A New Retrieval Scheme to Determine Tropospheric Composition from the Global Ozone Monitoring Experiment

by

Lara Nicole Gunn

Submitted in accordance with the requirements for the degree of Doctor of Philosophy.

University of Leeds

School of Earth and Environment

September 2008.

Chapter 3 of this thesis contains work that has previous been published as part of a joint publication (Gil et al., 2008) which can be found in Appendix 3. The subject of the publication is the comparison of NO_2 long–term trends at Izana, Tenerife. The publication contained a ground based study (M. Gil and M. Yela), satellite validation (A. Richter) and a comparison with chemical transport model data using the SLIMCAT model (L. Gunn - author). The publication includes results from two of the three model runs used in Chapter 3.

Abstract

The retrieval of tropospheric trace gas abundances from nadir–viewing satellite observations is challenging and this thesis explores new methods to improve on previous approaches. A new quantitative retrieval scheme has been developed and applied to observations from the Global Ozone Monitoring Experiment (GOME). The retrieval is framed as an inverse problem and includes two novel steps. First, the stratospheric component of the total vertical column is constrained using a chemical transport model (CTM) with chemical data assimilation. Second, coincident cloud and aerosol information from the Along-Track Scanning Radiometer (ATSR–2) is used to provide a more accurate model for the atmospheric state.

In order to provide the stratospheric constraint, observations of long-lived tracers from the Halogen Occultation Experiment (HALOE) have been assimilated into the SLIMCAT 3-D CTM. These are the first multiannual simulations of a full chemistry CTM with data assimilation. The assimilation model is shown to agree well with independent observations. Additionally, the multiannual simulation shows good agreement with long-term mid-latitude stratospheric column NO_2 observations in the 1990s indicating that discrepancies in previous SLIMCAT runs were caused by transport errors in the ERA-40 winds used to force the model.

The new Retrieval Model has been developed and thoroughly tested to determine the sensitivity of the retrieved columns to various model, cloud and aerosol parameters. It has been shown that the retrieval is most sensitive to cloud optical depth, followed by cloud top height and surface reflectance. It is also shown that the retrieval is sensitive to the stratospheric column. Realistic representations of these parameters are therefore needed for an accurate retrieval. Combining ATSR–2 and GOME data involves averaging the data from the ATSR–2 pixels onto the GOME pixel. Different averaging techniques have been assessed and a two-zone averaging (one cloudy, one cloud free) has been found to be the optimum method.

A full sensitivity analysis of the Retrieval Model highlighted that different random and systematic errors will affect the overall retrieval. Random error on the measurement and the calculated stratospheric column have been shown to be highly significant, with both sources causing errors in the retrieved tropospheric column of between 10–60% for most conditions (errors become very large for high optically thick clouds) for NO_2 and BrO. Systematic errors in the cloud parameters (errors in the ATSR–2 parameters) lead to errors in the retrieved tropospheric column of $\pm 20\%$ except at high cloud optical depths. The improvement of the systematic error by including cloud has been shown to be highly significant, emphasising the importance of including cloud in a Retrieval Model of this type.

The Retrieval Model has been used to produce example tropospheric vertical column data for NO_2 and BrO for July 1997 and November 1996, respectively. The NO_2 retrievals showed the expected enhancements for industrial emissions (3–20×10^{15} molecules cm^{-2}) and biomass burning emissions (3–10×10^{15} molecules cm^{-2}) while the BrO retrievals show a global background of tropospheric BrO (0.5–3×10^{13} molecules cm^{-2}). Errors for these two retrievals have been quantified. For NO_2 the errors were slightly smaller than those seen in previous quantitative retrievals with random errors of 0.2–6×10^{15} molecules cm^{-2}, with smaller values over remote areas (oceans) and larger values over enhanced NO_2, and with systematic errors ranging from -1–+0.1×10^{15} molecules cm^{-2}, with larger errors over areas of enhanced NO_2. BrO errors were quite large with random errors ranging from 0–3×10^{13} molecules cm^{-2}, which is comparable to the total tropospheric column that is retrieved, and systematic errors ranging from -0.15–+0.1×10^{13} molecules cm^{-2}. This is the first quantitative retrieval of tropospheric BrO and therefore the errors help to highlight areas where the retrieval needs to be improved further. One such source of error is shown to be the stratospheric correction, especially as variations in stratospheric column BrO correlate with variations in tropospheric BrO derived in other studies.

Acknowledgements

Contents

List of Acronyms

AERONET	AErosol RObotic NETwork
AMF	Air Mass Factor
AMF_G	Geometric Air Mass Factor
ATSR	Along Track Scanning Radiometer
AATSR	Advanced Along Track Scanning Radiometer
AVHRR	Advanced Very High Resolution Radiometer
BUV	Backscatter Ultra Violet
CMAQ	Community Multiscale Air Quality
CTM	Chemistry Transport Model
DOAS	Differential Optical Absorption Spectroscopy
ECMWF	European Centre for Medium Range Weather Forecasts
ENVISAT	Environmental Satellite
ERS	European Research Satellite
ESA	European Space Agency
ESD	Estimated Standard Deviation
FRESCO	Fast Retrieval Scheme for Clouds from the Oxygen A-band
GOME	Global Ozone Monitoring Experiment
HALOE	Halogen Occulation Experiment
IPCC	Intergovermental Panel for Climate Change
IPT	Image Processing Technique
MetOP	Meterological Satellite
MSG	MeteoSat Second Generation
NDVI	Normalized Difference Vegetation Index
OMI	Ozone Monitoring Instrument
ORCA	Optical Cloud Recognition Algorithm
PAR	Photosynthetically Active Radition
PDF	Probability Denisty Function
RAL	Rutherford Appleton Laboratory
RSG	Remote Sensing Group
SAR	Synthetic Aperture Radar
SBUV	Solar Backscatter Ultra Violet
SCD	Slant Column Density
SCIAMACHY	Scanning Imaging Absorption Spectrometer for Atmospheric CHartographY
SEVIRI	Spinning Enhanced Visible and Infrared Imager
SST	Sea Surface Temperature
STFC	Science and Technology Facilities Council
TOMS	Total Ozone Mapping Spectrometer
UARS	Upper Atmosphere Research Satellite
UT	Universal Time
UV	Ultra Violet

VCD	Vertical Column Density
VIS	Visible
VOC	Volatile Organic Compounds

List of Symbols

I	Intensity
λ	Wavelength
σ	Scattering cross–section
c	Concentration – number density
s	Distance
τ	Optical depth
S	Slant column density
y	Measurement
ϵ	Error
N_p	Number of Photons
V	Vertical column density
l	Photon path length
z	Altitude
E	Energy
θ	Angle
Ω	Solid angle
A	Area
t	time
F_λ	Irradiance
ϕ	Azimuth angle
k_{ext}	Extinction coefficent
k_{sca}	Scattering coefficient
k_{abs}	Absorption coefficient
ω_0	Single scattering coefficient
P	Phase function
ϕ	Bi–directional distribution function
r	radius
x	Size parameter
$\underline{V}$	Velocity vector
$\overline{P}$	Production rate
L	Loss rate
f	Forward function
y	Measurement vector
x	State vector
b	Retrieval model parameters (that are not retrieved)
ϵ	Error vector
R	Retrieval model
$\mathbf{x}_a$	*A priori*
$\hat{b}$	Best estimate of **b**
F	Forward model
c	a constant

P	Probability density function
$\mathbf{S}_a$	Covariance of the *a priori*
$\mathbf{S}_y$	Covariance of the measurement
$\mathbf{S}_x$	Covariance of the state
$\hat{x}$	Expected value of $\mathbf{x}$
$\mathbf{K}$	Weighting function matrix
χ^2	Cost function
$\mathbf{x}_0$	Initial guess at the state vector
Γ_i	Damping factor
$\mathbf{I}$	Unity matrix
$\mathbf{G}$	Gain matrix
$\mathbf{A}$	Averaging kernel matrix
$\mathbf{A}_p$	Profile averaging kernel
$\mathbf{A}_s$	Stratospheric profile averaging kernel
$\mathbf{A}_t$	Tropospheric profile averaging kernel
$\mathbf{A}_b$	Boundary layer profile averaging kernel
x_{sf}	Tropospheric scaling factor
R	Radiance
$\mathbf{K}_G$	Weight function matrix output by GOMETRAN
r(λ)	Sun–normalised radiance
f	Cloud fraction
n$_{ret}$	Number of retrievals
θ_s	Solar zenith angle
α_c	Cloud top reflectance
α	Surface reflectance
y$_{scd}$	Measurement - slant column density
y$_{scd_g}$	Measurement – GOME slant column density
y$_{scd_s}$	Measurement – SLIMCAT slant column density

List of Figures

Chapter 1

Introduction

1.1 Motivation and Aims

The troposphere is an essential part of the Earth's life support system. It is the lowest layer in the atmosphere (from the surface, to 15 km in the tropics and 8 km at the poles) and contains a variety of chemical species which are subject to different transport, production and destruction processes. A primary goal of tropospheric chemistry research is to measure and understand the response of tropospheric composition to natural and anthropogenic perturbations and to develop the capability to predict future change, due to a changing climate or human development.

Many key tropospheric trace gases have low concentrations and large variability, making them difficult to measure. Most tropospheric trace gas concentration data (e.g. O_3, NO_2, SO_2 profiles, total column and surface concentrations) have come from surface sites e.g. Junfraujoch, Switzerland or Lauder, New Zealand, and aircraft campaigns e.g. ITOP (Ravetta et al., 2007) and INTEX–A (Singh et al., 2006), in the past, meaning that only specific locations and, in the case of aircraft missions specific times are observed. Now, satellite instruments are able to provide global observations, in some cases on a daily basis, of many tropospheric trace gases including total column amounts and profiles (e.g. Singh and Jacob, 2000).

Initially nadir viewing ultra violet/visible (UV/VIS) satellite instruments focused on measuring vertical and total column ozone (O_3). The first such instrument was the Backscatter Ultraviolet (BUV or SBUV) instrument onboard the NASA NIMBUS 4 satellite (Heath et al., 1973) (launched in April 1970), whose primary aim was to measure O_3 profiles. However, due to unreliability (instrumental errors), in 1979 another instrument was launched called the Total Ozone Monitoring Spectrometer (TOMS) on the NASA NIMBUS 7 satellite (McPeters et al., 1998). TOMS was highly successful and has the longest running data set of total column O_3 (1979-1992), covering the period when the Antarctic O_3 hole was discovered (Solomon, 1999). The TOMS instrument, although sensitive to surface O_3, lacked the sensitivity to differentiate between the troposphere and the stratosphere. The next step was to develop an instrument

that could measure height–resolved O_3 and monitor more atmospheric species. This lead to the Global Ozone Monitoring Experiment (GOME) onboard the ESA ERS–2 (launched in 1995). GOME measured across a much larger wavelength region (240–790 nm), enabling the further retrieval of a number of UV/VIS species (Burrows et al., 1999). Since the success of GOME there have been another three similar instruments; the SCanning Imaging Absorption spectroMeter for Atmospheric ChartograpHY (SCIAMACHY) onboard the ESA ENVISAT satellite, the Ozone Monitoring Experiment (OMI) onboard the NASA AURA satellite and recently the second GOME instrument (GOME-2) onboard the ESA EUMETSAT polar platform, MetOp.

Despite the number of UV/VIS instruments that obtain tropospheric trace gas concentrations there is still a significant gap in the quantitative analysis of these measurements. The overall aim of this thesis is to develop a quantitative retrieval scheme for determining tropospheric NO_2 and BrO from the GOME instrument and therefore demonstrate the full potential of these observations.

The following sections review the steps necessary to produce a retrieval scheme and describe how each step has been achieved in previous studies. Section 1.2 describes the satellite instruments used in this thesis, these are GOME and the second Along Track Scanning Radiometer (ATSR–2). Section 1.3 gives an overview of the principal atmospheric chemistry involving NO_2 and BrO (which includes O_3 chemistry), the target species in this thesis. Section 1.4 describes the previous retrieval methods that have been applied to nadir-viewing UV/VIS spectrometers such as GOME. Section 1.5 summarises the previous work done using the GOME column information. Section 1.6 introduces the GRAPE aerosol/cloud dataset used in this work, and Section 1.7 lists the overall objectives of this thesis.

1.2   Instrumentation

The data used in this thesis has been produced using two instruments: GOME and ATSR–2. These instruments are onboard the ESA ERS–2 satellite which was launched in April 1995. ERS–2 is in a sun-synchronous orbit at an altitude of 780 km and 98.5° of inclination. The sub-satellite speed is 7 km s^{-1}, which corresponds an orbital period of 100 minutes and consequently the satellite completes 14 orbits per day. ERS–2 crosses the equator at a local time of 10:30 on the day side flying north to south. Figure 1.1 shows the ERS–2 satellite and the position of the two instruments.

1.2.1   Global Ozone Monitoring Experiment (GOME)

The GOME instrument is an across-track scanning optical spectrometer measuring in the 240–790 nm spectral range with a moderate spectral resolution of 0.2 nm in the ultra violet and 0.4 nm in the visible and near infrared (Burrows et al., 1999). The full normal scanning width is 960 km, using three ground pixels of size 40×320 km^2 which together provide global coverage every 3 days, after 42 orbits. The instrument scans across the three pixels, each in 1.5 seconds
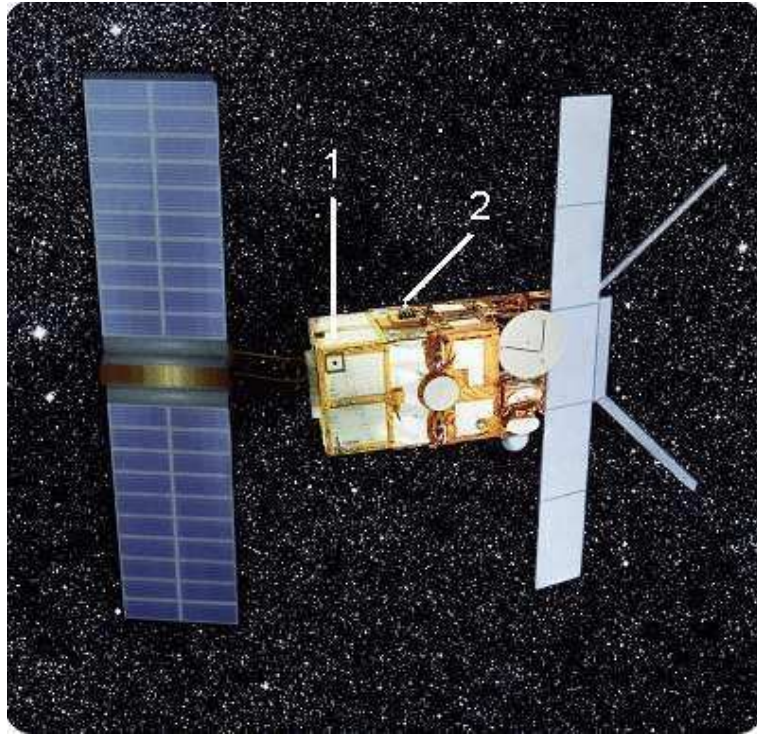
Figure 1.1: *The second European Research Satellite (ERS–2). The position of GOME (1) and ATSR–2 (2) are indicated. The image has been adapted from http: //www.naturalsciences.be*

and then completes a back scan in 1.5 seconds (see Figure 1.3).

Figure 1.2 illustrates the optics of the GOME instrument. Light back–scattered from the Earth is viewed via a scan mirror (near the nadir view port), which allows the instrument's instantaneous field-of-view (IFOV) to be scanned across the satellite ground track (i.e. viewing off-nadir). From the scan mirror, light passes through a telescope to the spectrometer entrance slit. From the slit, light is divided into 4 channels: A pre-disperser prism allows channels 1-2 (below 405 nm) to be separated from channels 3-4 and a second prism splits channels 1 (235–316 nm) and 2 (311–405 nm). Light to channels 3 (395–611 nm) and 4 (605–793 nm) is divided by a beam splitter. The spectrum in each of the channels is generated by a diffraction grating. The spectra are recorded on four linear Reticon Si-diode arrays with 1024 spectral elements each. The spectra are recorded from 240 nm to 790 nm (Weber et al., 1998).

The scene viewed by GOME, at an instant, is determined by the IFOV (Figure 1.3), which is defined by the spectrometer entrance slit and telescope. The narrow, dispersion dimension of the slit maps onto the across-track axis, and the long dimension is mapped along track (sketched in Figure 1.3). On the Earth's surface, the IFOV for GOME is 4 km wide in the across-track direction and 40 km along-track.

GOME measures backscattered radiance from the Earth (nadir view port) and solar irradiance (solar diffuser plate). As the backscattered spectra will contain spectral figures caused by atmospheric parameters that will not appear in the solar irradiance spectra, the ratio of these
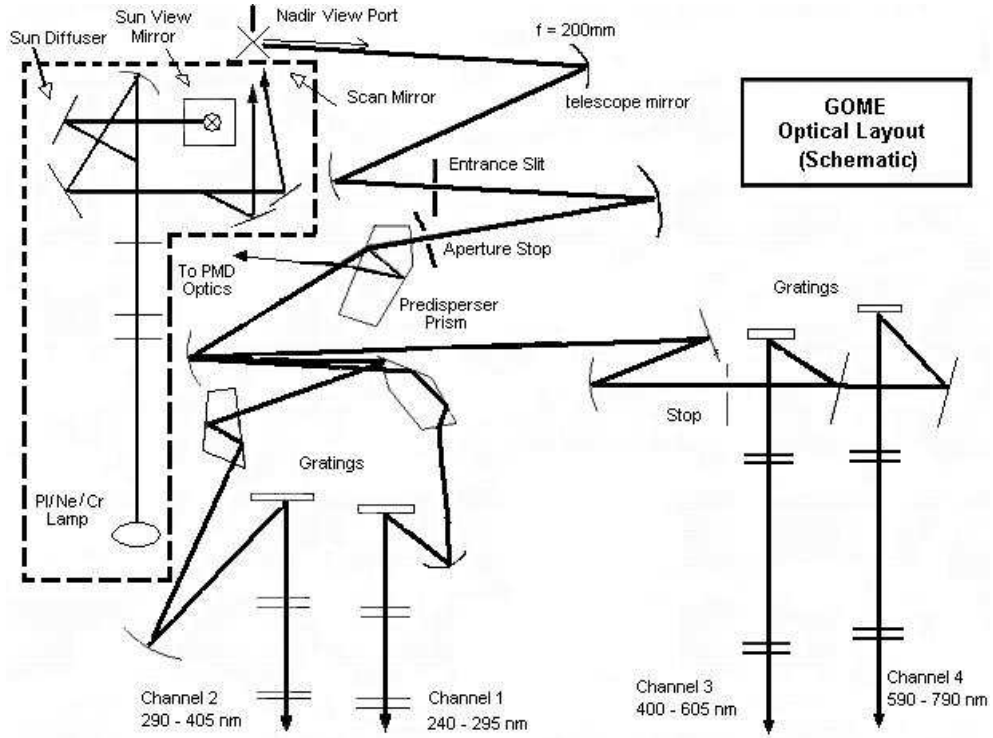
Figure 1.2: *Schematic of the GOME optics taken from Weber et al. (1998).*

two spectra contains information on these atmospheric parameters such as spectral signatures of a number of trace gases. Examples of this include NO_2, BrO, H_2O, O_3 and O_2 as well as NO_3, OClO and NO above 40 km, SO_2 under polluted conditions and following volcanic eruptions and HCHO under polluted conditions (Burrows et al., 1999). In general only one piece of information is available regarding the vertically integrated total amount of gas in the atmosphere. This is usually referred to as the total column amount of gas (units of molecules cm^{-2}). In the case of O_3, height–resolved information can be retrieved by exploiting the large gradient in O_3 optical depth across the Hartley–Huggins bands (e.g. see Munro et al., 1998). The secondary objective of GOME was to measure cloud radiative properties such as reflectance and cloud top height as well as properties such as surface reflectance, aerosol vertical optical depth and type, as well as solar UV irradiance variability which is related to the 11-year solar cycle.

GOME Diffuser Plate

The diffuser plate on the GOME instrument (see Figure 1.2) is used to record the solar spectrum on a roughly daily basis. This is then used to calculate the sun–normalised radiance. However, there is an imperfection present in the diffuser plate (often called the diffuser plate artefact) resulting in spectrally structured errors in the observed measurement of the solar spectrum. These errors vary from day to day (as they are dependent on viewing angle). Only the relative variations in the spectral error can be deduced from in-flight data, not the absolute error (e.g. Richter and Wagner, 1998). Therefore, in order to correctly determine this error assumptions
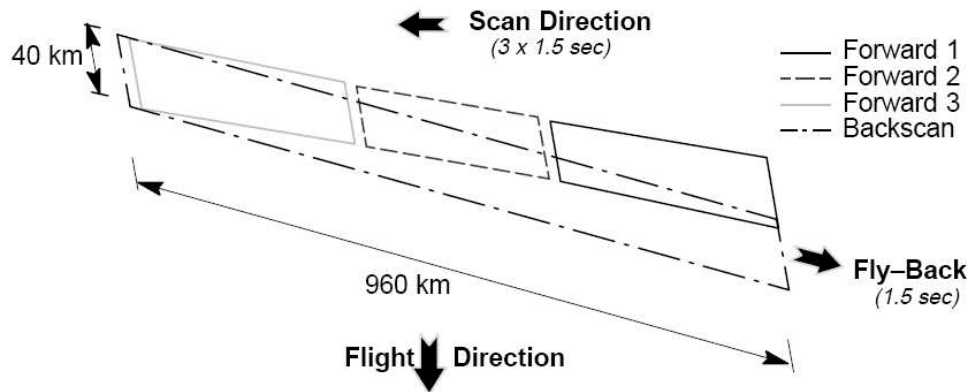
Figure 1.3: *Diagram illustrating the GOME scan across the ground pixels. Taken from Loyola et al. (2004).*

about the distribution of the trace gases need to be introduced. There are three main methods by which it is possible to correct for the GOME diffuser plate artefact (Richter and Wagner, 1998):

1. The solar spectrum can be replaced with a single irradiance measurement for a longer time series. The erroneous variations will no longer exist with this approach, however instrumental changes can introduce drifts in the time series. This method gives acceptable results for NO_2 at mid and high latitudes, but not for NO_2 at low latitudes and not for any one of the smaller absorbers (BrO, OClO, SO_2, HCHO).

2. A background spectra of Earth–shine radiance can be used instead of a solar spectrum. An unknown amount of absorption in the background spectrum is assumed and added to the derived columns. For OClO this is acceptable where high sun scenarios with negligible OClO can be found in every orbit, but not for other absorbers.

3. For some absorbers the assumption can be made that concentrations in equatorial regions are constant or well known from models. This is a reasonable assumption as the model can be validated and within a specific longitude range the value will be reasonably accurate. This can be compared to the slant columns to determine the size of the variation.

Method 3 is used in this study as the assumption here is the most valid and the model used to calculate the concentrations in the equatorial regions can be validated with ground-based observations. The application and impact of this diffuser plate artifact correction will be discussed in Chapter 6.

GOME Sensitivity to the Troposphere

The advantage of a backscatter instrument such as GOME is that (depending on the wavelength) some fraction of the photons that are observed by the instrument will have sampled the

troposphere. This can be shown by looking at the averaging kernels of the retrieval. This is investigated in Chapter 4, where Figure 4.8 shows the averaging kernels relevant to a retrieval of NO_2 at 430 nm for different surface reflectances. This figure shows that photons arriving at the satellite will sample the troposphere, however the signal is not specific to the troposphere and therefore the whole atmosphere needs to be considered when understanding the signal measured by the instrument.

1.2.2 Along Track Scanning Radiometer 2 (ATSR–2)

ATSR–2 is part of a series of instruments designed to measure sea-surface temperatures with an accuracy sufficient for climate science. The first instrument in this series was ATSR which is onboard ERS–1 (launched in 1991). A third instrument, the Advanced Along Track Scanning Radiometer (AATSR) is onboard ENVISAT (launched 2002).

The unique ability of these instruments is the dual viewing mode, as shown in Figure 1.4. A single scan sweeps around from nadir to a position about 900 km along the satellite track. Several minutes after acquiring a forward view of the scene the satellite will be directly above the scene and the nadir view is acquired. This enables two different atmospheric paths to be observed allowing rigorous atmospheric corrections to be calculated (Mutlow et al., 1999).
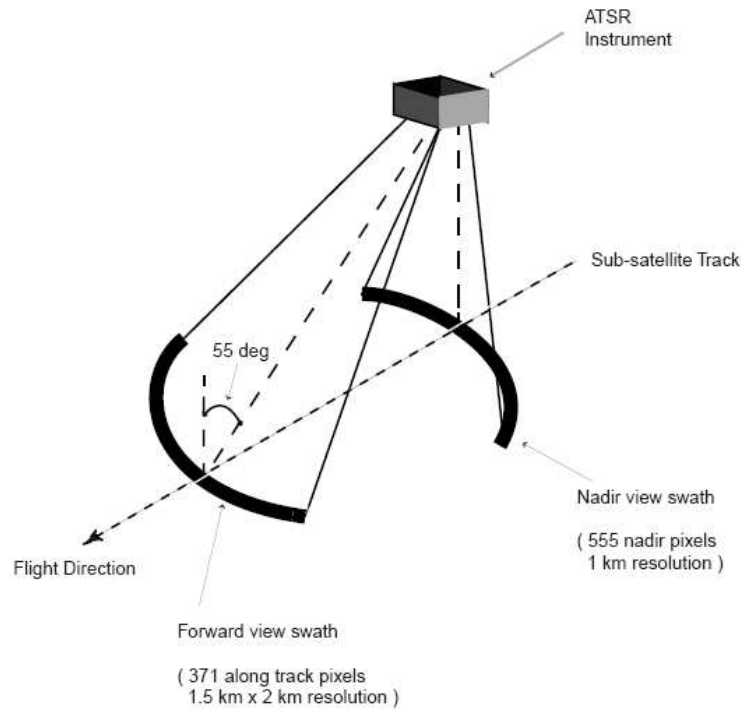
Figure 1.4: *A schematic to show the dual viewing model of the ATSR–2 instrument. Taken from Mutlow et al. (1999).*

ATSR–2 has two 500 km–wide curved swaths. The 'across nadir' swath has a resolution of 555 pixels and the 'forward swath' has a resolution of 371 pixels. It has 7 spectral channels each of

which are chosen to retrieve specific parameters from the atmosphere. These are summarised in Table 1.1.

Feature	Wavelength (μm)
Chlorophyll	0.55
Vegetation Index	0.67
Vegetation Index	0.87
Cloud Clearing	1.6
SST retrieval	3.7
SST retrieval	10.8
SST retrieval	12.0

Table 1.1: *The 7 different spectral channels on ATSR–2 along with the atmospheric features that are retrieved by these channels. Adapted from Mutlow et al. (1999).*

1.2.3   Combining GOME and ATSR–2 Data

A key aspect of this work is to exploit data from ATSR–2 in order to improve the accuracy of GOME retrievals. GOME and ATSR–2 are on the same platform and take measurements at the same location and time (see Figure 1.5). They can therefore be used simultaneously to improve the quality of the retrieval. Figure 1.5 shows the 0.55 μm channel ATSR–2 ground pixels overlaid on the GOME ground pixels. This shows that over land the ATSR–2 pixels only cover the central (nadir) GOME pixel but over the oceans the coverage is partial. The poor coverage over the oceans is due to an operational issue where the Synthetic Apperture Radar, or SAR, instrument (also onboard ERS-2) was allowed more bandwidth over the oceans and this was taken away from the ATSR–2 instrument in this spectral channel. The consequence of the small ATSR–2 ground scene is that global coverage when considering both instruments takes around 9 days.
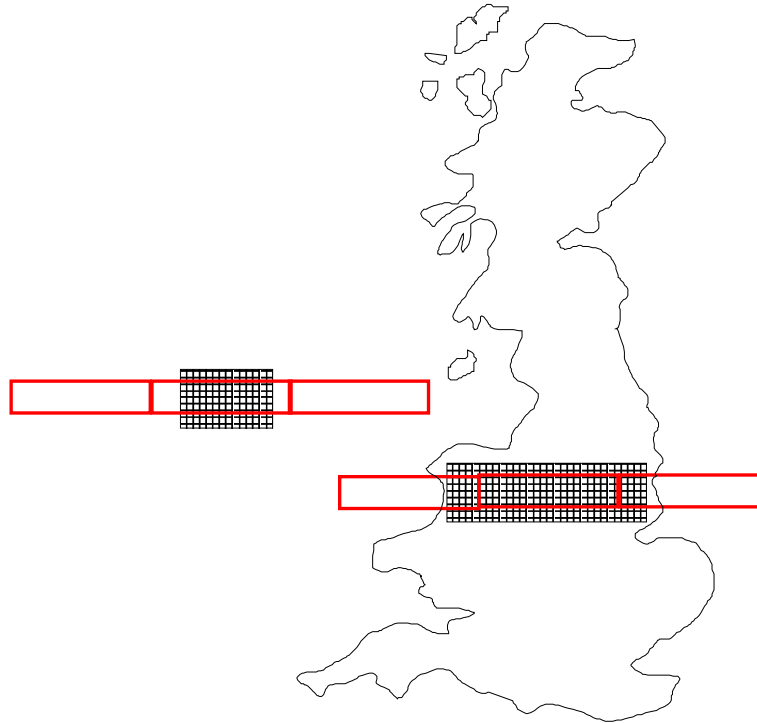
Figure 1.5: *An illustration to show the three GOME ground pixels (in red) overlaid with the array of smaller 0.5 µm channel ATSR–2 pixels (in black). Over land the ATSR–2 scene is larger and completely covers the GOME nadir (centre) pixel. However, over the oceans the ATSR–2 ground scene is smaller and only covers part of the GOME nadir pixel.*

1.2.4   MetOp

The retrieval scheme developed within this thesis could be applied to instruments on other satellites. An alternative would be the Meteorological Operational Satellite (MetOp). This is the first European polar orbiting satellite whose primary focus is operational meteorology (see `http://www.esa.int`). MetOp-A was launched in October 2008 and is the first of three satellites that will be launched every five years giving a continuous data record until 2020. The MetOp satellite has a payload of thirteen instruments including seven heritage instruments. These will record temperature and humidity, ocean surface wind speed and direction and O_3 and other trace gas concentrations. Among the instruments on-board MetOp are GOME-2 and the Advanced Very High Resolution Radiometer (AVHRR).

The GOME-2 instrument (Callies et al., 2000) is a follow on from GOME with a number of improvements. Firstly, the ground pixel is four times smaller at 80×40 km^2, secondly, global coverage is achieved every day (c.f. three days for GOME), and thirdly, the instrument has more information on the polarisation state due to better polarisation measurement devices. AVHRR measures reflected solar energy and radiated thermal energy from the land, sea and clouds at six spectral bands between 0.58-12.5 μm, this is similar to ATSR–2.

1.3   Atmospheric Chemistry

This section describes the atmospheric chemistry of the troposphere and stratosphere that is relevant to this work. This includes the chemistry of NO_x, BrO_x (NO_x and BrO_x are the generic terms for the mono–nitrogen and mono–bromine oxides, e.g. NO_x includes NO and NO_2) and O_3. The final part of this section discusses atmospheric chemical modelling.

1.3.1   Tropospheric Chemistry

A principal reason why tropospheric chemistry is important is because it is the area of the atmosphere in which we live. Many gases that are released by anthropogenic activity into the atmosphere remain in the lowest few kilometres and react to form other species which can be harmful to life. The following section will discuss the chemistry of the two species of interest in this thesis, NO_2 and BrO.

Nitrogen Dioxide

Nitrogen dioxide (NO_2) plays a key role in tropospheric chemistry. The photolysis of NO_2 is the primary production mechanism for O_3 (see Wayne, 2002).

$$NO_2 + h\nu \longrightarrow O + NO \qquad (1.1)$$

$$O + O_2 + M \longrightarrow O_3 + M \tag{1.2}$$

$$NO + RO_2 \longrightarrow NO_2 + RO \tag{1.3}$$

where R is a hydrocarbon. The rate determining step in this cycle (the nitrogen cycle) is the reaction 1.3. The nitrogen cycle, and therefore the abundance of tropospheric O_3, relies on sunlight (hv) and hydrocarbons (R). The abundance of O_3 is critically important in determining the oxidising capacity of the troposphere, because O_3 is the primary source of hydroxyl radicals (OH) and these dominate the chemistry of the troposphere during daytime.

$$O_3 + hv \longrightarrow O(^1D) + O_2(^1\Delta_g) \tag{1.4}$$

$$O(^1D) + H_2O \longrightarrow OH + OH \tag{1.5}$$

The O_3 budget in the upper troposphere is also affected by the transport of O_3 from the stratosphere.

There are a number of sources of NO_2 both natural and anthropogenic. The natural sources include forest fires and electric storms while the anthropogenic sources include high temperature combustion of oil, gas, and coal (see Wayne, 2002). NO_2 is a toxic gas and at high concentrations (40-100 $\mu g/m^3$) can cause lung problems. There have also been suggested links between NO_2 and childhood asthma (World Health Organisation Working Group, 2003).

In polluted atmospheres, i.e. where there are high concentrations of NO_2, a sink for OH is given by

$$NO_2 + OH + M \longrightarrow HNO_3 + M \tag{1.6}$$
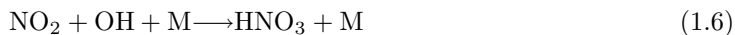
Nitric acid (HNO_3) can dissolve in clouds and be removed from the troposphere in the form of rain. This is an important removal process for NO_x.

There are two main reasons for measuring global distributions of NO_2. Firstly, to monitor the growth of NO_2 in the atmosphere, and secondly, to better understand and quantify its sources and sinks. The global distributions of NO_2 from surface measurements and in situ aircraft measurements are accurate but have poor spatial distribution (Singh and Jacob, 2000) due to limited global coverage. Satellite measurements can clearly provide global distributions.

Bromine Monoxide

Reactive halogen oxides such as BrO, ClO and IO, are known to play a role in the destruction of tropospheric O_3 (Barrie et al., 1988), which will lead to less OH radicals and consequently a reduction in the oxidising capacity of the troposphere. The tropospheric concentration of BrO is not well known and therefore more detailed, quantitative measurements are required in order

to better understand its contribution to tropospheric chemistry.

There are two catalytic chemical mechanisms that have been suggested for the destruction of tropospheric O_3 (Stutz, 2001).

Cycle A

$$Br + O_3 \longrightarrow BrO + O_2 \tag{1.7}$$

$$Y + O_3 \longrightarrow YO + O_2 \tag{1.8}$$

$$BrO + YO \longrightarrow Br + Y + O_2 \tag{1.9}$$

with the net reaction,

$$2O_3 \longrightarrow 3O_2 \tag{1.10}$$

Cycle B

$$BrO + HO_2 \longrightarrow HOBr + O_2 \tag{1.11}$$

$$HOBr + h\nu \longrightarrow Br + OH \tag{1.12}$$

$$OH + CO, \ O_3 \ or \ VOC \longrightarrow HO_2 + \ products \tag{1.13}$$

with the net reaction,

$$2O_3 \longrightarrow 3O_2 \tag{1.14}$$
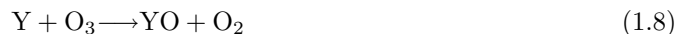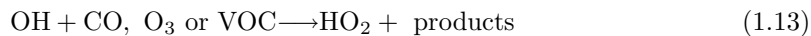
where Y is the halogen Br, I or Cl.

However, a recycling of BrO_x is necessary in order to maintain the high levels needed to account for the observed O_3 loss. The conversion to BrO_x from the relatively long-lived HBr achieves this. The production of HBr occurs by the reactions between Br and HCHO, or to a lesser extent, with other carbonyls.

$$Br + HCHO \ (+O_2) \longrightarrow HBr + CO + HO_2 \tag{1.15}$$
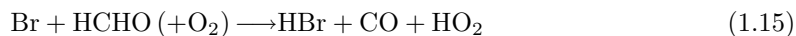
HBr dissolves in aerosols to release bromide (Br^-) in solution. Similarly, the other non-radical bromine reservoirs, $BrONO_2$ (produced from NO_2 + BrO) and HOBr (produced from HO_2 + BrO) are taken up by the aerosol, where $BrONO_2$ hydrolyses to HOBr. In acidic aerosols,

which are observed in the Arctic spring, HOBr reacts with Br^- to produce Br_2. This escapes into the gas phase and is photolysed (Evans et al., 2003).

$$Br^- + HOBr + H^+ \longrightarrow Br_2 + H_2O \qquad (1.16)$$

$$Br_2 + hv \longrightarrow 2Br \qquad (1.17)$$
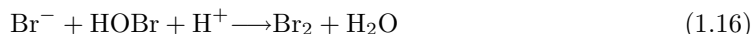
While the chemistry of tropospheric O_3 depletion is seemingly well understood, the amount of BrO needed to account for the observed O_3 depletion is harder to explain. The main processes that result in the release of reactive bromine are the chemical reactions that occur in the snow pack and on top of fresh (1 year old) sea–ice. When halides freeze they become concentrated on the surface of sea–ice and remain in a highly concentrated brine even down to temperatures of 230 K (Koop et al., 2000). Frost flowers (ice crystals) form from this highly concentrated brine and therefore are thought to be one of the more substantial sources of BrO.

There are three main ways in which this reactive bromine could be released into the atmosphere (Barrie and Platt, 1997). The depletion of halides from these sources is thought to be by emission into the atmosphere from the breakup of sea–ice in the late spring. The loss of salt from sea–ice can also be explained by rain as it washes the salt off the sea–ice releasing it into the atmosphere. Another explanation is that the salt is denser than the surrounding ice and therefore slowly sinks through the ice.

Other lesser sources of reactive Br gases are the reactions of biogenic organobromine compounds found in the marine atmosphere. Sturges et al. (1992) observed 10 pptv of biogenic organobromine compounds. On their own these emissions are too small to produce BrO concentrations large enough to affect O_3. Another source could be sea salt aerosol, but this also seems too small to account for the concentrations of BrO needed to explain the amount of O_3 depletion.

Evidently the tropospheric concentration of BrO is not well known, making quantitative global observations (which are available from satellites) key in improving our understanding.

1.3.2 Stratospheric Chemistry

The chemistry of the stratosphere is dominated by O_3. In the stratosphere O_3 is photolysed by UV radiation in the wavelength range of 200–310 nm, and therefore shelters the Earth's surface from radiation which is biologically active and can harm life. Also, the photolysis of O_3 leads to stratospheric heating which sets up the temperature structure of the atmosphere, which is primarily responsible for the radiative heating/cooling balance in the atmosphere (Wayne, 2002).

O_3 is primarily produced in the tropical stratosphere and is transported pole-ward by the large–scale Brewer Dobson circulation. It is continuously destroyed by the presence of catalysts (oxides of hydrogen, nitrogen, chlorine and bromine). These catalysts originate from natural

and anthropogenic tropospheric sources such as HO_x from natural sources of methane (CH_4) and water vapour (H_2O), NO_x from natural and anthropogenic N_2O, chlorine from anthropogenic chlorofluorocarbons (CFCs) and naturally produced methyl chloride and bromine from natural and anthropogenic sources of methyl bromide and halons. The catalytic efficiency of these species is determined by a set of coupled reactions which convert active oxides to catalytically inactive temporary reservoirs, e.g. HNO_3, HCl, $ClONO_2$, $HOCl$, $HOBr$, and $BrONO_2$. In the lower stratosphere the balance between catalytic oxides and temporary reservoirs is strongly affected by stratospheric aerosols, and also by polar stratospheric clouds (PSCs) as heterogeneous reactions occur on their surfaces.

Over the past 40 years stratospheric O_3 has been an active area of research, more recently spurred on by the discovery of the Antarctic O_3 hole (Solomon, 1999). Ozone loss in the Antarctic spring occurs after denitrification has occurred. In polar winter cold temperatures allow large particles to form on HNO_3 which then fall into the troposphere. With no nitrogen available to convert ClO_x to $ClONO_2$, O_3 destruction cannot be slowed and therefore the O_3 hole is formed.

Further stratospheric chemistry reactions relevant to NO_x partitioning will be discussed in Chapter 3.

1.3.3 Atmospheric Chemistry Modelling

Numerical models play an important role in our understanding of the atmosphere. Our present understanding of the current atmospheric state is advanced through model validation using observations from the real atmosphere. Models are a mathematical description of our current knowledge of different processes in the atmosphere and their interactions and impacts on parameters that we can observe. Using the best knowledge of how these parameters should evolve, a model can be compared directly with observations made of the atmosphere. Such comparisons allow validation of our understanding of atmospheric processes that govern the current state of the atmosphere and can also identify missing processes and deficiencies in our understanding and allow us to improve models accordingly.

A validated model, that is well understood and represents our best understanding of the atmosphere, can then be used to test hypotheses. The model allows an investigation into the atmospheric impact of the change in the rate of a processes or the introduction of a new process that is believed to be important.

Chemical transport models (CTMs) e.g. GEOS-chem (Bey et al., 2001), TOMCAT (Chipperfield, 2006), CLaMS (McKenna et al., 2002), are used to investigate the atmospheric processes responsible for the spatial and temporal concentrations of trace gases in the atmosphere. These distributions depend on the source strength of trace gases, the dynamic processes that transport air around the atmosphere and the chemical processes that create and destroy these gases. CTMs are 'off-line' models, i.e. the model winds are calculated separately from the chemistry-advection simulation itself. CTMs use winds and temperatures output from general circulation

models (GCMs) or from operational weather forecast models, to predict the advection of air in their model atmospheres. This approach is useful since the simulation is linked to real meteorology, with operational analysed winds incorporating assimilated global observations from radiosondes, surface sites and satellites. This also means that sensitivities to the chemical treatment can be investigated, with a 'best estimate' of the atmospheric dynamics for the time period of study. CTMs are also computationally cheaper to run than more complex GCMs, since expensive dynamical calculations do not need to be performed. The advection scheme in CTMs is driven by winds, which represent the atmospheric motion at the scale of the resolution of the model. For current CTM runs, this is typically of the order of $5° \times 5°$ or less. In CTMs it is therefore necessary to parameterise the processes that transport air on smaller scales. These are planetary boundary layer turbulence and convective transport.

1.4 Standard Retrieval Algorithms for GOME

This section discusses the standard retrieval algorithm that is usually used to derive trace gas total column amounts from the GOME instrument. Some of the steps in the algorithm differ between schemes but these will be highlighted.

1.4.1 Differential Optical Absorption Spectroscopy

Differential Optical Absorption Spectroscopy (DOAS) is a method used to determine concentrations of trace gases by measuring their specific narrow band absorption structures (see Perner and Platt, 1979; Platt, 1994). The absorbers dealt with in this thesis are optically thin. The differential absorption of trace gases is based on Beer's law (Liou, 2002) where the incremental increase of radiance $dR(\lambda)$, at wavelength λ through a slant path distance ds is proportional to the coefficient of absorption $\sigma(\lambda)$ times the incident radiance $R(\lambda)$ and the absorber column amount $c(s)ds$:

$$dR(\lambda) = -R(\lambda)\sigma(\lambda)c(s)ds \tag{1.18}$$

A wavelength window is chosen for the retrieval such that absorption and scattering by other trace gases, and surface reflectance affects are weakly dependent on wavelength. Therefore,

$$R(\lambda) = \frac{R_0(\lambda)r(\lambda)}{\pi} \exp(-\tau(\lambda)) \tag{1.19}$$

where $r(\lambda)$ is the sun–normalised radiance and $R_0(\lambda)r(\lambda)/\pi$ is the radiance that would be observed in the absence of the minor absorber. R_0 is assumed to be weakly wavelength dependent such that it can be fitted by a low order polynomial, and τ is the optical depth of the minor absorber,

14

$$\tau(\lambda) = \int_0^s \sigma(s, \lambda)c(s)ds \qquad (1.20)$$

where σ is the cross section which is independent of position along the path and c is the concentration of the minor trace gas. The DOAS measurement can be defined by

$$y(\lambda) = \ln(r(\lambda)) = \ln(R_0) - \tau(\lambda) + \epsilon(\lambda) \qquad (1.21)$$

where $y(\lambda)$ is the measurement, r is the sun–normalised radiance $(\pi R(\lambda)/R_0(\lambda))$ and $\epsilon(\lambda)$ is the measurement error.

$$y(\lambda) = \ln(R_0) - \sigma(\lambda) \int_0^s c(s)ds + \epsilon(\lambda) \qquad (1.22)$$

$$y(\lambda) = \ln(R_0) - \sigma(\lambda)S + \epsilon(\lambda) \qquad (1.23)$$

where S is the slant column density. Note that S is independent of wavelength and $\sigma(\lambda)$ is constant (independent of scene and path). S can be estimated by a least-squares fit of $y(\lambda)$ using known (laboratory measured) absorption cross sections with errors $\epsilon(\lambda)$. A polynomial of low order is also fitted to deal with the $\ln(R_0)$ term.

Now, if we consider S:

$$S = \int_0^s c(s)ds \qquad (1.24)$$

For scattered light the path s is different for each photon and the slant column is

$$S = \frac{1}{N_p} \sum_{i=0}^{N_p} \int_0^{s_i} c(s_i)ds_i \qquad (1.25)$$

where i indicates the ith photon of the number of photons (N_p) which in the absence of the absorption would be observed by the sensor. For each photon the vertical coordinate can be converted to z (altitude), and the sum can be taken inside the integral as it is no longer dependent on i, then

$$Sdz = \int_0^z c(z)dz \frac{1}{N_p} \sum_{i=0}^{N_p} ds_i \qquad (1.26)$$

which simplifies to

$$l(z) = \frac{1}{N_p} \left(\sum_{i=0}^{N_p} l_i(z) \right) \tag{1.27}$$

where $l_i(z)$ is $ds_i(s)/dz$ and ds_i is the total distance travelled by the photon i at any stage during the scattered path, through an infinitesimally thin layer (dz) about z. $l(z)$ is the mean photon path length per unit vertical distance (km/km), considering all photons which would reach the sensor in the absence of absorption by the trace gas. This can be calculated using a radiative transfer model, assuming the scattering and absorption properties of the atmosphere in the absence of the scatterer can be defined.

The vertical column density V is

$$V = \int_0^z c(z) dz \tag{1.28}$$

In a standard DOAS retrieval the ratio of $S/V = AMF$ where AMF is the air-mass factor and is computed using a radiative transfer model and deals with scattering in the atmosphere and absorption by species other than the target trace gas. This depends on profile shape of the minor trace gas but not the absolute amount, since a constant factor cancels in the ratio. The total column can be estimated by:

$$\text{Vertical Column Density} = \frac{\text{Slant Column Density}}{\text{Air-Mass Factor}} \tag{1.29}$$

In the absence of atmospheric scattering a geometric AMF (AMF_G) can be defined which is a simple function of the solar zenith angle θ_S and the satellite viewing angle θ_V:

$$\text{AMF}_G = \sec\theta_S + \sec\theta_V \tag{1.30}$$

In the UV/VIS region of the spectrum Rayleigh and Mie scattering interfere with the geometric viewing of the atmosphere. Therefore cloud, aerosol and surface parameters need to be incorporated in radiative transfer equations in order to determine the true AMF (Palmer et al., 2001).

1.4.2 Air Mass Factor Calculations

The AMF calculation stage of the retrieval involves running a radiative transfer model to calculate the Rayleigh and Mie scattering that occurred in the atmosphere at the time of the measurement. Previous retrieval algorithms differ in how they represent these scattering processes.

Clouds

Most previous retrieval studies for GOME use cloud parameters from one of the GOME cloud retrieval algorithms. These use the normalised spectra of the oxygen A-band for clear and cloudy pixels. The most commonly used GOME cloud datasets are the Fast Retrieval Scheme for Cloud Observables (FRESCO) (see Koelemeijer et al., 2001), and the Cloud Retrieval Algorithm for GOME (CRAG) (see Von Bargen et al., 1998).

The FRESCO algorithm retrieves the cloud top pressures and the effective cloud fractions from calibrated level-1 reflectivity data from the GOME spectrum of the oxygen A-band (between 758-775 nm) (Van der A and Wang, 2007). Three 1-nm-wide parts of the oxygen A-band spectra are used in the FRESCO retrieval at 758 nm (no absorption), 761 nm (strong absorption) and 765 nm (moderate absorption). Assuming the prior knowledge that reflectivity outside the oxygen A-band is almost independent of the cloud top pressure and is mainly dependent on the cloud fraction, cloud optical thickness and surface albedo, the reflectivities inside the bands are used to derive cloud pressure. An effective cloud fraction and cloud pressure are derived for each GOME pixel using the non-linear least-squares fitting of a measured spectrum to a simulated spectrum. Cloud albedo is assumed to be 0.8 and is used to determine cloud fraction. This is, therefore, an effective cloud fraction.

The CRAG algorithm has two stages. The first stage is to determine the cloud cover using a Polarisation Measurement Device (PMD) algorithm. The basic idea of the PMD algorithm is that while the Earth's surface reflects light with a strong spectral surface-dependent signature, clouds are almost perfect scatterers in the visible region. There are different ways to use this information. For example, the PMD Cloud Recognition Algorithm (PCRA) uses the three spectral regions of GOME, blue, green and red, and uses a flexible threshold technique depending on the underlying Earth surface. The Optical Cloud Recognition Algorithm (OCRA) determines the fractional cloud cover based on a comparison of the reflectances of each PMD band with the reflectances of a pre-determined cloud–free composite.

Figure 1.6 shows a hurricane observed by NOAA's GEOS-8 satellite and the same hurricane as observed by GOME using the PCRA. The figure shows that the GOME algorithm can capture the cloud variation in the system well, even at the coarser resolution of the GOME instrument.

The second stage of CRAG is to use the Revised Cloud Fitting Algorithm (RCFA). This part of the algorithm is similar to FRESCO as it uses radiances in and around the oxygen A-band to determine the cloud top height and the cloud optical thickness.

The use of these cloud parameters also varies between the schemes. The most accurate AMF calculations incorporate the cloud parameters in the radiative transfer model and simulate the radiative transfer through the cloud (Martin et al., 2002). A number of retrieval schemes do not run radiative transfer calculations in the cloud. Instead these schemes assume that a cloud has an albedo of 0.8 and run the radiative transfer model with the surface at cloud top height with a reflectance of 0.8. This is known as cloud weighting and is used in the Tropospheric
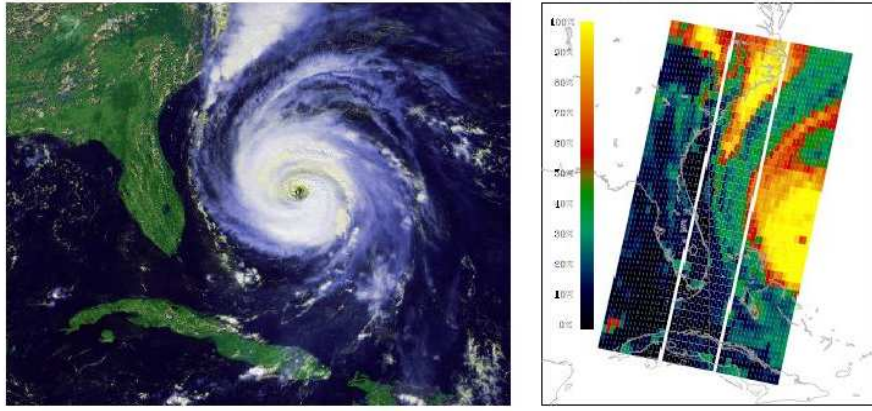
Figure 1.6: *Hurricane FRAN as seen by the NOAA's GOES–8 satellite at 17:15 UTC on 4 September 1996 (left image) and as detected by GOME measurements at 16:04 UTC using the PCRA (right image). Taken from Von Bargen et al. (1998).*

Emission Monitoring Internet Service retrieval[1] (TEMIS) (Wang et al., 2008). The use of a 1–D radiative transfer model to simulate the radiative transfer is a more sophisticated method for determining the light path through the atmosphere. In the ideal situation more dimensions would be used to calculate the radiative transfer in the atmosphere, but this is not possible without a 3–D version of the instrument.

Aerosols

Information about the aerosol in the scene is usually obtained from look–up tables and climatologies. These contain aerosol properties that have been recorded in projects such as the Aerosol Robotic Network (AERONET) or the Optical Properties of Aerosols and Clouds (OPAC) (Hess et al., 1998). Most studies use a number of general aerosol types: Urban-industrial, biomass, desert dust and marine aerosol. Aerosol models are used to predict the location of aerosols and these aerosol properties can be used in radiative transfer calculations to obtain the AMF.

Surface Reflectance

Surface reflectance can be retrieved from GOME by combining the radiance measurements from different wavelengths. Surface reflectance is extremely important in AMF calculations as higher reflectance will cause larger AMFs (Boersma et al., 2004) and more multiple scattering. Richter and Burrows (2001) used a few months of clear sky observations to create a map of surface reflectance, then applied this to the retrieval algorithm. Most studies use the surface reflectance dataset developed at KNMI (Koelemeijer et al., 2003b) which calculated the average Minimum Lambertian Equivalent Radiance (MLER) for a number of wavelengths from five years of GOME spectral radiances. As surface reflectance is dependent on surface type the surface is assumed

---

[1]http://www.temis.nl

to be lambertian in nature (isotropic) and therefore a Lambertian Equivalent Reflectance is fitted in place of surface reflectance. The MLER is retrieved because only Rayleigh scattering is considered. Therefore the influence of aerosol is not account therefore making the reflectance a minimum value.

1.4.3 Stratospheric Component

The DOAS method is used to find the total vertical column of an absorber. To obtain the tropospheric vertical column the stratospheric vertical column needs to be subtracted. There are many different methods used to achieve this and they are explained in Section 1.4.3. The calculation of the stratospheric component of a slant column can be carried out in a number of ways, differing between species.

NO_2 stratospheric column calculation

There are four main methods that have been used to calculate the stratospheric NO_2 column (Boersma et al., 2004).

1. The image processing technique by Leue et al. (2001). Initially land masses and cloud–free pixels are masked out leaving only cloudy pixels. Then it is assumed (depending on the amount of cloud in the pixel) that the NO_2 in these pixels is purely stratospheric i.e. the cloud obscures the NO_2 beneath it from the instrument. An image processing filter is then used to interpolate across the gaps. This process for contains a number of errors. Firstly, the assumption that the NO_2 column above a cloud is purely stratospheric does not always hold as lightning is a significant source of tropospheric NO_2 (Egorova et al., 1999) and cloudy pixels in the tropics are likely to be caused by thunderstorms. Secondly, the presence of the cloud will artificially enhance the NO_2 column that is observed. If the cloud is not modelled correctly the observed NO_2 can easily be in error.

2. The reference sector method used by Richter and Burrows (2001); Martin et al. (2002). Over the remote Pacific Ocean it is assumed that there is no tropospheric NO_2, as NO_2 is mostly anthropogenic or caused by biomass burning. Therefore, the column here can be taken as purely stratospheric and independent of longitude. This improves upon the image processing technique as the assumption is more valid, however, it still assumes that the stratospheric NO_2 is constant with longitude. This is a reasonable assumption in equatorial regions and mid-latitudes as the stratospheric NO_2 is dependent on photolysis and the amount of sunlight is constant with longitude. However, stratospheric dynamics can still cause some variability with longitude. In and near the polar vortex this assumption does not hold as there is more turbulent atmospheric motion. Martin et al. (2002) improved this method by assuming that the reference sectors were not completely stratospheric and used a tropospheric CTM to calculate the small tropospheric column. This did not, however, address the issue with zonal asymmetries in the stratosphere.

3. The use of stratospheric CTM by Richter and Burrows (2001). This method uses a stratospheric CTM to calculate the stratospheric column. The CTM uses known chemistry of the stratosphere to determine the NO_2 stratospheric column. This method has the advantage of taking into account stratospheric dynamics but the retrieval will be dependent on errors in the CTM.

4. The use of chemical data assimilation within a CTM used by Van der A et al. (2006). This method takes an NO_2 profile from a chemistry transport model then, using averaging kernels (see Section 1.2.1), the GOME NO_2 is added to the stratospheric part of the profile. The drawbacks with this method are that it uses GOME NO_2 measurements to constrain GOME NO_2 measurements and therefore does not use independent data. This can introduce systematic biases to the results.

The new retrieval algorithm developed in the thesis also uses an assimilation scheme to constrain the stratosphere but here it is done by assimilating long-lived tracers rather than the short-lived NO_2 (see Chapter 3). Table 1.2 is taken from Boersma et al. (2004), and summarises the errors associated with a number of different stratospheric NO_2 vertical column density calculations from different methods.

Authors	Method	Uncertainty in VCD_{trop}
Leue et al. (2001)	IPF	0.08–0.5 $\times 10^{15}$ molec. cm^{-2}
Richter and Burrows (2002)	Ref. sector	1 $\times 10^{15}$ molec. cm^{-2}
Martin et al (2002)	Ref. sector	0.2 $\times 10^{15}$ molec. cm^{-2}
Richter (2002)	CTM Ref.	(not given)
Boersma et al (2002)	Ref sector	0.45 $\times 10^{15}$ molec. cm^{-2}
Boersma et al (2004)	assimilation	0.2 $\times 10^{15}$ molec. cm^{-2}

Table 1.2: *Error in retrieved VCD_{trop} caused by the error in stratospheric VCD (taken from Boersma et al., 2004).*

The more sophisticated methods for calculating the stratospheric NO_2 vertical column density are those used by Martin et al. (2002) and Boersma et al. (2004). Martin et al. (2002) used the reference sector method combined with measurements from a CTM to account for any possible NO_2 over the remote areas of the Pacific. This enabled a more accurate calculation than Richter and Burrows (2001) where the reference sector method was used on its own. The method of Boersma et al. (2004) also reproduces stratospheric NO_2 to a high accuracy. However, the use of GOME NO_2 to constrain the model can bias the results. A better method would be to use NO_2 from a different instrument, with specific sensitivity to the stratosphere, to the constrain the stratosphere.

BrO Stratospheric Column Calculation

The chemistry of stratospheric BrO is fairly well understood as it is an important species in stratospheric O_3 depletion. There are a number of studies that have investigated its spatial distribution and concentration (e.g. Hegels et al., 1998; Wagner and Platt, 1998).

There are two main methods for calculating the BrO stratospheric column used in previous studies.

1. A simple stratospheric profile is assumed. The mixing ratio below 10 km is assumed to be 0 then a linear increase is imposed in the region 10-30 km, above which a constant mixing ratio is assumed. Mixing ratio concentrations above 30 km are around 12 ppt (Van Roozendael et al., 1999).

2. A method similar to the reference sector method for NO_2. As BrO is known to be roughly constant with longitude, a location where there is thought to be no tropospheric BrO can be used to determine the stratospheric column (the whole column) and this can be removed from the whole longitudinal range to leave the tropospheric column (Richter et al., 1998).

There are problems with both of these methods. Firstly, the BrO profile is not uniform so assuming a linearly increasing profile is potentially inaccurate. Secondly, there is increasing evidence that there is an omnipresence of BrO in the free troposphere (e.g. Evans et al., 2003). Therefore, there is no location on the globe where the slant column can be assumed to be purely stratospheric.

1.5 Scientific Application of GOME Tropospheric Data

1.5.1 Retrievals of Nitrogen Dioxide

There have been three main areas of interest involving retrieved total tropospheric NO_2 columns from GOME. The first concerns the enhancement of tropospheric NO_2 due to anthropogenic activity or biomass burning. Richter and Burrows (2001) looked at a case study for the summer of 1997 over Africa and investigated NO_2 enhancements observed by GOME which were thought to be due to biomass burning events. By comparing with ATSR-2 (the Along Track Scanning Radiometer collocated with GOME) fire counts they were able to distinguish between anthropogenic and biomass burning events for clear sky pixels. Then, by looking at cloudy pixels, they found NO_2 enhancements which they attributed to lightning–induced NO_2. Although there are some interesting results in this study, the results are qualitative rather than quantitative. The errors in the measurements are discussed and are stated to be due to the uncertainty in the NO_2 fit, uncertainty in the separation of the troposphere and the stratosphere, uncertainty in the light path and uncertainty in cloud cover. They state that their results are a lower estimate on the tropospheric NO_2 enhancement with an error of a factor of 2. There are a lot of shortcomings in this study. Firstly, although only clear pixels are used to identify NO_2 produced from biomass burning there will be a large amount of aerosol present in the plumes that are associated with the fires and this should be accounted for when calculating the light path. Secondly, for the cloudy pixels the assumption is made that the light does not penetrate into the lower troposphere. However, this is not always true (see Chapter 4) and the radiative

transfer within and below the cloud needs to be consider in order to obtain a more accurate NO_2 abundance due to lightning.

China and east Asia is another location where NO_2 enhancements are observed due to anthropogenic activity. There have been some important GOME studies which investigate the seasonal trends in these NO_2 enhancements. Richter et al. (2005) used retrieved NO_2 tropospheric columns from GOME and SCIAMACHY for 1996–2004 to look at the trends in anthropogenic NO_2. They observed a decrease in concentrations over some parts of Europe and the USA, but a highly significant increase of about 50% with an accelerating trend in annual growth rate, over the industrial areas of China. Their results showed a larger increase than those of bottom-up inventories (Streets, 2003). Although this study produced some interesting results as in their previous work (Richter and Burrows, 2001), their results are not well quantified and therefore could be significantly improved if the retrieval errors were known. Van der A et al. (2006) also investigated the trend and seasonal variation of NO_2 over China. Their retrieval scheme used a chemical data assimilation scheme within an atmospheric CTM to calculate the stratospheric NO_2 column. A cloud weighting scheme (described in Section 1.4.2) allowed the use of cloudy GOME pixels. They found a yearly increase in NO_2 over Shanghai of $20 \pm 6\%$ per year (1996 – 2005). The seasonal pattern varies between east and west China with a winter maximum found in eastern China, as expected due to anthropogenic activity. The maximum in western China occurred in summertime indicating natural emissions, which was expected due to the lack of anthropogenic activity there. The study by Van der A et al. (2006) did make some quantitative estimates of NO_2 trends. However, the cloud weighting scheme used in their retrieval introduced errors that could be removed by including radiative transfer within clouds.

The second area of interest is the comparison of GOME measurements to model calculations. Figure 1.7 is taken from Martin et al. (2002) and shows the comparisons between GOME and the GEOS-CHEM 3-D model. They showed that GOME NO_2 over the United States is 18% higher than GEOS–CHEM, although this is within the 35% error expected from the model. However, over South Africa, industrial USA and Europe, GOME columns are significantly higher than GEOS–CHEM while over the southern USA, Eastern Asia and biomass burning, areas GOME columns are significantly lower. Martin et al. (2002) attribute this to the aerosol scheme used. Uno et al. (2006) used the Community Multi Scale Air Quality (CMAQ) CTM with a year-by-year emissions inventory (REAS) to compare to the GOME retrieved NO_2 over eastern Asia. They show that the model underestimates GOME columns by 2-4% over polluted industrial regions (central eastern China, Korea and Hong Kong). Over Japan GOME and CMAQ show good agreement with respect to interannual variation with no increasing trend. Over China they show a 10-11% increase per year from GOME and 8-9% increase in CMAQ. The retrieval scheme used in Uno et al. (2006) is one of the better NO_2 schemes as it uses the full cloud information from the CRAG GOME cloud algorithm in order to calculate more accurate air mass factors. However, GOME cloud retrievals have been shown (see Figure 1.13) to not be as accurate as those from ATSR-2 and therefore the new retrieval algorithm developed in this thesis should be more accurate.

Schaub et al. (2006) used ground-based observations from instruments situated in the Alps to
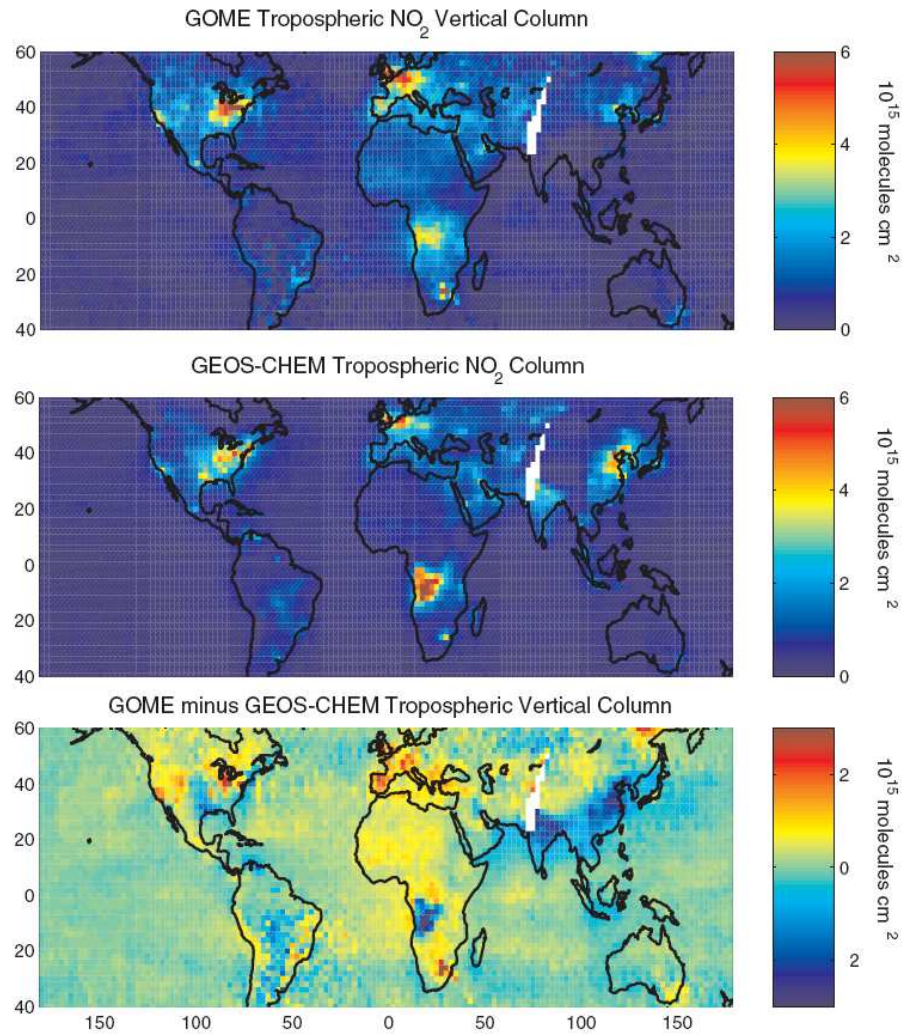
Figure 1.7: *Mean tropospheric NO₂ vertical columns (molecules cm⁻²) from GOME for July 1996 taken from (Martin et al., 2002). The top panel shows the GOME tropospheric NO₂ column, the middle panel shows the GEOS-CHEM modelled tropospheric NO₂ column and the bottom panel shows the difference between the two.*

compare to GOME NO_2. They showed that GOME overestimates the ground based observations of NO_2. Velders et al. (2001) used the MOZART and IMAGES CTMs to compare to GOME NO_2 columns and their results are summarised in Table 1.3.

These three studies all show that the GOME NO_2 retrieval algorithms calculate the spatial distribution of NO_2 well, however, the magnitude of the NO_2 column is less accurate. This is likely to be due to two reasons. Firstly, the stratospheric columns are not accurate enough, and secondly, the treatment of clouds is not done to a high degree of accuracy. Both of these cause inaccuracies in the retrieved amounts of NO_2.

REGION	MOZART	IMAGES
USA	0.18 – 0.15 smaller	2 – 3 smaller
Europe	0.7 – 2 smaller	1.5 – 4 smaller
Asia	1.5 – 3.5 smaller	3 – 5 smaller

Table 1.3: *Summary of GOME NO_2 comparison to two CTMs (Velders et al., 2001).*

The third area of interest has been to use GOME NO_2 to infer the NO_2 emissions and lightning sources. Boersma et al. (2005) quantified the source strength and 3-D distribution of lightning-produced nitrogen oxides and estimated a global NO_x production in the range of 1.1 – 6.4 Tg [N] in 1997. Beirle et al. (2006) looked at a large convective system over the Gulf of Mexico and with the knowledge of the number of cloud-ground flashes calculated a value of 1.3 kg [N] per flash of NO_x. This led to a global estimate of 1.7 (0.6–4.7) Tg [N] per year. It is important when carrying out NO_2 retrievals from an instrument such as GOME to be able to identify NO_x in the upper troposphere from lightning production. An estimate of the tropospheric profile of NO_2 is necessary in calculating the AMF. Boersma et al. (2005) uses a quantitative retrieval algorithm however, they use cloud weighting which will introduce significant errors into their retrieval.

NO_2 retrievals and Associated Errors

The first stage of an NO_2 retrieval is to determine the slant column density (unlike other retrieval schemes this new scheme in this thesis starts with previously retrieved slant columns). Figure 1.8 shows the NO_2 slant column density reported by the GOME data product (GDP)[2] and the associated errors. Figure 1.8(a) shows that slant column density is highly dependent on solar zenith angle. The figure shows this as the slant column density varies less with longitude than latitude. At the poles values reach $\sim 60 \times 10^{15}$ molecules cm^{-2} decreasing to $\sim 3 \times 10^{15}$ molecules cm^{-2} at the equator.

Figure 1.8(b) shows that the error in the slant column density. The error reported in this figure is most due to photon noise and is the random error on the slant column. Systematic errors in the slant column density are due to cross section errors, spectral calibration, temperature and the solar diffuser plate (Boersma et al., 2004). The solar diffuser plate error is large and is not

---

[2]The GOME data product is the data from GOME that has been determined by the European Space Agency (ESA). For further information is `http://earth.esa.int`
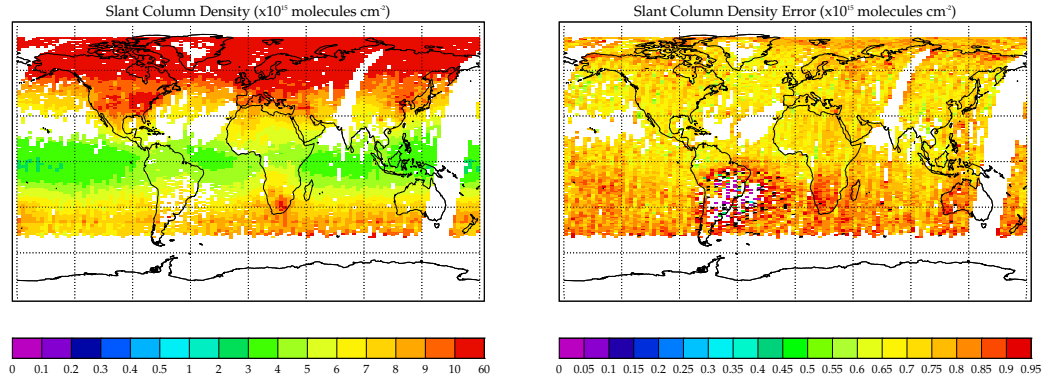
Figure 1.8: *Global total slant column NO$_2$ for July 1997 as retrieved from the GOME instrument and reported by the GOME data product (GDP) (a) total slant column density (1×10^{15} molecules cm^{-2}) (b) error in the slant column density (1×10^{15} molecules cm^{-2}).*
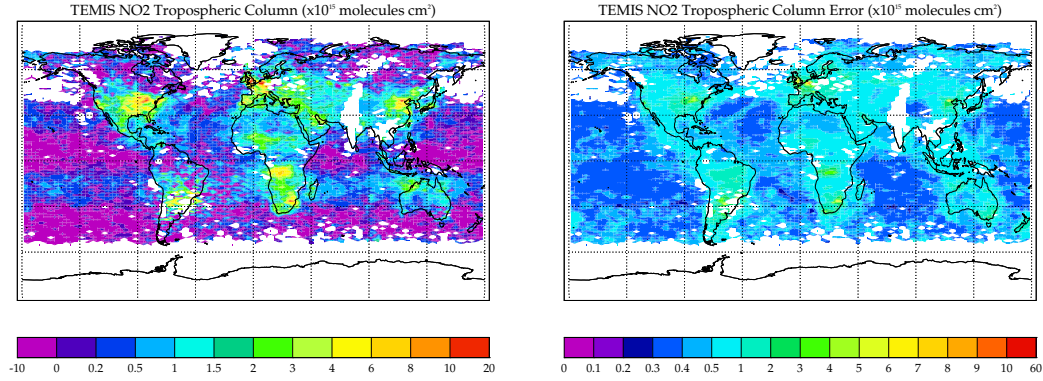
Figure 1.9: *Global total tropospheric vertical column NO$_2$ for July 1997 as retrieved from the GOME instrument by the Tropospheric Emission Monitoring Internet Service (TEMIS) part of the Dutch Royal Meteorological Department (KNMI) (a) tropospheric vertical column density (1 × 10^{15} molecules cm^{-2}) (b) error in the tropospheric column density (1 × 10^{15} molecules cm^{-2}).*

included in the value plotted here. Large errors off the east coast of South America are due to the Southern Anomaly. This is caused by interference with from the Van Allen Belt. Other systematic errors are relatively small.

Figure 1.9 shows the total tropospheric NO$_2$ vertical column retrieved from GOME by KNMI and the error. The total tropospheric NO$_2$ shows maximum values of 2×10^{16} molecules cm^{-2} in industrial areas (i.e. Western Europe, Eastern USA, California and China) and biomass burning plumes show enhancements of 4×10^{15} molecules cm^{-2}. The figure also shows a reported retrieval error of $\pm$ 2-5 $\times 10^{14}$ molecules cm^{-2} over the remote areas and $\pm$ 5-10 $\times$ 10^{15} molecules cm^{-2} over areas of enhanced NO$_2$. This is roughly an error of close to 100%.

1.5.2 Bromine Monoxide

A key area of BrO research in the troposphere is the observation of enhanced BrO around sea ice. The first estimates of BrO from satellite observations (Hegels et al., 1998) described the slant columns seen from GOME. An omnipresence of stratospheric BrO was observed suggesting a significant sink of stratospheric O_3. However, Chance (1998) used geometric air mass factors to obtain vertical column densities, only to realise that there was too much BrO present in the column for it all to be stratospheric. This suggested that there was a significant source at some locations in the troposphere. Further calculations were then performed using radiative transfer schemes to show that the estimates of BrO in the stratosphere were accurate above 15 km, although the method of using geometric airmass factors underestimated the tropospheric column significantly except in areas of low surface albedo and high solar zenith angle. This is due to poor penetration of the light at these wavelengths.

The next stage of this research was to determine the temporal and spatial extent of these BrO enhancements. Wagner and Platt (1998) carried out an experiment using two AMFs, one for the stratospheric part of the column (assuming a maximum BrO concentration at 17 km in altitude) and one for the tropospheric part (assuming a constant mixing ratio concentration between 0-1 km). The study attributed a 30% error to the column, due to errors in the cross section, AMF and in the modelled BrO absorption. The results of this study showed that tropospheric air masses enriched with BrO are always situated close to sea ice and typically extend over areas of about 300-2000 km lasting up to 1 to 3 days. Richter et al. (1998) carried out a complementary experiment using a number of different methods. A similar method to Wagner and Platt (1998) was used where an AMF was calculated using a simple stratospheric profile with a linear increase of mixing ratios from 10-30 km and constant above. This lead to the underestimation of the tropospheric enhancements.

Figure 1.10 shows 3–day composites of tropospheric BrO vertical columns for February to July 1997, taken from Richter et al. (1998). The figure shows the enhancements in BrO that were observed over the Hudson bay from the middle of February until the middle of May (these include interruptions). These enhancements were determined to be tropospheric columns due to the location (if they had been stratospheric they should have been mixed due to the vortex), the larger values were observed on clear days when the lower troposphere was more visible and the enhancements coincided with similar measurements observed from ground-based measurements observed at Ny Alesund.

Hollwedel et al. (2004) looked at the year-to-year variation of springtime polar tropospheric BrO from GOME. Their results showed that there was a difference in spatial and temporal variation and magnitude between the hemisphere, which can be explained due to land and ocean distribution. Their results also showed the spatial distribution is similar from year-to-year although magnitudes varied. There was, however, a 10% increase since 1996–2001, this could be due to there being less sea-ice and more one year old sea-ice due to a warmer climate. The results also show an increase in the tropospheric background BrO. These studies showed how the GOME BrO retrievals have been used to study tropospheric BrO enhancement due
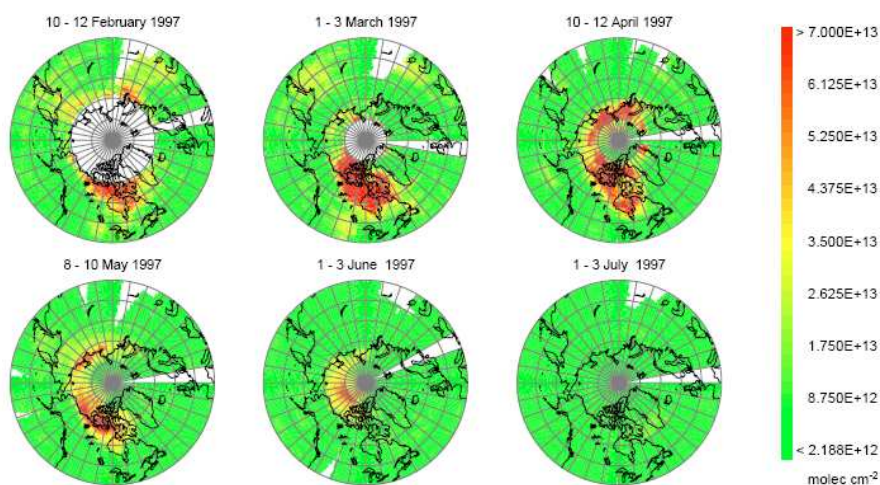
Figure 1.10: *Selected 3–day composites of tropospheric BrO from GOME. Values given are excess tropospheric vertical columns (molecules cm^{-2}) derived from the measurements. Taken from Richter et al. (1998).*

to the break up of sea ice. However, the retrieval schemes rely on numerous assumptions and therefore the results showed the spatial variation in these enhancements but not the magnitude. A more detailed retrieval scheme, with better assumptions about the BrO tropospheric profile, should lead to a more accurate estimate of the magnitude of tropospheric BrO in these cases.

The second area of research involving tropospheric BrO is the evidence suggesting an omnipresence of BrO in the free troposphere. Van Roozendael et al. (2002) compared GOME BrO measurements to ground-based and balloon observations and modelled BrO. Initial results showed that the balloon and the SLIMCAT 3D CTM vertical column results differed by 1-3 $\times 10^{13}$ molecules cm^{-2}, which was larger than the estimated error. This was suggested to be tropospheric BrO as balloon measurements have a low sensitivity and SLIMCAT assumed no tropospheric BrO. Although these two reasons are valid they rely on the quality of the retrieval technique for GOME. Further experiments were carried out by comparing zenith sky measurements. These included Langley plots, where a linear relationship between air mass factor and slant column is assumed and the vertical column is derived from the slope of the correlation curve. These results showed that the measurements taken at twilight agree better with GOME measurements than those taken at noon. This is another piece of evidence for tropospheric BrO, as twilight measurements are more sensitive to the troposphere. This study achieved a high degree of accuracy due to the combinations of a number of different measurements. Although it only covers a small time period it demonstrates the potential over the GOME observations if they are used in conjunction with other measurements.

BrO retrievals and Associated Errors

Unlike NO_2 retrievals, there are no truly quantitative retrieval algorithms that have been developed for BrO. Some studies, such as Wagner and Platt (1998), quantify errors due to slant

27
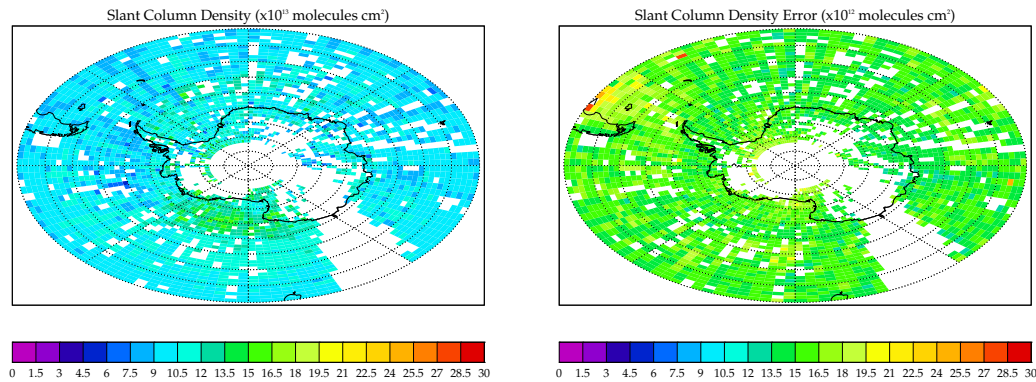
Figure 1.11: *Southern hemisphere slant column density BrO for November 1996 as retrieved from the GOME instrument by the Smithsonian Astrophysics Observatory Group (SAO) (data obtained from Dr Kelly Chance) (a) slant column density (1×10^{13} molecules cm^{-2}) (b) error in the slant column density (1×10^{12} molecules cm^{-2}).*

column retrieval but as most BrO retrievals are performed for clear sky pixels only and make gross assumptions about the stratospheric loading, it would be difficult to quantify any errors here and the retrievals are therefore qualitative in nature. Figure 1.11 shows the slant column density errors for the SAO BrO product which is used in this retrieval model.

The errors come from errors similar sources to those for NO_2 slant columns, cross section errors, spectral calibration errors and temperature errors. The BrO slant column errors are around 50% of the slant column which is a lot smaller than those reported for NO_2. Figure 1.12 shows the BrO tropospheric column for the southern hemisphere for November 1996 from the Bremen group (data from Dr Andreas Richter). This figure has better coverage than the slant column density data that will be used to run this retrieval model. However, this group do not report errors in their retrieval, but do state that this is not a quantitative retrieval scheme.

The new retrieval scheme developed in this thesis will aim to better the retrieval of BrO by approaching the problem in a more sophisticated manner. This should significantly improve upon previous BrO retrievals from GOME.

1.6 GRAPE Retrievals

The new retrieval scheme developed in this work uses cloud and aerosol data from the ATSR–2 instrument which will come from the Global Retrieval of ATSR Parameters and Evaluation (GRAPE) project. The GRAPE project developed a state-of-the-art dataset of cloud and aerosol properties from the ATSR–2 instrument[3].

---

[3]The GRAPE project was a NERC–funded project that ran from 2002–2005 between the University of Oxford and the Rutherford Appleton Laboratory. The principal investigator was Dr Don Grainger, *http* : *//www.atm.ox.ac.uk/project/grape*
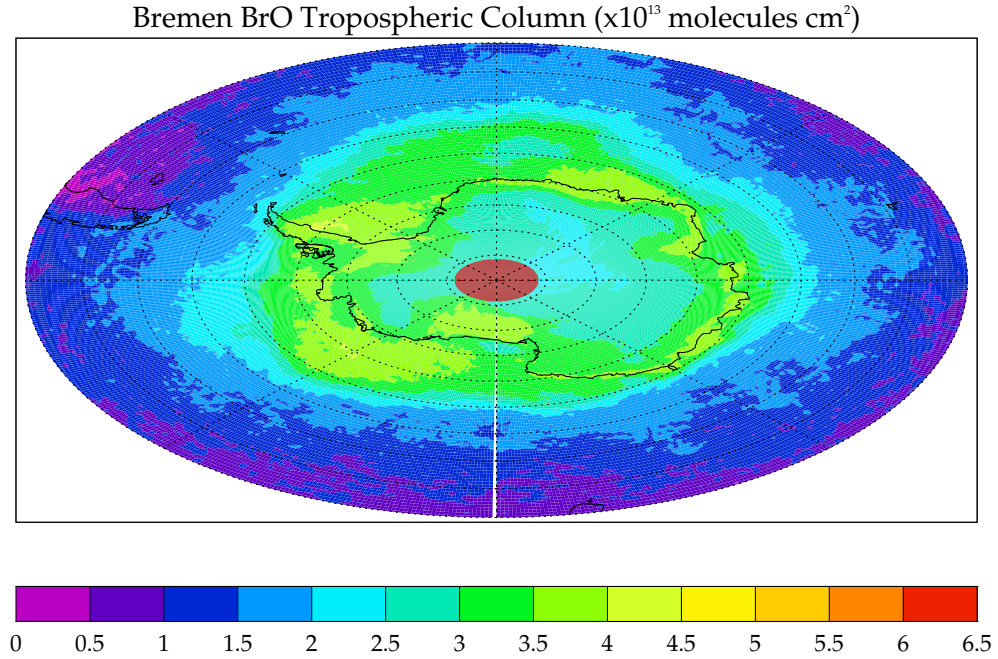
Figure 1.12: *Southern hemisphere tropospheric vertical column density of BrO (molecules cm^{-2}) for November 1996. Data provided by Dr Andreas Richter.*

1.6.1   Cloud Parameters

There are 7 cloud parameters that were retrieved in the GRAPE cloud dataset. These were:

1. Cloud Optical Depth

2. Cloud Phase

3. Cloud Particle Size

4. Cloud Top Pressure

5. Cloud Fraction

6. Cloud Water Path

7. Cloud Top Temperature

along with associated error measurements. The relivant parameters are defined in Chapter 2. The dataset was produced using a method developed for Meteosat Second Generation SE-VIRI measurements and tested on ATSR-2 data (Watts et al., 1998). The algorithm used to retrieve the cloud parameters uses Optimal Estimation (described in Chapter 4) to solve the cloud retrieval problem. In this method all available measurements and *a priori* information

are combined using fast radiative transfer and known error characteristics to fit the ATSR–2 measurements to known cloud optical models (see Chapter 2). The radiative transfer includes cloud, atmosphere and surface effects and assumes a plane-parallel single layer cloud (the 'model cloud').

For clouds there are two types; liquid and ice. For liquid clouds the droplets are spherical where the parameters can be described by Mie Theory (this is explained in more detail in Chapter 2). For ice clouds, larger non-spherical particles are assumed and the scattering properties are found using a combination of geometric optics and a T-matrix method for smaller crystals (Baran et al., 2001). Phase function (also described in Chapter 2) calculations are time consuming so they are pre–calculated and stored in look–up tables.

The advantage of using ATSR–2 cloud parameters over GOME is that the ATSR–2 ground pixels is 1 km× 1 km, which is a better resolution than the GOME 20 km×40 km (for the polarisation measurement devices used to derive cloud information). Figure 1.13 shows cloud top pressure and reflectivity for ERS–2 orbit 1337 for GOME using the FRESCO algorithm and from ATSR–2. It can be seen that the cloud top pressures are different and as ATSR–2 will be more accurate it is the better instrument to use.
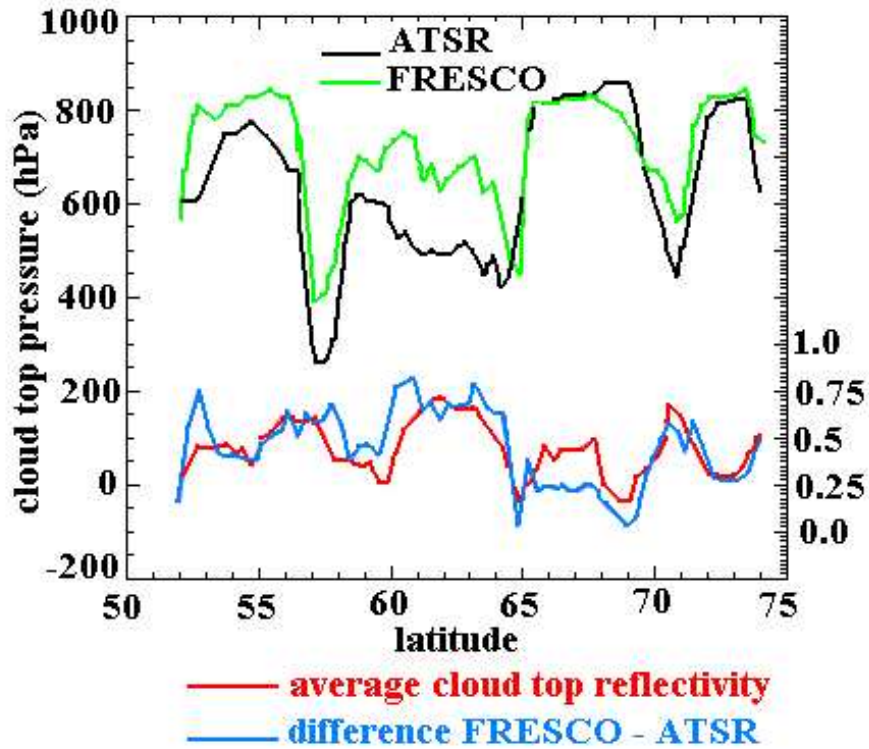
Figure 1.13: *Cloud top pressures from FRESCO, KNMI cloud retrieval for GOME (green) and ATSR–2 (black), as well as their difference, in hPa. In addition the average cloud top reflectivity is shown. The data are from ERS–2 orbit 1337 (taken from Koelemeijer and Stammes, 1998).*

Figure 1.14 shows the different cloud fractions retrieved from ATSR–2 and by the different GOME cloud retrieval algorithms for an example orbit (7105). The figure shows that all

30

schemes fit the ATSR–2 cloud fractions well for cloud fractions above 0.8, but not well below this. Overall, this figure shows that the ATSR–2 observations are not reproduced well from GOME observations and highlight that the use of ATSR–2 cloud parameters will be more accurate than those from GOME.
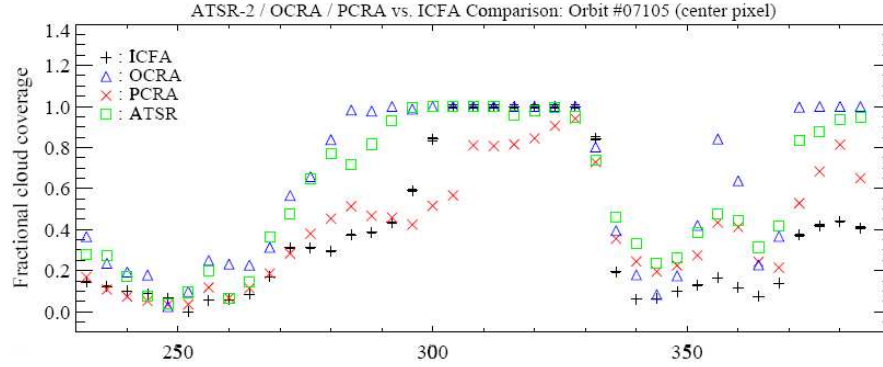
Figure 1.14: *Fractional cloud cover for ATSR–2 orbit number 7105 against pixel number. Results for ATSR–2 (green quadrangle), ICFA (black cross), OCRA (blue triangle), and PCRA (red cross). Taken from Von Bargen et al. (1998).*

1.6.2 Aerosol

To retrieve aerosol properties within GRAPE a similar approach is used to that for cloud. The properties retrieved were:

1. Aerosol particle size

2. Aerosol optical depth

3. Aerosol single scattering albedo

The GRAPE aerosol parameter retrieval relies on a number of assumptions about the characteristics of the aerosol. These assumptions are used as *a priori* estimates and then optimal estimation (see Chapter 4) is used to fit the radiances measured by the ATSR–2 instrument to the optical models (see Chapter 2). The assumptions used for aerosols are described in detail in Marsh et al. (2004). To retrieve the aerosol optical depth assumptions are made about the aerosol refractive index, the size distribution, the phase function[4] and the vertical distribution of the aerosol load. Assumptions about the aerosol type are taken from the Optical Properties of Aerosol and Clouds (OPAC), presented by (Hess et al., 1998).

---

[4]Phase functions were explicitly calculated for the optical properties by Dr. E. Carboni at the University of Oxford as part of the GRAPE project.

1.7 Aims

The main aim of this thesis is to develop a new retrieval scheme to determine tropospheric composition from nadir viewing UV/VIS spectrometers. This chapter has described the basic theory behind existing retrievals from the GOME instrument and highlighted where improvements can be made in order to produce a more accurate scheme.

The main difference between the new scheme developed within this thesis and previous methods is that it uses cloud and aerosol information from the ATSR–2 instrument, which is better than any previous cloud characterisation as it is based on higher resolution data (1 km$\times$ 1 km), and the use of SLIMCAT with chemical data assimilation, which will produce a better representation of the stratosphere. Also the retrieval is framed as an inverse problem which allows for a detailed error analysis. The retrieval scheme also retrieves surface reflectance at the appropriate wavelength as this controls the photon path lengths which control the total column density.

The retrieval scheme measurement is the slant column density (assumed to be a direct measurement) meaning there is no need to fit the radiance spectra. The scheme recognises that almost all of the information on a given trace gas (other than O_3) in the GOME measurement is contained by the slant column of the gas and a single sun–normalised radiance in the fitting window. The scheme developed here therefore uses slant column observations from other sources and focuses on proper treatment of the radiative transfer required to estimate the tropospheric vertical column.

A single sun–normalised radiance in the fitting window helps to characterise the photon path by fitting the surface reflectance. A single slant column is a good equivalent to fitting a full GOME spectrum but avoids a lot of the difficulty which has been handled well already by many other groups mentioned earlier. The advantage of using an optimal method is that it is theoretically sound and based on well understood radiative transfer calculations.

The objectives of this thesis are therefore to:

1. Develop a retrieval model that combines the retrieved slant column densities from GOME and the ATSR–2 (GRAPE) cloud and aerosol data to calculate a tropospheric vertical column density.

   - Determine the stratospheric component of the trace gases of interest (NO_2 and BrO). This will be achieved by using the SLIMCAT CTM constrained by chemical data assimilation of long–lived tracers from the HALOE instrument (described in Chapter 3).

   - Develop a method for combining the GOME and ATSR–2 data so that the true atmospheric scattering profile can be modelled.

2. Carry out a detailed error analysis on the retrieval scheme to determine the errors associated with these types of calculation.

3. Evaluate and compare the new retrieval scheme to existing schemes.

Throughout this thesis the errors due to different assumptions and input parameters to the retrieval model will be quantified and then, in Chapter 6, the impact of these upon the final retrieval product will be determined. In this chapter a number of error sources have been addressed. These are summarised in Table 1.4.

Species	Parameter	Error
NO_2	SCD error	5–9×10^{15} mol. cm^{-2}
BrO	SCD error	1.5–1.8×10^{13} mol. cm^{-2}
NO_2	Previous study VCD_{trop} error	0.5–10×10^{15} mol. cm^{-2}
BrO	Previous study VCD_{trop} error	na

Table 1.4: *The errors associated with assumptions or parameters that will be used in the new retrieval model developed in this thesis for NO_2 and BrO (see text for discussion).*

The slant column density errors for NO_2 are between 5–$\times 10^{15}$ molecules cm^{-2} and for BrO are between 1.5–18×10^{13} molecules cm^{-2}. The errors are reported from each pixel individually. In the case of NO_2 the best estimate of the error in previous studies is from the TEMIS retrieval algorithm where an error in tropospheric VCD of 0.5–10×10^{15} molecules cm^{-2} is reported (see Boersma et al., 2004). For BrO there are no such estimates of errors as the retrieval algorithms are not as sophisticated i.e. they do not model cloud and aerosol well.

This thesis will describe the theoretical basis, application, validation and initial scientific exploitation of a new retrieval scheme to produce tropospheric trace gas column amounts from the GOME instrument. Chapter 2 describes the radiative transfer processes which are important when interpreting the information from the observations from GOME. Chapter 3 describes the 3-D chemical transport model (SLIMCAT) used to determine the stratospheric component of the trace gases total column which is retrieved from the GOME instrument. This chapter also describes the chemical data assimilation scheme used in SLIMCAT as well as describing the error sources. Chapter 4 outlines the basic concepts of inverse theory as applied in the context of this work. The Retrieval Model developed in this thesis will be described and tested using the forward model and will show how the retrieval model is capable of reproducing the radiative transfer in the atmosphere. Chapter 5 describes a comprehensive error analysis that has been carried out on the retrieval scheme. Chapter 6 presents some initial scientific results from the retrieval scheme focusing on comparisons with previous, less quantitative, retrieval schemes and Chapter 7 contains the conclusions and suggestions for further work.

Chapter 2

Radiative Transfer

Radiation in the Earth's atmosphere can interact with its constituents and hence cause changes to temperature and trace gas concentrations. This thesis is concerned with the measurement of radiation that has undergone a number of radiative transfer processes within the lower atmosphere (the troposphere). It is therefore important to comprehend and model the radiative transfer equation in order to correctly interpret the radiation received by the satellite.

This chapter describes the theory of radiative transfer relevant to this study and introduces the radiative transfer model used. Section 2.1 introduces the basic theory underlying radiative transfer models and Section 2.2 discusses the Radiative Transfer Equation and the assumptions used to solve it. Section 2.3 describes the main interactions that radiation undergoes in the atmosphere, namely, scattering, absorption, emission and surface reflectance. Section 2.4 summarises how the cloud and aerosol parameters are derived and what effect they have on radiative transfer and Section 2.5 introduces the radiative transfer model.

2.1   Introduction

2.1.1   Radiance

Radiance is the amount of energy δE, passing through an arbitrary area δA, with a direction of solid angle element $\delta \Omega$, over time δt, and has a wavelength interval of $\delta \lambda$ (Liou, 2002). This is given by

$$\delta E = I_\lambda \cos \theta_v \; \delta A \; \delta \Omega \; \delta t \; \delta \lambda, \tag{2.1}$$

where θ_v is the angle between the beam and the normal to the element surface dA. This is illustrated in Figure 2.1. Equation 2.1 can be used to define the intensity I_λ, of a pencil beam of radiation.
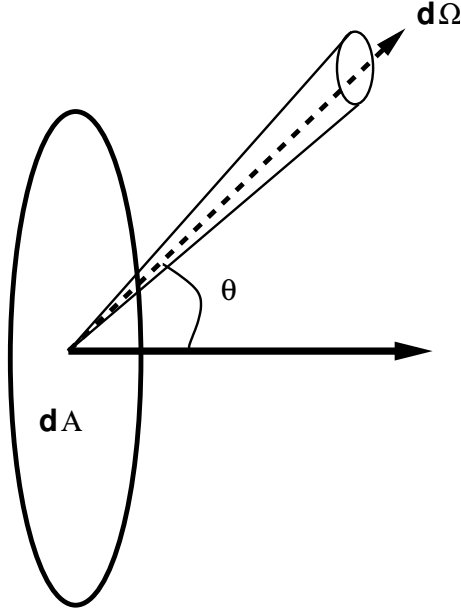
Figure 2.1: *An illustration of a beam of radiant energy passing through an element with an arbitrary area dA, confined to an element of solid angle dΩ, traveling in the direction of $\cos\theta_v$ with respect to the normal through dA.*

2.1.2   Irradiance

The irradiance F_λ, is the total radiant energy incident upon, or exiting from, a surface, per unit area, per unit time, per unit wavelength interval, and is given by

$$F_\lambda = \frac{1}{\pi} \int_0^{2\pi} \int_0^{\frac{\pi}{2}} I_{\lambda_v} \cos(\theta_v) \sin(\theta_v) \delta\theta_v \delta\phi, \tag{2.2}$$

where ϕ is the azimuth angle about the normal to the plane.

2.1.3   Extinction, absorption and scattering coefficients

The intensity, I_λ, of a radiant energy beam will be lessened by its interaction with molecules in the atmosphere. This reduces it to $I_\lambda + \partial I_\lambda$, where

$$\partial I_\lambda = -k_{ext} I_\lambda \partial s, \tag{2.3}$$

∂s is a small distance traveled by the beam and k_{ext} is the volume extinction coefficient (a measure of how strongly a volume of the atmosphere absorbs/scatters radiation at a given wavelength). The extinction coefficient can be represented by contributions from the absorption and scattering coefficients

35

$$k_{ext} = k_{abs} + k_{sca}, \tag{2.4}$$

where k_{abs} and k_{sca} are the absorption and scattering coefficient, respectively. These quantities represent the amount of radiant energy that will be affected by the interaction with the atmosphere. Another important property is the single scattering albedo, this is a measure of the efficiency of a medium or particle to scatter radiation and is defined by

$$\omega_0 = k_{sca}/k_{ext}. \tag{2.5}$$

2.2 Radiative Transfer Equation

The radiative transfer equation, Equation 2.6, states that as a beam of radiation travels through a medium it losses energy by absorption and energy is redistributed by scattering. The equation is given by

$$\frac{dI_\lambda(s,\theta,\phi)}{ds} = -k_{abs}(s)I_\lambda(s,\theta,\phi) + k_{sca}(s)I_\lambda(s,\theta,\phi), \tag{2.6}$$

where s is distance and ϕ and θ are shown on Figure 2.2, which illustrates a simple path of radiation through the atmosphere, it shows a beam of radiant energy entering the atmosphere where it interacts with a single particle in the atmosphere and undergoes a single scattering process. The beam then interacts with a second particle where the beam is added to by another radiant beam that has been influenced by multiple scattering. Figure 2.2 illustrates the different angles that are considered.

Also there is a contribution to the beam from multiple scattering. Adding this component and simplifying Equation 2.6, the scalar approximation for the radiative transfer equation becomes

$$\frac{dI_\lambda(s,\theta,\phi)}{ds} = -k_{ext}(s)I_\lambda(s,\theta,\phi) + \frac{k_{sca}(s)}{4\pi}$$

$$\int_0^{2\pi}\int_0^\pi P(s,\theta',\phi';\theta,\phi)\ I_\lambda(s,\theta',\phi')\ sin(\theta')\ d\theta'd\phi'. \tag{2.7}$$

where P is the phase function. The phase function is an important parameter when considering a multiple scattering radiative transfer scheme as it describes the angular distribution of the scattered radiation. Depending on the size and shape of the molecule the phase functions will vary according to
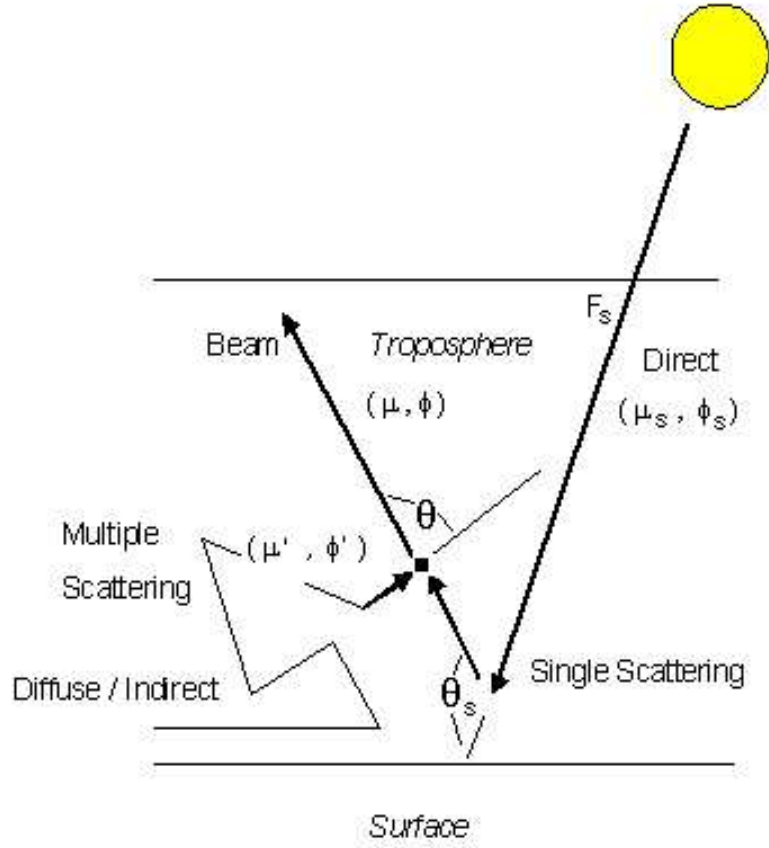
Figure 2.2: *An illustration of how radiation can interact in the atmosphere.*

$$\int_0^{2\pi} \int_0^{\pi} \frac{P\cos\theta}{4\pi} \sin\theta d\theta d\phi = 1, \tag{2.8}$$

where P is the phase function. Also, emission is neglected here as when compared to the solar background term it is negligible.

2.2.1 Assumptions

Without simplification the full vector form of the radiative transfer equation is difficult and computationally expensive to solve. Therefore, a number of assumptions and approximations are needed to simplify the equation and reduce the computational cost, while still retaining an accurate description of the radiative field.

Scalar Approximation

The orientation of the oscillations of an electromagnetic wave are described by its polarisation. For sunlight, the orientation of the polarisation of the wave is random, resulting in the radiation beam being unpolarised, i.e. on average the plane has no preferred orientation. Due to interactions between light and matter it is possible for light to become polarised, i.e. where a preferred orientation is developed (Hecht, 1987). In this thesis the radiation is assumed to be unpolarised and no process in the atmosphere or at the surface of the Earth will act to polarise the radiation. This is known as the scalar approximation.

The scalar approximation version of the radiative transfer equation is given by Equation 2.7. There are a number of reasons why it is possible to use the scalar approximation. Firstly, sunlight is unpolarised. Secondly, the GOME instrument corrects for polarisation, so the polarisation state of the radiation when it is detected is not an issue, except for residual errors, which will be low in DOAS (see Kerridge et al., 2004). Although scattering at the surface does not polarise radiation, scattering in the atmosphere does (especially Rayleigh scattering). Once radiation becomes polarised by scattering in the atmosphere this affects subsequent scattering, so affects the propagation of photons in the atmosphere. However, this becomes less important with multiple scattering (as polarisation becomes random again). In summary polarisation is only important for ~ 2 scatterings (with atmospheric polarisation scatter), and any interaction with the surface removes polarisation.

In the NO_2 region (430 nm) single scattering and surface reflectance are important, or multiple scattering in clouds. Therefore, the scalar approximation will be a reasonable assumption. In the O_3 region (~ 300 nm), which is similar to BrO, Rayleigh scattering is dominant and therefore the scalar approximation will introduce errors into the radiative transfer calculation. A study by Latter (2004) showed errors in the retrieved tropospheric O_3 column of $\sim 10\%$ for high solar zenith and azimuth angles. This error will be comparable to the BrO random error due to the scalar approximation (as similar radiative transfer processes dominate BrO as O_3) and will be an upper limit for the NO_2 random error.

Plane Parallel Approximation

The plane parallel approximation assumes that the atmosphere can be considered to vary in only one dimension. This is valid for most viewing angles except at high solar zenith angles, where pseudo spherical coordinates are needed (Rozanov et al., 1997). This is chosen to be the z coordinate and corresponds to altitude above the surface of the Earth. The intensity can therefore be expressed as a function of z, the viewing direction defined by the solar zenith angle θ, and the relative azimuth angle ϕ, only. Scattering, extinction and absorption coefficients are also functions of z, for a given viewing direction, where $ds = dz/cos(\theta)$.

$$\mu \frac{dI_\lambda(z, \mu, \phi)}{dz} = -k_{ext}(z)I_\lambda(\mu, \phi) +$$

$$\frac{k_{sca}(z)}{4\pi} \int_0^{2\pi} \int_{-1}^1 P(z, \mu'\phi', \mu, \phi)I_\lambda(z, \mu', \phi')d\mu'd\phi, \tag{2.9}$$

where $\mu = \cos\theta$.

2.2.2   Solving the Radiative Transfer Equation

The plane-parallel, scalar radiative transfer equation, Equation 2.9 can be solved numerically by

1. Discretising the atmosphere in the vertical and considering thin layers with fixed optical properties,

2. Separating the angular dependence of the radiation into independent components,

3. Defining boundary conditions.

At the top of the atmosphere, z_T, the down-welling radiance can be given by

$$I_-(z_T, \mu, \phi) = \delta(\mu + \mu_0, \phi + \phi_0)\pi F_\lambda, \tag{2.10}$$

where πF_λ is the solar irradiance, μ_0 and ϕ_0 are the cosine of the solar zenith angle and the azimuth angles, respectively, and $\delta(x, y)$ is the delta function. At the top of the atmosphere there is no diffuse radiation, only plane parallel solar irradiance from the direction (μ_0, ϕ_0).

At the surface, where $z = 0$, the upwelling radiance can be given by

$$I_+(0, \mu, \phi) = \int_0^{2\pi} \int_0^{-1} \rho(\mu, \phi, \mu', \phi')I_-(0, \mu', \phi') \, d\mu'd\phi' \tag{2.11}$$

where ρ is the bi-directional distribution function or BRDF and this describes the scattering and absorption by the surface.

There are various numerically methods that have been developed to solve this equation e.g. discrete ordinate (DISORT, Spurr et al., 2001) or finite differencing (GOMETRAN, Rozanov et al., 1997). In this work the latter method has been used, using the optimised code developed at RAL (see Siddans, 2003).

2.3 Atmospheric and Surface Optical Properties

2.3.1 Surface Properties

Surface reflectance is characterised by the Bidirectional Reflectance Distribution Function (BRDF) which is a function that defines how light is reflected at an opaque surface. Albedo is the integral of this over all directions. In most studies albedo is assumed to be Lambertian (meaning it is isotropic), which is a specific model of BRDF. In this work a Lambertian surface is used which is equivalent to albedo. This means that the Lambertian equivalent reflectance (LER) gives the correct reflectance for the GOME observing geometry. This will only lead to errors if there are multiple reflections at the surface, such as with polarisation.

Koelemeijer et al. (2003b) used GOME sun–normalised radiance to retrieve the Minimum Lambertian-Equivalent Reflectance (MLER) and created a climatology of surface reflectance from this data for a number of wavelengths in the GOME spectral range. This assumes Rayleigh scattering above the surface, but not scattering due to aerosol (Mie scattering). This climatology has been used in a number of GOME retrieval schemes, e.g. (Boersma et al., 2005), and this thesis uses this MLER as an initial guess for the surface LER (see Chapter 4).

The plots in Figures 2.3 and 2.4 show MLER for January, April, July and October 1997, for the wavelength of 430 nm, which is the NO_2 retrieval wavelength, and 340 nm, which is the BrO retrieval wavelength. In Figure 2.3 the higher surface albedo is yellow/red and through the year it is possible to see the melting of the sea ice in the northern and southern hemispheres. Also, there is a change in the surface reflectance in Eurasia and Canada where there is more snow and ice in January. This decreases in April when the northern hemisphere spring starts and disappears in the middle of summer, around July, only to become colder again in October. Another noticeable change in surface reflectance throughout the year is over Africa and South America. This could be due to surface reflectance changes due to biomass burning. However, this is more likely to be due to increased aerosol produced in the biomass burning events (Myhre et al., 2005; Eck et al., 2003). This change shows a noticeable decrease in the surface reflectance that coincides with the biomass burning events that begin in mid–Africa (around the Congo) in April and move south (noticeable in the July plots). Figure 2.4 shows these same features although the surface reflectance values are different due to the different wavelengths.

2.3.2 Trace Gas Absorption

Absorption is the process where energy is removed from the beam of radiant energy and used to change the state of the absorber. The nature of the absorption depends upon the particle that the beam of radiant energy is interacting with and the wavelength of the radiation. Figure 2.5 shows the optical depth of relevant trace gases in the wavelength range of GOME. The figure shows a strong absorption due to O_3 at wavelengths less than 350 nm and that Rayleigh scattering has a greater effect at shorter wavelengths. It shows that NO_2 is the next strongest absorber at wavelengths less than 500 nm. The other trace gas studied in this work is BrO
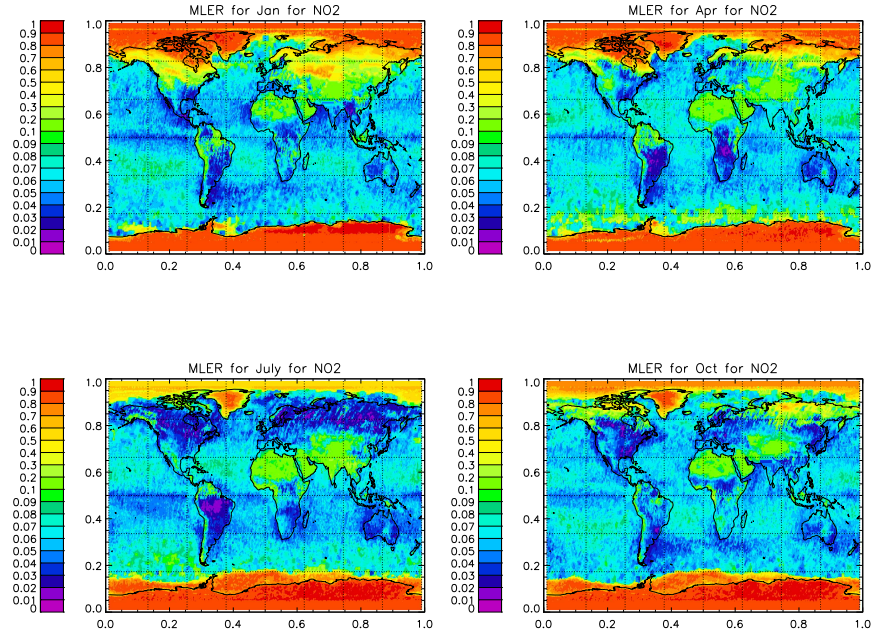
Figure 2.3: *The global plots of Minimum Lambertian-Equivalent Reflectance at 430 nm in January, April, July and October 1997.*
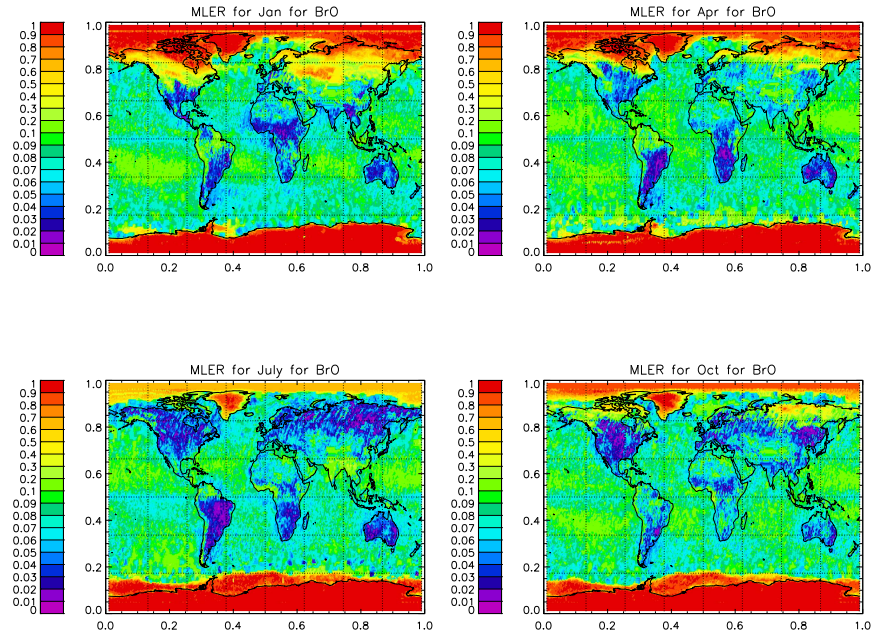
Figure 2.4: *The global plots of Minimum Lambertian-Equivalent Reflectance at 340 nm in January, April, July and October 1997.*

which has absorption features between 300 nm and 380 nm.
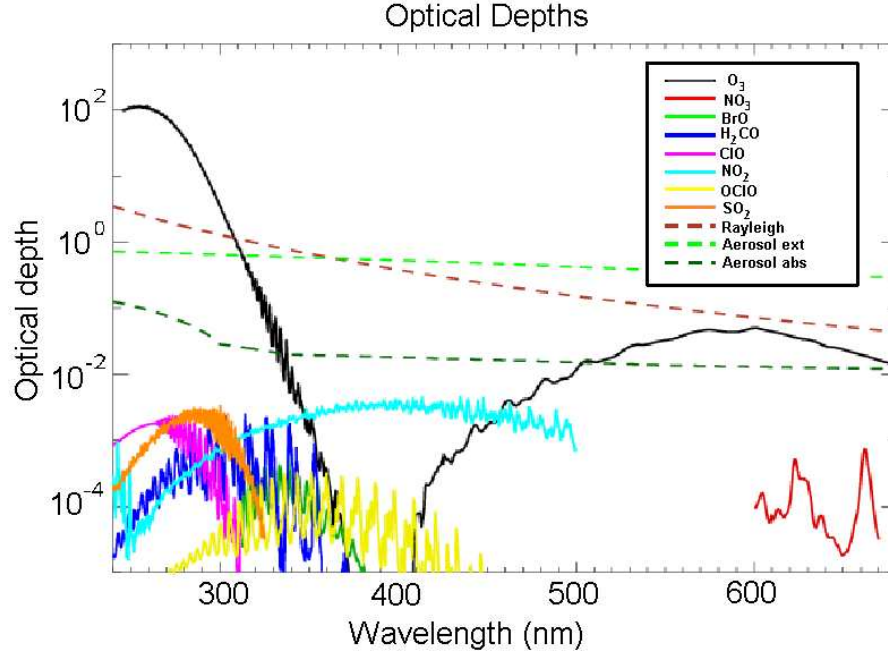
Figure 2.5: *Vertical optical depths of trace gases in the GOME spectral range (240–700 nm). This is taken from Siddans (2003).*

2.3.3   Scattering Particles

Scattering occurs when a beam of radiant energy is diverted from its initial direction by the interaction with matter. All particles in the atmosphere can affect radiation in this way although the amount of scattering that occurs depends upon size of the particle and the wavelength of the radiation. It should also be noted that scattering can cause changes in polarisation (as suggested in Section 2.2.1). All particles have an effective radius r, this leads to the definition of a size parameter x, which is given by

$$x = \frac{2\pi r}{\lambda},\tag{2.12}$$

where λ is the wavelength of the incident radiation. The size parameter can be used to determine the type of scattering that will occur. Figure 2.6 shows how incident radiation wavelength and particle radii determine the type of scattering that occurs.

Figure 2.6 shows that air molecules have effective radii of between 10^4–10^5 and therefore solar radiation will be Rayleigh scattered by its interaction with them. However, smoke particles and cloud droplets, with radii in the order of 10^{-2}–10^{-1} μm, cause Mie scattering of solar radiation as the particles are larger.
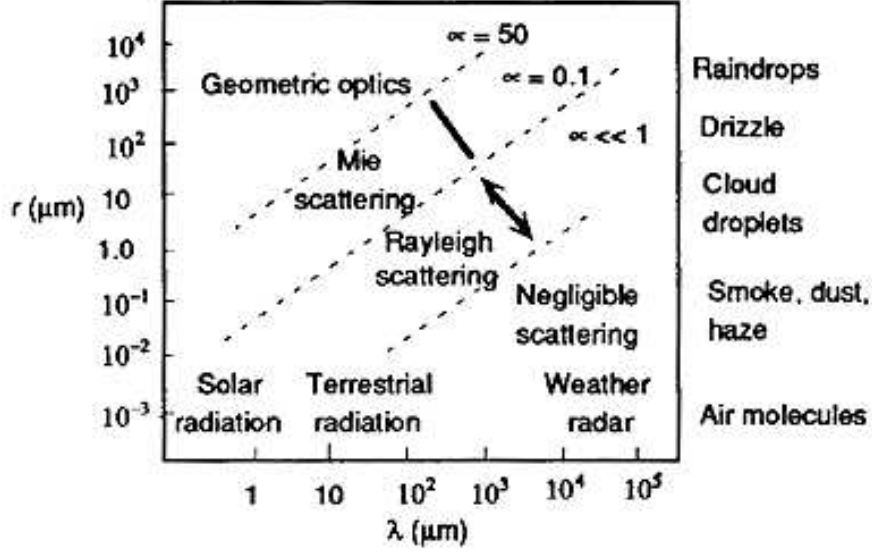
Figure 2.6: *A graph of the wavelength of incident radiation against particle radii. Regions of different types of scattering are separated by the dashed lines. Also the region of radiation, Solar, Terrestrial and Weather radar are shown.*

Mie Scattering

The mathematical theory of Mie scattering starts with Maxwell's equation for electromagnetic radiation and is used for all scattering theory. For particles where r is much smaller than λ the solution can be approximated to Rayleigh scattering (see Section 2.3.3), and for particles where r is much larger than λ the solution can be simplified to geometric scattering (see Liou (2002)).

For particles where r is comparable to λ a full Mie solution is needed. The intensity scattered by a particle as a function of direction is given by

$$I(\theta) = I_0 \Omega_{eff} \frac{P(\theta)}{4\pi} = I_0 \left(\frac{\sigma_s}{a^2}\right) \left(\frac{P(\theta)}{4\pi}\right) \tag{2.13}$$

where I_0 is the incident intensity, P is the phase function, Ω_{eff} is the effective solid angle upon which scattering occurs, a is the distance between the particle and the observer, σ_s is the scattering cross section and 4π is the solid angle for the entire spherical space (Liou, 2002). The solution of this using Mie theory, is a computationally accurate method for spheres (Mie, 1908).

43

Rayleigh Scattering

Rayleigh scattering is the elastic scattering of light and other electromagnetic radiation by particles much smaller than the wavelength of the light. Rayleigh scattering of sunlight in clear atmospheres is the main reason why the sky is blue. The angular distribution of Rayleigh scattering is governed by the phase function

$$P(\cos\theta) = \frac{3}{4}(1 + \cos^2\theta) \tag{2.14}$$

where P is the phase function and θ is the scattering angle and the scattering is symmetric in the plane normal to the incident direction of the light. Also the forward scatter equals the backwards scatter. Integrating over the sphere surrounding the particle gives the Rayleigh scattering cross section

$$\sigma_s = \frac{2\pi^5}{3}\frac{d^6}{\lambda^4}\left(\frac{n^2-1}{n^2+2}\right)^2, \tag{2.15}$$

where σ_s is the scattering coefficient, d is the diameter of the particle and n is the refractive index. For this equation it can be seen that there is a power of four wavelength dependence with Rayleigh scattering.

2.4 Optical Models used to Characterise Cloud and Aerosol

In this work the two important features in the atmosphere which need to be correctly modelled, in optical terms, are clouds and aerosols. Both cloud and aerosol are extremely variable, therefore optical models are used to approximate their characteristics. In the GRAPE retrieval a specific cloud or aerosol optical model is assumed and the retrieval scheme then fits the observations to the optical model. To limit the error in this new retrieval scheme the same optical models are used to describe the cloud and aerosol as were used for GRAPE. The following subsections discuss the optical properties of liquid cloud, ice cloud and aerosol in terms of their extinction profiles.

It is important to note that it is not possible to capture all this variability in clouds and aerosol as firstly, only 1-D radiative transfer is considered here (meaning 3–D structures such as clouds will not be treated as accurately as possible), and secondly, only a specific composition of aerosol is considered, whereas in reality composition will vary.
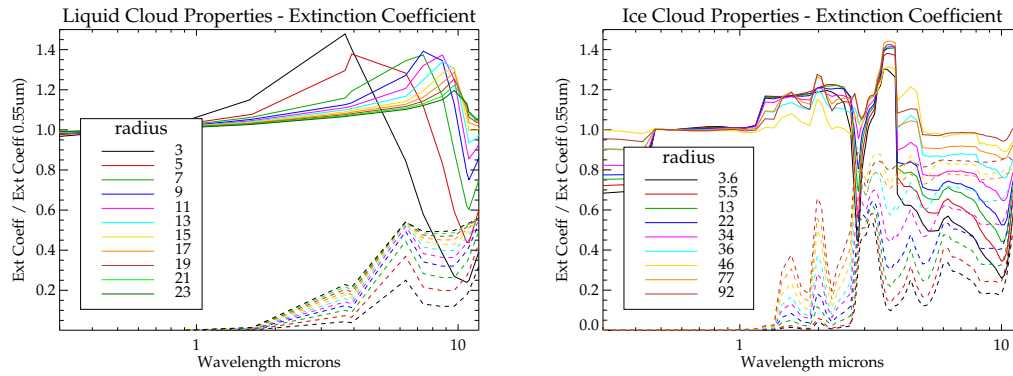
Figure 2.7: *(a) The liquid cloud optical properties over wavelengths 0.3-12 microns. The solid line shows the extinction coefficient/extinction coefficient at 0.55 microns and the dashed line shows the absorption coefficient/extinction coefficient at 0.55 microns. (b) is same a figure (a) but for ice cloud.*

2.4.1 Liquid Cloud Properties

Figure 2.7(a) shows liquid cloud extinction and absorption coefficient[1] (normalised to extinction coefficient at 0.55 microns) with respect to wavelength for a range of particle sizes. The figure shows that the extinction coefficient below 1 micron does not vary as much as it does at the longer wavelengths, above 1 micron). Also, liquid clouds absorb radiation in the longer wavelength range and scatter radiation in the shorter wavelength range, below 1 micron.

2.4.2 Ice Cloud Properties

Figure 2.7(b) shows the ice cloud extinction and absorption coefficient[2] (normalised by extinction coefficient at 0.55 microns) with respect to wavelength for a range of particle sizes. The figure shows that ice cloud optical properties are more irregular with respect to wavelength than liquid clouds. Ice clouds are also strong scatterers in the shorter wavelength ranges, below 1 micron. It can be seen that at wavelength ranges that will be used in this work (i.e. UV/VIS - below 1 μm) that ice particles will scatter nearly all the radiation incident upon them. At longer wavelengths radiation will be both scattered and absorbed.

2.4.3 Aerosols

Figure 2.8(a) shows the extinction and absorption coefficients (normalised to extinction coefficient at 0.55 microns) with respect to wavelength for the Antarctic aerosol for a range of aerosol particle sizes. The figure shows that the relationship between extinction coefficient and wavelength is nearly linear and is therefore easy to calculate. Large particles, with effective

---

[1]Liquid water cloud optical properties were originally provided by P. Watts previously at RAL, now EU-METSAT.

[2]GRAPE ice cloud optical properties were originally provided by A. Baran at the UK Met Office.
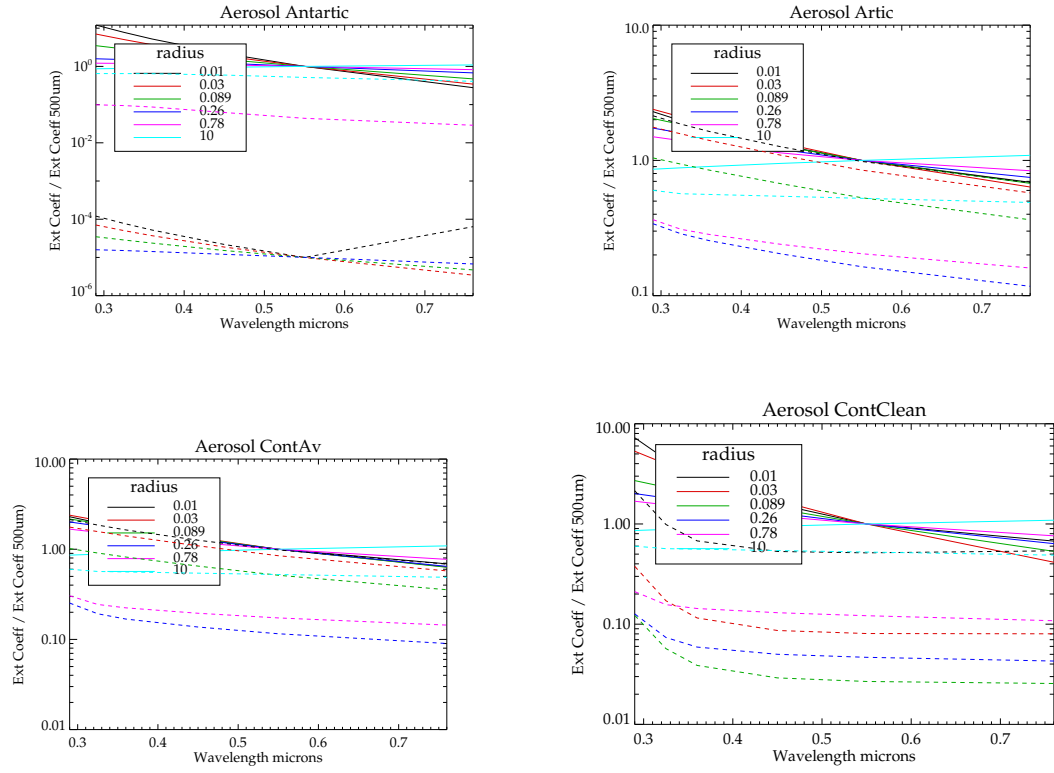
Figure 2.8: *The aerosol properties for the wavelengths 0.3-0.8 µm. The solid line shows the (extinction coefficient)/(extinction coefficient at 0.55 µm) and the dashed line shows the (absorption coefficient)/(extinction coefficient at 0.55 µm).*

radii 0.7–10 µm, absorb and scatter. The smaller particles will mostly scatter. These smaller particles will cause Rayleigh scattering.
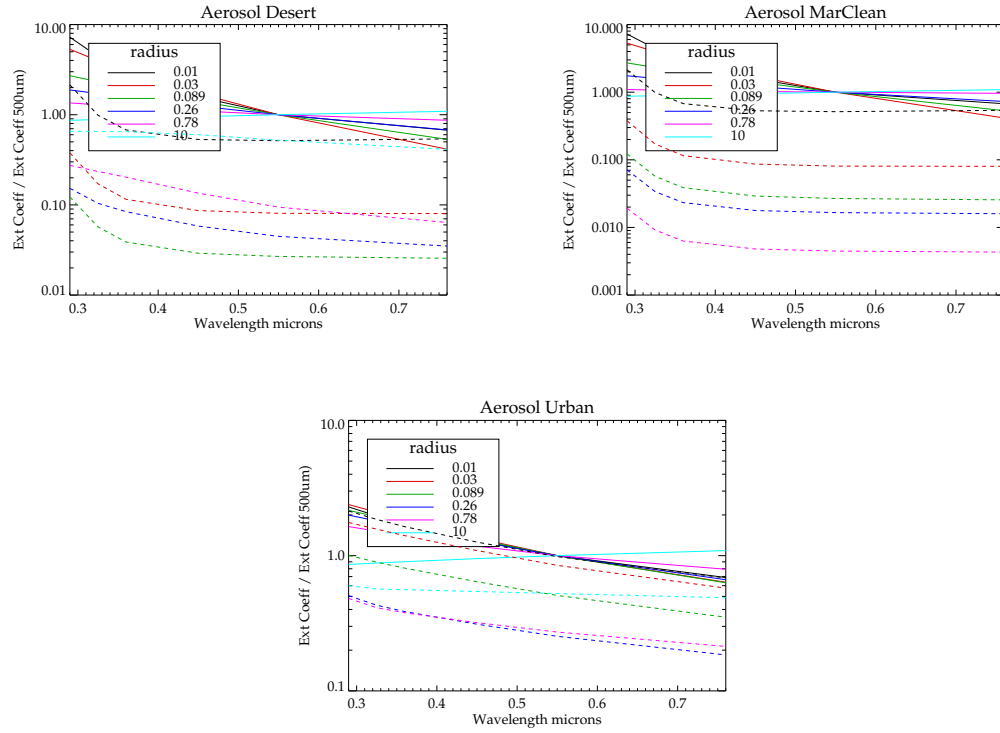
Figure 2.9: *The aerosol properties for the wavelengths 0.3-0.8 μm. The solid line shows the (extinction coefficient)/(extinction coefficient at 0.55 μm) and the dashed line shows the (absorption coefficient)/(extinction coefficient at 0.55 μm).*

Figure 2.8(b), shows that the Arctic aerosol has different optical properties to the Antarctic aerosol. The figure shows that the relationship between extinction coefficient and wavelength is also nearly linear and this increases with decreasing wavelength. Arctic aerosol is seen to be more absorbing than Antarctic aerosol because the very small particle sizes (0.01-0.02 μm) are more absorbing than the larger size particles (0.7-10 μm). The medium size particles (0.1-0.5 μm) will affect the radiation through scattering.

Figure 2.8(c), is for the continental average aerosol. It is similar to Figure 2.8(b) and shows that the smaller and larger particles are more absorbing, whereas the medium sized particles are more scattering.

Figure 2.8(d), is for continental clean aerosol. This figure shows that at lower wavelengths of the order 0.3 μm, the absorption coefficient increases, this means that there is less scattering by the continental clean aerosol in this wavelength range. The relationship between extinction coefficient and wavelength is also nearly linear.

Figure 2.9(a), is for the desert aerosol. The figure shows a near linear relationship between extinction coefficient and wavelength for the largest and smallest particles. Therefore, for the other particles they mostly scatter radiation.

Figure 2.9(b), is for the maritime clean aerosol. The relationship between extinction coefficient and wavelength is nearly linear and at lower wavelengths, around 0.3 μm, the absorption coefficient becomes large meaning there is more absorption. The largest particle size absorption coefficient is not shown on this plot as it is too small. However, the figure shows that increasing the particle size decreases the absorption coefficient.

Figure 2.9(c), is for the urban aerosol. The figure shows that the relationship between extinction coefficient and wavelength is nearly linear and that the urban aerosol is highly absorbing.

2.5   Radiative Transfer Model

The radiative transfer equation is solved computationally in this thesis using the RAL radiative transfer code. This is based on an optimised version of GOMETRAN (Rozanov et al., 1997). The radiative transfer model takes input profiles of temperature, pressure, trace-gas volume mixing ratios, cloud properties (ice or liquid), aerosol amount and effective radius. The model computes optical properties using cross-section data and GRAPE cloud/aerosol optical models (described in Section 2.4). It then uses the finite differencing method to solve the scalar radiative transfer equation for a plane parallel atmosphere and approximates the angular distribution of the radiant field using stream angles[3]. The RAL radiative transfer code is run in this thesis with 15 stream angles which is the optimum number required to solve the radiative transfer equation in this case (see Chapter 4 for more details).

---

[3]streams angles are used to represent the different directions in which the multiply scattered radiation originates.

The radiative transfer model uses the Delta–M method (Wiscombe, 1977), which approximates highly forward scattering mediums such that the fraction of scattered energy residing in the forward direction is removed from the scattering parameters and is treated as though no scattering has occurred.

2.6 Summary

In this chapter the key radiative transfer processes affecting the propagation of photons in the UV/VIS spectral range have been described. These processes include multiple scattering (Rayleigh and Mie) and absorption (by trace gases, clouds, aerosols and the surface). Understanding these processes allows the radiative transfer equation to be defined. However, it is a complex equation when all radiative processes are considered, therefore, in order to solve the equation a number of approximations are made. These include the scalar approximation (for unpolarised light), the plane–parallel atmosphere (assumes atmosphere varies in one dimension which is set to altitude) and the omission of emission (due to this contribution being negligible in comparison to the solar background).

One of the main aims for this thesis is to combine data from the GOME and ATSR–2 instrument in order to solve the radiative transfer equation more accurately. This chapter described the optical properties of cloud and aerosol i.e. the GRAPE optical properties, and shows how they vary in the UV/VIS/IR wavelengths.

The radiative transfer model GOMETRAN was also introduced. The GOMETRAN model has been optimised to allow for a faster, more accurate radiative transfer calculation by introducing the ability to add predetermined phase functions and by using the Delta–M approximation to approximate highly forward scattered radiation.

There are a number of assumptions introduced in this chapter which will affect the final retrieved tropospheric vertical column density. These have been added to Table 1.4 and are now summarised in Table 2.1.

Species	Parameter	Error
NO_2	SCD error	5–9×10^{15} mol. cm^{-2}
BrO	SCD error	1.5–1.8×10^{13} mol. cm^{-2}
NO_2	Previous study VCD$_{trop}$ error	0.5–10×10^{15} mol. cm^{-2}
BrO	Previous study VCD$_{trop}$ error	na
NO_2	Scalar approximation	upper limit 10% of VCD$_{trop}$
BrO	Scalar approximation	10% of VCD$_{trop}$
NO_2 / BrO	Error due to cloud and aerosol optical models	to be quantified in Chapter 4
NO_2 /BrO	Plane–parallel assumption	only significant at high sza

Table 2.1: *The errors associated with assumptions or parameters that will be used in the new retrieval model developed in this thesis for NO_2 and BrO.*

The table states that the errors of importance from this chapter are firstly due to the scalar approximation. This has been shown to cause an upper limit of 10% error in the retrieved

vertical column density. Secondly, is the error associated with the plane–parallel assumption. This is only valid for high solar zenith angles. Thirdly, the error that will be introduced due to errors in the optical properties of clouds and aerosol. The impact of these errors will be quantified in Chapters 4 and 5.

Chapter 3

The SLIMCAT 3-D CTM

3.1 Introduction

In this chapter the SLIMCAT three-dimensional chemical transport model (CTM) is described. Its major application for this work is to accurately describe the chemistry within the stratosphere, i.e. to quantify the relevant UV–absorbing trace gases. The gases of primary interest in this work are NO_2 and BrO, which are present in larger concentrations in the stratosphere than in the troposphere. Therefore, it is important that the stratospheric contribution is estimated accurately. This can then be removed from the total column, and hence the total tropospheric column given by the satellite measurement can be estimated with a greater degree of accuracy.

In this work the stratospheric column is calculated using the SLIMCAT 3-D CTM, which is constrained using chemical data assimilation. This involves running long–term simulations of the CTM which include the first multi-annual runs with chemical data assimilation. The model runs are compared with a range of observations to test the impact of the chemical data assimilation scheme. This also provides the opportunity to analyse the long-term trends of stratospheric column NO_2 and O_3 (key stratospheric species), using a model where chemical data assimilation can constrain the long-lived tracers.

Section 3.2 introduces the SLIMCAT 3-D CTM after which Section 3.3 describes the chemical data assimilation and how it is incorporated into the model. Section 3.4 describes the SLIMCAT runs and Section 3.5 discusses the effects of the chemical data assimilation. Section 3.6 looks at the long-term trends of stratospheric NO_2 and O_3 and Section 3.7 investigates the errors in the SLIMCAT total column measurements.

3.2 The SLIMCAT Model

SLIMCAT is a global 'off-line', 3-D stratospheric chemical transport model. It has been used for a number of studies of stratospheric chemistry and has been shown to perform well when simulating key chemistry and transport processes (e.g. Chipperfield, 2006; Feng et al., 2007). This section briefly describes the basic model; more details can be found in Chipperfield (1999).

3.2.1 Model Setup

SLIMCAT is an Eulerian model, i.e. the model domain is divided into a fixed grid of cells. The model works by calculating the rate of change of tracer concentration, c, in each grid cell, according to the wind velocity advecting air through each face of the cell and the chemical loss and production within the cell. The equation that is key to this, and all other chemical models, is the continuity equation given by

$$\frac{dc}{dt} + \nabla \bullet (c\underline{V}) = P - Lc, \tag{3.1}$$

where V is the velocity vector across a cell and P and L are the chemical production and lose rates, respectively. The model horizontal grid spacing is variable and its choice depends on available computational power. The longitudinal spacing is regular although the latitude spacing can be irregular. The SLIMCAT model uses a hybrid $\sigma - \theta$ vertical coordinate, where the definition of the levels changes with altitude. A hybrid coordinate system allows the model to use terrain–following coordinates near the surface and potential temperature (θ) coordinates above a reference level (e.g. $\theta = 350K$). θ coordinates give a better separation of horizontal and vertical motion in the stratosphere.

The term 'off-line' is used as SLIMCAT takes wind and temperature fields from reanalyses. SLIMCAT runs in this work use ECMWF ERA-40 (Uppala et al., 2005) reanalyse and operational winds. Vertical advection in the θ-level domain is calculated from diabatic heating rates using a radiation scheme which gives a good representation of vertical transport and age-of-air with ERA–40 analyses (see Chipperfield, 2006; Monge-Sanz et al., 2007).

3.2.2 The SLIMCAT Stratospheric Chemistry Scheme

The SLIMCAT model contains a stratospheric chemistry scheme that is described in Chipperfield (1999) and Feng et al. (2007). The scheme includes a detailed description of gas–phase stratospheric chemistry and a treatment of heterogeneous reactions on liquid aerosols, nitric acid trihydrate (NAT) and ice. A list of the chemical species calculated by the model is given in Table 3.1. A list of reactions can be found in Chipperfield (1999).

The SLIMCAT model runs performed in this chapter contain two new tracers for bromine source gases. These are a halon and a very short–lived species or VSLS. These are initialised

Table 3.1: *The chemical species contained in the SLIMCAT model.*

Category	Species
Coupled short-lived species	O_x (= O_3 + O(^{3}P) + O(^{1}D)), H_2O_2
	NO_x (= N + NO + NO_2), NO_3, N_2O_5,
	HNO_3, HO_2NO_2,
	ClO_x (=Cl + ClO + $2Cl_2O_2$), $ClONO_2$,
	HCl, HOCl, OClO,
	BrO_x (=Br + BrO), $BrONO_2$, BrCl,
	HBr, HOBr, HCHO, CH_3OOH
Steady state	H, OH, HO_2, CH_3, CH_3O_2,
	CH_3O, HCO
Source gases	CH_4, N_2O, CO, H_2O, $CFCl_3$,
	CF_2Cl_2, CH_3Br, VSLS, Halon
Fixed	O_2, N_2, H_2

using data from H1211 and CH_2Br_2 (Feng et al., 2007). Most CTMs base their bromine loading solely on long–lived species. However, it has been shown by WMO (2002), for example, that such CTMs underestimate the stratospheric bromine loading and therefore its impact on O_3. With the addition of these short–lived bromine species results are improved (e.g. Feng et al., 2007).

3.2.3 Motivation for SLIMCAT Model Runs

In this work the SLIMCAT model, constrained by chemical data assimilation, is used to perform multi–annual simulations of stratospheric chemistry. The output from the model is used in two ways. Firstly, in order to determine the stratospheric vertical column density (VCD), to a high degree of accuracy such that it can be removed from the total VCD retrieved from the GOME observations, and secondly, to obtain accurate profiles of the stratospheric UV–absorbing species relevant for the atmospheric retrieval (see Figure 2.5).

3.3 Chemical Data Assimilation

Data assimilation is an analysis technique whereby observations can be accumulated and added to a model. The ability to directly assimilate the measurements of chemical constituents into models can provide a number of significant benefits, particularly with regards to assessing the consistency between measurements and theory. The theory of data assimilation is well established as it has been used for a number of years in weather forecast models (e.g. Riishojgaard, 1996). However, in recent years it has been expanded into global chemistry models (e.g. Elbern and Ebel, 1997; Levelt et al., 1998; Khattatov et al., 2000; Chipperfield et al., 2002).

Over the past twenty-five years measurements of chemical species in the atmosphere have increased due, in part, to the increase in the number of observational satellites. This has led to

the need for techniques to interpret and validate measurements produced. Data assimilation provides a quantitative comparison between model results and measurements and can allow the missing chemistry and chemical balance in models to be assessed. Data assimilation also provides information about the expected distribution of the chemical species at locations and times not covered by observations and can also be used for satellite validation.

There are two main methods of data assimilation, variational and sequential (see Figure 3.1). Sequential assimilation only considers observations that have been made in the past, up until the time of assimilation. When a measurement, represented by a dot in the upper part of Figure 3.1, is encountered, a new analysed field value is computed. This is represented by a jump in the simulation. Through sequential assimilation the observations are used to continuously 'nudge' the model towards reality. Variational assimilation, labelled 4–D var in Figure 3.1, is a more complicated procedure. Observations made in the past and future are used to constrain the model. The lower part of Figure 3.1 shows that for the first guess the first observation is used. However, the first guess track does not match the other observations so when the next observation is added, the model uses this information to recalculate the previous estimate of the field track. This allows the model to track the observations better. Variational assimilation is computationally expensive as model values are recalculated many times.

3.3.1 SLIMCAT Assimilation Scheme

The SLIMCAT model uses a sub–optimal Kalman filter (a sequential assimilation scheme), previously described by Khattatov et al. (2000). In this work the scheme assimilates four long–lived chemical species, CH_4, H_2O, HCl and O_3, from observations made by the Halogen Occultation Experiment (HALOE, a solar occultation instrument onboard the Upper Atmosphere Research Satellite (UARS)).

The main effect of sequential chemical data assimilation on the SLIMCAT model is to correct for any model transport errors. As all tracers in the model are overwritten at the surface and advected into the stratosphere, problems with transport will cause the distribution of tracers to be wrong, e.g. incorrect age-of-air.

The overall aim of this part of the work was to obtain stratospheric profiles of NO_2 and BrO to a high degree of accuracy. It is also important to get an accurate profile of O_3 as, although it was not retrieved in this thesis, it has significant absorption features in the wavelengths that are used in this study and therefore an incorrect profile will introduce errors into the retrieval. In order to allow the data assimilation to affect other species in the model, the other long–lived tracers and the shorter–lived species, constraints are imposed to ensure self consistency.

The first constraint is on the modelled long–lived tracers. Plumb and Ko (1992) showed that there are compact correlations between the mixing ratios of long–lived tracers in the stratosphere. Figure 3.2 shows the correlations of long–lived tracers in the free running model (Run **A**, see Section 3.4). These relationships can be used to constrain the model allowing the assimilated CH_4 tracer to impact upon all the other long–lived tracers in the model.
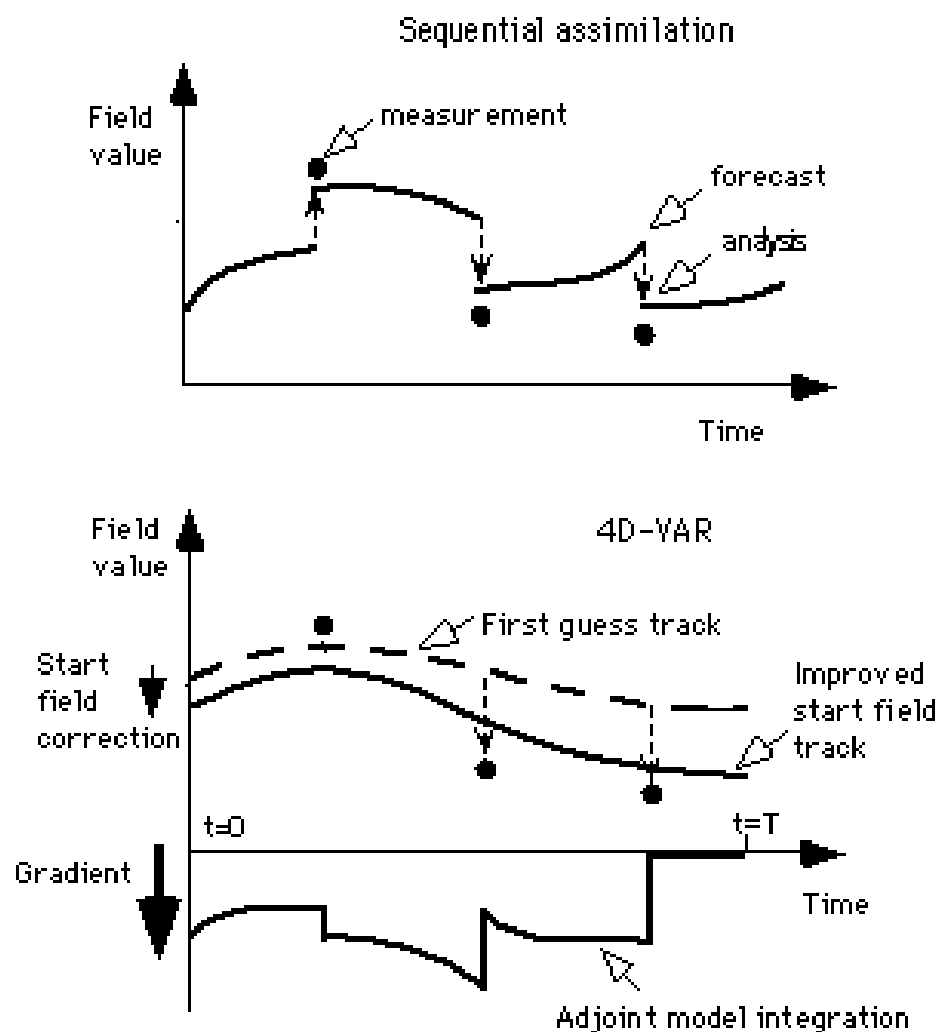
Figure 3.1: *An illustration of sequential (top) and variational (bottom) assimilation schemes. The plots show how a field value i.e. a chemical concentration, is changed over time under the different assimilation schemes. The figure has been taken from Eskes et al. (1998).*

55
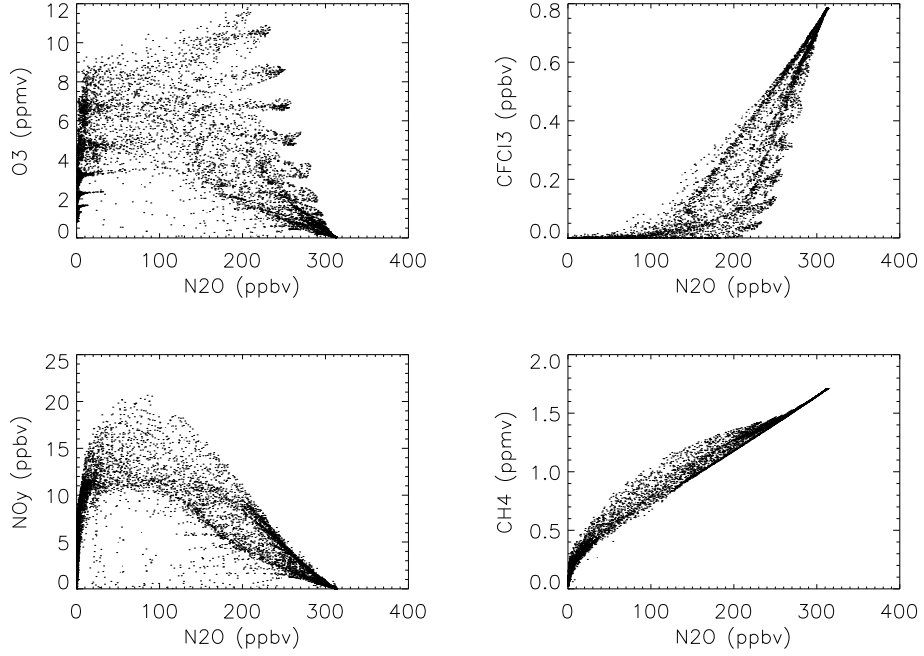
Figure 3.2: *Correlation plots of tracers produced by the SLIMCAT model for the period November 1991 to January 1992 in a run without chemical data assimilation (Run A). Plot (a) is O_3 versus N_2O, plot (b) is $CFCl_3$ versus N_2O, plot (c) is NO_y versus N_2O and plot (d) is CH_4 versus N_2O.*

The second constraint is on HCl. It is known from measurements of atmospheric halocarbons concentrations, e.g. CFCs, that there is a limit on Cl_y of around 3.6 ppbv for 1990 conditions (e.g. WMO, 2002). Therefore assimilated HCl is constrained so that its value does not exceed that expected from the modelled Cl_y.

These correlations are preserved in the SLIMCAT model runs using chemical data assimilation and by imposing these additional constraints. Figure 3.3 shows the correlations from the assimilation run for the same period and species as Figure 3.2. The figure shows that the correlations have indeed been preserved in the assimilation run, i.e. the plots are very similar.

3.4 Model Runs

In this work the SLIMCAT model was run with chemical data assimilation from 1991 to 2004 (which includes the GOME period). In order to understand how the chemical data assimilation affected the model output, three experiments were performed. The experiments described here have been run with a model resolution of $7.5° \times 7.5°$ with 24 vertical levels from the surface to around 60 km. The tunable parameters for the assimilation scheme are outlined in Chipperfield et al. (2002). The model was initialised on January 1, 1977, partly to allow for spin-up, using output from a 2-D model and was integrated until 2004 using a series of experiments. Table 3.2 describes three model runs that have been performed.
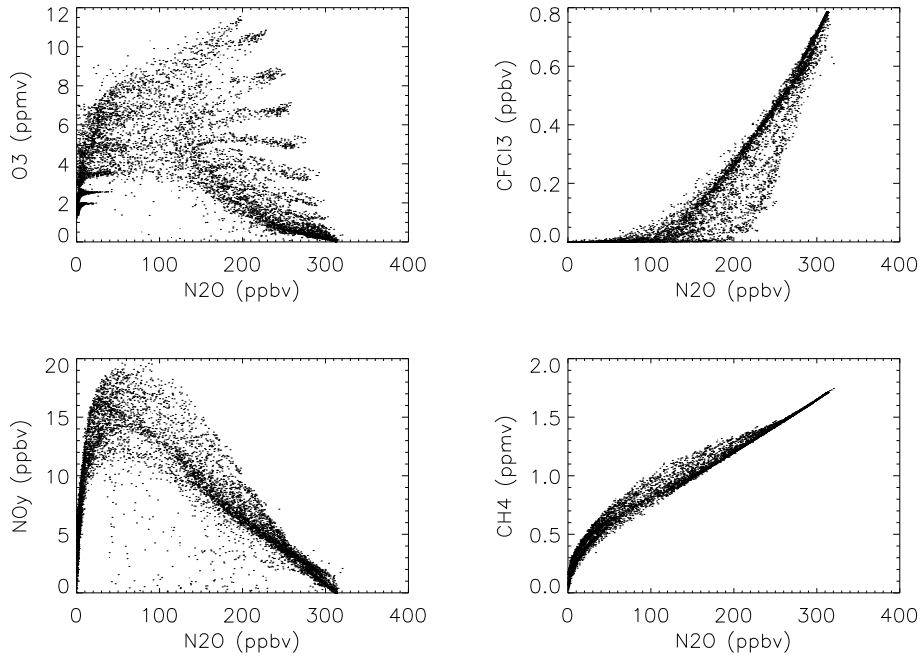
Figure 3.3: *As Figure 3.2 but for the SLIMCAT model with chemical data assimilation (Run B).*

Table 3.2: *SLIMCAT 3-D Model Assimilation Runs*

Run	Date	Notes
A	1977–2004	Basic model run
B	1991–2004	As **A** but with assimilation of HALOE CH_4, H_2O, HCl, and O_3
C	1991–2004	As **B** but without the assimilation of O_3

The first run (**A**) is the basic model or free–running model that does not include any chemical data assimilation. The second run (**B**) is the same as run **A** until 1991 when the run then uses the chemical data assimilation scheme. The chemical data assimilation scheme takes measurements of CH_4, H_2O, HCl, and O_3 from the HALOE instrument.

Figure 3.4 shows the coverage of the HALOE instrument. Although the coverage is sparse this is not a problem as the species assimilated are long–lived and will carry information about the atmosphere long after the data has been assimilated. The third run (**C**), is the same as model run **B** except that O_3 is not assimilated. Run **C** is used to diagnose the chemical impact of assimilating O_3 on other species, for example NO_2.
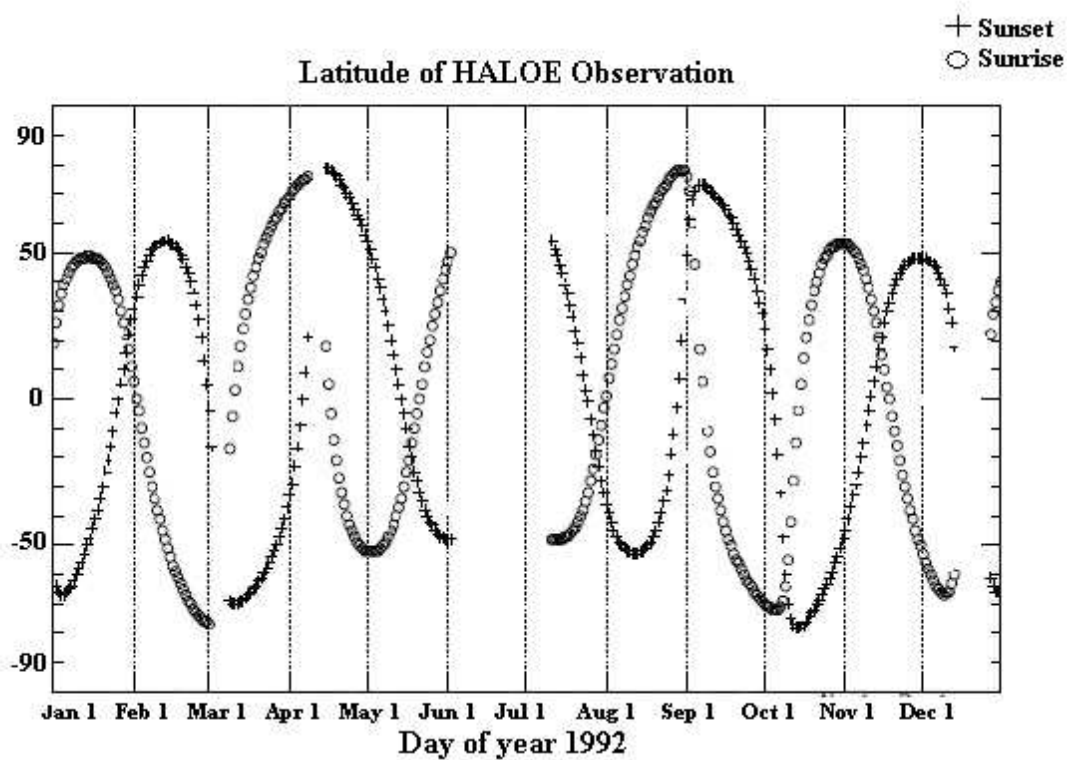
Figure 3.4: *Example coverage of the HALOE sunrise and sunset observations for 1992. (Figure taken from http: //haloedata.larc.nasa.gov).*

3.5 Effects of Chemical Data Assimilation

This section uses independent observations to evaluate the model runs with chemical data assimilation. Firstly, the long–lived tracers are discussed, followed by O_3. Section 3.5.3 then looks at the reservoir species and Section 3.5.4 the short–lived species.

3.5.1 Long–Lived Tracers

As the SLIMCAT chemical data assimilation scheme works by assimilating long–lived tracers, model runs **B** and **C** are expected to be more realistic than model run **A**. Figure 3.5 shows the zonal mean distribution of CH_4 and H_2O on January 30, 1992 for the runs **A** and **B**. The most notable difference between the two runs is that run **B** has an increased horizontal gradient in concentration in the subtropics (at $30°$ N and $30°$ S). This feature has been observed in other data assimilation studies and is known to be due to problems in the tropical dynamics of the analysis models (e.g. Schoeberl et al., 2003). It has been noted that the winds used to force 'off–line' models, have a too strong horizontal transport in the subtropics, this in turn affects the stratosphere as the extratropical age-of-air is too young. The chemical assimilation improves upon this by correcting the species concentrations, i.e. correcting for errors in the transport. The results here differ from those observed in the previous study of Chipperfield et al. (2002), where the chemical data assimilation scheme was originally implemented in SLIMCAT. There the increased gradients in the subtropics were more notable.

The present study uses ECMWF ERA-40 winds whereas Chipperfield et al. (2002) used UK Met Office (UKMO) UARS analyses (Swinbank and O'Neil, 1994). The newer ECMWF winds, although still with their problems, are more realistic.

Figure 3.6 shows the zonal mean plots for the total inorganic chlorine, Cl_y, and HCl in runs **A** and **B** on January 30, 1992. $CFCl_3$ and CF_2Cl_2 are constrained in the model by their correlation to CH_4. The subsequent balance of total organic and inorganic chlorine produces a zonal mean distribution that resembles that of CH_4 in Figure 3.5.

The zonal mean plots shown in Figure 3.6 show the expected results, i.e. that runs **B** and **C** produce more realistic results than run **A**. Note that results from run **C** are very similar to run **B** (not shown). However, it is necessary to further validate the model runs with independent data. This is achieved by comparing the model runs with data from the Atmospheric Trace Molecular Spectroscopy Experiment (ATMOS), a solar occultation instrument that was launched on-board three Space Shuttle flights in 1992, 1993, and 1995. The instrument carried out profile measurements at sunrise and sunset of a number of trace gases in the stratosphere (Gunson, 1996).

Figure 3.8 shows the comparison of five long–lived tracers, the four species that have been assimilated into the model CH_4, HCl, O_3 and H_2O, and also N_2O which is updated in the model by its correlation to CH_4 (see Figure 3.2). The locations of these observations are shown in Figure 3.7 and are listed in Table 3.3.
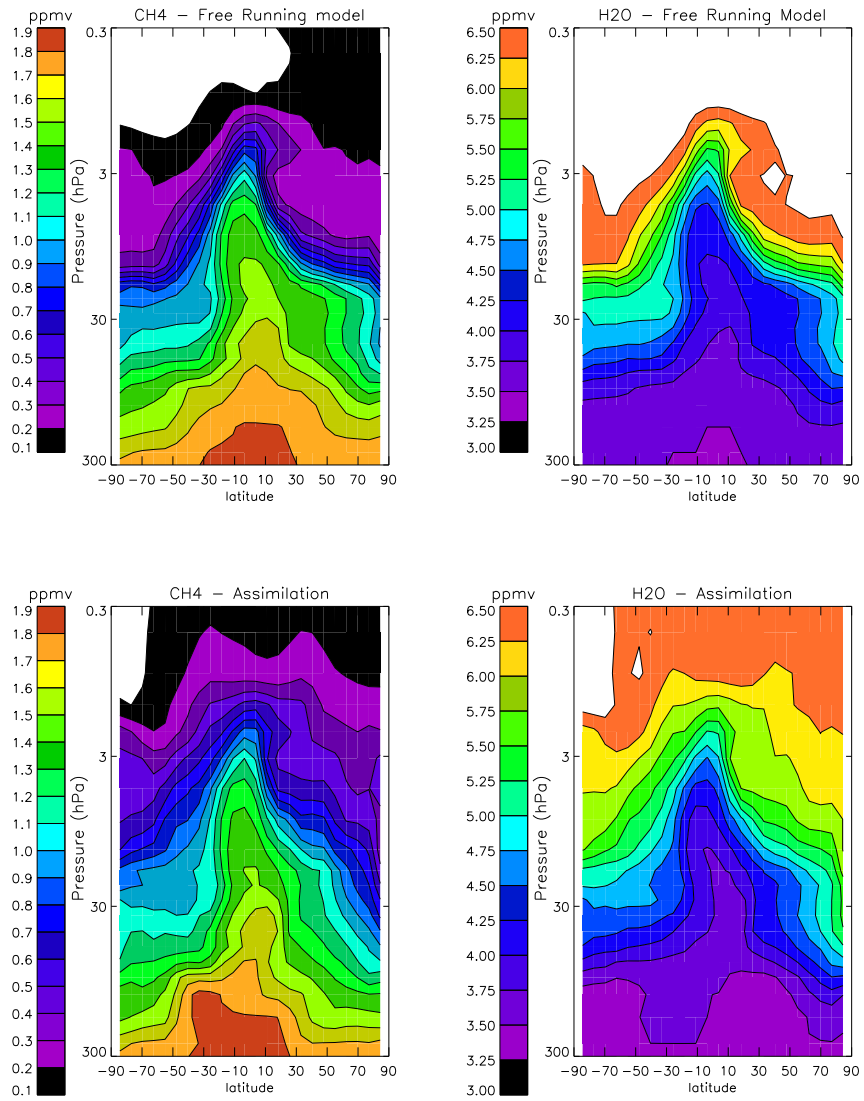
Figure 3.5: *The zonal mean distribution of CH$_4$ (ppmv) and H$_2$O (ppmv) on January 30, 1992 for SLIMCAT runs. The upper plots show results from run* **A** *and the lower plots show results from run* **B**.
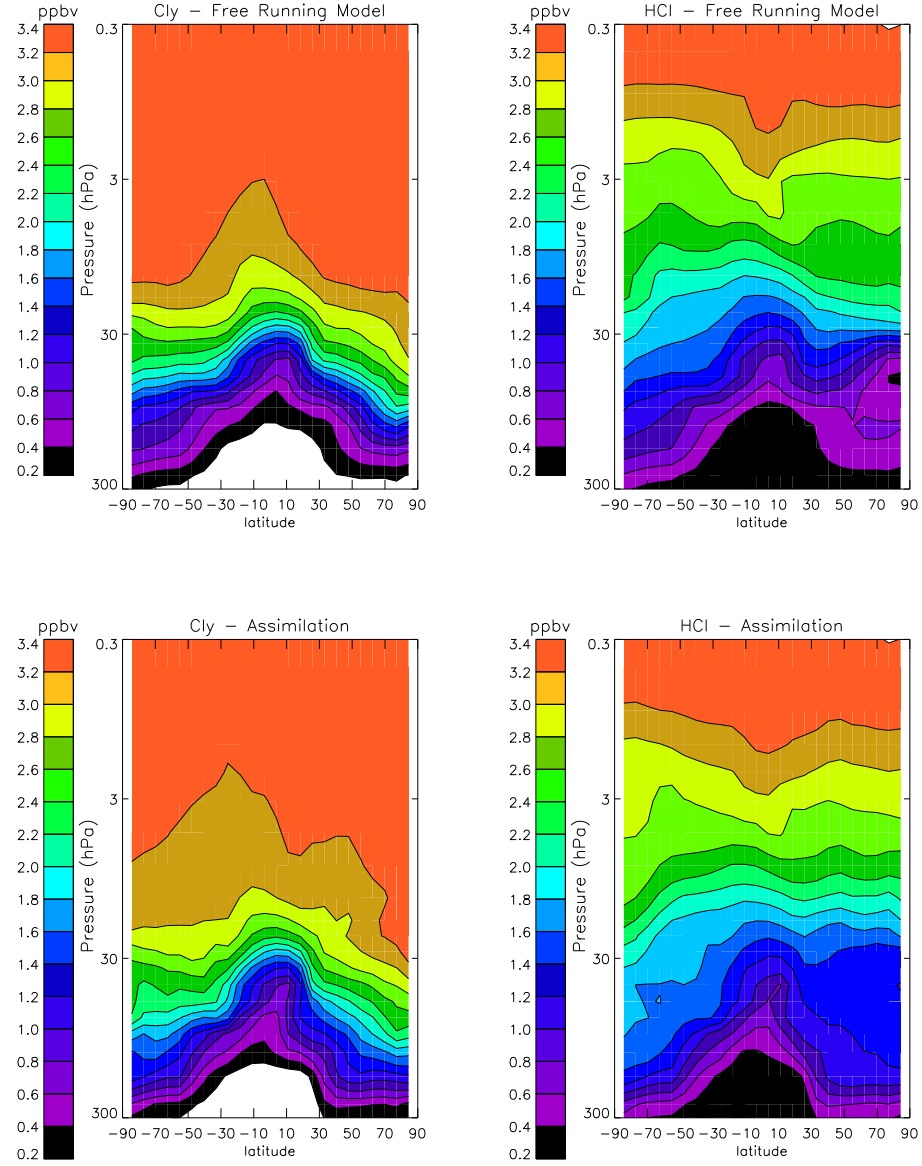
Figure 3.6: *The zonal mean distribution of Cl_y (ppbv) and HCl (ppbv) on January 30, 1992 for SLIMCAT runs. The upper plots show results from run* **A** *and the lower plots show results from run* **B**.

Table 3.3: *Locations of ATMOS profiles at sunset and sunrise in March 1994 used in this study.*

Date	Sunrise/Sunset	Number	Longitude (°E)	Latitude (°N)
08.03.1994	Sunset	11	7.7	-47.6
12.03.1994	Sunset	49	-130.8	38.2
12.03.1994	Sunset	55	94.2	36.3
09.03.1994	Sunrise	17	105.3	-67.8
12.03.1994	Sunrise	41	21.3	-70.4
12.03.1994	Sunrise	44	-46.7	-70.4
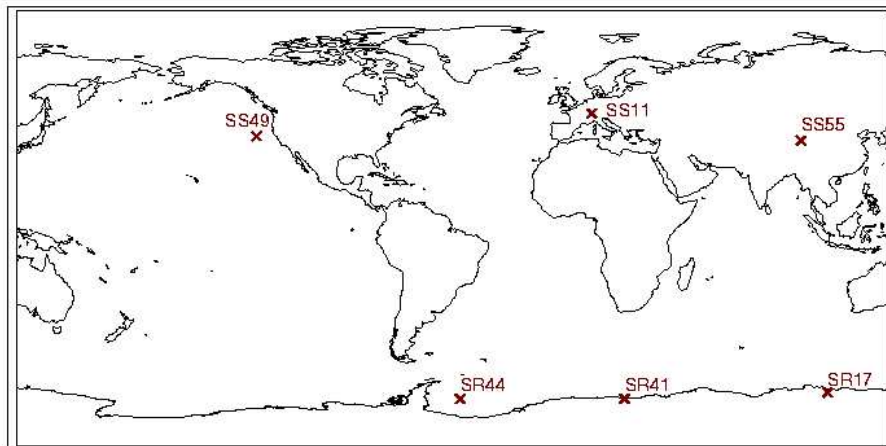
Figure 3.7: *The locations of the different ATMOS measurements used in these comparisons (see Table 3.3).*

As the model outputs data every day at 0 GMT, a 0–D stacked box model based on the same SLIMCAT chemical model was run. This was initialised from the 3-D model output for each day to obtain a model profile at the same local time as the observation. Figure 3.8 shows comparisons of the long–lived tracers, HCl and O_3 with ATMOS observations in November 1994. Figure 3.8 shows that the modelled CH_4 and H_2O profiles from run **A** fit reasonably well with the ATMOS observations; however, the assimilation of the HALOE data corrects the remaining model–observation differences. Although it is not assimilated, the preservation of the model's long–lived tracer–tracer correlations means that in runs **B** and **C** the model N_2O profile is also in much better agreement with the ATMOS data.
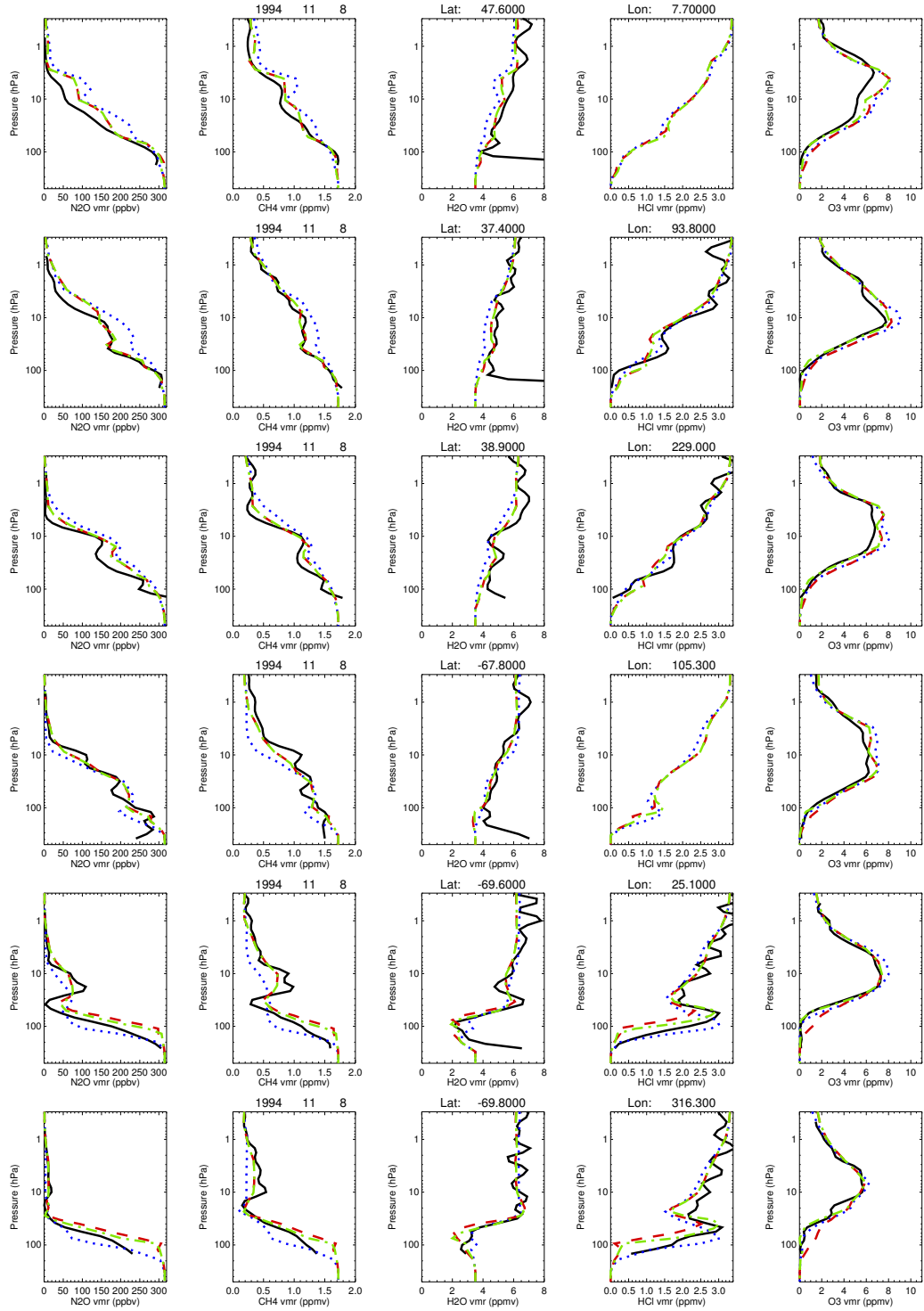
Figure 3.8: *Comparison of ATMOS v3 profiles (black solid line) of (a) N_2O (ppbv), (b) CH_4 (ppmv), (c) H_2O (ppmv), (d) HCl (ppmv) and (e) O_3 (ppmv) with the SLIMCAT runs **A** (blue dotted line), **B** (red dashed line) and **C** (green dashed–dotted line) for 6 profiles in late March 1992 between $17.5°N$ and $51.6°S$. Each plot has species volume mixing ratio on the x-axis and atmospheric pressure (hPa) on the y-axis. The location and date of each observation is printed above the rows of plots.*
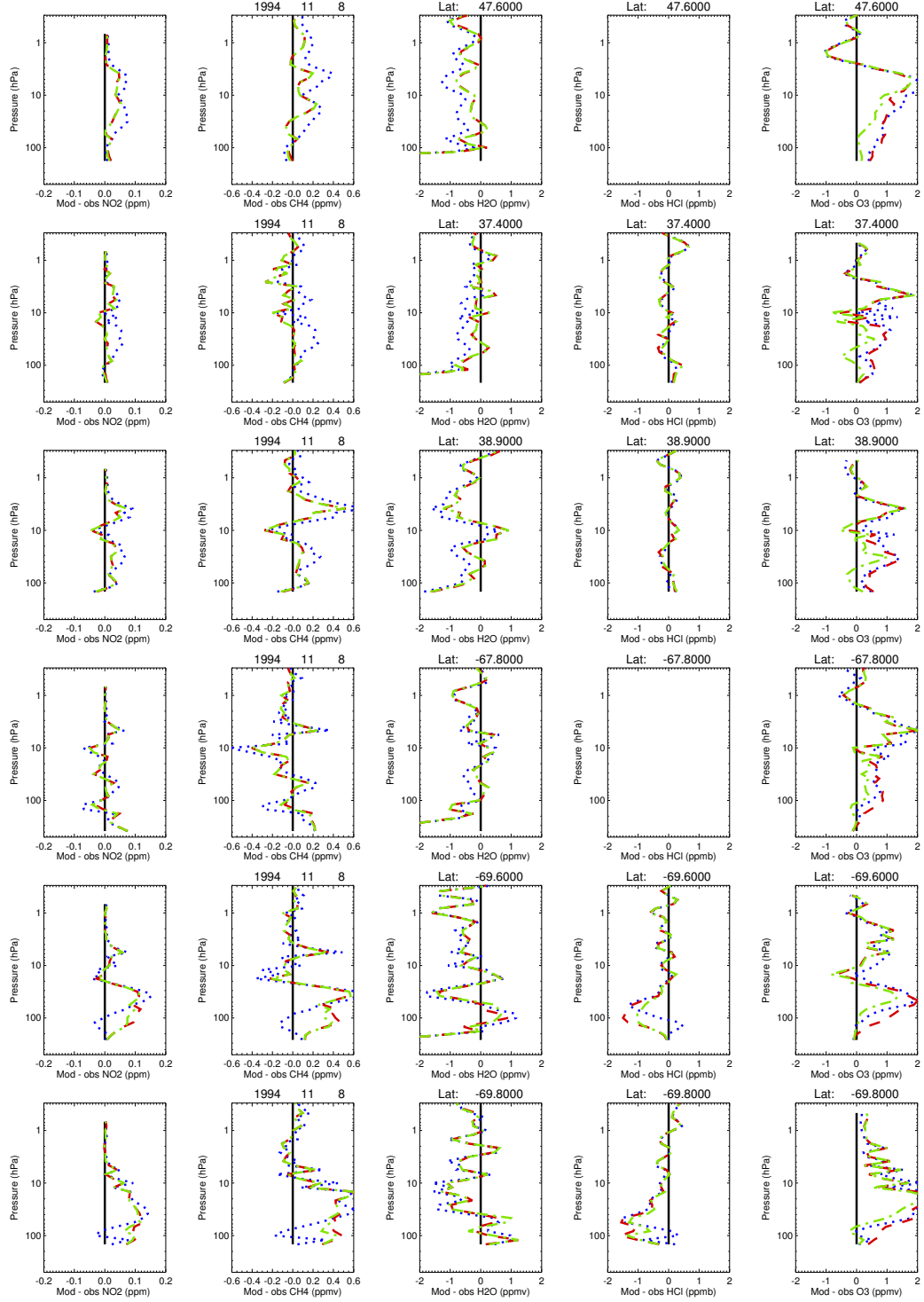
Figure 3.9: *As Figure 3.8 but for the differences between SLIMCAT and ATMOS v3 profiles (i.e. model-obs).*
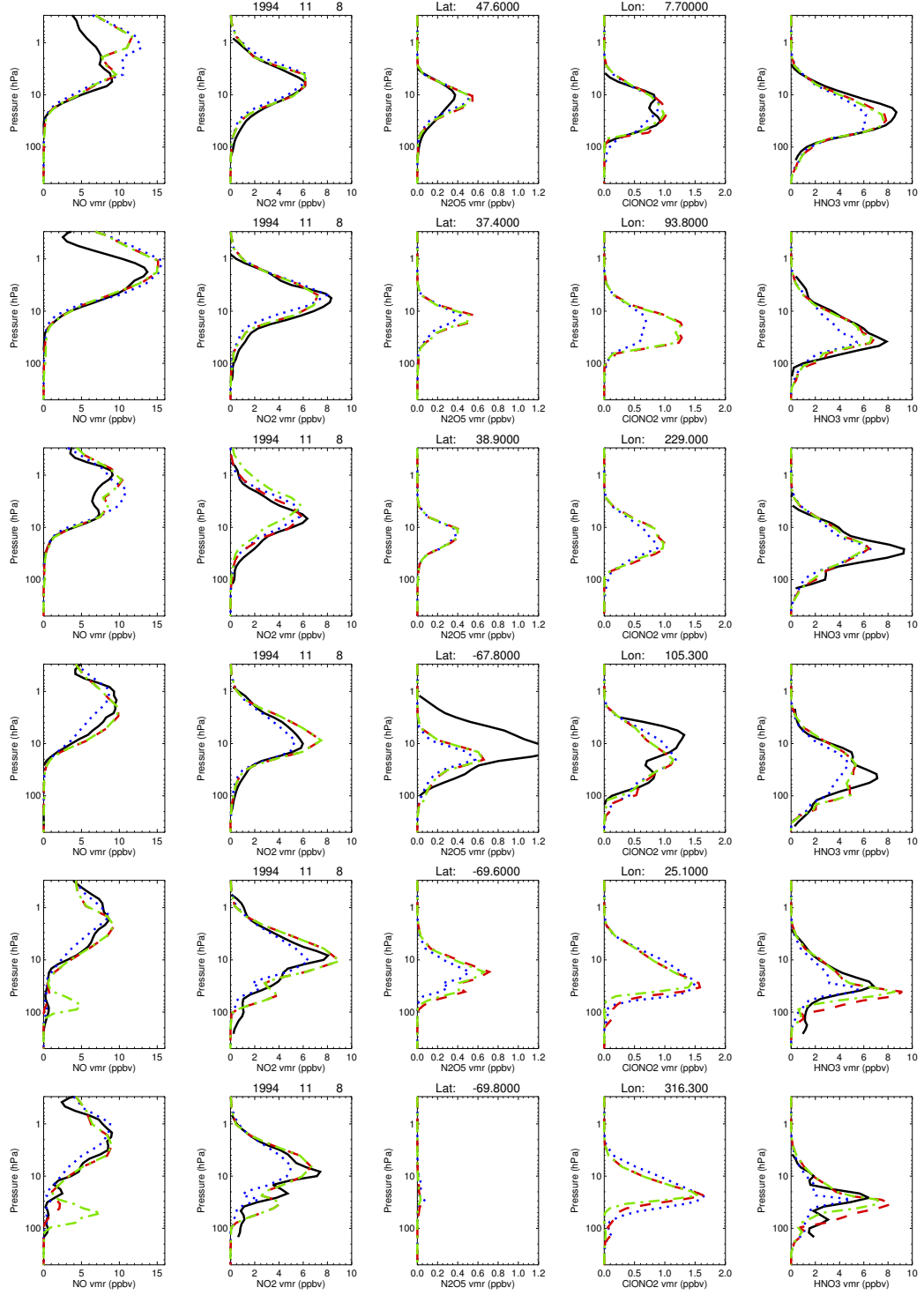
Figure 3.10: *Comparison of ATMOS v3 profiles of (a) NO (ppbv), (b) NO$_2$ (ppbv), (c) N$_2$O$_5$ (ppbv), (d) ClONO$_2$ (ppbv) and (e) HNO$_3$ (ppbv) with the SLIMCAT runs* **A** *(blue dotted line),* **B** *(red dashed line) and* **C** *(green dashed–dotted line) for 6 profiles in late March 1992 between 17.5° N and 51.6° S.*
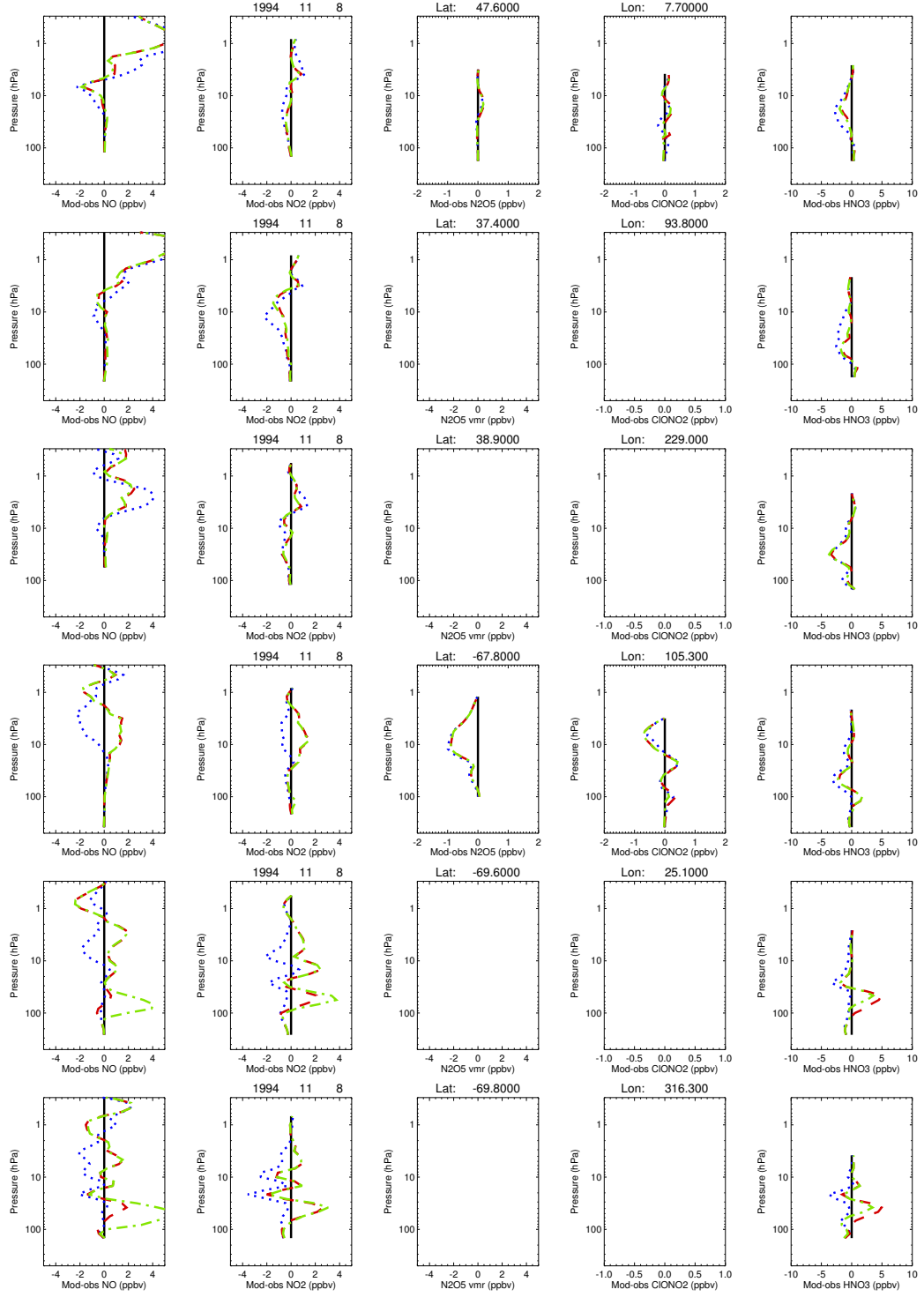
Figure 3.11: *As Figure 3.10 but for the differences between SLIMCAT and ATMOS v3 profiles (i.e. model-obs). Note blanck panels indeicate that no ATMOS observations are available for this species at this location.*

From Figure 3.8 it can be concluded that the chemical data assimilation scheme has corrected for the model transport errors that are known to exist when using the ECMWF ERA40 winds (Schoeberl et al., 2003) (shown by the zonal mean plots in Figure 3.5). The preservation of compact correlations has performed well and improved the fit of other long–lived tracers within the model to independent data, e.g. the ATMOS N_2O comparisons.

3.5.2 Ozone

Figure 3.12 shows the zonal mean plots of O_3, for run **A**, the basic model, and run **B**, the chemical assimilation model for January 30, 1992. The assimilated model, on the right, produces a more confined region of high O_3 concentrations in the mid-latitudes, with concentrations of O_3 in the northern hemisphere from 50-90° between 30–3 hPa being lower by 1–2 ppmv than in the basic model run.
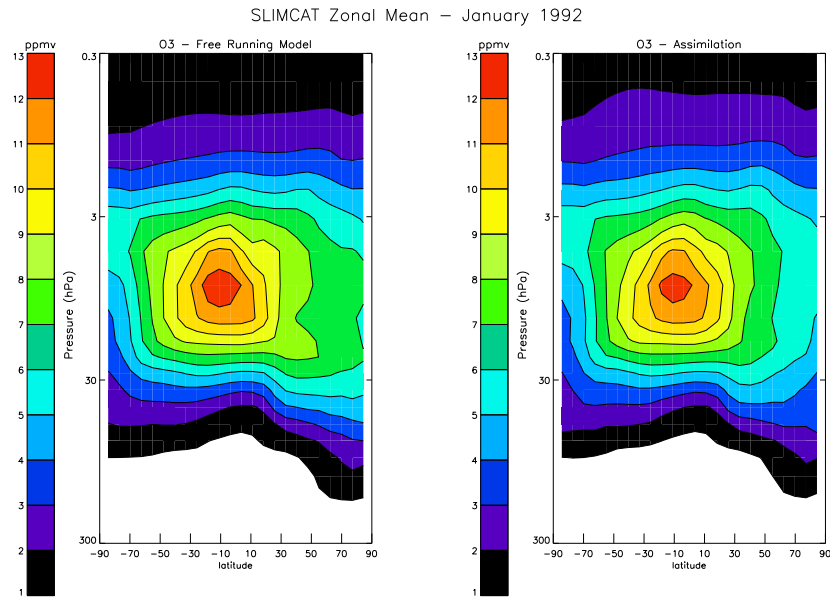
Figure 3.12: *The zonal mean distribution of O_3 on January 30, 1992 for the basic SLIMCAT model,* **A** *(left plot) and the chemical assimilation model,* **B** *(right plot).*

Figure 3.8 showed comparisons of O_3 with ATMOS (far right column of plots). Between 5–30 hPa, where O_3 is relatively long–lived, run **B**, which assimilates O_3, agrees better with the ATMOS data. Above this altitude the short lifetime of O_3 means that there is essentially no direct impact of assimilating O_3 in a sequential scheme. O_3 in the lower stratosphere (where O_3 starts to decrease in concentration, at about 100-300 hPa) is long–lived and therefore, as it more dependent on transport processes, is harder to model. Run **B** does not fit well to ATMOS in this region. This will be investigated further by comparing the model runs to ozone sonde data, for which the measurements are more frequent and the instruments are more accurate (the errors in measurements are smaller).
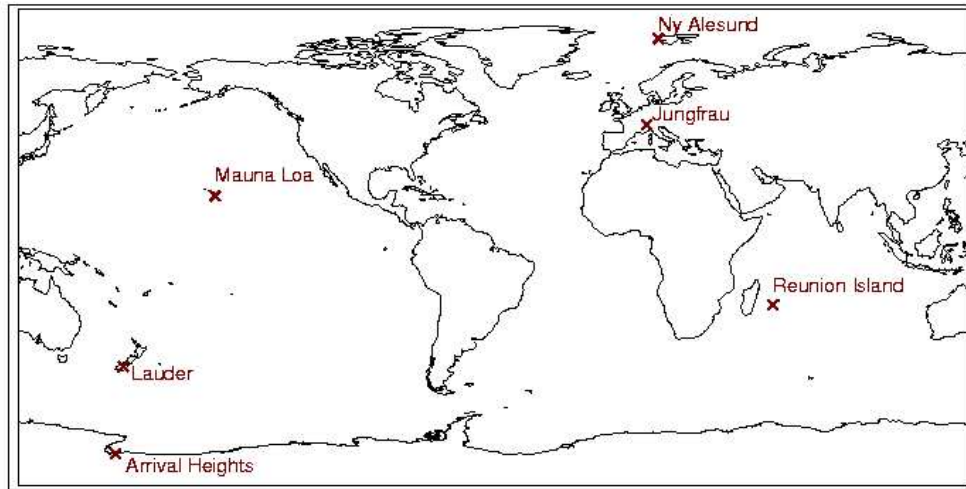
Figure 3.13: *Locations of the key NDACC ground based and O$_3$ sondes sites used in these comparisons.*

An ozone sonde is a balloon-borne instrument that continuously estimates O$_3$ concentration as it ascends up through the atmosphere. Ozone sondes typically reach heights of around 30 km. The ozone sonde data used here has been taken from the international Network for the Detection of Atmospheric Composition Change or NDACC (`http://www.ndacc.org`). NDACC consists of over 70 high quality, remote sensing research stations, each observing the physical and chemical state of the stratosphere and upper troposphere. Figure 3.13 shows some of the remote NDACC locations. These include Jungfraujoch, Switzerland (48°N) and Lauder, New Zealand (45°S).

Figures 3.14 and 3.15 show the observed monthly mean profiles of O$_3$ at the Jungfraujoch and Lauder sites, respectively (different to the ATMOS plots where the comparisons are specific to each measurement). The plots run from January (top left) through to December (lower right) and also show the basic model, **A**, the chemically assimilated model, **B**, the chemically assimilated without ozone model, **C**. The ozone sonde observations are taken on a nearly daily basis and the error bars are the standard deviations of the monthly observations at specific heights in the atmosphere. Runs **A** and **C** in these plots both fail to reproduce observed concentrations of O$_3$ between 300 and 100 hPa. This the area of the atmosphere referred to as the UTLS (Upper Troposphere Lower Stratosphere). The dynamics in this area of the atmosphere are poorly understood and therefore many models do not reproduce realistic transport here (both through horizontal mixing in the lower stratosphere and through convection and advection in the upper troposphere). The comparisons for the two runs **A** and **C** show an overestimation in O$_3$ in this area which is constistent with previous findings in Schoeberl et al. (2003), where the horizontal transport in the lower stratosphere is thought to be too strong. However, run **B** fits the observations much better and within the error, (having a lower O$_3$ concentration in the UTLS) which fits with previous hypotheses that the assimilation will correct for transport errors. Another reason for the poor agreement in this area is that SLIMCAT is a stratospheric model and therefore any comparison below the tropopause will be subject to error as the model

assumes rapid vertical mixing in the troposphere. This could explain why below 400 hPa, i.e. in the lower troposphere, the concentration of O_3 is lower than the observations.

The estimate of the amount of O_3 in the lower stratosphere is important as this is where the highest concentrations are found. The model estimate will affect the total column amount as well as the estimate of NO_2, as O_3 and NO_2 are closely coupled (Figure 3.2). The coupling between O_3 and NO_2 is due to

$$NO + O_3 \longrightarrow NO_2 + O_2 \tag{3.2}$$

and at night time

$$NO_2 + O_3 \longrightarrow NO_3 + O_2 \tag{3.3}$$

$$NO_3 + NO_2 + M \longrightarrow N_2O_5 + M \tag{3.4}$$
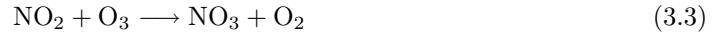
These reactions suggest that an overestimation of O_3 will contribute to an overestimation of the amount of NO_2 in the lower stratosphere. However, for NO_2 this will not have a large effect as the bulk of the column is located between 20 and 25 km.
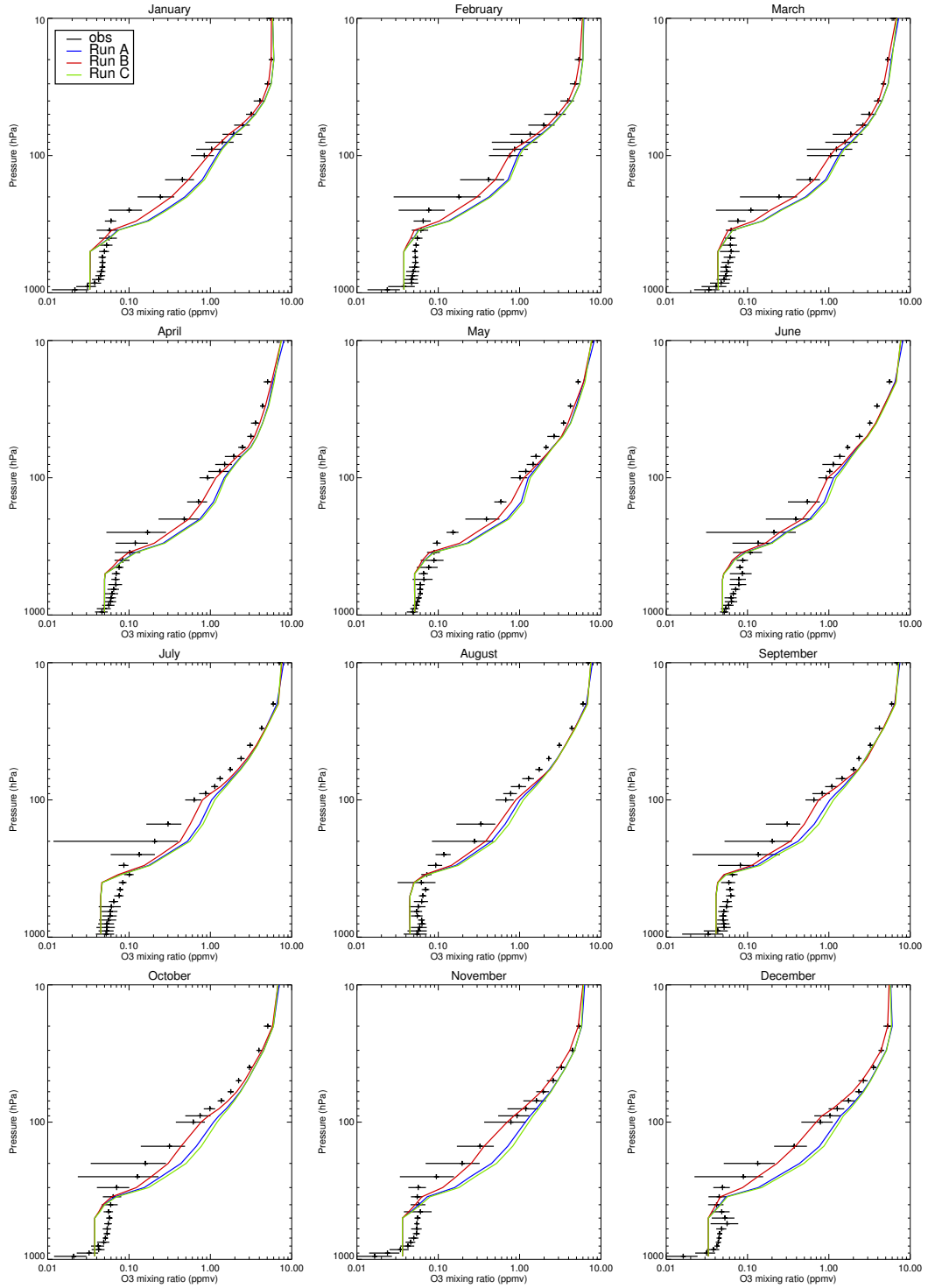
Figure 3.14: *Monthly mean ozone sonde profiles and results from 3 SLIMCAT model runs at Jungfrau-joch Switzland (48° N). The errors bars show standard deviation of the monthly mean. The vertical axis is the atmospheric pressure (hPa) and the horizontal axis is the O_3 mixing ratio (ppmv). The black lines are the observations, the blue line is the basic model, **A**, the red line is the chemically assimilated model, **B** and the yellow line is the chemically assimilated without ozone model, **C**.*
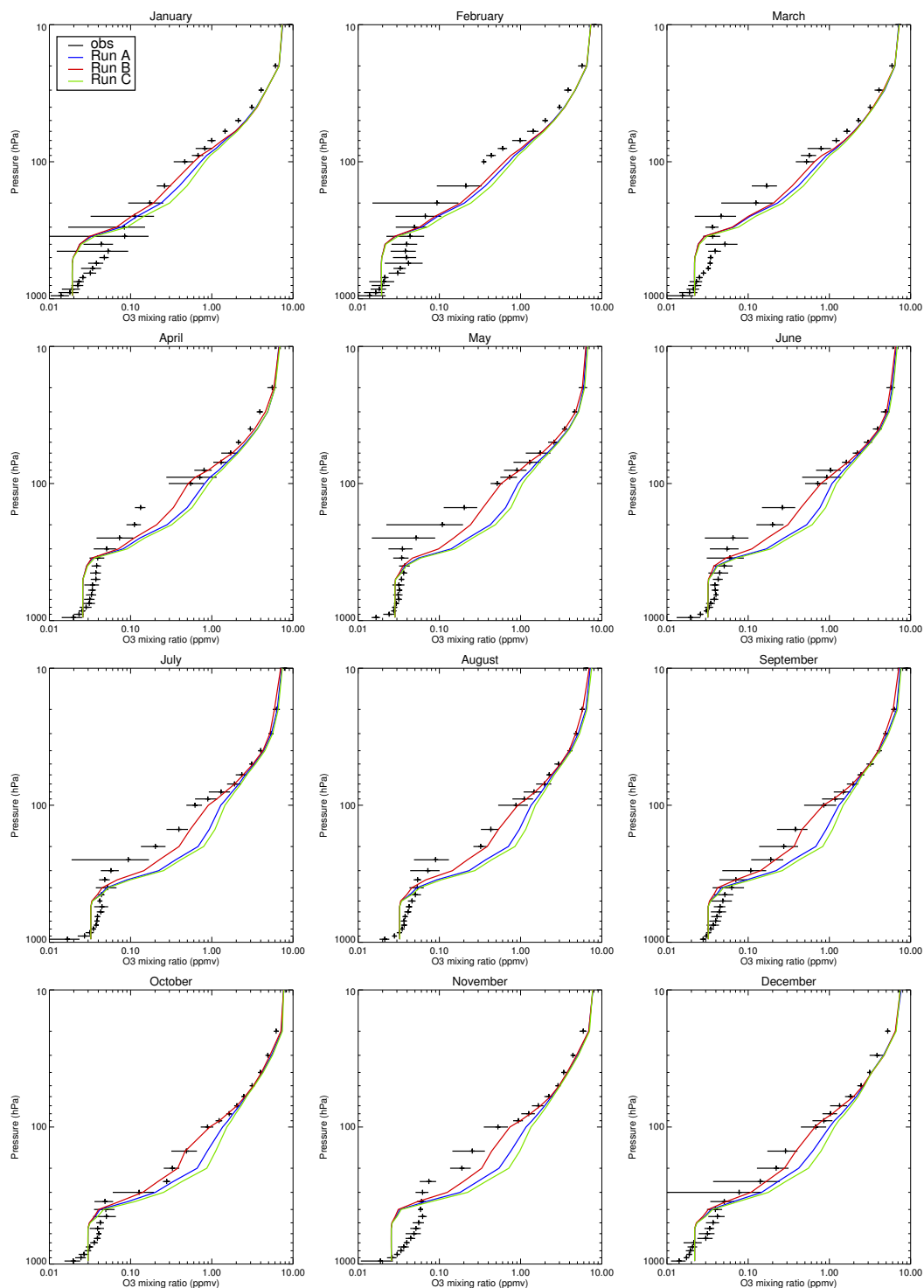
Figure 3.15: *As Figure 3.14 but for Lauder, New Zealand (45°S).*

3.5.3 Reservoir Species

To determine why O_3 is overestimated in the lower stratosphere this section addresses other related key species in the model. These are not directly assimilated, but will be indirectly affected by the chemical data assimilation. Figure 3.16 shows how the amount of $ClONO_2$ varies over two years using the three SLIMCAT runs. This is shown for six NDACC sites: Ny Alesund (79°N), Jungfraujoch (47°N), Mauna Loa (20°N), Reunion Island (22°S), Lauder (45°S) and Arrival Heights (78°S). These observations have been taken using Fourier Transform Infrared Spectroscopy (FTIR) instruments and the locations are shown in Figure 3.13. The Lauder plot in Figure 3.16 shows a good agreement between the data and observations for the model runs **B** and **C**. The comparison at Jungfraujoch agrees better for run **A** in 1997. However, all runs overestimate $ClONO_2$ in 1998, and at Arrival Heights run **A** produces better results when compared to observations.

These results are inconsistent with expectations because the model runs **B** and **C** do not always produce the best agreement with the observed results. However, this could be due to a number of reasons. The partitioning of the Cl_y species could be inaccurate or there could be some issues with the HALOE HCl during this time period.

Figure 3.17 is a corresponding plot for HNO_3. In the northern hemisphere, and at Lauder the comparisons with observations show better agreement with run **B** and **C**. At Arrival Heights run **B** and **C** overestimate the amount of HNO_3.

Figure 3.10 showed the profile comparisons of short–lived and reservoir species along with the ATMOS observations. The reservoir species of interest here are N_2O_5, HNO_3 and $ClONO_2$, these are shown in the last three columns of Figure 3.10. The figure shows that for species which are not directly assimilated run **B** and **C** agree better with ATMOS. This shows that the constraints placed with the model are valid and that the model chemistry performs well. Figure 3.11 shows the same data as Figure 3.10 but as differences between SLIMCAT and ATMOS. This highlights that though overall the assimilation scheme is better above about 50 hPa, it is not as good in the lower stratosphere upper troposphere (below about 50 hPa). This is most notable for HNO_3 at 69.9 S, 25.1 E where the free running model fits the ATMOS observation well below 50 hPa, whereas the two assimilation runs do not. This is the area of the atmosphere that is poorly represented by models as small errors in tropopause height and error in stratoshere–troposphere exchange will cause discrepancies in the concentrations of trace gases.

N_2O_5 is only compared at two locations due to the lack of data, and it was shown that it was greatly underestimated by the model at 67.8°S and slightly overestimated at 47.6°N. N_2O_5 is controlled by two reactions, its daytime photolysis described by the Equation 3.3 and night–time production by NO_2 and O_3, described by the Equations 3.4 and 3.5.

In Figures 3.14 and 3.15, O_3 is overestimated in the lower stratosphere and this could cause too much N_2O_5 to be produced during the night, this could be one reason for the overestimation of N_2O_5 by the SLIMCAT model. However, this does not explain the underestimation. There
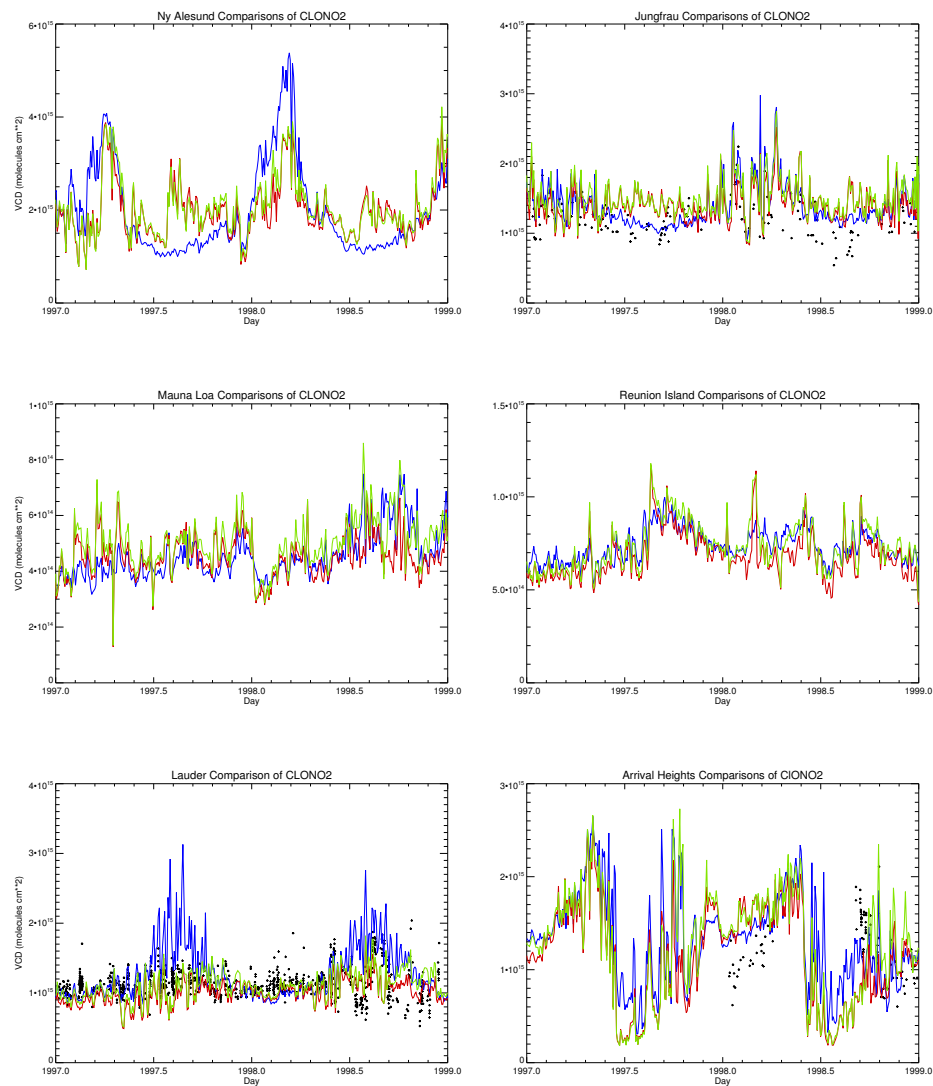
Figure 3.16: *The SLIMCAT model runs and observations of column ClONO$_2$ (molecules cm^{-2}) at 6 NDACC stations, Ny Alesund (79° N), Jungfraujoch (47° N), Mauna Loa (20° N), Reunion Island (22° S), Lauder (45° S) and Arrival Heights (78° S). The black crosses are the observations, the blue line is the basic model, **A**, the red line is the chemically assimilated model, **B** and the green line is the chemically assimilated without ozone model, **C**.*

73
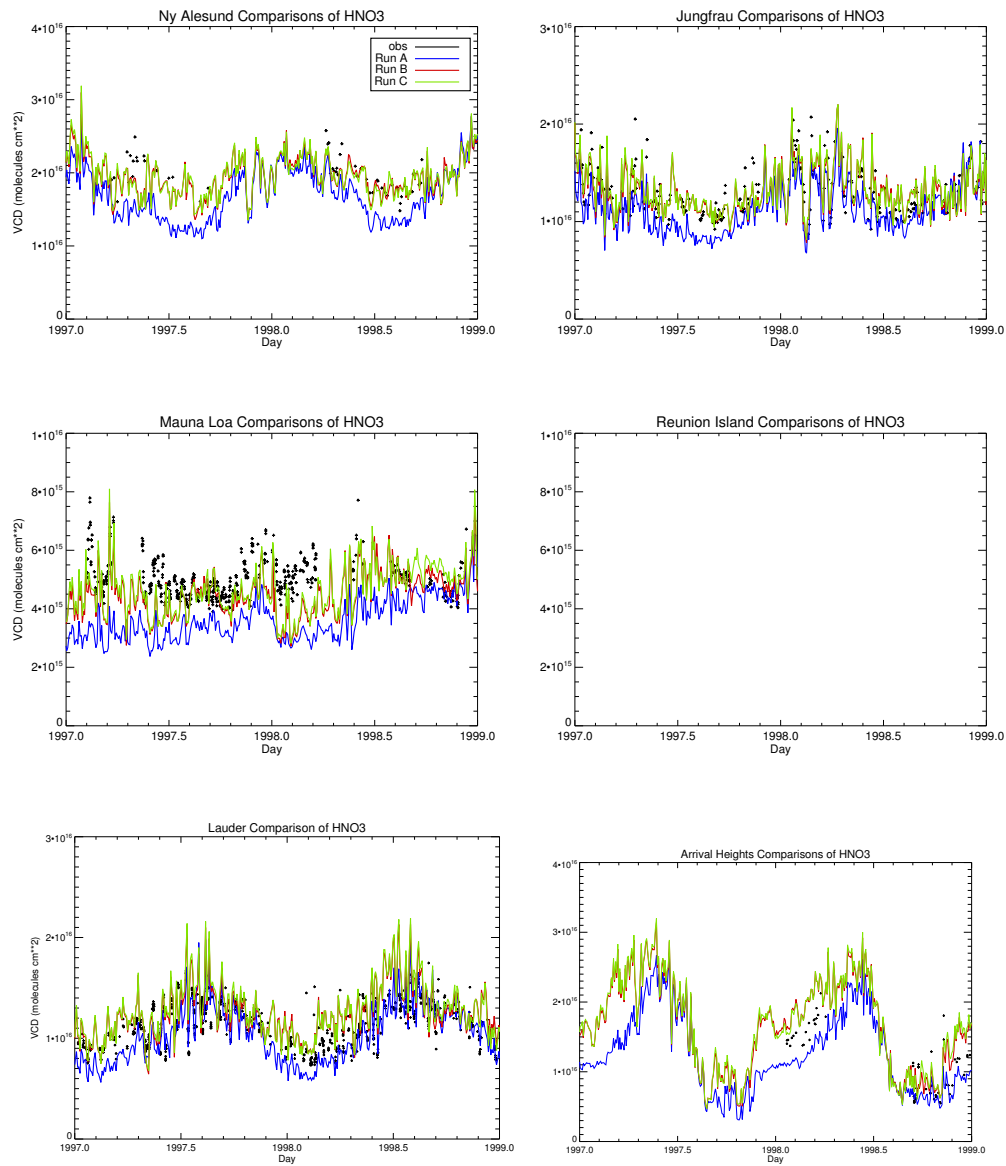
Figure 3.17: *As Figure 3.16 but for HNO₃.*

74

is a need for a more detailed analysis here, but, as it is difficult to measure N_2O_5 in order to make a critical comparison, there is not enough data. The ATMOS comparisons with $ClONO_2$ and HNO_3 show better agreement for the chemically assimilated runs **B** and **C**, although there are only a few comparisons in the case of $ClONO_2$.

3.5.4   Short-lived species

In the SLIMCAT model runs that have chemical data assimilation the short–lived species are not modified directly. In all runs the abundances of short–lived species are determined by their chemical partitioning within long–lived chemical families. Therefore, if the assimilation scheme improves the long–lived tracers it would be expected that the estimation of short–lived species will also become more realistic. This, of course, depends on the model having a realistic treatment of fast chemistry.

Figure 3.10 showed a comparison of the short–lived species, NO and NO_2, with ATMOS data. The assimilation process affects their distribution by modifying the abundance of the long–lived total family NO_y through the tracer–tracer correlations (as shown by N_2O and CH_4 in Figure 3.7). N_2O is the major source of NO_y in the stratosphere and as this is constrained through its correlation with CH_4, CH_4 assimilation will also impact NO_y. This is then partitioned into NO and NO_2 as well as other species. This chemical partitioning relies on the model chemistry scheme but uses a more accurate estimate of NO_y, from the assimilation scheme, to calculate the contributions of NO and NO_2. Figure 3.11 highlights the differences between ATMOS and the SLIMCAT runs and shows that for NO and NO_2 the agreement with runs **B** and **C** is better.

For NO and NO_2 located below 10 hPa the assimilation runs **B** and **C** show improved agreement with the ATMOS profiles. This is the region which will be most important for the column contributions. In general, NO and NO_2 have been shown to be more realistic in runs **B** and **C**. This suggests that the model NO_x fast chemistry is accurate. This is consistent with other studies which have shown good agreement between photochemical box models and in–situ stratospheric NO_x data (Lary et al., 1991).

3.6 Long–Term Trends

3.6.1 NO$_2$

Section 3.5 showed that chemical data assimilation in the model can improve the modelled distributions of key NO$_y$ species and can also correct for transport errors produced by analysed winds. This knowledge can now be used to investigate the implications for long–term modelled NO$_2$ variations. Model studies of long–term trends in NO$_2$ are of interest as previous studies (e.g. WMO, 2002) have shown that stratospheric NO$_2$ long–term trends do not agree with observations.

The source of stratospheric NO$_y$, and therefore NO$_2$, is N$_2$O. Liley et al. (2000) observed an increasing stratospheric NO$_2$ trend at Lauder of 5±1% per decade, where as N$_2$O had been increasing at a rate of only 2.4% per decade from 1986-2005 (WMO, 2002). This increasing trend is believed due to the increasing stratospheric NO$_y$ due to increasing N$_2$O and also the decreasing aerosol following the eruption of Mt Pinatubo in 1991. There is also a contribution from the effect of decreasing O$_3$ (Fish et al., 2000; McLinden et al., 2001).

Figures 3.18 and 3.19 show the comparison of modelled sunrise and sunset column NO$_2$ for 1992–2000 for 4 NDACC stations, Jungfraujoch, Issyk–Kuhl, Tenerife and Lauder. The comparison of the basic model, run **A**, with the observations is quite poor especially in the early 1990s. The model captures the annual cycle but significantly underestimates the magnitude of the summertime maximum, this was by $\sim$50% at Lauder in 1992/93. Column NO$_2$ produced in run **A** does increase during this period and by 1999/2000 the agreement is much better. Evidently the modelled NO$_2$ trend in run **A** is too large.

The annual cycle of NO$_2$ at Tenerife was studied in more detail and results were published in Gil et al. (2008) (see Appendix B). The free-running model (run **A**) underestimated the observed column by up to 12% at sunrise and 25% at sunset. However, runs **B** and **C** improved this greatly.

Both of the model runs which include chemical data assimilation, **B** and **C**, show a much better agreement throughout the whole period. The improvement is especially marked in the early 1990s. The modelled summertime maximum is now in good agreement with the observations. Moreover, the modelled NO$_2$ trend over this period agrees with the observations. The improvement in modelled NO$_2$ in runs **B** and **C** is due to improvements in the modelled distribution of NO$_y$ via its correlation with the long–lived tracer CH$_4$.
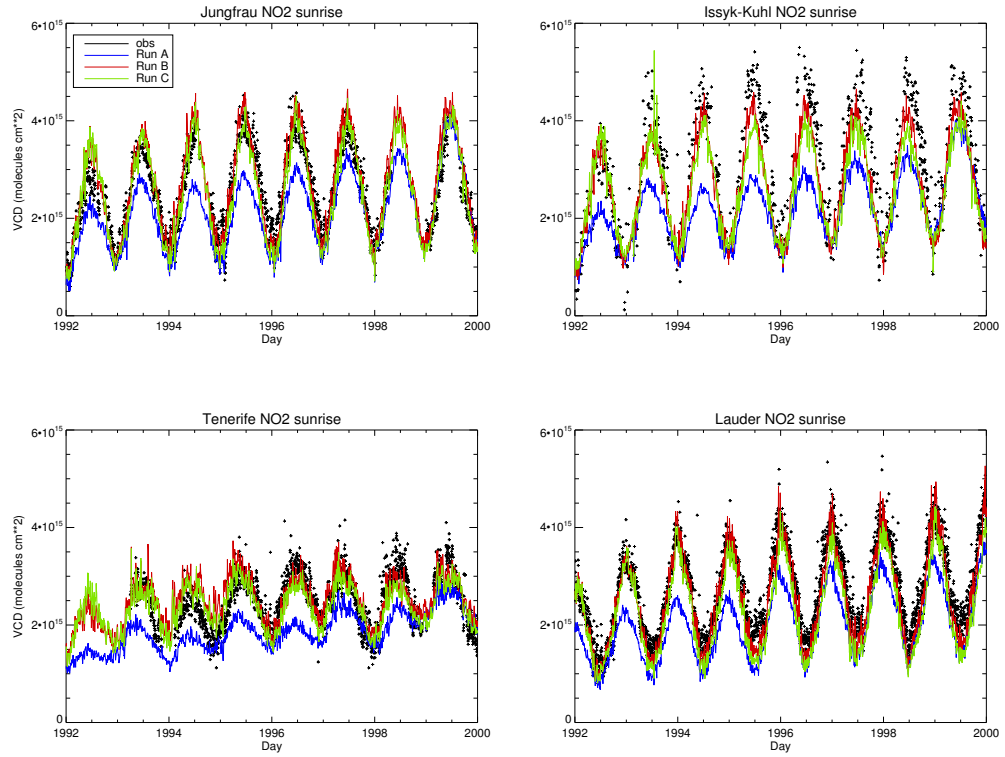
Figure 3.18: *The vertical column NO_2 (molecules cm^{-2}) from 1992 to 2000 for ground based DOAS sunrise measurements along with SLIMCAT model runs. The blue line is the basic model, **A**, the red line is the chemically assimilated model, **B** and the yellow line is the chemically assimilated without ozone model, **C**.*
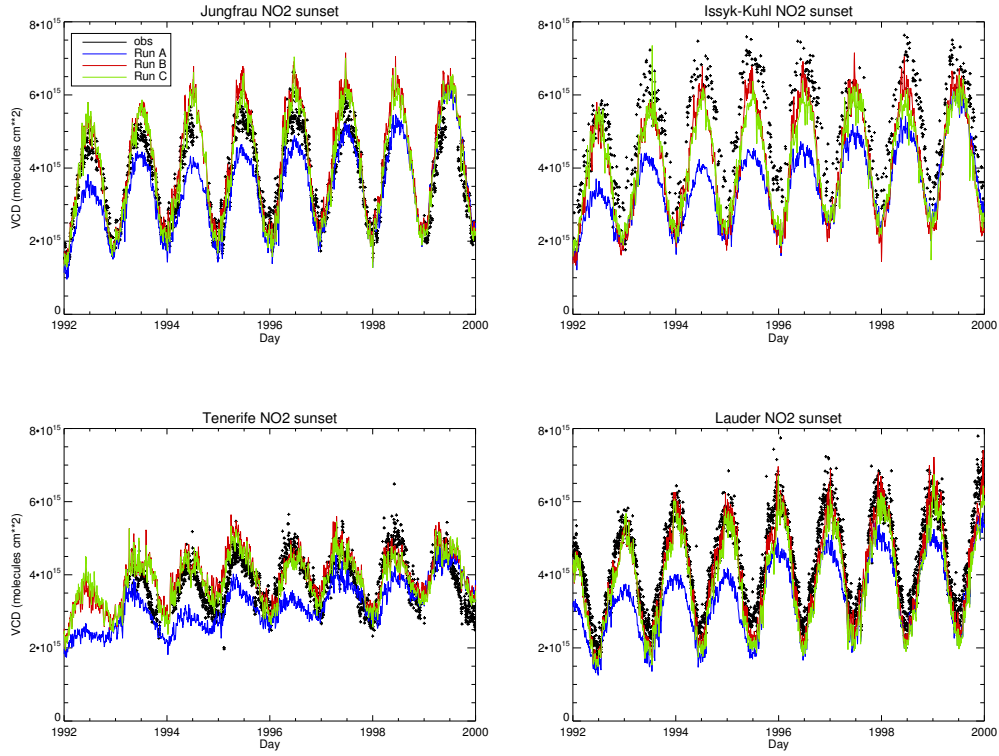
Figure 3.19: *As Figure 3.18 but for sunset.*

The difference in column NO_2 between runs **B** and **C** is small. This indicates that the impact of changing model O_3 does affect the daytime abundance, i.e. via the $[NO]:[NO_2]$ partitioning. However, the O_3 field produced by the free–running model, **A**, is realistic enough in the important altitude region, 20–30km, that any improvement produced using assimilation only causes small changes to the column NO_2.

The night time conversion of NO_2 to N_2O_5 is determined by Reactions 3.4 and 3.5 and N_2O_5 itself can thermally decompose to reform NO_2 and NO_3, this reaction would be given by

$$N_2O_5 + M \longrightarrow NO_3 + NO_2 + M \tag{3.5}$$

If the rate of Reaction 3.4 is fast at night, and Reaction 3.5 is not important, NO_3 can be placed in a night time, steady state and the rate of loss of NO_2 becomes

$$\frac{-d[NO_2]}{dt} = -2k_1[NO_2][O_3], \tag{3.6}$$

where k_1 is the rate of Reaction 3.3. The overall rate of decay of NO_2 between sunset and sunrise can then be written as
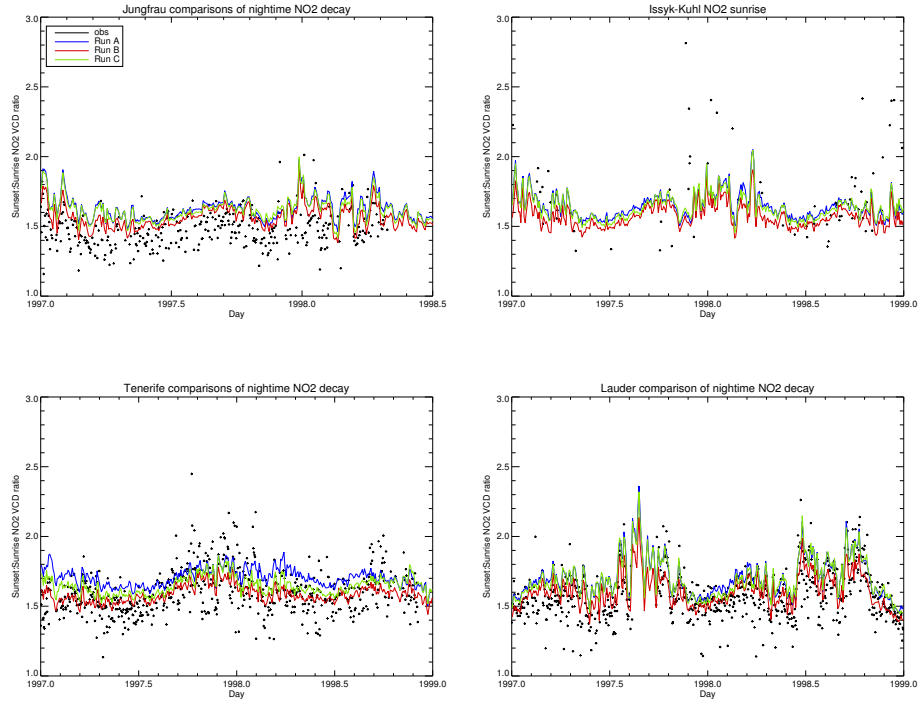
Figure 3.20: *Comparison of NO$_2$ night time decay from 1992–1998 from ground–based DOAS measurements with SLIMCAT runs* **A**, **B** *and* **C**.

$$[NO_2]_{sr} = [NO_2]_{ss} \exp(-2k_1[NO_2][O_3]). \qquad (3.7)$$

Therefore, the ratio of sunset:sunrise NO$_2$ is determined by atmospheric temperature (via k_1) and the concentration of O$_3$. Figure 3.20 shows the observed ratio of column sunset:sunrise NO$_2$ for 4 stations from 1997–1998, along with results from the 3 model runs. The modelled ratio is very similar between runs **A** and **C**. All model runs used the same temperature fields and the O$_3$ fields in these 2 runs in the altitude region important for the NO$_2$ column are also similar. The modelled ratio from run **B**, which assimilates O$_3$, is slightly smaller. However, this does not impact the agreement with the observations which in all cases is reasonable. The worst agreement is at Jungfraujoch where the model ratio tends to overestimate the observations. The overall good agreement, however, indicates that this aspect of NO$_2$ chemistry, i.e. the nighttime decay, is well modelled. This supports the argument that the improvements seen in model column NO$_2$ in Figure 3.18 and 3.19 is due to changes in NO$_y$, and not a chemical effect.

Important results from Figure 3.18 and 3.19 are the implications for the cause of the modelled NO$_2$ trend. Given the poor agreement of run **A** it would have been tempting to question whether the modelled aerosol trend was realistic. However, all runs use the same aerosol values and can reproduce the observations well. Overall the results point to the cause of the poor agreement in run **A** being principally due to the transport related errors in the forcing ERA–40 winds.

3.6.2 O_3

The results shown in Figures 3.14 and 3.15 are of interest when considering long–term changes in column O_3. Analysed wind fields are used to quantify the role of dynamics in long–term changes in column O_3 (e.g. WMO, 2002). This can be via their use to force long–term CTM simulations or via direct diagnosis of dynamical quantities. The results presented here add a further note of caution to the process of taking long–term reanalysis datasets and using them to derive trends. Even when created with a constant assimilation system, the reanalyses may have discontinuities when different datasets become available for assimilation.

The results in Feng et al. (2007) show that the modelled mid–latitude O_3 trends from the basic model would not be realistic in the 1990s. Feng et al. (2007) presented long–term O_3 variations from a run very similar to run **A**. The long–term O_3 change was expressed as a percentage anomaly relative to 1980 (shown in Figure 3.21). This figure shows that while the model captured the overall long–term change from 1980–2002 there were periods where the model disagreed significantly, e.g. an overestimate in the late 1980s and an underestimate in the early 1990s. Indeed, Figure 3.21 of Feng et al. (2007) shows that model run **A** has a too strong increase in column O_3 from 1993 to around 1998. As we are now able to assimilate a long–lived tracer from 1991 onwards we can interpret this underestimate of O_3 in the early 1990s as being due at least in part to errors in the long–term variations in ERA–40 analyses. Unfortunately, we cannot repeat the 'anomaly' analysis of Feng et al. (2007), where O_3 changes are expressed as differences will respect to 1980, as we do not have observation to constrain the assimilation runs back this far.
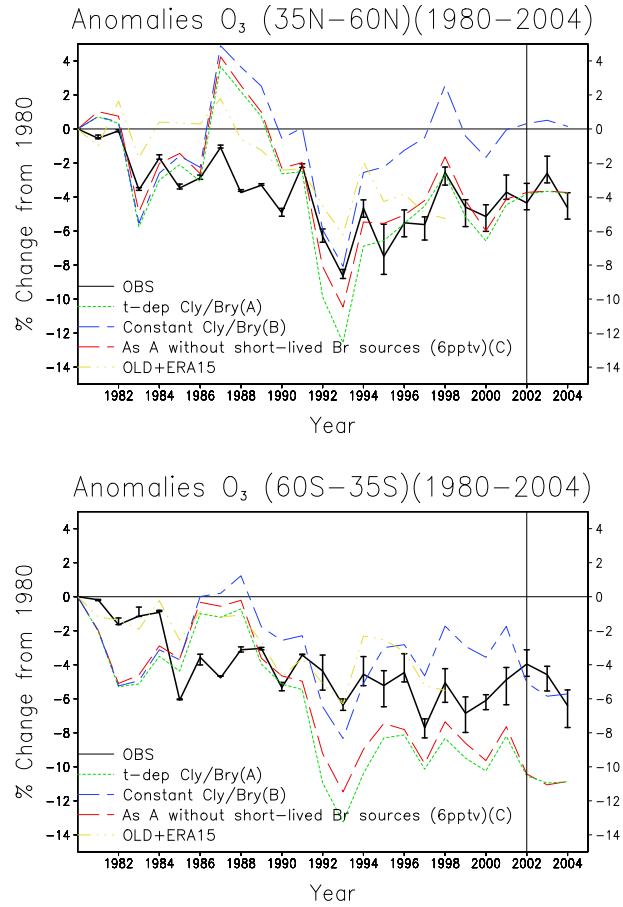
Figure 3.21: *Figure 4 from (Feng et al., 2007). Changes in satellite observed zonal mean, annual mean column O_3 compared to 1980 weight averaged within latitude bands (a) $35°N$–$60°N$ (b) $35°S$–$60°S$. The run 't-dep Cl_y/Br_y (A)' is very similar to run **A** in this thesis.*

81

3.7 Quantifying SLIMCAT errors

Although the previous sections have shown that the chemical data assimilation scheme within SLIMCAT agrees well with ground–based and satellite observations for long–lived and short lived species, it is necessary, for this thesis, to correct the model for any discrepancies in the total column. This is required as any error in the stratospheric column will cause an error in the retrieved tropospheric column. In order to obtain a correction factor for the SLIMCAT stratospheric total columns as a function of latitude, the ground–based comparisons have been used.

3.7.1 NO$_2$ Correction

In the case of the NO$_2$ the comparisons carried out at 6 ground–based stations (Ny Aselund, Junfraujoch, Mauna Loa, Reunion Island, Lauder and Arrival Heights) have been analysed to obtain a monthly mean correction factor for each latitude (to correct for SLIMCAT systematic error). This assumes that the error is longitudinally invariant and the correction can be linearly interpolated onto a 10° latitude grid to obtain a smoothed correction factor.

Figure 3.22 shows the SLIMCAT total column NO$_2$ for July 1997 after the systematic error has been removed, the random error in the NO$_2$ column (the standard deviation of the data) and the percentage random error in the column NO$_2$.

A number of previous studies assume that the NO$_2$ distribution in the stratosphere is zonally symmetric. As can be seen in the top left panel of Figure 3.22 this is not the case. One of the main improvements in this retrieval scheme compared to previous studies is the treatment of the stratosphere. The error in the NO$_2$ stratospheric column has been added to Table 1.2 and is shown in Table 3.4. The errors in the stratospheric column in this thesis are comparable and slightly larger than those reported in previous studies. However, the error reported from this work are purely random and will reduce when considering monthly means. Also the work carried out in this thesis uses a more sophisticated method for determining the stratosphere, and it will be seen in Chapters 5 and 6, that this results in a smaller overall error in the retrieved tropospheric column.

3.7.2 BrO Correction

As stated in Section 3.2.3 the SLIMCAT model run used in this work has two new species a halon and a very short lived species (VSLS). Comparisons of previous SLIMCAT model runs with ground-based BrO showed that the stratospheric Br$_y$ was too low as BrO accounts for half the Br$_y$ (Sinnhuber et al., 2002). The introduction of the two new species in the model scales the Br$_y$ loading to fit the ground-based measurements of BrO. The errors in the stratospheric BrO column will therefore be equal to those stated in arriving at these conclusions (10%) and these values are used here.
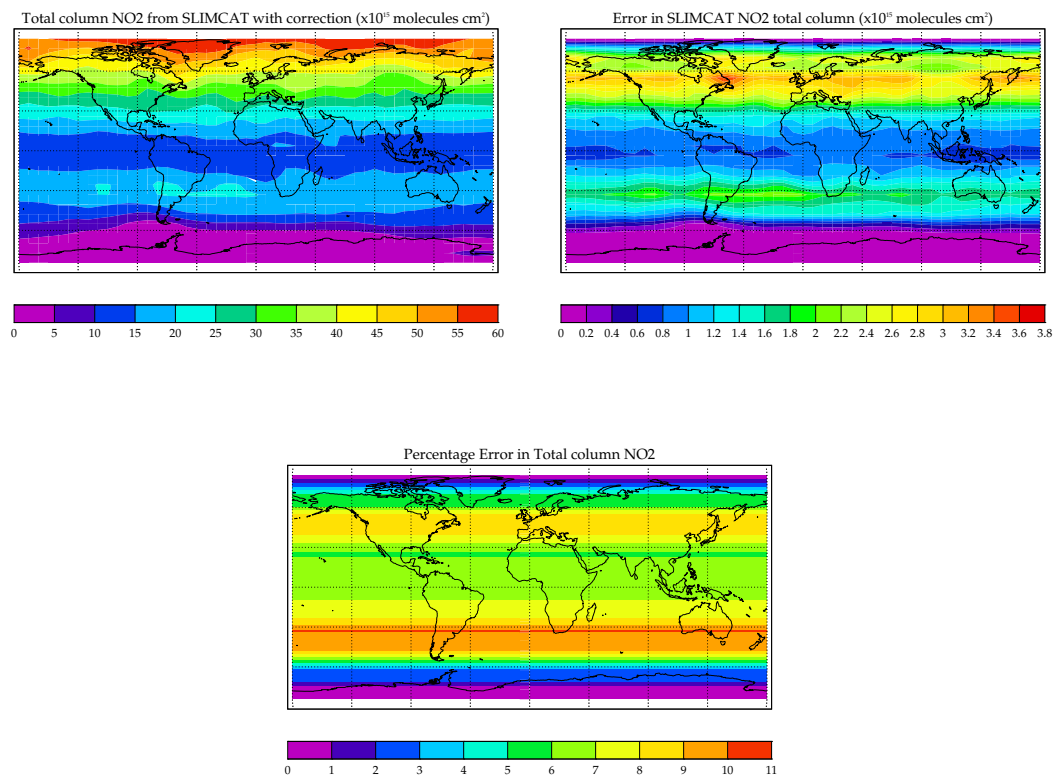
Figure 3.22: *SLIMCAT NO$_2$ total column for July 1997, (a) corrected SLIMCAT (1×10^{16} molecules cm^{-2}) (b) error in the corrected SLIMCAT total column (1×10^{15} molecules cm^{-2}) (c) percentage error in the corrected SLIMCAT total column.*

Authors	Method	Uncertainty
Leue et al. (2001)	IPF	$0.08–0.5 \times 10^{15}$ molec. cm^{-2}
Richter and Burrows (2002)	Ref. sector	1×10^{15} molec. cm^{-2}
Martin et al (2002)	Ref. sector	0.2×10^{15} molec. cm^{-2}
Richter (2002)	CTM Ref.	(not given)
Boersma et al (2002)	Ref sector	0.45×10^{15} molec. cm^{-2}
Boersma et al (2004)	assimilation	0.2×10^{15} molec. cm^{-2}
This Thesis (2008/2009)	assimilation	$0.1–3 \times 10^{15}$ molec. cm^{-2}

Table 3.4: *Errors associated with different stratospheric column density calculations (taken from Boersma et al., 2004).*

3.8 Summary

This chapter has aimed to model accurately specific trace gases in the stratosphere (namely NO_2, BrO and O_3). To achieve this the SLIMCAT 3-D CTM constrained with chemical data assimilation has been used. Three model experiments have been run and comparisons with multiple observational datasets have been performed.

Profile measurements from the ATMOS space–based instrument have been used for model comparisons. These comparisons show that long–lived tracers are improved through the chemical data assimilation, as would be expected. Also N_2O, which is not directly assimilated by constrained in the model by its correlation to CH_4, is also improved. Reservoir species ($ClONO_2$, HNO_3 and N_2O_5) and short–lived species (NO and NO_2) are constrained through partitioning within their chemical families NO_y and Cl_y. The comparisons with these species show that the chemical assimilation scheme has also improved these species. This leads to the conclusion that the model stratospheric chemistry is largely correct, but the model transport causes the largest discrepancies.

Ground–based comparisons of total column amounts have been performed for a number of NDACC sites that have FTIR measurements of reservoir species (HNO_3 and $ClONO_2$) and long–term NO_2 from DOAS datasets. For HNO_3 the comparisons in the northern hemisphere and Lauder show better agreement for runs with chemical data assimilation however for Arrival Heights these runs overestimate the amount of HNO_3.

Long–term trends in NO_2 and O_3 were investigated. The basic model significantly underestimates the observed column NO_2 from mid-latitude ground–based sites before about 1998. As the model agreement is worse in the early 1990s, it overestimates the increasing trend later in the decade. However, when the model assimilates HALOE CH_4 data both comparisons are greatly improved. These results have implications for the use of analysed winds in determining the dynamical contribution to column O_3 changes, especially when using them to force a transport model. Clearly, modelled column O_3 changes will be sensitive to errors in the analysed winds. Unfortunately, the availability of long-lived tracer data for assimilation does not extend over the long period typically used for stratospheric O_3 trend studies (e.g. 1980 to present day).

Assimilation of the long–lived species corrects for transport errors in the model (i.e. from the

ERA-40 reanalyses) and allows a more direct test of the model's chemistry. With this constraint, it appears that the current understanding of atmospheric chemistry can then quantitatively reproduce observed long-term column NO_2 changes. However, the period over which this study can be performed is limited by the availability of suitable tracer observations to assimilate.

The ground–based comparisons have been used to correct the SLIMCAT total column amounts that will be used in the retrieval scheme developed within this thesis, i.e. for species NO_2 and BrO. For NO_2 the correction factor has been calculated and it has been shown how it can be applied to the stratospheric column. For BrO no correction factor is necessary as BrO has been altered in the SLIMCAT run by adding a new halon and very short–lived species to account for Br_y discrepancies previously observed.

This chapter has described the method for calculating the stratospheric column which is used in the retrieval model developed in this thesis to remove the stratospheric part of the total column and to leave the tropospheric column. Errors in the stratospheric column have been found by validating the results of the SLIMCAT model run with chemical data assimilation with ground based measurements. The error in the stratospheric column is summarised in Table 3.5.

Species	Parameter	Error
NO_2	SCD error	5–9×10^{15} mol. cm^{-2}
BrO	SCD error	1.5–1.8×10^{13} mol. cm^{-2}
NO_2	Previous study VCD_{trop} error	0.5–10×10^{15} mol. cm^{-2}
BrO	Previous study VCD_{trop} error	na
NO_2	Scalar approximation	upper limit 10% of VCD_{trop}
BrO	Scalar approximation	10% of VCD_{trop}
NO_2 / BrO	Error due to cloud and aerosol optical models	to be quantified in Chapter 4
NO_2 / BrO	Plane–parallel assumption	only significant at high sza
NO_2	Stratospheric Col.	0.5–7% (strat. col.)
BrO	Stratospheric Col.	10% (strat. col.)

Table 3.5: *The errors associated with assumptions or parameters that will be used in the new retrieval model developed in this thesis for NO_2 and BrO.*

The impact of the error in the stratospheric column on the retrieved tropospheric vertical column density will be quantitative discussed in Chapter 4 and 5.

Chapter 4

Retrieval Model–Development and Testing

This chapter describes the retrieval model that has been developed in this thesis. Chapter 2 introduced the concept of radiative transfer in the atmosphere and briefly summarised how it can be modelled if the relevant atmospheric parameters are known. In this thesis, radiative transfer needs to be 'inversely modelled'. This is the purpose of the retrieval model and its development is outlined in this chapter.

This chapter also describes a number of tests that have been performed in order to better understand which atmospheric and model parameters are the most sensitivity to changes (or errors). The final part of this chapter investigates how ATSR–2 and GOME data can best be combined.

Section 1.4.1 explained how the GOME instrument provides measurements of radiances, in the form of spectra, which are used to calculate the slant column densities of different atmospheric species, e.g. NO_2 or BrO. The aim of the retrieval model is to take these slant column densities, along with other measurements of the atmosphere (cloud, aerosol and surface Lambertian Equivalent Reflectance (LER)), and infer from them information about the state of the atmosphere, in particular the tropospheric vertical column densities of the relevant absorbing species. This chapter will outline how this has been achieved. Section 4.1 introduces retrieval theory and derives the important mathematics which will help in the understanding of this problem. Section 4.2 describes the retrieval model that has been developed in this thesis. Section 4.3 introduces the 'Forward Model', which has been developed for the retrieval model and illustrates some key aspects of the inverse problems to be solved. Finally, Sections 4.4 and 4.5 discuss how the GOME and ATSR-2 data can be combined to allow for the best representation of the data. The performance of the retrieval model, i.e. the impact of errors in the model parameters, is characterised in Chapter 5.

It is important to note the retrieval model developed in this thesis is not based on the aspects

of the optimal estimation theory that will be described in Section 4.1. Instead a special case, non-linear least squares fitting, is used that does not contain any priori constraint. However, it is still important to understand retrieval theory, firstly as it is used in the GRAPE retrieval scheme, and secondly as there are key diagnostics used in this retrieval which are taken from optimal estimation theory.

4.1 Retrieval Theory

It is possible, using the radiative transfer equation (Section 2.1.5), for the radiative field of the Earth's atmosphere to be calculated. However, in reality the parameters used in the calculation have errors associated with them, usually caused by errors in observations (instrumental errors). The relationship between the observed measurements and the true atmospheric state can be described by the forward function, $\mathbf{f}$ (Rodgers, 2000). The equation that relates the true and observed measurements is given by

$$\mathbf{y} = \mathbf{f}(\mathbf{x}, \mathbf{b}) + \epsilon, \tag{4.1}$$

where $\mathbf{x}$ is the state vector (containing the parameters that are to be retrieved), $\mathbf{y}$ is the measurement vector (containing the measured parameters), ϵ is the measurement error described by the covariance and $\mathbf{b}$ represents the parameters that affect $\mathbf{f}$ but are not retrieved. An inverse or retrieval model is used to find the expected value of the state $\hat{\mathbf{x}}$, which would result in the measurement $\mathbf{y}$, by adjusting $\mathbf{x}$ until observations are satisfactorily matched by the forward function. Therefore,

$$\hat{\mathbf{x}} = \mathbf{R}(\mathbf{y}, \hat{\mathbf{b}}, \mathbf{x_a}, \mathbf{c}), \tag{4.2}$$

where $\mathbf{R}$ is the retrieval model, $\mathbf{x_a}$ is the *a priori* of $\mathbf{x}$, i.e. some previous knowledge of the state, $\hat{\mathbf{b}}$ is the best estimate of the true values of $\mathbf{b}$ and c is a constant which includes the covariance of the *a priori*. In practice the forward function $\mathbf{f}$ can be too complicated to model so it is approximated with the Forward Model, $\mathbf{F}$.

4.1.1 Bayes' Theorem

A useful way of looking at an inverse problem is to use Bayes' theorem. This approach relates the *a priori* probability density function (PDF) of the state and the PDF of the measurement given the true state, to the PDF of the true state given the measurement.

1. $P(\mathbf{x})$ is the prior PDF of the state $\mathbf{x}$. This defines where $\mathbf{x}$ lies before the measurement is taken and is described by $\mathbf{S_a}$, the covariance of the *a priori*.

2. $P(\mathbf{y})$ is the prior PDF of the measurement. This is the PDF of the measurement before it is taken.

3. $P(\mathbf{y} \mid \mathbf{x})$ is the conditional PDF of $\mathbf{y}$ given $\mathbf{x}$, or where $\mathbf{y}$ lies given $\mathbf{x}$. This is described by $\mathbf{S_y}$, the covariance of the measurement.

4. $P(\mathbf{x} \mid \mathbf{y})$ is the conditional PDF of $\mathbf{x}$ given $\mathbf{y}$, or where $\mathbf{x}$ lies given $\mathbf{y}$. This is described by $\mathbf{S_x}$, the covariance of the state. This is what is required to solve an inverse problem i.e. the state vector, $\mathbf{x}$.

Bayes' theorem is described by the equation

$$P(\mathbf{x} \mid \mathbf{y}) = \frac{\mathbf{P(y \mid x)P(x)}}{\mathbf{P(y)}}. \tag{4.3}$$

Here, the right-hand-side is the PDF of the state given the measurement. This can then be used to find a more accurate version of the state vector. The PDFs are assumed to be Gaussian and hence for a single-element state vector the PDF can be described by

$$P(\mathbf{x}) = \frac{\mathbf{1}}{\mathbf{(2\pi)^{1/2}\Delta x}} \exp \left\{ \frac{\mathbf{x} - \hat{\mathbf{x}}}{\mathbf{\Delta \hat{x}}} \right\}^{\mathbf{2}}. \tag{4.4}$$

It can also be assumed that the experimental error has a Gaussian distribution. This would be given by

$$-lnP(\mathbf{y} \mid \mathbf{x}) = \frac{\mathbf{1}}{\mathbf{2}}(\mathbf{y} - \mathbf{F(x)})^{\mathbf{T}} \, \mathbf{S_y^{-1}} \, (\mathbf{y} - \mathbf{F(x)}) + \mathbf{c}, \tag{4.5}$$

where c is a constant and $\mathbf{F(x)}$ is the Forward Model and can be written as

$$\mathbf{y} = \mathbf{F(x)} + \epsilon \tag{4.6}$$

and $\mathbf{S_y}$ is the covariance matrix of the experimental error ϵ. If the *a priori* PDF is also assumed to be Gaussian then

$$-lnP(\mathbf{x}) = \frac{\mathbf{1}}{\mathbf{2}}(\mathbf{x} - \mathbf{x_a})^{\mathbf{T}} \, \mathbf{S_a^{-1}} \, (\mathbf{x} - \mathbf{x_a}) + \mathbf{c}, \tag{4.7}$$

where $\mathbf{S_a}$ is the covariance of $\mathbf{x_a}$, the *a priori* state. Therefore, using Bayes' theorem, equation 4.3, the distribution of the state vector can be given as

$$-2lnP(\mathbf{x} \mid \mathbf{y}) = (\mathbf{y} - \mathbf{F(x)})^{\mathbf{T}} \, \mathbf{S_y^{-1}} \, (\mathbf{y} - \mathbf{F(x)}) + (\mathbf{x} - \mathbf{x_a})^{\mathbf{T}} \, \mathbf{S_a^{-1}} \, (\mathbf{x} - \mathbf{x_a}) + \mathbf{c}. \tag{4.8}$$

If the problem is considered to be linear then the Forward Model can be represented by the weighting function matrix $\mathbf{K}$, this relationship is given by

$$\mathbf{F}(\mathbf{x}) = \mathbf{Kx}. \tag{4.9}$$

Equation 4.8 therefore reduces to

$$-2lnP(\mathbf{x} \mid \mathbf{y}) = (\mathbf{y} - \mathbf{Kx})^{\mathbf{T}} \mathbf{S_y^{-1}} (\mathbf{y} - \mathbf{Kx}) + (\mathbf{x} - \mathbf{x_a})^{\mathbf{T}} \mathbf{S_a^{-1}} (\mathbf{x} - \mathbf{x_a}) + \mathbf{c}. \tag{4.10}$$

The general form of a PDF is a quadratic function of $\mathbf{x}$ and is given by

$$-2lnP(\mathbf{x}) = (\mathbf{x} - \hat{\mathbf{x}})^{\mathbf{T}} \mathbf{S_x^{-1}} (\mathbf{x} - \hat{\mathbf{x}}) + \mathbf{c}. \tag{4.11}$$

Using Bayes' equation we can therefore write

$$(\mathbf{x} - \hat{\mathbf{x}})^{\mathbf{T}}\mathbf{S_x^{-1}}(\mathbf{x} - \hat{\mathbf{x}}) = (\mathbf{y} - \mathbf{Kx})^{\mathbf{T}} \mathbf{S_y^{-1}} (\mathbf{y} - \mathbf{Kx}) + (\mathbf{x} - \mathbf{x_a})^{\mathbf{T}} \mathbf{S_a^{-1}} (\mathbf{x} - \mathbf{x_a}) . \tag{4.12}$$

Equating the linear and quadratic terms of $\mathbf{x}$ in Equation 4.12 produces the relationships

$$\mathbf{S_x} = (\mathbf{S_a^{-1}} + \mathbf{K^TS_y^{-1}K})^{-1} \tag{4.13}$$

and

$$\hat{\mathbf{x}} = \mathbf{x_a} + (\mathbf{S_a^{-1}} + \mathbf{K_TS_y^{-1}K})^{-1} \left(\mathbf{K^TS_y^{-1}}(\mathbf{y} - \mathbf{Kx_a})\right) . \tag{4.14}$$

$\hat{\mathbf{x}}$ is the expected value of $\mathbf{x}$ and is equivalent to the maximum likelihood estimate of the true state. The optimal estimation technique seeks to determine this version of the state, from all the possible states consistent with $\mathbf{S_x}$. The maximum likelihood solution for Bayes theorem minimises the cost function

$$\chi^2 = (\mathbf{y} - \mathbf{F}(\hat{\mathbf{x}}))^{\mathbf{T}}\mathbf{S_y^{-1}}(\mathbf{y} - \mathbf{F}(\hat{\mathbf{x}})) + (\mathbf{x} - \mathbf{x_a})^{\mathbf{T}}\mathbf{S_a^{-1}}(\mathbf{x} - \mathbf{x_a}) \tag{4.15}$$

4.1.2 A Non-Linear Problem

Equation 4.14 shows how to solve a linear problem. A non-linear problem, such as the one in this thesis, can be solved by linearising about a given state and iterating. The problem in this thesis is known as a moderately non-linear problem (Rodgers, 2000). In order to linearise the problem, Equation 4.6 can therefore be written as

$$y = F(x_0) + K(x - x_0) + \epsilon \qquad (4.16)$$

Assuming that the model is linear for $\mathbf{x}$, Equation 4.14 becomes

$$\hat{\mathbf{x}} = \mathbf{x_0} + (\mathbf{S_a^{-1}} + \mathbf{K^T S_y^{-1} K})^{-1} \left(\mathbf{K^T S_y^{-1}} (\mathbf{y} - \mathbf{F(x_0)}) - \mathbf{S_a^{-1}} (\mathbf{x_0} - \mathbf{x_a}) \right). \qquad (4.17)$$

An iteration procedure is now required to find the solution. Applying a simple fixed point iteration scheme to Equation 4.17 gives

$$\mathbf{x_{i+1}} = \mathbf{x_i} + (\mathbf{S_a^{-1}} + \mathbf{K^T S_y^{-1} K})^{-1} \left(\mathbf{K^T S_y^{-1}} (\mathbf{y} - \mathbf{F(x_i)}) - \mathbf{S_a^{-1}} (\mathbf{x_i} - \mathbf{x_a}) \right), \qquad (4.18)$$

where $\hat{\mathbf{x}}$ has become $\mathbf{x_{i+1}}$, an approximation for the retrieved state at the next step and $\mathbf{x_0}$ is the initial estimate of $\mathbf{x}$. If a problem is not highly non-linear, this Newtonian iteration scheme should minimise the Bayesian cost function.

4.1.3 Convergence

To perform the iteration an initial estimate has to be made, $\mathbf{x_0}$, and the iteration continues until a final solution is produced, $\hat{\mathbf{x}}$. It is necessary when iterating for convergence criteria to be set, so that it is known when convergence is reached.

1. The reduction of the cost function is smaller than a chosen value.

2. The gradient of the cost function is close to 0.

3. The size of the change in the state is small.

The test used to check for convergence in thesis work and in GRAPE, is to calculate the absolute difference between successive cost function evaluations. Convergence is assumed when this difference decreases below a predetermined tolerance. In this case the value of the tolerance will be 1. The convergence test therefore takes the form

$$| \chi_{i+1}^2 - \chi_i^2 | < 1. \qquad (4.19)$$

Once Equation 4.19 has been satisfied, convergence is assumed and the iteration procedure stops.

4.1.4 Levenberg-Marquardt Iteration

In non-linear cases the Newtonian iteration may never converge as the steps can be too large. However, to avoid this the Levenberg-Marquardt iteration can be used. This scheme reduces the magnitude of increments of the estimate state, i.e. slows down the retrieval, so that the steps are small enough that within the range of the retrieval it is linear - hence the solution is slowly but surely approached (and the method allows steps to get bigger, approaching the speed of the Newtonian iteration as the retrieval approaches the solution). The Levenberg-Marquardt iteration achieves this by introducing a damping factor γ, into Equation 4.18. This produces

$$\mathbf{x_{i+1}} = \mathbf{x_i} + (\mathbf{S_a^{-1}} + \mathbf{K^T S_y^{-1} K} + \gamma_i \mathbf{I})^{-1}[\mathbf{K^T S_y^{-1}}(\mathbf{y} - \mathbf{F(x_i)}) - \mathbf{S_a^{-1}(x_i - x_a)}], \qquad (4.20)$$

where $\mathbf{I}$ is the identity matrix and γ_i is a damping factor that is adjusted after every step of the iteration. Using a starting value of γ_0 the residual sum of the squares is calculated, this is repeated with a value of γ_0/η, where η is a factor that is used to decrease the damping component and is taken as $\eta > 1$. If the residual sum of squares is decreased, i.e. the solution is improved, the damping factor is decreased by a factor of η. If the solution is degraded then the damping factor is increased by a factor of η. This continues until the convergence criteria given in Equation 4.19 is satisfied.

4.1.5 Linear Retrieval Diagnostics

A number of useful retrieval diagnostics can be derived from the optimal estimation method which characterise the quality of the solution, under the assumption that the retrieval is linear about the solution. The square–roots of the diagonal elements of $\mathbf{S_x}$ give the expected a posteriori precision of the estimated state. This vector is referred to here as the a posteriori error.

Differentiation of Equation 4.15 (also equating the estimated solution with the expectation of the PDF, $\hat{x} = <x>$) gives the sensitivity of the retrieval to perturbations in the measurement. This is referred to as the Gain Matrix:

$$\mathbf{G} = (\mathbf{K^T S_\epsilon^{-1} K} + \mathbf{S_a^{-1}})^{-1} \mathbf{K^T S_\epsilon^{-1}} \qquad (4.21)$$

As $\mathbf{K}$ is the weighting function matrix and describes the sensitivity of the measurement to the state, the sensitivity of the retrieval to perturbations in the truth (with elements $\partial\hat{\mathbf{x}}_i/\partial\mathbf{x_j}$) can be written:

$$\mathbf{A} = \mathbf{G_y K} \qquad (4.22)$$

This is often referred to as the averaging kernel matrix.

4.1.6 Non-linear Least Squares Fitting

As no *a priori* information is used in this retrieval scheme Equation 4.15 can be simplified to a non-linear least squares fit by removing the *a priori*:

$$\mathbf{x_{i+1}} = \mathbf{x_i} + \left(\mathbf{K^T S_y^{-1} K}\right)^{-1} \mathbf{K^T S_y^{-1}}(\mathbf{y} - \mathbf{F(x_i)}) \tag{4.23}$$

with the simplified retrieval diagnostics:

$$\mathbf{x_{i+1}} = \mathbf{x_i} + \left(\mathbf{K^T S_y^{-1} K} + \gamma_i \mathbf{I}\right)^{-1} \mathbf{K^T S_y^{-1}}(\mathbf{y} - \mathbf{F(x_i)}), \tag{4.24}$$

$$\mathbf{G} = (\mathbf{K^T S_\epsilon^{-1} K})^{-1} \mathbf{K^T S_\epsilon^{-1}} \tag{4.25}$$

$$\mathbf{S_x} = (\mathbf{K^T S_y^{-1} K})^{-1}. \tag{4.26}$$

Equation 4.26 is the simplified version of *a posteriori* error (Equation 4.13). Hereafter, the square root of the diagonals of S_x will be referred to as the Estimated Standard Deviation (ESD).

4.2 Retrieval Model Definition

The previous section described the mathematics and theory behind retrievals. This section will describe the retrieval method and model developed within this thesis.

The aim of this retrieval model is to reproduce the GOME measurements of the slant column density and, in doing so, infer the tropospheric vertical column density. This problem is formulated as a retrieval, in which the measurement is a GOME measured slant column and one sun–normalised radiance in a specific fitting window (see Chapter 1), and the state vector is the tropospheric scaling factor and the surface LER. A Forward Model is required to predict the slant columns based on an estimated state.

The tropospheric scaling factor (x_{sf}) is the amount by which the troposphere is adjusted by the retrieval until the measured and observed slant columns match. This is illustrated in Figure 4.1. The figure shows an exaggerated illustration of how the concentration of species varies with altitude. The solid black line represents a profile that is assumed to be representative of the state of the atmosphere and the red dashed line represents the new profile that fits with the observations after the retrieval has been carried out. The tropospheric scaling factor, x, has been calculated using the assumption that the stratospheric slant column is accurate (within a diagnosed error) and any discrepancy in the slant column is therefore contained in the troposphere. It is also assumed that the profile shape in the troposphere is correct (implications of this assumption will be addressed in the following chapters).

Figure 4.2 is a flow chart that summarises the different stages of the Retrieval Model.
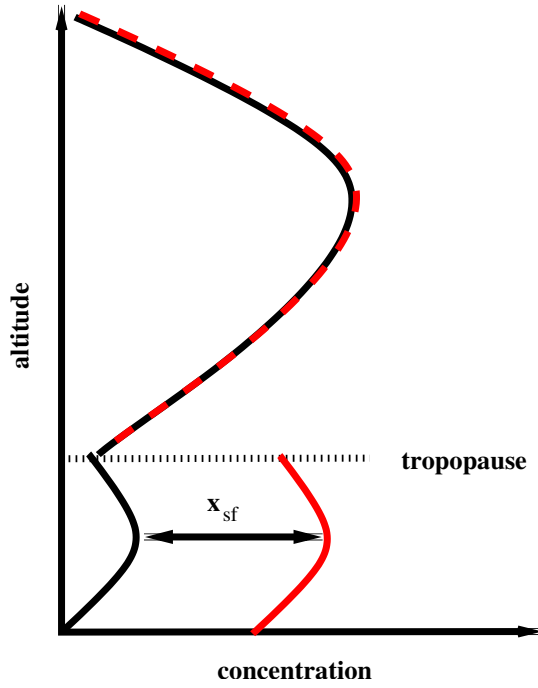
Figure 4.1: *Illustration of the tropospheric scaling factor (x_{sf}), where the x–axis is the concentration of an atmospheric species and the y–axis is altitude in the atmosphere. The black line shows an example of an atmospheric species concentration with height in the atmosphere, and the red line shows the new atmospheric concentration of this species after the tropospheric scaling factor has been applied.*

1. Take the GRAPE data containing cloud and aerosol parameters i.e. optical depth and particle size, for a specific orbit.

2. Match a GRAPE orbit to the appropriate GOME orbit and select only co located pixels.

3. Then pixel, by pixel,

   (a) Stratospheric and tropospheric profiles of pressure, temperature and relevant chemical species number density are obtained from SLIMCAT and TOMCAT at the correct location (described in Section 4.2.3).

   (b) GRAPE cloud and aerosol data are interpreted into extinction profiles (see Section 4.2.3).

   (c) Profiles are split into atmospheric layers.

   (d) An initial state vector is estimated.

   (e) GOME scene is divided into zones (see Section 4.5).

   (f) Forward Model is run for each zone.

   (g) Radiative transfer model is run, resulting in the sun–normalised radiance and the photon path lengths (see Section 4.2.4).

   (h) Photon path lengths are used to calculate the slant column density, resulting in a modelled measurement vector containing the slant column density and sun–normalised radiance.

93

(i) Rerun Forward Model for next zone.

(j) Combine measurement vector for each zone to give measurements appropriate to whole pixel (see Section 4.4).

(k) Fit the modelled measurement vector to the observed measurement vector using non-linear least squares fitting, and calculate new state vector.

(l) Iterate until the best possible estimate of the state is found.

4. Final state vector is used to calculate a tropospheric vertical column density of the relevant species.
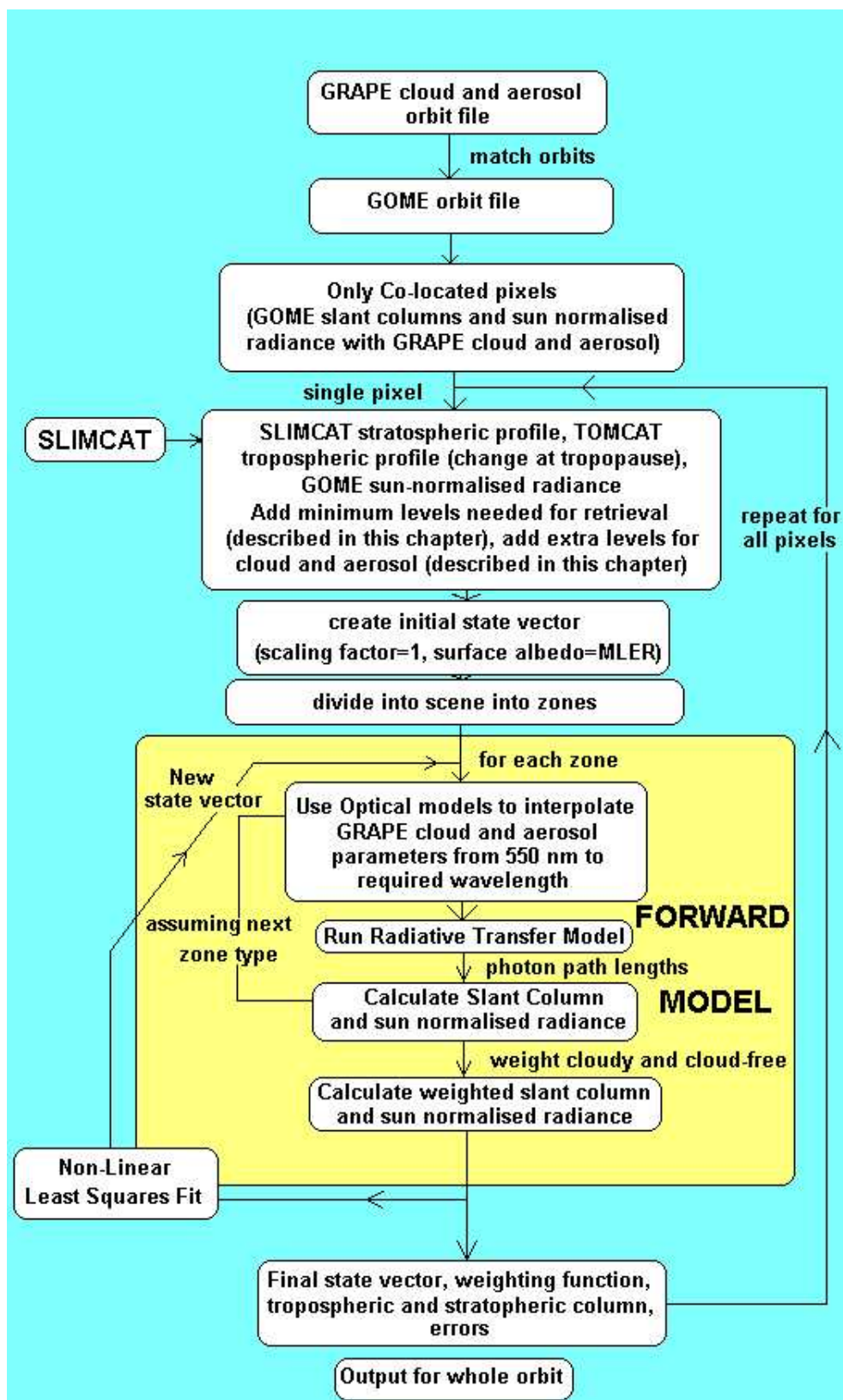
Figure 4.2: *A flow chart summarising the different stages incorporated into the Retrieval Model.*

4.2.1 Measurement Vector

The measurement vector contains the observations that are to be inversely modelled. It contains the slant column density and one sun–normalised radiance from the GOME instrument, these are taken from the GOME Data Product (GDP) for NO_2 and from the Harvard algorithm for BrO. The measurement vector error is described by the covariance matrix $\mathbf{S_y}$. The errors in the slant column density arise from random errors due to noise and systematic errors due to fitting errors (including temperature errors and cross section errors). The errors in the sun–normalised radiance come from instrumental errors (see Chapter 1).

GOME slant columns

The GOME slant columns are a level 2 product. A level 2 product is calculated using the level 1 product. The calculation of slant column density is achieved through the first stage of DOAS and was briefly described in Chapter 1. The slant column errors associated with the DOAS fitting are an important part of this new retrieval scheme and were reported in Chapter 1.

GOME sun–normalised radiance

The GOME sun–normalised radiance is a level 1 product, i.e. it can be directly determined from the GOME measurements. It is the ratio of the Earthshine radiance to the solar radiance. At a wavelength appropriate for representing scattering within the slant column fit window, a single sun–normalised radiance can be used to retrieve surface reflectance. The errors in sun-normalised radiance also come from the level 1 product.

4.2.2 State Vector

The state vector contains the initial parameters that the retrieval model calculates. For this work these parameters are a tropospheric scaling factor, defined earlier, and the surface LER. The initial guess for the state vector, x_0, is such that the scaling factor is assumed to be 1 and the surface LER is taken as the retrieved Minimum Lambertian Effective Reflectance (MLER), from GOME (Koelemeijer et al., 2003b), for the location and time of the measurement.

4.2.3 Model Parameters

A unique aspect within this new retrieval scheme is that there are exclusive Forward Model input parameters for every GOME measurement. Where possible no climatologies or look–up tables are used although some aerosol and cloud radiative properties and the initial guess at the surface LER do come from those sources. Not relying as strongly on look–up tables or climatologies should result in a more accurate scheme because errors are specific to the input

parameters and not to a mean error, which could be unrepresentative. The following section summarises the Forward Model input parameters.

Atmospheric Profiles

As well as the target species (NO_2 and BrO) it is also important to include profiles of O_3 as it will affect the path of radiation through the atmosphere (as it is not optically thin). The other species in the atmosphere can be thought of as optically thin. These species are present in the lower two layers of the atmosphere, the troposphere and stratosphere. The atmospheric profiles used in this work are calculated using the appropriate output generated by validated chemical transport models (CTMs).

Stratospheric Profiles

Stratospheric profiles are calculated using the SLIMCAT, 'offline', 3–D CTM forced by ECMWF ERA–40 re-analysis, as described in Chapter 3. The model was run for the GOME period, 1995 onwards. SLIMCAT includes chemical data assimilation of long–lived tracers. SLIMCAT output fields were stored every day at the local time of the GOME overpass. Figure 4.3 shows the local time of the GOME overpass with latitude.

It is important to sample the model at the correct local time due to the diurnal cycles of some of the species of interest. Figure 4.4 shows the NO_2 total vertical column if the model is sampled at the same time everywhere 12UT (the dashed line). This is compared with the total vertical column if NO_2 is sampled at the correct local time (the solid line). As NO_2 has a diurnal cycle, Figure 4.4 shows how NO_2 can vary if the model is sampled at the wrong time of day. The figure shows that there are large differences when the model is sampled at different times of the day for varying latitudes.

The SLIMCAT output fields have been extensively evaluated (see Chapter 3) meaning that the modelled stratosphere can be assumed to be accurate with a known quantitative estimate of the random error.

Tropospheric Profiles

Tropospheric monthly mean profiles for NO_2 were calculated using the 'off-line' TOMCAT 3–D CTM. The model was run on a $5.6° \times 5.6°$ grid with 31 hybrid σ–ρ levels up to 30 km in altitude. The TOMCAT model was run for 20 months with a spin-up time of 6 months from the 20th April 2000. Profiles are calculated for 2001 (Arnold et al., 2005). The sub-grid scale transport is performed using the Tiedtke convection scheme (Tiedtke, 1989; Stockwell and Chipperfield, 1999) and the Holtslag and Boville (1993) parameterisation for the turbulent mixing in the boundary layer follows the method suggested by Wang et al. (1999).

The tropospheric chemistry run of TOMCAT used to obtain the tropospheric profiles in this
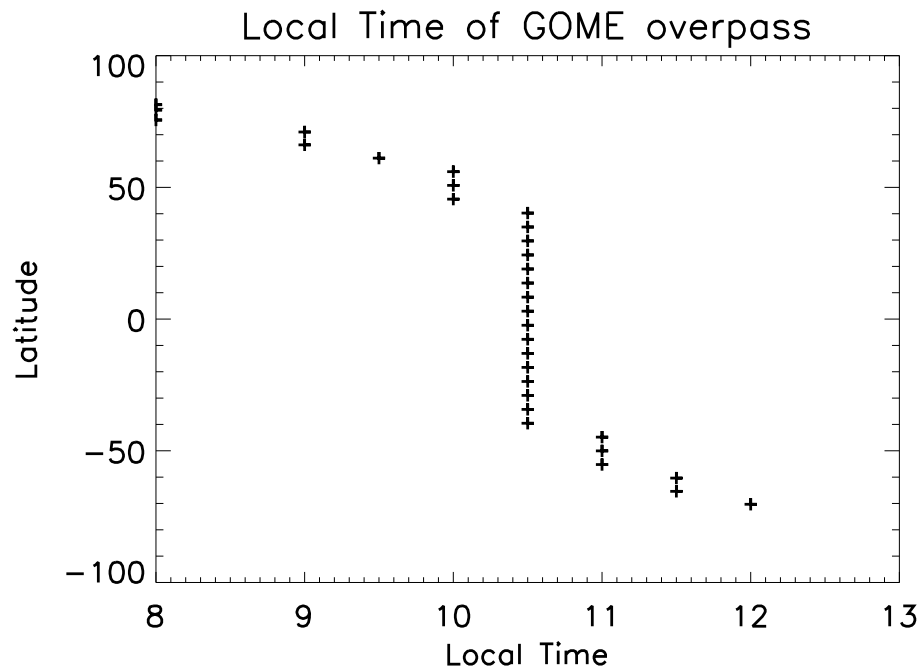
Figure 4.3: *The local time of an overpass as a function of the latitude for the ERS-2 satellite. The cross symbols show the model sampling times on a 7.5° grid, which correspond to 30 minute steps.*
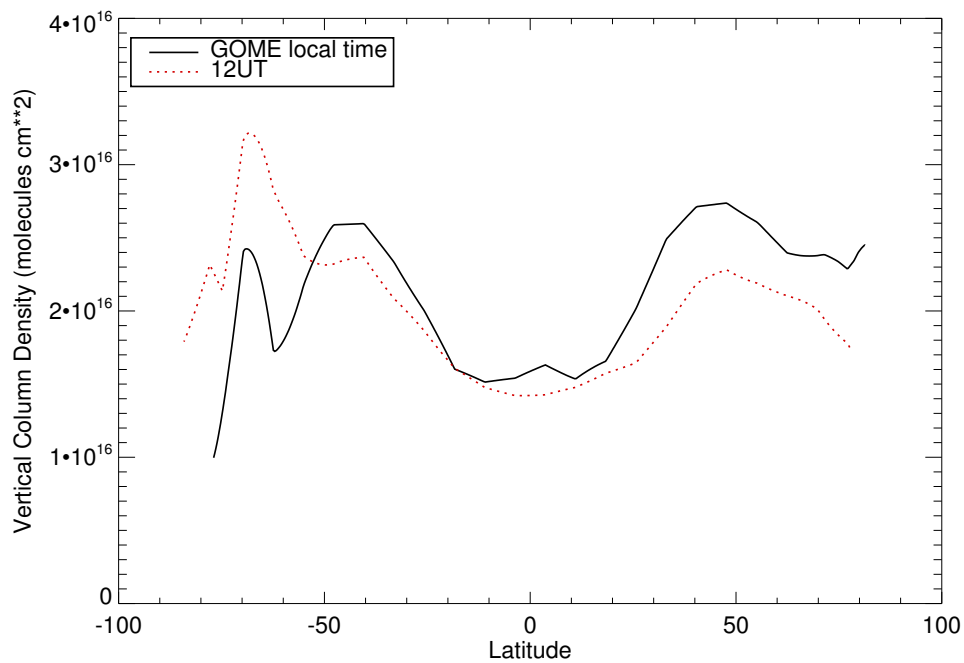
Figure 4.4: *The vertical column densities of NO_2 (molecules cm^{-2}) as a function of latitude. This is done when SLIMCAT has been sampled at the correct local time of the GOME overpass (black solid line) and when SLIMCAT has been sampled at 12 UT (red dotted line). This is for 8th October 1996.*

Table 4.1: *A table summarising the cloud and aerosol radiative parameters. Where the parameter comes from and a description of the parameter.*

Parameter	GRAPE/Assumed	Description
Cloud OPD	GRAPE	Cloud optical depth is retrieved in GRAPE
Cloud Thickness	Assumed	GRAPE retrieval assuming infinitely thin cloud
Cloud Particle Size	GRAPE	Cloud particle size retrieved in GRAPE
Cloud Top Height	GRAPE	Cloud top height determined using cloud top temperature/ECMWF
Cloud Phase	GRAPE	Cloud phase (water or ice) is retrieved in GRAPE
Cloud Extinction Profile	Assumed	Tested in Section 4.4
Aerosol Type	Assumed	Aerosol type assumed in OPAC
Aerosol Particle Size	GRAPE	Aerosol particle size retrieved in GRAPE
Aerosol OPD	GRAPE	Aerosol optical depth retrieved in GRAPE
Aerosol Extinction Profile	Assumed	Taken from OPAC

work, used 23 advected tracers. These included important tropospheric species such as O_3, NO_x, CO, HO_x, CH_4 and VOC's. The model did not include isoprene chemistry, but accounts for this by artificially enhancing the CO emissions. The surface emissions are based upon the IPCC (Intergovernmental Panel for Climate Change) scenarios. This model also includes lightning parameterisations and emissions from aircraft.

Cloud and Aerosol

The GRAPE dataset is a state-of-the-art retrieval of ATSR–2 cloud and aerosol parameters and was introduced in Chapter 2. GRAPE data has been averaged to fit the GOME pixel (see Section 4.5) resolution producing an accurate measurement of cloud and aerosol that can be used in the retrieval process. To run the Forward Model these cloud and aerosol parameters need to be interpreted into extinction coefficient profiles (where the extinction coefficient is equal to the sum of the absorption and scattering coefficients). For clouds, the cloud top height is taken from GRAPE, then the cloud is assumed to be 1 km thick. The optical depth is used to calculate the extinction of the cloud. The particle size and phase (liquid or ice) are also taken from GRAPE. For aerosol, the aerosol type (which is assumed from OPAC Hess et al., 1998, due to geo–location) is used to obtain an extinction profile shape. The optical depth (taken from GRAPE) is used to scale the profile to the correct value.

Table 4.1 summarises the cloud and aerosol parameters used in the retrieval model, whether they are taken from GRAPE or are assumed, and briefly describes each parameter. For more detail on a description of these properties see Chapter 1.

4.2.4 Forward Model

The Forward Model, as described earlier, is an approximation of the forward function which uses a state vector to describe the measurement vector. The Forward Model contains two steps, the radiative transfer model and the slant column calculation.

Radiative Transfer Model

The Radiative Transfer Model (RTM), used in this work is an optimised version of GOMETRAN (described in Chapter 2). The RTM outputs the top of the atmosphere (TOA) sun–normalised radiance i.e. that seen by the instrument, and the weighting function for the change in radiance due to the small change in the concentration of a species per layer $(c_j dr/dc_j)$, which is also multiplied by the concentration in the layer. For the slant column density to be derived, photon path lengths need to be calculated. The following describes how photon path length can be determined from this weighting function.

From Equation 1.19 and the definition of sun-normalised radiance,

$$\frac{dR(\lambda)}{dc_j} = -r(\lambda)\frac{d\tau(\lambda)}{dc_j} \tag{4.27}$$

where $r(\lambda)$ is the sun–normalised radiance, c_j is the species concentration and τ is the optical depth along the light path. The weighting function output from GOMETRAN ($\mathbf{K_G}$ is the derivative of R with respect to a 100% change in concentration at a particular level, and assuming that σ is constant in s, it can be seen that

$$\mathbf{K}_G = -r(\lambda)c_j\frac{d\tau(\lambda)}{dc_j} \tag{4.28}$$

where j is a model layer. If

$$\tau = \sigma \int_o^z c(z)l(z)dz \tag{4.29}$$

where l is photon path length, and the equation has been discretised into layers

$$\tau = \sigma \sum_{j=1}^{N_l} (c_j l_j \delta z_j) \tag{4.30}$$

therefore,

$$\frac{d\tau}{dc_j} = \sigma l_j \delta z_j = -\frac{K_G}{r(\lambda)c_j} \tag{4.31}$$

So

$$l_j = -\frac{K_G}{r(\lambda)c_j\sigma\delta z_j} \tag{4.32}$$

and $c_j\sigma = k_a$ (i.e. the absorption coefficient).

$$l_j = \frac{\mathbf{K}_G}{r(\lambda)k_a\delta z_j} \tag{4.33}$$

This is how photon path length is derived from the GOMETRAN weighting function.

Slant Column Calculation

Equation 1.27 shows the calculation of the slant column density (S).

$$S = \int_0^{z_{TOA}} c(z)l(z)dz \tag{4.34}$$

However, in this retrieval scheme the tropospheric part of the slant column density is calculated with the scaling factor x_{sf}:

$$S = x_{sf}\int_0^{z_{trop}} c(z)l(z)dz + \int_{z_{trop}}^{z_{TOA}} c(z)l(z)dz \tag{4.35}$$

where z_{trop} is the level at the top of the tropopause and z_{TOA} is the level at the top of the atmosphere.

4.2.5   Vertical Column Density

The vertical column density is calculated from the scaling factor and the target species profile that was used to initialise the retrieval model.

$$V_{trop} = x_{sf}\int_0^{z_{trop}} c(z)\delta z \tag{4.36}$$

4.3   Photon Propagation

This section describes a number of tests that have been performed on the Forward Model to determine its response to different scattering problems and to illustrate how photon propagation is affected by different atmospheric conditions. The testing is achieved by plotting photon path lengths from the Forward Model averaged over a number of equally spaced layers. Photon path
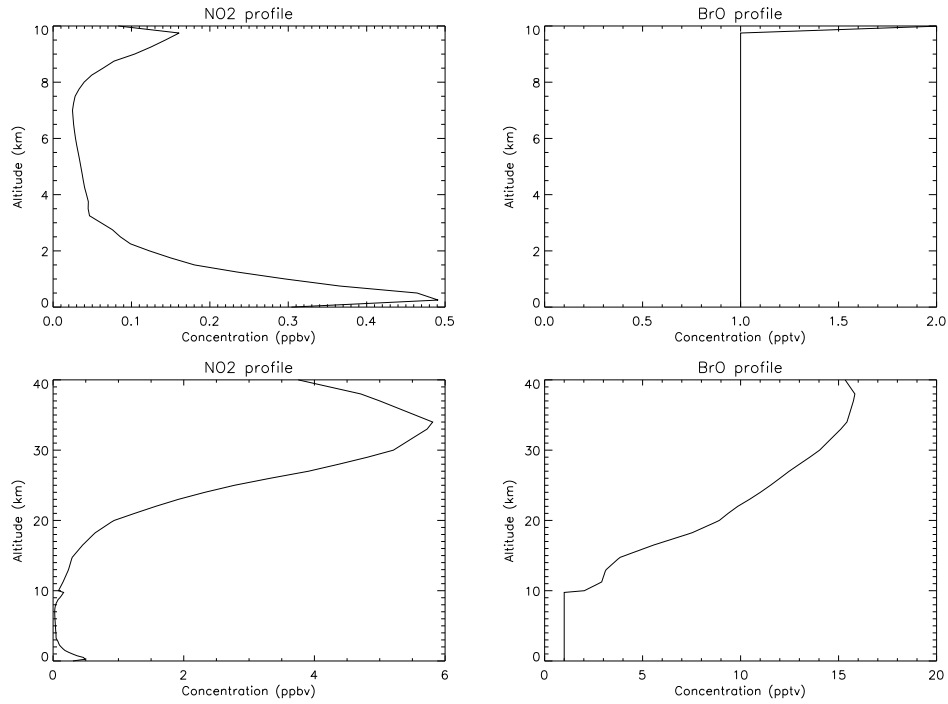
Figure 4.5: *NO$_2$ profile in the bottom 10 km of the atmosphere, used in the tests described in Section 4.3.*

length is described in Section 4.2.4, where the photon path length is the mean path of a photon in a layer that is observed by the sensor (i.e. GOME). Therefore, if the photon path length is long then the observed photon has experienced multiple scattering within a specific layer. Also the photon path length is related to the slant column through Equation 1.27, these plots can be used to qualitatively discuss the implications of errors in different atmospheric parameters.

To aid the testing process, a scenario has been setup for July at 55° N, 0° E which gives a solar zenith angle of 39° (described fully in Section 5.2). SLIMCAT/TOMCAT chemical profiles for the location and date of the scenario have been used. The TOMCAT data are monthly means, however the SLIMCAT stratospheric profiles are taken for the middle of the month (15th). The tests were run at 430 nm and 350 nm to simulate a retrieval in the NO$_2$ and BrO fitting windows respectively.

The profiles of NO$_2$ and BrO used in these photon path length tests are shown in Figure 4.5. The figure shows the profile for just the troposphere and then from the surface to 40 km so that the profile in the stratosphere can be examined also. NO$_2$ has a large peak in the stratosphere and a slightly lower peak in the boundary layer (bottom 2 km), whereas BrO also has a peak in the stratosphere, but as little is known about the profile of BrO in the troposphere, a constant concentration is used.

In order to also gain some quantitative information about the photon path lengths the following plots have information on the sensitivity of each retrieval to (1) a small change in species concentration in the stratosphere, (2) a small change in species concentration in the troposphere

102

and (3) a small change in species concentration in the boundary layer. These are described by what will be referred to as the profile averaging kernels.

4.3.1 Profile Averaging Kernel

In this retrieval scheme the profile averaging kernel, $\mathbf{A_{p_j}}$, is of interest (Eskes and Boersma, 2003). $\mathbf{A_{p_j}}$ is the sensitivity of the retrieved tropospheric vertical column (V_t) to a fractional change in the concentration of the target species within a particular layer (c_j). To determine the value of A_{p_j}, Equation 4.22 can be used to obtain dx_{xf}/dS by taking the appropriate elements of the retrieval gain matrix.

$$\mathbf{A}_{p_j} = \frac{dx_{sf}}{dS} \frac{dS}{dc_j} \tag{4.37}$$

where dx_{sf}/dS is the first elements of the gain matrix and dS/dc_j can be taken from the radiative transfer model

$$\mathbf{A}_{p_j} = \frac{dx_{sf}}{dS} l_j \delta z_j, \tag{4.38}$$

where $\mathbf{A}_{p_j}$ is the profile averaging kernel for a specific model level.

4.3.2 Surface Lambertian Equivalent Reflectance

Figure 4.6 shows how different specified surface LER values affect the photon path lengths derived from the Forward Model for NO_2. The figure shows the average photon path length (km/km) every 0.5 km in the lower atmosphere (surface to 10 km) for a range of surface LER values. At low surface LER (0.01–0.03) there is more absorption and less reflection at the surface. This means the photon path length shortens near the surface, due to a decrease in the number of photons that are observed in this layer by the sensor. At high surface LER (0.4–0.6) there is less absorption and more reflection at the surface, therefore photon path lengths are longer at the surface. The green line in Figure 4.6 shows the geometric photon path length. The geometric photon path length represents the path that a photon would take if there was no atmosphere (no interaction). Photon path lengths will tend to the geometric photon path length at higher altitudes, as the atmosphere becomes thinner and so less scattering occurs. Figure 4.7 is the same as Figure 4.6 but shows the photon path lengths up to 40 km in the atmosphere, and shows how the photon path lengths converge to the geometric photon path length.

Figure 4.6 also highlights the relationship between surface LER values and the retrieved tropospheric VCD. This can be explained by Equation 1.27, where modelled slant column density (SCD) is calculated using the photon path length. A large error in the surface LER in the Forward Model would affect the fitted SCD and therefore the tropospheric VCD. This effect
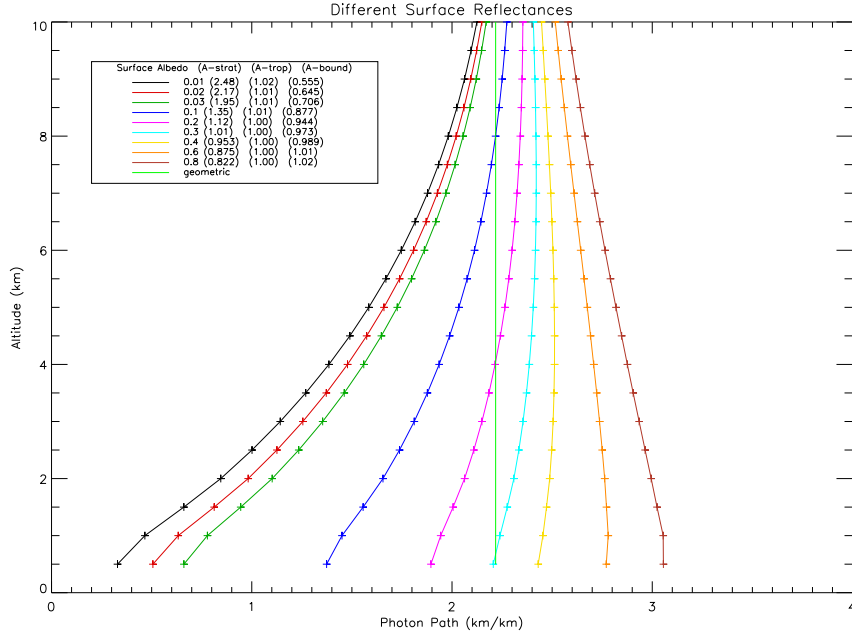
Figure 4.6: *The modelled photon path lengths (km/km) through the lower atmosphere at different surface LER for the Forward Model run in July at 55° N, 0° E, the wavelength is 430 nm. The legend identifies the different surface LER sampled and the stratospheric ($\mathbf{A_s}$), the tropospheric ($\mathbf{A_t}$) and the boundary layer ($\mathbf{A_b}$) averaging kernels for each surface LER.*
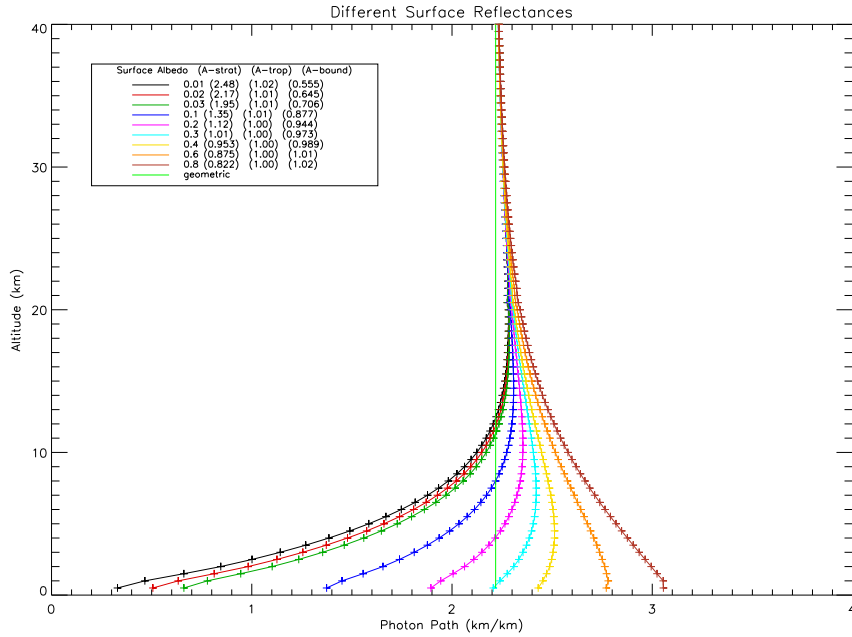
Figure 4.7: *As Figure 4.7 but from the surface to 40 km.*

104
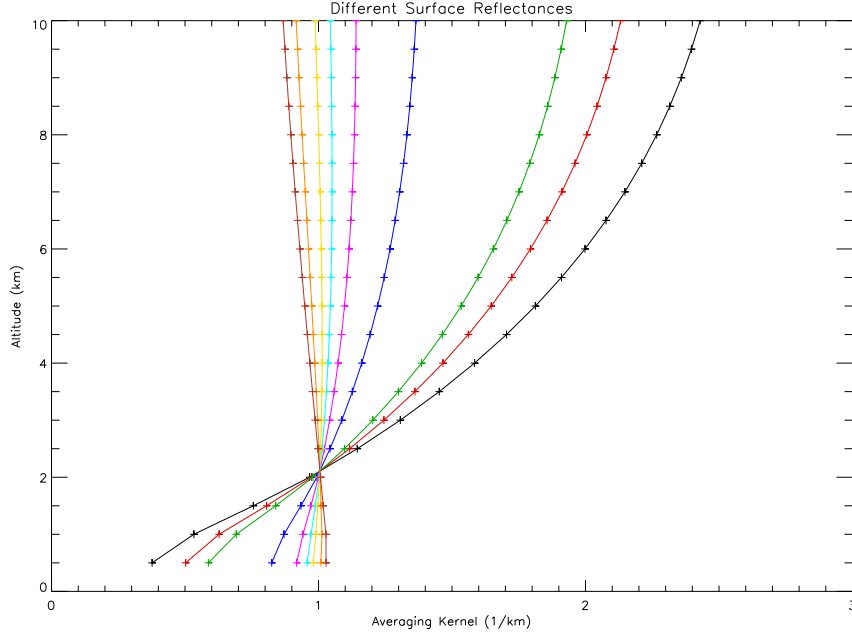
Figure 4.8: *The averaging kernels through the lower atmosphere (per km) at different surface LER for the Forward Model run in July at $55°N$, $0°E$, for the wavelength of 430 nm. The legend identifies the different surface LER sampled and the stratospheric ($\mathbf{A_s}$), the tropospheric ($\mathbf{A_t}$) and the boundary layer ($\mathbf{A_b}$) profile averaging kernel for each surface LER.*

will be discussed in Chapter 5.

Figure 4.8 shows the profile averaging kernels for every layer, $\mathbf{A_p}$, for different specified surface LER. As discussed earlier, the value of $\mathbf{A_p}$ describes the sensitivity of the retrieved tropospheric VCD to a change in concentration of the target species within a specific atmospheric layer. This figure shows that for low surface LER values $\mathbf{A_p}$ is small below 2 km and then increases to large values $\sim$2. This profile averaging kernel is shows where the retrieval scheme observes the target species in the atmosphere. From Figure 4.5 it can be seen that the NO_2 profile in the troposphere is mostly in the lowest 2 km. However, $\mathbf{A_p}$ tells us that the retrieval is weighting the retrieved tropospheric VCD to the upper troposphere, i.e. the retrieval is insensitive to the bottom 2 km, and overly sensitive from 2-10 km. For higher surface LER this is reversed, however to a much lesser degree, e.g. the values of $\mathbf{A_p}$ are closer to 1. Overall, these profile averaging kernels are not only useful for providing information on layer by layer sensitivity to the retrieved tropospheric VCD, but if there are averaged over a number of layers (i.e. the whole of the troposphere) they can be used to determine the sensitivity to errors in, for example, the troposphere. The values printed in the legends in Figure 4.6 show the profile averaging kernels for the stratosphere $\mathbf{A_s}$, the troposphere $\mathbf{A_t}$ and the boundary layer $\mathbf{A_b}$.

The stratospheric profile averaging kernel, $\mathbf{A_s}$, describes the sensitivity of the retrieved tropospheric VCD to a change in concentration of the target species (here NO_2) in the stratosphere, where a value of 1 for the profile averaging kernel is a perfect response (i.e. for an increase in

concentration of the target species in the stratosphere of say 1×10^{15} molecules cm^{-2}, there will be a 1×10^{15} molecules cm^{-2} increase in the concentration of retrieved tropospheric VCD of the target species). Figure 4.6 shows that $\mathbf{A_s}$ for NO$_2$ varies from 1.12–2.48 for surface LER less than 0.3, with the value of $\mathbf{A_s}$ decreasing with increasing surface LER. This means that a uniform increase in stratospheric NO$_2$ concentration of 1×10^{15} molecules cm^{-2} (i.e. per layer) will result in an increase in the retrieved tropospheric NO$_2$ VCD of 1.12–2.48$\times 10^{15}$ molecules cm^{-2}. For higher values of the surface LER (above 0.3) $\mathbf{A_s}$ varies from 0.82–0.95. This means that an increase in the concentration of NO$_2$ in the stratospheric of 1×10^{15} molecules cm^{-2} will result in an increase in the retrieved tropospheric NO$_2$ VCD of 0.82–0.95$\times 10^{15}$ molecules cm^{-2}. If a change in stratospheric NO$_2$ concentration is thought of as an error in the stratospheric total column (i.e. an error in the SLIMCAT calculated stratosphere) then it can be seen that an error in the stratospheric column will have a significant impact on the retrieved tropospheric NO$_2$ VCD. Therefore, $\mathbf{A_s}$ highlights the importance of calculating the stratospheric loading as accurately as possible, as large values of $\mathbf{A_s}$ mean there will be a large impact of the retrieved tropospheric VCD[1].

The tropospheric profile averaging kernel, $\mathbf{A_t}$, describes the sensitivity of the retrieved tropospheric NO$_2$ VCD to a change in concentration of the target species in the troposphere. The values of $\mathbf{A_t}$ are smaller than $\mathbf{A_s}$ for NO$_2$. For all surface LER, values of $\mathbf{A_t}$ for NO$_2$ vary from 1.00-1.02. This means a 1×10^{15} molecules cm^{-2} increase in tropospheric NO$_2$ concentration will result in a 1.–1.02$\times 10^{15}$ molecules cm^{-2} increase in the retrieved tropospheric VCD. This is an interesting result because although from Figure 4.5 it can be seen that the NO$_2$ mostly resides in the bottom 2 km of the atmosphere, increasing the concentration of NO$_2$ will not affect the retrieved tropospheric VCD significantly for changing surface LER.

The boundary layer profile averaging kernel, $\mathbf{A_b}$, describes the sensitivity of the retrieved tropospheric VCD to a change in concentration in the boundary layer. As NO$_2$ and BrO commonly reside in the boundary layer, it is important that the retrieval model is sensitive to this part of the atmosphere. For surface LER values between 0.01–0.3, $\mathbf{A_b}$ is low (0.55–0.97). This means an increase in NO$_2$ concentration of 1×10^{15} molecules cm^{-2} throughout the boundary layer will result in an increase of 0.55–0.97$\times 10^{15}$ molecules cm^{-2}. Above a surface LER of 0.3 the value of $\mathbf{A_b}$ is ∼1, therefore a perfect response.

One of the novel parts of this new retrieval scheme is that it uses a tropospheric scaling factor to determining the tropospheric VCD and aims to fit the modelled SCD to the observed SCD by scaling the troposphere by a uniform amount for the target trace gas. In doing so, the scheme will infer the amount of the trace gas in the troposphere. One problem with this approach however, is that for a number of tropospheric species, scaling the troposphere uniformly may not be very accurate. For example, Figure 4.5 shows that NO$_2$ mostly resides in the boundary layer, therefore it might be better to scale the boundary layer rather than the whole troposphere. As the values of $\mathbf{A_t}$ shown in Figure 4.6 are all close to 1 this shows that the tropospheric scaling factor is suitable in this case. It is important to note that the profile averaging kernels will

---

[1]It should be noted that these tests only deal with changes in retrieved tropospheric VCD, whereas the final retrieval model will also include the retrieval of surface LER which will help to reduce the overall error.

be different in every scenario and that values are dependent on a number of Forward Model parameters e.g. profile shape and solar zenith angle. The values of $\mathbf{A_b}$ show that for low surface LER (below 0.3) the retrieval is not sensitive to changes in the boundary layer and therefore scaling the troposphere instead will act to fit the tropospheric VCD better.

Overall, profile averaging kernels for the stratosphere, troposphere and boundary layer provide important information. The stratospheric profile averaging kernels show that the retrieval is very sensitive to changes in concentration in the stratosphere for NO_2 and therefore highlights the importance of calculating the stratospheric loading of NO_2 and associated errors. The tropospheric profile averaging kernel shows that the use of a tropospheric scaling factor is reasonable for NO_2 in cases where no cloud is present, as the retrieval has a near perfect response to a change in tropospheric concentration. The boundary layer profile averaging kernel highlights that at low surface LER the retrieval is not sensitive to the boundary layer. Therefore, it is important to have the best possible knowledge of the profile shape in the boundary layer to provide the most accurate retrieval.

As stated above, profile averaging kernels are specific to many Forward Model parameters. Figures 4.9 and 4.10 show the photon path lengths and profile averaging kernels for BrO. The first point to note is that from Figure 4.5 it can be seen that the profile of BrO in the troposphere is very different to that of NO_2, with concentrations of BrO in the parts per trillion (pptv) in contrast of NO_2 parts per billion (ppbv). Secondly studying Figure 4.9 the photon path lengths vary more for BrO than NO_2. Chapters 1 and 2 discussed Rayleigh scattering and its impact of retrievals in the UV. The variation in the photon path lengths here will partly be due to Rayleigh scattering having more of an effect at 350 nm (the wavelength of BrO retrievals) than at 430 nm (the wavelength of NO_2 retrievals).

The stratospheric profile averaging kernels for BrO, $\mathbf{A_s}$, show that for BrO the retrieved tropospheric VCD is less sensitive to changes in the concentration in the stratosphere than for NO_2. However, the values are still large, varying from 1.17-1.65 for surface LER values of 0.3–0.01 and from 0.8–0.9 for surface LER values of 0.8–0.4. The tropospheric profile averaging kernel for BrO, $\mathbf{A_t}$ is very different from $\mathbf{A_t}$, for NO_2, with values ranging from 0.6–1 for all surface LER values ($\mathbf{A_t}$ increasing with increasing surface LER). This means that for BrO retrievals the retrieved tropospheric VCD is not sensitive to changes in BrO concentration in the troposphere. This will be a large source of error in the overall retrieval of BrO. The BrO profile used to initiate the retrieval model is likely to be higher in error (due to lack of knowledge of BrO profiles in the troposphere) than NO_2. Therefore, the high values of $\mathbf{A_t}$ will result in larger errors compared to NO_2. For example at a surface LER of 0.01 if the tropospheric BrO profile needs to be corrected by 5×10^{13} molecules cm^{-2}, this would result in a tropospheric VCD change of only 3×10^{13} molecules cm^{-2}. This is even more evident in the values of $\mathbf{A_b}$ which are between 0.27–1.05 for all surface LERs (increasing $\mathbf{A_t}$ values with increasing surface LER). This means that the retrieval of tropospheric BrO is insensitive to the boundary layer, which is where most of the BrO is thought to reside.

Figure 4.10 shows the profile averaging kernels for the BrO retrieval at different surface LER values for every layer. The figure shows that a perfect response to a change in BrO concentration

107
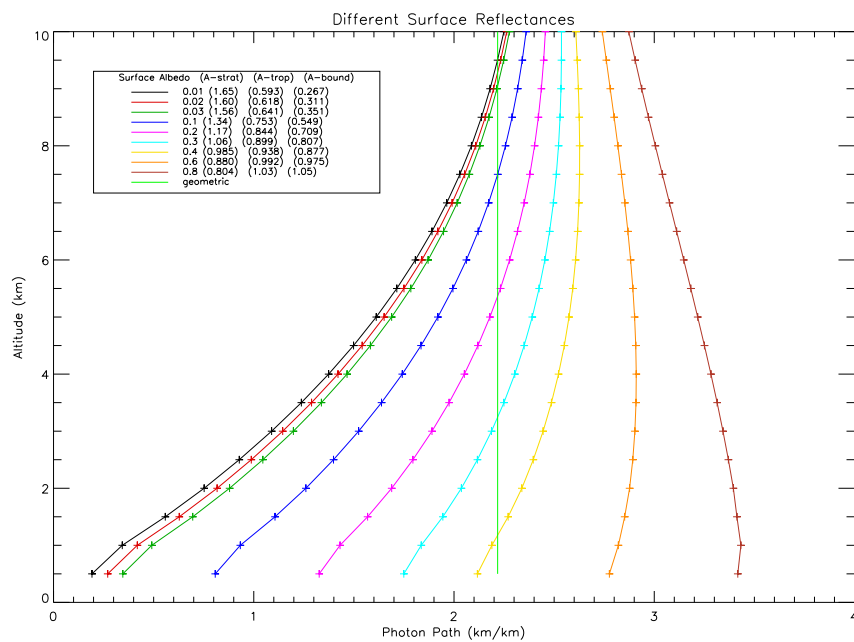
Figure 4.9: *As Figure 4.6 but for BrO.*

is observed at ∼4 km, whereas this was ∼2 km for NO₂. This helps in the understanding of why BrO retrievals are less sensitive to the boundary layer than NO₂ (i.e. the profile averaging kernels are much smaller at the surface for BrO than fro NO₂).
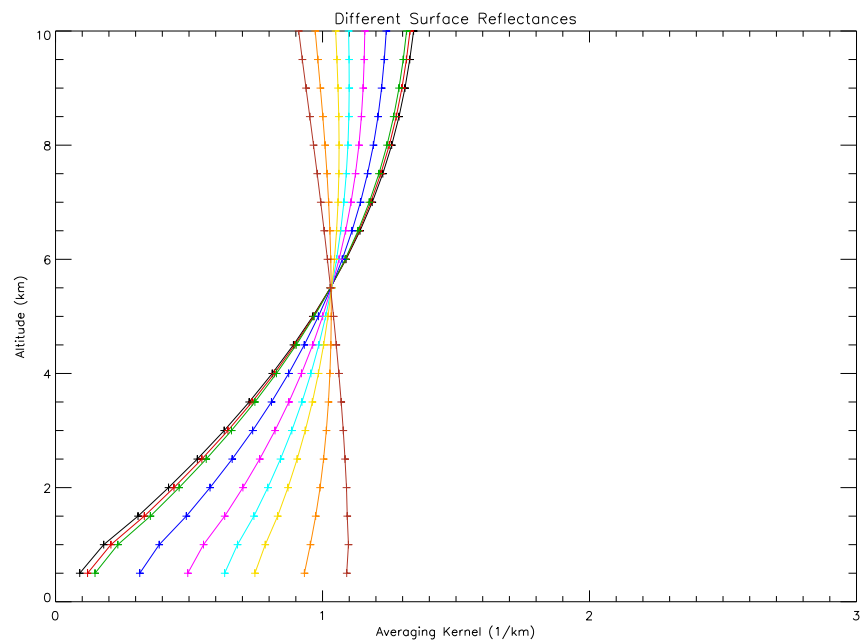
Figure 4.10: *As Figure 4.8 but for BrO.*

Overall, Figures 4.9 and 4.10 show that BrO retrievals have low sensitivity to changes in BrO concentration in the troposphere. However, they are less affected by errors in the stratosphere than NO_2 retrievals are.

4.3.3 Tropospheric Profile Shape

Figures 4.11 shows $\mathbf{A_p}$ when varying the tropospheric profile shape of NO_2 (430 nm). The figure shows that values of $\mathbf{A_p}$ are different depending on the profile shape of the species. This can be seen in Equation 4.37 where $\mathbf{A_p}$ is dependent on c_i, the concentration at each level (i.e. the profile shape). The profile dependence of $\mathbf{A_p}$ highlights that the value of $\mathbf{A_p}$ will be specific to each retrieval (as tropospheric profile shape varies pixel by pixel) and therefore the values of $\mathbf{A_p}$ discussed above are not relevant in all cases, only to the specific scenario being discussed. Figure 4.11 also demonstrates the importance of profile shape on the retrieval. The figure shows that, especially for low surface LER (top righthand plot (b)), profile shape error will cause large discrepancies in tropospheric VCD. This will be discussed quantitatively in Chapter 5.

4.3.4 Clouds

Clouds have a large impact on radiation in the atmosphere. Therefore, it is important to quantify the effect of clouds in the Forward Model. The Forward Model was run for three surface LERs, a low value, over oceans (0.02), a medium value, over land (0.1) and a high value, over snow/ice (0.8). The simulated cloud is a water cloud with particle size of 30 μm with a cloud top height of 5 km with a geometric depth of 2 km. All figures are for NO_2 with a profile shown in Figure 4.5. The impact on BrO will be described but will not be shown.

Cloud Optical Depth

The plots in Figure 4.12 show how the cloud optical depth affects the photon path lengths. Figure 4.12(a) shows the photon path lengths at a low surface LER of 0.02. The black line shows the photon path lengths when there is no cloud. Following this line down through the atmosphere the photon path lengths are shortened with decreasing altitude, starting at around 2 km/km and ending up at 0.5 km/km in the lowest layer (0.5 km). For clouds with an optical depth of below 0.1 the photon path lengths are similar to when there is no cloud present (i.e. the cloud has little affect on the photons). However, as the optical depth reaches 0.1 the photons become perturbed by the cloud (the photons are scattered by the cloud particles). At a cloud optical depth of 0.3 the first impact of the cloud is observed, where the photon path lengths above the cloud increase. This is due to an increase in the amount of multiple scattering above the cloud. As the cloud optical depth increases the photon path lengths just above the cloud increase, up to almost 4 km/km in the case of the cloud with an optical depth of 30. Inside and below the cloud the photon path lengths decrease with increasing cloud optical depth as there are less photons traveling through the cloud and back to the satellite. At a cloud optical depth
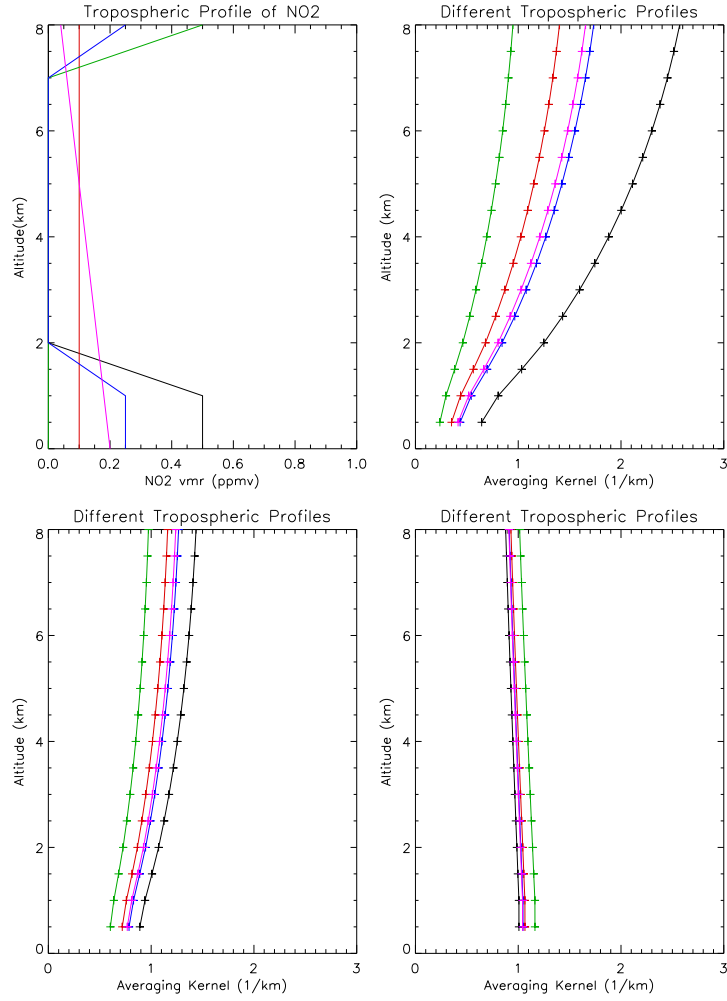
Figure 4.11: *Averaging kernels for changing the tropospheric profile shape for NO_2. Panel (a) volume mixing ratio (vmr) profile shape, (b) profile averaging kernels with altitude for surface LER 0.02, (c) for surface LER 0.1, and (d) for surface LER 0.8.*

of 30 it is hard for photons to penetrate the cloud and those that do are likely to be absorbed by the surface as the reflectance is so low.

The profile averaging kernels, $\mathbf{A_s}$, $\mathbf{A_t}$ and $\mathbf{A_b}$ are all highly variable for different cloud optical depths. The values of $\mathbf{A_s}$ are similar to those seen in Figure 4.6 for low surface LER with low optical depths. The values increase at the same point where the photon path lengths start to change (around a cloud optical depth of 0.1). The value of $\mathbf{A_s}$ is very high at 5.09 for a cloud optical depth of 30. As Figure 4.12(a) shows, the photon path lengths above the cloud increase, and this means they will increase in the stratosphere also.

The values of $\mathbf{A_t}$ for low surface LER show that for low cloud optical depth there is a perfect response in the retrieved tropospheric VCD to a change in concentration in the troposphere. However, above cloud optical depths of 0.3 the retrieved tropospheric VCD becomes insensitive to changes in the concentration in the troposphere. This value of $\mathbf{A_t}$ decreases to 0.06 for cloud optical depths of 30. A similar effect is observed in the value of $\mathbf{A_b}$ where the value decreases to 0.002 at an optical depth of 30. This means that the retrieved tropospheric VCD is highly insensitive to changes in concentration in the boundary layer. From Figure 4.5 it can be seen that the majority of NO_2 resides in the boundary layer, therefore these values of $\mathbf{A_t}$ and $\mathbf{A_b}$ are understandable, as from Figure 4.12(a) it can be seen that the photon path lengths become very short in the boundary layer, if a cloud with an optical depth of 30 is present. This is because the photons do not penetrate through the cloud. Therefore, it would be expected that the retrieval would not see the lower layers of the atmosphere; those beneath the cloud. In reality the instrument (GOME) will not have measured the surface either. In many retrieval studies, if an optically thick cloud is present a *ghost column* can be fitted (e.g. Martin et al., 2002). This relies on having good knowledge of the NO_2 (or other target species) under the cloud. This new retrieval scheme will calculate something similar to a ghost column, as the scaling factor will scale the original tropospheric profile to fit the observations. However, both these methods rely strongly on the original tropospheric profile of NO_2 (or other target species) and therefore bias the final retrieved tropospheric column accordingly. The implications of these large profile averaging kernels will be discussed further in Chapter 5.

Figure 4.12(b) shows the photon path lengths at a surface LER of 0.1. The black line shows the photon path lengths when there is no cloud. Following this line down through the atmosphere it can be seen that the photon path lengths decrease at a slower rate with decreasing altitude than in Figure 4.12(a). Their values decrease from around 2 km/km to 1.4 km/km. The relationship between cloud optical depth and photon path length at this surface LER is similar to that at a reflectance of 0.02 but with a shallower drop in photon path lengths. This is mirrored in the profile averaging kernels which show that if the surface LER is slightly higher, tropospheric VCD is not as strongly sensitivity to changes in concentration.

Similar results are mimicked in the values of $\mathbf{A_s}$, $\mathbf{A_t}$ and $\mathbf{A_b}$ to the photon path lengths. The values behave in the same way as at a surface LER of 0.02, but to a slightly lesser extent. However, at the higher cloud optical depths the values of the profile averaging kernels are similar at both surface LERs. The stratospheric profile averaging kernel values are still very high, ranging from 1.36–5.07, which once again suggests that the stratospheric NO_2 column needs to be as accurate as possible with all errors quantified. The values of the tropospheric profile averaging kernels range from 1.01–0.07. This is also mimicked in the values of $\mathbf{A_b}$ which vary between 0.87–0.006. These values of the profile averaging kernels, mean that any errors in the cloud optical depth will cause errors in the retrieved tropospheric VCD.

Figure 4.12(c) shows the photon path lengths for a high surface LER of 0.8. Again the black line shows the photon path lengths when there is no cloud. Following this line through the atmosphere it can be seen that the photons travel through the atmosphere where their photon path lengths are increased as they decrease in altitude. This is due to the reflection of the photons by the surface. The photons behave in a slightly different way at high surface LER as there are more photons in the boundary layer due to reflection at the surface. Similar features are seen in the photon path lengths as with Figures 4.12(a) and 4.12(b) (increasing above the cloud with optical depth and decreasing below the cloud with optical depth). However, the averaging kernels are much closer to 1 here as the retrieval becomes more sensitive to changes in the concentration.

The values of $\mathbf{A_s}$ for high surface LER vary from 0.8–3.64 with increasing optical depth. However, the values decrease from 0.8 for no cloud to 0.72 at a cloud optical depth 0.6, then the value increases with increasing optical depth. This can be explained by studying the photon path lengths in Figure 4.12(c). The figure shows the no cloud photon paths through the atmosphere slightly increase with decreasing altitude, then the red line (cloud optical depth 0.01) shows slightly longer path lengths. The photon path lengths in the lower troposphere (below the cloud 4 km) increase until the yellow line (cloud optical depth 1) when the photon path lengths start to decrease with decreasing altitude below the cloud (below 4 km). The value of $\mathbf{A_s}$ for high surface LER and varying cloud optical depth shows that the retrieved tropospheric VCD is not as overly sensitive to changes in the stratospheric concentration as lower surface LER are, and therefore will be not be as strongly affected by errors in the stratospheric column.

Both the values of $\mathbf{A_t}$ and $\mathbf{A_b}$ are close to 1 until a cloud optical depth of 4 is reached. At a cloud optical depth of 30 these values reduce to $\mathbf{A_t}$–0.4 and $\mathbf{A_b}$–0.3. This means that when clouds are optically thick, (above 3-4) even at high surface LER, where there are more photons

and consequently more signal (higher sensitivity), the retrieved tropospheric VCD is insensitive to changes in concentration in the troposphere and boundary layer.

Figure 4.12 has shown that different cloud optical depths will produce very different photon path lengths which vary not only with changing optical depth, but also with changing surface LER. The values of the profile averaging kernels show that cloud optical depth will cause significant errors on retrieved tropospheric VCD, in terms of the stratosphere, troposphere and the boundary layer. The errors due to surface LER will be minimised as the surface LER is retrieved alongside the tropospheric scaling factor. This in turn will help to minimise the errors due to cloud (as when a cloud is present the cloud top reflectance will be retrieved). However, errors in the cloud parameters used to model the scene will be very important in this retrieval scheme. These will be discussed and quantified in Chapter 5.
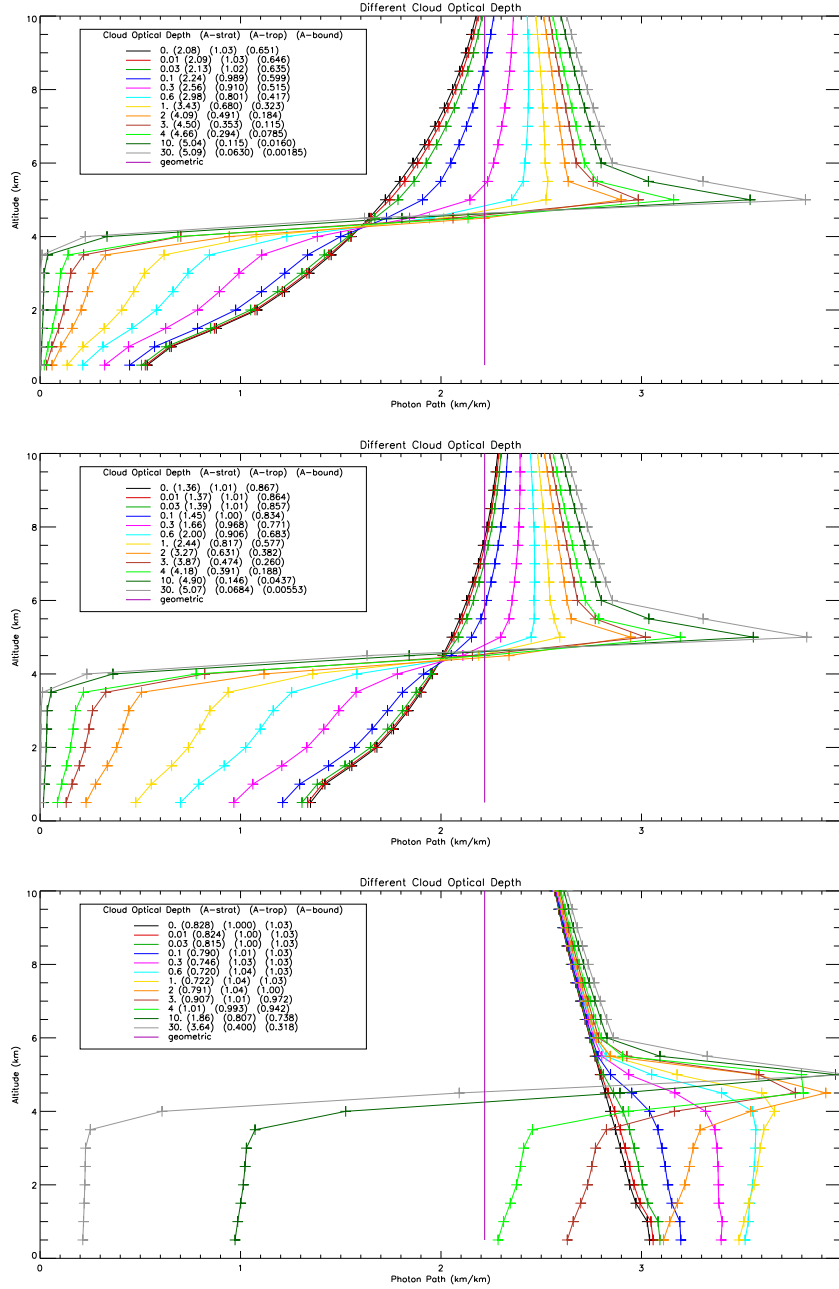
Figure 4.12: *The modelled photon path lengths (km/km) through the lower atmosphere when cloud optical depth is changed for (a) a surface LER of 0.02, (b) a surface LER of 0.1, and (c) a surface LER of 0.8. The legend indicates the different cloud optical depths and then the total tropospheric photon path normalised with the total tropospheric geometric photon path and the total boundary layer photon path normalised with the total boundary layer geometric photon path. The cloud is a water cloud, 1 km thick with the cloud top at a height of 3 km. The Forward Model was run for July at 55° N.*

Cloud Top Height

Figure 4.13 shows how cloud top height affects the photon path lengths. This figure is for a cloud optical depth of 30, which was shown in Figure 4.12 to create short photon path lengths under the cloud. Also it was shown that for clouds with optical depths of 30 the retrieved tropospheric VCD is highly sensitive to changes in concentration of the species at all levels in the atmosphere. By changing the height of the cloud top, more or less of the troposphere is obscured from the instrument. If the cloud top height is low, more of the troposphere is visible and therefore the values of $\mathbf{A_s}$, $\mathbf{A_t}$ and $\mathbf{A_b}$ are closer to 1. As the cloud top increases with altitude the values of $\mathbf{A_s}$ become much greater than 1, meaning that the retrieval puts more weight on the stratosphere than the troposphere. This is due to longer photon path lengths above the cloud and therefore in the stratosphere (at high cloud top height). The values of $\mathbf{A_t}$ and $\mathbf{A_b}$ will decrease with increasing cloud top height (especially here where the cloud optical depth is 30). This is because more of the troposphere is becoming obscured from the satellite. Similar variations in the profile averaging kernels are observed for varying surface LER as for cloud optical depth, i.e. they improve with increasing surface LER, due to the increase in the number of photons in the lower atmospheric layer.

Overall the figure shows that cloud top height will significantly affect the retrieval, especially for high cloud optical depths. Therefore, errors in cloud top height need to be accurately quantified.
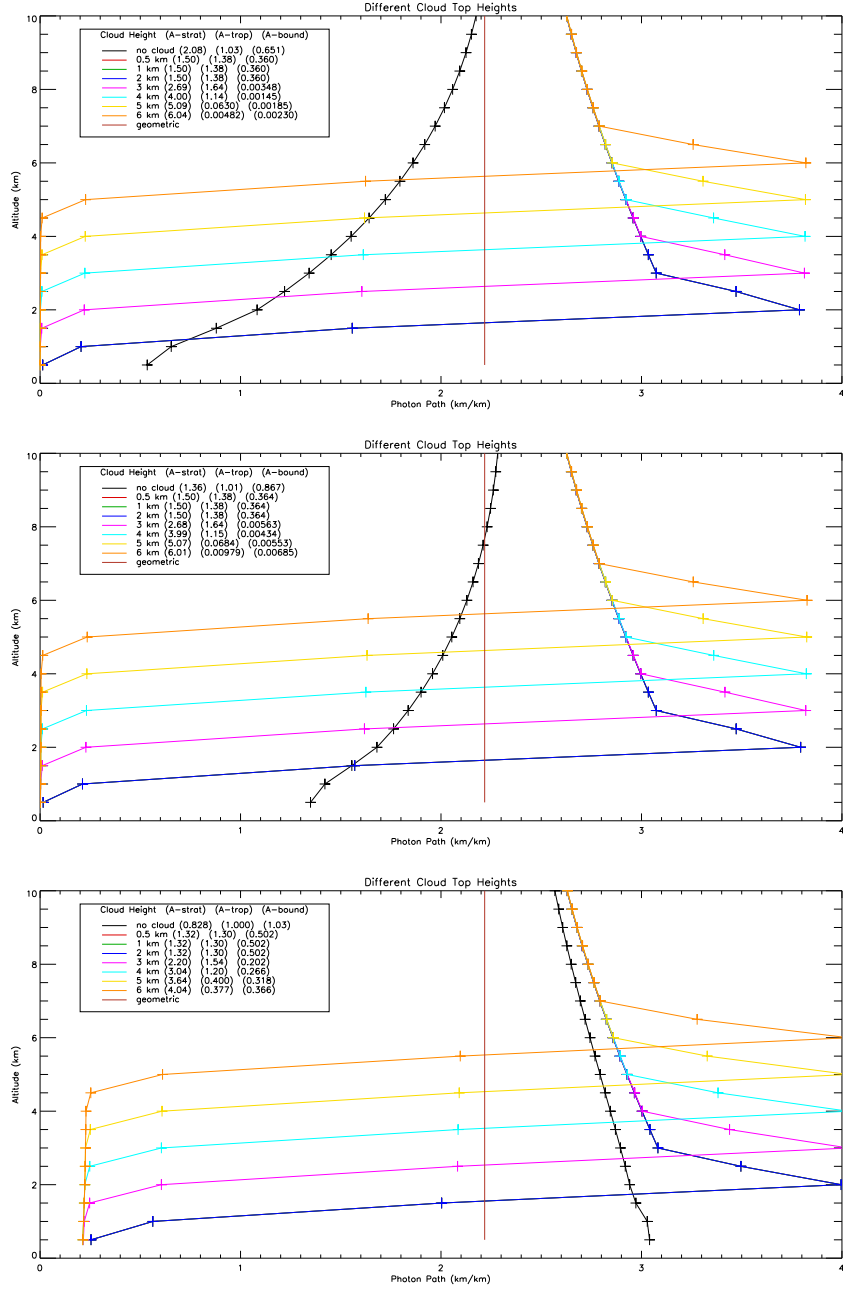
Figure 4.13: *The same as Figure 4.12, but for different cloud top heights, for a cloud with an optical depth of 30. (a) is for a surface LER of 0.02, (b) is for a surface LER of 0.1 and (c) is for a surface LER of 0.8.*

Shape of Single–Layer Cloud

Figure 4.14 shows how the cloud extinction profile affects the photon path lengths. Figure 4.14(a) shows cloud extinction with altitudes. Figures 4.14(b), 4.14(c) and 4.14(d) show the photon path lengths for every 200 m. This allows changes within the cloud to be observed. These plots show that changing the cloud vertical profile does not significantly effect the photon path lengths at any surface LER. This is also shown in the profile averaging kernels which remain similar for each different profile.

As this figure shows that profile shape has little impact on the photon path lengths or the profile averaging kernels a simple triangular profile (1) will be used in this retrieval model. This will have a similar affect at all wavelengths.

Multiple Layered Clouds

An interesting quantity to test in the Forward Model is the presence of multiple layer clouds. The GRAPE cloud data does not retrieve multiple cloud layers, as there is no optical model for this and it would be difficult to know when to apply such a model. However, it has been shown that a large measurement cost on the GRAPE retrieval signals that the scene does not fit the optical model well EUMETSAT (2007), which could suggest a multiple cloud layer. The Forward Model in this thesis can simulate multiple cloud layers. Figures 4.15 and 4.16 show multiple cloud layers with optical depths of 30 and 1, respectively. Figures 4.15(a) and 4.16(a) show the cloud profile shape and the extinction with altitude. The Forward Model has been tested with three different multiple cloud layer situations. Each cloud has a top at a height of 8 km and an equal optical depth from the cloud top to bottom. The higher cloud has a depth of 2 km and the lower cloud is at a height of 3 km and has a depth of 2 km. The scenario changes so that the first test is where the higher cloud is twice as optically thick as the lower cloud. For the second state the clouds have equal optical depths and the third state is when the lower cloud has twice the optical depth of the higher cloud[2]. Also shown on these figures is the cloud that GRAPE would retrieve in this case and the no cloud case. The GRAPE cloud simulation of what GRAPE would have retrieved in the case where the multiple cloud is actually observed. Figures 4.15 and 4.16 show that photon path lengths are significantly altered by the presence of a multiple cloud layer at both high and low optical depths. This is with the exception of when the cloud is optically thin and there is a high surface LER (as there are enough photons to correctly characterise the scene). The profile averaging kernels also agree with this result and for example, $\mathbf{A_t}$ ranges from 0.432 for cloud type 1 (with an optical thick cloud at a high level, and an optically thinner cloud at the lower level) to 0.87 for the GRAPE cloud that would have been observed in this case.

The values of the $\mathbf{A_s}$ are all very different but illustrate that having an optically thick cloud in the scene makes the retrieved tropospheric VCD very sensitive to changes in stratospheric concentration (i.e. $\mathbf{A_s}$ is much greater than 1). The value of $\mathbf{A_s}$ is different for high surface

---

[2]It should be noted that these scenarios are not based on realistic cloud profiles but have been designed to demonstrate the impact of a multiple cloud layer.
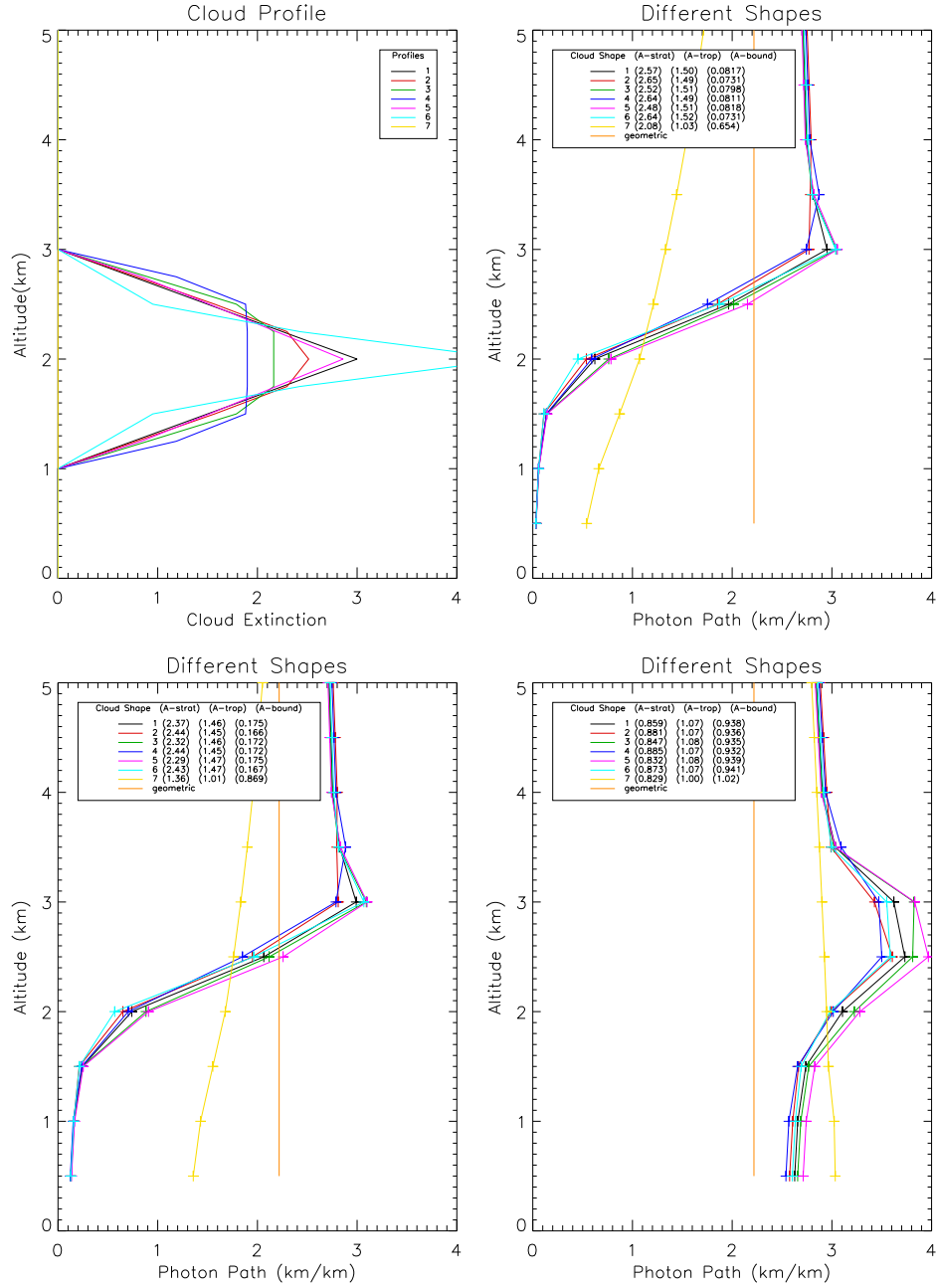
Figure 4.14: *The modelled photon path lengths (km/km) through the lower atmosphere when the mod-elled cloud shape is changed. Plot (a) shows the cloud extinction with altitude. The cloud is at 3 km with a thickness of 2 km and an optical depth of 3. Plots (b), (c) and (d) are as figure 4.12 but for different cloud shapes.*

LER with a low cloud optical depth. Here the value of $\mathbf{A_s}$ is less than 1 and the values are all similar to those that would be seen if the scene was modelled as a GRAPE cloud retrieval instead. This is expected as was shown in Figure 4.12, where for a low cloud optical depth at high surface LER the values of $\mathbf{A_s}$ are close to 1 for low cloud optical depths.

The values of the $\mathbf{A_t}$ are also very different for the different multiple cloud layers. The value of $\mathbf{A_t}$ varies for each test. For low cloud optical depths, at low surface LER the values vary from 0.99–1.34 for different layered clouds, with a value of 0.75 for the cloud observed by GRAPE. At higher surface LER there is not as much variation. $\mathbf{A_t}$ varies from 1.01–1.25 for surface LER of 0.1, with the GRAPE cloud being 0.87, and values being equal at a surface LER of 0.8. This has also been seen in Figure 4.14 where a cloud optical depth of 2 was shown to not significantly affect the photon path lengths and values of $\mathbf{A_t}$ at the higher surface LER. However, at higher cloud optical depths the differences are much greater (also as seen in Figure 4.14). For a surface LER of 0.02 the values of $\mathbf{A_t}$ vary from 0.411–1.55 for the different multiply layered clouds, with the GRAPE cloud being 0.749. This variation diminishes with increasing surface LER. These values of $\mathbf{A_t}$ mean that where possible it is best to ignore multiple cloud layer scene as firstly the distribution of the cloud within the layers will be unknown, and secondly this will lead to a difficulty in determining the errors due to the multiple cloud layers. Values of $\mathbf{A_b}$ reiterate the point that the boundary layer is not observed through an optically thick cloud and therefore errors can be large when an optically thick cloud is present.
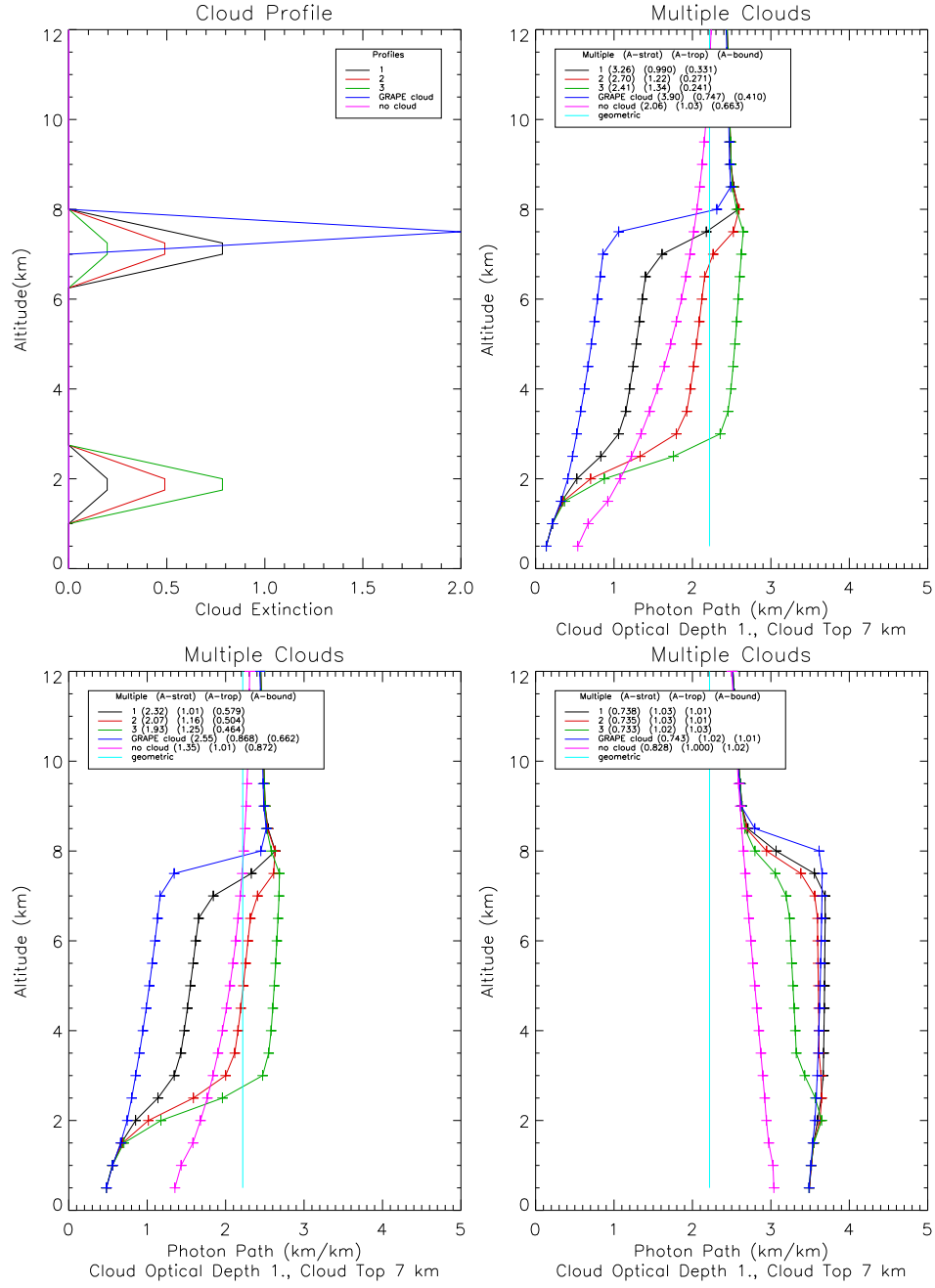
Figure 4.15: *The modelled photon path lengths (km/km) through the lower atmosphere when multiple clouds are of different shapes are present. Plot (a) shows the cloud profiles through the lower atmosphere in terms of their extinction, the cloud has an optical depth of 30. Panels (b)–(d) are the same as figure 4.12, but for multiple cloud extinction profiles.*
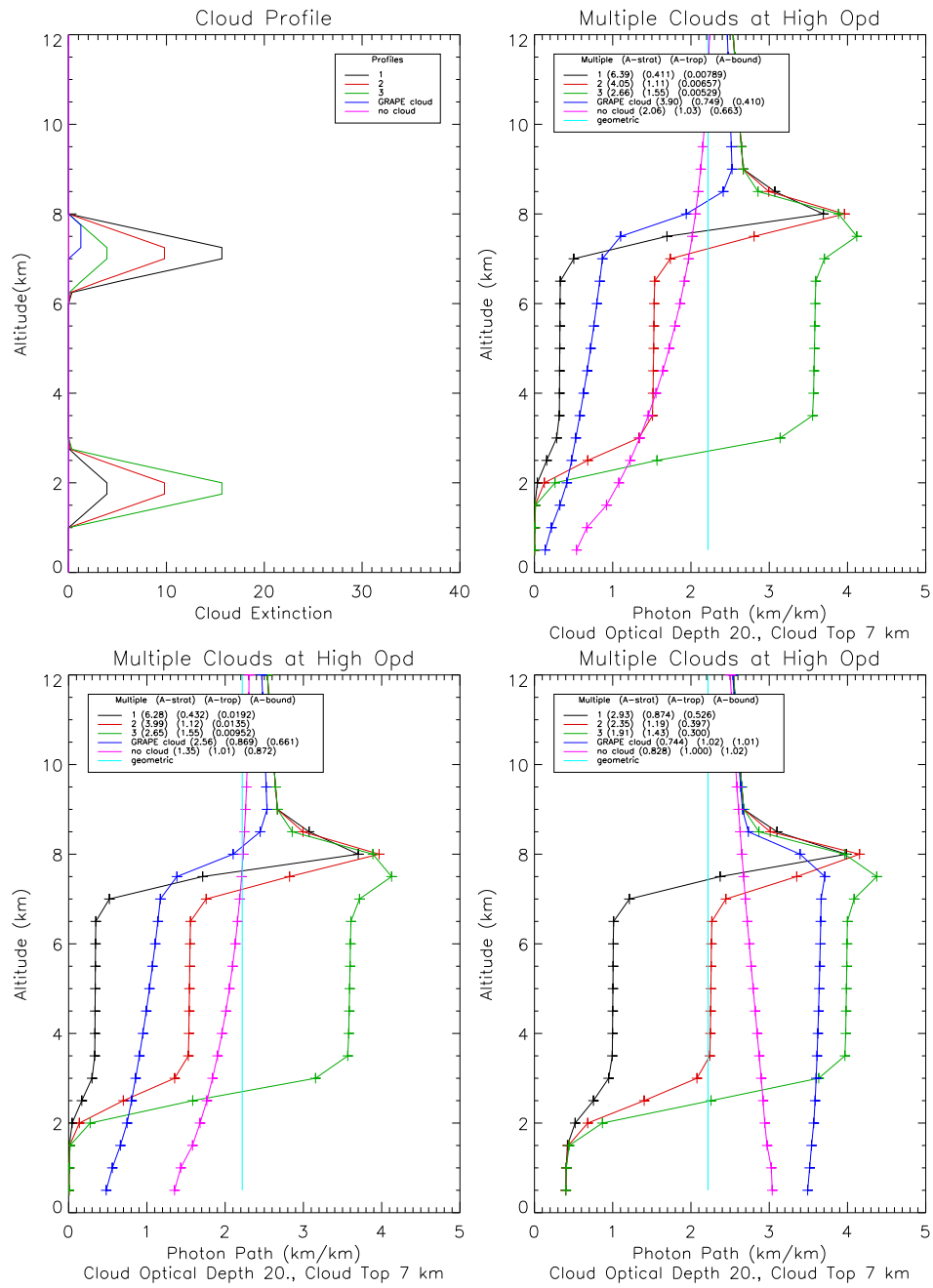
Figure 4.16: *As Figure 4.15 but for a cloud optical depth of 1.*

In this work the clouds are modelled with the same parameters that the GRAPE retrieval has used. However, these Figures 4.15 and 4.16 show that for multiply layered clouds will introduce significant errors to the retrieved tropospheric VCD. As explained earlier, it has been shown however, that if a larger measurement cost is associated with the GRAPE retrieval this can be assumed to be multiple cloud scene. Therefore as a large error will be introduced if these multiple clouds are present in this new retrieval scheme, GRAPE cloud parameters that have been retrieved but have a large cost will be ignored. Further implications of this will be discussed in Section 4.5, and in Chapter 5.

Once again this effect will be similar for different wavelengths, so cloud parameters with large costs will be ignored for all retrievals.

Cloud Geometric Depth

Figures 4.17 and 4.18 show how changing the geometric cloud depth affects the photon path lengths. In the GRAPE retrieval the clouds all have the same layer depth (infinitely thin) so this retrieval scheme models cloud with a similar cloud geometric depth to GRAPE (500 m). The figures show that the geometric depth of a cloud does not impact the photon path length at low cloud optical depth, but can do at high cloud optical depth. However, the profile averaging kernels in Figure 4.18 are similar for different geometrical depths, therefore the assumption of a cloud geometric depth of 500 m will be used in this retrieval scheme.

Although not shown in the section, cloud particle size and cloud type (i.e. ice or water) have little affect on the retrieved tropospheric VCD, and the photon path lengths are similar under different conditions. In the following section it will be shown that aerosol particle size can affect the photon path lengths. However, as discussed in Chapter 2, smaller particles experience more variations in their scattering. For example, at small particle sizes, radiation is both back and forward scattered. However, at large particle size (i.e. clouds) the radiation is mostly forward scattered. Therefore there is less variation in the scattering for cloud particle sizes, therefore less error.
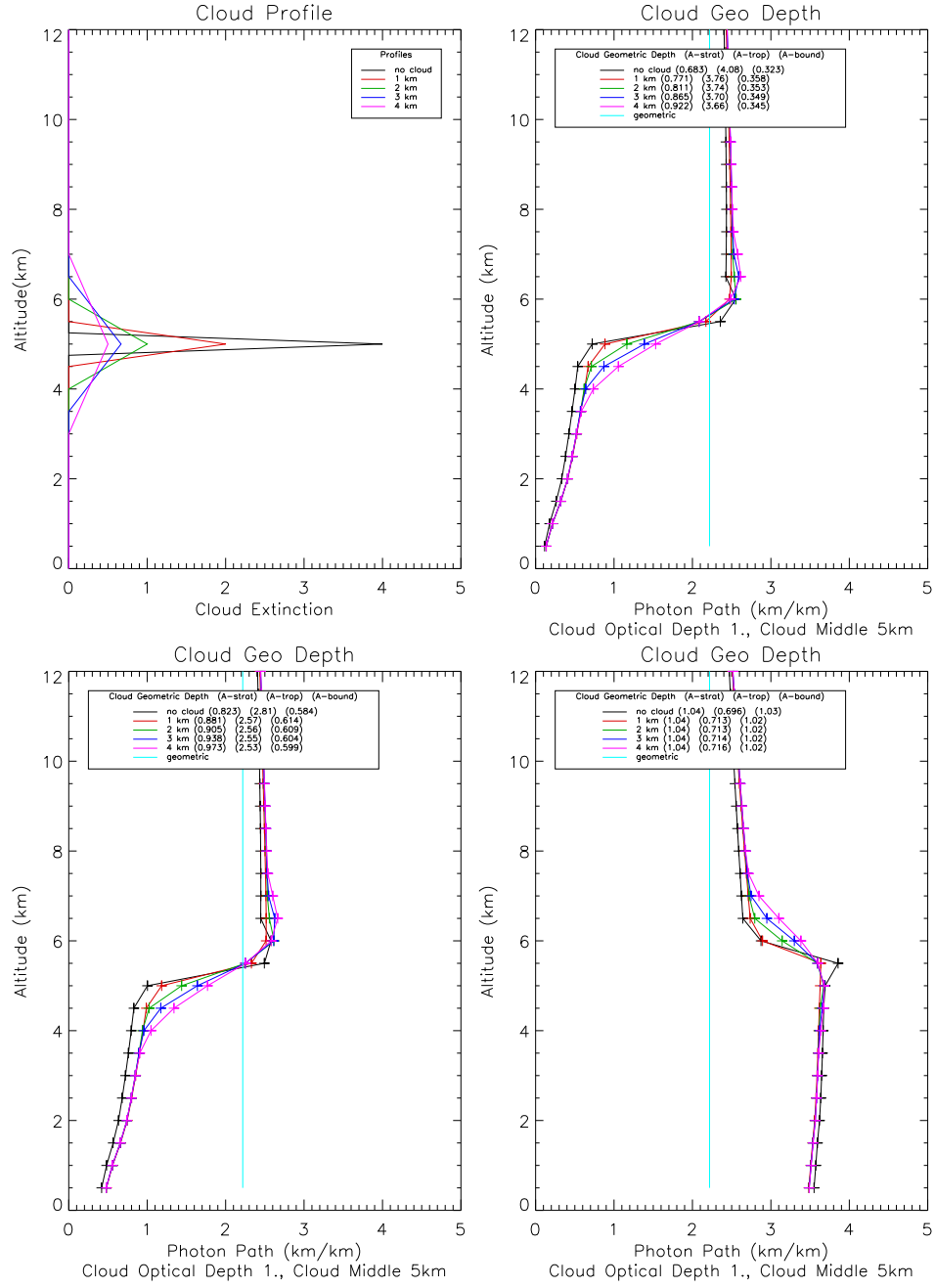
Figure 4.17: *The modelled photon path lengths (km/km) through the lower atmosphere when the clouds geometric depth is changed. Plot (a) shows the cloud profiles through the lower atmosphere in terms of their extinction, the cloud has an optical depth of 1. Plots (b), (c) and (d) are the same as figure 4.12, but for changing geometric cloud depth.*
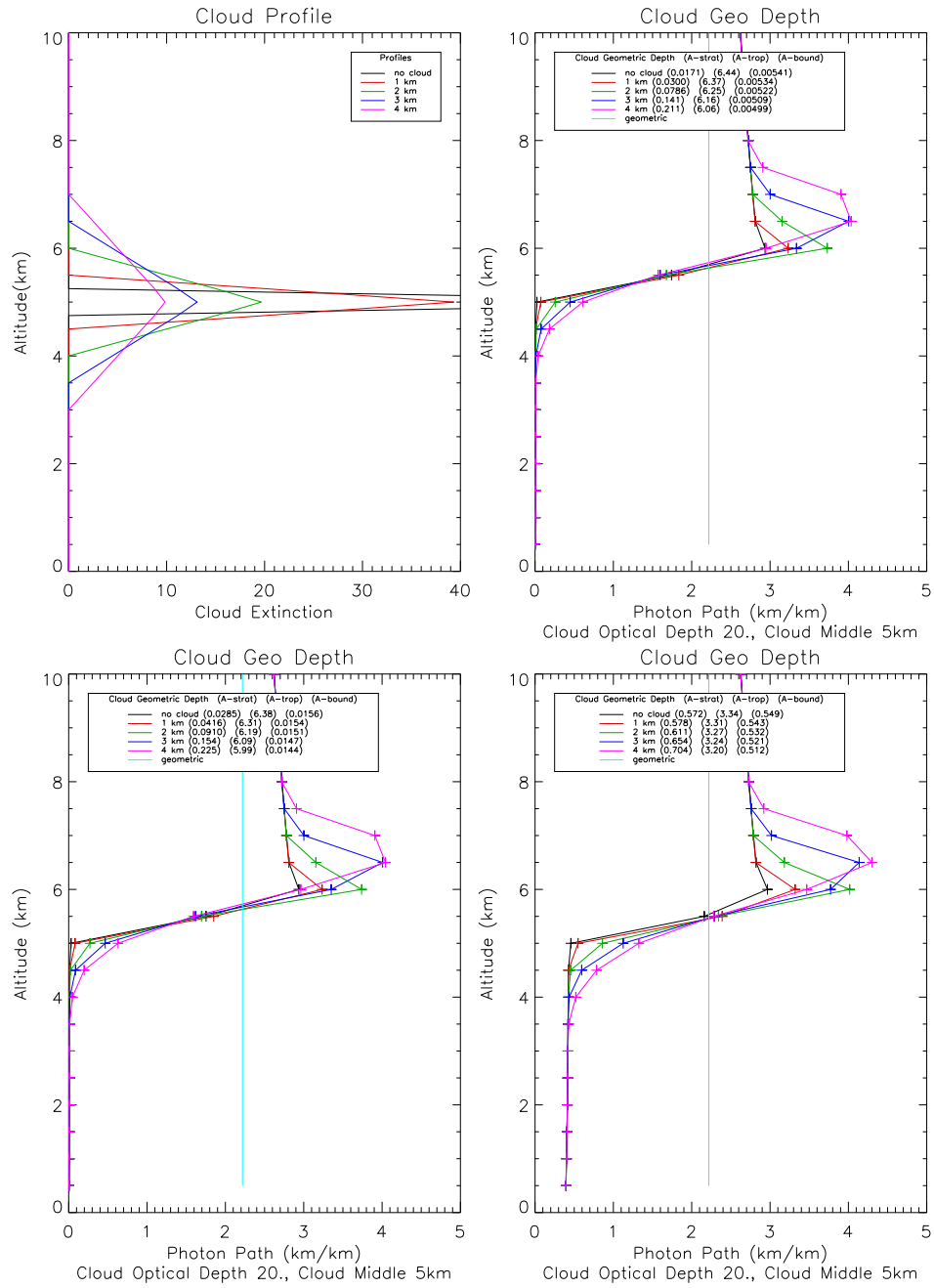
Figure 4.18: *As figure 4.17, but for a cloud optical depth of 30.*

125

4.3.5 Aerosol

Aerosol is present in both the troposphere and stratosphere. In general it has a lower optical depth than clouds and therefore when a cloud is in the scene it will be less important than the cloud errors. Aerosol will have an effect in clear sky conditions though and therefore this needs to be tested. For the following tests the Forward Model has been run for aerosol parameters that are seen in nature and for ranges that will affect the specific wavelengths used in the retrievals within this thesis (e.g. UV/VIS radiation). From Mie theory it can be shown that aerosols with effective radii between 0.05 and 1.0 μm will affect the radiation of the relevant wavelengths (350 and 430 nm) (Van de Hulst, 1981). Also, a number of ground based studies (see Raes et al., 2000; Dubovik et al., 2002) have shown that aerosol optical depth varies from 0.05-0.1 for the natural background conditions, to 0.5 for natural anthropogenic background (generally over cities) and 1. for pollution events (biomass burning events, power station plumes).

Figure 4.19 shows how aerosol optical depth affects photon path lengths. The figure shows similar results to those of the cloud optical depth but with a smaller effect due to the lower optical depth. The photon path lengths decrease slowly with decreasing altitude until the top of the boundary layer, where the aerosol concentration is highest. For these tests, aerosol is assumed to be continental in source and has a profile described in Chapter 2. At the top of the boundary layer (set to 2 km) a small amount of reflection is observed, this is not as much as with clouds due to different scattering properties of aerosol and cloud particles, and due to the significantly lower optical depth. Once in the boundary layer, below 2 km, the photons are scattered. It should be noted that the decrease in photon path length is due to the low surface LER.

Values of $\mathbf{A_s}$ are similar at surface LER values above 0.1. At the low surface LER values though, the values vary from 1.6–2 for aerosol optical depths of 1–0.1. This can be explained by looking at the photon path lengths. It can be seen that at the higher optical depths the photon path lengths are longer, at around 2 km in altitude. This is due to multiple scattering in the aerosol layer. At lower aerosol optical depths this is not seen and the value of $\mathbf{A_s}$ is similar to that seen when there is no cloud or aerosol present but the surface LER is low. These values of $\mathbf{A_s}$ highlight that even for scenes with just aerosol, if the surface LER is low the stratosphere needs to be modelled very accurately in order to get an accurate value of the tropospheric VCD.
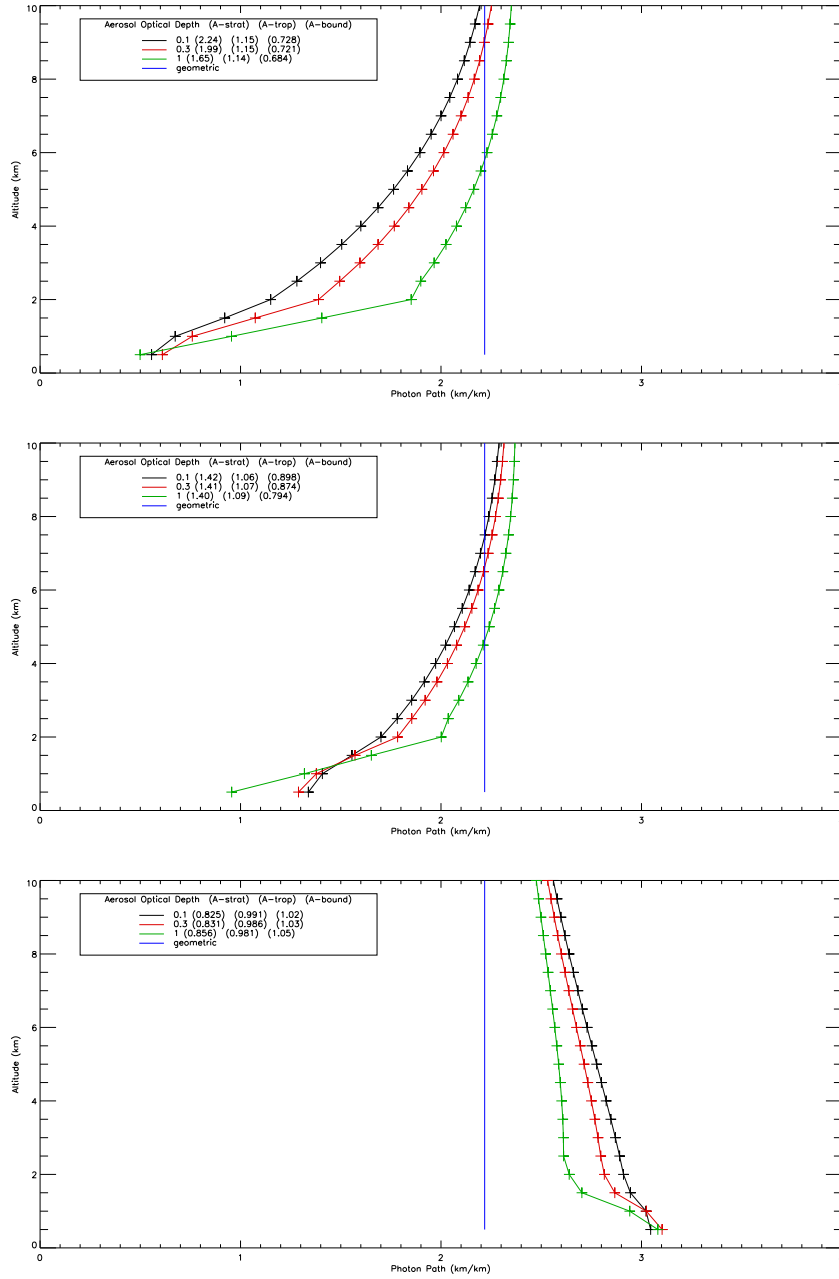
Figure 4.19: *The modelled photon path lengths (km/km) through the lower atmosphere when the aerosol optical depth is changed. Note that the aerosol optical depths are larger than expected to allow the model changes to be enhanced. Plots (a), (b) and (c) are for different surface LER, the top plot is for a reflectance of 0.02, the central plot 0.1, and the lower plot 0.8. The legend identifies the different aerosol optical depths for the total tropospheric photon path normalised with the total tropospheric geometric photon path and the total boundary layer photon path normalised with the total boundary layer geometric photon path. The aerosol type is clean continental with a layer of 1 km at the bottom of the atmosphere. The particle sizes are assumed to be 55 μm. The Forward Model was run for July at 55° N.*

127

Values of $\mathbf{A_t}$ and $\mathbf{A_b}$ are similar for all aerosol optical depths, highlighting that aerosol optical depth. In Figure 4.6 it was seen that in a clear sky retrieval $\mathbf{A_t}$ was close to 1. In Figure 4.19 $\mathbf{A_t}$ is greater than 1 suggesting that modelling aerosol will improve the retrieval. However, the value of $\mathbf{A_s}$ and $\mathbf{A_b}$ are relatively similar when varying aerosol optical depth showing that errors in the modelled optical depth will not impact the value of the retrieved tropospheric VCD.

Figure 4.20 shows how the aerosol particle size affects the photon path lengths. The aerosol particle size will have a different affect on photon path length to cloud particle size as they are smaller, i.e. more Rayleigh and Mie scattering as oppose to geometric optics in clouds (see Chapter 2). With aerosol parameters the Forward Model behaves slightly differently. This is because the particle size is defined at 550 nm and then this affects the optical depth which will change at different wavelengths (see Chapter 2). However, the particle size variation also has a small effect on the photon path lengths, which is more noticeable at lower surface LER. The values of $\mathbf{A_s}$ at low surface LER are also the most variant (as was for aerosol optical depth) varying from 1.29–1.79 from aerosol particles sizes of 0.05–1 μm for a surface LER of 0.01 and from 1.2–1.49 for the same size range but for a surface LER of 0.1. This shows that the larger particle size causes more scattering in the lower layers of the atmosphere and therefore allows for more of the target species to be observed. As more scattering occurs at low wavelengths, (there is more Rayleigh scattering in the BrO wavelength range than the NO_2 wavelength range) error in the BrO retrieval relating to particle size are not as important as they are in the NO_2 wavelength range. However, aerosols will not significantly affect the retrieved tropospheric VCD except in clear sky conditions, and then the errors are much smaller than those observed due to cloud.

The value of $\mathbf{A_s}$ is highly variant and, as in other plots, this highlights that it is important to model the stratospheric component as well as possible. The value of $\mathbf{A_t}$ varies by 0.08 from particle sizes of 0.05–1 for low surface LER values and by 0.4 at high surface LER values.
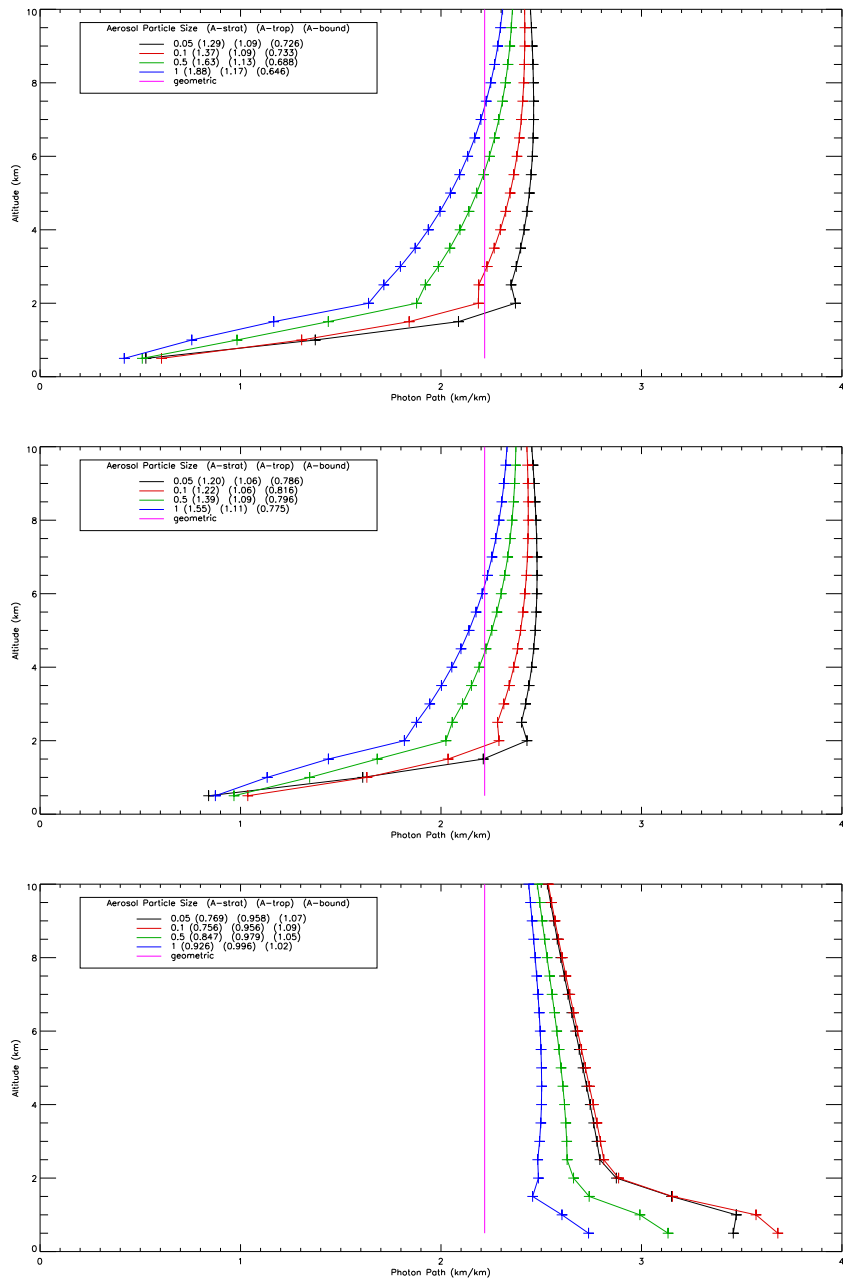
Figure 4.20: *As Figure 4.12 but for changing aerosol particle size and with an aerosol layer optical depth of 0.05.*

Figure 4.21 shows how the aerosol type affects photon path lengths. The Forward Model was run under the same conditions as the run with changing surface LER but with an aerosol layer. The aerosol particle size of 0.55 μm is used. The different aerosol types have different scattering properties, these were discussed in Chapter 2.

Figure 4.21 shows that different aerosol types produce different photon path lengths. The values of $\mathbf{A_s}$ vary largely with different aerosol types, however, for $\mathbf{A_t}$ and $\mathbf{A_b}$ this variation is less. This means that the stratosphere needs to be accurate if the aerosol type is not known. However, the retrieved tropospheric VCD will not be as greatly affected by errors in the aerosol type. It should be noted that the aerosol type and hence the aerosol extinction profile and properties are calculated from look–up tables (see Chapter 2) and it is hard to predict the error in this parameter.

4.3.6   Different Scenarios

The satellite viewing angle will significantly affect the photon propagation. This is shown in Figure 4.22. Here the different scenarios are defined in Table 5.2 (describes the different solar zenith angle for each scenario) and Figure 5.1 (which shows the profile shapes for each scenario). In Figure 4.22 the green line shows scenario July 55°N which is assumed in the other photon propagation figures in this section. As Table 5.2 shows there are 4 scenarios with low solar zenith angles, 30°–40°, (April 5°N, July 5°N and 55°N, and October 5°N). Low solar zenith angle means that the photon path lengths are shorter as the radiation comes in closer to the nadir (solar zenith angle of 0) and therefore spends less time in each layer. July 75°N has a solar zenith angle of 53° and high solar zenith angles, 60°–85° are observed for scenarios October 55°N, 75°N and 75°S. For high solar zenith angles the photon path lengths are longer as the radiation spends longer in each layer. This is seen in Figure 4.22.
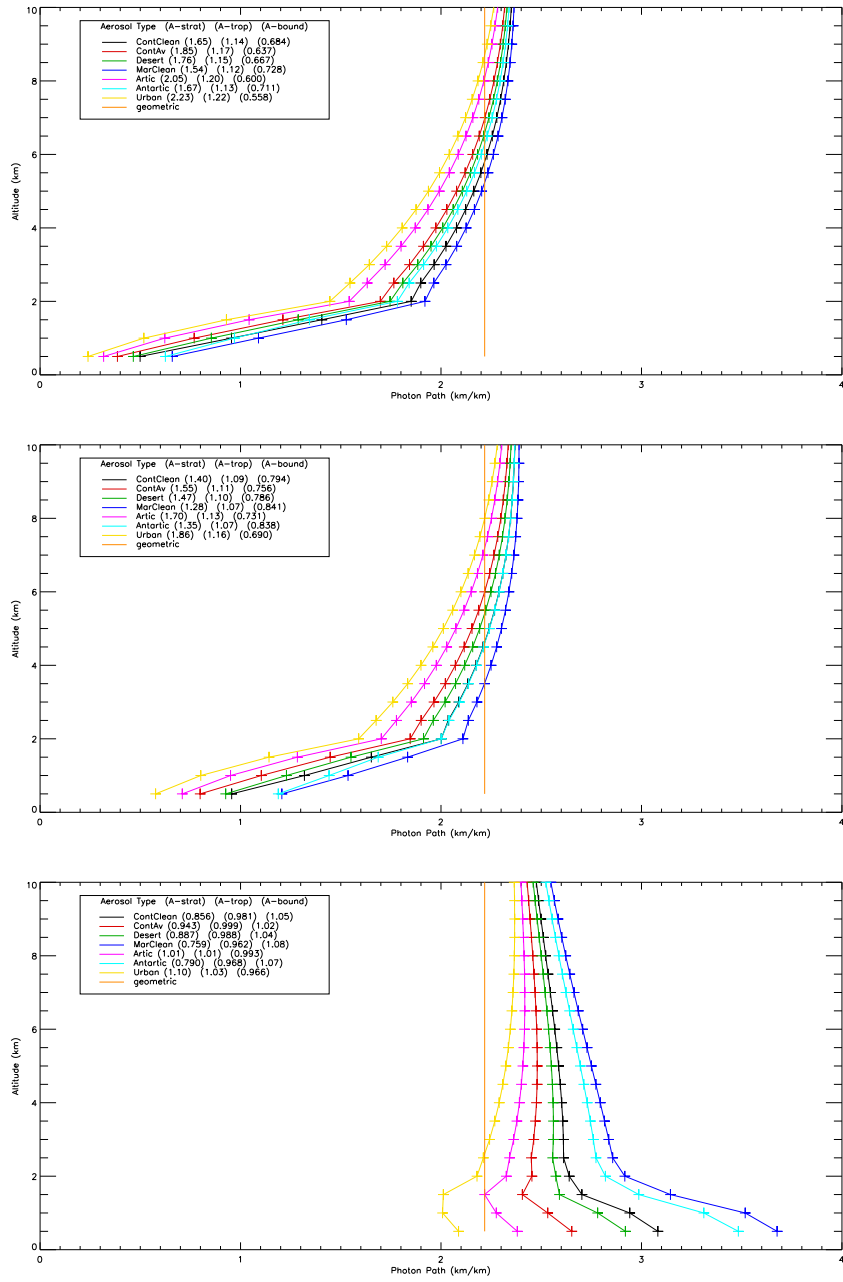
Figure 4.21: *As Figure 4.12 but for changing aerosol type.*
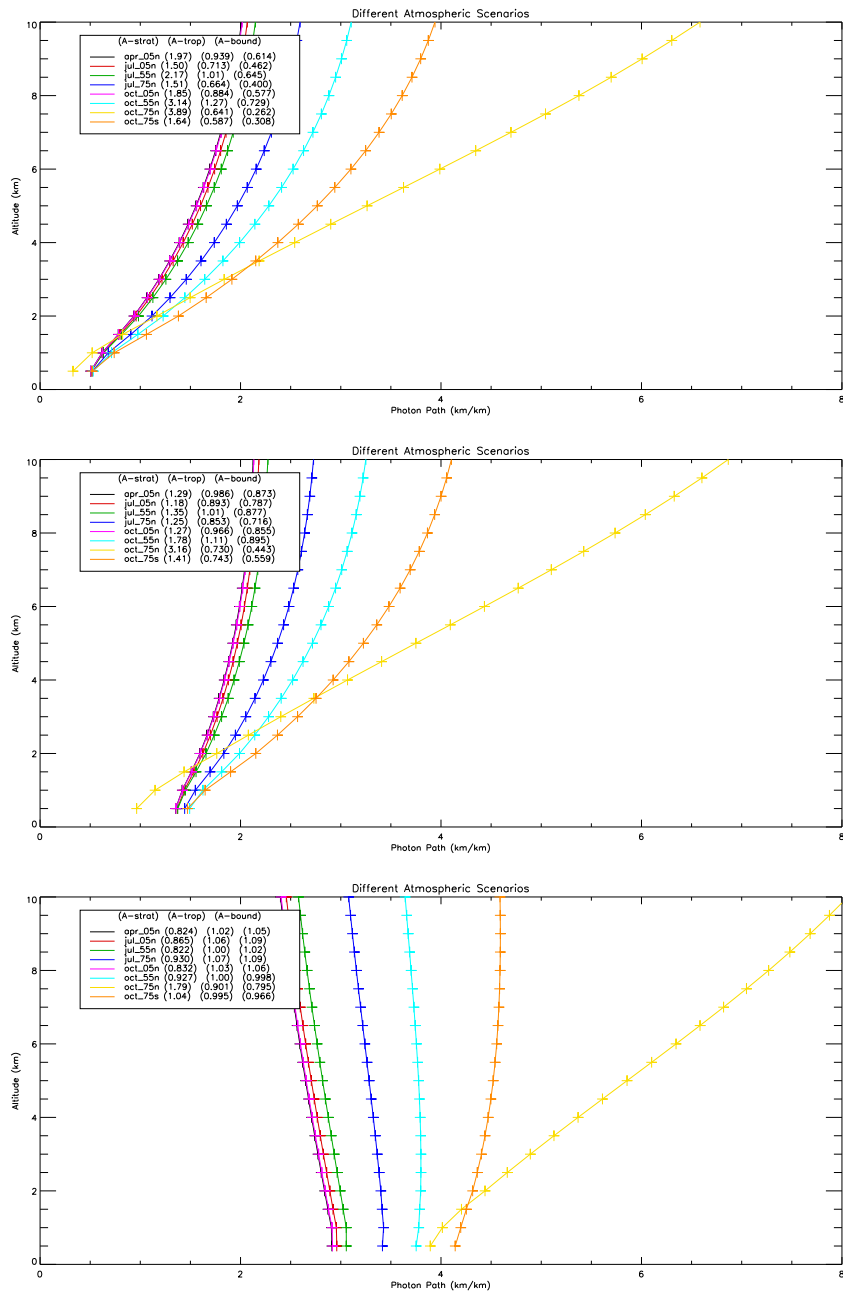
Figure 4.22: *As Figure 4.12 but changing the scenario (viewing angle).*

4.3.7 Computational Constraints

Number of Streams

Figure 4.23 shows how the photon path lengths vary when the number of streams are changed. The number of streams is a measure of the number of scattering angles used to solve for multiple scattering. The figure shows how many streams are needed to produce the most realistic radiative transfer calculation.

Figure 4.23(a) shows the photon paths at a surface LER of 0.02. The figure shows that the minimum number of streams needed is around 3. This is consistent with Figures 4.23(b) and 4.23(c). However, the presence of cloud will mean more streams are needed. This is observed in Figure 4.24 which shows that around 15 streams are needed to solve the radiative transfer equation. This will be used in this retrieval scheme.

Layer Depth

Figure 4.25 shows the layer depth needed for the radiative transfer equation to be solved.

Figure 4.23(a) shows layer depth used at a surface LER of 0.02. The figure shows that the minimum layer depth needed is around 0.5 km. This is consistent with Figures 4.23(b) and 4.23(c). However, the presence of cloud will mean layer depth needs to be decreased. In this work a layer depth of 0.25 km is used in the bottom 20 km of the atmosphere, then a layer depth of 1 km is used between 20–60 km. When a cloud is present the layer depth is deceased around the cloud, such that it is scaled according to the optical depth (an atmospheric layer is not allowed to have an optical depth greater than 2).

Use of a Single Wavelength within the Fitting Window

One of the main assumptions used it the retrieval model developed within this thesis (and used in retrieval schemes developed by other groups (Martin et al., 2002), is the use of a single wavelength to determine the radiative transfer within the fitting window. Figure 4.26 shows the photon path lengths for different wavelengths. The figure shows that there is little difference in the photon path lengths at different wavelengths. Also the averaging kernels show little variation which also highlights that the use of a single wavelength will not produce a significant error in the retrieved tropospheric VCD.
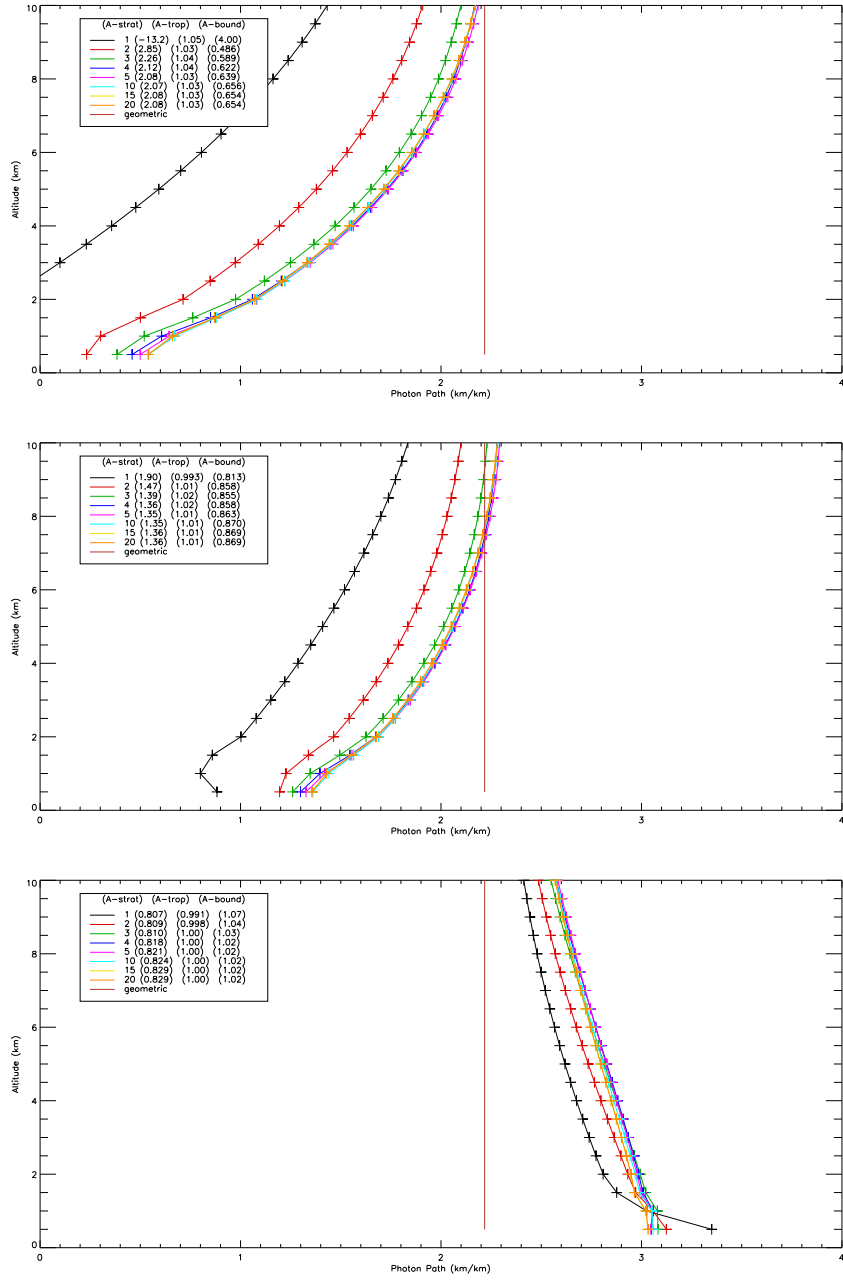
Figure 4.23: *As Figure 4.6 but for varying the number of streams.*

134
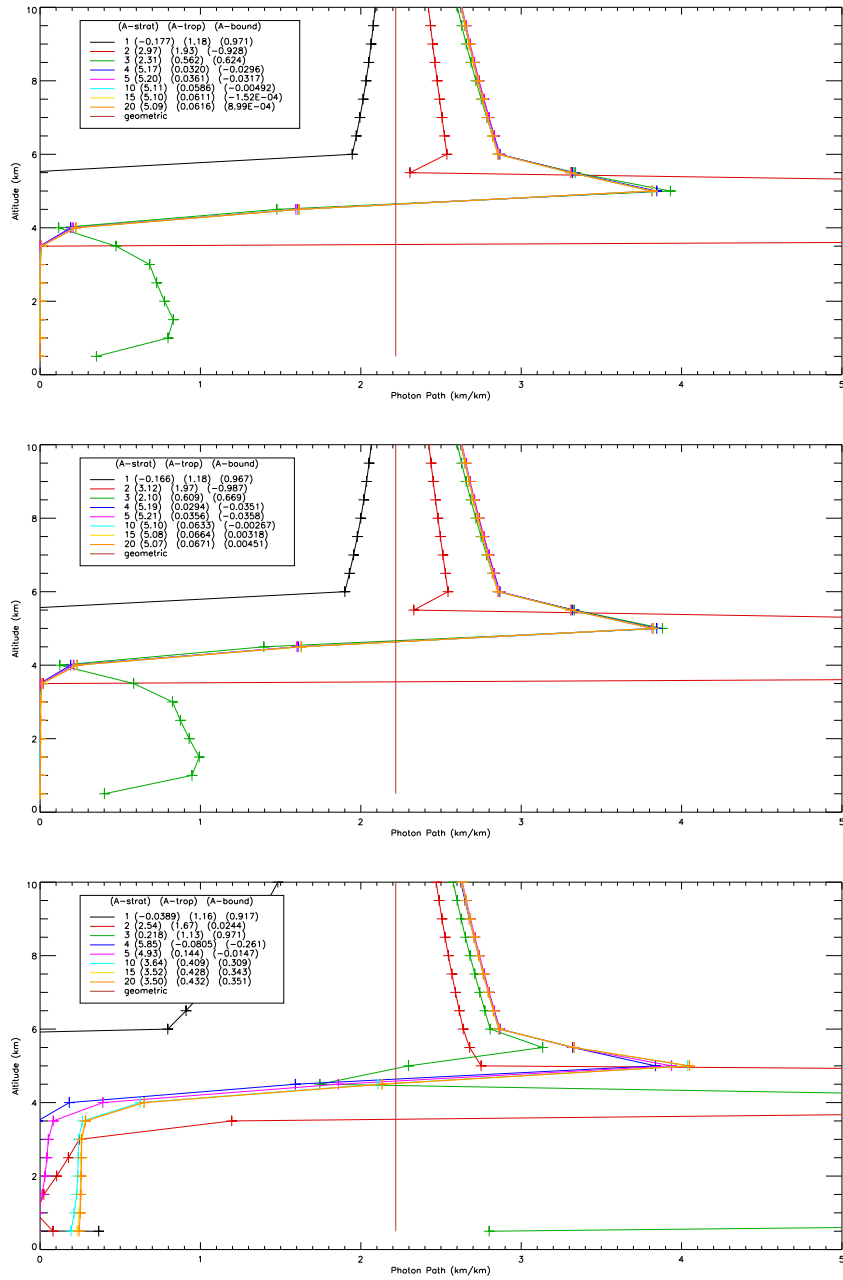
Figure 4.24: *As Figure 4.23 but for varying the number of streams when an optically thick cloud is present.*
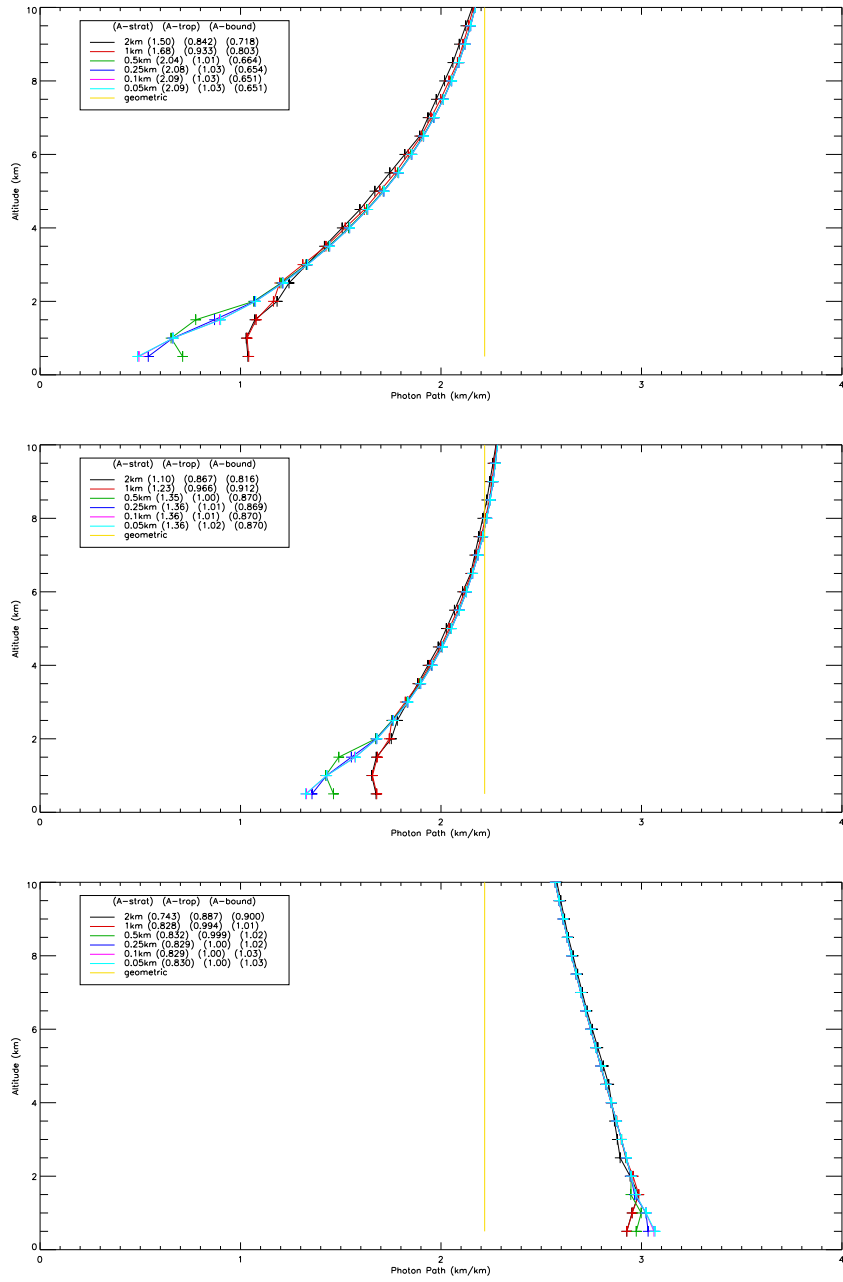
Figure 4.25: *As Figure 4.6 but for varying the layer depth.*
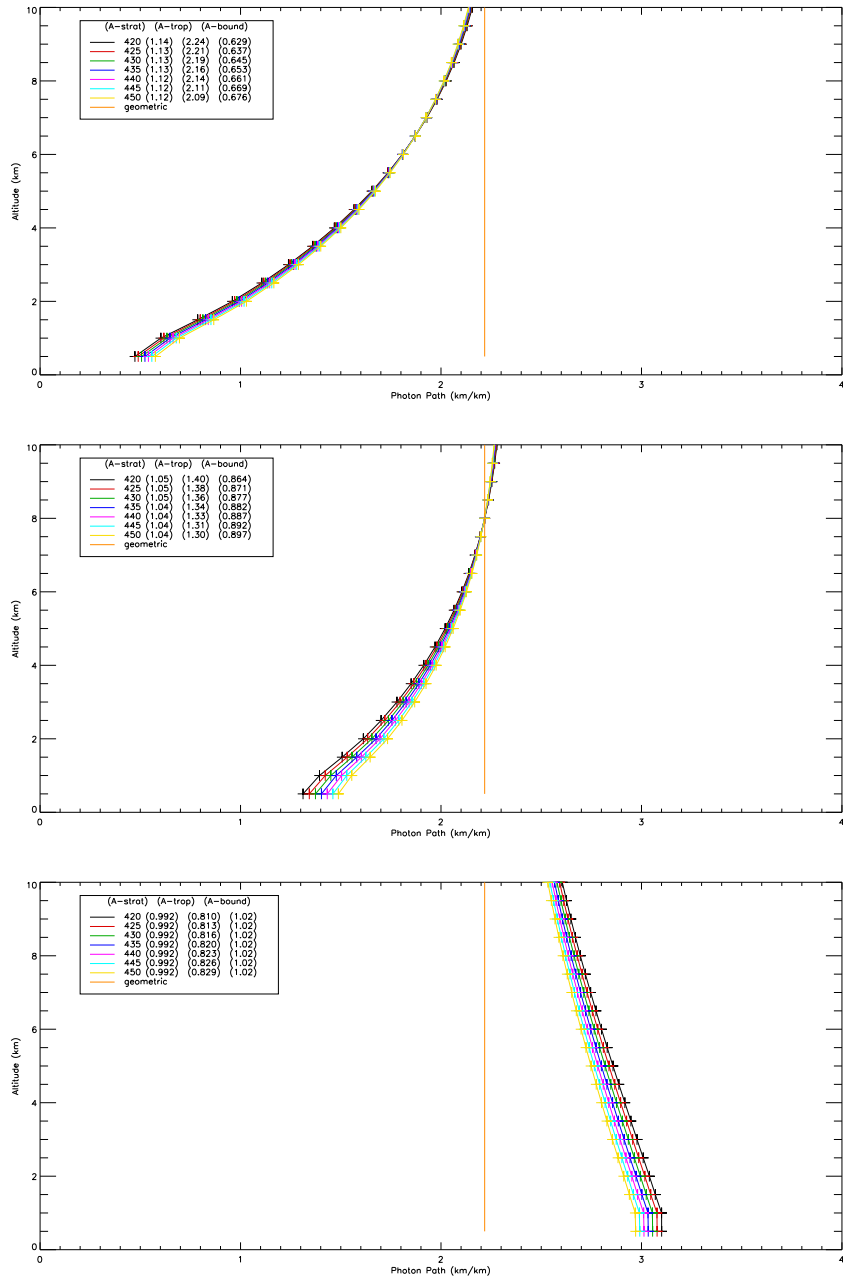
Figure 4.26: *As Figure 4.6 but for varying the wavelength of the retrieval.*

4.3.8 Summary of Forward Model Tests

Table 4.2 summarises the finding of this section. The main conclusions where that photon path lengths vary widely with changing cloud optical depth and cloud top height. Also multi–layered clouds have very different photon path lengths. This means it is necessary to model these three components as carefully as possible. The testing also provided insight to the model parameters. The number of streams necessary to run the retrieval model is around 15 and this will be used in the final retrieval runs. Also as surface reflectance causes highly varying photon path lengths this will be retrieved alongside tropospheric scaling factor. Another important parameter that has been determined is the number and size of layers. This has been found to be 0.25 km layers from 0–20 km, then 1 km to 60 km.

Table 4.2: *Summary of Forward Model Test Findings*

Forward Model Parameter	Summary
Lambertian Equivalent Reflectance	Different surface LERs cause large differences in the observed photon path lengths. Will cause error in tropospheric column if reflectance is in error.
Cloud Optical Depth	Different cloud optical depths cause large differences in the observed photon path lengths. Errors in cloud optical depth will cause errors in the tropospheric column. Assessed further in Chapter 5.
Cloud Top Height	Different cloud top heights cause large differences in the observed photon paths lengths. Errors in cloud top height will cause errors in the vertical column density. Assessed further in Chapter 5.
Shape of single–layer cloud	This does not significantly affect the photon paths and therefore a simple symmetrical cloud profile will be used in the retrieval.
Multi–Layered Cloud	Multiply layered clouds will significantly affect the photon paths. GRAPE does not retrieve multiple clouds and therefore the impact of this needs to be addressed.
Cloud Geometric Thickness	This does not significantly impact the photon paths over the range that has been simulated. Cloud geometric thickness of 1 km will be used in the retrieval.
Cloud Particle Size	Figure not shown but there is no change in photon path length for changing cloud particle size between 5–100 μm.
Cloud Type	Figure not shown but photon path lengths are not affected by change type from water to ice.
Aerosol Optical Depth	Photon path lengths only differ at optical depths of 1 and above. These are not common for aerosol and therefore the aerosol optical depth will have little effect on the slant column retrieval.
Aerosol Particle Size	Photon paths are not affected by changing aerosol particle size that have been tested (0.05–1 μm).
Aerosol Type	Photon path lengths are altered for different aerosol types. This is difficult to quantify as aerosol type is assumed, however it will be assessed in Chapter 5.
Tropospheric Species Shape	Photon paths are not affected by changing tropospheric species profile shape, however the slant column density will be through Equation 4.34. This will be addressed in Chapter 5.
Number of Streams	Optimum number of streams chosen is 15.
Number of Atmospheric Levels	Optimum chosen to be 0.25 km layers up to 20 km, thereafter 1 km to 60 km.
Different Wavelengths	Range of wavelengths within the fitting window for NO_2 were chosen. This showed little variation in the photon paths or averaging kernels, negligible affect on the retrieved tropospheric column (Martin et al., 2002).

Table 4.3 summarises the fractional error in the retrieved tropospheric VCD as calculated from the tropospheric profile averaging kernels. These are upper limits on the errors as the profile averaging kernels do not consider the improvement which will be achieved by retrieving surface reflectance alongside the tropospheric column amount (which happens in the final retrieval model).

Table 4.3: *Summary of fractional errors in retrieved tropospheric VCD as calculated from the values of the tropospheric profile averaging kernels. These values are upper limits as in final retrieval runs the surface LER will be retrieved alongside tropospheric profile therefore lowering the error. The true value would be the realistic parameter, i.e. that which is true of the atmosphere at say the time of measurement, and the assumed value is the true value plus an error, i.e. a model value. Using these different values should result in different retrievals. The difference between the tropospheric VCD that is calculated in each case will give some indication of the sensitivity of the retrieval to an error in each specific parameter. The term GRAPE cloud, refers to the cloud that GRAPE would have simulated, i.e. GRAPE will represent a multi–layer cloud as a single layer.*

Parameter	True Measurement	Assumed Measurement	Surface Reflectance	Fractional Tropospheric VCD Error
Surface LER	0.1	0.01	na	0.01
NO$_2$	0.1	0.8	na	-0.01
Surface LER	0.1	0.01	na	0.21
BrO	0.1	0.8	na	-0.33
Tropospheric	1	2	0.02	-0.34
Species Profile	1	2	0.1	-0.21
	1	2	0.8	0.06
Cloud Optical	10	30	0.02	-0.45
Depth	10	30	0.1	-0.55
	10	30	0.8	-0.5
	3	4	0.02	-0.16
	3	4	0.1	-0.18
	3	4	0.8	-0.02
Cloud Top	3 km	4 km	0.02	-0.44
Height	3 km	4 km	0.1	-0.44
	3 km	4 km	0.8	-0.28
Multi-Layered	3	GRAPE cloud	0.02	-0.44
Cloud	3	GRAPE cloud	0.1	-0.30
Low Optical Depth	3	GRAPE cloud	0.8	1
Multi-Layered	3	GRAPE cloud	0.02	-0.52
Cloud	3	GRAPE cloud	0.1	-0.44
High Optical Depth	3	GRAPE cloud	0.8	-0.29
Aerosol Optical	0.3	1	0.02	-0.02
Depth	0.3	1	0.1	-0.02
	0.3	1	0.8	-0.01

4.4   Spatial Cloud Distribution

From Section 4.3 it was seen that clouds will significantly affect the path of radiation through the atmosphere (e.g. the photon path lengths) and consequently will affect the tropospheric VCD that is retrieved. It is therefore important that cloud parameters are modelled in the best possible way. One limitation of the GOME instrument is its large ground pixel 40 km × 320 km, especially when characterising the cloud as cloud generally varies on much smaller scales than this. For cloud retrievals from GOME the polarisation measurement devices (PMDs) are used, which have a better spatial resolution with a ground pixel size at 20 km × 40 km. Although this resolution is better for cloud retrievals, the ATSR–2 instrument (from which cloud data is used in this work - GRAPE) has even better spacial resolution at 1 km × 1 km, which results in around 300–500 ATSR–2 pixels within one GOME pixel (see Figure 1.5), depending on the viewing angle of the satellite. Although the cloud data is now more accurate this information needs to be averaged into the GOME scene. Ideally, the GOME pixel would be sub–divided so that each GRAPE pixel is individually represented within a GOME pixel. In reality, this is computationally prohibitively expensive and so it is necessary to average the GRAPE information while providing the best possible fit to the real data. Different techniques for carrying out this averaging are described in Section 4.5.

Dividing the GOME pixel into different scene types is simple as each section of the GOME pixel just needs to be weighted (Richter and Burrows, 2001), i.e. for sun–normalised radiance

$$r = \sum_{i=1}^{N} r_i f_i \qquad (4.39)$$

where N is the number of scene types, r_i is the sun–normalised radiance in the scene type, and f is the fraction of the pixel covered by the scene type. This equation is a good approximation of r, as it does not include 3–D affects it is not exact. Should note that, as stated in Section 1.2, to obtain r in 3–D would require a 3–D instrument, therefore 1–D is the best possible approximation for the level of sophistication of the instrument.

4.5 Using GRAPE data to Characterise GOME Field-of-View

The GRAPE data set contains retrieved cloud and aerosol parameters from ATSR–2. As stated in Section 4.4 it is important to average the GRAPE cloud information in the optimum way to allow for the most accurate representation of the GOME scene. The following section assesses a number of different averaging techniques that can be used to achieve this.

4.5.1 Averaging into Cloud zones

Three methods were examined for the averaging of the GRAPE parameters. The first and computationally cheapest method was to take the pixels that GRAPE flagged as cloudy and clear and apply two scene types, one with cloud and aerosol and one with just aerosol. This is shown in Figure 4.27(a). The figure shows three colors (purple–no cloud, green–liquid cloud, red–ice cloud) which identify where cloud and cloud–free pixels are present. It can be seen in this figure that there is a large cloud that cover the left hand-side of the GOME pixel.
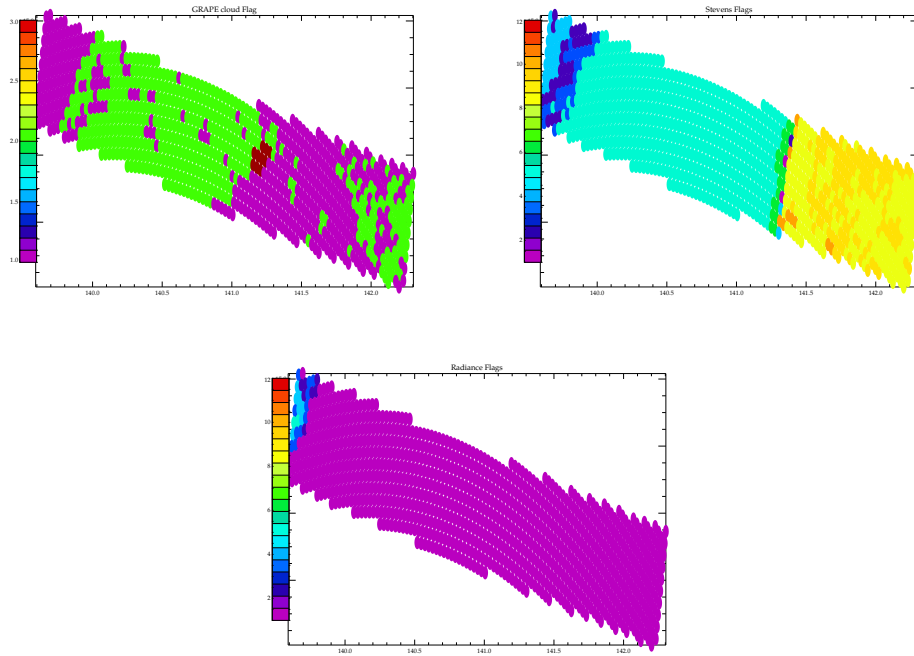
Figure 4.27: *Plots of a GOME pixel with the different averaging flags overlaid. Plot (a) is the ATSR cloud flags overlayed, plot (b) is as (a), but with the Stevens flags overlayed and plot (c) is as (a) but with radiance flags overlayed.*

The second method is to average the GRAPE parameters into different scene types using the Stevens Zones (personal communication with the Remote Sensing Group (RSG) group at the Rutherford Appleton Laboratory (RAL)). The Stevens Zones were developed by Alan Stevens (RAL) and are based on the Normalised Difference Vegetation Index (NDVI). The NDVI uses the relationship between the red and near infrared reflectances to determine the vegetation type. The basic theory behind this is that plants absorb radiation in the Photosynthetically Active Radiation (PAR), spectral range to perform photosynthesis. Therefore, live green plants appear dark in the PAR and bright in the near-infrared. Figure 4.28 shows how the plants under different conditions absorb and reflect light differently. A healthy plant with plenty of water will have a higher NDVI index whereas a plant that has less water or too much sunlight, i.e. in the desert, will have a lower NDVI index. The Stevens Zones uses the NDVI index but takes it one step further by including clouds and ice, which can be identified in a similar way as they appear bright in the red and visible part of the spectrum, and dark in the near infrared. The Stevens Zones are found by looking at the correlation between the following two ratios
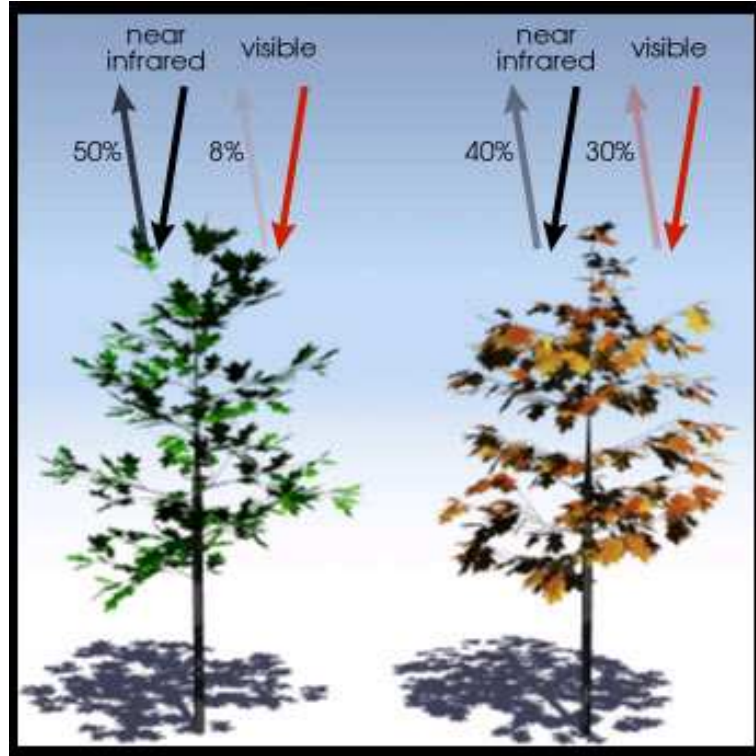
Figure 4.28: *An illustration of how plants absorb and reflect light under different conditions. The healthy plants on the left reflects 50% near infrared and 8% visible light in contrast to the unhealthy plant which reflects only 40% near infrared and 30% visible light. This difference can be used to determine the type of vegetation that is observed. This image is taken from the website http: //earthobservatory.nasa.gov.*

$$\frac{V87 - V67}{V87 + V67}, \qquad \frac{V67 - V55}{V67 + V55}, \tag{4.40}$$

where $V87$ is the reflectance at 870 nm, $V67$ is the reflectance at 670 nm and $V55$ is the

reflectance at 550 nm. This correlation can be used to determine the zone type. An example region is shown in figure 4.29.
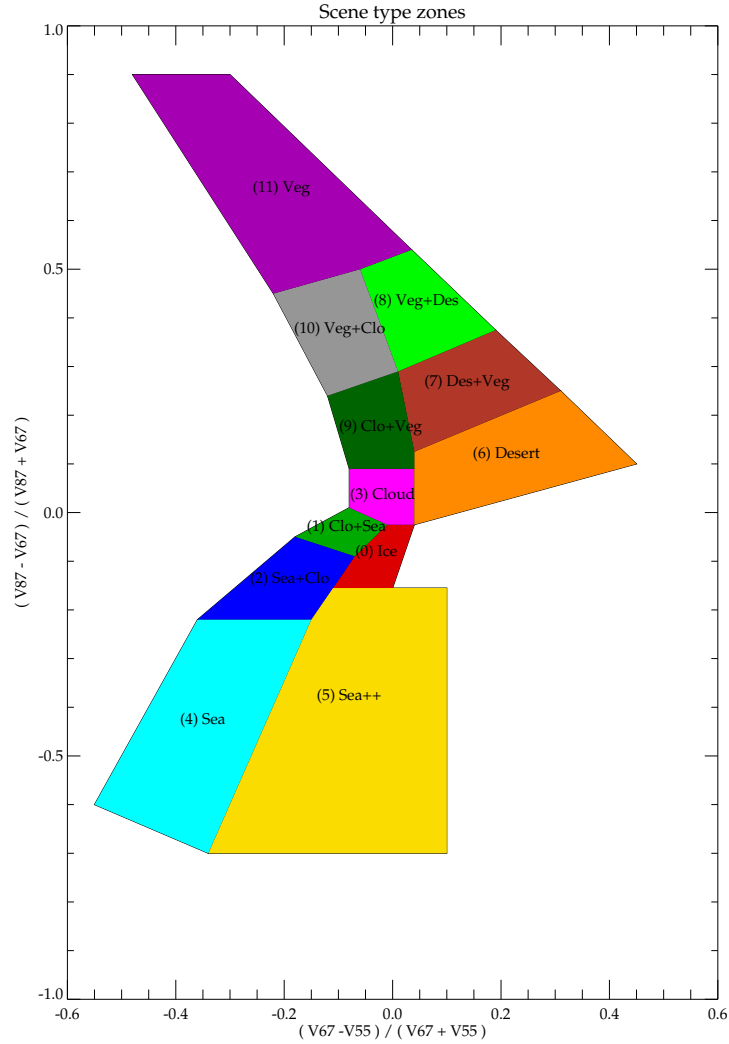
Figure 4.29: *An example of the Alan Stevens Zones. The x–axis is the ratio of the radiance in the 670 nm and 550 nm wavelength bands and the y–axis is the ratio of the radiance in the 870 nm and the 670 nm wavelength bands. The correlation of these two numbers provides information about the scene types.*

Figure 4.27(b) shows how the Stevens Zones can be used to average the GRAPE pixels onto a GOME pixel. The figure shows that in this particular pixel there are four Stevens Zones, cloud, cloud and vegetation, vegetation and cloud, and vegetation. Figure 4.27(b) shows a more detailed view of the GRAPE pixels within the GOME pixel. The same cloud that was observed in Figure 4.27(a) is observed in the lefthand side of the pixel, but the remain GRAPE pixels are divided differently, where the blue pixels have cloud and sea within them, and the yellow pixes are over vegetation and desert. It can already be seen that this is a better classification of the scene as more details on the surface types are highlighted.

The third method is to average the GRAPE parameters in twelve radiance bins. The GRAPE

Table 4.4: *Forward Model runs to test the different GRAPE averaging techniques*

Name	Run	Description
E	2 zone	Uses two zones a cloudy and cloud free
F	Stevens	Uses the Alan Stevens cloud zones to average GRAPE cloudy pixels
G	Radiance	Uses radiance averages to average the GRAPE cloudy pixels
H	All	Every GRAPE cloudy pixel is treated separately

radiances at 550 nm are used to determine the bins for the ATSR–2 pixel. Figure 4.27(c) shows how the radiance averaging zones can be used to average the GRAPE pixels onto a GOME pixel. This method is similar to the Stevens Zones as they both use radiances to determine a zone type. However, as we are using parameters in the 550 nm range in the retrieval model averaging over just these radiances could be more accurate. When comparing figures 4.27(b) and 4.27(c) it can be seen that there are some significant differences. For example the Stevens Zones captures more detail in the surface types and partially cloudy scenes, whereas the radiance averaging captures more variability in the cloudy pixels. Figure 4.27(c) shows here that the radiances are only different over to sea. This shows that the Stevens Zones has picked up more variability than the radiance averaging. However, this will not be the case for all pixels.

A number of tests were performed on each averaging technique to determine which was most suitable.

Sun–Normalised Radiance

Four different runs of the Forward Model were performed for a single orbit on July 8th 1997 to test the Forward Model's ability to recreate the observed sun–normalised radiance seen by the GOME instrument. The model was run at 430 nm which is the within the wavelength needed for an NO_2 retrieval. The four runs are summarised in Table 4.4.

The computationally expensive run **H** is necessary to determine the best possible result from the Forward Model. It would be unreasonable to run the model this way for a full run as it is very computationally expensive. The averaging technique that best fits the run **H** is the technique that will be used to average the GRAPE data in the final retrieval run.

Figure 4.30 shows a comparison of the sun-normalised radiance (modelled and observed) along the orbit with respect to latitude. The figure shows that the modelled sun-normalised radiance is larger than the observed GOME sun-normalised radiance for all averaging techniques. All figures show similar agreement with the Forward Model reproducing peaks and troughs well. However the true agreement is hard to see, therefore correlations have been plotted which better describe the relationship have been plotted in Figure 4.31.

Figure 4.31 shows the best fit between observed and modelled GOME sun-normalised radiance

at 430 nm for all four runs. Run **H** shows the best fit with a correlation coefficient of 0.939. Although the best fit after this is run **F** with a correlation coefficient of 0.782, this is almost the same of the other two runs with run **E** have a correlation coefficient of 0.75 and run **G** having a correlation coefficient of 0.78. From this figure it can be seen that using all pixels (run **H**) is the best possible approach to averaging the GRAPE parameters in the GOME pixel. However, as this is computationally prohibitively expensive another test can be performed to see how well the remaining 3 runs fit with the sun-normalised radiance derived in run **H**.

Figure 4.32 shows the correlation of modelled sun-normalised radiances in the 4 runs to the modelled sun-normalised radiance in run **H**. The correlation coefficients are all very similar, run **E** - 0.97, run **F** - 0.97, run **G** - 0.96 and run **H** - 1 show that run **E** which is the most computationally simple method will be as good as the other 3 methods when averaging the GRAPE cloud parameters onto the GOME pixel.
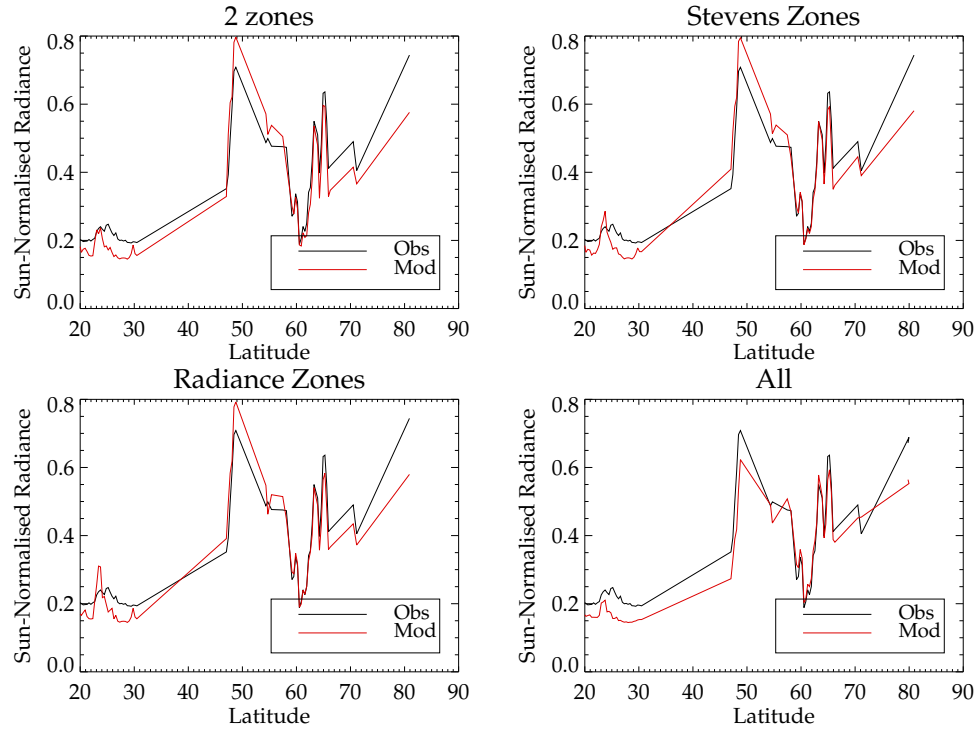
Figure 4.30: *The sun–normalised radiance along the orbit track for 1st orbit on July 8th 1997. The x–axis is the latitude of the measurements (in degrees) and the y–axis is the sun-normalised radiance. The red line is the model sun-normalised radiance at 430 nm, the black line is the sun-normalised radiance observed from GOME at 430 nm.*
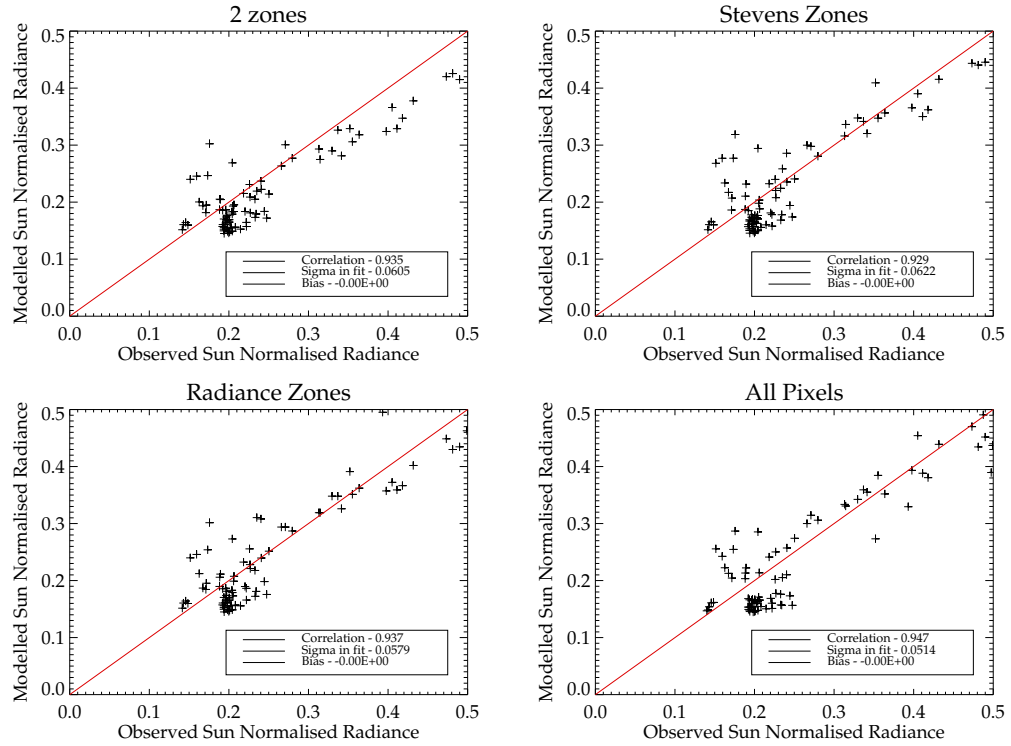
Figure 4.31: *Scatter plots of modelled sun-normalised radiance and GOME observed sun-normalised radiance for the 4 averaging techniques (a) 2 zones, (b) Stevens Zones, (c) Radiance zones and (d) All pixels. The red line shows the 1:1 line, i.e. the optimal result. The black line is the line of best fit with the legend showing the correlation coefficient, the line of best fit y-intercept, the gradient of the line of best fit, and the standard deviation of the points to the line of best fit.*
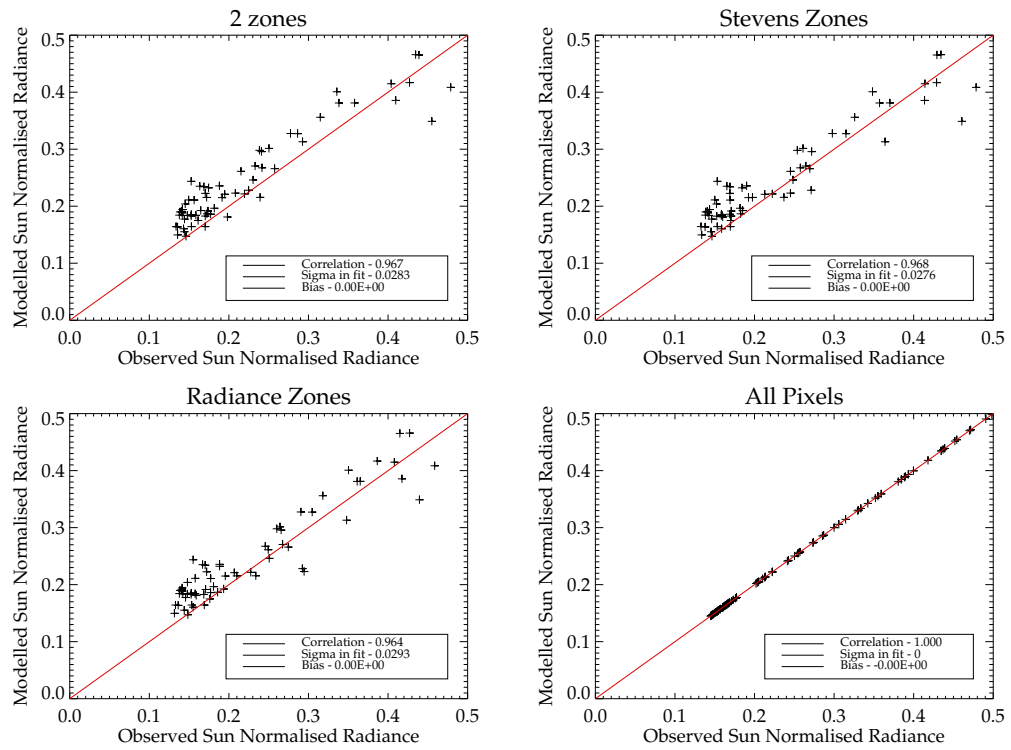
Figure 4.32: *As Figure 4.31 but when scattering modelled sun-normalised radiances to those modelled in run* **H**.

Photon Paths

Although the ability of the model to reproduce sun-normalised radiance is important it is also necessary to examine how the model reproduces the photon path lengths as these are used to calculate the tropospheric vertical column density. Figure 4.33 shows the tropospheric photon path length normalised by the geometric tropospheric photon path length (km). The geometric tropospheric photon path length is the photon path length if there is no scattering, i.e. if there is no atmosphere. Figure 4.33 shows that the tropospheric photon path lengths for each run are very similar, however run **H** produces slightly smaller values than runs **E**, **F** and **G**.
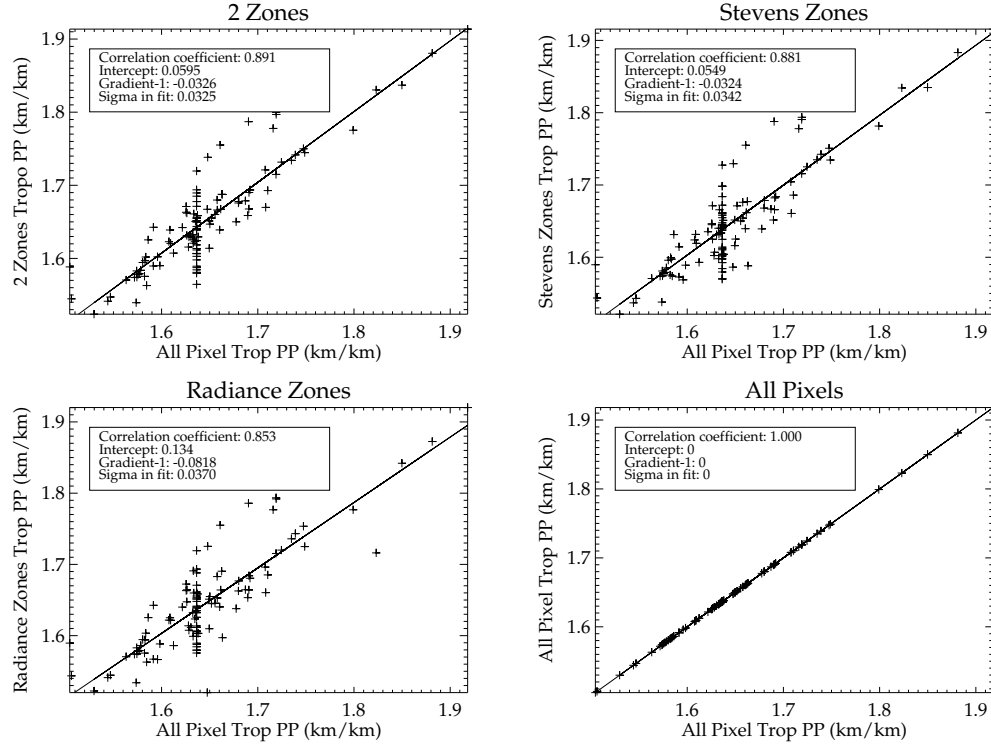
Figure 4.33: *The tropospheric photon path length normalised by the geometric tropospheric photon path length along the orbit track for the first orbit of GOME on July 7th 1997. The x–axis is the latitude and the y–axis is the average tropospheric photon path length per 1 km layer (cm). The black line is run* **E**, *the red line is run* **F**, *the green line is run* **G** *and the blue line is run* **H**.

The tropospheric trace gases studied here can have large concentrations in the boundary layer, it is therefore important to see how the model represents this part of the atmosphere. Figure 4.34 shows the boundary layer photon path lengths (per km) for each run.

In summary, after the three tests looking at sun–normalised radiance, tropospheric photon path length and boundary path photon path lengths the best fit averaging technique to run **H** is run **F**. As all the techniques fit well with run **H**, as the two zone approach is the least computationally expensive, this is the method that will be used to average the GRAPE data onto the GOME pixel in the final retrieval runs in Chapter 6.
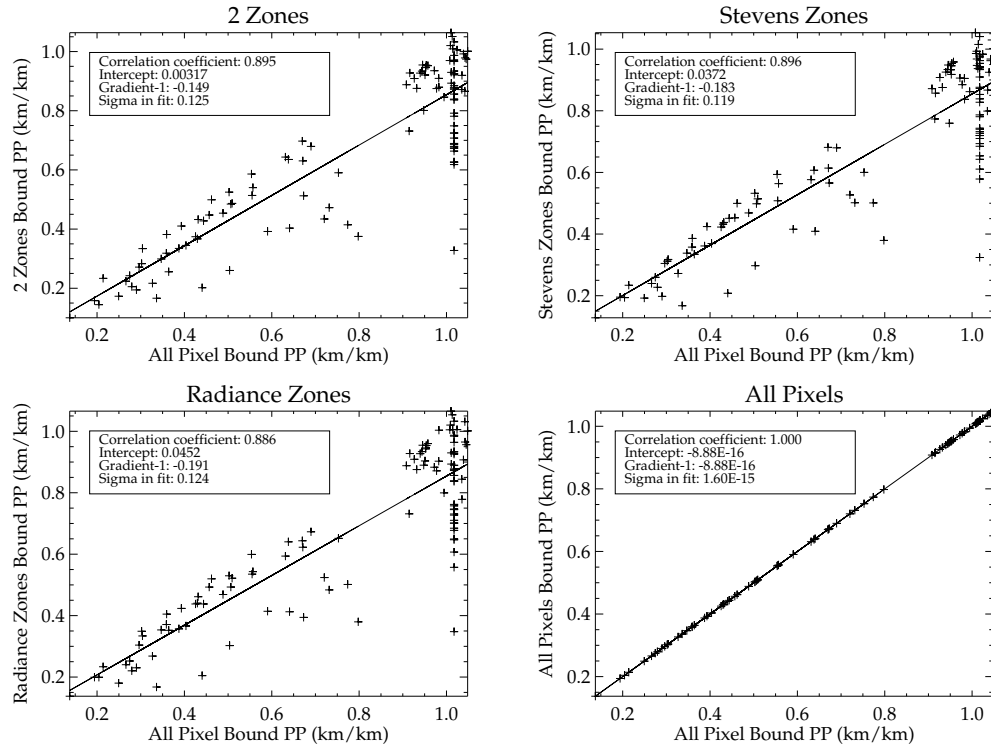
Figure 4.34: *The same as figure 4.33, but for the boundary layer photon path (per km).*

4.5.2 Additional Averaging Techniques

As was concluded in Section 4.5.1 the best averaging technique when considering all the pixels in the scene is to just break the scene into two zones (cloud and cloud free). However, as was shown in the Section 4.3, photon path lengths are strongly effected by cloud properties i.e. cloud optical depth and cloud top height (km), and therefore so will be sun-normalised radiance.

Although Figure 4.30 showed that run **H** has reasonably good agreement between observed and modelled sun-normalised radiance there were still a number of discrepancies between the radiances. One reason for this could be the use of every GRAPE pixel within the GOME pixel to gather as much cloud and aerosol information as possible. In reality for pixels with high cloud optical depth, cloud fraction or cloud top height the retrieval is likely to have been high in error and therefore sun-normalised radiance will be too. Another reason for this could be because not every cloud fits the optical model assumed in the optimal estimation retrieval used in GRAPE, i.e. the use of an infinitely thin cloud. It is possible to analyse how well the GRAPE modelled atmosphere fits with the measurement by looking at the final measurement cost. In an ideal world the cost from the retrieval would be 5, equal to the number of retrieval parameters, however, a cost of 20 would suggest that the retrieval is converged and represented the scene well[3]. The maximum measurement cost can be extremely high up to 300 if the retrieval struggles to converge. The high measurement costs have been associated with multiple cloud layers. As has been seen in Section 4.3.4 if a single layer is assumed the scene will be hard to represent if the true state is a multiple cloud and therefore this would explain the high cost.

Figure 4.35 shows the correlations of observed versus modelled sun-normalised radiance of a number of averaging techniques relating to the removal of pixels that are high in error. The plots in the right column have a maximum cost of 30 and fit the show overall that this results in a better fit with observations and modelled results. The correlation also improves down the rows as more pixels are removed that will be high in error due to high cloud fractions or cloud optical depths. Overall these plots show that the GOME observed modelled sun–normalised radiance can be reasonably well produced if pixels that are high in error are removed. The impact of this upon the retrieval scheme will be discussed in Section 5.7.

Figure 4.36 is the same as Figure 4.35 but for BrO (retrieval at 350 nm). Firstly the plots show that sun–normalised radiance modelled at this wavelength fits better with observed GOME sun–normalised radiance. Secondly the removal of pixels with high measurement cost does not

---

[3]Personal Communication Caroline Poulsen RAL
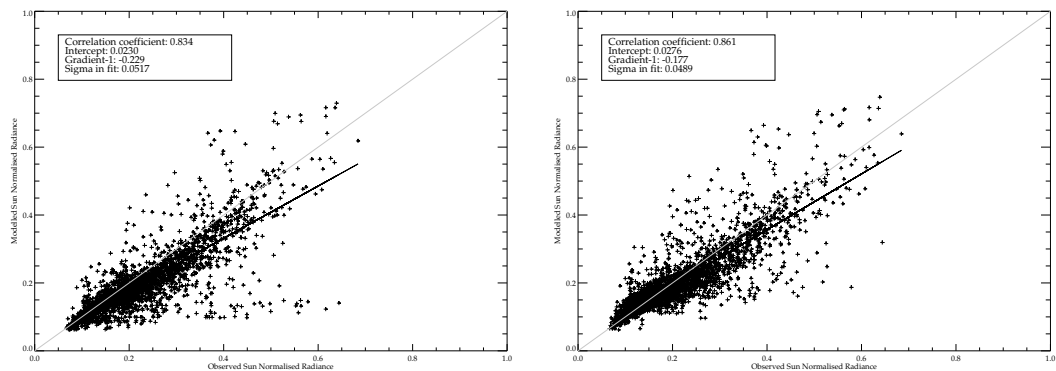
Figure 4.35: *Scatter plots of observed versus modelled sun normalised for NO₂ retrieval for many GOME orbits (a) for all pixels, (b) only for pixels with a measurement cost less than 30.*

152
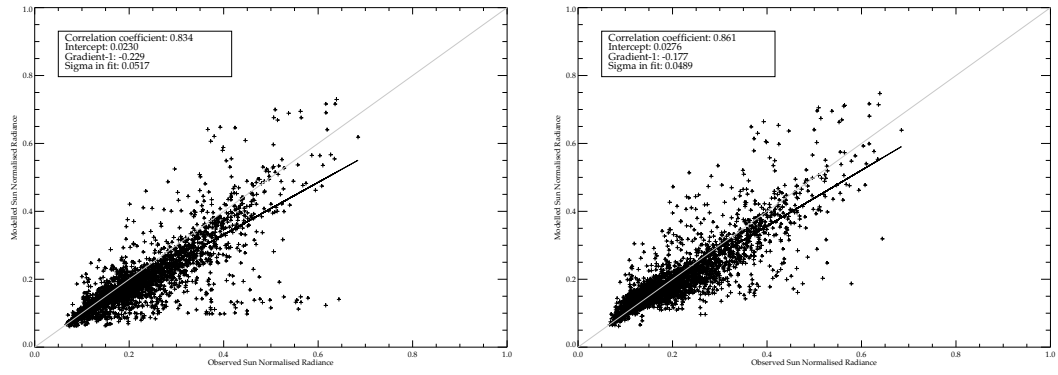
Figure 4.36: *As Figure 4.35 but for BrO.*

improve this except when considering the cloud fraction of 50% or less and cloud optical depth of 20 or less. The impact of this upon the retrieval scheme will be discussed in Section 5.7.

4.6 Summary

This chapter has described the theory needed to understand an inverse problem and has outlined the Retrieval Model developed in this work. This new Retrieval Model differs substantially from previous studies (e.g. Wagner and Platt (1998), Richter and Burrows (2001) and Van der A et al. (2006)) due to the following

- The stratospheric component of the total vertical column is calculated using the SLIMCAT global 3-D CTM, constrained further by chemical data assimilation.

- The retrieval model uses a full information based algorithm and retrieves two parameters, the tropospheric scaling factor and the surface Lambertain Equivalent Reflectance (LER).

- The cloud and aerosol parameters are taken from ATSR-2 measurements averaged onto the GOME pixel.

The Forward Model tests show that the radiative transfer model (GOMETRAN) is capable of producing photon path lengths which are expected from the theory and reproducing expected averaging kernels (Eskes and Boersma, 2003). In the presence of clouds it is not possible to retrieve information in the boundary layer for optically thick clouds, or for clouds with high cloud top heights. However, at high surface LER some information can be retrieved with optically thick clouds (however, this is an unlikely scene representation as the surface has to be ice). The tests have shown that the modelled cloud vertical profile makes little difference to the photon path lengths unless there are multiple cloud layers present, when there is a significant difference. Cloud particle size has been shown to not have a significant affect on the tropospheric photon paths.

The tests on aerosol parameters show that the photon path lengths are not significantly altered with changing particle size, type, or optical depth within the tested ranges. In pollution events the values of the aerosol parameters will became comparable to cloud and the affects of aerosol will become similar to those of cloud. However, it will be difficult for the GRAPE retrieval to differentiate between cloud and aerosol in pollution events, although there will be high measurement costs associated with such retrievals and these can be removed from the retrieval if necessary (Arnold, 2007).

The Forward Model tests are not only important for checking the validity of the radiative transfer scheme, but they have highlighted the parameters that are most likely to cause significant errors in the retrieval. These have been shown to be cloud parameters, surface LER, and the stratospheric column amount. These were shown to be significant sources of error in previous studies also (e.g. Boersma et al., 2004). Chapter 5 will further investigate how errors in these parameters will affect the retrieval.

When looking at how to average the GRAPE cloud and aerosol parameters onto the GOME pixel it was noted that ATSR–2 cloud radiances can be significantly higher than GOME radiances. Although a number of GRAPE cloud averaging techniques have been tested the simple two zone

averaging (cloudy and non cloudy) is the least computationally expensive and produces similar results to the more complicated, time consuming methods. The use of a maximum measurement cost in the averaging of the cloud parameters does improve the representation of the GOME radiances. Therefore, the Retrieval Model will be run with a maximum measurement cost of 30.

This chapter has described how the Forward Model parameters will affect the retrieved tropospheric VCD by analysing photon path lengths and determining averaging kernels. Therefore Table 4.5 is a new table that summarises the important sources of error in the retrieval. It can be seen in the table that the previous sources of error remain and now errors due to clouds and profile shape have been added.

Species	Parameter	Error
NO_2	SCD error	5–9×10^{15} mol. cm^{-2}
BrO	SCD error	1.5–1.8×10^{13} mol. cm^{-2}
NO_2	Previous study VCD_{trop} error	0.5–10×10^{15} mol. cm^{-2}
BrO	Previous study VCD_{trop} error	na
NO_2	Scalar approximation	upper limit 10% of VCD_{trop}
BrO	Scalar approximation	10% of VCD_{trop}
NO_2 / BrO	Error due to cloud and aerosol optical models	to be quantified in Chapter 4
NO_2 / BrO	Plane–parallel assumption	only significant at high sza
NO_2	Stratospheric Col.	0.5–7% (strat. col.)
BrO	Stratospheric Col.	10% (strat. col.)
NO_2 / BrO	Cloud and Aerosol Parameters	Derived in Chapter 5
NO_2 / BrO	Multi-cloud Layers	Large error therefore multi-layered clouds will be removed where it is possible by considering the cost of the GRAPE retrieval
NO_2 / BrO	Concentration Profile Shape	Large errors, but difficult to quantify as true profile will be unknown

Table 4.5: *The errors associated with assumptions or parameters that will be used in the new retrieval model developed in this thesis for NO_2 and BrO.*

The quantitative impact of the new errors (clouds, aerosol and profile shape) were determined in this chapter and summarised in Table 4.3. It was also shown, in this table, that the stratospheric column error will have a large impact on the retrieved tropospheric VCD. This was determined from the stratospheric profile averaging kernel. Further quantification of the errors will be described in Chapter 5.

Chapter 5

Estimating Retrieval Errors

5.1   Introduction

This chapter discusses the model factors that affect the retrieval of tropospheric trace gas columns from the GOME instrument. It is important when developing an atmospheric retrieval algorithm to understand the sources of error and to quantify their importance in the retrieval process. In order to achieve this a number of test scenarios have been set-up that represent geophysical conditions, both temporally and spatially, for the GOME instrument. These scenarios are then used to test this new retrieval scheme.

There are two types of error considered in this chapter

1. The estimated standard deviation (ESD) – the uncertainty in the retrieval resulting from the measurement error quantified in the retrieval via the measurement covariance.

2. Error due to Forward Model parameters (not including state or measurement covariance) – which are determined by running a non-linear retrieval in which a measurement is simulated using a defined 'true state' and the retrieval is performed using a second 'retrieved state'.

ESD is assumed to be purely random (as this is an underlying assumption of optimal estimation or least squares fitting, Rodgers, 2000). However, this is not entirely true within the context of this thesis[1]. The Forward Model parameter errors can be both random and systematic. Random errors are always present in a measurement and are unpredictable. An example of random error could be when taking a measurement of the same quantity (i.e. the length of a piece of string). If the measurement is taken 20 times the mean of the results would be the measurement and the standard deviation of the measurements would be the random error (due to human error). Random error can be large however, it can be reduced by averaging to

---

[1]Errors that are specified in $\mathbf{S_y}$ in this work are not only due to noise, but also chosen to cover some systematic errors. See Chapter 1.

form e.g. monthly, spatially averaged means. Systematic error can also be unpredictable but typically is constant over some time or spatial scale. An example of systematic error would be a calibration error, where every measurement is offset by 10%. Systematic error cannot be reduced by averaging. The random and systematic errors will be discussed in more detail in the following sections.

The errors derived in the following chapter will be used to derive the overall systematic and random error on the retrieved tropospheric columns which are shown in Chapter 6. This will be achieved by scaling the errors that are determined in this chapter. This relies on the relationships between each parameter being well understood. For example, each error will be determined for a number of atmospheric parameters (cloud fraction, cloud top height, cloud optical depth, solar zenith angle and surface LER). Therefore, the interpolation or scaling of the errors to different values of these atmospheric parameters needs to be understood to allow for the correct error to be assigned to a specific pixel in the final retrieval. For small changes in the atmospheric parameters this relationship will be assumed to be linear. However, as will be shown in this chapter, this is not the case for all atmospheric parameters and therefore a different relationship will need to be applied.

Section 5.2 describes the different test scenarios and how they have been set-up while Section 5.3 discusses the reasons behind the decision to retrieve surface Lambertian Equivalent Reflectance (LER) along with tropospheric scaling factor (x_{sf}). Section 5.4 looks at the ESD (error due to random noise) in the retrieval scheme and Section 5.5 investigates the errors associated with not retrieving cloud when it is present. Section 5.6 discusses the impact of cloud parameter errors on the retrieved tropospheric VCD, both random and systematic, and Section 5.7 investigates the impact of errors in the GRAPE aerosol parameters on the retrieved tropospheric VCD. Section 5.8 investigates how errors in the stratospheric column impact the final retrieved tropospheric VCD and Section 5.9 discusses the other sources of error in the retrieval.

5.2   Scenario Setup

Twenty four scenarios were set-up to represent different global locations, seasons and land types based on those used in Kerridge et al. (2004). Overall the scenarios covered:

- 4 seasons (January, April, July, October)
- 3 latitudes (5°N, 55°N, and 75°N)
- 2 surface LERs (0.05 - sea/land and 0.8 - ice)

However, due to the large number of scenarios these were reduced down to 6 scenarios for this work. These were chosen to cover different viewing angles of the satellite and seasons (April 5°N, July 55°N and April 75°N) and covering two surface LER's (0.05 and 0.8). Each scenario used realistic trace gas and temperature/pressure profiles which were taken from the

SLIMCAT/TOMCAT model runs used to initialise the retrieval model. The stratospheric trace gas and the temperature/pressure profiles were taken from a 1997 model run of SLIMCAT with chemical data assimilation (Chapter 3) and the tropospheric profiles were monthly means taken from a 2001 run of TOMCAT. Figure 5.1 shows the atmospheric trace gas profiles used for the different scenarios. The viewing geometry for the scenarios is taken from Kerridge et al. (2004) and is summarised in Table 5.1.
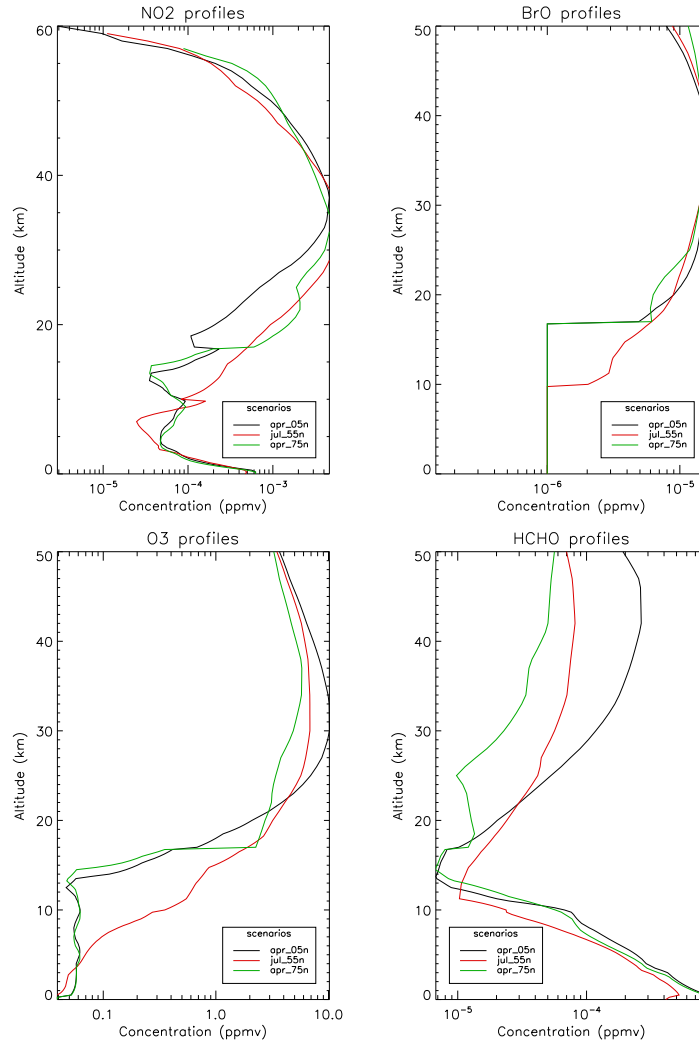
Figure 5.1: *The atmospheric profiles (ppmv) of the different tropospheric trace gases (NO$_2$, BrO, O$_3$ and HCHO) that are used for the different test scenarios. The legend shows the scenario that each profile is used for.*

Table 5.1: *A table summarising the tropospheric column amounts (1×10^{15} molecules cm^{-2} for NO_2 and 1×10^{12} molecules cm^{-2} for BrO), errors in the slant columns (1×10^{15} molecules cm^{-2} for NO_2 and 1×10^{13} molecules cm^{-2} for BrO) and the sun–normalised radiance for each scenario for each surface LER.*

Scenario	Solar Zenith Angle (deg)	Surface LER	Species NO_2/BrO	Tropospheric Column	Slant Column Error	Sun–Norm. Radiance
April 5N	36.7	0.05	NO_2	2.1	0.8	0.120
April 5N	36.7	0.05	BrO	1.3	1.0	0.209
April 5N	36.7	0.8	NO_2	2.1	0.3	0.760
April 5N	36.7	0.8	BrO	1.3	8.0	0.760
July 55N	39.0	0.05	NO_2	1.9	0.4	0.085
July 55N	39.0	0.05	BrO	1.3	12.8	0.151
July 55N	39.0	0.8	NO_2	1.9	0.2	0.469
July 55N	39.0	0.8	BrO	1.3	7.8	0.461
April 75N	65.4	0.05	NO_2	2.2	0.5	0.065
April 75N	65.4	0.05	BrO	1.4	0.1	0.113
April 75N	65.4	0.8	NO_2	2.2	0.3	0.317
April 75N	65.4	0.8	BrO	1.4	1.4	0.309

The measurements (slant column density and sun–normalised radiance) used in this testing process where calculated by running only one iteration of the Forward Model with realistic atmospheric parameters and viewing angles. Measurement errors are an important part of any retrieval scheme. Errors in the slant columns used in the final retrieval model are taken from those reported along with the retrieved slant column (from GDP or SAO). These are both systematic and random, as explained in Chapter 1. For the scenarios used in this sensitivity study, slant column errors are taken from Kerridge et al. (2004), and are derived from a noise model (Van der A, 1997). Therefore, it is assumed that these errors are purely random. This is chosen because the slant column error is mostly due to photon noise (i.e. is well represented by the noise model). Systematic errors in the slant column density can be ignored as they will be removed by the diffuser plate correction (see Chapters 1 and 6). The slant column density random errors are for cloud–free scenes (taken from, Kerridge et al., 2004) and are summarised in Table 5.1.

The other measurement used to initialise the model is sun-normalised radiance (from which surface LER is retrieved). The reported error in sun–normalised radiance is very small (usually less than 1%). 1% is chosen to represent the estimated error in the forward model (e.g. neglection of the ring effect) and errors in radiometric calibration of instrument (Siddans, 2003).

5.3   Surface Lambertian Equivalent Reflectance Retrieval

Previous studies have shown that errors in surface LER can cause discrepancies in the retrieval of atmospheric constituents (e.g. Boersma et al., 2004). This new retrieval scheme derives two pieces of information; the tropospheric scaling factor (x_{sf}) and the surface LER. This will help lower the overall error on the retrieval. This section argues why it is necessary to retrieve two pieces of information in the retrieval scheme and discusses the impact of surface LER errors if it were not derived alongside x_{sf}.

In order to prove that it is necessary to retrieve surface LER along with x_{sf} the retrieval scheme was run assuming surface LER was perfectly known. This forced the retrieval model to treat the initial guess at the surface LER as the final answer and to only perform the retrieval on the slant column therefore only calculating x_{sf}. The retrieval scheme was run for the 3 scenarios to simulate a range of solar zenith angles and seasons. The measurement was simulated with a 'true' value of the surface LER of 0.1 and then a retrieval was performed for a 'retrieved' range of surface LERs (from 0 − 1), to calculate the fractional error in the retrieved tropospheric VCD. Figure 5.2 shows the ratio of the 'true' tropospheric VCD to the 'retrieved' tropospheric VCD. The figure shows that where 'true' and 'retrieved' are equal (i.e. a value of 0.1) the ratio is 1. As noted in previous studies (Boersma et al., 2004), if the surface LER is in error, there is a large impact on the 'retrieved' tropospheric VCD, for all scenarios. Below a LER of 0.1 the error increases rapidly, i.e. an error in LER of ±50% would result in a fractional error in the tropospheric VCD of -0.08–+0.15 (-8% –+15%) for the April 75°N case and -0.1–+0.2 (-10% –+20%) for the April 5°N and July 55°N. It is important to note that in this work radiation is in the UV/VIS spectrum and surface LER values are typically around 0.1 and below, unless over snow and ice. Figure 5.2 shows that errors for this range of surface LER values are higher.

For pixels where there is cloud and aerosol, the surface LER fitting will also accommodate other errors, namely due to cloud and aerosol changing the reflectance of the scene.

5.4   Estimated Standard Deviation

The Estimated Standard Deviation (ESD) is the estimated error on the retrieval due to the assumed measurement error and is calculated using Equation 4.26. The ESD can also be thought of as the minimum error on the retrieval, i.e. if every other parameter in the retrieval model was perfect the ESD would be the random error on the retrieval product.

There are a few important points to note in this section and chapter. Firstly, in this retrieval scheme there are two measurements, slant column density and sun–normalised radiance. In this chapter, only the slant column density is discussed, because errors in the observed radiance are expected to be sufficiently characterised by the measurement covariance and hence the ESD. Hereafter, when referring to the measurement it should be assumed that this refers to the slant column density unless stated. Secondly, the tropospheric column is retrieved below 8 km in all the following tests. This allows for the highest cloud top to be at the top of the troposphere and
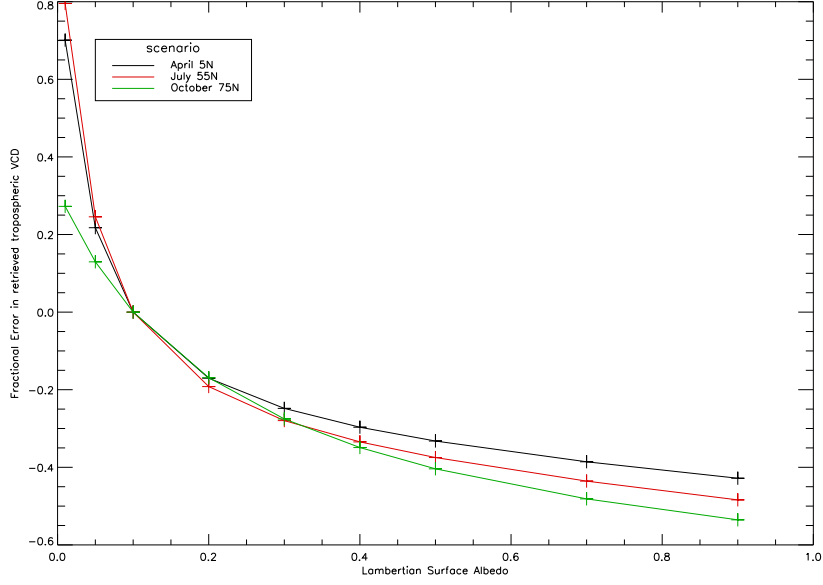
Figure 5.2: *The ratio of the 'true' tropospheric VCD (with surface LER of 0.1) to the 'retrieved' tropospheric VCD (at varying surface LER). The retrieval model was run for 3 scenarios April 5° N, July 55° N and April 75° N and a full tropospheric scaling factor was retrieved. The + marks show the different surface LERs that were sampled.*

therefore a case where the troposphere is obscured totally by the cloud can be studied. Thirdly, the following tests use the same tropospheric species profile in both the 'true' and 'retrieved' states. In reality there will be errors in this profile. However, this will be difficult to quantify as the 'true' profile will always be unknown.

The ESD is calculated from weighting functions computed assuming the state vector is equal to the truth. The measurement errors for the slant column were taken from Kerridge et al. (2004) and the sun–normalised radiance errors were assumed to be $1\%^2$. The slant column error in Kerridge et al. (2004) is for cloud–free scenes. This can be scaled to other levels of radiance (e.g. produced by cloud in the scene), assuming errors are limited by photon noise on the signal, as follows:

$$\epsilon_c = \epsilon_{cf} \sqrt{\frac{r_{cf}}{r_c}} \tag{5.1}$$

where ϵ_c is the new ESD, ϵ_{cf} is the original ESD, r_{cf} is the cloud-free sun–normalised radiance and r_c is the cloudy sun–normalised radiance.

In Section 4.4, the profile averaging kernels were assessed for the stratosphere, the troposphere and the boundary layer. In this section the scaling factor is applied to either the troposphere (the bottom 8 km of the atmosphere) and just the boundary layer (bottom 2 km of the atmosphere). For some tropospheric trace gases the majority of the species resides in the boundary layer. It may therefore be more accurate to only scale the boundary layer rather than the whole

---

[2]A discussion of the source of the slant column errors can be found in Chapter 1

troposphere[3]. This would only be applicable at locations where there are enhancements in the column due to surface emissions, e.g. for NO_2 over cities (anthropogenic pollution) or for BrO over frost flowers (source of BrO in the boundary layer) and, in practice, this is hard to diagnose.

It should be noted that errors stated here are for one retrieval, whereas for monthly means the errors will decrease to:

$$\epsilon_{mm} = \sqrt{\frac{\epsilon^2}{n_{ret}}} \qquad (5.2)$$

where ϵ_{mm} is the new ESD, ϵ is the ESD for each retrieval (assuming the error is equal for each pixel) and n_{ret} is the number of retrievals. With the combined GOME and ATSR–2 instruments, daily coverage is achieved roughly every 9 days. As stated earlier, Equation 5.2 only applies for random error.

Figure 5.3 shows the ESD in the tropospheric NO_2 VCD for the 3 scenarios at surface LER's of 0.05 and 0.8. In this (and the subsequent figures) there are 6 plots. The top two being for April 5°N at 0.05 and 0.8, the middle two being for July 55°N and the bottom two being for April 75°N. Each plot then has two panels. The first is for a cloud with an optical depth of 30 and shows how errors change for varying cloud fraction and cloud top height. The second panel is for a cloud fraction of 1 and shows how errors change with varying cloud optical depth and cloud top height. Firstly (from Figure 5.3), the figure shows that the fractional error in tropospheric VCD increases with increasing cloud top height, cloud optical depth and cloud fraction for all values of surface LER. This is what is expected as when cloud is present in the scene the retrieval is more complicated and therefore the error will increase. Secondly, the fractional error in the retrieved tropospheric VCD decreases with increasing surface LER. This is because the measurement error will be smaller at high surface LER's as the signal received by the satellite will be larger at high surface LER's. Also, the photon path lengths in the low troposphere are longer at higher surface LER.

Thirdly, the fractional error in the tropospheric VCD increases with increasing solar zenith angle[4]. This relationship is due to the fact that with increasing solar zenith angle the photon path lengths in the lower troposphere become shorter relative to the photon path lengths in the upper troposphere, i.e. the retrieval becomes less sensitive to the lower troposphere with increasing solar zenith angle (see Figure 4.22).

Finally, for clouds that are in the boundary layer and have large cloud optical depths, the fractional error in tropospheric VCD decreases. Looking at Figure 4.12 it can be seen that the photon path lengths increase above a cloud with an optical depth of greater than 3. Also the

---

[3]NO_2 and BrO are often produced at the surface. As their lifetimes are short they will react quickly and therefore, only low concentrations will escape the boundary layer. However, both gases are found outside the boundary layer through cycling, or production outside the boundary layer. It is difficult to estimate where the profile will be purely in the boundary.

[4]Scenario April 5°N solar zenith angle 38°, July 55° solar zenith angle 39° and April 75° solar zenith angle 65°.

NO$_2$ profile (see Figure 5.1) shows that there is a large amount of NO$_2$ in the boundary (i.e. near the cloud). As the tropospheric VCD is the sum of photon path lengths multiplied by the concentration of the species in a layer, this decrease in error is due to the tropospheric VCD being artificially enhanced by the increase in the photon path lengths. There may also be a slight increase in the signal that the satellite observes due to the increase in photon path lengths above the cloud. Although this is good because the ESD is slightly decreased, the final retrieval is actually biased by this affect and is in fact not sensitive to the boundary layer. Therefore, the tropospheric VCD that is retrieved is highly dependent on the original profile that was used to initiate the retrieval model. The values of ESD vary from 10–60% for low solar zenith angles except for extremely high cloud top heights, cloud fractions and cloud optical depths (cloud top heights higher than 6 km) where it can grow to over 100%. For higher surface LER the ESD is lower varying from 5–40%.

In Appendix A, Figure A.1 shows the ESD in the tropospheric VCD when considering boundary layer scaling rather than tropospheric scaling. The figure shows that the error behaves in the same way as in Figure 5.3 but the errors are larger.

Figure 5.4 is similar to Figure 5.3 but for BrO. The figure shows the same pattern in the errors but the values are different due to the different wavelength and the different profile shape. This is also due to the different amounts of scattering. For example, there is more Rayleigh scattering in the BrO fitting window and therefore the photon path lengths are longer. The ESD for BrO, for low surface LER, is generally around 60% with values of greater than 100% for high cloud top height, large cloud optical depth and high cloud fraction. For cloud in the boundary layer the ESD becomes small at around 10% for high optical depths. For high surface LER, the ESD for BrO is between 10–40%. In Appendix A Figure A.2 shows the ESD in the tropospheric VCD when considering boundary layer scaling rather than tropospheric scaling for BrO. The errors are similar to those shown in Figure 5.4 but the values are different due to the different wavelength and profile shape.
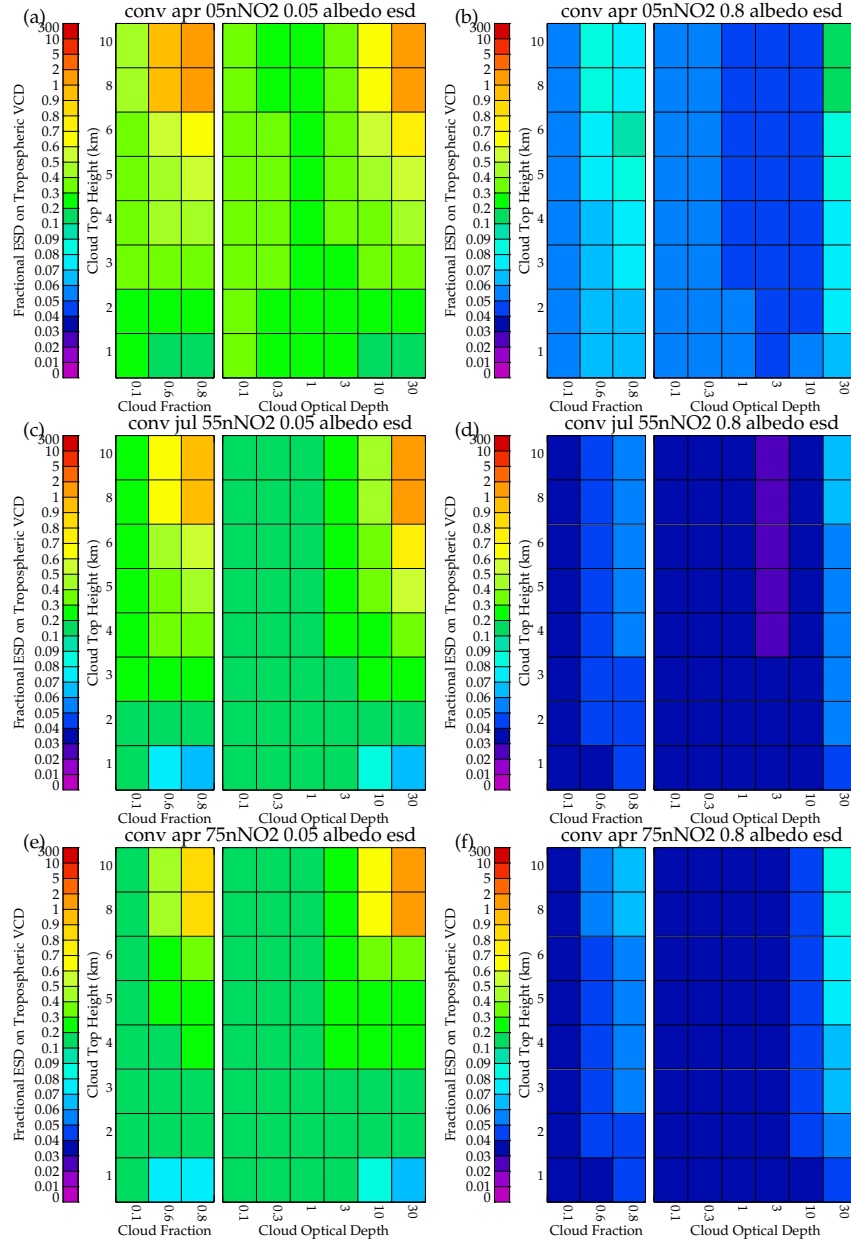
Figure 5.3: *The ESD in the total tropospheric NO_2 VCD (molecules cm^{-2}) (a) for April $5°N$ at a surface LER of 0.05, (b) for April $5°N$ at a surface LER of 0.8, (c) for July $55°N$ at a surface of 0.05, (d) for July $55°N$ at a surface LER of 0.8, (e) for April $75°$ at a surface LER of 0.05 and (f) for April $75°$ at a surface LER 0.8. In each panel there are two plots, the left plot shows the error when cloud fraction and cloud height are varied, for a cloud with an optical depth of 30. The right plot shows the error when cloud optical depth and cloud height are varied for a cloud fraction of 1.*
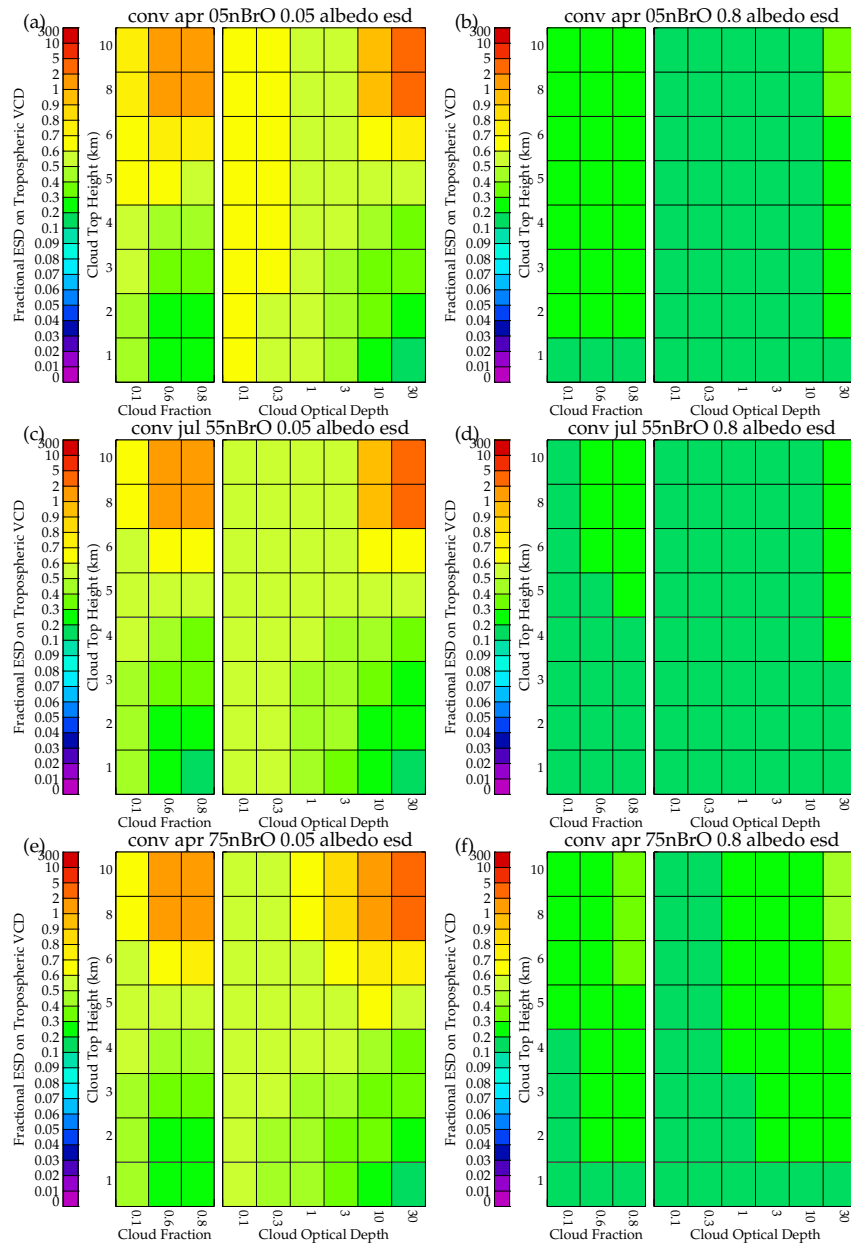
Figure 5.4: *As Figure 5.3 but for BrO.*

5.5 Modelling Cloud

As noted throughout this thesis, one of the most important sources of error in a UV–VIS atmospheric trace gas retrieval is cloud. This is particularly important for the GOME instrument as the ground pixel is large (40 km×320 km), meaning there are not many cloud–free ground pixels. Some retrieval schemes ignore cloudy scenes (Richter et al., 1998) meaning their coverage is limited, while others (Chance, 1998) treat scenes with low cloud fractions (less than 20%) as clear sky, which biases the results, as cloud causes enhancements in the sun–normalised radiance observed by the instrument and therefore changes the differential structure which affects the slant column.

Due to the discrepancies caused by cloud it is necessary to investigate the impact of errors that are associated with not modelling clouds when they are present. To carry this out a new test was run, where the 'true' state has a cloud and the 'retrieved' state is cloud–free. The fractional error in the resulting tropospheric VCDs (V_t) is:

$$\Delta V_t = \frac{V_{t_m} - V_{t_0}}{V_{t_0}} \qquad (5.3)$$

where ΔV_t is the change in the V_t (or the error in V_t), V_{t_0} is the 'true' or initial V_t and V_{t_m} is the 'retrieved' or modelled V_t.

Figure 5.5 shows the error (assuming cloud-free when cloud is present) in tropospheric NO_2 VCD for different scenarios and surface LER's. Firstly, the fractional error in the tropospheric VCD increases with increasing cloud optical depth, cloud top height and cloud fraction. The error varies from 5% at low cloud optical depth, cloud top height and cloud fraction to up to 100% at high cloud optical depth, high cloud top height and for large cloud fraction. The trend in the errors due to not modelling cloud are similar to the ESD, but the values are larger. These larger values means that the error can be improved by including cloud in the retrieval.

Figure 5.5 shows a number of other interesting features. Firstly, for high surface LER, and low solar zenith angles at high cloud top heights at around an optical depth of 3 and for cloud fraction of 1, the errors are positive whereas everywhere else they are negative. A positive error means that the 'retrieved' tropospheric column (cloud–free) is larger than the 'true' tropospheric column (cloudy). In most cases the cloud–free slant column will be lower as the surface LER is so much lower than that of the cloud top (which is around 0.8 Wang et al., 2008). However, at higher surface LER the surface has the same LER to the cloud and therefore if the cloud slightly reduces the signal received by the satellite the errors will be positive.

Secondly, there is an increase in the error in the tropospheric VCD for cloud optical depths of 3. As seen in Figure 4.12 a cloud optical depth of 3 is the key point where photon path length significantly increase above the cloud due to increased reflection. Thirdly, the errors although small, vary from positive to negative at low cloud top height and low optical depth. This is due to the same affect as in the first point, but as the NO_2 profile is largest in the bottom 2 km of

the atmosphere, this too will influence the value of the error.

Figure 5.6 shows the same results for BrO. Errors also show that the retrieval of BrO can also be improved by including cloud in the retrieval.
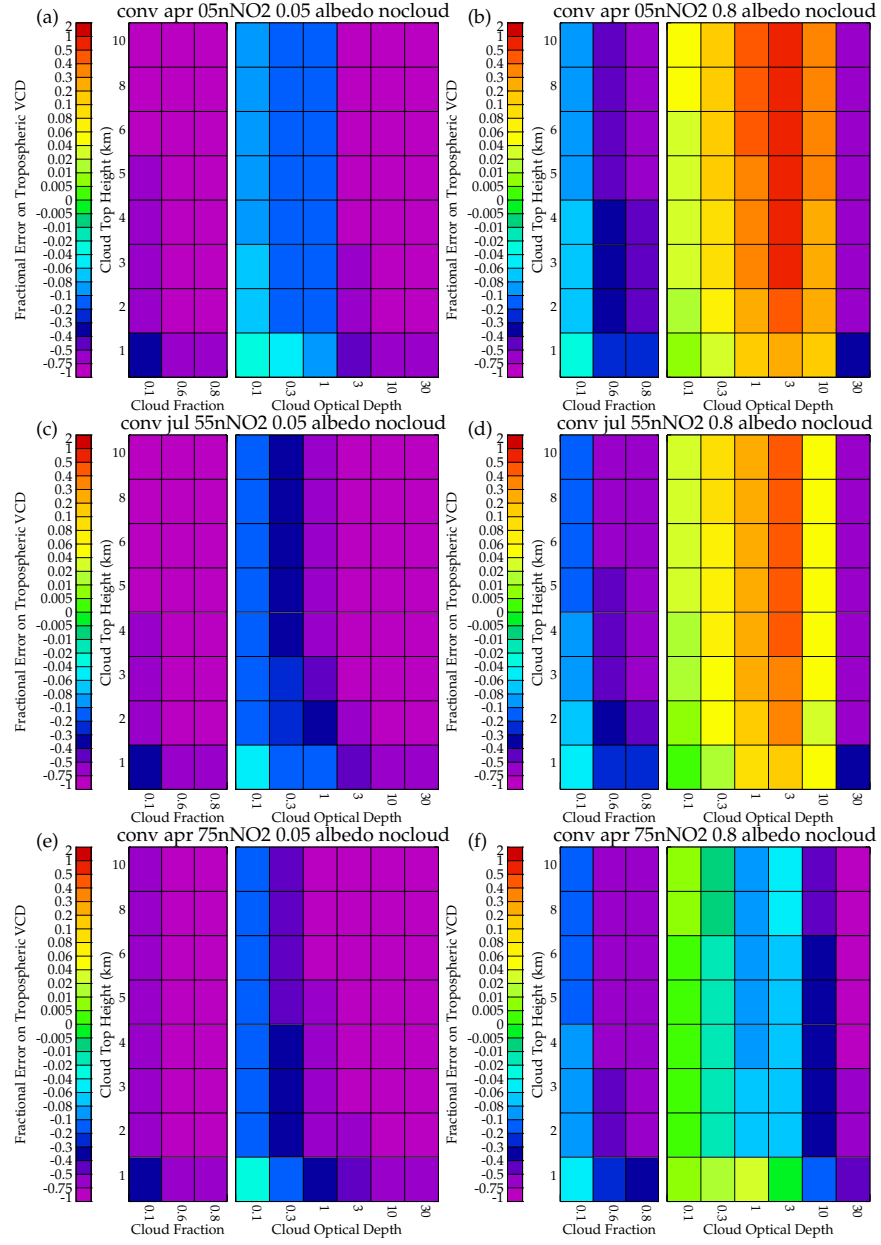
Figure 5.5: *The error in tropospheric NO$_2$ VCD (molecules cm^{-2}) when assuming a cloud–free scene when cloud is present (a) for April 5°N at a surface LER of 0.05, (b) for April 5°N at a surface LER of 0.8, (c) for July 55°N at a surface of 0.05, (d) for July 55°N at a surface LER of 0.8, (e) for April 75° at a surface LER of 0.05 and (f) for April 75° at a surface LER 0.8. The plots are in the same format as in Figure 5.3.*
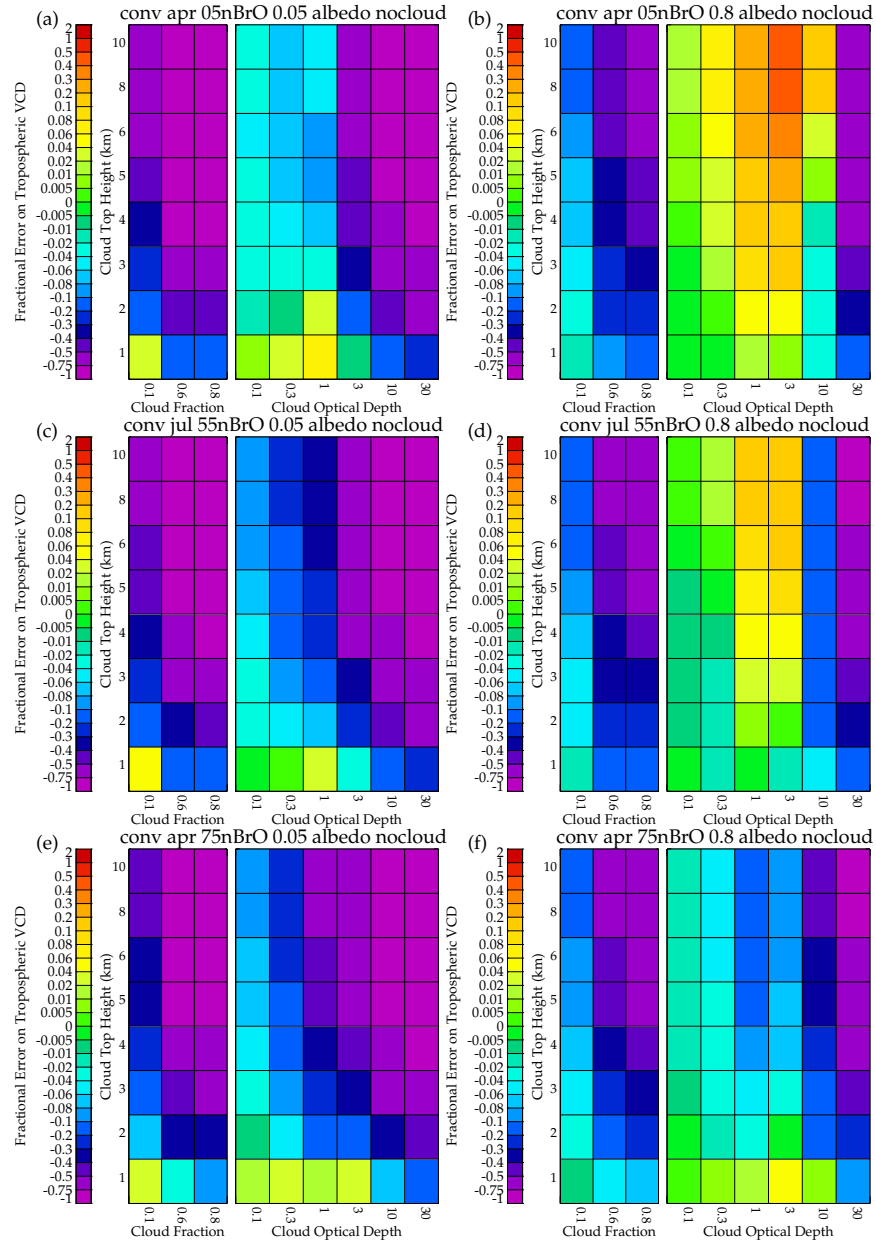
Figure 5.6: *As Figure 5.5 but for BrO.*

169

5.6 Impact of Errors in Cloud Parameters

5.6.1 Sensitivity to Cloud Parameter Errors

The following section shows how the retrieved tropospheric VCD responds to errors in the cloud parameters. The method for carrying out the following tests is described below:

- A measurement (slant column density and sun–normalised radiance), along with associated measurement error (determined as described above) was calculated from a 'true' state for a number of different atmospheric states with varying cloud optical depths, cloud top heights and cloud fractions. The cloud particle size in these tests was 10 μm for a liquid water cloud. An aerosol layer was placed in the boundary layer (bottom 2 km) with an optical depth of 0.5 and it was assumed that the aerosol type was clean continental with an average particle size of 0.5 μm.

- A run of the Forward Model calculates the tropospheric VCD for this atmospheric state.

- A retrieval is performed from the simulated measurement using an assumed, erroneous, cloud state which is defined by changing either the cloud top height (+500 m), the cloud fraction (+0.05), the cloud optical depth (+0.5) or the cloud particle size (+2 μm).

- The error on the retrieved tropospheric VCD is the difference between the 'true' and the 'retrieved' tropospheric VCD and is plotted in the following figures.

- For cloud fraction if the 'true' cloud fraction was 1 then -0.05 was used. It is then assumed that a -0.05 error and a +0.05 error in cloud fraction will have the same affect on the tropospheric VCD. This means the difference between the 'true' and 'modelled' tropospheric VCD for a -0.05 error in cloud fraction, can be multiplied by -1 to give the error due to a +0.05 cloud fraction error.

The cloud parameters that are tested here were determined in Chapter 4.

Cloud Optical Depth

Figure 5.7 shows the impact of an error in cloud optical depth of +0.5 upon the 'retrieved' NO_2 tropospheric VCD. The figure shows that the error in the retrieved tropospheric VCD due to a small error in the cloud optical depth increases with cloud optical depth, cloud fraction and cloud top height and that for low cloud top height the error decreases with increasing cloud optical depth. These are the expected results, as firstly increasing the cloud optical depth will mean there is more scattering above and within the cloud therefore, increasing the photon path lengths and allowing the tropospheric VCD to be larger (increasing the difference between the 'true' and 'retrieved'). However, at low cloud top heights, where the error decreases with increasing cloud optical depth, a change in the optical depth of the cloud makes little difference to the overall error as the photon paths are already very short here (i.e. there is not much NO_2

that is being seen). Therefore, increasing the optical depth is going to make little difference to the photon path lengths and therefore little difference to the retrieved tropospheric VCD. Although the error here is small, as explained in Chapter 4, the profile averaging kernel suggests that the retrieval is insensitive to the boundary layer when there is an optically thick cloud present, meaning that the retrieval is relying on the original tropospheric profile which was used to initiate the retrieval.

Figure 5.7 also shows a positive error at high surface LER at an optical depth of 1–3. This was explained in the previous section, and is caused by the 'retrieved' tropospheric column becoming larger than the 'true' tropospheric column. This is due to the increase in the photon path lengths due to the increased amount of multiple scattering introduced at a cloud optical depth of 3. Also the negative values are large due to the difference in reflectance of the surface and the cloud top reflectance. At high surface reflectance this affect is reversed and the errors are mostly positive.

Figure 5.7 shows errors in the retrieved tropospheric VCD due to a small error in the cloud optical depth of between 1–30% (except for at high cloud optical depth where errors are between 50–100%) for low surface LER and 1–10% for high surface LER (difference due to increased signal at higher surface LER). If these values are compared to the ESD they are comparable in value and therefore it is necessary to quantify the cloud optical depth errors from GRAPE in order to calculate a realistic error estimate for the retrieval.
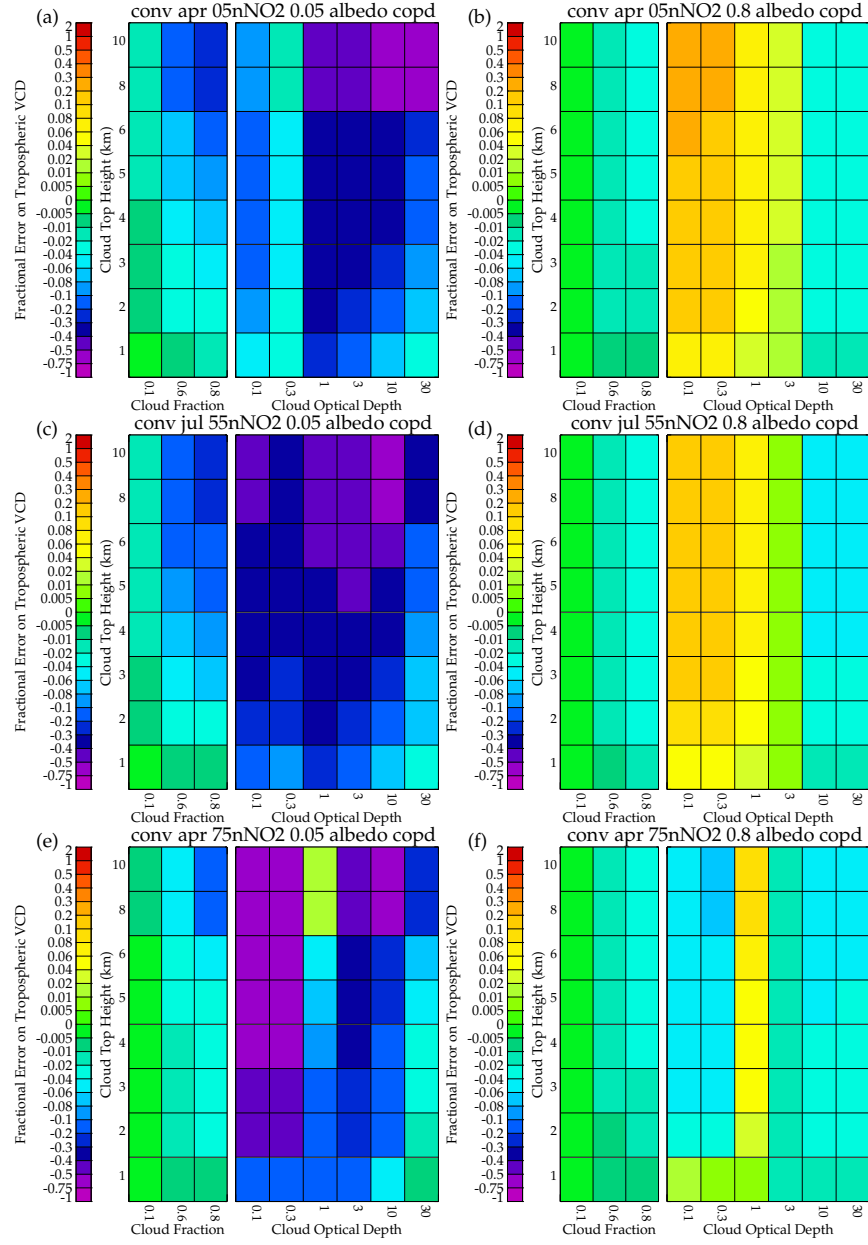
Figure 5.7: *The fractional error in the NO_2 tropospheric VCD (molecules cm^{-2}) due to a +0.5 change in cloud optical depth (a) for April 5°N at a surface LER of 0.05, (b) for April 5°N at a surface LER of 0.8, (c) for July 55°N at a surface of 0.05, (d) for July 55°N at a surface LER of 0.8, (e) for April 75° at a surface LER of 0.05 and (f) for April 75° at a surface LER 0.8. The panels are the same as in Figure 5.3.*

In Appendix A Figure A.3 shows the impact of an error of 0.5 in cloud optical depth for BrO. The figure shows the same affect as seen in Figure 5.7.

Cloud Fraction

Figure 5.8 shows the impact of a 0.05 error in cloud fraction upon the retrieved NO_2 tropospheric VCD. The figure shows that error in the tropospheric VCD increases with increasing cloud optical depth, cloud fraction and cloud top height. At high solar zenith angles and optical depths around 3, the errors are small but positive. This is due to the slight increase in the signal received by the satellite by increasing the cloud fraction[5].

Overall, at low surface LER error in the retrieved tropospheric VCD due to cloud fraction errors can become large with values up to 100% for low surface LER at high cloud top height and high optical depth. At lower cloud optical depths the errors are between 0–40% of the retrieved tropospheric VCD. At higher surface LER values the maximum error is only around 20%. This is due to the increased signal.

In Appendix A Figure A.4 shows the impact of a 0.05 error in cloud fraction for BrO. The figure shows the same affect as seen in Figure 5.8.

---

[5]It should be noted that for the cloud fraction of 1, where it was stated in the method that -0.05 was used in place of +0.05 (as cloud fraction cannot be great than 1) this error has been scaled to a +0.05 case.
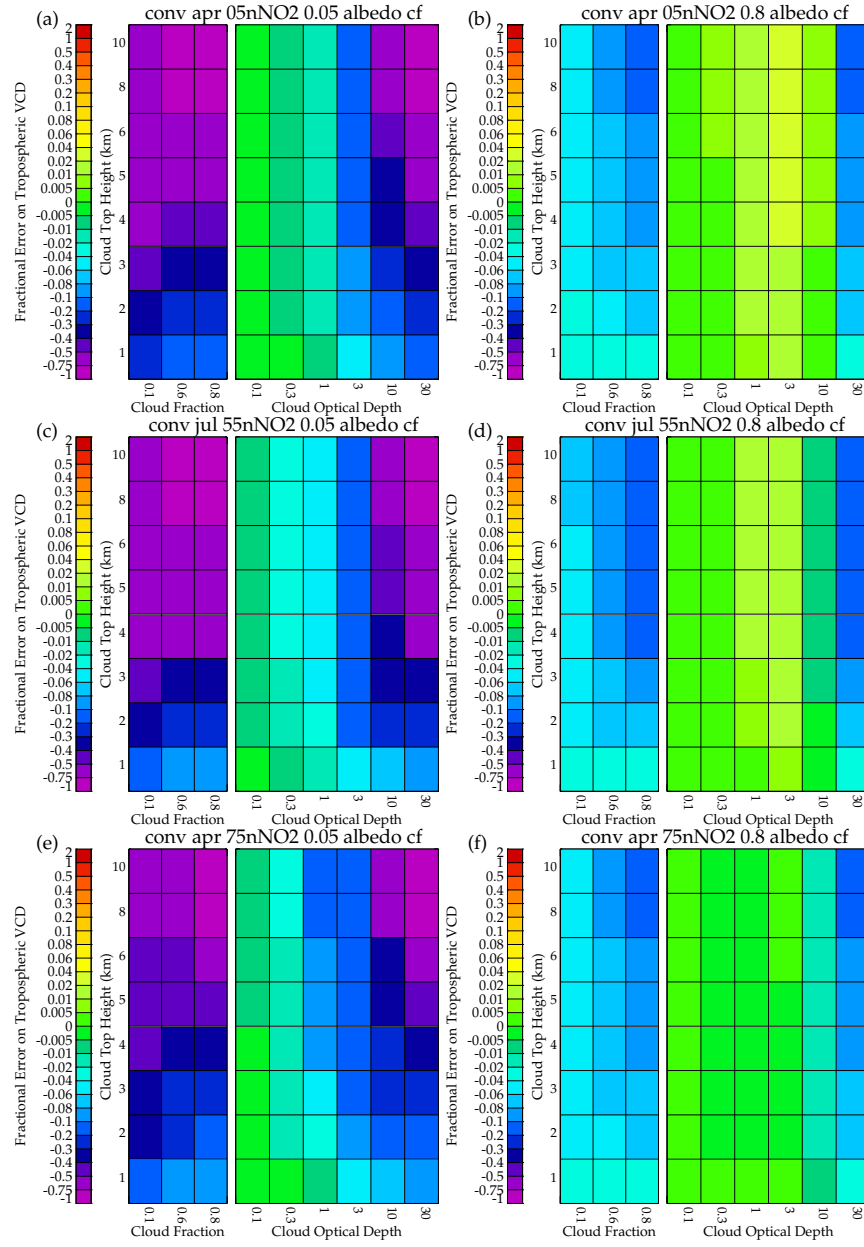
Figure 5.8: *As Figure 5.7 but for a +0.05 error in cloud fraction for NO$_2$.*

Cloud Particle Size

Figure 5.9 shows the impact of a +5 μm error in particle effective radius assuming a liquid cloud with a particle size of 10 μm on the retrieved NO_2 tropospheric VCD. The figure shows that an error in particle size affects the tropospheric VCD in the same way as an error in cloud optical depth and cloud fraction. However, particle size has a much smaller affect on the tropospheric VCD except at high optical depths. Looking at Figure 2.7(a) it can be seen that changing the particle size for a liquid cloud will change the extinction (which is the optical depth). It can therefore be expected that Figure 5.9 is similar to Figure 5.7.
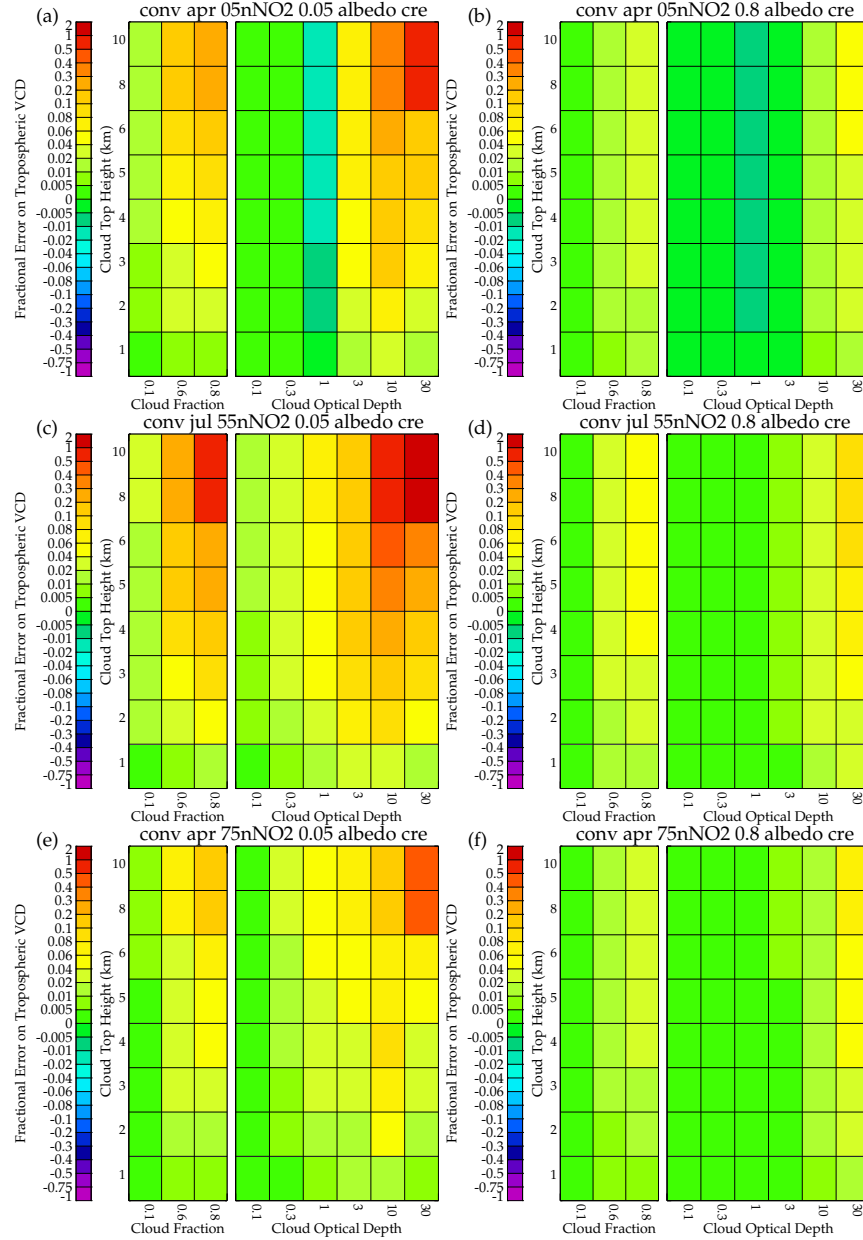
Figure 5.9: *As Figure 5.7 but for a +5 μm error in particle effective radius for a liquid cloud.*

This relationship is the same for BrO. In Appendix A, Figure A.5 shows slightly larger errors fro BrO with a similar trend as observed in cloud optical depth and cloud fraction errors.

Figure 5.10 shows the impact of a +10 μm error in particle effective radius assuming an ice cloud with a particle size of 60 μm. Looking at Figure 2.7(b) it can be seen that for an ice cloud the optical depth varies less with changing particle size. This is observed in Figure 5.10 where if compared to Figure 5.9 it can be seen that the errors are much smaller for ice clouds than liquid clouds.
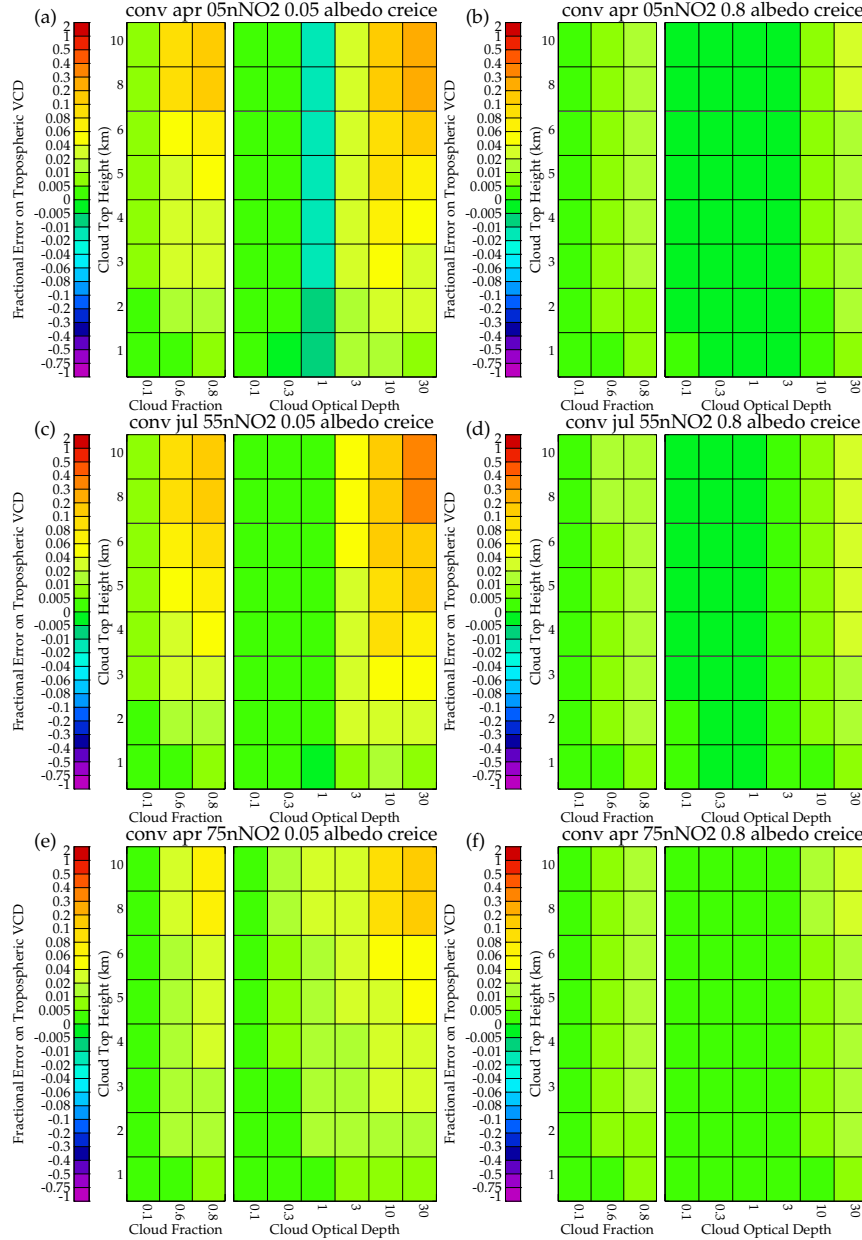
Figure 5.10: *As Figure 5.7 but for a +10 μm error in particle effective radius for an ice cloud.*

In Appendix A, Figure A.6 the same results as Figure A.5.

Cloud Top Height

Figure 5.11 shows the impact of a +500 m error in cloud top height on the NO_2 tropospheric VCD. The figure shows that for an error in cloud top height the error decreases with increasing cloud top height and also the error increases with increasing cloud top height above a cloud top height of 4–6 km and an optical depth of 3 for low surface LER. The decrease in the error with cloud top height is due to more of the trace gas being present in the measurement with the decreasing of the cloud top height. The increase in the error at high optical depths above 4 km is due to the fact that the measurement at this point will be very small as the cloud is obscuring nearly all of the troposphere (see Figure 4.9) and therefore if the cloud top is moved 500 m higher then even less of a small amount will be observed.

For high surface LER the errors are smaller due to the increased signal, however for low cloud top height at optical depths of between 0.3–10 the error become positive. This is due to the 'retrieved' slant column being larger than the 'true' slant column due to the increase in the photon path lengths above the cloud. By increasing the cloud top height in these cases the retrieval becomes more sensitive to the troposphere.

In Appendix A Figure A.7 shows the impact of a +0.5 km error in cloud top height for BrO. As the profile shape in BrO is constant throughout the troposphere there is not much dependence of the error trend on cloud top height here and the error just increases with increasing cloud optical depth and fraction. This shows that the relationship between cloud top height error is profile shape dependent. The value of +500 m was chosen here as this is the random error in cloud top height for GRAPE found through ground based comparisons with Chilbolton Radar[6].

Overall the trends in the cloud parameter errors are the same, i.e. the errors increase with cloud fraction and optical depth, there is an increase in the errors at an optical depth of 3. However, cloud top height error is more significant at low cloud top height. This is mostly due to the profile of the target species, as most of the gas residues in the bottom 2 km which is where the error in will be more important.

---

[6]personal communication with Dr Caroline Pouslen (at RAL).
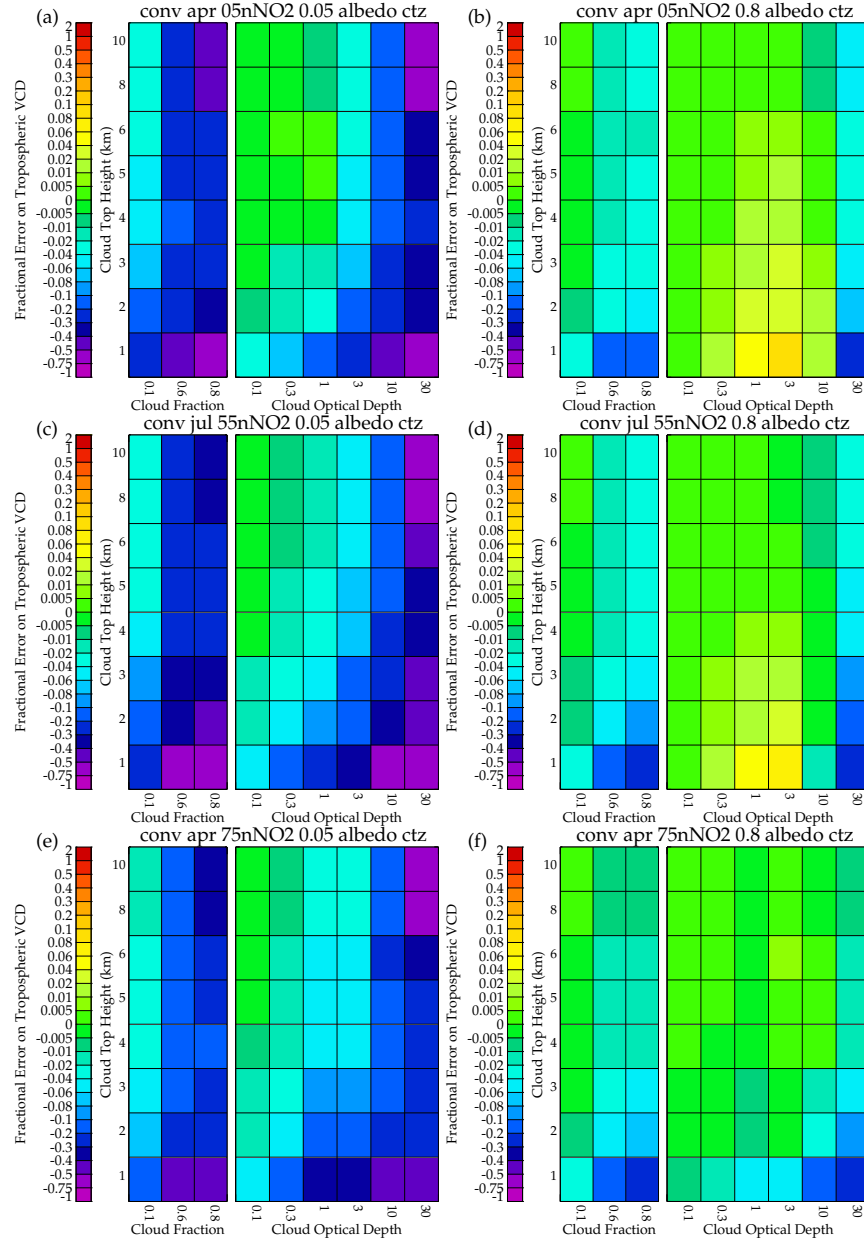
Figure 5.11: *As Figure 5.7 but for a +500 km error in cloud top height.*

5.6.2 Sensitivity to GRAPE errors

This section discusses the errors in the GRAPE parameters, systematic and random. Chapter 4 showed that cloud parameters such as cloud optical depth and cloud top height will cause errors in the retrieval scheme. It is therefore important to not only have a measure of the reported measurement error from the GRAPE retrieval, but to also have a measure of the systematic error in the GRAPE parameters. To calculate the systematic error in the GRAPE parameters they need to be compared to independent observations.

GRAPE reported random errors

The GRAPE random errors come from the measurement covariance calculated in the GRAPE retrieval. This is the equivalent ESD for each GRAPE parameter (cloud optical depth, cloud fraction, cloud particle size, and cloud top height). Every ATSR–2 pixel has been used and the errors have been averaged into a number groups, depending on the cloud optical depth of the pixel. This has been calculated for 4 months.

Figure 5.12 shows the reported measurement error in cloud optical depth retrieved within GRAPE, averaged at different cloud optical depths. The figure shows that for very low cloud optical depths (less than 1) the error is around ±1. It then decreases and becomes negligible, around ±0.25, before increasing at higher cloud optical depths (around 100 and above) where the error become very large, growing to $\sim \pm1$.

Figure 5.13 shows the reported measurement error in cloud top height retrieved within GRAPE averaged at different cloud optical depths. The figure shows large error (±8 km) for low optical depths (around 1). This is because the cloud is too thin for a precise determination of the cloud top height. The error then decreases for optical depths between 1 and 80, after which the error increases again to ±3 km for optical depths larger than 100.

Figure 5.14 shows the reported retrieval error in cloud particle size retrieved within GRAPE averaged at different cloud optical depths. The figure shows large error (±40–50 μm) for low optical depths (around 1). Error then decreases for optical depths between 1 and 80.

Figure 5.15 shows the reported measurement error in cloud fraction retrieved within GRAPE averaged at different cloud optical depths. The figure shows that the errors are around 0.1 for all cloud optical depths. The error decreases slightly with increasing optical depth. The low optical depth causes higher errors as there are less photons interacting with the cloud, therefore providing less information.

In summary, the GRAPE reported retrieval error for the cloud parameters has shown that at low optical depths (less than 1) cloud parameters are uncertain due to the cloud being too tenuous for precise retrieval of particle size or altitude. However, for low optical depths the cloud will not impact the retrieval (as will be shown in the previous section) and therefore this large error will become insignificant. Interestingly, Boersma et al. (2004) use ATSR–2 cloud parameters as the true values to compare their GOME cloud retrievals too. Therefore,
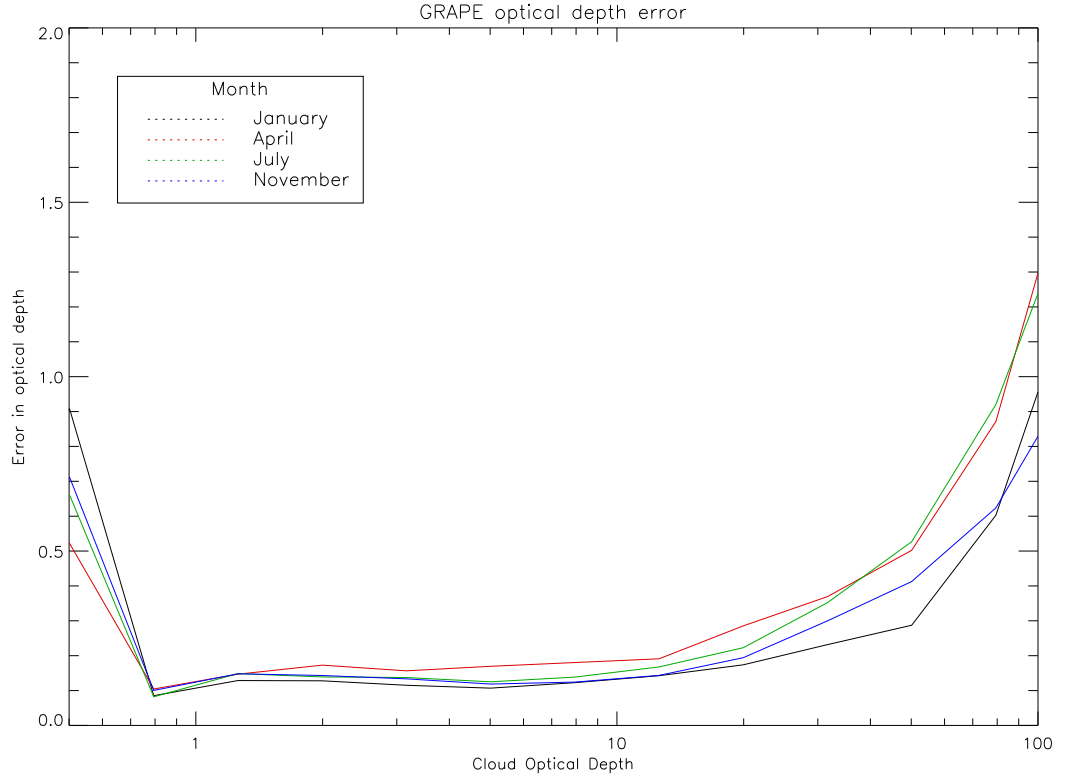
Figure 5.12: *Average GRAPE cloud optical depth error with respect to the different cloud optical depths for 4 months, January, April, July and November 1997.*

their reported error is smaller than that reported here, but their assumption is not valid (that ATSR–2 cloud data is perfect).

For truly random error it should be possible to decrease the overall error by averaging. This is only possible however, if the errors are completely uncorrelated. For the GRAPE retrieval it is therefore necessary to determine the scale on which the error is truly random. It could be assumed that the retrieval errors are random from pixel to pixel, which would mean that the error on the retrieval would be easy to calculate. However, it is more likely that the retrieval errors reported by GRAPE may be partially systematic in similar cloud conditions. Therefore, it may be that the reported retrieval error within GRAPE varies on maybe a 40 km×40 km scale rather than 1 km×1 km scale (pixel-by-pixel).
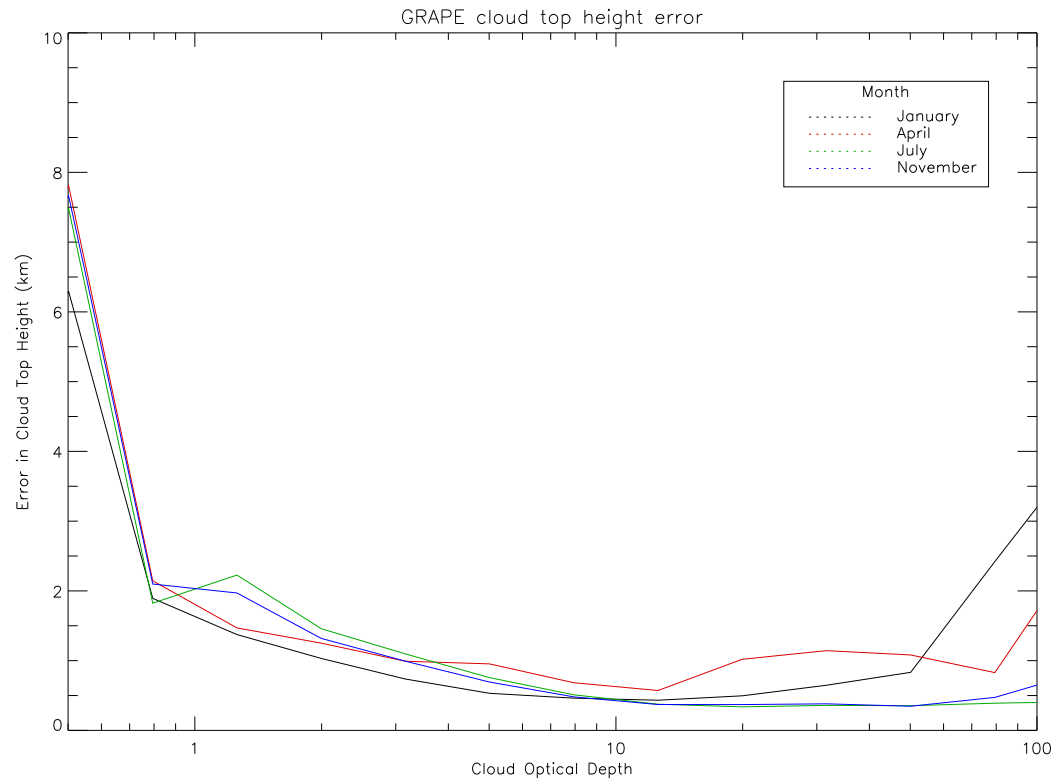
Figure 5.13: *As Figure 5.12 but for cloud top height.*
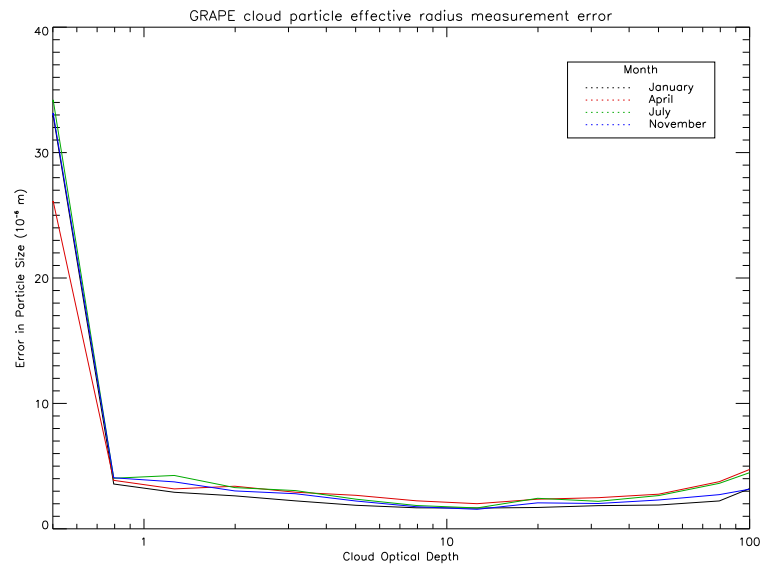
Figure 5.14: *Same as Figure 5.12 but for cloud particle size (μm).*
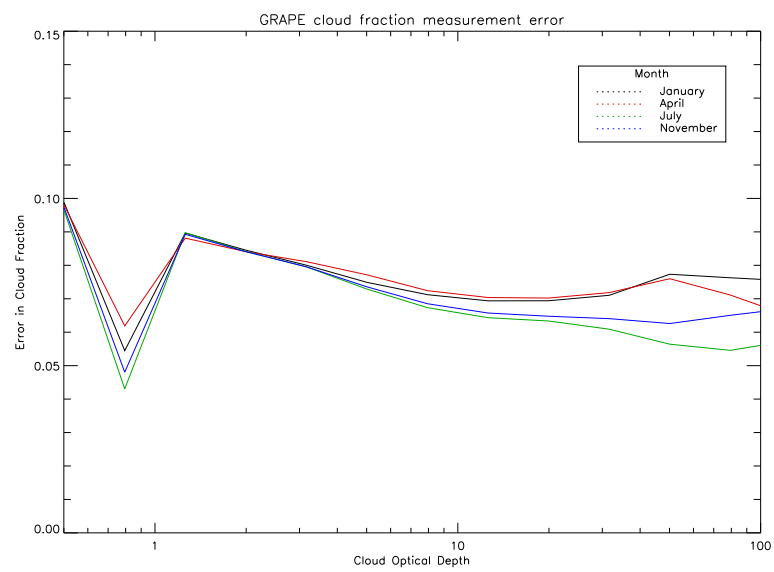
Figure 5.15: *Same as Figure 5.12 but for cloud fraction.*

5.6.3 Error in Tropospheric VCD caused by the combined affect of errors in GRAPE

The reported random errors in the GRAPE cloud parameters have been taken from Section 5.6.2 and using data from the figures in Section 5.6.1 and assuming that the errors can be scaled linearly (which is not the case for cloud fraction) the error in the tropospheric VCD for the reported GRAPE random error can be determined. For example, if the GRAPE cloud optical depth error is 0.1 at an optical depth of 1. The error in the tropospheric VCD for cloud optical depth error in Section 5.6.1 was calculated at an optical depth error of 0.5, therefore if scaled linearly the resulting error in tropospheric VCD due to the reported GRAPE cloud optical depth error would be that value divided by 5. For cloud fraction the errors cannot be scaled linearly. This is shown in Figure 5.8. The error study has been performed for cloud fraction errors of 0.025, 0.05, 0.075, 0.1 and 0.125. These have then all been scaled linearly to 0.1. The figure highlights that the cloud fraction errors cannot be scaled linear (as if they could all these lines would be the same), however a relationship between these test results can be used to scale the cloud fraction errors in a more appropriate way. Therefore, a linear interpolation was used to determine a relationship that describes how the tropospheric VCD error caused by errors in cloud fraction scales with increasing cloud fraction error.

The following figures show the combined affect of the reported random error in the cloud parameters observed by GRAPE on the retrieved tropospheric VCD. These errors need to be combined as they are all correlated. The impact of the reported GRAPE random errors is determined by linearly scaling the errors reported in Section 5.6.1 (expect for cloud fraction which is interpolated according to the results reported in Section 5.6.1). In order to combine random errors the following relationship is used

$$\epsilon_{random} = \sqrt{(\epsilon_{n_0}^2 + \epsilon_{n_1}^2 ... + \epsilon_{n_i}^2)} \tag{5.4}$$

where ϵ_{random} is the overall random error and ϵ_n is the individual random errors (i.e. cloud optical depth, cloud fraction, cloud particle size and cloud top height).

Figures 5.17 and 5.18 show the combined affect of the cloud fraction/cloud optical depth error (increasing error with increasing cloud top height and cloud fraction) and the cloud top height error (decreasing error with decreasing cloud top height). The error in tropospheric VCD reported in these figures is very large (often up to and above 100%). The error reported here is for a single GRAPE pixel and therefore this error will be reduced when averaging the GRAPE data within the GOME pixel. If these figures are compared to Figure 5.7 and 5.8, it can be seen that cloud fraction and cloud optical depth are the dominant sources of error. However, as stated above, it is difficult to determine the scale on which these error are random and therefore it may be more advantageous to determine the errors on GRAPE in a different way. Another method for determining this error is described in the next section.

It is possible that errors in GRAPE parameters are correlated. For example, if GRAPE over-

estimates the cloud optical depth, it will probably underestimate the cloud fraction (these would compensate to give similar radiances and tend to cancel in the tropospheric VCD error) (Boersma et al., 2004). However, correlations are not reported and therefore, it is not possible to comment on this here.
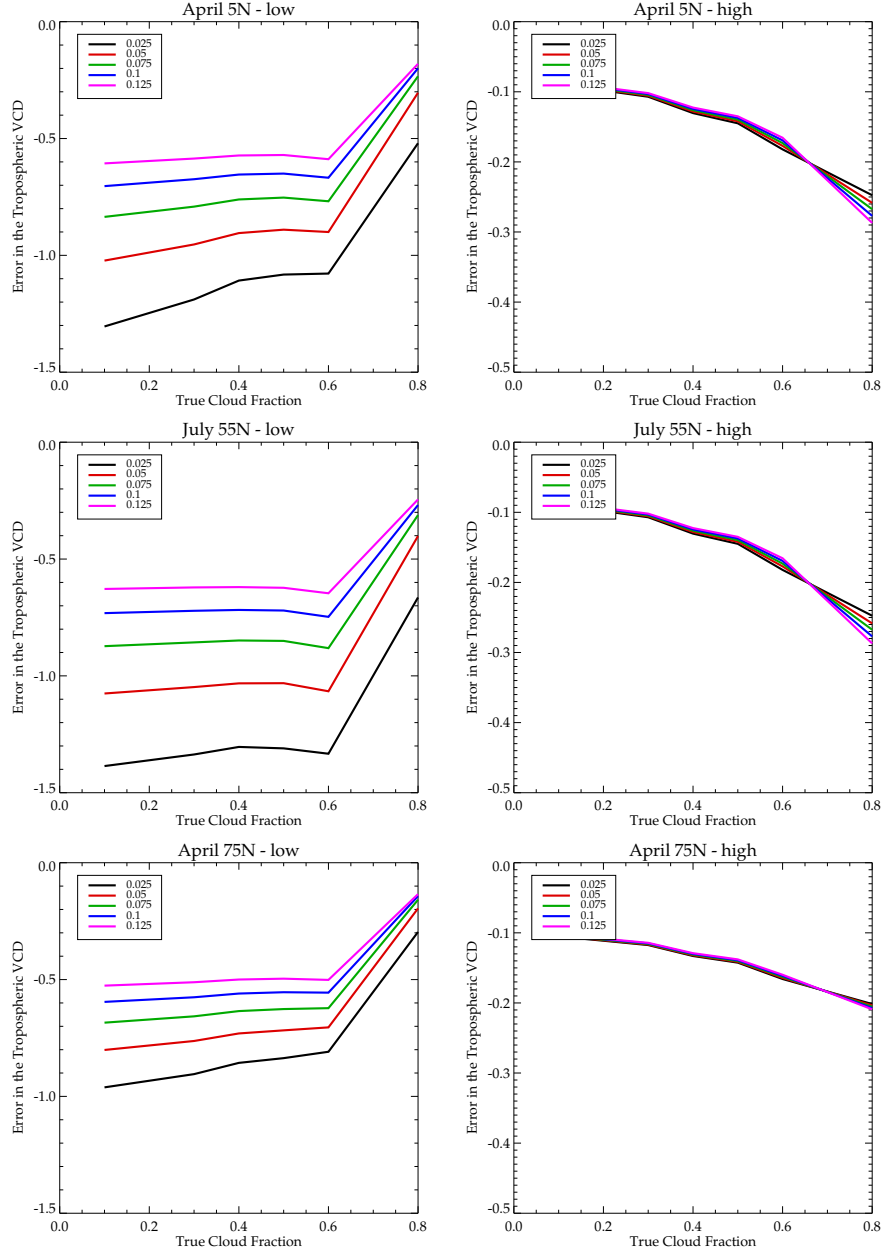
Figure 5.16: *The fractional error in NO$_2$ tropospheric VCD for a cloud fraction error of +0.1 when linearly scaled from 5 different cloud fraction errors (+0.025, +0.05, +0.075, +0.1, and +0.125) for a cloud top height of 4 km (above the boundary layer) and a cloud optical depth of 30 for NO$_2$ for (a) low surface LER for April 5° N, (b) high surface LER for April 5° N, (c) low surface LER for July 55° N, (d) high surface LER for July 55° N, (e) low surface LER for April 75° N and (f) high surface LER for April 75° N.*
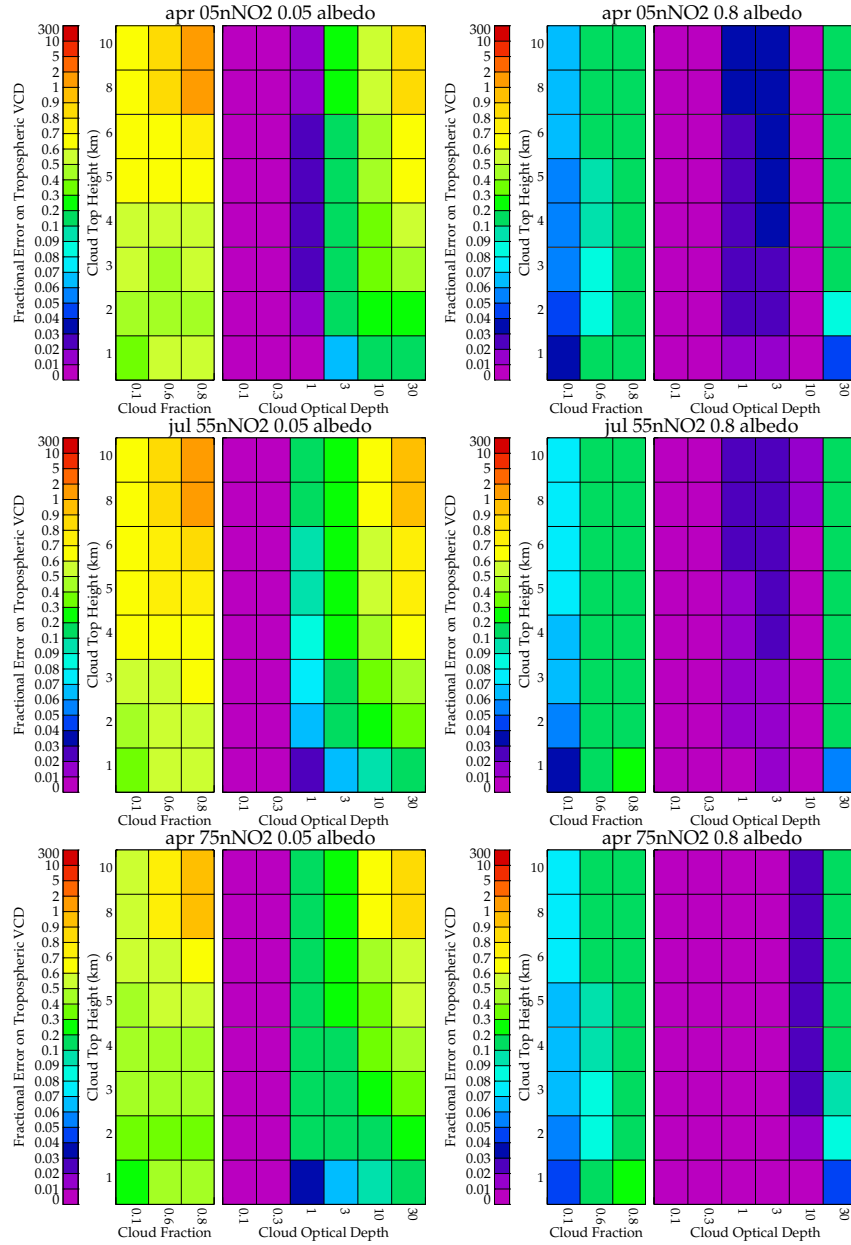
Figure 5.17: *The combined random error due to the reported GRAPE retrieval error for tropospheric VCD NO$_2$ (molecules cm^{-2}) (a) for April 5° N at a surface LER of 0.05, (b) for April 5° N at a surface LER of 0.8, (c) for July 55° N at a surface of 0.05, (d) for July 55° N at a surface LER of 0.8, (e) for April 75° at a surface LER of 0.05 and (f) for April 75° at a surface LER 0.8. In each panel there are two plots, the left plot shows the error when cloud fraction and cloud height are varied, for a cloud with an optical depth of 30. The right plot shows the error when cloud optical depth and cloud height are varied for a cloud fraction of 1.*
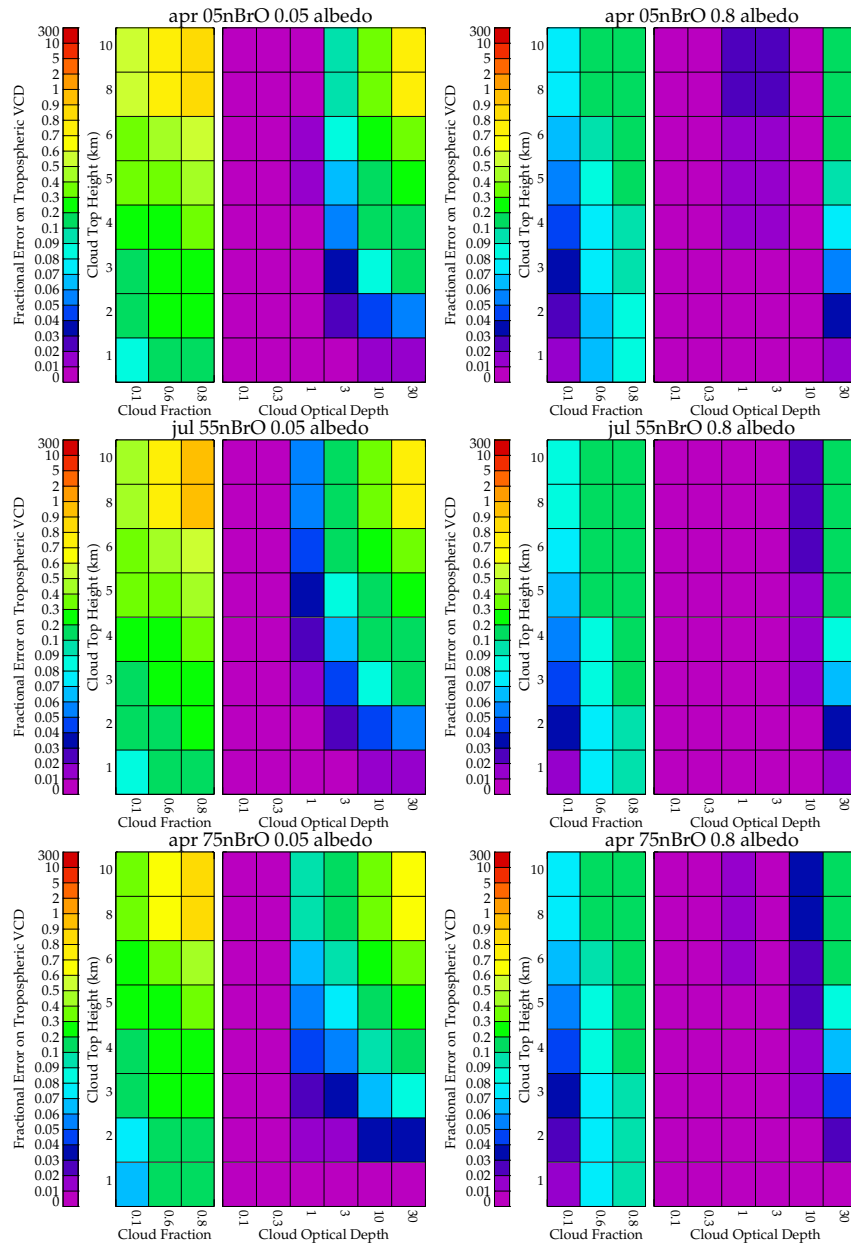
Figure 5.18: *As Figure 5.17 but for BrO.*

5.6.4 GRAPE systematic and random Error for sun–normalised radiance comparisons

Currently, there has been no validation of cloud parameters from GRAPE that could be used in this work to give systematic errors. Therefore, a comparison of modelled and observed sun–normalised radiance has been performed (see Section 4.6) and this is used to derive an 'effective' cloud fraction error. This section describes the methodology used to derive these errors.

Figures 4.29 and 4.30 showed a comparison of modelled and observed sun–normalised radiances at 430 and 350 nm respectively. Sun–normalised radiance is calculated in this new retrieval scheme by taking measurements of cloud and aerosol optical depth at 550 nm (from GRAPE) and interpreting this to the wavelength of the retrieval using the same optical models as were used in the GRAPE retrieval (see Section 2.4). Other cloud and aerosol parameters are not considered to be wavelength specific. Differences between the radiances observed by the GOME instrument and those modelled by the retrieval scheme will be due mostly to cloud parameter errors (in cloud optical depth, cloud fraction and cloud particle size[7]). However, to a lesser extent, they will also be due to aerosol parameter errors and to errors in the optical models used to convert the optical depths to different wavelengths. Therefore, an error in sun–normalised radiance (which can be derived from Figures 4.29 and 4.30) can be used to determine an error in the cloud parameters from GRAPE.

It is difficult to separate the sun–normalised radiance error into the three individual components i.e. cloud fraction, cloud optical depth and cloud particle size. As Figures 5.7–5.10 showed, the impact of errors in cloud optical depth, cloud fraction and cloud particle size on the retrieved tropospheric VCD are all quantitatively similar. Koelemeijer et al. (2001) also use the assumption that cloud optical depth, cloud fraction and cloud particle size can be combined. It is therefore possible to converted the error in sun–normalised radiance to an 'effective' error in cloud fraction (cloud optical depth could also be used but cloud fraction is chosen here). This can be assumed to be the error due to all three cloud parameters (with a contribution due to the optical models and aerosol parameters).

The errors in sun–normalised radiance can be converted into an 'effective' cloud fraction error using the following relationship (taken from Richter and Burrows, 2001):

$$r = \cos\theta(\alpha_c f + \alpha(1 - f)) \tag{5.5}$$

where r is the sun-normalised radiance, θ is the solar zenith angle and f is the cloud fraction, α is the surface LER (in the scenarios use 0.05 and 0.8) and α_c is the cloud top reflectance (assumed to be 0.8 (Koelemeijer et al., 2001)). This equation can be rearranged for the cloud fraction:

---

[7]Cloud top height will not affect radiance
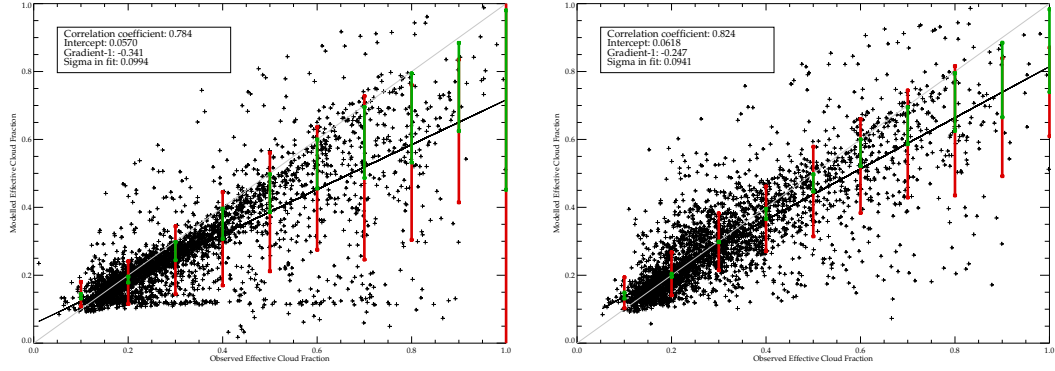
Figure 5.19: *Scatter plots of observed versus modelled sun normalised for NO$_2$ retrieval for many GOME orbits (a) for all pixels, (b)only for pixels with a measurement cost less than 30. The red line shows the mean difference between the observations and the modelled data, and the green line shows the standard deviation of the modelled data for different cloud fraction bins.*

$$f = \frac{(r/\cos\theta) - \alpha}{\alpha_c - \alpha} \tag{5.6}$$

Figures 5.19 and 5.20 are similar to Figures 4.29 and 4.30, but the sun–normalised radiance has been converted to an 'effective' cloud fraction so that the systematic and random errors can be taken directly from the figures. The red bar shows the mean difference between the observations and the model (the systematic error) and the green bar shows the standard deviation of the measurements (the random error). The advantage to using the errors reported in these figures over those reported in the GRAPE retrieval scheme are that these errors provide an independent estimate of the random error, as it applies to the GOME pixel scale. They also provide an estimate of the systematic error that might pertain to this use of the GRAPE data, taking into account errors other than those explicitly included in the GRAPE state vector (e.g. errors in the optical models).

The errors have been averaged for different cloud fraction bins as at low cloud fraction the error will be smaller than at high cloud fraction (as has been shown throughout this thesis).

In calculating the random and systematic error the scenarios that use a maximum measurement cost are used. This significantly reduces the errors and also removes multi-layered cloud situations (EUMETSAT, 2007). The value of 30 for the measurement cost was choose due to tests carried out in Chapter 4.

5.6.5 Error in Tropospheric VCD caused by the random and systematic GRAPE errors

Figure 5.21 and 5.22 shows the affect of the systematic 'effective' cloud fraction error on the retrieved NO$_2$ and BrO as a function of cloud fraction, cloud top height and cloud optical
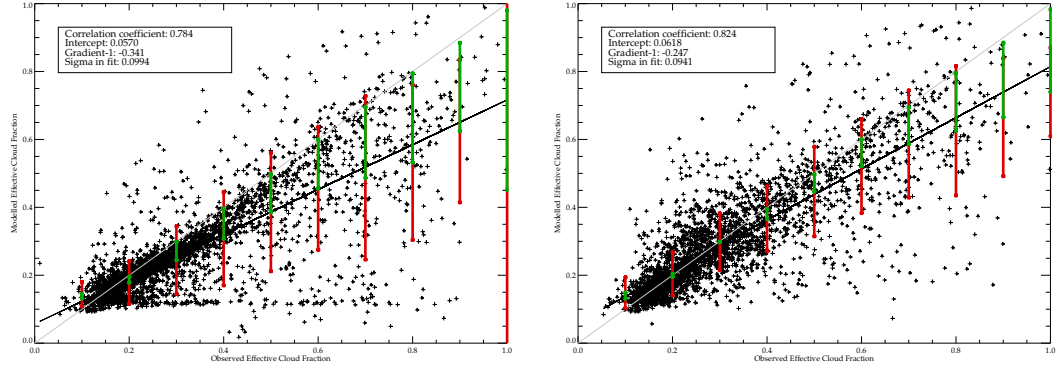
Figure 5.20: *As Figure 5.19 but for BrO.*

depth. The figures show that except for at low surface LER at high optical depths (10–30), the retrieved NO_2 and BrO tropospheric VCD varies from 0–40%. Tropospheric VCD errors are reported here for a particular magnitude of the optical depth error. The effect of estimated optical depth errors which are expected from the GRAPE retrieval are described in Section 5.6. However, it is important to consider the systematic errors as they cannot be reduced in monthly means and therefore the systematic errors will remain large, whereas ESD random error will be reduced.

190
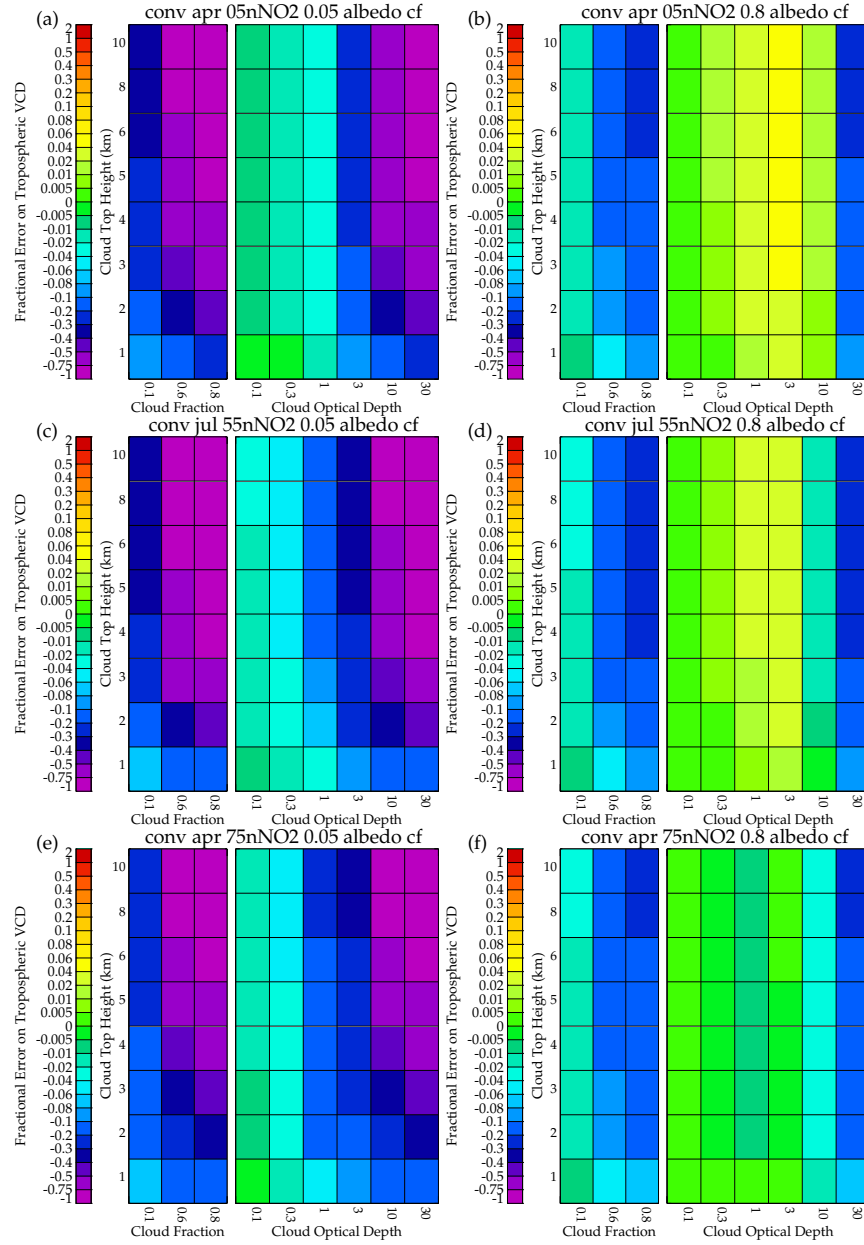
Figure 5.21: *Impact of the GRAPE systematic error in 'effective' cloud fraction on the retrieved NO₂ tropospheric VCD. The figure is the same format as Figure 5.3.*
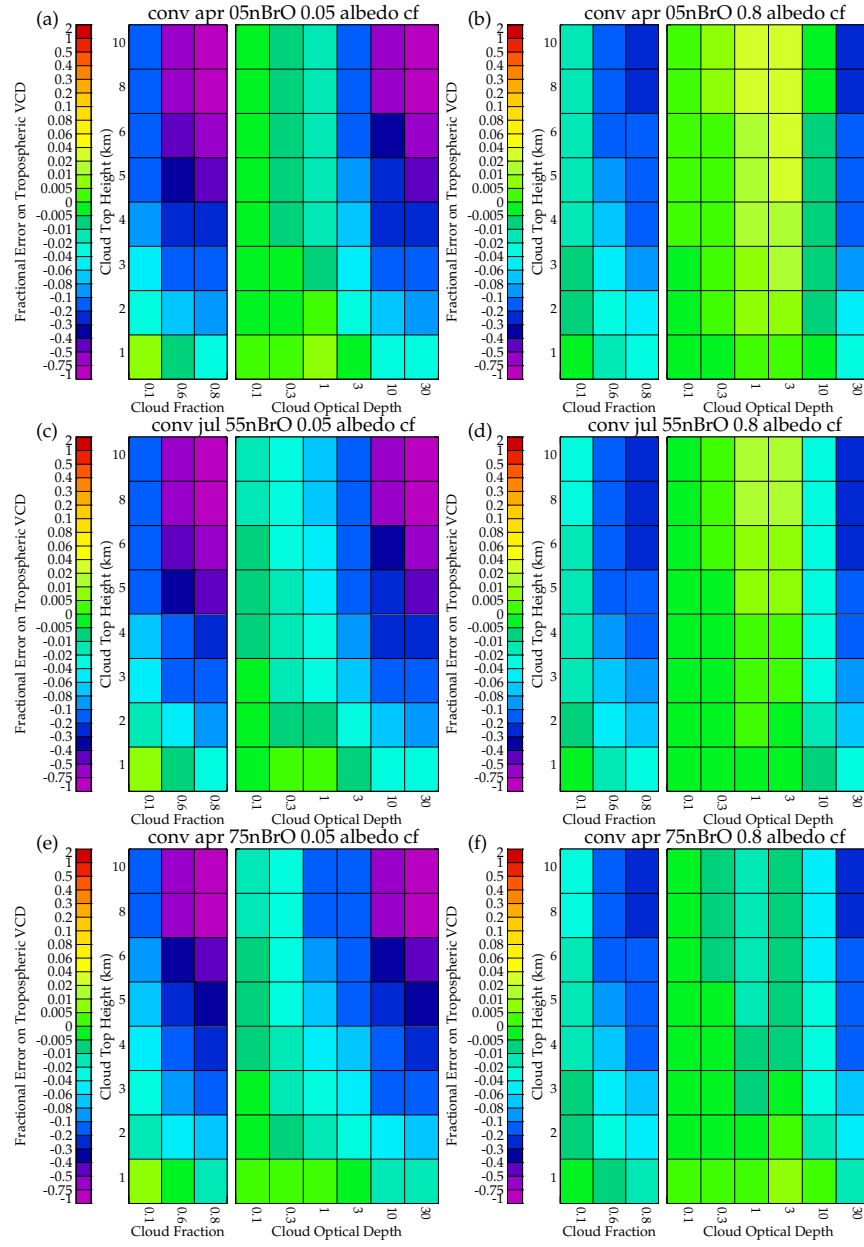
Figure 5.22: *As Figure 5.21 but for BrO.*

Figure 5.23 and 5.24 shows the affect of the sum of the random 'effective' cloud fraction and the random cloud top height error on the retrieved NO_2 and BrO as a function of cloud fraction, cloud top height and cloud optical depth. The figures show that except for at low surface LER at high optical depths (10–30), the retrieved NO_2 and BrO tropospheric VCD varies from 0–+100%. If these figures are compared to the ESD in Figures 5.3 and 5.4 it can be seen that the ESD is slightly larger if not the same of the error due to GRAPE random error. Therefore, it is unnecessary to consider the GRAPE random error.
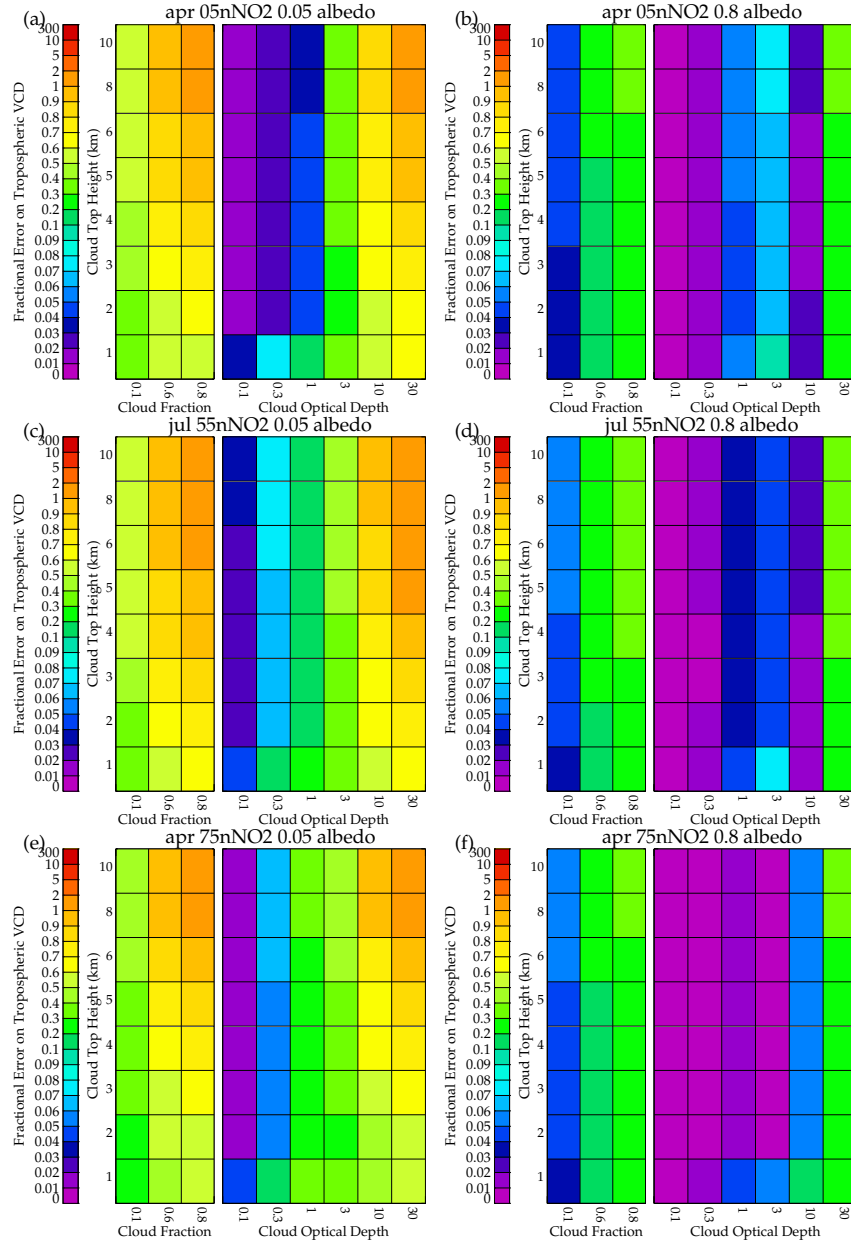
Figure 5.23: *Impact of the GRAPE random error in 'effective' cloud fraction on the retrieved NO₂ tropospheric VCD. The figure is the same format as Figure 5.3.*
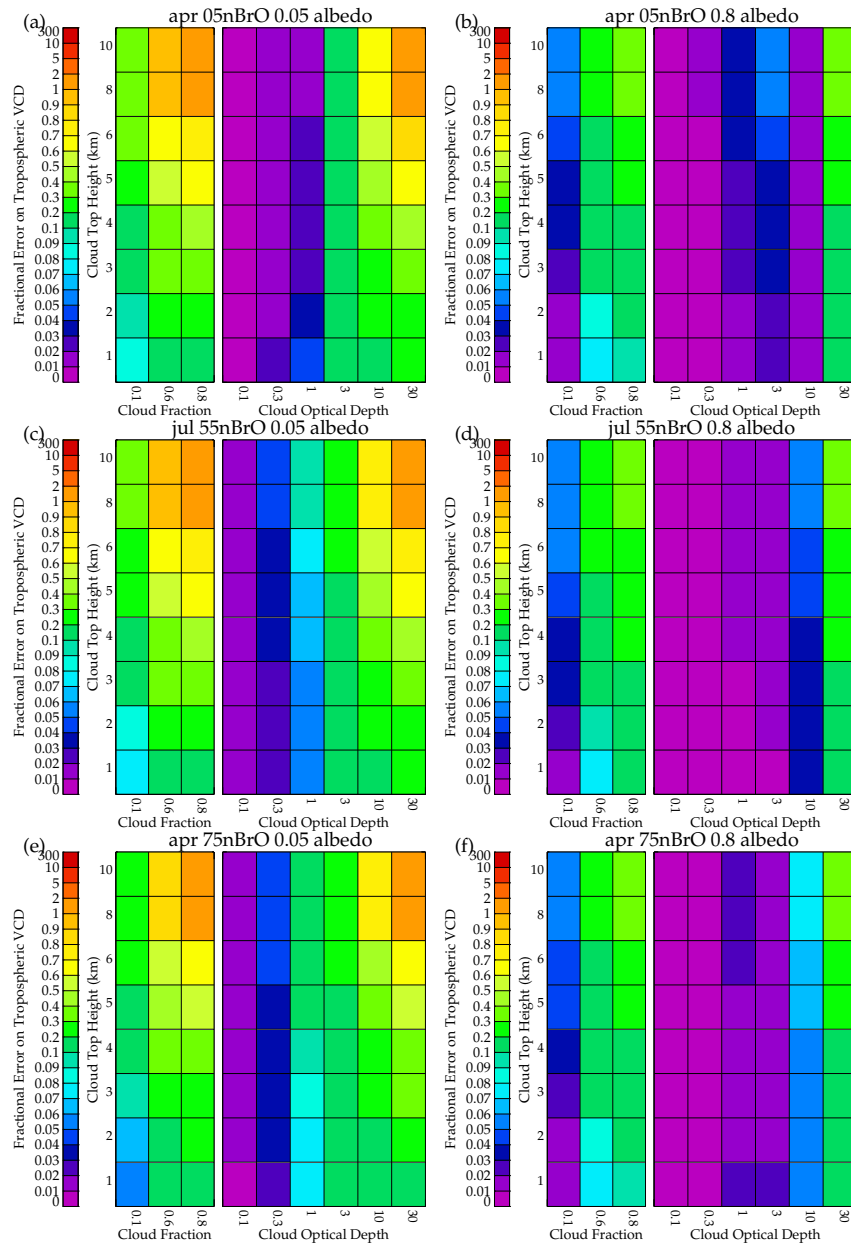
Figure 5.24: *As Figure 5.23 but for BrO.*

5.6.6 Scaling Errors when Tropospheric Concentration Changes

The overall aim of this section was to determine a model of how systematic and random errors in the GRAPE cloud parameters affect the retrieved tropospheric column. However, for each scenario a fixed tropospheric concentration was assumed. Therefore, when scaling the errors derived in this section to those for the final retrieval, the tropospheric column will vary. Figure 5.25 shows how the error in the tropospheric VCD due to cloud fraction error varies with changing tropospheric concentration. The figure shows that the fractional error in the retrieved tropospheric column varies slightly with changing tropospheric concentration, however, this is only around 1% for a 50% increase in the concentration. Therefore, the fractional error is the same for all values of the tropospheric concentration.

This test assumed that the tropospheric profile did not change. Chapter 4 showed that changing tropospheric profile shape can cause changes in the tropospheric VCD. However, in reality it is not possible to know the true profile and it is therefore assumed that the TOMCAT profile (which is used in the retrieval model) is correct.
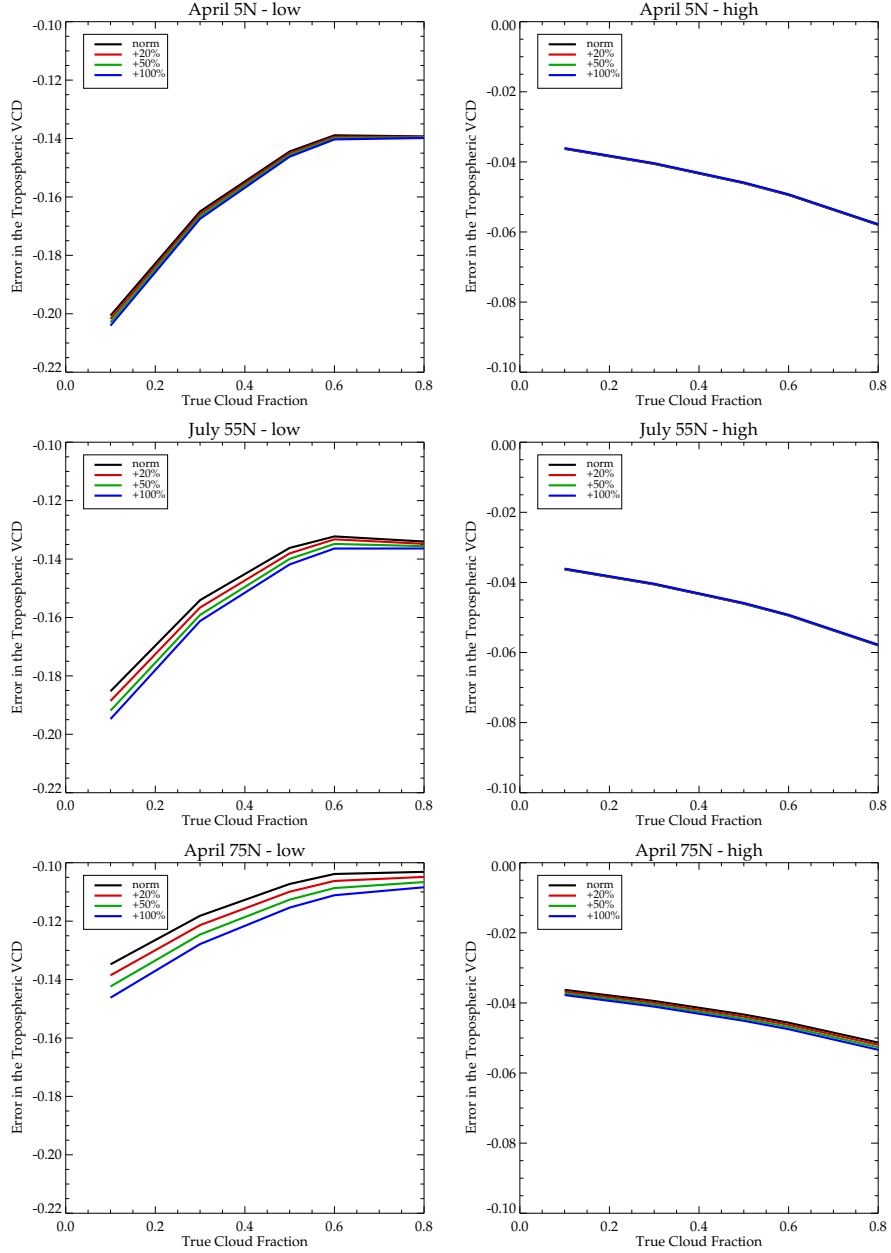
Figure 5.25: *The fractional error in NO_2 tropospheric VCD for a cloud fraction error of 0.05 for the original tropospheric column, and then for a 20%, 50% and 100% increase in the tropospheric column. The cloud top height here is 4 km (above the boundary layer) and a cloud optical depth of 30 for NO_2 for (a) low surface LER for April $5°N$, (b) high surface LER for April $5°N$, (c) low surface LER for July $55°N$, (d) high surface LER for July $55°N$, (e) low surface LER for April $75°N$ and (f) high surface LER for April $75°N$.*

5.7 GRAPE Aerosol Parameter Error

Aerosol retrievals from GRAPE have large errors associated with them especially over land. Definitive errors in the aerosol retrieval were not available for this work. However, an upper limit corresponding to the natural variability of aerosol can be taken as a starting point. In Section 4.3 it was shown that aerosols with effective radii from 0.05 μm upwards will affect the radiation of the wavelengths used in this thesis (350 and 430 nm). Also, from a number of ground based studies (see Raes et al., 2000; Dubovik et al., 2002) show that aerosol optical depth varies from 0.05-0.1 for the natural background conditions, to 0.5 for pollution plumes and 1. for pollution events. This natural variability was introduced in Section 4.3, where it was stated that

- Aerosol optical depth in the wavelength of interest (UV/VIS) will vary from 0.05 for normal back conditions, to 1 for pollution events.

- Aerosol particle sizes that will impact this wavelength (Rayleigh and Mie scatterers) will vary from 0.05 μm to 1μm.

Although this error is small compared to cloud parameter errors it will be corrected for in the 'effective' cloud fraction calculation as it will affect the sun–normalised radiance calculation.

5.8 Impact Stratospheric errors on Retrieved Tropospheric VCD

It was shown in Chapter 4 by looking at the stratospheric profile averaging kernels that errors in the stratospheric column can have a large affect on the retrieval. Also previous work such as Boersma et al. (2004) show the errors in the stratospheric column can cause large errors in the retrieved tropospheric VCD. Although the systematic error in the stratospheric column can be removed by comparing with independent ground based data, the overall random error in the column is still around 10% for both NO_2 and BrO. However, errors in the stratospheric column will not scale in the same way as other Forward Model parameters as it is dependent on the tropospheric VCD.

Figures 5.26 and 5.27 show the error in the tropospheric VCD due to the reported random error in the stratospheric component for the scenarios described throughout this chapter. As Chapter 3 showed that systematic error in the SLIMCAT stratospheric column could be removed therefore leaving random errors in stratosphere. The figures show that error in the retrieved tropospheric VCD due to the random error in the stratosphere are comparable or slightly smaller than those reported in the ESD (Figures 5.3 and 5.4). Therefore, the overall random error, will be comparable to the stratospheric column error with a small increase due to the slant column error (ESD).
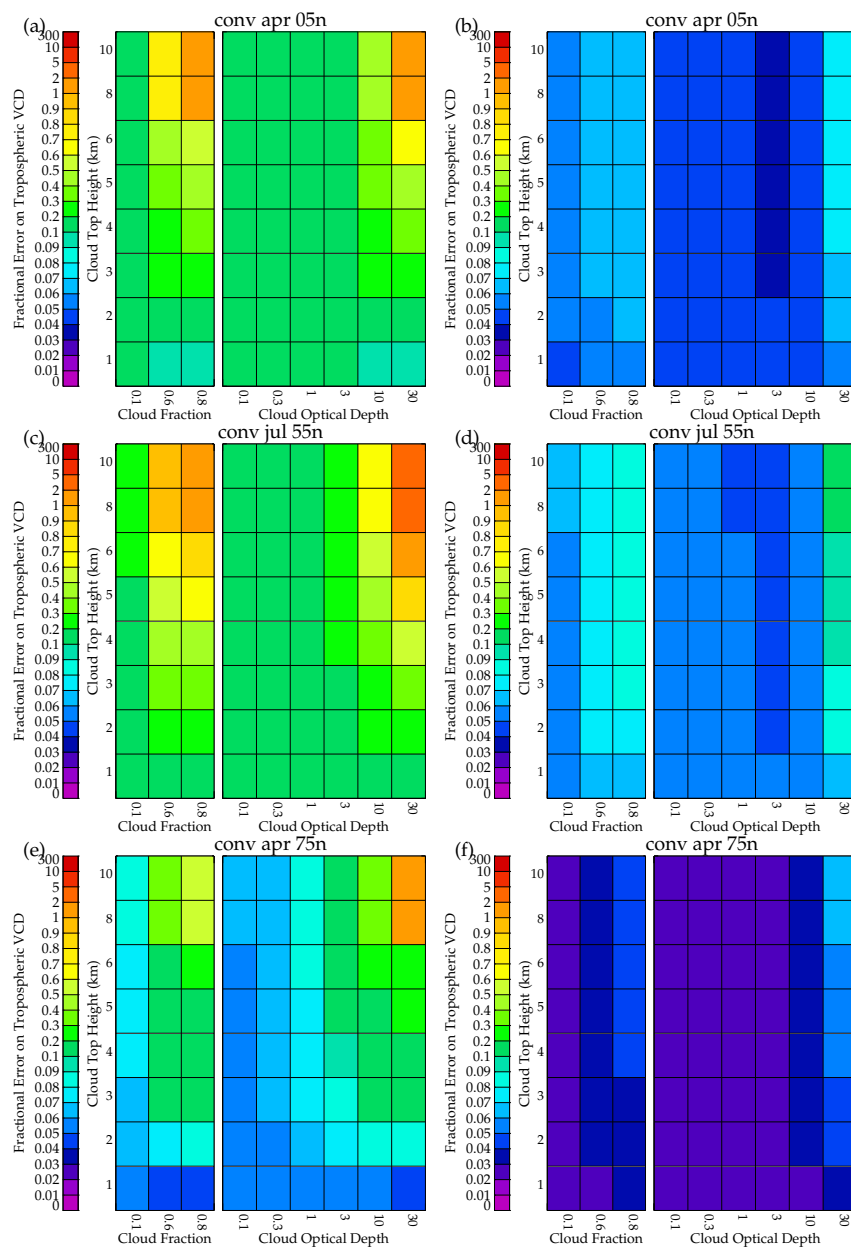
Figure 5.26: *As Figure 5.7 but for random error in the stratospheric NO$_2$ VCD.*
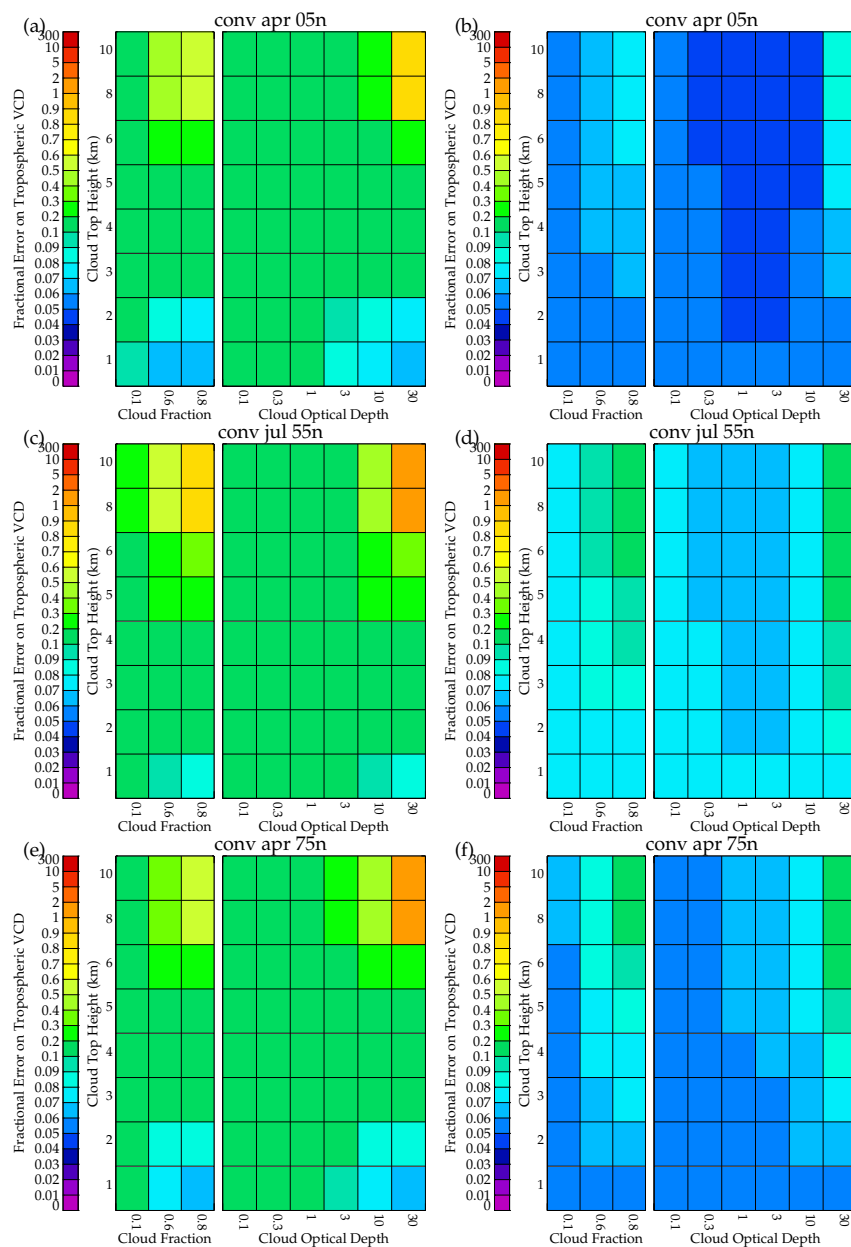
Figure 5.27: *As Figure 5.26 but for stratospheric BrO.*

5.9 Other Errors

1. Systematic slant column errors have been corrected for via ground based validation and are considered small compared to that introduced by the diffuser plate artefact. Residual errors will also be reduced by the diffuser plate correction.

2. Scalar approximation errors are around a maximum of 10% of the retrieved tropospheric VCD. As this will be a random error, when compared to the value of the ESD this is not be significant.

3. Multi–layer cloud errors will be mostly removed by the use of only low measurement cost pixels. However, some will still remain. These will give high sun–normalised radiance errors and therefore will be accounted for in the 'effective' cloud fraction error. It should be noted that as it is not possible to determine which pixels contain multi–layered clouds, it can only be done by speculation and therefore any that remain will just be treated as single layers.

4. Profile errors are hard to quantify as the 'true' profile is not known. It is therefore important to remember that the profile shape will be in error and that error is not quantified here. When looking at the final retrieval product, the profile averaging kernel for each retrieval can be studied and this will give some information on how sensitive the satellite was to certain parts of the profile (see Chapter 4).

5.10 Summary

This chapter has investigated the errors associated with the new retrieval. Also the justification for the use of two parameters in the state vector (not only tropospheric scaling factor but also surface LER) was discussed. Here it was shown that the error on the retrieved vertical column would be large if surface LER was not retrieved. Therefore, by combining it within the retrieval the error can be reduced. This would also help to improve the retrieval of cloudy scene by adjusting the overall reflectance of the scene (i.e. not just the surface but also the cloud reflectance).

It has been shown that the ESD increases with increasing cloud optical depth, cloud top height and cloud faction. Also, for species where the majority of the tropospheric column lies in the boundary layer e.g. NO_2, for high optically thick clouds the retrieval is not sensitive to the boundary layer and therefore the errors will be very large (over 50% of the tropospheric VCD).

When investigating where cloud modelling will improve the accuracy of the retrieval the errors show that when clouds are not explicitly modelled the systematic errors are large. Therefore, the retrieval can be improved by including cloud.

The GRAPE errors, as they apply to this retrieval scheme, both random and systematic, have been estimated. The GRAPE systematic and random errors have been derived by looking at the error in the observed sun-normalised radiance (that seen by GOME) and the modelled

sun-normalised radiance. This difference between these can be attributed to systematic errors in GRAPE cloud parameters and in the error due to the extrapolation of GRAPE parameters from 550 nm to the wavelength of the retrieval (either 430 nm or 350 nm). This sun-normalised radiance difference can then be related to a cloud fraction error which can be used to determine the impact of the retrieved tropospheric VCD.

Impact of the different errors on the retrieved tropospheric vertical column has been investigated in this Chapter. The main parameters that will affect the Forward Model are all correlated and therefore a look-up-table of the errors needed to be created (one for random and one for systematic error). This showed that the error due to systematic errors in cloud and aerosol parameters in GRAPE would result in errors of 10-30% in most cases but it can be higher up to 100% if the cloud fraction and cloud optical depth are large.

It has been shown overall that the main sources of error that will affect the retrieval are:

- Forward Model Parameters
    - Cloud Parameters - optical depth, fraction and top height.
    - SLIMCAT stratospheric errors.
- Slant Column Density Errors or measurement errors

Now it is possible to take Table 4.5 and determine which sources of error will significantly impact the retrieval.

Slant Column Density Error

For NO_2 the slant column density error (taken from previous work) was between 5–9×10^{15} molecules cm^{-2}. It has been shown via the ESD tests that this will significantly affect the final retrieved tropospheric column, especially in the presence of clouds. For BrO the slant column density error was between 1.5–1.8×10^{13} molecules cm^{-2} which was also shown to significantly affect the final retrieved tropospheric column. It should be noted that this is a random error and will be average down in a monthly mean.

Scalar Approximation Error

Previous studies (Latter, 2004) showed that scalar approximation error will result in a random error on the tropospheric VCD of $\sim$10% for BrO and an upper limit of $\sim$10% for NO_2.

Plane–parallel Assumption

Kerridge et al. (2004) states that the error due to the plane–parallel assumption is negligible for all scenarios simulated in their study (which are similar to the scenarios used here) if

the pseudo–spherical approximation is used in the radiative transfer. This pseudo–spherical approximation is included in GOMETRAN and therefore it can be assumed here that the plane–parallel approximation error is insignificant.

Stratospheric Column Errors

For both NO_2 and BrO the stratospheric column errors are similar to the values of the ESD and therefore are significant in terms of the overall retrieval.

Cloud and Aerosol Parameters Errors

It was first shown in Chapter 4 that clouds will significantly affect the retrieved tropospheric VCD if they are present. It was shown in this chapter that systematic errors in the GRAPE parameters are significant and are the major source of systematic error on the final retrieval product. However, random errors on the retrieved tropospheric VCD due to random errors in the GRAPE data are smaller than those due to ESD and the stratosphere.

Multi–cloud Layers

It was shown in Chapter 4 that it is not possible to determine where multi-cloud layers are retrieved. However, high measurement cost in the GRAPE retrieval signals that the atmospheric state did not fit the optical model (which is a single layer cloud) and therefore this is likely to be due to multi–cloud layers or thick clouds. These high measurement cost GRAPE retrievals will not be used in the final retrieval run to help to lower the overall error on the retrieval. It should be noted that this is supposition and not all multi–layer clouds will be detected. A subject of further work here could be to combine GRAPE information with the O_2 A–band type cloud retrieval (used by other GOME retrievals). Although GRAPE should be better in providing sub–GOME–pixel information on cloud, however the A–band does measure the overall impact of cloud on the light path (see Siddans et al., 2007).

Tropospheric Species Profile Shape

In was shown in Chapter 4 that tropospheric species profile shape could introduce large errors if the profile that is used to initiate the retrieval model is in error. However, this is difficult to quantify as the true profile shape will always be unknown. However, the species profile shape comes from a validated Chemical Transport Model (CTM) and therefore it can be assumed that the profile is reasonably accurate.

In conclusion, this chapter showed that clouds, slant column density (measurement) errors and stratospheric column errors have a significant effect on the retrieved tropospheric VCD and that all these parameters need to be considered when quantifying the errors on the retrieval in

Chapter 6.

Chapter 6

Retrieval Results

6.1 Introduction

This chapter presents example results from the Retrieval Model. The results focus on two particular months in order to illustrate the performance of the scheme. For NO_2 July 1997 and for BrO November 1996 are investigated.

Section 6.2 shows how the diffuser artefact can be calculated and corrected for. Section 6.3 explains how the systematic and random errors are interpolated in the final retrievals. Section 6.4 shows retrieval results for case study 1: Global tropospheric NO_2 for July 1997 along with a description of associated errors. Section 6.5 presents case study 2: Tropospheric BrO for November 1996 along with associated errors.

6.2 Diffuser Plate Correction

The diffuser plate artefact and methods for removing it were discussed in Section 1.2. This section will reintroduce the method that is applied in this new retrieval scheme and will show how it is calculated and removed.

6.2.1 Calculating the Diffuser Plate Correction

As described in Section 1.2 the diffuser plate correction used in this new retrieval scheme is similar to the reference sector method used to remove the stratospheric component of the total column[1]. As the artefact will manifest itself by producing day-to-day variability in the total column amount, if a location where the concentration is known to be constant is studied,

---

[1]where a part of the atmosphere assumed to be remote i.e. the middle of the Pacific ocean, is assumed to be purely stratospheric and therefore used to determine the stratospheric component of the total column.

variations can be assumed to be due to the diffuser plate error and can be accounted for and removed from the total column. It is assumed that in the middle of the Pacific Ocean (20°S to 20°N, 200°E to 240°E) the total column is uniform if measured at the same local time every day. For NO_2 and BrO it is reasonable to assume that the column over this area is constant and almost purely stratospheric with a small amount residing in the troposphere.

The diffuser plate artefact is present in the solar spectra that are used to calculate the sun–normalised radiances. The solar spectra are measured on roughly a daily basis, therefore a different diffuser plate correction needs to be determined for every solar spectra (Richter and Wagner, 1998). The diffuser plate correction is found by averaging the GOME slant columns over the Pacific ocean for every measurement taken using the same solar spectra, and then this is compared to the column calculated from the SLIMCAT stratospheric chemistry model plus the TOMCAT tropospheric chemistry model (where the stratosphere, represented by SLIMCAT, and the troposphere, represented by TOMCAT, are separated by the tropopause which varies from 15 km at the equator to 8 km at the poles). Therefore, making the assumption that the chemistry transport models are correct within some diagnosed error (see Chapter 3)[2]. The correction is calculated from cloud–free GOME data only. This minimises errors in modelling the slant column which are introduced by cloud.

Figure 6.1 shows the averaged GOME NO_2 slant column for each solar spectrum day and the model calculated column (CTM slant column) with associated errors (only due to stratospheric errors from Chapter 3). Assuming that the model is correct, the slant column varies widely from day–to–day, suggesting a large error that varies on this time scale. This is the affect of the diffuser plate artefact. Figure 6.2 is similar to Figure 6.1 but for BrO. Both figures show that the diffuser plate artefact causes a large variation in both NO_2 and BrO and highlights that this day-to-day (solar spectrum to solar spectrum) variability needs to be removed.

---

[2]TOMCAT errors will be negligible as the tropospheric column is very small over the Pacific and is insignificant compared to the stratospheric column amount. The SLIMCAT random error discussed in Chapter 5 will be the main source of error here.
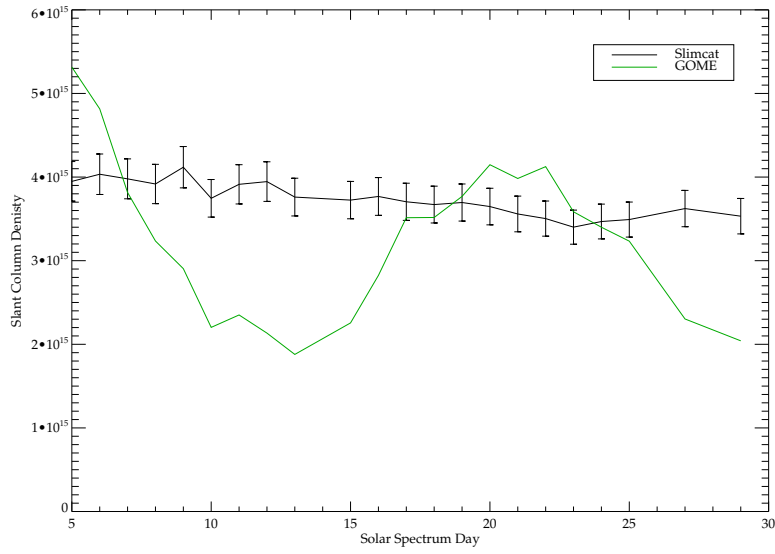
Figure 6.1: *The average NO$_2$ slant column (molecules cm^{-2}) for July 1997 over the Pacific (200° E–240° E) at the equator (-20° N–20° N) for every solar reference spectrum (along the x-axis). The black line shows the SLIMCAT slant column and the green line shows the GOME slant column.*
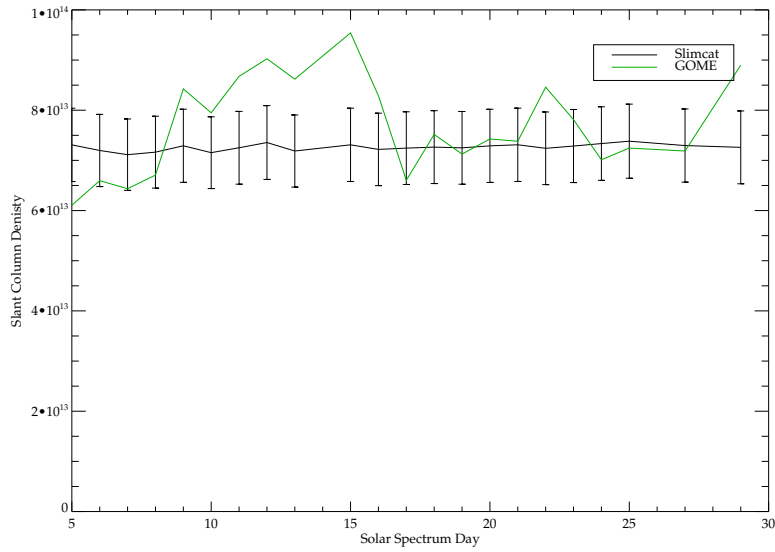
Figure 6.2: *As Figure 6.1 but for BrO.*

Uncorrected Tropospheric Column NO2 (x10¹⁵ molecules cm⁻²)

Uncorrected Tropospheric Column BrO (x10¹³ molecules cm⁻²)

(a)

(b)
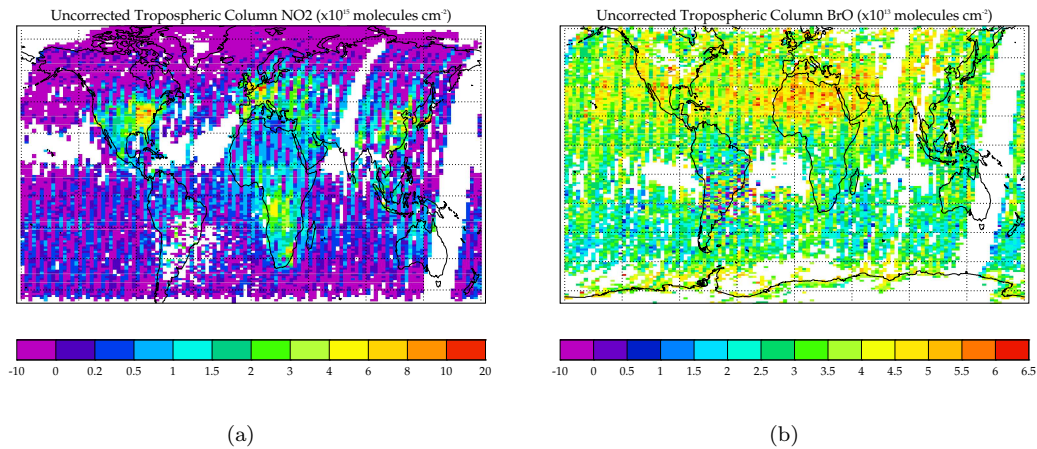
Figure 6.3: *(a) Weighted monthly mean uncorrected NO_2 slant column density for July 1997 from GDP slant columns ($\times 10^{15}$ molecules cm^{-2}), (b) weighted monthly mean uncorrected BrO slant column density for November 1996 ($\times 10^{13}$ molecules cm^{-2}).*

6.2.2   Applying the Diffuser Plate Correction

The diffuser plate correction can be applied in this retrieval scheme after the retrieval has been performed using the gain matrix, which describes how a change in the measurements affects the retrieval parameters (see Chapter 4).

$$\delta \mathbf{x} = \mathbf{G} \delta \mathbf{y} \qquad (6.1)$$

The gain matrix is therefore computed from the retrieval model output and used to correct the retrievals. As the retrieval is perfectly linear with respect to the slant column (only non linear with respect to sun–normalised radiance and surface LER) this becomes a simple linear problem.

Figure 6.3 shows the GOME tropospheric VCD for NO_2 and BrO (output from a run of the retrieval model where cloud is retrieved, discussed in more detail in Section 6.3) before the diffuser plate correction has been applied. The figure shows that the original NO_2 tropospheric VCD is strongly affected by the diffuser plate artefact. This manifests itself in the 'striping'[3] that can be seen in the figure. This 'striping' is due to the fact that global coverage is only achieved every 3 days (for GOME alone). Therefore, at least 3 different solar spectra are used in calculating these slant columns and each will have a different error associated with them. This will produce different biases for each day and therefore, a 'striping'. The corrected NO_2 tropospheric VCD can be seen in Figure 6.4(a) (where 'striping' is reduced).

---

[3]The term 'striping' is used because if the Figure 6.3 is studied for NO_2, looking in the direction of longitude, the slant column varies from blue to purple in a repetitive manner, which gives the impression of 'stripes'.

6.3 Determining Overall Errors, Systematic and Random, on the Final Retrieval

In Chapter 5 the key sources of error were determined. There are two different systematic errors used in this chapter. The first systematic error is relevant for a retrieval that explicitly models clouds. These systematic errors arise from the GRAPE systematic error in 'effective' cloud fraction that was derived from the sun–normalised radiance. The effect of these errors on the retrieved tropospheric VCD were estimated for

- solar zenith angles – 36.9, 39, 65.4,

- cloud fraction – 0.1, 0.6, 0.8,

- cloud optical depth – 0.3, 3. 30.,

- surface LER – 0.05, 0.8,

- cloud top height (km) – 1, 5, 8,

and stored in look–up–tables. To determine the error on the overall retrieval, these errors are then scaled accordingly for each GOME pixel, as discussed in Chapter 5 (i.e. linear except for cloud fraction). The second systematic error is relevant for a retrieval where cloud is not explicitly modelled. These systematic errors arise from the errors estimated in Section 5.5, where the errors associated with not modelling cloud when it is present in the scene were investigated. These errors were then stored in a look–up–table and then scaled onto the GOME pixels according to the solar zenith angle, cloud fraction, cloud optical depth, surface LER and cloud top height. The final error that is used here is the random error. The random error is the sum of the stratospheric column random error and the error due to the slant column error. The error due to the slant column error does not need to be estimated as it is determined in the retrieval run for each pixel and is stored in the covariance matrix ($\mathbf{S_x}$). The stratospheric column error is estimated in the same way to the systematic errors and is stored in look–up–tables and interpolated for each GOME pixel. The two random error sources can be added together according to Equation 5.4.

6.4 Case Study 1: Global NO$_2$ July 1997

The following section describes weighted monthly mean global NO$_2$ tropospheric vertical column density (VCD) produced by applying the retrieval scheme developed in this thesis. The retrieval has been performed using NO$_2$ slant column densities and sun–normalised radiances from the GDP dataset (described in Chapter 4) for two different runs of the scheme; one without using cloud in the radiative transfer part of the retrieval and therefore only using cloud–free pixels or pixels where cloud fraction is less than 20%, and the other using cloud and aerosol data from the GRAPE dataset. In all figures only the descending nodes are included. Figure B.1 shows

the global coverage if only cloud–free scene are considered. This illustrates why it is necessary to include cloud fractions up to 20% in the retrieval where cloud to not explicitly modelled.

There are a number of constraints that have been placed upon the data when forming the weighted monthly mean. These are

- the error on the measured slant column must not exceed 2×10^{15} molecules cm^{-2},

- the measurement cost on a single GRAPE retrieval must not exceed 30. This will highlight a cloud or aerosol retrieval that did not fit the optical model well and might indicate multi-layer or multi type - liquid and ice - clouds,

- at least half (50%) of the GRAPE pixels must be used to represent the GOME pixel,

- retrieved surface LER must not exceed 1 - this can occur within the retrieval but highlights a problem with the cloud parameters.

The monthly weighted mean is calculated by binning the data onto a $3°$ by $3°$ grid and then weighting each retrieval by the random error. This means that more weight will be given to retrievals that have a low error associated with them, i.e. are more accurate.

6.4.1 New NO_2 Retrieval Results

Figure 6.4(a) shows the monthly mean NO_2 tropospheric VCD for the retrieval in which cloud is modelled using GRAPE data. The first point to note is the lack of global coverage as there are some areas in which there is no data. One area of limited coverage is over the oceans (Pacific and Atlantic). These are due to either sun glint[4] or due to the lack of GRAPE pixels due to the reduced swath of ATSR–2 over the ocean (as described in Chapter 1). Figure 6.4(a) shows a number of features expected in tropospheric NO_2. Firstly, anthropogenic enhancements are observed over the USA, Europe and south east Asia, with between 3–20×10^{15} molecules cm^{-2}. Secondly, areas of NO_2 produced by biomass burning are observed over southern Africa, with concentrations of between 3–10×10^{15} molecules cm^{-2}. However, there are negative values observed over the oceans. These are due to errors in the diffuser plate correction, the stratospheric column and random noise.

Figure 6.5(a) shows the monthly mean NO_2 tropospheric VCD retrieved without explicitly including cloud in the radiative transfer (though some of its impact will be accommodated via the retrieval of LER). To provide slightly better coverage GOME pixels that GRAPE would be reported to have cloud fractions less than 20% are included along with cloud–free pixels. It can be seen that the retrieval column concentrations are a lot lower than in Figure 6.4(a) by around 5–10×10^{15} molecules cm^{-2}. As some cloud is still included in this figure, this highlights why it is necessary to include cloud in the retrieval. If a cloud is present but not considered it will mean that the radiative transfer equation will lower the total column. This is because the

---

[4]Sun glint is the specular reflection from the ocean surface.

cloud will obscure the lower atmosphere and therefore lower the tropospheric VCD (especially as most NO_2 resides in the boundary layer).

Figures 6.4(b) and 6.4(c) show the monthly mean fractional and absolute systematic error for the retrieval with GRAPE cloud, respectively, while, Figures 6.5(b) and 6.5(c) show the same for the retrieval without cloud, respectively (random error will be the same between these retrievals). The systematic error is reasonably small when cloud is included, varying from -25%–+5% in most cases. As cloud is more prevalent in the tropics in the summer months (July when the retrieval was performed) this is the area where the errors are largest, i.e. where cloud fraction is highest (see Figure B.3). In the retrieval without cloud, Figure 6.5 , the systematic error is a lot larger than in the Figure 6.4 varying from -70% –+5%. If Figures 6.4(a) and 6.4(c) are compared to Figures 6.5(a) and 6.5(c) it can be seen that the lower concentrations in the NO_2 can be attributed to this systematic error, i.e. by the neglection of cloud.

Previous studies (Leue et al., 2001; Richter and Burrows, 2001; Boersma et al., 2004) showed that cloud errors would cause large overall errors on the retrieved tropospheric VCD (between 50–100%). The errors reported here are different as previous studies, which used cloud parameters from the oxygen A-band of GOME. This has been suggested to be less accurate than cloud that is retrieved from the ATSR–2 instrument (see Chapter 1)[5]. An interesting study, that will be talked about more in Chapter 7, would be to compare this scheme directly with those based on the parameters retrieved from GOME measurements in the O_2 A–band.

Figures 6.4(d) and 6.4(e) show the monthly mean fractional and absolute random error, respectively. As stated before this is the same for both the retrieval with and without cloud. Figures B.2(a) and B.2(b) show the monthly mean random error due to only the ESD. If these are compared it can be seen that ESD is the dominant source of random error. As the random error is therefore dominated by the measurement error (error on the slant column) this is higher over the oceans, which is shown in Figure 6.4(d). However, in absolute terms (as the tropospheric VCD is larger over land) the error is larger over land. The random absolute error varies from 2-6 $\times 10^{15}$ molecules cm^{-2} over locations of enhanced NO_2, i.e. over industrial regions and the areas of biomass burning. Over remote areas the errors are almost equal to the tropospheric VCD.

Monthly mean NO_2 tropospheric VCDs from GOME can be determine to an accuracy of between 20–50% over land (where the column is larger). Over remote areas the NO_2 tropospheric VCD is not as accurate with errors around 100%. One reason for this could be that the NO_2 concentration is low over the oceans. This means that the noise on the measured slant column will be higher. Also the surface LER will be very low (below 0.1 for oceans), which will make the retrieval error larger (shown in Chapter 5).

---

[5]It should be noted that although GRAPE will represent the sub–GOME–pixel cloud better, the O_2 A–band cloud retrieval will measure the overall impact of the cloud on the light path and therefore may represent some quantities better than GRAPE as no averaging the data is required (Siddans et al., 2007).
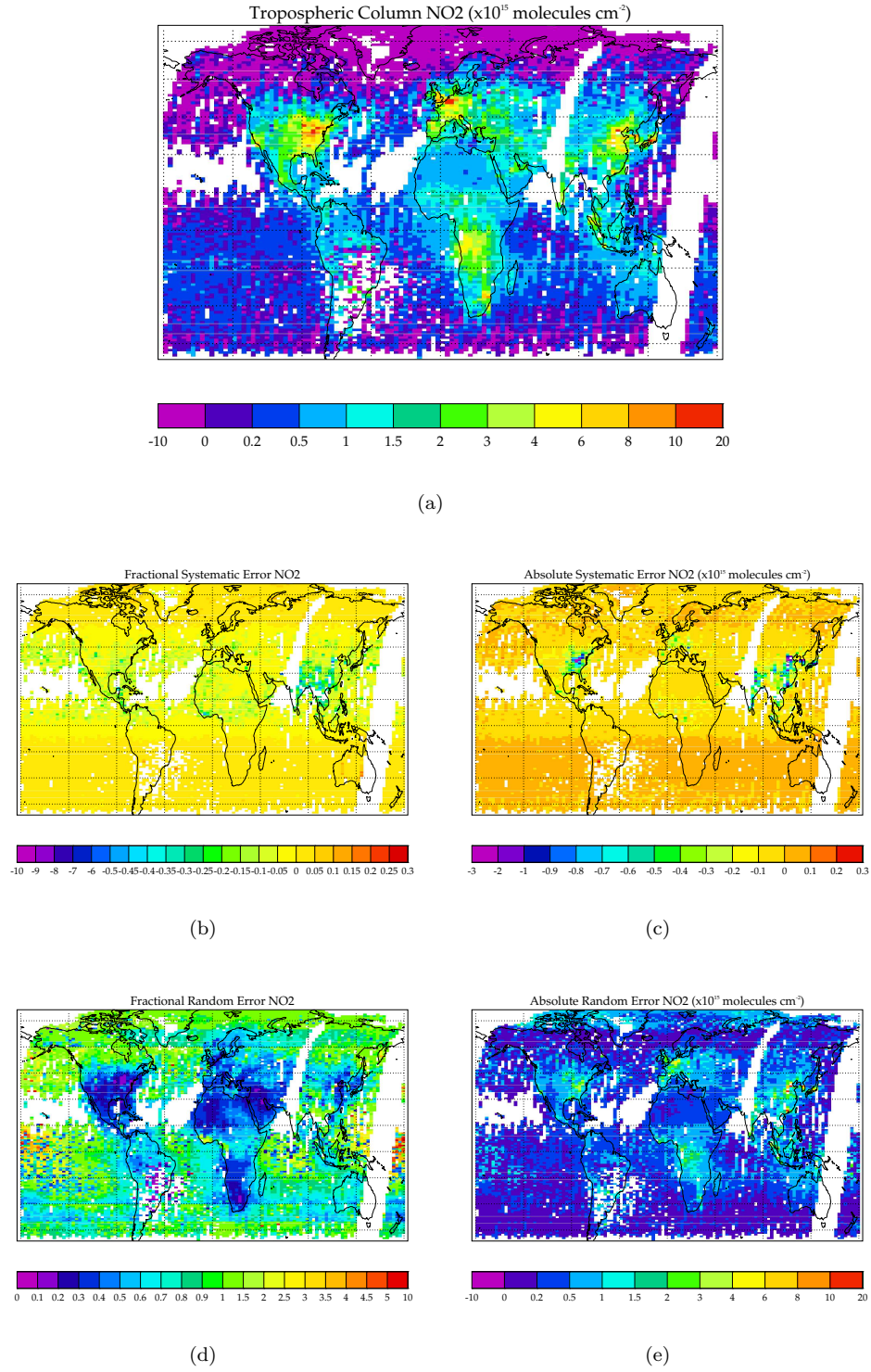
Figure 6.4: *Weighted monthly mean NO_2 for July 1997 from GDP slant columns for (a) tropospheric VCD ($\times 10^{15}$ molecules cm^{-2}), (b) fractional systematic error on the retrieved tropospheric VCD, (c) absolute systematic error in the retrieved tropospheric VCD ($\times 10^{15}$ molecules cm^{-2}), (d) fractional random in the retrieved tropospheric VCD, (e) absolute random error in the retrieved tropospheric VCD ($\times 10^{15}$ molecules cm^{-2}).*
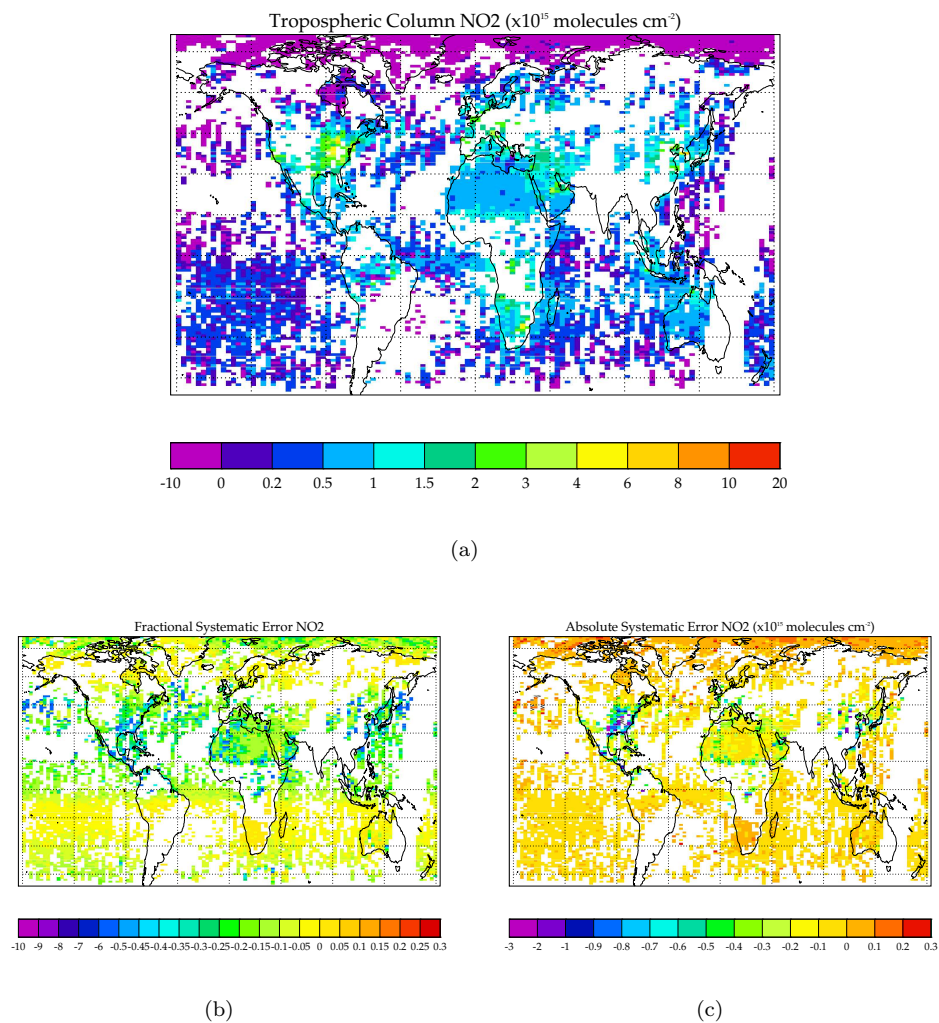
Figure 6.5: *Weighted monthly mean NO_2 for July 1997 from GDP slant columns when cloud is not explicitly modelled for (a) tropospheric VCD ($\times 10^{15}$ molecules cm^{-2}), (b) fractional systematic error on the retrieved tropospheric VCD, (c) fractional random error in the retrieved tropospheric VCD.*

In general there are some interesting points that can be drawn from these figures.

- By modelling cloud in a more sophisticated manner it is possible to get accurate retrievals of NO_2 even in partially cloudy scene, which is important for GOME given that there are very few cloud–free scenes,

- Systematic errors from the GRAPE cloud algorithm are small enough to not strongly affect the retrieval. An interesting comparison for further work here, would be to look at how this compares to systematic errors from other cloud retrieval algorithms (such as CRAG and FRESCO). As, in theory, ATSR–2 should have better cloud parameters than these schemes the errors there may be more significant.

- When not explicitly modelling cloud in a retrieval (when it is present), NO_2 tropospheric VCD is underestimated due to large systematic errors.

- Random error, where the measurement error is the dominate form, is reasonable large.

6.4.2 Comparison with TEMIS NO_2 Retrieval Scheme

Figure 6.6 shows the TEMIS NO_2 from July 1997 and associated errors[6]. The figure also shows the difference between this new retrieval scheme and the TEMIS retrieval. Studying Figure 6.6(a) and comparing to Figure 6.4(a) there are common NO_2 geographical patterns observed. Firstly, the areas of enhanced NO_2 are present in similar locations, e.g. USA, Europe, South Asia and biomass burning in southern Africa. Secondly, the concentrations in the two retrieval schemes are similar. Thirdly, there is more coverage in the TEMIS retrieval. This is because TEMIS does not rely on ATSR–2 data and can therefore exploit the full GOME swath. Finally, there is an increase in the concentration along the full longitude at about 30°S. Interestingly, Figure 6.6(c) shows that this is similar in the figures. However, it is slightly reduced in the scheme developed here. This is likely to be due to the stratospheric correction. This is because at this latitude the zonal symmetry in the stratosphere (which is assumed in most other retrieval schemes) will not hold, as it is the edge of the polar vortex. If this is not modelled accurately it will cause discrepancies in the tropospheric column.

Expanding on the second point, generally where there is enhanced NO_2 the new retrieval scheme has slightly higher values. This would be expected due to the better treatment of cloud. However, there are two exceptions to this. First, over South–East Asia and second, over South Africa. Over South–East Asia there are some optically thick clouds (see Figure B.3). These may not cause issues in this new retrieval scheme, but for the TEMIS retrieval the high cloud optical depth could introduce biases into the retrieval and explain the larger concentrations here.

Interestingly, there have been two studies (Richter et al., 2005; Van der A et al., 2006) which state that NO_2 satellite retrievals over South–East China overestimate the NO_2 that is expected

---

[6]Should be noted that TEMIS use a difference algorithm to get their slant columns, however the same fitting window is used which should limit the difference between the two retrieval schemes.

due to emission calculations. The poor treatment of cloud in these studies may be improved in this new retrieval scheme. Therefore, by repeating these studies using this new scheme the conclusions may change. This would have to be the subject of further work as the retrieval model has not been run for long enough the study.

For the second exception, the lower concentrations over South Africa, the reason for this is a little hard to see. The most obvious reason is that the new retrieval scheme (developed here in this thesis) may not have seen the enhanced NO_2 due to the GRAPE retrieval not being able to distinguish between smoke and cloud (as there is a biomass burning plume here). Also, Figure B.3 shows that although the cloud optical depth of South Africa is around 1 (low) the cloud top height is low and the fraction is high. In the TEMIS retrieval, where cloud is treated as an elevated reflecting surface, the high cloud fraction would lead to the surface having a higher reflectance then the surface. It may therefore be that the TEMIS retrieval saw more NO_2 than there really is. However, both these points are speculation and further analysis is needed to discover a reason for the discrepancy.

Figure 6.6(b) shows the error in the TEMIS NO_2 retrieval. Comparing Figure 6.6(b) with Figure 6.4(c) and 6.4(e), the systematic and random error in the new retrieval scheme, it can be seen that random errors in the new retrieval scheme are similar to those reported in the TEMIS retrieval. This is what would be expected if the error is dominated by noise on the measurements.

Figure 6.6(c) shows the difference between the TEMIS retrieval and the new retrieval scheme. The figure shows that generally the new retrieval scheme retrieves slightly higher values of NO_2, as has been discussed above. As stated in Chapter 1, the TEMIS retrieval scheme treats cloud as an elevated reflective surface. This will introduce biases into the retrieval. The more sophisticated treatment of cloud provided by this new retrieval scheme should give better results and hence mean that the new retrieval scheme has improved upon previous retrievals. However, this cannot be confirmed without validation of the results.

As the reported errors from the two retrieval schemes appear to be quite similar the question of whether this new retrieval scheme is an improvement on previous scheme arise. This new retrieval is an improved scheme, due to the more sophisticated method that is used. However, although the TEMIS retrieval scheme (and other schemes) state that their retrievals are quantitative they do not incorporate all error sources and therefore it is difficult to compare to. Validation of the scheme used in this work would help to determine which scheme is best. Although it should be noted that validation of satellite retrievals such as this is hard as there are not many datasets to compare to (Boersma et al., 2004). This is also an area of further work beyond this study.

Previous Retrieval Tropospheric Column NO2 (x10¹⁵ molecules cm⁻²)

(a)

Previous Retrieval Tropospheric Column Error NO2 (x10¹⁵ molecules cm⁻²)

Difference Modelled - Previous NO2 molecules cm⁻²
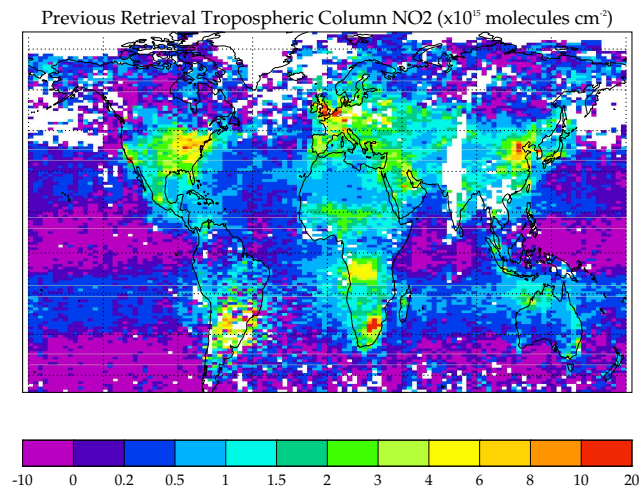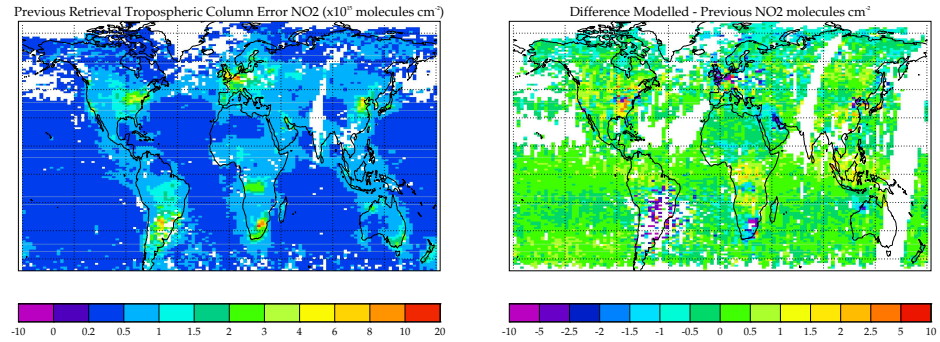
(b)

(c)

Figure 6.6: *Weighted monthly mean for July 1997 (a) TEMIS retrieval of NO_2, (b) errors in the TEMIS retrieval of NO_2 and (c) difference between TEMIS retrieval and the new retrieval scheme.*

6.5 Case Study 2: Tropospheric BrO in November 1996

The following section will describe a retrieval that has been performed to produce global weighted monthly mean tropospheric BrO vertical column densities (VCD), using the new retrieval scheme developed within this thesis. The retrieval has been performed using BrO slant column densities and sun–normalised radiances from the SAO dataset (described in Chapter 1) for two different runs of the scheme; one without using cloud in the radiative transfer part of the retrieval and therefore only using cloud–free pixels or pixels where cloud fraction is less than 20%, and the other using cloud from the GRAPE dataset. In all figures only the descending nodes are included. Figure B.1(b) shows the global coverage if only cloud–free scene were considered. This illustrates why it is necessary to include cloud fractions up to 20% in the retrieval where cloud to not explicitly modelled. The same constraints have been applied to this data as were for NO_2.

6.5.1 New BrO Retrieval Results

Figure 6.7(a) shows the monthly mean BrO tropospheric VCD from the run with cloud from the GRAPE dataset. The first point to note is that the retrieval does suggest that there is an omnipresence of BrO in the troposphere, with between 0.5–1.5×10^{13} molecules cm^{-2} at the equator and in the southern hemisphere, and between 1–3×10^{13} molecules cm^{-2} in the northern hemisphere and at the south pole. The second point to note is the poor coverage over the Antarctica. This is highlighted in Figure B.4(a) which shows the BrO column over the the south pole for the same run. The coverage here is very poor. The apparent enhancement in the BrO column due to the break up of sea ice can be seen, as the BrO column retrieved here is higher than that elsewhere in the southern hemisphere. The poor coverage is due to the lack of GRAPE data. This is due to the difficulty to retrieve cloud and aerosol at such extreme solar zenith angles and over such a bright surface. Figure 6.8(a) shows the monthly mean BrO retrieval when cloud is not explicitly modelled. The vertical column BrO is similar to that in Figure 6.7(a) but the coverage is not as full due to the lack of pixels with clouds. However, the concentrations of BrO are similar (not as in the NO_2 plots).

Figures 6.7(b) and 6.7(c) show the monthly mean fractional and absolute systematic error for the retrieval with cloud, respectively and Figures 6.8(b) and 6.8(c) show the monthly mean fractional and absolute systematic error for the retrieval without cloud, respectively. First, it should be noted that, as in Chapter 5, the systematic errors for BrO are higher than they were for NO_2. This is most due to the lower concentrations of BrO in the troposphere and the larger error in the stratospheric column. Second, the systematic error here is most notable near the equator but varies from -20% –+5%. In the no cloud retrieval, although the actual retrieved tropospheric column is similar to that when cloud is retrieved, the systematic errors are very large, ranging from -50% – +10%. This also highlights that not including cloud in the retrieval will produce significant biases in the total column that is retrieved.

Figures 6.7(d) and 6.7(e) show the monthly mean fractional and absolute total random error

on the BrO retrieval, respectively. The random error ranges from 30–70% in most locations with higher errors ~100% over the south anomaly.
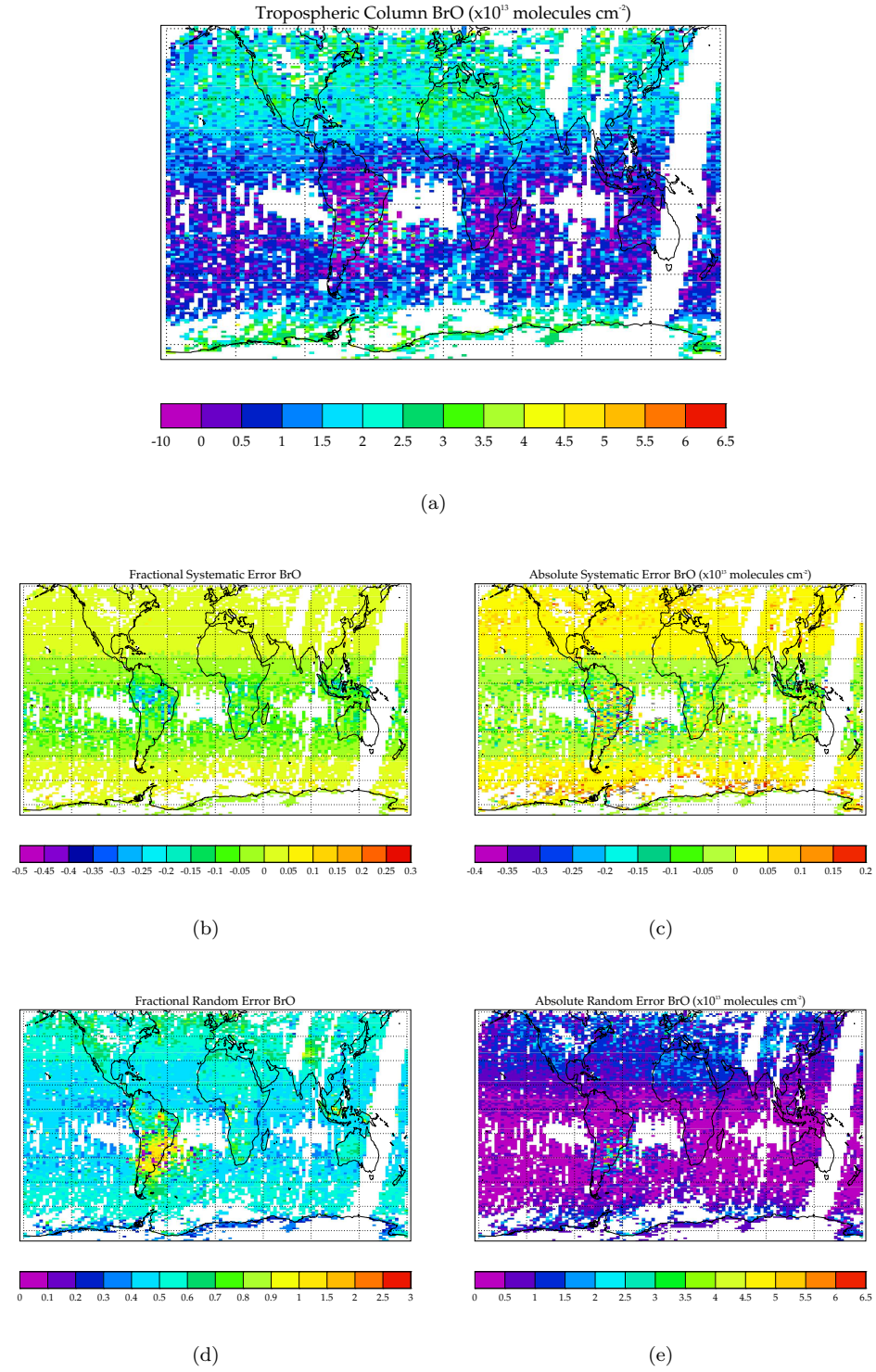
Figure 6.7: *Weighted monthly mean BrO for November from SAO slant columns for (a) tropospheric VCD ($\times 10^{13}$ molecules cm^{-2}), (b) fractional systematic error on the retrieved tropospheric VCD, (c) absolute systematic error in the retrieved tropospheric VCD ($\times 10^{13}$ molecules cm^{-2}), (d) fractional random in the retrieved tropospheric VCD, (e) absolute random error in the retrieved tropospheric VCD ($\times 10^{13}$ molecules cm^{-2}).*

219
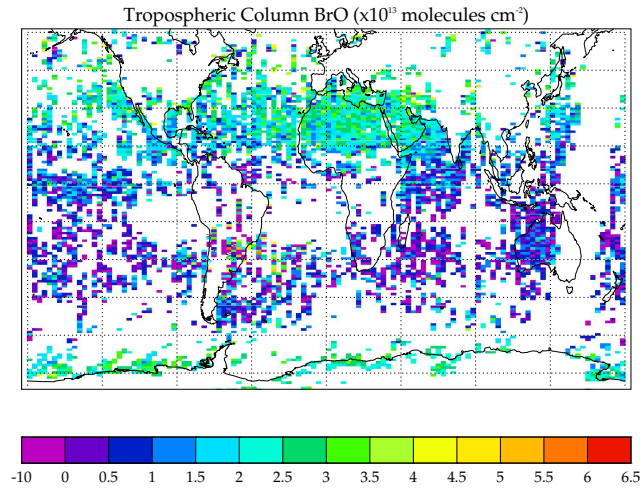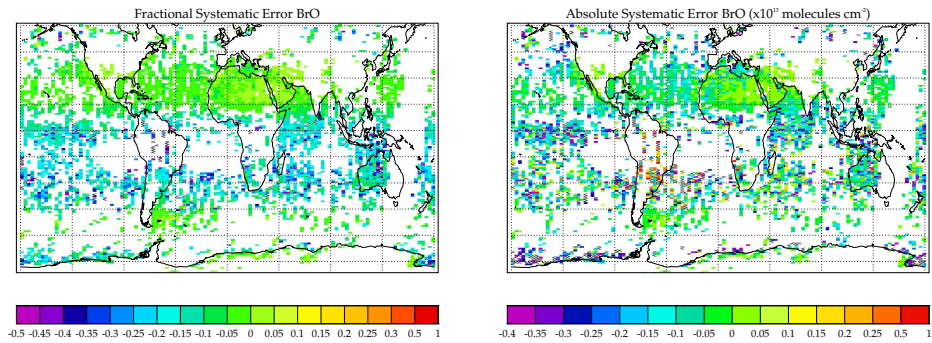
Figure 6.8: *Weighted monthly mean BrO for November 1996 from SAO slant columns when not explicitly modelling cloud for (a) tropospheric VCD ($\times 10^{13}$ molecules cm^{-2}), (b) fractional systematic error on the retrieved tropospheric VCD, (c) absolute systematic error in the retrieved tropospheric VCD ($\times 10^{13}$ molecules cm^{-2}).*

6.5.2 Comparison with Bremen BrO Retrieval Scheme

Figure 6.9 shows the global monthly mean BrO[7] from November 1996 from a previous retrieval by the University of Bremen and the difference between this new retrieval scheme and the previous retrieval. Figure 6.9(a) shows that the BrO column varies from 0–1×10^{13} molecules cm^{-2} in the southern hemisphere and from 1.5–2.5×10^{13} molecules cm^{-2}, with values between 3–5×10^{13} molecules cm^{-2} over Antarctica. The global coverage is much better in the Bremen retrieval than in this new scheme due to there being more GOME pixels available when not using collocated ATSR–2 pixels.

Figure 6.9(b) shows that the new retrieval scheme has slightly higher BrO columns than the previous retrieval, in the northern hemisphere, by around 0–2×10^{13} molecules cm^{-2} and slightly lower BrO columns in the southern hemisphere. There are a number of reason why the new retrieval scheme produces more tropospheric BrO than the Bremen scheme. Firstly, the tropopause could be wrong, therefore meaning that some stratospheric BrO is included in the tropospheric column. Secondly, the new retrieval scheme treats clouds in a more sophisticated manner, that leads to a more accurate value of the BrO tropospheric loading. Also the stratospheric BrO in each case will be different as the Bremen retrieval assumes a zonal systemic linearly increasing BrO stratosphere (see Richter et al., 1998). It is difficult to determine which retrieval scheme produces more accurate results without validation, which will be performed in further work. However, this new retrieval scheme provides a more sophisticated retrieval with a quantitative error analysis which has not been performed before. Therefore, with further validation this will provide more insight to tropospheric BrO retrieval from GOME.

| (a) | (b) |
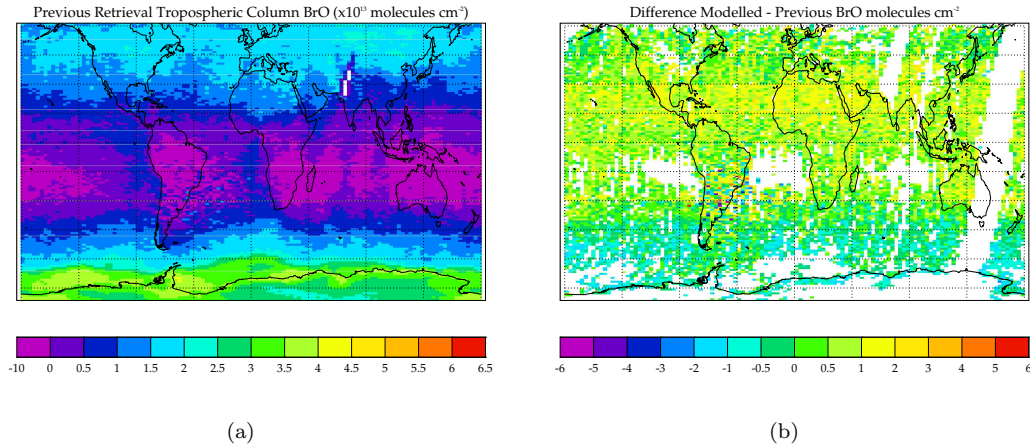
Figure 6.9: *Weighted monthly mean BrO for November 1996 ($\times 10^{13}$ molecules cm$^\circ$) (a) from the Bremen retrieval scheme, (b) difference between this new retrieval scheme and the Bremen retrieval scheme.*

In most previous retrievals of BrO the stratospheric component is assumed to be zonally symmetric. However, in reality over the poles stratospheric BrO is not zonally symmetric. This could be a significant source of error in the BrO retrievals. Figure 6.10 shows two plots. The

[7]supplied by the Andreas Richter at the University of Bremen

first is a plot of SLIMCAT total column BrO minus the zonal mean SLIMCAT column. The second plot shows the tropospheric BrO retrieved from GOME (retrieved by the Bremen retrieval group). Both plots are monthly means for March 1997. If the two plots are compared it can be seen that in removing a zonal mean stratospheric column (i.e. assuming that the stratosphere only varies with latitude not longitude) similar variations in the BrO column are observed as is seen in the GOME retrieval, even though in the SLIMCAT model there is no tropospheric BrO. This is mostly evident over Hudson bay area (Canada). Over the south pole this is not as evident (this is not shown), however this highlights the importance of interpreting the stratospheric component as accurately as possible.

(a)   (b)
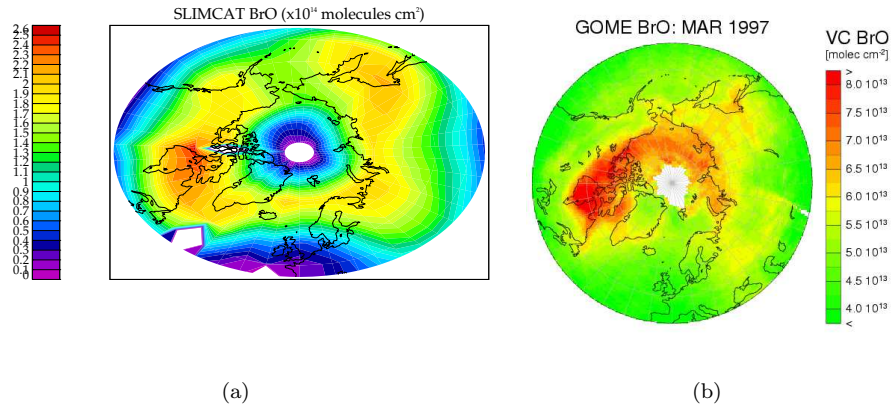
Figure 6.10: *Monthly mean BrO for March 1997 (a) SLIMCAT BrO minus the zonal mean stratospheric column, (b) GOME retrieved tropospheric BrO (taken from `http: //www.doas-bremen.de`).*

6.6 Conclusion

This chapter has shown how the diffuser plate artefact can be calculated and corrected for. It was shown that this causes large variations in both the NO_2 and BrO slant columns and when compared to CTM results with associated errors the variation is still much larger than the errors. The method for calculating a correction for this bias was shown, where a remote area of the troposphere is used to calculate diffuser plate correction for each solar spectra. It is possible to do this as the diffuser plate artefact is constant for each solar spectra. The correction was also shown to remove the 'striping' in the monthly mean figures which is how the artefact can be seen due to the different corrections needed from day to day.

The first case study (NO_2 global distribution) showed that by including cloud in the retrieval the retrieved tropospheric NO_2 is similar to that seen by previous studies. The error analysis showed that included cloud significantly improved the systematic biases and that the main source of error in this retrieval scheme is the random noise on the measured slant column. As it is the random error that is large for NO_2 retrievals, an interesting question is why is it important to include a better representation of cloud and the stratosphere in these retrievals. The answer is that cloud will significantly bias the overall retrieved tropospheric column and therefore the results will not be as accurate, i.e. the radiative transfer calculation will be in error, if the wrong cloud parameters are used. In the case of the stratospheric component and the errors associated with it, it should be noted that the random error in the stratospheric component was not very large as a lot of work was done to reduce this. Therefore, in other retrieval schemes this is not the case, and therefore highlights why this new retrieval scheme is more sophisticated than those that have been used previously.

In the case of global BrO retrievals it is possible to retrieve BrO concentrations that are similar to those retrieved by other schemes. The main features observed are the omnipresence of BrO in the troposphere of between 1–$3{\times}10^{13}$ molecules cm^{-2}, and the enhancement due to the break up of sea–ice. However, coverage over Antarctica is poor due to the lack of GRAPE data and therefore it is difficult to draw any major conclusions here. It was also noted that the systematic errors vary from -50% –+10% with random errors varying up to over 100%.

Further work on BrO would involve trying to reduces these errors to enable a more accurate retrieval of BrO.

Chapter 7

Summary and Conclusions

7.1 Motivation

The aim of this thesis was to develop and test a new retrieval scheme for determining tropospheric trace gas composition from the Global Ozone Monitoring Experiment (GOME). This has been achieved by framing the retrieval as an inverse problem.

The two main novel steps in this method were to:

- Determine the stratospheric component of the trace gas of interest by using a Chemical Transport Model (CTM) constrained by chemical data assimilation of long–lived tracers.

- Combine cloud and aerosol data from the Along Track Scanning Radiometer (ATSR–2) instrument to provide a more accurate model for the atmospheric state.

This thesis has described the theoretical basis, the implementation of the theory, the initial validation, and exploitation of the results of this retrieval scheme.

The main difference between the new scheme developed within this thesis and previous methods (see Chapter 1) is that the retrieval is framed as an inverse problem. In principle this would be a profile retrieval, e.g. NO_2 from the GOME spectra. However, in reality the lack of vertical resolution means that it is only possible to retrieve a tropospheric total column. This scheme also retrieves surface Lambertian Equivalent Reflectance (LER) at the appropriate wavelength, as this controls the path of radiation through the atmosphere.

The retrieval scheme measurement is the slant column density (assumed to be a direct measurement) meaning there is no need to fit the radiance spectra. A single sun–normalised radiance in the fitting window for a slant column is a good equivalent to fitting a full GOME spectrum but avoids a lot of the difficulty which has been handled well already by many other groups (see Chapter 1). As this retrieval problem only has two measurements, and two highly independent

unknowns, it does not require any optimal estimation *a priori* constraint but it is non-linear due to the effect of surface LER on the photon path lengths. The cloud and aerosol parameters used in the retrieval are taken from the GRAPE dataset. The same optical models are used in this retrieval as are used in the GRAPE retrieval (see Chapter 2) to allow the same representation of the scene. The stratospheric component of the total column is calculated from the SLIMCAT 3–D CTM, constrained by chemical data assimilation of stratospheric species (see Chapter 3).

The advantage of using this method is that it is theoretically sound and based on well understood radiative transfer calculations. The modelled errors are propagated through the method.

7.2   Results

Chapter 3 used a CTM constrained by chemical data assimilation in order to determine an accurate stratospheric component of the total trace gas columns observed from the GOME instrument. The chemical data assimilation was shown to improve the distribution of long–lived tracers, which impacted upon the short–lived and reservoir species, improving these too. For this work NO_2, which is a short–lived species, was investigated and it was shown that the chemical assimilation improved both its total column and profile in the stratosphere, when compared to independent data. Also, a long–term CTM run (1991–2003) was used to investigate how the chemical data assimilation impacted the long–term trend in NO_2. Comparisons with long–term ground based datasets showed that the chemical data assimilation scheme acted to improve the long–term trend from the CTM. From this it was possible to conclude that the previous observed discrepancies in long–term trends in NO_2 in CTMs was due to transport problems (which the assimilation scheme acts to improve). In this work, this is due to inaccuracies in ECMWF ERA-40 data. Also the impact of the chemical data assimilation on O_3 was investigated. Although the scheme assimilates O_3, the agreement with ozone–sondes shows that O_3 is overestimated near the tropopause at high latitudes. This is partly due to the coverage of the assimilation data (where there is little HALOE data at the higher latitudes) and model transport in the Upper Troposphere Lower Stratosphere (UTLS) region. Furthermore, in Chapter 3 the errors in the stratospheric column are quantified by comparing the NO_2 stratospheric column with ground based measurements to determine a bias. This bias was then used with the retrieval results to obtain a more accurate tropospheric column. For BrO there was no bias calculation needed as the CTM was forced so that it agreed with the work of Sinnhuber et al. (2002), which reported an error of $\pm 10\%$ in the concentration of BrO.

Chapter 4 discussed the development of the Retrieval Model and tested the Forward Model. The Forward Model testing involved analysing the photon path lengths that the model would reproduce for different atmospheric conditions and highlighted how different Forward Model parameters would affect the Retrieval Model. The main parameters that would affect the results of the Forward Model were shown to be cloud optical depth, cloud top height, cloud fraction, multi-layered clouds and surface LER. Also, Chapter 4 investigated the method by

which the GRAPE parameters could be averaged onto the GOME pixel. Although many complex averaging techniques were tested the most basic technique (averaging the scene into two zones; one cloudy and one cloud free) performed just as well as the other techniques but was computationally less expensive. Further to this the GRAPE parameters themselves have a measurement cost for each pixel. This is a measure of how well the retrieved parameters fit the optical models used in the retrieval. The measurement cost was used to obtain the best possible fit of the GRAPE parameters in the GOME pixel. This means that the GRAPE pixels were averaged such that only pixels with a maximum measurement cost of 30 were used in the retrieval.

Chapter 5 discussed the errors associated with the retrieval scheme. Firstly, the Estimated Standard Deviation (ESD) was determined. This is the minimum random error in the retrieval parameters (here only considering the tropospheric column) and comes from the measurement error. The ESD was small for low cloud fraction scenarios. However, for high cloud optical depths, cloud top height and cloud fractions this errors grows to up to 100%. This chapter investigated whether it is possible to improve the retrieval by modelling cloud, by determining the error on the retrieved column due to not modelling cloud. It was shown that the retrieval could be improved by correctly modelling cloud. Finally, this chapter investigated the impact of forward model parameter errors both random and systematic upon the final retrieved column. Chapter 4 showed that cloud parameters and the stratospheric column have the largest impact on the tropospheric VCD. Following on from that, Chapter 5 performed detailed analysis of the errors relevant to the scheme. The GRAPE systematic and random errors were estimated from sun–normalised radiance comparisons (modelled and observed). The differences between modelled and observed sun–normalised radiances are mostly due to cloud errors (fraction, optical depth and particle size) but also had a small contribution from the cloud and aerosol optical models and aerosol parameter errors. This analysis showed that cloud systematic errors where significant however, the cloud random errors were smaller and therefore not important. Stratospheric column errors were random as the systematic error was removed in Chapter 3. It was shown in Chapter 4 that stratospheric column errors where more significant when the scene is cloudy. When the stratospheric error was converted into tropospheric column errors it was also shown that the errors were smaller than the ESD. Chapter 5 therefore concluded that systematic cloud errors, and the ESD and stratospheric random errors were the significant errors on the final retrieval product.

Chapter 6 discussed the diffuser plate artefact and showed that it could be corrected for easily after the Retrieval Model has been run. Example NO_2 retrievals were performed for July 1997. Comparisons with a well established GOME retrieval scheme showed that NO_2 tropospheric VCD retrieved by this new scheme generally produced larger maximum values in polluted areas. This was explained by the more sophisticated treatment of cloud in the scheme.

In the case of BrO two retrievals were also discussed, investigating a retrieval with and without cloud. The values of the retrieved BrO were similar between the two runs, suggesting that cloud does not bias the results in the same way as for NO_2. However, the overall systematic error when not explicitly modelling cloud is very high, as for NO_2. Overall, the BrO retrievals were

high in error and therefore it is unlikely that any definitive conclusions could be drawn from the results without further investigation. The importance of the stratospheric correction was highlighted here, as it was shown that stratospheric BrO has similar spacial patterns to previous retrievals of tropospheric BrO. Further improvements that could be added will be discussed in the final section.

7.3 Future Work

The potential for a retrieval scheme of this type is large. The initial testing of this Retrieval Model has shown that the results are comparable to previous retrieval schemes and appear to change the overall results in the direction that is expected i.e. poor treatment of cloud leads to an underestimation of the NO_2 tropospheric column. Therefore, the next stage of this work would be to run the Retrieval Model for the full GOME dataset and then compare to independent observations.

In this work the species NO_2 and BrO have been investigated. These is much work that can be carried out with quantitative retrievals of these gases including studies of NO_2 trends due to anthropogenic pollution, studies of NO_2 from biomass burning events, the enhancement of Antarctic BrO in the spring and the source of tropospheric BrO in the free boundary. However, there is also a need to expand this scheme to other gases. There is the potential to develop this scheme for gases such as O_3 and HCHO, which have reasonably large tropospheric columns, but also for OClO and SO_2 which have very small tropospheric columns.

As stated in the previous section, BrO retrievals in this work have had large errors associated with them suggesting that little information can truly been drawn from the results without further validation. However, if the stratospheric column could be determined to a greater degree of accuracy this would improve the retrieval.

Another area of interest would be to use cloud information from GOME (from the O_2 A band) to compare to the GRAPE cloud and investigate the impact of this upon the retrieval vertical columns. Also it should be noted that there were systematic differences found between the GOME radiances and those simulated based on the GRAPE data. Investigating the reasons behind these discrepancies and trying to improve the agreement would also be an area of further research.

As well as comparing the A–band derived cloud from GOME, it would be good to consider combining information from the A–band with that from ATSR–2. If a consistent fit to both ATSR–2 and GOME A–band spectra could be obtained then that would represent the best available cloud/aerosol representation for the GOME scene. Cost (of simultaneous ATSR–2/GOME fit) could be used to ensure that scattering profiles are consistent with all measurements and so presumably trace–gas retrieval reliability. Use of GOME might also improve the quality of the ATSR–2 cloud/aerosol products themselves.

As the title of this work suggests, there are other nadir viewing spectrometers similar to GOME;

SCIAMACHY, OMI and GOME2, to which this retrieval scheme could be applied. The large ground pixel of GOME makes the coverage of this scheme low. However, for the younger instruments their ground pixels are smaller and therefore should produce better coverage. The wealth of new data from these spectrometers will justify continued efforts in improving the quantitative retrieval of tropospheric gases.

Appendix A

Sensitivity Study Figures
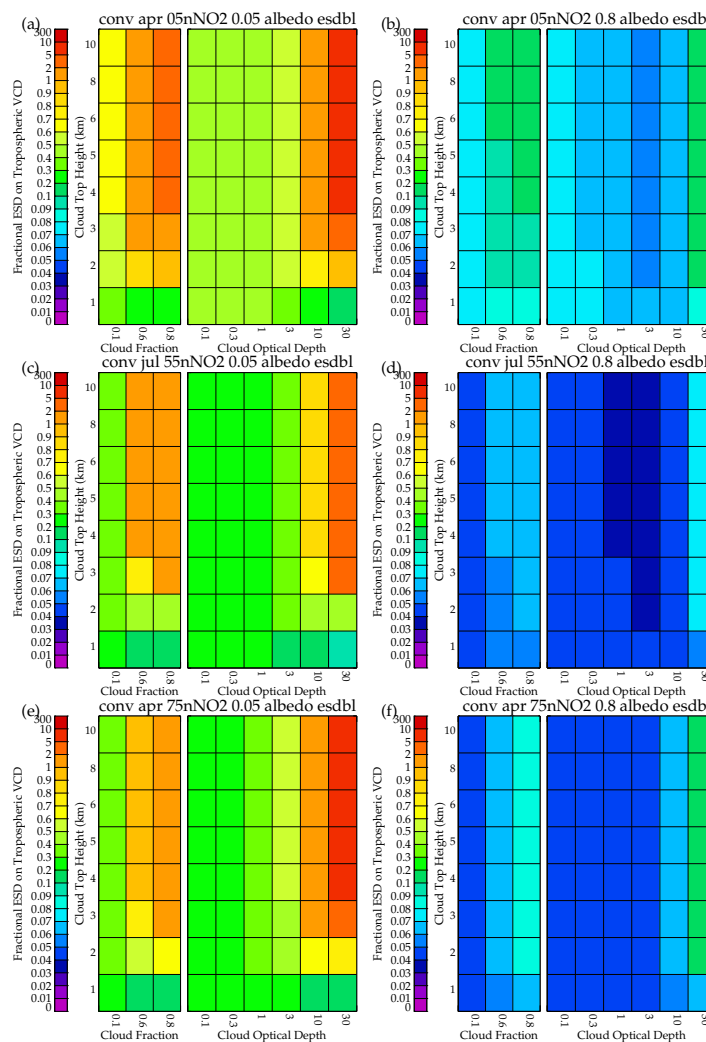
Figure A.1: *As Figure 5.3 but for boundary layer scaling.*
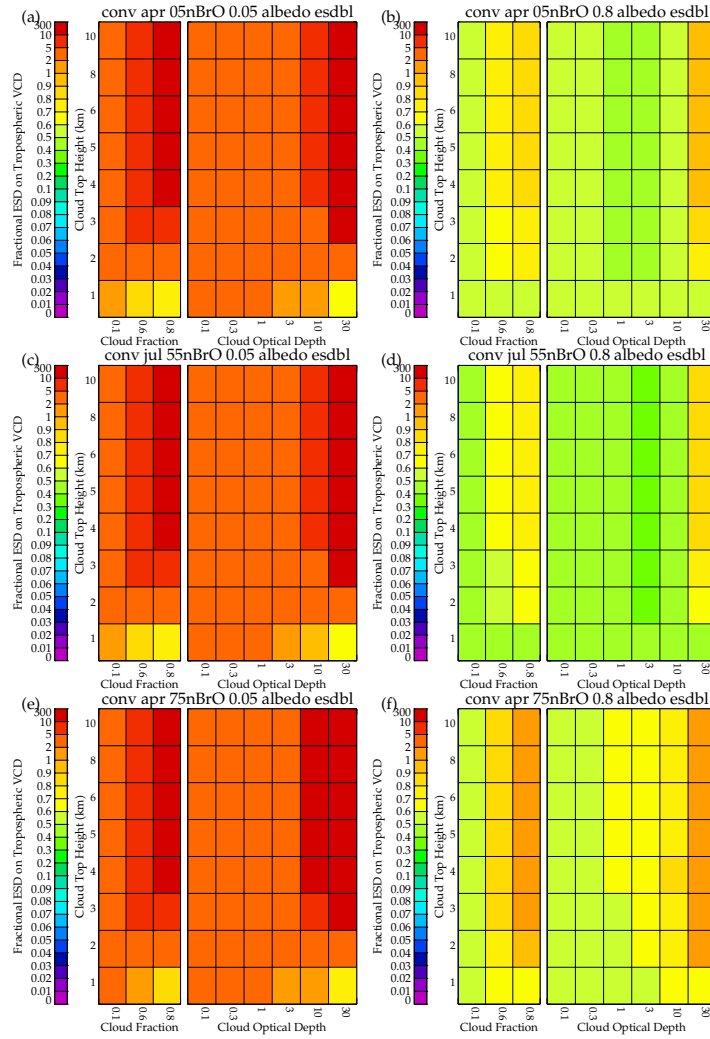
Figure A.2: *As Figure A.1 but for BrO.*
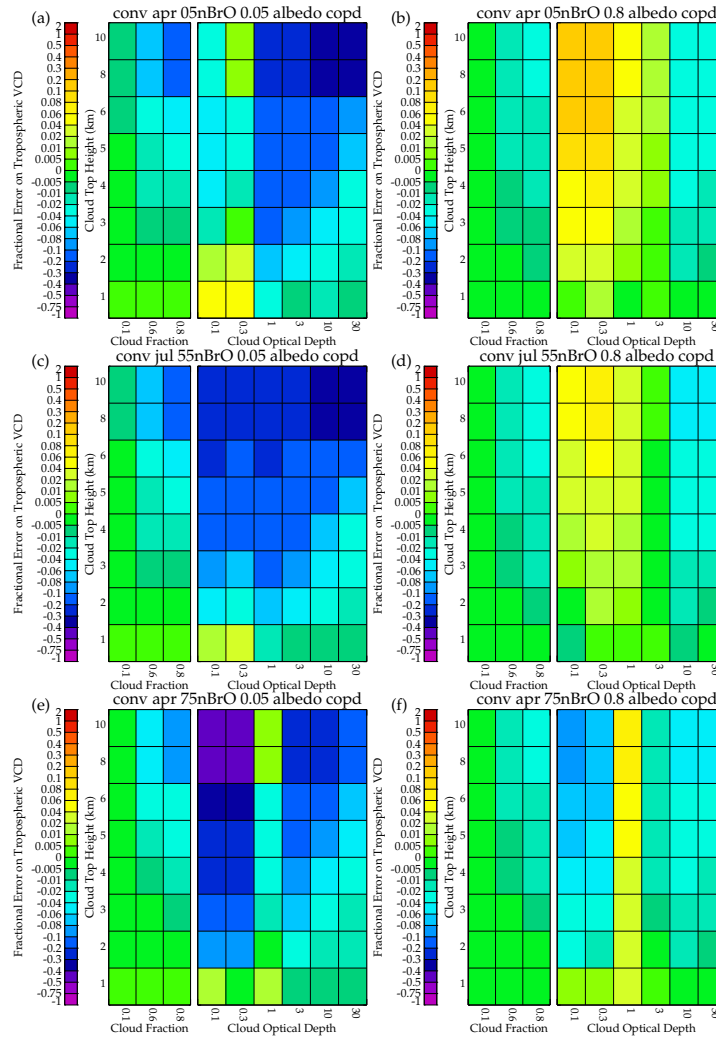
Figure A.3: *The fractional error in the tropospheric VCD for BrO (molecules cm^{-2}) for a -2 error in cloud optical depth for (a) October 5° N for a surface LER 0.05, (b) October 5° N surface LER 0.8, (c) October 55° N for a surface LER 0.05, (d) October 55° N for a surface LER 0.8, (e) October 75° N for a surface LER 0.05, and (f) October 75° N for a surface LER 0.8. In each panel there are two plots, the left plot shows the error when cloud fraction and cloud height are varied, for a cloud with an optical depth of 30. The right plot shows the error when cloud optical depth and cloud height are varied for a cloud fraction of 1.*
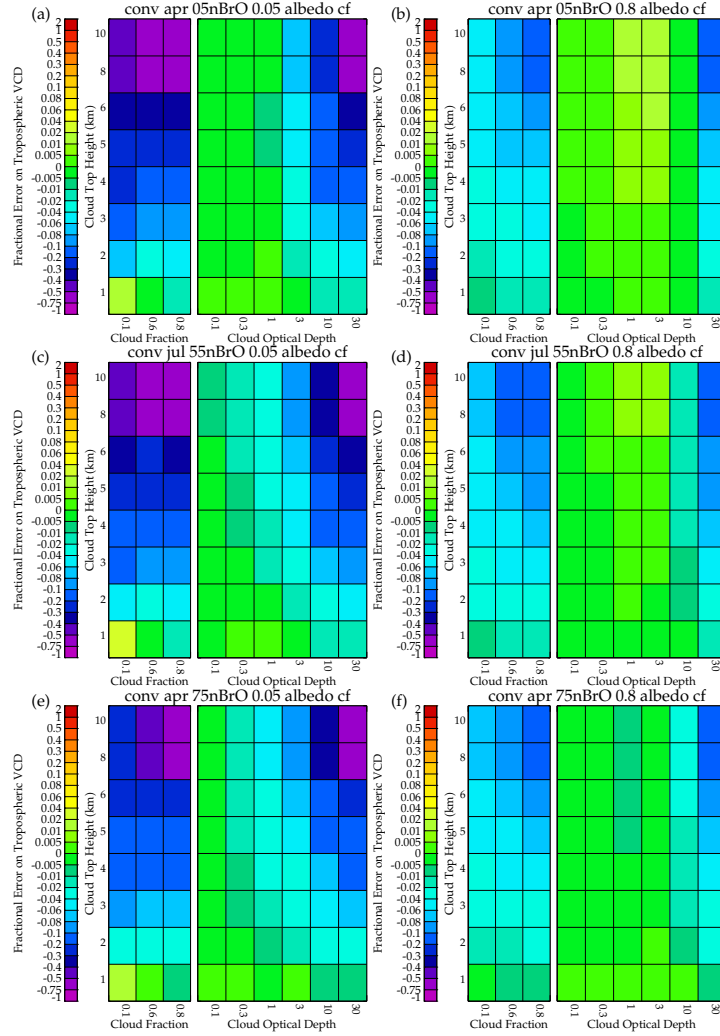
Figure A.4: *As Figure A.3 but for a -0.1 error in cloud fraction.*
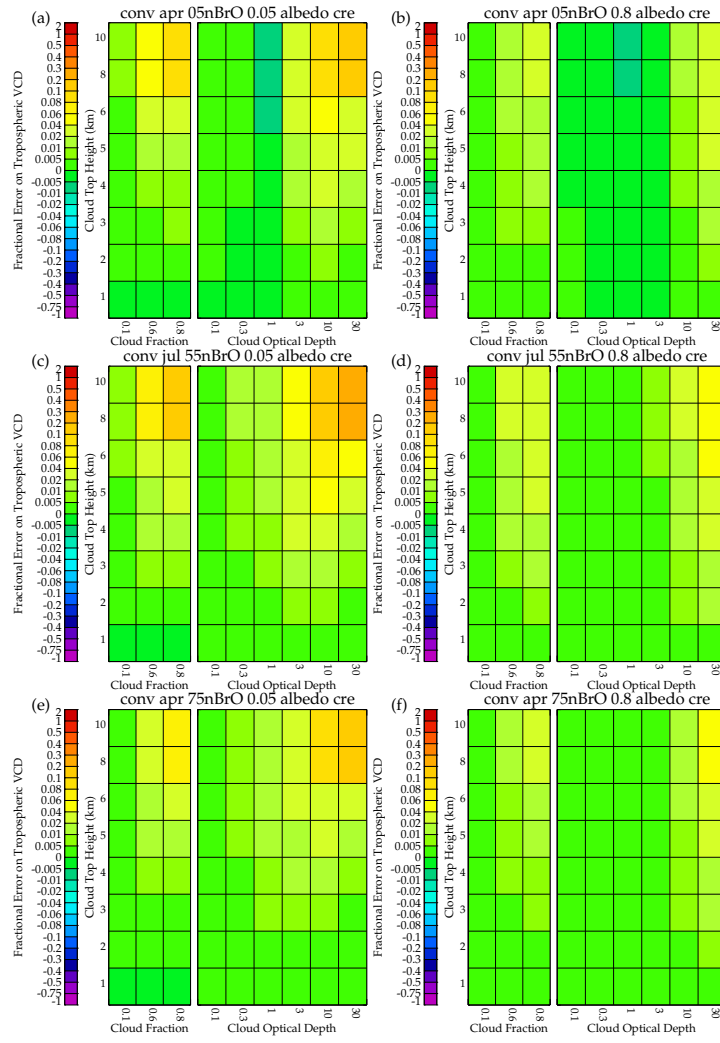
Figure A.5: *As Figure A.3 but for a 5µm error in liquid cloud particle size.*
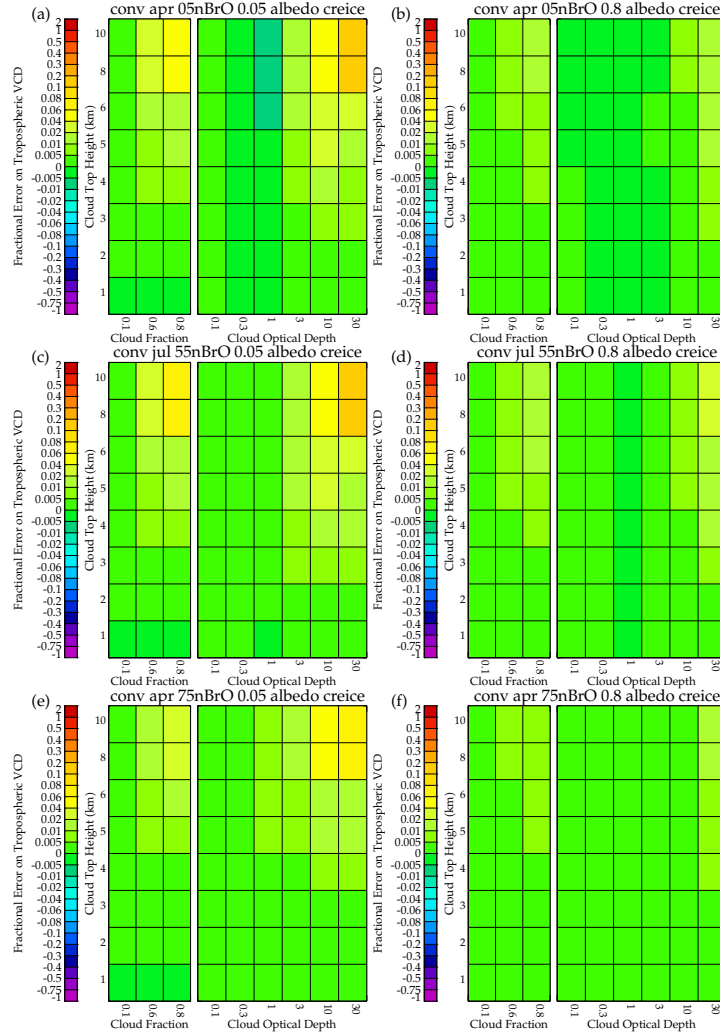
Figure A.6: *As Figure A.3 but for a 5µm error in an ice cloud particle size.*
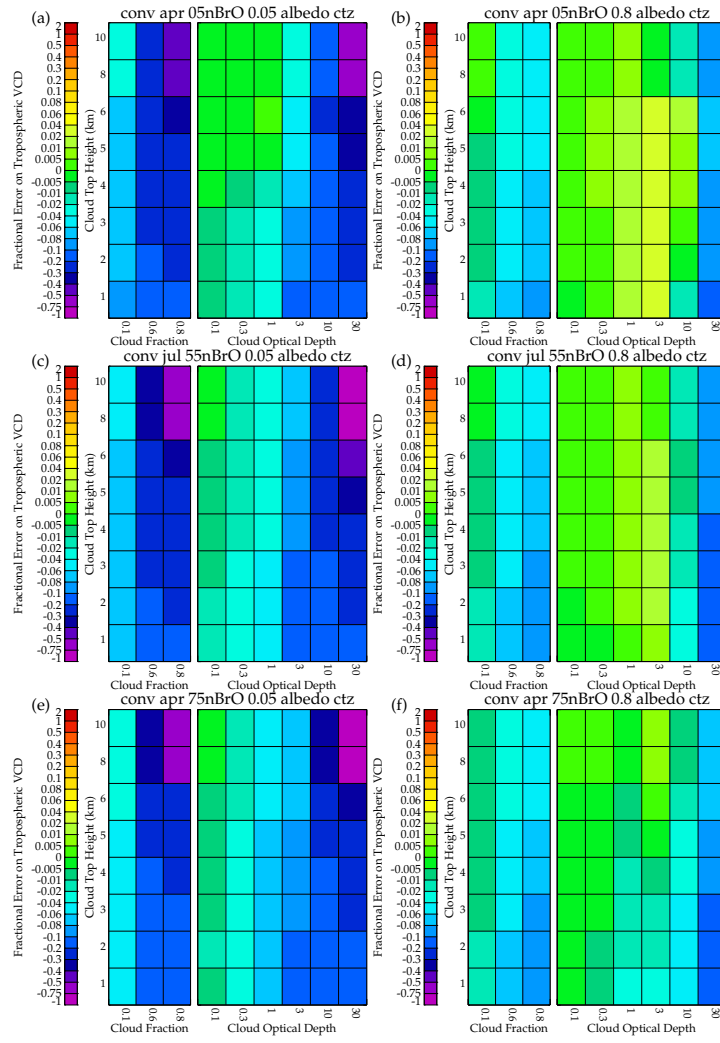
Figure A.7: *As Figure A.3 but for a 0.5 km error in cloud top height.*

Appendix B

Retrieval Results Figures
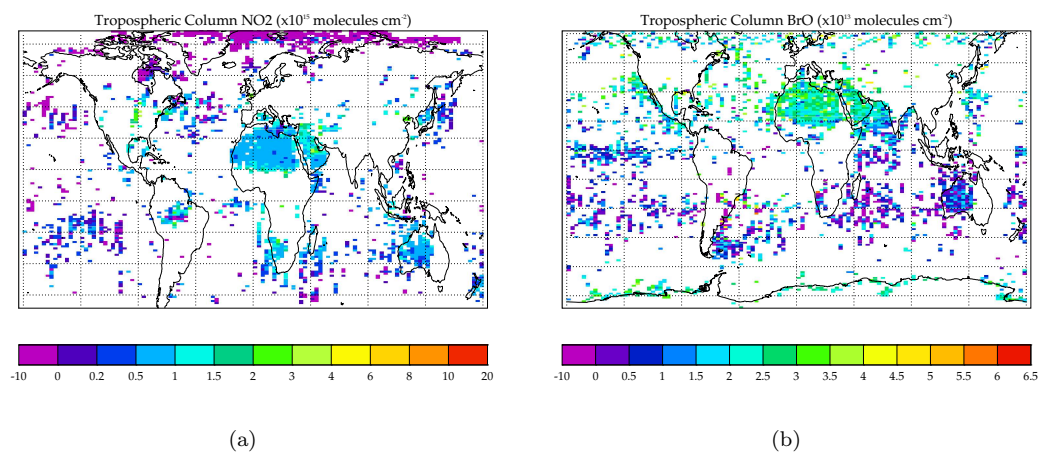
Figure B.1: *Weighted monthly mean (a) NO$_2$ tropospheric VCD for July 1997 with only cloud–free scenes (b) BrO tropospheric VCD for November 1996 with only cloud–free scenes.*
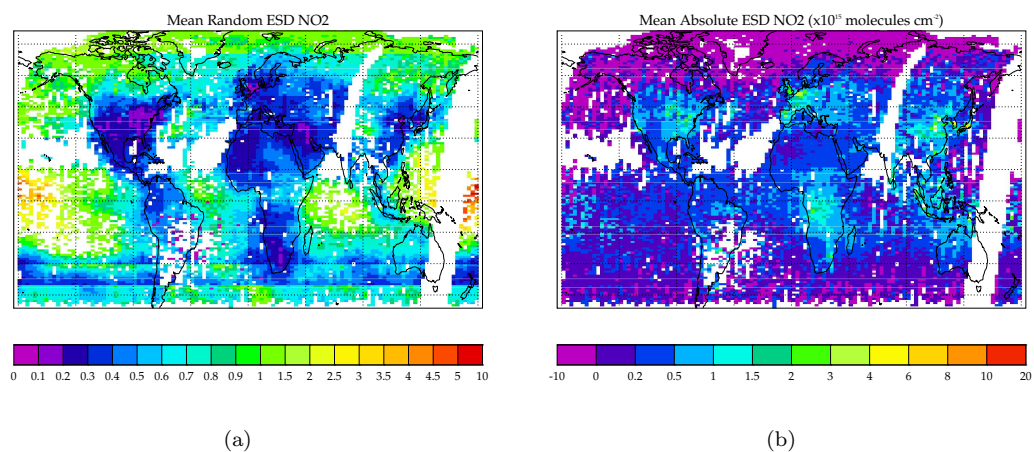
Figure B.2: *Weighted monthly mean errors in NO$_2$ tropospheric VCD for July 1997 (a) fractional measurement error, (b) absolute measurement error.*

Mean Cloud Optical Depth

(a)

Mean Cloud Top Height

(b)
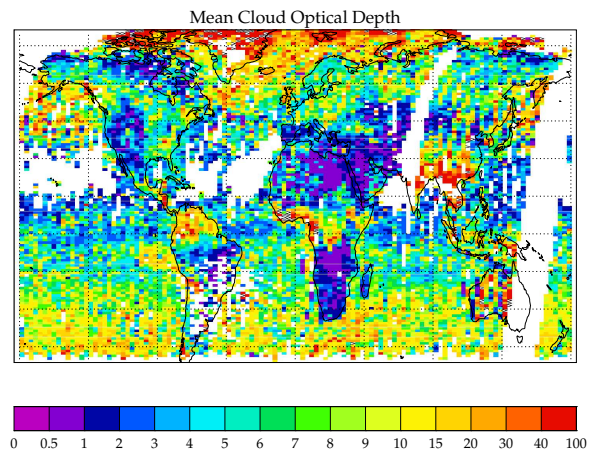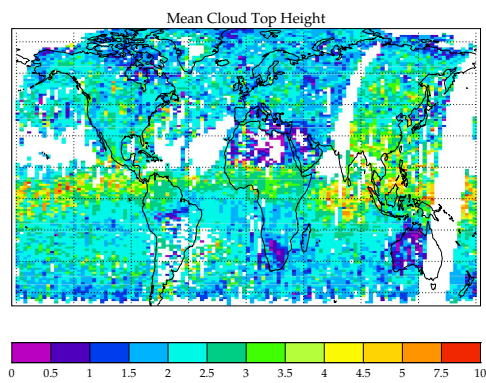
Mean Cloud Fraction
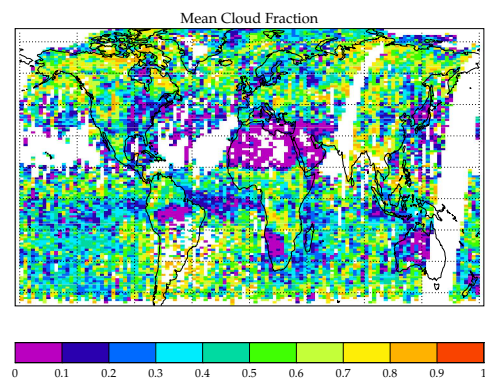
(c)

Figure B.3: *Weighted monthly mean cloud parameters for July 1997 (a) cloud optical depth, (b) cloud top height (km), (c) cloud fraction.*
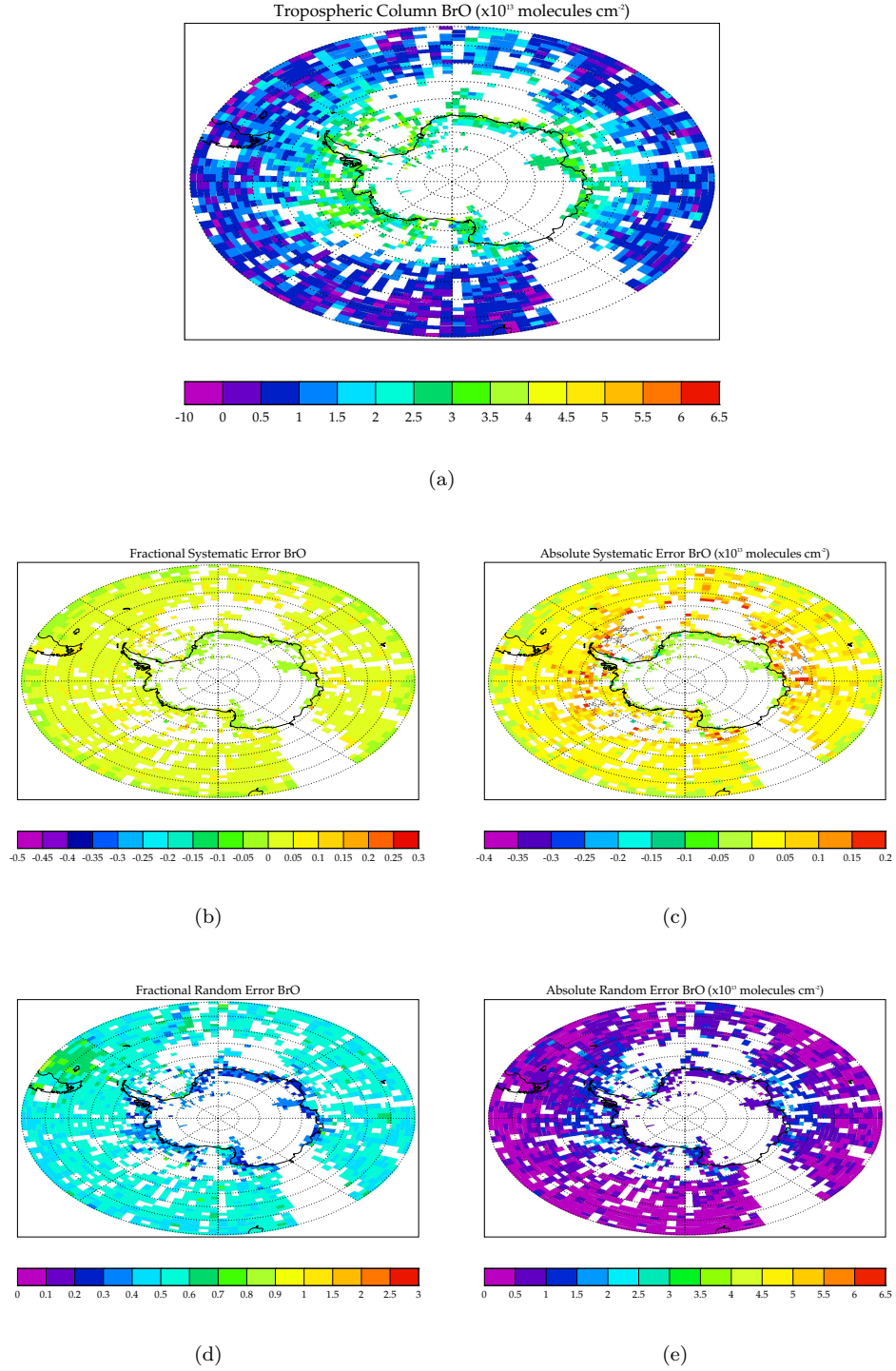
Figure B.4: *Weighted monthly mean BrO for November 1996 from SAO slant columns for the south pole for (a) tropospheric VCD ($\times 10^{13}$ molecules cm^{-2}), (b) fractional systematic error on the retrieved tropospheric VCD ($\times 10^{15}$ molecules cm^{-2}), (c) absolute systematic error in the retrieved tropospheric VCD ($\times 10^{13}$ molecules cm^{-2}), (d) fractional ESD in the retrieved tropospheric VCD, (e) absolute ESD in the retrieved tropospheric VCD ($\times 10^{13}$ molecules cm^{-2}), (e) fractional random in the retrieved tropospheric VCD, (g) absolute random error in the retrieved tropospheric VCD ($\times 10^{13}$ molecules cm^{-2}).*
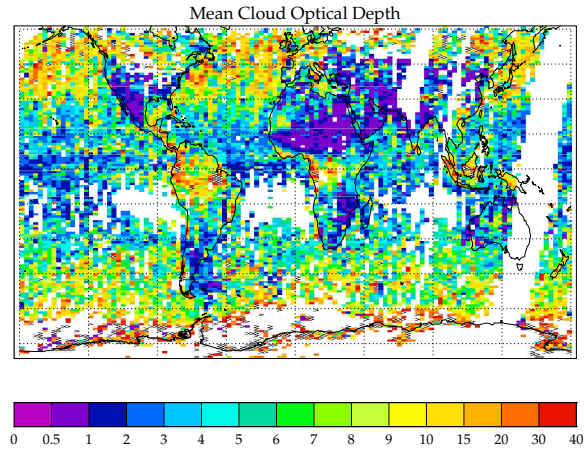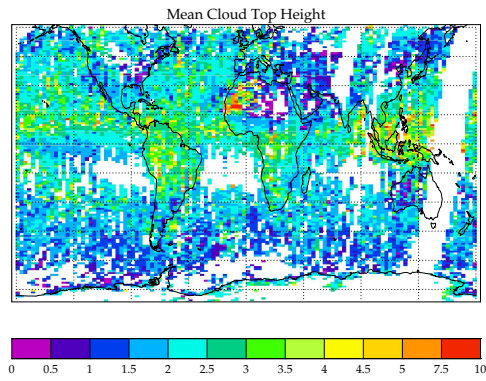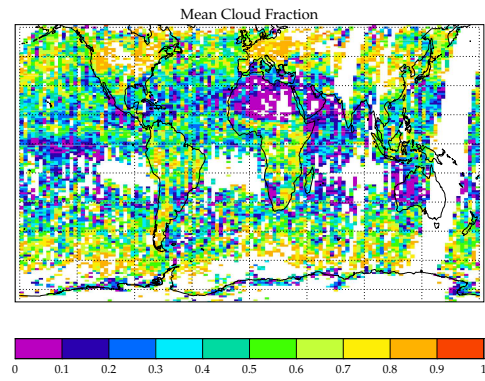
Figure B.5: *Weighted monthly mean cloud parameters for July 1997 (a) cloud optical depth, (b) cloud top height (km), (c) cloud fraction.*

References

Abbot, D. S., P. I. Palmer, R. V. Martin, K. Chance, D. J. Jacob, and A. Guenther (2003). Seasonal and interannual variability of North American isoprene emissions as determined by formaldehyde column measurements from space. *Geophys. Res. Lett. 30*(17), 1887. doi:10.1029/2003GL017336.

Agency, E. S. Image of European Research Satellite. Available at: `http://www.esa.int`.

Arnold, C. (2007). Analysis of GRAPE data. Available at: `http://www.atm.oc.ac.uk/group/eodg/mphys_ reports/2007_ Arnold.pdf`.

Arnold, S. R., M. P. Chipperfield, and M. A. Blitz (2005). A three-dimensional model study of the effect of new temperature-dependent quantum yields for acetone photolysis. *J. Geophys. Res. 110*(D10S40). doi:10.1029/2005JD005998.

Austin, J. (1992). Toward the four-dimensional assimilation of stratospheric chemical constituents. *J. Geophys. Res. 97*(D2), 2,569–2,588.

Baran, A. J., P. N. Francis, L. C. Labonnote, and M. Doutriaux-Boucher (2001). A scattering phase function for ice cloud: Tests of applicability using aircraft and satellite multi–angle multi–wavelength radiance measurements of cirrus. *Q.J.R. Meterol. Soc 127*, 2,395–2,416.

Barkstrom, B. (1976). A finite difference method of solving anisotropic scattering problems. *J. Quant. Spectrosc. Radiat. Transfer 16*, 725 – 739.

Barrie, L. and U. Platt (1997). Artic tropospheric chemistry: an overview. *TELLUS 49B*, 450–454.

Barrie, L. A., J. W. Bottenheim, R. C. Schnell, P. J. Crutzen, and R. A. Rasmussen (1988). Ozone destruction and photochemical reactions at polar sunrise in the lower Arctic atmosphere. *Nature 334*, 138–141.

Beirle, S., N. Spichtinger, A. Stohl, K. L. Cummins, T. Turner, D. Boccippio, O. R. Cooper, M. Wenig, M. Grzegorski, U. Platt, and T. Wagner (2006). Estimating the NOx produced by lightning from GOME and NLDN data: A case study in the Gulf of Mexico. *Atmos. Chem. Phys. 6*, 1,075 – 1,089.

Bey, I., D. J. Jacob, R. M. Yantosca, J. A. Logan, B. D. Field, A. M. Fiore, Q. Li, H. Liu, L. J. Mickeley, and M. G. Schultz (2001). Global modelling of tropospheric chemistry with assimilated meteorology: Model description and evaluation. *J. Geophys. Res. 106*, 23,073–23,096.

Boersma, K. F., H. J. Eskes, and E. J. Brinksma (2004). Error analysis for tropospheric NO_2 retrieval from space. *J. Geophys. Res. 109*. D04311, doi:10,1029/2003JD003962.

Boersma, K. F., H. J. Eskes, E. W. Meijer, and H. M. Kelder (2005). Estimates of lightning NOx production from GOME satellite observations. *Atmos. Chem. Phys. 5*, 2,311 – 2,331.

Borrell, P. M., J. P. Burrows, and U. Platt (2004). *Sounding the Troposphere from Space.* Verlag, Berlin, Heidelberg, New York: Springer.

Bovensmann, H., J. P. Burrows, M. Buchwitz, J. Frerick, S. Noel, V. V. Rozanov, K. V. Chance, and A. P. H. Goede (1999). SCIAMACHY: Mission objectives and measurement modes. *J. Atmos. Sci. 56*(2), 127–147.

Brewer, A. W., C. T. McElroy, and J. R. Kerr (1973). Nitrogen dioxide concentrations in the atmosphere. *Nature 246*, 129–131.

Burrows, J. P., M. Weber, M. Buchwitz, V. Rozanov, A. Ladstatter-Weiβenmayer, A. Richter, R. Debeek, R. Hoogen, K. Bramstedt, K. Eichman, M. Eisenger, and D. Perner (1999). The Global Ozone Monitoring Experiment (GOME): Mission concept and first scientific results. *J. Atmos. Sci. 56*(2), 151–175.

Cahalan, R. F., W. Ridgway, W. J. Wiscombe, T. L. Bell, and J. B. Snider (1994). The albedo of fractal stratocumulus clouds. *J. Atmos. Sci. 51*.

Callies, J., E. Corpaccioli, M. Eisinger, A. Hahne, and A. Lefebvre (2000). Gome–2 – Metop's Second-Generation Sensor for Operational Ozone Monitoring. *ESA Bulletin 102*, 28–36.

Chance, K., P. I. Palmer, R. J. D. Spurr, R. V. Martin, T. P. Kurosu, and D. J. Jacob (2000). Satellite observations of formaldehyde over North America from GOME. *Geophys. Res. Lett. 27*, 3,461–3,464.

Chance, K. V. (1998). Analysis of BrO measurements from the Global Ozone Monitoring Experiment. *Geophys. Res. Lett. 25*(2), 3,335–3,338.

Chipperfield, M. P. (1999). Multiannual simulations with a three-dimensional chemical transport model. *J. Geophys. Res. 104*(D1), 1,781–1,805.

Chipperfield, M. P. (2006). New version of the TOMCAT/SLIMCAT off-line chemical transport model: Intercomparison of stratospheric tracer experiments. *Q.J.R. Meterol. Soc 132*, 1,179–1,203.

Chipperfield, M. P., B. V. Khattatov, and D. J. Lary (2002). Sequential assimilation of stratospheric chemical observations in a three-dimensional model. *J. Geophys. Res. 107*(D21), 4585. doi:10.1029/2002JD002110.

Chipperfield, M. P., M. L. Santee, L. Froidvaux, G. L. Manney, W. G. Read, J. W. Waters, A. E. Roche, and J. M. Russell (1996). Analysis of UARS data in the southern polar vortex in September 1992 using a chemical transport model. *J. Geophys. Res. 101*(D13), 18,861–18,881.

Clerbaux, C., J. Hadji-Lazaro, D. Hauglustaine, and G. Megie (2001). Assimilation of carbon dioxide measured from satellite in a three-dimensional chemistry-transport model. *J. Geophys. Res. 106*(D14), 15,385–15,394.

Coldewey-Egbers, M., M. Weber, L. N. Lamsal, R. de Beek, M. Buchwitz, and J. P. Burrows (2004). Total ozone retrieval from GOME UV spectral data using the weighting function DOAS approach. *Atmos. Chem. Phys. 4*, 4,916–4,944.

deBeek, R., R. Hoogen, V. Rozanov, and J. Burrows (1998). Ozone profile retrieval from GOME satellite data I: Alogorithm description. Technical document.

deBeek, R., R. Hoogen, V. Rozanov, and J. P. Burrows (1999). Ozone profile retrieval from GOME satellite data I: Algorithm description and first validation. *J. Geophys. Res. 104*, 8,263–8,280.

Di Giuseppe, F. and A. M. Tompkins (2003). Effect of spatial organization on solar radiative transfer in three-dimensional idealized stratocumulus cloud fields. *J. Atmos. Sci. 60*(15), 1,774–1,794.

Dubovik, O., B. Holben, T. F. Eck, A. Smirnov, Y. J. Kaufman, M. D. King, D. Tanre, and I. Slutsker (2002). Variability of absorption and optical properties of key aerosol types observed in worldwide locations. *Journal of Atmospheric Sciences 59*, 590–608.

Eck, T. F., B. N. Holben, D. E. Ward, M. M. Mukelabai, O. Dubovik, A. Smirnov, J. S. Schafer, N. C. Hsu, S. J. Piketh, A. Queface, J. Le Roux, R. J. Swap, and I. Slutsker (2003). Variability of biomass burning aerosol optical characteristics in southern Africa during the SAFARI 2000 dry season campaign and a comparison of single scattering albedo estimates from radiometric measurements. *J. Geophys. Res. 108*(D13). doi:10.1029/2002JD002321.

Egorova, T., V. Zubov, S. Jagovkina, and E. Rozanov (1999). Lightning production of NO_x and Ozone. *Phys. Chem. Earth 24*(5), 473–479.

Eisinger, M. and J. P. Burrows (1998). Tropospheric sulfur dioxide observed by the ERS-2 GOME instrument. *Geophys. Res. Lett. 25*, 4,177–4,180.

Elbern, H. and H. S. A. Ebel (1997). Variational data assimilation for tropospheric chemistry modelling. *J. Geophys. Res. 102*(D13), 15,967–15,985.

EPA (1990–1996). National Air Pollutant Emission Trends. Available at: `http://www.epa.gov/ttn/chief/trends/`.

Errera, Q. and D. Fonteyn (2001). Four-dimensional variational chemical assimilation of CRISTA stratospheric measurements. *J. Geophys. Res. 106*(D11), 12,253–12,265.

Eskes, H. J. and K. F. Boersma (2003). Averaging kerneld for DOAS total–column satellite retrievals. *Atmos. Chem. Phys. 3*, 1285 — 1291.

Eskes, H. J., A. J. M. Piters, M. A. F. Allart, and H. Kelder (1998). Variational data assimilation: How to extract more information from GOME total ozone data. *Earth Observation Quarterly–GOME special 58*, 35–38.

Eskes, H. J., A. J. M. Piters, P. F. Levelt, M. A. Allaart, and H. M. Kelder (1999). Variational assimilation of GOME total-column ozone satellite data in a 2D latitude-longitude tracer-transport model. *J. Atmos. Sci. 56*(20), 3,560–3,572.

EUMETSAT (2007). Study to Consolidate the UVS Mission Requirements for the Oxygen A–band. Available at: `http://www.eumetsat.int`.

Evans, M. J., D. J. Jacob, E. Atlas, C. A. Cantrell, F. Eisele, F. Flocke, A. Fried, R. L. Mauldin, B. A. Ridley, B. Wert, R. Talbot, B. Heikes, J. Snow, J. Walega, A. J. Weinheimer, and J. Dibb (2003). Coupled evolution of BrOx-ClOx-HOx-NOx chemistry during bromine-catalyzed ozone depletion events in the Arctic boundary layer. *J. Geophys. Res. 108*(D4), 8368. doi:10.1029/2002JD002732.

Fahey, D. W., K. K. Kelly, D. M. Murphy, M. H. Proffitt, C. S. Eubank, M. K. W. Ko, G. V. Ferry, M. Loewenstein, and K. R. Chan (1989). Measurements of nitric oxide and total reactive nitrogen in the antarctic stratosphere: Observations and chemical implications. *J. Geophys. Res. 94*, 16,665–16,681.

Feng, W., M. P. Chipperfield, M. Dorf, K. Pfeilsticker, and P. Ricaud (2007). Mid-latitude ozone changes: studies with a 3-D CTM forced by ERA40 analyses. *Atmos. Chem. Phys. 7*, 2,357–2,369.

Fish, D. J., H. K. Roscoe, and P. V. Johnston (2000). Possible causes of stratospheric NO_2 trends observed at Lauder, New Zealand. *Geophys. Res. Lett. 27*, 3,313–3,316.

Fisher, M. and D. J. Lary (1995). Lagrangian 4-dimensional variational data ssimilation of chemical species. *Q.J.R. Meterol. Soc 121*, 1,681 – 1,704.

Fishman, J., C. E. Watson, J. C. Larsen, and J. A. Logan (1990). Distribution of tropospheric ozone determined from satellite data. *J. Geophys. Res. 95*, 3,599–3,617.

Fishman, J., A. E. Wozniak, and J. K. Creilson (2003). Global distribution of tropospheric ozone from satellite measurements using the empirically corrected tropospheric ozone residual technique: Identification of the regional aspects of air pollution. *Atmos. Chem. Phys. 3*, 893–907.

Fu, Q., M. C. Cribb, H. W. Barker, S. K. Krueger, and A. Grossman (2000). Cloud geometry effects on atmospheric solar absorption. *J. Atmos. Sci. 57*(8), 1,156–1,168.

Fu, T. M., D. J. Jacob, P. I. Palmer, K. Chance, Y. X. Wang, B. Barletta, D. R. Blake, J. C. Stanton, and M. J. Pilling (2007). Space-based formaldehyde measurements as constraints on volatile organic compound emissions in east and south Asia and implications for ozone. *J. Geophys. Res. 112*.

Gil, M., M. Yela, L. N. Gunn, A. Richter, I. Alonso, M. P. Chipperfield, E. Cuevas, J. Iglesias, M. Navarro, O. Puentedura, and S. Rodríguez (2008). NO_2 Climatology in the northern subtropical region: Diurnal, seasonal and interannual variability. *Atmos. Chem. Phys. 8*, 1,635–1,648.

Gunson, M. R. (1996). The atmospheric trace molecule spectroscopy (ATMOS) experiment: Deployment on the ATLAS space shuttle missions. *Geophys. Res. Lett. 23*, 2,333–2,336.

Hasekamp, O. P. and J. Landgraf (2001). Ozone profile retrieval from backscattered ultraviolet radiances: The inverse problem solved by regularization. *J. Geophys. Res. 106*(D8), 8,077 – 8,088.

Hasekamp, O. P. and J. Landgraf (2002). Tropospheric ozone information from satellite-based polarization measurements. *J. Geophys. Res. 107*(D17), 4326. doi:10.1029/2001JD001346.

Hausmann, M. and U. Platt (1994). Spectroscopic measurement of bromine oxide and ozone in the high Arctic during Polar Sunrise Experiment 1992. *J. Geophys. Res. 99*, 25,399–25,413.

Health Organization, W. (2003). Health aspects of air pollution with particulate matter, ozone and nitrogen dioxide.

Heath, D. F., C. L. Mateer, and A. J. Kreuger (1973). The NIMBUS–4 BUV atmospheric ozone experiement – Two year's operation. *Pure and Applied Geophysics 106-108*, 1,238–1,253.

Hecht, E. (1987). *Optics* (2nd Edition ed.). Addison–Wesley. ISBN.

Hegels, E., P. J. Crutzen, T. Klupfel, D. Perner, and J. P. Burrows (1998). Global distribution of atmospheric bromine-monoxide from GOME on earth observing satellite ERS-2. *Geophys. Res. Lett. 25*, 3,127–3,130.

Hess, M., P. Koepke, and I. Schult (1998). Optical Properties of Aerosols and Clouds: The software OPAC. *Bull. Am. Met. Soc. 79*, 831–844.

Hild, L., A. Richter, V. Rozanov, and J. P. Burrows (2002). Air mass factor calculations for GOME measurements of lightning-produced no.2. *Adv. Space Res. 29*, 1,685–1,690.

Hollwedel, J., M. Wenig, S. Beirle, S. Kraus, S. Kühl, W. Wilms-Grabe, U. Platt, and T. Wagner (2004). Year-to-year variations of spring time polar tropospheric BrO as seen by GOME. *Adv. Space Res. 34*(4), 804–808.

Holtslag, A. A. M. and B. Boville (1993). Local versus nonlocal boundary layer diffusion in a global climate model. *Climate 6*, 1,825–1,842.

Jacobson, M. Z. (1999). *Fundamentals of Atmospheric Modelling*. The Edinburgh Building, Cambridge, CB2 2RU, UK: Cambridge University Press.

Jaegle, L., L. Steinberger, R. V. Martin, and K. Chance (2005). Global partitioning of NO_x sources using satellite observations: Relative roles of fossil fuel combustion, biomass burning and soil emissions. *Faraday Discussions 130: Atmospheric Chemistry*.

Kerridge, B. J. K., R. Siddans, B. L. Latter, J. P. Burrows, M. Weber, R. De Beek, I. Aben, and W. Hartman (2004). EUMETSAT Final Report. GOME2 error assessment. Contract EUM/CO/01/901/DK.

Khattatov, B. V., J. C. Gille, L. V. Lyjak, G. P. Brasseur, V. L. Dvortsov, A. E. Roche, and J. W. Waters (1999). Assimilation of photochemically active species and a case analysis of UARS data. *J. Geophys. Res. 104*, 18,715–18,737.

Khattatov, B. V., J. F. Lamarque, L. V. Lyjak, R. Menard, P. Levelt, X. Tie, G. P. Brasseur, and J. C. Gille (2000). Assimilation of satellite observations of long-lived chemical species in global chemistry transport models. *J. Geophys. Res. 105*(D23), 29,135–29,144.

Kiehl, J. T. and K. E. Trenberth (1997). Earth's annual mean global energy budget. *Bulletin of the American Meteorological Society 78*, 197–208.

Kim, J. H., M. J. Newchurch, and K. Han (2001). Distribution of tropical tropospheric ozone determined by the scan-angle method applied to TOMS measurements. *J. Atmos. Sci. 58*, 2,699–2,708.

Koelemeijer, R. and P. Stammes (1998). Cloud fraction and cloud top pressure retrieval from GOME compared with ATSR-2 measurements. *J. Quant. Spectrosc. Radiat. Transfer 60*(6), 943–961.

Koelemeijer, R. B. A., J. F. de Haan, and P. Stammes (2003a). A database of spectral reflectivity in the range 335-772 nm derived from 5.5 years of GOME observations. *J. Geophys. Res. 108*, 4070. doi:10.1029/2002JD002429.

Koelemeijer, R. B. A., J. F. de Haan, and P. Stammes (2003b). A database of spectral surface reflectivity in the range 335-772 nm derived from 5.5 years of GOME observations. *J. Geophys. Res. 108*(D2). doi:10,1029/2002JD002429.

Koelemeijer, R. B. A., P. Stammes, J. W. Hovenier, and J. F. de Haan (2001). A fast method for retrieval of cloud parameters using oxygen A band measurements from the Global Ozone Monitoring Experiment. *Geophys. Res. Lett. 106*, 3,475–3,490.

Koop, T., A. Kapilashrami, L. Molina, and M. Molina (2000). Phase transitions of sea-salt/water mixtures at low temperatures: Implications for ozone chemistry in the polar marine boundary layer. *Geophys. Res. Lett. 105*, 26,393–26,402.

Kreher, K., P. V. Johnston, S. W. Wood, B. Nardi, and U. Platt (1997). Ground-based measurements of tropospheric and stratospheric BrO at Arrival Heights, Antarctica. *J. Geophys. Res. 24*, 3,021–3,024.

Kunhikrishnan, T., M. G. Lawrence, R. von Kuhlmann, A. Richter, A. Ladstatter-Weiβenmayer, and J. P. Burrows (2004). Analysis of tropospheric NOx over Asia using the model of atmospheric transport and chemistry (MATCH-MPIC) and GOME-satellite observations. *Atmos. Environ. 38*, 581–596.

Lary, D. J., J. A. Pyle, C. R. Webster, and R. D. May (1991). The BLISS measurements of NO_2: some new insights. *18*(12), 2,261–2,263.

Latter, B. (2004). *Investigation of UV radiative transfer and instrument errors and their impact on ozone profile retrieval for the Global Ozone Monitoring Experiment.* Ph. D. thesis, University of Reading.

Leue, C., M. Wenig, T. Wagner, O. Klimm, U. Platt, and B. Jähne (2001). Quantitative analysis of NOx emissions from Global Ozone Monitoring Experiment satellite image sequences. *J. Geophys. Res. 106*, 5,493–5,505.

Levelt, P. F., B. V. Khattatov, J. C. Gille, G. P. Brasseur, X. X. Tie, and J. W. Waters (1998). Assimilation of MLS ozone measurements in the global three-dimensional chemistry transport model ROSE. *Geophys. Res. Lett. 25*(24), 4,493–4,496.

Liley, J. B., P. V. Johnston, R. L. Mackenzie, A. J. Thomas, and I. Boyd (2000). Stratospheric NO_2 variations from a long time series at Lauder, New Zealand. *J. Geophys. Res. 105*(11), 11,633–11,640.

Liou, K. N. (2002). *An Introduction to Atmospheric Radiation.* 84 Theobalds Road, London, WC1X 8RR, UK: Academic Press.

Loyola, D., W. Thomas, and W. Balzer (2004). Product Specification Document of the GOME Data Processor. `http://earth.esa.int/ers/gome`.

Lui, X., K. Chance, and T. P. Kurosu (2007). Improved ozone profile retrievals from GOME data with degradation correction in reflectance. *Atmos. Chem. Phys. 7*, 1,575–1,583.

Marsh, S. H., S. M. Dean, R. G. Grainger, A. L. Quinjano, G. E. Thomas, and B. N. Lawerence (2004). An Optimal Estimation Aerosol Retrieval Scheme For ATSR-2. AOPP memorandum. available at :http://www.atm.ox.ac.uk/main/Science/technical/2007.1.pdf.

Marshak, A., A. Davis, W. Wiscombe, and R. Cahalan (1997). Inhomogeneity effects on cloud shortwave absorption measurements: Two-aircraft simulations. *J. Geophys. Res. 102*(D14), 16,619–16,637.

Martin, R. V., K. Chance, D. J. Jacob, T. P. Kurosu, R. J. D. Spurr, E. Buscela, J. F. Gleason, P. I. Palmer, I. Bey, A. M. Fiore, Q. Li, R. M. Yantosca, and R. B. A. Koelemeijer (2002). An improved retrieval of tropospheric nitrogen dioxide from GOME. *J. Geophys. Res. 107*, 4437. doi:10.1029/2001JD001027.

Martin, R. V., D. J. Jacob, K. Chance, T. P. Kurosu, P. I. Palmer, and M. J. Evans (2003). Global inventory of nitrogen oxide emissions constrained by space-based observations of NO$_2$ columns. *J. Geophys. Res. 108*, 4537. doi:10.1029/2003JD003453.

McKenna, D. S., P. Konopka, J. U. Grooβ, G. Gunther, R. Muller, R. Spang, D. Offermann, and Y. Orsolini (2002). A new Chemical Langrangian Model of the Stratosphere (CLaMS) 1. Formulation of advection and mixing. *J. Geophys. Res. 107*(D16). 10.1029/2000JD000114.

McLinden, C. A., S. C. Olsen, M. J. Prather, and J. B. Liley (2001). Understanding trends in NO$_y$ and NO$_2$. *J. Geophys. Res. 106*, 27,787–27,793.

McPeters, R. D., A. J. Krueger, P. K. Bhartia, and J. R. Herman (1998). TOMS data Products User's Guide. Technical report, Goddard Space Flight Centre, Greenbelt, Maryland.

Meerkoetter, R., B. Wissinger, and G. Seckmeyer (1997). Surface UV from ERS-2/GOME and NOAA/AVHRR data: A case study. *Geophys. Res. Lett. 24*(15), 1,939–1,942.

Meijer, Y. J., R. J. Van der A, R. F. Van Oss, D. P. J. Swart, H. M. Kelder, and P. V. Johnston (2003). Global Ozone Monitoring Experiment ozone profile characterisation using interpretation tools and lidar measurements for intercomparison. *J. Geophys. Res. 108*(D23), 4723. doi:10,1029/2003JD003498.

Menard, R. and L. P. Chang (2000). Assimilation of stratospheric chemical tracer observations using a kalman filter. Part 2: X2-Validated results and analysis of variance and correlation dynamics. *Monthly Weather Review 128*, 2,671–2,686.

Menard, R., S. E. Cohn, L. P. Chang, and P. M. Lyster (2000). Assimilation of stratospheric chemical tracer observations using a kalman filter. Part 1: Formulation. *Mon Wea Rev 128*, 2,654–2,671.

Meyer-Arnek, J., A. Ladstatter-Weiβenmayer, A. Richter, F. Wittrock, and J. P. Burrows (2005). A study of the trace gas columns of O$_3$, NO$_2$ and HCHO over Africa in September 1997. *Faraday Discussion 130: Atmospheric Chemistry*.

Mie, G. (1908). Contributions of optics of turbid media, particularly of colloidal metal solutions. *Ann. Phys. 25*(3), 377–455.

Monge-Sanz, B., M. P. Chipperfield, A. Simmons, and S. Uppala (2007). Mean age of air transport in a CTM: Comparison of different ECMWF analyses. *Geophys. Res. Lett. 34*. L04801,doi:10.1029/2006GL028515.

Munger, J. W., S. M. Fan, P. S. Bakwin, M. L. Goulden, A. H. Goldstein, A. S. Coleman, and S. C. Wofsy (1998). Regional budgets for nitrogen oxides from continental sources: Variations of rates for oxidation and deposition with season and distance from source regions. *J. Geophys. Res. 103*, 8,355–8,368.

Munro, R., R. Siddans, W. J. Redburn, and B. J. Kerridge (1998). Direct measurements of tropospheric ozone from space. *Nature 392*, 168–171.

Mutlow, C., J. Murray, P. Bailey, A. Birks, and D. Smith (1999). ATSR -1/2 User Guide. `http://www.atsr.rl.ac.uk/documentation/docs/userguide/index.shtml`.

Myhre, G., Y. Govaerts, J. M. Haywood, T. K. Berntsen, and A. Lattanzio (2005). Radiative effect of surface albedo change from biomass burning. *Geophys. Res. Lett. 32*. doi:10.1029/2005GL022897.

N$\ddot{u}\beta$, J. H., A. Richter, B. V. Sinuhuber, F. Wittrock, J. P. Burrows, U. Niemeier, C. Granier, J. F. Muller, N. Savage, K. S. Law, J. A. Pyle, and J. G. J. Olivier (2002). GOME NO_2 retrieval with model profiles. Available at: `www.doas-bremen.de`.

Palmer, P. I., D. J. Jacob, K. Chance, R. V. Martin, R. J. D. Spurr, T. P. Kurosu, I. Bey, R. Yantosca, A. Fiore, and Q. Li (2001). Air mass factor formulation for spectroscopic measurements from satellites: Application to formaldehyde retrievals from the Global Ozone Monitoring Experiment. *J. Geophys. Res. 106*, 14,539–14,550.

Palmer, P. I., D. J. Jacob, A. M. Fiore, R. V. Martin, K. Chance, and T. P. Kurosu (2003). Mapping isoprene emissions over North America using formaldehyde column observations from space. *J. Geophys. Res. 108*(D6), 10–1029.

Perner, D. and U. Platt (1979). Detection of nitrous acid in the atmsphere by differential optical absorption. *J. Geophys. Res. 6*, 917 – 920.

Platt, U. (1994). *Air Monitoring by Spectroscopic Techniques*. Wiley.

Platt, U. and D. Perner (1980). Direct measurements of atmospheric CH^2O, HNO^2, O^3, NO^2, and SO^2 by Differential Optical Absorption in the near UV. *J. Geophys. Res. Oceans Atmos* (85), 7,453–7,458.

Plumb, R. A. and M. K. W. Ko (1992). Interrelationships between mixing ratios of long–lived stratopsheric constituents. *J. Geophys. Res. 97*, 10,145–10,156.

Poulsen, C., C. Mutlow, and S. Tjemkes (2004). EUMETSAT Cloud resolution study. Technical document.

Poulsen, C., R. Siddans, S. Dean, D. Grainger, G. Thomas, and P. Watts (2002). *Cloud Parameter retrieval from ATSR–2*. EUMETSAT Meterological Satellite Conference.

Raes, F., R. Van Dingnen, E. Vignati, J. Wilson, J.-P. Putaud, J. H. Seinfeld, and P. Adams (2000). Formation and cycling of aerosols in the global troposphere. *Atmospheric Environment 34*, 4,215–4,240.

Ravetta, F., G. Ancellet, A. Colette, and H. Schlager (2007). Long-range transport and tropospheric ozone variability in the western Mediterranean region during the Intercontinental Transport of Ozone and Precursors (ITOP-2004) campaign.

Richter, A. and J. P. Burrows (2001). Tropospheric NO_2 from GOME measurements. *Adv Space Res. 29*, 1,673–1,683.

Richter, A., J. P. Burrows, H. NuB, C. Granier, and U. Niemeier (2005). Increase in tropospheric nitrogen dioxide over China observed from space. *Nature 437*. doi:10.1038/nature04092.

Richter, A., M. Eisenger, A. Ladstatter-Weiβenmayer, and J. P. Burrows (1999). DOAS zenith sky observations: 2 seasonal variation of BrO over Bremen (53°N 1994-1995. *J. Atmos. Chem. 32*, 83–99.

Richter, A., M. Eisenger, F. Wittrock, and J. P. Burrows (1998). Measurements of halogen oxides by GOME. *Earth Observation Quarterly 58*, 19–20.

Richter, A. and T. Wagner (1998). GOME diffuser plate structures. Available at: `http://www.iup.bremen.de.gome`.

Richter, A., A. F. Wittrock, A. Ladstβtter-Weissenmayer, and J. P. Burrows (2002). GOME measurements of stratospheric and tropospheric BrO. *Adv. Space Res. 29*, 1,673–1,683.

Richter, A., F. Wittrock, M. Eisinger, and J. P. Burrows (1998). GOME observations of tropospheric BrO in northern hemisphere spring and summer 1997. *Geophys. Res. Lett. 25*, 2,683–2,686.

Rihan, F. A. and C. G. Collier (2003). Four-dimensional data assimilation and numerical weather prediction. Available at: `http://www.ties.salforf.ac.uk/frihan/`.

Riishojgaard, L. P. (1996). On four-dimensional variational assimilation of ozone data in weather-prediction models. *Q.J.R. Meterol. Soc 122*, 1,545–1,571.

Rinsland, C. P., E. Mahieu, R. Zander, P. Demoulin, J. Forrer, and B. Buchmann. Free tropospheric CO, C2H6, and HCN above central Europe: Recent measurements from the Jungfraujoch including the detection of elevated columns during 1998. *J. Geophys. Res. 105*(24).

Rodgers, C. (2000). *Inverse Methods for Atmospheric Sounding*. The Edinburgh Building, Cambridge, CB2 2RU, UK: Cambridge University Press.

Rozanov, V. V., D. Diebel, R. J. D. Spurr, and J. P. Burrows (1997). GOMETRAN: A radiative transfer model for the satellite project GOME, the plane-parallel version. *J. Geophys. Res. 102*, 16,683 – 16,695.

Schaub, D., K. F. Boersma, J. W. Kaiser, A. K. Weiss, D. Folini, H. J. Eskes, and B. Buchmann (2006). Comparison of GOME tropospheric NO_2 columns with NO_2 profiles deduced from ground based in situ measuremnets. *Atmos. Chem. Phys. 6*, 3,211 – 3,229.

Schoeberl, M. R., A. R. Douglass, Z. Zhu, and S. Pawson (2003). A comparison of the lower stratospheric age spectra derived from a general circulation model and two data assimilation systems. *J. Geophys. Res. 108*, 4113. doi:10.1029/2002JD002652.

Schofield, R., K. Kreher, B. J. Connor, P. V. Johnston, A. Thomas, D. Shooter, M. P. Chipperfield, C. D. Rodgers, and G. H. Mount (2004). Retrieved tropospheric and stratospheric BrO columns over Lauder, New Zealand. *J. Geophys. Res. 109*, D14304. doi:10.1029/2003JD004463.

Siddans, R. (2003). *Height Resolved Ozone Retrievals from the Global Ozone Monitoring Experiment*. Ph. D. thesis, University of Reading.

Siddans, R., B. Latter, and B. J. K. Kerridge (2007). Study to Consolidate the UVS Mission Requirements for the Oxygen A–band. Contract EUM/CO/05/1411/SAT.

Singh, H. B., W. H. Brune, J. H. Crawford, D. J. Jacob, and P. B. Russell (2006). Overview of the Summer 2004 Intercontinental Chemical Transport Experiment-North America (INTEX-A). *J. Geophys. Res. 111*(D24). doi:10.1029/2006JD007905.

Singh, H. B. and D. J. Jacob (2000). Future Directions: Satellite observations of tropospheric chemistry. *Atmos. Env. 34*, 4,399–4,401.

Sinnhuber, B. M., D. W. Arlander, H. Bovensmann, J. P. Burrows, M. P. Chipperfield, C. F. Enell, U. Frieβ, F. Hendrick, P. V. Johnston, R. L. Jones, K. Kreher, N. Mohamed-Tarin, R. Muller, K. Pfeilsticker, U. Platt, J. P. Pommereau, I. Pundt, A. Richter, A. M. South, K. K. Tornkvist, M. Van Roozendael, T. Wagner, and F. Wittrock (2002). Comparison of measurements and model calculations of stratospheric bromine monoxide. *J. Geophys. Res. 107*. doi:10.1029/2001JD000940.

Sinreich, R., U. Freiβ, T. Wagner, and U. Platt (2005). Multi axis differential optical absorption spectroscopy (MAX-DOAS) of gas and aersol distributions. *Faraday Discussions 130: Atmospheric Chemistry*.

Solomon, A., A. Schmeltekopf, and R. Sanders (1987). On the interpretation of zenith sky absorption measurements. *J. Geophys. Res. 92*(D7), 8,311–8,319.

Solomon, S. (1999). Stratospheric ozone depletion: A review of concepts and history. *Rev. Geophys 37*(3), 275–316.

Spracklen, D. V., K. J. Pringle, K. S. Carslaw, M. P. Chipperfield, and G. W. Mann (2005). A global off-line model of size-resolved aerosol microphysics: Model development and prediction of aerosol properties. *Atmos. Chem. Phys. 5*, 2,227–2,252.

Spurr, R. J. D., T. P. Kurosu, and K. V. Chance (2001). A linearized discrete ordinate radiative transfer model for atmospheric remote–sensing retrieval. *Journal of Quantiative Spectroscopy and Radiative Transfer 68*, 689 – 735.

Stockwell, D. Z. and M. P. Chipperfield (1999). A tropospheric chemical transport model: Development and validation of the model transport schemes. *Q.J.R. Meterol. Soc 125*, 1,747–1,783.

Streets, D. G. (2003). An inventory of gaseous and primary aerosol emissions in Asia in the year 2000. *J. Geophys. Res. 108*(D21). doi:10.1029/2002JD003093.

Sturges, W. T., G. F. Cota, and P. T. Buckley (1992). Bromoform emission from Arctic ice algae. *Nature 358*, 660–662.

Stutz, J. (2001). Influence of halogen oxides on tropospheric ozone. Available at: `http://www.atmos.ucla.edu/ jochen/research/hox/hox.html`.

Swinbank, R. and A. O'Neil (1994). A stratosphere-troposphere data assimilation system. *Mon. Wea Rev 122*, 686–702.

TEMIS (2009). Tropospheric Emission Monitoring Internet Service. `http://www.temis.nl/`. viewed 9th February 2009.

Tiedtke, M. (1989). A comprehensive mass flux scheme for cumulus parameterization in large-scale models. *Mon. Wea Rev 117*, 1,779–1,800.

Tuinder, O. N. E., R. de Winter-Sorkina, and P. J. H. Builtjes (2004). Retrieval methods of effective cloud cover from the GOME instrument: an intercomparison. *Atmos. Chem. Phys. 4*, 255–273.

Uno, I., Y. He, T. Ohara, R. Yamaji, J. I. Kurokawa, M. Katayama, Z. Wang, K. Noguchi, S. Hayashida, A. Richter, and J. P. Burrows (2006). Systematic analysis of interannual and seasonal variations of model – simulated tropospheric NO_2 in Asia and comparison with GOME – satellite data. *Atmos. Chem. Phys. 6*, 11,181 – 11,207.

Uppala, S. M., P. W. KÅllberg, A. J. Simmons, U. Andrae, V. Da Costa Bechtold, M. Fiorino, J. K. Gibson, J. Haseler, A. Hernandez, G. A. Kelly, X. Li, K. Onogi, S. Saarinen, N. Sokka, R. P. Allan, E. Andersson, K. Arpe, M. A. Balmaseda, A. C. M. Beljaars, L. Van De Berg, J. Bidlot, N. Bormann, S. Caires, F. Chevallier, A. Dethof, M. Dragosavac, M. Fisher, M. Fuente, S. Hagemann, E. Holm, B. J. Hoskins, L. Isaksen, P. A. E. M. Janssen, R. Jenne, A. P. Mcnally, J. F. Mahfouf, J. J. Morcrette, N. A. Rayner, R. W. Saunders, P. Simon, A. Sterl, K. E. Trenberth, A. Untch, D. Vasiljevic, P. Viterbo, and J. Woollen (2005). The ERA–40 Re–analysis. *Q.J.R. Meteorol. Soc 131*, 2961–3012. doi:10.1256/qj.04.176.

Van de Hulst, H. C. (1981). *Light scattering by small particles*. Courier Dover Publication.

Van der A, R. J. (1997). *SCIAMACHY: Observational noise and errors on atmospheric profiles*. Ph. D. thesis, Space Reseach Organization of the Netherlands.

Van der A, R. J., D. H. Peters, H. Eskes, K. F. Boersma, M. Van Roozendael, I. De Smedt, and H. M. Kelder (2006). Detection of the trend and seasonal variation in tropopsheric NO_2 over China. *J. Geophys. Res. 111*.

Van Der A, R. J., R. F. Van Oss, A. J. M. Piters, and J. P. F. Fortuin (2002). Ozone profile retrieval from recalibrated Global Ozone Monitoring Experiment data. *J. Geophys. Res. 107*(D15). doi:10,1029/2001JD000696.

Van der A, R. J. and P. Wang (2007, June). Production Specification Document, FRESCO. Technical Document TEM/PSD2/002, TEMIS.

Van Roozendael, M., T. Wagner, A. Richter, I. Pundt, D. W. Arlander, J. P. Burrows, M. Chipperfield, C. Fayt, P. V. Johnston, J. C. Lambert, K. Kreher, K. Pfeilsticker,

U. Platt, J. P. Pommereau, S. B. M, K. K. Tornkvist, and F. Wittrock (2002). Intercomparison of BrO measurements from ERS-2 GOME, ground-based and balloon platforms. *Adv Space Res. 29*, 1,661–1,666.

Van Roozendael, M., T. Wagner, A. Richter, I. Pundt, D. W. Arlander, J. P. Burrows, M. Chipperfield, C. Fayt, P. V. Johnston, J. C. Lambert, K. Kreherg, K. Pfeilsticker, U. Platt, J. P. Pommereau, B. M. Sinnhuber, K. K. Trønkviste, and F. Wittrock (1999). Intercomparison of BrO measurements from ERS-2 GOME, ground-based and balloon platforms. *Adv Space Res. 29*(11), 1,661–1,666. doi:10.1016/S0273-1177(02)00098-4.

Velders, G. J. M., C. Granier, R. W. Portmann, K. Pfeilsticker, M. Wenig, T. Wagner, U. Platt, A. Richter, and J. P. Burrows (2001). Global tropospheric NO_2 column distributions: Comparing three – dimensional model calculations with GOME measurements. *J. Geophys. Res. 106*, 12,643 – 12,660.

Von Bargen, A., T. P. Kurosu, K. Chance, D. Loyola, B. Aberle, and R. J. Spurr (1998). CRAG- Cloud Retrieval Algorithm for GOME. Technical document.

Von Glasow, R. and P. J. Crutzen (2004). Tropospheric Halogen Chemistry. *Treatise on Geochemistry 4*, 21–64.

Von Glasow, R., R. Von Kuhlmann, M. G. Lawrence, U. Platt, and P. J. Crutzen (2004). Impact of reactive bromine chemistry in the troposphere. *Atmos. Chem. Phys. 4*, 2,481–2,497.

Vountas, M., V. Rozanov, and J. Burrows (1998). Ring Effect: Impact of rotational raman scattering on radiative transfer in Earth's atmosphere. *J.Quart.Spec.Rad.Tran 60*(6), 943–961.

Wagner, T. and U. Platt (1998). Satellite mapping of enhanced BrO concentrations in the troposphere. *Nature 395*, 486–490.

Wang, K. Y., A. Pyle, M. G. Sanderson, and C. Bridgeman (1999). Implementation of a convective atmospheric boundary atmospheric boundary layer scheme in a tropospheric chemistry transport model. *J. Geophys. Res. 104*, 23,729–23,745.

Wang, P., P. Stammes, R. Van der A, G. Pinardi, and M. Van Roozendael (2008). FRESCO+: An improved O2 A–band cloud retrieval algorithm for tropospheric trace gas retrievals. *Atmos. Chem. Phys. 8*.

Watts, P. D., C. T. Mutlow, A. J. Baran, and A. M. Zavody (1998). Study on cloud properties derived from METEOSAT Second Generation observations. *Technical Report ITT 97/181.*

Wayne, R. P. (2002). *Chemistry of Atmospheres, The third edition.* Great Clarendon Street, Oxford, OX2 6DP: Oxford University Press.

Weber, M., J. P. Burrows, and R. P. Cebula (1998). GOME Solar UV/VIS Irradiance Measurements between 1995 and 1997 - First Results on Proxy Solar Activity Studies. *Solar Physics 177*, 63–77.

Wiscombe, W. J. (1977). The Delta–M:Rapid yet accurate radiative flux calculations for strongly asymmetric phase functions. *J. Atmos. Sci. 34*, 1,408–1,422.

WMO (2002). Scientic assessment of ozone depletion.

World Health Organisation Working Group (2003). Health Aspects of Air Pollution with Particular Matter, Ozone and Nitrogen Dioxide. `http://www.euro.who.int/document/e79097.pdf`. viewed 2nd February 2009.

Yang, X., R. A. Cox, J. Warwick, J. A. Pyle, G. D. Carver, F. M. O'Conner, and N. H. Savage (2005). Tropospheric bromine chemistry and its impact on ozone: A model study. *J. Geophys. Res. 110*, D23311.