorithm for polarization in the SASKTRAN radiative transfer framework. *Journal of Quantitative Spectroscopy & Radiative Transfer*, *199*, 1–11. doi:10.1016/j.jqsrt.2017.05.016

Fabry, C., & Buisson, H. (1921). A study of the ultra-violet end of the solar spectrum. *The Astrophysical Journal*, *54*(5), 297–323. doi:10.1086/142647

Farman, J. C., Gardiner, B. G., & Shanklin, J. D. (1985). Large losses of total ozone in Antarctica reveal seasonal ClO_x/NO_x interaction. *Nature*, *315*, 207–210. doi:10.1038/315207a0

Flynn, L. E., Seftor, C. J., Larsen, J. C., & Xu, P. (2006). The Ozone Mapping and Profiler Suite. In J. J. Qu, W. Gao, M. Kafatos, R. E. Murphy, & V. V. Salomonson (Eds.), *Earth science satellite remote sensing* (pp. 279–296). Beijing, China: Tsinghua University Press.

Grainger, J. F., & Ring, J. (1962). Anomalous Fraunhofer line profiles. *Nature*, *193*, 762–762. doi:10.1038/193762a0

Griffin, D., Zhao, X., McLinden, C. A., Boersma, F., Bourassa, A., Dammers, E., Degenstein, D., Eskes, H., Fehr, L., Fioletov, V., Hayden, K., Kharol, S. K., Li, S.-M., Makar, P., Martin, R. V., Mihele, C., Mittermeier, R. L., Krotkov, N., Sneep, M., Lamsal, L. N., ter Linden, M., van Geffen, J., Veefkind, P., & Wolde, M. (2018). High-resolution mapping of nitrogen dioxide with TROPOMI: First results and validation over the Canadian oil sands. *Geophysical Research Letters*, *46*(2), 1049–1060. doi:10.1029/2018GL081095

Gulde, S. T., Kolm, M. G., Smith, D. J., Maurer, R., Courrèges-Lacoste, G. B., Sallusti, M., & Bagnasco, G. (2014). Sentinel 4: A geostationary imaging UVN spectrometer for air quality monitoring: Status of design, performance and development. In Z. Sodnik, B. Cugny, & N. Karafolas (Eds.), *International conference on space optics - ICSO 2014*, Tenerife, Canary Islands, Spain.

Han, D., Chen, L., Su, L., Tao, J., Li, S., Yu, C., Zhang, Y., & He, B. (2011). A convolution algorithm to calculate differential cross sections of the Ring effect in the earth's atmosphere based on rotational Raman scattering. *Science China Earth Sciences*, *54*(9), 1407–1412. doi:10.1007/s11430-011-4226-x

Hedin, A. E. (1991). Extension of the MSIS thermosphere model into the middle and lower atmosphere. *Journal of Geophysical Research*, *96*(A2), 1159–1172. doi:10.1029/90JA02125

Hu, L., Keller, C. A., Long, M. S., Sherwen, T., Auer, B., Da Silva, A., Nielsen, J. E., Pawson, S., Thompson, M. A., Trayanov, A. L., Travis, K. R., Grange, S. K., Evans, M. J., & Jacob, D. J. (2018). Global simulation of tropospheric chemistry at 12.5 km resolution: Performance and evaluation of the GEOS-Chem chemical module (v10-1) within the NASA GEOS Earth system model (GEOS-5 ESM). *Geoscientific Model Development*, *11*, 4603–4620. doi:10.5194/gmd-11-4603-2018

Kaplan, L. D. (1959). Inference of atmospheric structure from remote radiation measurements. *Journal of the Optical Society of America*, *49*(10), 1004–1007. doi:10.1364/JOSA.49.001004

Kim, J., Jeong, U., Ahn, M.-H., Kim, J. H., Park, R. J., Lee, H., Song, C. H., Choi, Y.-S., Lee, K.-H., Yoo, J.-M., Jeong, M.-J., Park, S. K., Lee, K.-M., Song, C.-K., Kim, S.-W., Kim, Y., Kim, S.-W., Kim, M., Go, S., Liu, X., Chance, K., Miller, C. C., Al-Saadi, J., Veihelmann, B., Bhartia, P. K., Torres, O., Abad, G. G., Haffner, D. P., Ko, D. H., ... Choi, Y. (2019). New era of air quality monitoring from space: Geostationary Environment Monitoring Spectrometer (GEMS). *Bulletin of the American Meteorogical Society*. doi:10.1175/BAMS-D-18-0013.1

Koelemeijer, R. B. A., de Haan, R. F., & Stammes, P. (2003). A database of spectral surface reflectivity in the range 335-772 nm derived from 5.5 years of GOME observations. *Journal of Geophysical Research*, *108*. doi:10.1029/2002JD002429

Kokhanovsky, A. A., & Breon, F.-M. (2012). Validation of an analytical snow brdf model using parasol multi-angular and multispectral observations. *IEEE Geoscience and Remote Sensing Letters*, *9*(5), 928–932. doi:10.1109/LGRS.2012.2185775

Levelt, P. F., van den Oord, G. H. J., Dobber, M. R., Malkki, A., Visser, H., de Vries, J., Stammes, P., Lundell, J. O. V., & Saari, H. (2006). The Ozone Monitoring Instrument. *IEEE Transactions on Geoscience and Remote Sensing*, *44*(5), 1093–1101. doi:10.1109/TGRS.2006.872333

Levenberg, K. (1944). A method for the solution of certain non-linear problems in least squares. *Quarterly of Applied Mathematics*, *2*, 164–168. doi:10.1090/qam/10666

Llewellyn, E. J., Lloyd, N. D., Degenstein, D. A., Gattinger, R. L., Petelina, S. V., Bourassa, A. E., Wiensz, J. T., Ivanov, E. V., McDade, I. C., Solheim, B. H., McConnell, J. C., Haley, C. S., von Savigny, C., Sioris, C. E., McLinden, C. A., Griffioen, E., Kaminski, J., Evans, W. F. J., Puckrin, E., Strong, K., Wehrle, V., Hum, R. H., Kendall, D. J. W., Matsushita, J., Murtagh, D. P., Brohede, S., Stegman, J., Witt, G., Barnes, G., ... Nordh, L. (2004). The OSIRIS instrument on the Odin spacecraft. *Canadian Journal of Physics*, *82*, 411–422. doi:10.1139/P04-005

Marchenko, S., Krotkov, N. A., Lamsal, L. N., Celarier, E. A., Swartz, W. H., & Bucsela, E. J. (2015). Revising the slant column density retrieval of nitrogen dioxide observed by the Ozone Monitoring Instrument. *Journal of Geophysical Research: Atmospheres*, *120*, 5670–5692. doi:10.1002/2014JD022913

Marquardt, D. W. (1963). An algorithm for least-squares estimation of non-linear parameters. *Journal of the Society for Industrial and Applied Mathematics*, *11*(2), 431–441. doi:10.1137/0111030

Martin, R. V., Chance, K., Jacob, D. J., Kurosu, T. P., Spurr, R. J. D., Bucsela, E., Gleason, J. F., Palmer, P. I., Bey, I., Fiore, A. M., Li, Q., Yantosca, R. M., & Koelemeijer, R. B. A. (2002). An improved retrieval of tropospheric nitrogen dioxide from GOME. *Journal of Geophysical Research*, *107*. doi:10.1029/2001JD001027

McPeters, R. D., Labow, G. J., & Logan, J. A. (2007). Ozone climatological profiles for satellite retrieval algorithms. *Journal of Geophysical Research: Atmospheres*, *112*(D5). doi:10.1029/2005JD006823

Molina, M. J., & Rowland, F. S. (1974). Stratospheric sink for chlorofluormethanes: Chlorine atom-catalysed destruction of ozone. *Nature*, *249*, 810–812. doi:10.1038/249810a0

More, J. J. (1978). The levenberg-marquardt algorithm: Implementation and theory. In G. A. Watson (Ed.), *Numerical analysis* (pp. 105–116). Heidelberg, Germany: Springer Berlin Heidelberg.

Munro, R., Lang, R., Klaes, D., Poli, G., Retscher, C., Lindstrot, R., Huckle, R., Lacan, A., Grzegorski, M., Holdak, A., Kokhanovsky, A., Livschitz, J., & Eisinger, M. (2016). The GOME-2 instrument on the Metop series of satellites: Instrument design, calibration, and level 1 data processing - an overview. *Atmospheric Measurement Techniques*, *9*, 1279–1301. doi:10.5194/amt-9-1279-2016

Nicks, D. (2014). Tempo instrument update. Presentation slides from the 2014 TEMPO Science Team Meeting, National Institute of Aerospace, Hampton, VA. Retrieved May 15, 2019, from http://tempo.si.edu/presentations.html

Nicks, D. (2016). Tempo instrument update. Presentation slides from the 2016 TEMPO Science Team Meeting, National Museum of the American Indian, Washington, DC. Retrieved May 15, 2019, from http://tempo.si.edu/presentations.html

Northon, K. (2019). Commercial space ride secured for NASA's new air pollution sensor. Retrieved October 4, 2019, from https://www.nasa.gov/press-release/commercial-space-ride-secured-for-nasa-s-new-air-pollution-sensor

Platt, U., & Stutz, J. (2008). *Differential optical absorption spectroscopy principles and applications*. Heidelberg, Germany: Springer-Verlag.

Rozanov, V. V., & Rozanov, A. V. (2010). Differential optical absorption spectroscopy (DOAS) and air mass factor concept for a multiply scattering vertically inhomogeneous medium: Theoretical considerations. *Atmospheric Measurement Techniques*, *3*, 751–780. doi:10.5194/amt-3-751-2010

Schuster, A. (1905). Radiation through a foggy atmosphere. *The Astrophysical Journal*, *21*(1), 1–22. doi:10.1086/141186

Spurr, R. J. D., Kurosu, T. P., & Chance, K. (2001). A linearized discrete ordinate radiative transfer model for atmospheric remote-sensing retrieval. *Journal of Quantitative Spectroscopy & Radiative Transfer*, *68*, 689–735. doi:10.1016/S0022-4073(00)00055-8

van Geffen, J. H. G. M., Eskes, H. J., Boersma, K. F., Maasakkers, J. D., & Veefkind, J. P. (2019). TROPOMI ATBD of the total and tropospheric NO_2 data products.

Veefkind, J. P., Aben, L., McMullan, K., Forster, H., de Vries, J., Otter, G., Claas, J., Eskes, H. J., de Haan, J. F., Kleipool, Q., van Weele, M., Hasekamp, O., Hoogeveen, R., Landgraf, J., Snel, R., Tol, P., Ingmann, P., Voors, R., Kruizinga, B., Vink, R., Visser, H., & Levelt, P. F. (2012). TROPOMI on the ESA Sentinel-5 precursor: A GMES mission for global observations of the atmospheric composition for climate,

air quality and ozone layer applications. *Remote Sensing of Environment*, *120*, 70–83. doi:10.1016/j.rse.2011.09.027

Veefkind, J. P., de Haan, J. F., Brinksma, E. J., Kroon, M., & Levelt, P. F. (2006). Total ozone from the Ozone Monitoring Instrument (OMI) using the DOAS technique. *IEEE Transactions on Geoscience and Remote Sensing*, *44*(5), 1239–1244. doi:10.1109/TGRS.2006.871204

World Meteorogical Organization. (2018). *Scientific assessment of ozone depletion: 2018* (No. 58). Global Ozone Research and Monitoring Project. Geneva, Switzerland.

Yantosca, B. (2016). GEOS-Chem vertical grids. Retrieved May 1, 2019, from http://wiki.seas.harvard.edu/geos-chem/index.php/GEOS-Chem_vertical_grids

Zawada, D. J., Dueck, S. R., Rieger, L. A., Bourassa, A. E., Lloyd, N. D., & Degenstein, D. A. (2015). High-resolution and Monte Carlo additions to the SASKTRAN radiative transfer model. *Atmospheric Measurement Techniques*, *8*, 2609–2623. doi:10.5194/amt-8-2609-2015

Zoogman, P., Liu, X., Suleiman, R. M., Pennington, W. F., Flittner, D. E., Al-Saadi, J. A., Hilton, B. B., Nicks, D. K., Newchurch, M. J., Carr, J. L., Janz, S. J., Andraschko, M. R., Arola, A., Baker, B. D., Canova, B. P., Miller, C. C., Cohen, R. C., Davis, J. E., Dussault, M. E., Edwards, D. P., Fishman, J., Ghulam, A., Abad, G. G., Grutter, M., Herman, J. R., Houck, J., Jacob, D. J., Joiner, J., Kerridge, B. J., ... Chance, K. (2017). Tropospheric emissions: Monitoring of pollution (TEMPO). *Journal of Quantitative Spectroscopy & Radiative Transfer*, *186*, 17–39. doi:10.1016/j.jqsrt.2016.05.008

Appendix A

Normal Distributions

A.1 Normal Distribution Definition

Univariate

A normal distribution has the following PDF:

$$P(X = x) \propto \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right), \tag{A.1}$$

where X is the normal random variable, μ is its mean, and σ is its standard deviation. The standard normal distribution is the normal distribution with $\mu = 0$ and $\sigma = 1$, so that the PDF becomes

$$P(Z = z) \propto \exp\left(-\frac{z^2}{2}\right). \tag{A.2}$$

These two PDFs are related by the transformation $z = (x-\mu)/\sigma$, so that X can be sampled by first sampling Z and then calculating $X = \mu + \sigma Z$. Samples of these distributions can be characterized by how many standard deviations away from the mean they are. The probability of X falling within the one-sigma domain $(\mu-\sigma, \mu+\sigma)$, or equivalently the probability of Z falling within the domain $(-1, 1)$, is constant, approximately 68.4%. Similarly, the probability of falling within the two-sigma domain is approximately 95.5%, and within the three-sigma domain is approximately 99.7%.

Multivariate

An N-dimensional multivariate normal distribution has the following PDF:

$$P(\boldsymbol{X} = \boldsymbol{x}) \propto \exp\left(-\frac{1}{2}(\boldsymbol{x}-\boldsymbol{\mu})^T\Sigma^{-1}(\boldsymbol{x}-\boldsymbol{\mu})\right), \tag{A.3}$$

where $\boldsymbol{X}$ is the normal random variable, $\boldsymbol{\mu}$ is its mean, and Σ is its covariance matrix. Covariance matrices are always positive semidefinite, which means they have non-negative eigenvalues and a corresponding set of orthonormal eigenvectors can always be found. Therefore Σ can be diagonalized so that $\Sigma = SDS^T$ where the matrix D has all non-negative entries

which will be denoted σ_i^2, and the matrix S is unitary so that $S^{-1} = S^T$. The PDF can then be written as

$$P(\boldsymbol{X} = \boldsymbol{x}) \propto \exp\left(-\frac{1}{2}(\boldsymbol{x} - \boldsymbol{\mu})^T S D^{-1/2} D^{-1/2} S^T (\boldsymbol{x} - \boldsymbol{\mu})\right), \tag{A.4}$$

where $D^{-1/2}$ is a diagonal matrix with entries σ_i^{-1}. The substitution $\boldsymbol{z} = D^{-1/2} S^T (\boldsymbol{x} - \boldsymbol{\mu})$ produces a new PDF,

$$P(\boldsymbol{Z} = \boldsymbol{z}) \propto \exp\left(-\frac{1}{2}\boldsymbol{z}^T \boldsymbol{z}\right) = \prod_{i=1}^{N} \exp\left(-\frac{z_i^2}{2}\right), \tag{A.5}$$

which describes the standard normal multivariate distribution, where every element of $\boldsymbol{Z}$ is chosen from an independent standard normal distribution. Therefore $\boldsymbol{X}$ can be sampled by sampling $\boldsymbol{Z}$ and using the transformation $\boldsymbol{X} = \boldsymbol{\mu} + S D^{1/2} \boldsymbol{Z}$. Here the vector $\boldsymbol{Z}$ is stretched by $D^{1/2}$, which stretches it in the direction of the i^{th} axis by a factor of σ_i, and then it is rotated by S. The one-sigma domain of $\boldsymbol{Z}$ is the unit hypersphere, covering all points such that $\boldsymbol{z}^T \boldsymbol{z} < 1$, and this domain is similarly stretched by $D^{1/2}$ and rotated by S in order to produce a hyperellipsoid which defines the one-sigma domain of $\boldsymbol{X}$.

A.2 Propagation of Normal Distributions

Univariate

Consider a normal random variable X with mean μ and variance σ^2, such that $X = \mu + \sigma Z$ where Z is a standard normal random variable. The corresponding value of a derived quantity $Y = f(X)$ can be approximated with a first order Taylor expansion,

$$Y = f(X) = f(\mu + \sigma Z) \approx f(\mu) + \left.\frac{df}{dX}\right|_{X=\mu} \sigma Z\,, \tag{A.6}$$

so that Y is a normal random variable with mean and variance given by

$$\mu_Y = f(\mu) \qquad \sigma_Y^2 = \left(\left.\frac{df}{dX}\right|_{X=\mu}\right)^2 \sigma^2\,. \tag{A.7}$$

Multivariate

Consider the m-dimensional equivalent of this process, where a derived quantity $Y = f(\boldsymbol{X})$ is to be calculated from a multivariate normal random variable $\boldsymbol{X}$ with mean $\boldsymbol{\mu}$ and covariance matrix Σ. As discussed in Appendix A.1, $\boldsymbol{X}$ can be expressed as $\boldsymbol{X} = \boldsymbol{\mu} + S D^{1/2} \boldsymbol{Z}$, where

$\Sigma = SDS^T$ is the diagonalization of Σ and $\boldsymbol{Z}$ is a vector of independent standard normal random variables. The first order Taylor expansion is now

$$Y = f(\boldsymbol{X}) = f(\boldsymbol{\mu} + SD^{1/2}\boldsymbol{Z}) \approx f(\boldsymbol{\mu}) + \boldsymbol{\nabla} f(\boldsymbol{\mu})^T SD^{1/2}\boldsymbol{Z} = f(\boldsymbol{\mu}) + \boldsymbol{a}^T\boldsymbol{Z}\,, \tag{A.8}$$

where $\boldsymbol{a} \equiv D^{1/2}S^T\boldsymbol{\nabla} f(\boldsymbol{\mu})$. The sum of independent normal distributions is also a normal distribution, with variance equal to the sum of the variances of the original distributions. The dot product $\boldsymbol{a}^T\boldsymbol{Z}$ is the sum of m normal distributions with variances a_i^2, and is therefore equal to a single normal distribution with variance $|\boldsymbol{a}|^2$. This variance can be expressed in terms of the original covariance matrix Σ:

$$\sigma_Y^2 = |\boldsymbol{a}|^2 = \boldsymbol{a}^T\boldsymbol{a} = \boldsymbol{\nabla} f(\boldsymbol{\mu})^T SD^{1/2}D^{1/2}S^T\boldsymbol{\nabla} f(\boldsymbol{\mu}) = \boldsymbol{\nabla} f(\boldsymbol{\mu})^T\Sigma\boldsymbol{\nabla} f(\boldsymbol{\mu})\,. \tag{A.9}$$

Note that measurement errors are commonly interpreted as the standard deviation of a normal distribution, in which case this process describes the error propagation process for quantities derived from measurements. In a typical scenario, the measurements are independent, in which case Σ is diagonal and Equation (A.9) reduces to

$$\sigma_Y^2 = \sum_{i=1}^{m}\left(\frac{\partial f}{\partial X_i}\bigg|_{\boldsymbol{X}=\boldsymbol{\mu}}\sigma_i\right)^2\,, \tag{A.10}$$

where σ_i is the error corresponding to measurement X_i.

A.3 Ratio of Normal Distributions

A useful application of Equation (A.9) is the approximate description of the ratio of two normal distributions. The ratio of two normal distributions is not normal, although it is approximately normal if the distribution in the denominator has a standard deviation that is much smaller than its mean. Consider two dependent normal random variables X and Y with variances $\mathrm{Var}(X)$ and $\mathrm{Var}(Y)$ and with covariance $\mathrm{Cov}(X,Y)$, and also their ratio $R = X/Y$. Evaluating Equation (A.9) gives the result

$$\begin{aligned}\mathrm{Var}(R) &= \begin{pmatrix}\frac{1}{Y} & -\frac{X}{Y^2}\end{pmatrix}\begin{pmatrix}\mathrm{Var}(X) & \mathrm{Cov}(X,Y)\\ \mathrm{Cov}(X,Y) & \mathrm{Var}(Y)\end{pmatrix}\begin{pmatrix}\frac{1}{Y}\\ -\frac{X}{Y^2}\end{pmatrix}\\ &= \frac{X^2}{Y^2}\left(\frac{\mathrm{Var}(X)}{Y^2} - 2\frac{\mathrm{Cov}(X,Y)}{XY} + \frac{\mathrm{Var}(Y)}{Y^2}\right)\,.\end{aligned} \tag{A.11}$$

Appendix B

Monte Carlo Integration

B.1 Monte Carlo Integration

Monte Carlo integration is a technique which allows definite integrals to be estimated using a non-deterministic approach. Consider a continuous random variable $\boldsymbol{X}$ with PDF $p(\boldsymbol{x})$ on the domain D with measure V. Equation (B.1) states that the integral of the product of a function $f(\boldsymbol{x})$ and PDF $p(\boldsymbol{x})$ is equal to the expectation value of $f(\boldsymbol{X})$, denoted by $\langle f(\boldsymbol{x})\rangle_{\boldsymbol{x}\sim p(\boldsymbol{x})}$, multiplied by the measure of the domain.

$$I = \int_D p(\boldsymbol{x}) f(\boldsymbol{x}) d\boldsymbol{x} = V \langle f(\boldsymbol{x})\rangle_{\boldsymbol{x}\sim p(\boldsymbol{x})} \tag{B.1}$$

To estimate the value of I and its uncertainty σ_I, first take n samples $\boldsymbol{X_i}$ of $\boldsymbol{X}$ using inverse transform sampling as described in Appendix B.2, then evaluate $f_i = f(\boldsymbol{X_i})$ for each sample, and finally use f_i to estimate the mean and uncertainty as described in Appendix B.3.

B.2 Inverse Transform Sampling

Univariate

Inverse transform sampling is a method that allows an arbitrary PDF to be sampled using a random number generator. Consider a continuous random variable X with PDF $p(x)$ and cumulative distribution function (CDF) $P(x) = \int_{-\infty}^{x} p(x')dx'$. The probability integral transform states that the random variable $Y = P(X)$ has a uniform distribution, or in other words, a random variable X can be mapped to a uniform distribution using its CDF.

The inverse of this statement is necessarily true as well: a uniform random variable can be mapped to any random variable using the inverse CDF. Therefore a random variable X with CDF $P(x)$ can be sampled by using a random number generator to generate u_i from the uniform distribution uni$(0, 1)$, and calculating $x_i = P^{-1}(u_i)$. It is likely that an analytical expression for this inverse will not exist, in which case numerical methods must be used.

Multivariate

To use inverse transform sampling to sample a multi-dimensional PDF, some additional steps must be taken. Consider two continuous random variables X and Y with joint PDF $p(x, y)$.

First, one of the variables, y in this example but it could be either, must be averaged out by defining the marginal distribution function $p_X(x)$ and its CDF $P_X(x)$:

$$p_X(x) = \int_{-\infty}^{\infty} p(x,y)dy \qquad P_X(x) = \int_{-\infty}^{x}\int_{-\infty}^{\infty} p(x',y)dy\,dx' \,. \tag{B.2}$$

The marginal CDF $P_X(x)$ can be used to generate a sample x_i in the same manner as in the one dimensional case. To generate the corresponding sample y_i, the conditional probability distribution $p(y|x)$ and its CDF $P(y|x)$ must be defined:

$$p(y|x) = \frac{p(x,y)}{p_X(x)} \qquad P(y|x) = \frac{1}{p_X(x)}\int_{-\infty}^{y} p(x,y)dy \,. \tag{B.3}$$

The conditional CDF $P(y|x_i)$ can be used to generate the sample y_i as before. Note that if X and Y are independent, then $p(x,y) = p_X(x)p_Y(y)$ and x_i and y_i can be generated independently.

B.3 Estimating Population Mean And Variance

Univariate

Consider a random variable X which draws from a population with mean μ and variance σ^2:

$$\mu = \mathrm{E}[X] \qquad \sigma^2 = \mathrm{Var}(X) = \mathrm{E}\left[(X - \mathrm{E}[X])^2\right] = \mathrm{E}\left[X^2\right] - \mu^2 \,. \tag{B.4}$$

Now consider n independent samples X_i taken from this population. The sample mean μ_X and sample variance σ_X^2 are defined as follows:

$$\mu_X = \frac{1}{n}\sum_{i=1}^{n} X_i \qquad \sigma_X^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \mu_X)^2 = \frac{1}{n}\sum_{i=1}^{n} X_i^2 - \mu_X^2 \,. \tag{B.5}$$

The expectation value of the sample mean is equal to the population mean,

$$\mathrm{E}\left[\mu_X\right] = \mathrm{E}\left[\frac{1}{n}\sum_{i=1}^{n} X_i\right] = \frac{1}{n}\sum_{i=1}^{n}\mathrm{E}[X_i] = \frac{1}{n}\sum_{i=1}^{n}\mu = \mu \,, \tag{B.6}$$

therefore the sample mean is a good estimate of the population mean. The variance of the sample mean is equal to the variance of the population reduced by a factor of n,

$$
\begin{aligned}
\mathrm{Var}(\mu_X) &= \mathrm{E}\left[\mu_X^2\right] - \mathrm{E}[\mu_X]^2 \\
&= \mathrm{E}\left[\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right)^2\right] - \mu^2 \\
&= \mathrm{E}\left[\frac{1}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{n} X_i X_j\right] - \mu^2 \\
&= \frac{1}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{n} \mathrm{E}[X_i X_j] - \mu^2 \\
&= \frac{1}{n^2}\sum_{i=1}^{n} \mathrm{E}\left[X_i^2\right] + \frac{1}{n^2}\sum_{i=1}^{n}\sum_{j\neq i} \mathrm{E}[X_i X_j] - \mu^2 \\
&= \frac{n}{n^2}\left(\sigma^2 + \mu^2\right) + \frac{n(n-1)}{n^2}\left(\mu^2\right) - \mu^2 \\
&= \frac{\sigma^2}{n}\,.
\end{aligned}
\tag{B.7}
$$

Because X_i and X_j are independent samples, the expectation value $\mathrm{E}[X_i X_j]$ is equal to the product of the expectation value of each sample, both of which are equal to μ. The expectation value $\mathrm{E}\left[X_i^2\right]$ can be found to be equal to $\sigma^2 + \mu^2$ by rearranging the variance formula in Equation (B.4).

The sample variance σ_X^2 is a biased estimate of the population variance σ^2. This is because its expectation value is not equal to the population variance:

$$
\begin{aligned}
\mathrm{E}\left[\sigma_X^2\right] &= \mathrm{E}\left[\frac{1}{n}\sum_{i=1}^{n}(X_i - \mu_X)^2\right] = \frac{1}{n}\sum_{i=1}^{n}\mathrm{E}\left[\left(X_i - \frac{1}{n}\sum_{j=1}^{n} X_j\right)^2\right] \\
&= \frac{1}{n}\sum_{i=1}^{n}\mathrm{E}\left[X_i^2\right] - \frac{2}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{n}\mathrm{E}[X_i X_j] + \frac{1}{n^3}\sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{k=1}^{n}\mathrm{E}[X_j X_k] \\
&= \frac{1}{n}\sum_{i=1}^{n}\mathrm{E}\left[X_i^2\right] - \frac{1}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{n}\mathrm{E}[X_i X_j] \\
&= \frac{1}{n}\sum_{i=1}^{n}\mathrm{E}\left[X_i^2\right] - \frac{1}{n^2}\sum_{i=1}^{n}\mathrm{E}\left[X_i^2\right] - \frac{1}{n^2}\sum_{i=1}^{n}\sum_{j\neq i}\mathrm{E}[X_i X_j] \\
&= \left(\mu^2 + \sigma^2\right) - \frac{1}{n}\left(\mu^2 + \sigma^2\right) - \frac{n(n-1)}{n^2}\left(\mu^2\right) \\
&= \frac{n-1}{n}\sigma^2\,.
\end{aligned}
\tag{B.8}
$$

Therefore the quantity $n\sigma_X^2/(n-1)$ is a good estimate of the population variance, and combining this result with Equation (B.7), the quantity $\sigma_X^2/(n-1)$ is a good estimate of the variance of the sample mean. According to the central limit theorem, as the sample size n grows, the distribution of the sample mean μ_X becomes more like a normal distribution, regardless of the nature of the original population. Therefore the square root of the variance of the sample mean can be safely used as an estimate of the uncertainty of the sample mean.

In summary, an estimate μ_n of the mean of a population can be made by taking the mean of n samples, and the corresponding uncertainty σ_{μ_n} can be estimated by calculating the variance of those samples and dividing by $n-1$:

$$\mu_n = \frac{1}{n}\sum_{i=1}^{n} X_i \qquad \sigma_{\mu_n}^2 = \frac{1}{n-1}\left(\frac{1}{n}\sum_{i=1}^{n} X_i^2 - \mu_n^2\right). \tag{B.9}$$

Note that these values can be computed by keeping a running sum of X_i and X_i^2, avoiding the need to store every sample individually.

Multivariate

The same analysis is valid for a multivariate population. The random variable $\boldsymbol{X}$ is now a vector of length m, and its population has mean $\boldsymbol{\mu}$ and covariance matrix Σ:

$$\boldsymbol{\mu} = \mathrm{E}\left[\boldsymbol{X}\right] \qquad \Sigma = \mathrm{Cov}\left(\boldsymbol{X}\right) = \mathrm{E}\left[\left(\boldsymbol{X} - \mathrm{E}\left[\boldsymbol{X}\right]\right)\left(\boldsymbol{X} - \mathrm{E}\left[\boldsymbol{X}\right]\right)^T\right] = \mathrm{E}\left[\boldsymbol{X}\boldsymbol{X}^T\right] - \boldsymbol{\mu}\boldsymbol{\mu}^T. \tag{B.10}$$

The sample mean and sample covariance matrix are now defined as

$$\boldsymbol{\mu_X} = \frac{1}{n}\sum_{i=1}^{n} \boldsymbol{X_i} \qquad \Sigma_X = \frac{1}{n}\sum_{i=1}^{n}\left(\boldsymbol{X_i} - \boldsymbol{\mu_X}\right)\left(\boldsymbol{X_i} - \boldsymbol{\mu_X}\right)^T = \frac{1}{n}\sum_{i=1}^{n}\boldsymbol{X_i}\boldsymbol{X_i}^T - \boldsymbol{\mu_X}\boldsymbol{\mu_X}^T, \tag{B.11}$$

where n samples $\boldsymbol{X_i}$ have been taken from the population. Identities equivalent to those shown in Equations (B.6) through (B.8) take the following forms:

$$\mathrm{E}\left[\boldsymbol{\mu_X}\right] = \boldsymbol{\mu} \tag{B.12}$$

$$\mathrm{Cov}\left(\boldsymbol{\mu_X}\right) = \frac{\Sigma}{n} \tag{B.13}$$

$$\mathrm{E}\left[\Sigma_X\right] = \frac{n-1}{n}\Sigma\,. \tag{B.14}$$

Therefore the same conclusion can be made: an estimate $\boldsymbol{\mu_n}$ of the mean of the population can be made by taking the mean of n samples, and the covariance matrix $\Sigma_{\boldsymbol{\mu_n}}$ describing the uncertainty in this mean can be estimated by calculating the covariance of the samples and dividing by $n-1$:

$$\boldsymbol{\mu_n} = \frac{1}{n}\sum_{i=1}^{n}\boldsymbol{X_i} \qquad \Sigma_{\boldsymbol{\mu_n}} = \frac{1}{n-1}\left(\frac{1}{n}\sum_{i=1}^{n}\boldsymbol{X_i}\boldsymbol{X_i}^T - \boldsymbol{\mu_n}\boldsymbol{\mu_n}^T\right). \tag{B.15}$$

Note that, like the single variable case, running sums can be kept in order to negate the need to store every sample individually while retaining the ability to compute the covariance. In this case there are $m(m+3)/2$ running sums that must be kept: $X_{i,j}$ for $j = 1,\ldots,m$, and $X_{i,j}X_{i,k}$ for $j = 1,\ldots,m$ and $k = 1,\ldots,j$.

Sum of Multivariate

In the event that the sum of the elements of $\boldsymbol{X}$ is a quantity of interest, there are two possible methods to calculate the variance of the sum. Let the sum be called Y, so that $Y = \boldsymbol{u}^T\boldsymbol{X}$ where $\boldsymbol{u}$ is a vector containing ones. The sample mean is simply the sum of the elements of the vector sample mean:

$$\mu_Y = \boldsymbol{u}^T\boldsymbol{\mu}_X = \sum_{j=1}^{m}\mu_{X,j}. \tag{B.16}$$

To calculate the variance using the first method, consider Equation (A.9) where the gradient $\boldsymbol{\nabla} f$ is equal to $\boldsymbol{u}$ so that the product $\boldsymbol{u}^T\Sigma_X\boldsymbol{u}$ is simply the sum of the elements of Σ_X:

$$\sigma_Y^2 = \boldsymbol{u}^T\Sigma_X\boldsymbol{u} = \sum_{j=1}^{m}\sum_{k=1}^{m}\left(\frac{1}{n}\sum_{i=1}^{n}X_{i,j}X_{i,k} - \mu_{X,j}\mu_{X,k}\right). \tag{B.17}$$

For the second method method, running sums of the quantities $Y_i = \boldsymbol{u}^T\boldsymbol{X_i}$ and Y_i^2 must be kept, so that the mean and variance are given by

$$\mu_Y = \frac{1}{n}\sum_{i=1}^{n}Y_i \qquad \sigma_Y^2 = \frac{1}{n}\sum_{i=1}^{n}Y_i^2 - \mu_Y^2\,, \tag{B.18}$$

which is equivalent to the variance obtained in Equation (B.17):

$$
\begin{aligned}
\sigma_Y^2 &= \frac{1}{n}\sum_{i=1}^{n} Y_i^2 - \mu_Y^2 \\
&= \frac{1}{n}\sum_{i=1}^{n}\left(\sum_{j=1}^{m} X_{i,j}\right)^2 - \left(\sum_{j=1}^{m}\mu_{X,j}\right)^2 \\
&= \frac{1}{n}\sum_{i=1}^{n}\left(\sum_{j=1}^{m}\sum_{k=1}^{m} X_{i,j}X_{i,k}\right) - \sum_{j=1}^{m}\sum_{k=1}^{m}\mu_{X,j}\mu_{X,k} \\
&= \sum_{j=1}^{m}\sum_{k=1}^{m}\left(\frac{1}{n}\sum_{i=1}^{n} X_{i,j}X_{i,k} - \mu_{X,j}\mu_{X,k}\right) .
\end{aligned}
\tag{B.19}
$$

The second method is more straightforward and requires fewer running sums to be kept, but provides no information about the individual elements.

Appendix C

Levenberg-Marquardt Algorithm

C.1 Least Squares Optimization

Least squares optimization is a powerful tool for evaluating and understanding a set of measurements. One of the simplest cases is finding the line of best fit for data that is roughly linear. In this case, least squares optimization returns two import results: the slope and y-intercept of the line along with their uncertainties, and a measure of the linearity of the data. More complex models can also be used but the principle is the same: least squares optimization allows you to calculate optimal model parameters and their uncertainties, while also providing feedback on how well the model fits the data.

For example, take a simple model $f(x, \beta)$ with one parameter, β, being fit to N data points $\{x_k, y_k, \sigma_k\}$, where y_k are the measurements of the dependent variable at x_k, the corresponding independent variable values, and σ_k are the uncertainties associated with y_k. The optimal value of β is found by minimizing a cost function,

$$\Phi(\beta) = \sum_{k=1}^{N} \left(\frac{y_k - f(x_k, \beta)}{\sigma_k} \right)^2 \equiv \sum_{k=1}^{N} r_k^2(\beta) \,, \tag{C.1}$$

which consists of the sum of the squares of each data points deviation from the model evaluated at β. This cost function is not arbitrary, rather it is the result of the assumption that data points y_k are normally distributed around the model $f_k(\beta) = f(x_k, \beta)$ with standard deviations equal to the uncertainties σ_k. Under this assumption, the quantity r_k, called the residual, follows a standard normal distribution, and the probability $P_k(\beta)$ that a single measurement (x_k, y_k) came from a system described by $f(x, \beta)$ is proportional to the Gaussian,

$$P_k(\beta) \propto \exp\left(-\frac{(y_k - f_k(\beta))^2}{2\sigma_k^2} \right) \,, \tag{C.2}$$

and the probability $P(\beta)$ that the entire set of measurements came from a system described by $f(x, \beta)$ is proportional to the product of all these factors:

$$P(\beta) \propto \prod_{k=1}^{N} P_k = \exp\left(-\frac{1}{2}\sum_{k=1}^{N}\left(\frac{y_k - f_k(\beta)}{\sigma_k}\right)^2\right) = \exp\left(-\frac{1}{2}\Phi(\beta)\right). \tag{C.3}$$

From Equation (C.3) it is evident that minimizing the cost function $\Phi(\beta)$ is equivalent to maximizing the probability $P(\beta)$, or in other words picking the parameter β such that the model has the highest probability of describing the data. The optimal value of β will be denoted β_0. For simple models, β_0 can be found by setting the derivative of $\Phi(\beta)$ with respect to β to 0 and solving for β, but for more complex models recursive strategies such as the Levenberg-Marquardt Algorithm described in Section C.2 are employed.

Once β_0 has been found, an uncertainty can be estimated. Near β_0 the cost function can be approximated with a second order Taylor series,

$$\Phi(\beta) \approx \Phi(\beta_0) + \frac{d\Phi}{d\beta}(\beta - \beta_0) + \frac{1}{2}\frac{d^2\Phi}{d\beta^2}(\beta - \beta_0)^2, \tag{C.4}$$

where the derivatives are evaluated at $\beta = \beta_0$. Note that second term is actually equal to 0, since by definition the first derivative must be 0 at the minimum. If $\Phi(\beta)$ is approximately quadratic, then $P(\beta)$ is approximately normal, and we can find the standard deviation σ_β that describes this normal distribution, which can then be interpreted as the uncertainty in the optimal parameter β_0.

Substituting Equation (C.4) into Equation (C.3),

$$P(\beta) \propto \exp\left(-\frac{1}{2}\left(\Phi(\beta_0) + \frac{1}{2}\frac{d^2\Phi}{d\beta^2}(\beta - \beta_0)^2\right)\right) \propto \exp\left(-\frac{1}{4}\frac{d^2\Phi}{d\beta^2}(\beta - \beta_0)^2\right), \tag{C.5}$$

and comparing with the form of a normal distribution,

$$P(\beta) \propto \exp\left(-\frac{(\beta - \beta_0)^2}{2\sigma_\beta^2}\right), \tag{C.6}$$

it is clear that the uncertainty σ_β can be calculated from the second derivative of the cost function, or from derivatives of the model itself:

$$\frac{1}{\sigma_\beta^2} = \frac{1}{2}\frac{d^2\Phi}{d\beta^2} = \sum_{k=1}^{N}\frac{1}{\sigma_k^2}\left[\left(\frac{df_k}{d\beta}\right)^2 - \frac{d^2 f_k}{d\beta^2}\left(y_k - f_k(\beta)\right)\right]. \tag{C.7}$$

For a model $f(x, \boldsymbol{\beta})$ with M parameters $\boldsymbol{\beta}$, a generalization of Equation (C.7) describes elements of the inverse covariance matrix $\boldsymbol{\Sigma}_\beta^{-1}$:

$$\sigma_{ij}^{-1} = \frac{1}{2}\frac{\partial^2 \Phi}{\partial\beta_i \partial\beta_j} = \sum_{k=1}^{N} \frac{1}{\sigma_k^2}\left[\frac{\partial f_k}{\partial\beta_i}\frac{\partial f_k}{\partial\beta_j} - \frac{\partial^2 f_k}{\partial\beta_i\partial\beta_j}\left(y_k - f_k(\beta)\right)\right]. \tag{C.8}$$

This can be inverted to find the covariance matrix $\boldsymbol{\Sigma}_\beta$, and the square root of the diagonal elements of $\boldsymbol{\Sigma}_\beta$ serve as uncertainty estimates for $\boldsymbol{\beta}$. This is helpful for understanding that the uncertainty in $\boldsymbol{\beta}$ depends on the derivatives of the model, but in practice the covariance matrix may be estimated numerically as a part of the algorithm, avoiding the need to perform the expensive matrix inversion.

A helpful quantity for determining the quality of the fit is the reduced chi-squared statistic. A chi-squared (χ^2) distribution with k degrees of freedom is a probability distribution describing the sum of squares of k independent standard normal random variables. Given k samples of some distribution, the reduced chi-squared statistic χ_k^2, which is the chi-squared value of the samples χ^2 divided by the degrees of freedom k, contains information about the sampled distribution. If $\chi_k^2 \approx 1$, the distribution is approximately standard normal; if $\chi_k^2 > 1$, the distribution is wider ($\sigma > 1$); and if $\chi_k^2 < 1$ the distribution is narrower ($\sigma < 1$).

Note that the cost function $\Phi(\boldsymbol{\beta})$ is a χ^2 distribution, since the residuals $r_k(\boldsymbol{\beta})$ follow a standard normal distribution. Degrees of freedom are lost in the fitting process, so $\Phi(\boldsymbol{\beta})$ is a chi-squared distribution with $k = N - M$ degrees of freedom, where N is the number of data points and M is the number of model parameters contained in $\boldsymbol{\beta}$. Therefore the value

$$\chi_{N-M}^2 = \frac{1}{N-M}\Phi(\boldsymbol{\beta}) \tag{C.9}$$

can be compared to unity to determine the quality of the fit. If $\chi_{N-M}^2 > 1$, there are two plausible explanations: the errors σ_k have been underestimated, or the chosen model does not have enough flexibility to capture the trends in the data. Similarly, if $\chi_{N-M}^2 < 1$, either the errors σ_k have been overestimated, or the model has too much flexibility and is overfitting, essentially fitting to noise.

C.2 Levenberg-Marquardt Algorithm

The Levenberg-Marquardt algorithm is an iterative approach to solving a non-linear least squares problem, where the model $f(x, \boldsymbol{\beta})$ is non-linear with respect to $\boldsymbol{\beta}$. It combines the strengths of two separate non-linear least squares algorithms: the Gauss-Newton algorithm and the method of gradient descent.

In the Gauss-Newton algorithm, a value of $\boldsymbol{\beta}$ is perturbed by an amount $\boldsymbol{\delta}$ such that first order Taylor expansions of $f(x_k, \boldsymbol{\beta} + \boldsymbol{\delta})$ minimizes the cost function Φ. The perturbation $\boldsymbol{\delta}$ is found by taking the Taylor expansion,

$$f_k(\boldsymbol{\beta}+\boldsymbol{\delta}) \approx f_k(\boldsymbol{\beta}) + \sum_j \frac{\partial f_k}{\partial \beta_j}\delta_j\,, \tag{C.10}$$

plugging it into Equation (C.1), and setting the derivatives of Φ with respect to $\boldsymbol{\delta}$ to zero, which produces the matrix equation

$$\mathbf{J}^T\mathbf{J}\boldsymbol{\delta} = \mathbf{J}^T\boldsymbol{\Sigma}^{-1}(\boldsymbol{y}-\boldsymbol{f})\,, \tag{C.11}$$

where $\boldsymbol{y}$ is a vector containing measured data points y_k, $\boldsymbol{f}$ is a vector containing modelled data points $f_k(\boldsymbol{\beta})$, $\mathbf{J}$ is the Jacobian matrix containing derivatives evaluated at $\boldsymbol{\beta}$,

$$\mathbf{J} = \begin{pmatrix} \partial f_1/\partial\beta_1 & \partial f_1/\partial\beta_2 & \dots & \partial f_1/\partial\beta_M \\ \partial f_2/\partial\beta_1 & \partial f_2/\partial\beta_2 & \dots & \partial f_2/\partial\beta_M \\ \vdots & \vdots & \ddots & \vdots \\ \partial f_N/\partial\beta_1 & \partial f_N/\partial\beta_2 & \dots & \partial f_N/\partial\beta_M \end{pmatrix}\,, \tag{C.12}$$

and $\boldsymbol{\Sigma}$ is the measurement covariance matrix, which is diagonal for independent measurements:

$$\boldsymbol{\Sigma} = \begin{pmatrix} \sigma_1^2 & & 0 \\ & \ddots & \\ 0 & & \sigma_N^2 \end{pmatrix}\,. \tag{C.13}$$

Equation (C.11) can be solved for $\boldsymbol{\delta}$. This is highly effective when $f(x,\boldsymbol{\beta})$ is nearly linear or when $\boldsymbol{\beta}$ is already quite close to the optimal solution, but it can actually increase the value of the cost function $\Phi(\boldsymbol{\beta})$ if these conditions are not met. The contribution of Levenberg in 1944 (Levenberg, 1944) was to propose a scale factor to limit on the length of $\boldsymbol{\delta}$ to avoid overstepping the solution in these scenarios.

In the method of gradient descent, $\boldsymbol{\delta}$ is chosen in the direction opposite to the gradient of the cost function, or in other words in the direction of steepest descent. Given a small enough length of $\boldsymbol{\delta}$, this guarantees that $\boldsymbol{\beta}$ moves towards the solution at the minimum. This method is less likely to fail compared to the Gauss-Newton algorithm, but the method is very slow as you approach the solution. The gradient of the cost function is

$$\boldsymbol{\nabla}\Phi(\boldsymbol{\beta}) = -2\mathbf{J}^T\boldsymbol{\Sigma}^{-1}(\boldsymbol{y}-\boldsymbol{f})\,, \tag{C.14}$$

so the gradient descent method requires that

$$\boldsymbol{\delta} \propto \mathbf{J}^T \boldsymbol{\Sigma}^{-1} (\boldsymbol{y} - \boldsymbol{f}) \,. \tag{C.15}$$

Since the Gauss-Newton algorithm performs well near the solution, and the gradient descent method performs well far from the solution, a combination of the two makes intuitive sense. This was the contribution of Marquardt in 1963 (Marquardt, 1963), suggesting the following equation to solve for $\boldsymbol{\delta}$ which interpolates between the two methods:

$$(\mathbf{J}^T \mathbf{J} + \lambda \mathbf{I}) \boldsymbol{\delta} = \mathbf{J}^T \boldsymbol{\Sigma}^{-1} (\boldsymbol{y} - \boldsymbol{f}) \,. \tag{C.16}$$

When λ is small, this resembles Equation (C.11) (Gauss-Newton), and when λ is large, it resembles Equation (C.15) (gradient descent). If $\boldsymbol{\delta}$ calculated with small λ brings a satisfactory reduction in Φ, then it is assumed that the solution is nearby. If it does not bring a satisfactory reduction in Φ, then it is assumed that the solution is far away, and λ can be increased so that the next $\boldsymbol{\delta}$ more closely follows the gradient descent method.

Appendix D

Vertical Grids

Several different parameterizations of vertical space are commonly used in atmospheric science. Within the scope of this thesis, there are pressure grids used in the DOAS literature such as Bucsela et al. (Bucsela et al., 2013), the hybrid sigma-pressure grid used by the GEOS-5 ESM model, optical depth grids used by the discrete ordinates radiative transfer algorithm, and altitude grids used by SASKTRAN. This appendix outlines some of the theory and the methods that were used to convert between grids in this thesis.

D.1 Hybrid Sigma-Pressure Grids

HSP grids are used by global atmospheric models such as GEOS-5 ESM in order to efficiently account for changes in terrain elevation in the lower atmosphere while leaving the upper atmosphere unaffected. They have evolved from pure sigma levels $\sigma(L)$ which represent a fraction of the total change in pressure from the surface to the defined top of atmosphere, so that the true pressure $P(I, J, L)$ can be calculated as

$$P(I, J, L) = \sigma(L)P_{\text{surface}}(I, J) + (1 - \sigma(L))P_{\text{top}}\,, \tag{D.1}$$

where L is the index for the vertical sigma coordinate and I and J represent a horizontal location. However, it was found that pure sigma levels caused unphysical terrain signatures to appear in the stratosphere. The HSP grid was introduced to solve this problem, in which the lowermost levels are sigma coordinates, the uppermost levels are unchanging pressure coordinates, and the middle levels are a combination of the two. GEOS-5 uses a 72-level HSP grid, with 73 edges defined by the constants $A_p(L)$ and $B_p(L)$ displayed in Table D.1 (Yantosca, 2016). Given a surface pressure, the pressures corresponding to each edge can be calculated as follows:

$$P(I, J, L) = A_p(L) + B_p(L)P_{\text{surface}}(I, J)\,. \tag{D.2}$$

D.2 Integration on Pressure Grids

Vertical profiles of trace gas concentrations are typically specified either as VMRs on pressure grids, or as number densities on altitude grids. Both forms are easily integrated to compute VCDs. Calculating the VCD of a trace gas profile that is specified in number density on an

L	$A_p(L)$ (hPa)	$B_p(L)$	L	$A_p(L)$ (hPa)	$B_p(L)$
1	0.000000	1.000000	38	6.660341×10^{1}	0.000000
2	4.804826×10^{-2}	9.849520×10^{-1}	39	5.638791×10^{1}	0.000000
3	6.593752	9.634060×10^{-1}	40	4.764391×10^{1}	0.000000
4	1.313480×10^{1}	9.418650×10^{-1}	41	4.017541×10^{1}	0.000000
5	1.961311×10^{1}	9.203870×10^{-1}	42	3.381001×10^{1}	0.000000
6	2.609201×10^{1}	8.989080×10^{-1}	43	2.836781×10^{1}	0.000000
7	3.257081×10^{1}	8.774290×10^{-1}	44	2.373041×10^{1}	0.000000
8	3.898201×10^{1}	8.560180×10^{-1}	45	1.979160×10^{1}	0.000000
9	4.533901×10^{1}	8.346609×10^{-1}	46	1.645710×10^{1}	0.000000
10	5.169611×10^{1}	8.133039×10^{-1}	47	1.364340×10^{1}	0.000000
11	5.805321×10^{1}	7.919469×10^{-1}	48	1.127690×10^{1}	0.000000
12	6.436264×10^{1}	7.706375×10^{-1}	49	9.292942	0.000000
13	7.062198×10^{1}	7.493782×10^{-1}	50	7.619842	0.000000
14	7.883422×10^{1}	7.211660×10^{-1}	51	6.216801	0.000000
15	8.909992×10^{1}	6.858999×10^{-1}	52	5.046801	0.000000
16	9.936521×10^{1}	6.506349×10^{-1}	53	4.076571	0.000000
17	1.091817×10^{2}	6.158184×10^{-1}	54	3.276431	0.000000
18	1.189586×10^{2}	5.810415×10^{-1}	55	2.620211	0.000000
19	1.286959×10^{2}	5.463042×10^{-1}	56	2.084970	0.000000
20	1.429100×10^{2}	4.945902×10^{-1}	57	1.650790	0.000000
21	1.562600×10^{2}	4.437402×10^{-1}	58	1.300510	0.000000
22	1.696090×10^{2}	3.928911×10^{-1}	59	1.019440	0.000000
23	1.816190×10^{2}	3.433811×10^{-1}	60	7.951341×10^{-1}	0.000000
24	1.930970×10^{2}	2.944031×10^{-1}	61	6.167791×10^{-1}	0.000000
25	2.032590×10^{2}	2.467411×10^{-1}	62	4.758061×10^{-1}	0.000000
26	2.121500×10^{2}	2.003501×10^{-1}	63	3.650411×10^{-1}	0.000000
27	2.187760×10^{2}	1.562241×10^{-1}	64	2.785261×10^{-1}	0.000000
28	2.238980×10^{2}	1.136021×10^{-1}	65	2.113490×10^{-1}	0.000000
29	2.243630×10^{2}	6.372006×10^{-2}	66	1.594950×10^{-1}	0.000000
30	2.168650×10^{2}	2.801004×10^{-2}	67	1.197030×10^{-1}	0.000000
31	2.011920×10^{2}	6.960025×10^{-3}	68	8.934502×10^{-2}	0.000000
32	1.769300×10^{2}	8.175413×10^{-9}	69	6.600001×10^{-2}	0.000000
33	1.503930×10^{2}	0.000000	70	4.758501×10^{-2}	0.000000
34	1.278370×10^{2}	0.000000	71	3.270000×10^{-2}	0.000000
35	1.086630×10^{2}	0.000000	72	2.000000×10^{-2}	0.000000
36	9.236572×10^{1}	0.000000	73	1.000000×10^{-2}	0.000000
37	7.851231×10^{1}	0.000000			

Table D.1: Hybrid sigma-pressure grid definition used by GEOS-5.

altitude grid is intuitive: integrate the number density over altitude from the surface of the earth to the top of the atmosphere:

$$V = \int_0^H n_{\text{trace}}(z) dz \,. \tag{D.3}$$

The hydrostatic equation can be used to perform a change of variables from altitude to pressure. The equation reads

$$dp = -n\, m\, g\, dz \,, \tag{D.4}$$

stating that the difference in pressure dp across a small vertical distance dz is due to the weight of the air within dz, where n is the number density of the air, m is the mass of one air molecule, and g is acceleration due to gravity. In theory, pressure as a function of altitude could be calculated by combining this equation with the ideal gas law, given a temperature profile. In the present work, however, pressure and temperature profiles are acquired from an external source, typically from the MSIS-90 model. It is assumed that these profiles obey the hydrostatic equation and the ideal gas law to a good approximation, so that the pressure profile $p(z)$ and the hydrostatic equation can be used together to perform a change of variables on Equation (D.3):

$$V = \int_{p(0)}^{p(H)} -\frac{n_{\text{trace}}(p)}{m\, n(p)\, g(p)}\, dp \,. \tag{D.5}$$

Neglecting the altitude dependence of gravity and defining the VMR $\xi_{\text{trace}}(p) = n_{\text{trace}}(p)/n(p)$, this can be written as

$$V = \frac{1}{mg} \int_{p(H)}^{p(0)} \xi_{\text{trace}}(p) dp \,. \tag{D.6}$$

Note that the ratio $n_{\text{trace}}(p)/n(p)$ is actually the definition of the mixing ratio, which is only equivalent to the VMR under the assumption of ideal gas. The equivalence of Equations (D.3) and (D.6) depends on three assumptions: air is well described by the hydrostatic equation, air is well described by the ideal gas law, and the change in gravity with altitude is negligible.

D.3 Integration on Hybrid Sigma-Pressure Grids

Consider $\boldsymbol{\xi}$, a vector of length M containing elements ξ_j which represent a constant VMR within the layer j, where $j = 1$ represents the ground layer and $j = M$ represents the layer at the top of the defined atmosphere. Let the layer j have boundary pressures p_{j-1} and p_j, where p_{j-1} represents the boundary that is lower in altitude but higher in pressure. If $\boldsymbol{D}_p$ is

a vector of length M containing pressure differences $d_j = p_{j-1} - p_j$, then Equation (D.6) can be evaluated with the dot product

$$V = \frac{1}{mg} \boldsymbol{D}_p^T \boldsymbol{\xi} \,. \tag{D.7}$$

A variant of this method was used to calculate layer-specific VCDs V_i and their uncertainties, given NO_2 profile information from GEOS-5, which is defined on a HSP grid. Let $\boldsymbol{V}$ be a vector of length N containing elements V_i, where again $i = 1$ represents the lowest layer, and where layer i is defined by boundary altitudes z_{i-1} and z_i with z_{i-1} representing the lower boundary. In the present work, these layers have a uniform thickness of 500 m, matching the grid used to calculate box-AMF tables. An external pressure profile $p(z)$ was used to calculate the corresponding pressure boundaries $p_i = p(z_i)$, and the surface pressure p_0 from this external profile was used with Equation (D.2) to calculate the HSP pressure boundaries p_j. An $N \times M$ matrix D_p was defined to integrate the VMR profile over pressure while simultaneously sorting the contributions from each HSP layer into the correct 500 m layers. The matrix D_p contains elements d_{ij}, which represent the pressure overlap of 500 m layer i and HSP layer j:

$$d_{ij} = \max(0, \min(p_{i-1}, p_{j-1}) - \max(p_i, p_j)) \,, \tag{D.8}$$

which allows the 500 m layer VCDs to be computed as follows:

$$\boldsymbol{V}' = \frac{1}{mg} D_p \boldsymbol{\xi} \,. \tag{D.9}$$

At higher altitudes where the HSP layers become larger than the 500 m layers, the constant VMR within each HSP layer results in unphysical oscillations in the 500 m layer VCDs, as illustrated in Figure D.1. The oscillations are smoothed out using the following $N \times N$ matrix,

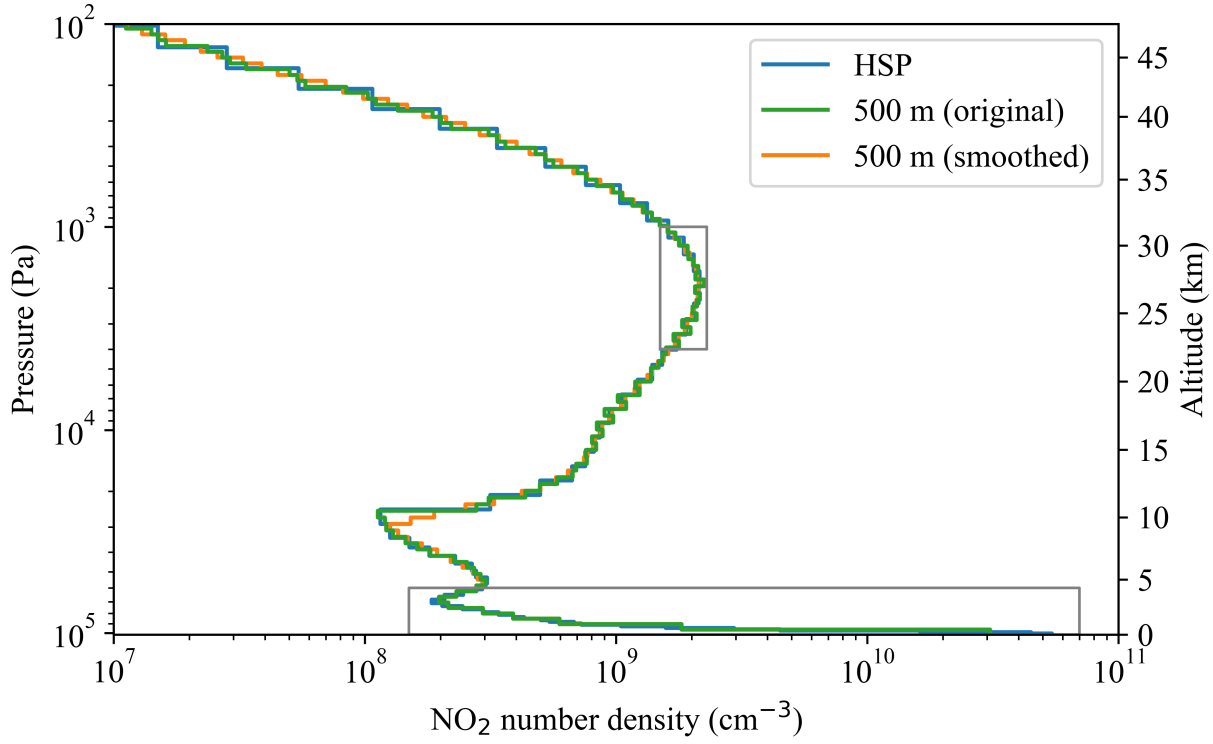
$$S_p = \begin{pmatrix} 1 & & & & & & & & \\ & \ddots & & & & & & 0 & \\ & & 1 & & & & & & \\ & & & 0.6 & 0.2 & 0.2 & & & \\ & & & 0.2 & 0.4 & 0.2 & 0.2 & & \\ & & & 0.2 & 0.2 & 0.2 & \ddots & \ddots & \\ & & & & 0.2 & \ddots & \ddots & \ddots & 0.2 \\ & & & & & \ddots & \ddots & 0.2 & 0.2 \\ & 0 & & & & & 0.2 & 0.2 & 0.4 \\ & & & & & & & 0.2 & 0.2 \end{pmatrix} \begin{matrix} \\ \\ \\ \\ \\ \\ 0.2 \\ 0.2 \\ 0.2 \\ 0.6 \end{matrix} \tag{D.10}$$

which leaves lower layers untouched and performs a five point moving average on higher layers without altering the total VCD. The moving average begins at the lowest 500 m layer that is smaller then one of its overlapping HSP layers. The smoothed 500 m layer VCDs are computed as follows:

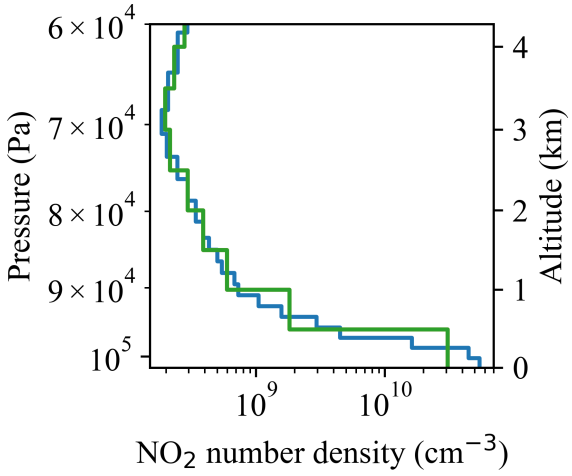
$$\boldsymbol{V} = S_p \boldsymbol{V}' = \frac{1}{mg} S_p D_p \boldsymbol{\xi} \,. \tag{D.11}$$

The covariance matrix Σ_ξ representing the uncertainty in $\boldsymbol{\xi}$ was similarly mapped to a covariance matrix Σ_V representing the uncertainty in $\boldsymbol{V}$ using the same method:

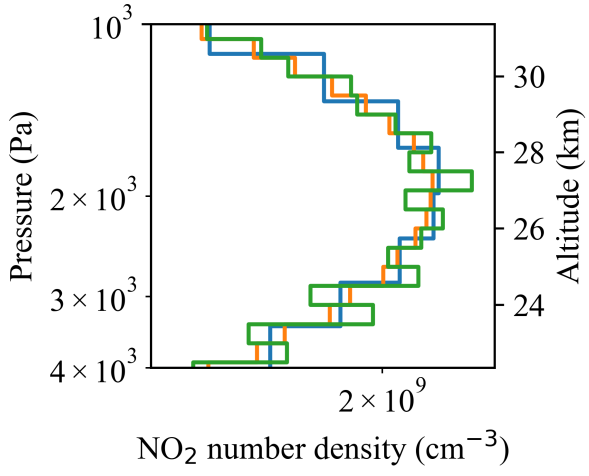
$$\Sigma_V = \left(\frac{1}{mg}\right)^2 S_p D_p \Sigma_\xi D_p^T S_p^T \,. \tag{D.12}$$



(a) NO_2 profile with boxed regions of interest matching the figures below.



(b) Low altitudes: HSP layers are smaller than the 500 m layers, so no smoothing operation is necessary.



(c) High altitudes: HSP layers are larger than the 500 m layers, resulting in unphysical oscillations which are reduced by a smoothing operation.

Figure D.1: An example NO_2 profile that explicitly shows the original HSP layers and the derived 500 m layers. For both layer types, number density values were acquired by dividing the layer-specific VCD by the thickness of the layer in cm.