Comparing remote sounding measurements of a variable stratosphere

by

Matthew Toohey

A thesis submitted in conformity with the requirements
for the degree of Doctor of Philosophy
Graduate Department of Physics
University of Toronto

Abstract

Comparing remote sounding measurements
of a variable stratosphere

Matthew Toohey
Doctor of Philosophy
Graduate Department of Physics
University of Toronto
2009

The measurement of trace gases through remote sounding techniques has led to a better understanding of the processes controlling the structure and variability of the stratosphere. Differences between measurements over space and time are due to atmospheric variability and instrument errors: thus, comparison of measurements can be used to test our knowledge of both.

Comparisons of measurements over long time periods are used to identify trends. Balloon-borne infrared emission radiometer instruments have been used to make measurements of midlatitude stratospheric HNO_3 spanning a period of twelve years. The timing of the measurements is notable, since they occur before and well after the eruption of Mt. Pinatubo, which significantly perturbed HNO_3 levels, complicating prior trend analyses. No significant differences are found between the HNO_3 retrievals, although large measurement uncertainties preclude any conclusion concerning trends.

Comparisons of measurements that are closely spaced in space and time are useful for satellite validation, where one aims to reduce the effect of atmospheric variability on the estimation of systematic and random errors. A novel technique for the estimation of systematic error, which differentiates between additive and multiplicative bias, is introduced. In a comparison of measurements by the ACE-FTS and Aura MLS instruments, significant multiplicative biases are identified and described.

In order to validate the reported random errors (RREs) of measurements, satellite validation studies often focus on measurements in the tropical stratosphere, where variability is weak. The scatter in tropical measurements can then be used as an upper limit on instrument precision. In an analysis of tropical measurements by the ACE-FTS, scatter is found to be roughly consistent with the RREs for H_2O and CO. The scatter in measurements of O_3, HNO_3, and N_2O, while larger than the reported random errors, is roughly consistent with the variability simulated in the Canadian Middle Atmosphere Model. This work implies that the random error of the ACE-FTS measurements is smaller than the weak natural variability of the tropical stratosphere.

Acknowledgements

My supervisor, Kim Strong, originally hired me as a Research Assistant to aid in the MANTRA 2002 field campaign. This turned out to be a turning point in my academic career, and life. For her support and guidance over the years since I am deeply grateful.

Ted Shepherd has had a large influence on my work, and I'm grateful for his guidance, and his seemingly never-ending stream of ideas. His enthusiasm for bridging model studies and observations has really helped define my perspective on atmospheric science.

This thesis has benefitted greatly from the insightful questions and guidance provided by my Ph.D. supervisory committee, which has included Jon Abbatt, Jim Drummond and Dylan Jones. I am also greatly appreciative for the guidance I received from my final exam committee, which included many of those mentioned above, as well as Kaley Walker, and my external examiner Michelle Santee.

For their particular help in the discussions, arguments, problem set sessions, etc. that make up ones life as a grad student, I thank James Anstey, Annemarie Fraser, Andreas Jonsson, Tobias Kerzenmacher, Jeff Taylor, Debra Wunch and Aldona Wiacek.

The MANTRA balloon campaigns are certainly a team effort, and so for the availability of MANTRA data, and for the education I received in the field, I must thank the complete MANTRA science team. For the use of data from balloon flights in the early 1990's I thank the Environment Canada researchers involved, including (but certainly not limited to) Wayne Evans, Clive Midwinter and Tom McElroy. I owe much to Ben Quine, whose work with the emission radiometer instruments defined the starting point of my Ph.D. work.

In this thesis, I have made substantial use of data from the ACE and MLS satellite missions. For making the data available, and for being approachable and supportive of my research, I would like to thank the ACE and MLS science teams.

All four MANTRA campaigns were supported by the Canadian Space Agency (CSA) and the Meteorological Service of Canada. MANTRA 1998 also received support from the Centre for Research in Earth and Space Technology, while MANTRA

2002 and 2004 were also supported by Natural Sciences and Engineering Research Council (NSERC) of Canada. The Atmospheric Chemistry Experiment (ACE) is a Canadian-led mission mainly supported by the CSA and NSERC. MLS is funded by the National Aeronautics and Space Administration.

I gratefully acknowledge the scholarship support of NSERC and the CSA, as well as that provided through a Walter C. Sumner Memorial Fellowship.

Finally, I would like to thank Susann Tegtmeier, who took on the difficult job of providing a balance of moral support, inspirational prodding and unflinching scientific criticism during the writing and editing of this thesis.

Dedication

For my girls.

Table of Contents

List of Figures

Glossary of acronyms

ACE	Atmospheric Chemistry Experiment
ACE-FTS	Atmospheric Chemistry Experiment - Fourier transform spectrometer
AES	Atmospheric Environment Service
CCM	Coupled chemistry-climate model
CMAM	Canadian Middle Atmosphere Model
CVF	Circular variable filter
FOV	Field-of-view
GHG	Greenhouse gas
IR	Infrared
MAD	Median absolute deviation
MAD*	Composite median absolute deviation
MANTRA	Middle Atmosphere Nitrogen Trend Assessment
MIPAS	Michelson Interferometer for Passive Atmospheric Sounding
MLS	Microwave Limb Sounder
NAT	Nitric acid trihydrate
OLS	Ordinary least-squares
PSC	Polar stratospheric cloud
PV	Potential vorticity
QBO	Quasi-Biennial Oscillation
RMS	Root-mean-square
RRE	Reported random error
SD	Standard deviation
SD*	Composite standard deviation
SST	Sea surface temperature
UV	Ultraviolet
VMR	Volume mixing ratio

Chapter 1

Introduction

1.1 Ozone and stratospheric trace gases

The atmosphere is composed primarily of molecular nitrogen (N_2), molecular oxygen (O_2) and argon (Ar), which together account for about 99.9% of the total atmosphere. The remaining 0.1% of the atmosphere consists of a large number of chemical species, known as trace gases. Due in large part to their radiative properties, i.e., their ability to absorb and emit radiation of different wavelengths, the impact of trace gases on the climate of the Earth can be substantial, even while their relative abundances are not.

One of the most important trace gases is ozone (O_3), which exists in its largest relative abundance in the stratosphere, between $\sim$15 and 50 km altitude. O_3 absorbs solar ultraviolet (UV) radiation with wavelengths between 200 and 300 nm, shielding the surface of the Earth from this radiation. The absorption of UV radiation by O_3 is critical because such radiation has enough energy to break chemical bonds in DNA or interfere in other ways with biological processes (Kondratyev and Varotsos, 2000, and references therein).

The study of O_3 has become a subject of public interest since it was noted that anthropogenic activities might lead to a gradual lowering of the amount of O_3 in the atmosphere (e.g., Broderick et al., 1982, and references therein). This reduction, in turn, would lead to enhanced UV radiation at the surface of the Earth, which would

be detrimental to various life forms at the surface. Scientific research into O_3 has been motivated in large part by the goal of better understanding the processes by which human activity affects the abundance of O_3.

O_3 is also an important component of the atmosphere for other reasons. Absorption of UV radiation by O_3 leads to heating of the stratosphere, which represents a significant source of energy to the atmosphere. The thermal structure of the stratosphere is determined primarily by this O_3 heating. Furthermore, O_3 is a highly reactive species, and is a principal actor in the chemistry of the stratosphere.

O_3 abundances are controlled by the combination of (1) radiative processes, since the production of O_3 is tied directly to the photolysis of O_2; (2) chemical processes, since O_3 is destroyed through a number of chemical reactions with other stratospheric trace gases; and (3) dynamical processes, since the movement of air can act as a local source or sink of O_3 in a particular location of the atmosphere. Thus, in order to understand and predict the impact of human activity on stratospheric O_3, it is important to study each of the processes that control O_3 abundances, and their complex interactions and feedbacks.

1.2 Observations of stratospheric trace gases

> Measurements form the base on which our understanding of the atmosphere is built. New or improved measurement capability or application of measurement techniques in new ways or new situations has often resulted in a significant expansion or modification in our understanding. Measurements provide information for testing models and validating theoretical concepts. In atmospheric chemistry, no measurements are more fundamental than measurements of the composition of the atmosphere. (Mankin et al., 1999)

Systematic measurement of stratospheric O_3 commenced with the organization and operation of a network of ground-based instruments by G. M. B. Dobson (Dobson and Harrison, 1926), and has continued to the present day through a number of

measurement techniques and platforms. Scientific understanding of the processes that control O_3 has progressed based on the measurement of O_3, as well as of the myriad other trace gas species in the stratosphere. Understanding the chemical processes important in controlling O_3 has relied upon the measurement of the various chemical species involved in O_3 chemistry. Measurements of stratospheric trace gases, and the relationships between trace gases, has driven the development of our understanding of the dynamical nature of the stratosphere, since the spatial distributions of trace gases are often affected by dynamical transport.

The techniques used to measure trace gas abundance are numerous (see e.g., Heard, 2006). This thesis focusses on measurements made through spectroscopic remote sounding techniques. Remote sounding measurements are made of the atmosphere located at a distance from the instrument. Spectroscopic methods are based on the fact that different chemical compounds absorb and emit electromagnetic radiation differently at specific wavelengths. The pattern of absorption with wavelength provides a "fingerprint" for the chemical compound, and the intensity of absorption depends on the amount of absorber. Spectroscopic remote sounding measurements are therefore produced through a two-step process, involving first the measurement of radiation, and secondly the "retrieval" of trace gas abundances. Remote sounding measurements can be made from the ground, from sub-orbital platforms such as balloons and aircraft, and from space-based platforms. Space-based measurements have the advantage of global (or near-global) coverage, from which the full three-dimensional distribution of trace gases can be constructed.

Two observational missions are of primary importance in this thesis, and are introduced below.

1.2.1 MANTRA

The Middle Atmosphere Nitrogen Trend Assessment (MANTRA) series of balloon flights was undertaken to investigate changes in the abundances of northern hemisphere mid-latitude stratospheric ozone, and of nitrogen and chlorine compounds that play a role in ozone chemistry (Strong et al., 2005). In each flight, a large helium

balloon carried a payload of atmospheric remote sounding instruments to a float altitude of approximately 35 km for approximately one day. During the flight, various instruments made measurements of the chemical composition and temperature structure of the atmosphere. The main balloon flights were the principal component of a longer campaign: other measurements were made by ground-based instruments, sonde flights, and occasional solo instrument balloon flights before and after the main flight. Four campaigns were carried out biennially between 1998 and 2004, all from Vanscoy, Saskatchewan, Canada (52°N, 107°W).

The scientific objectives of the MANTRA mission were:

1. to measure profiles of the relevant chemical species in the ozone budget,

2. to determine historical trends of these quantities from balloon campaigns held at mid-latitudes over the past 20 years,

3. to perform intercomparisons between instruments using different measurement techniques, and

4. to participate in the validation of satellite data.

This thesis concentrates on measurements collected by two emission radiometers included as part of the balloon payload on each of the MANTRA flights. The instruments measure thermal infrared emission by the atmosphere, from which it is possible to retrieve vertical profiles of nitric acid (HNO_3) abundance. The emission radiometers flown on the MANTRA flights are part of a large suite of identical instruments built in the 1980's, some of which were flown on balloon flights dating back to 1990. As a result, the emission radiometer raw data set spans a period of over a decade. In reference to the second stated objective of the MANTRA mission, comparison of retrieved HNO_3 profiles from the MANTRA-era measurements with the early 1990's measurements is used to assess possible long-term changes in atmospheric composition.

1.2.2 ACE

The Atmospheric Chemistry Experiment (ACE) is a Canadian-led satellite mission that is measuring the concentrations of atmospheric trace gases by absorption spectroscopy using the Sun as a light source. The principal goals of the ACE mission are to make measurements that will improve understanding of the chemical and dynamical processes that control the distribution of ozone in the upper troposphere and stratosphere, especially in the Arctic, and to explore the relationship between atmospheric chemistry and climate change (Bernath et al., 2005).

The primary ACE instrument is a high-resolution (0.02 cm^{-1}) infrared Fourier transform spectrometer (ACE-FTS), covering the mid-infrared spectral region (750 – 4400 cm^{-1}). During sunrise and sunset (from the perspective of the orbiting satellite), ACE-FTS measures sequences of atmospheric absorption spectra in the limb-viewing geometry with different slant paths and tangent heights; when these spectra are analyzed, the results are inverted into vertical profiles of atmospheric constituents. A high inclination (74°), circular low Earth orbit (650 km) gives ACE-FTS coverage of tropical, mid-latitude and polar regions. The orbit is specifically tailored to result in maximum measurement sampling density over polar regions during winter and spring.

1.3 Motivation and themes

As stated above, advancing understanding of the chemical and dynamical processes that control O_3 abundances, and of the changes in those processes over time, is the motivation behind the MANTRA and ACE missions, and many other observational missions. The primary "themes" explored in this thesis, within the greater field of O_3 science, are discussed below.

1.3.1 Comparison and validation of measurements

Three of the four stated aims of MANTRA directly involve comparison of measurements: comparison of measurements made over long time periods, comparison of

measurements made by different instruments on the MANTRA balloon payload, and comparison of measurements made by MANTRA instruments with those made by satellite instruments.

Comparison of measurements is performed for two primary reasons. On the one hand, comparison of measurements over time, or space, can be used to extract information about the atmosphere. For example, the study of temporal trends, or spatial distributions, involves the comparison of measurements. On the other hand, comparison of measurements may serve the purpose of learning more about the instruments themselves, e.g., potential problems, biases and errors present in the measurements made by any particular instrument.

These two motivations for comparing measurements are intimately related. The scientific interpretation of measurements depends heavily upon the quality of those measurements. For example, the ability to detect long-term trends is dependent on the uncertainties in the data. So too is the ability to forecast future trace gas abundances and distributions through the assimilation of data into models. At the same time, assessing the quality of measurements is (as will be explored in this thesis) in some sense limited by our knowledge of the natural processes of the atmosphere.

Another type of comparison which is of considerable importance, and is touched on briefly in this thesis, is that of the comparison of measurements with model results. Atmospheric models (such as coupled chemistry-climate models, or CCMs) represent multifaceted hypotheses of the workings of the atmosphere; therefore comparing model results with measurements amounts to hypothesis testing. While measurements can be used to validate models, model results can also be used to better understand measurements. Model fields do not suffer from sparse data density or from random measurement errors that can plague observational data, and can therefore be helpful for assessing the impact of these factors.

Model-measurement comparison is an integral component of data assimilation, the process through which measurements and models are used in combination to produce a best estimate of the state of the atmosphere. The success of a data assimilation system depends on both the quality of the measurements assimilated, and on the

correct knowledge of the quality of the data, which is used to properly weight the data. This point underscores the importance of the validation of remote sounding measurements, in terms of both the retrieved values and their errors.

1.3.2 Atmospheric variability

A common thread joining (and complicating) the MANTRA objectives of trend detection and data validation is the issue of atmospheric variability.

The dynamics of the atmosphere are inherently chaotic, and lead to variability on a wide range of time and length scales. Since it is often long-term changes, on the time scale of decades, that are of particular concern in atmospheric studies, variability on shorter time scales acts as "noise" in the detection of trend signals in the long-term record.

Short-term variability, on the time scale of days, is often filtered out in the analysis of atmospheric measurement and model data, through the calculation of, e.g., monthly means. In some situations, it may however be necessary to consider the effects of short-term variability. The emission radiometer instrument measurements, for instance, have produced a handful of measurements at a single location over a period of many years. In order to understand to what degree any differences between these measurements might represent long-term variations, it is important to consider the effects of short-term variability. As another example, the validation of satellite instrument data often uses "coincident measurements" closely spaced in location and time. Differences between coincident measurements due to short-term variability are a significant source of uncertainty in satellite validation studies. Finally, since the short-term variability of certain trace gas species is produced by dynamical motions, trace gas variability is an indirect measure of dynamics. Therefore, comparison of the short-term variability of measurements with that of model results may be useful in assessing model performance (Erbertseder et al., 2005; Erbertseder et al., 2006).

1.4 Thesis objectives and organization

The objectives of this thesis are:

1. to produce retrievals of HNO_3 from spectra recorded by the emission radiometer instruments during the MANTRA-era and 1990-era balloon flights, to assess any change in HNO_3 over the time span of the measurements, and to place the HNO_3 measurements within the context of other observational data sets; and

2. to explore ways of improving upon the standard methods of satellite validation, especially in terms of exploiting the information contained in the variability of measurements.

The material of the thesis begins in Chapter 2 with a brief overview of relevant concepts in stratospheric chemistry and dynamics. The climatological spatial distributions of trace gas species in the stratosphere, and the variability they show about these distributions, are described in Chapter 3. Chapter 4 focusses on the emission radiometer measurements: the retrieval of HNO_3 profiles, the investigation of any possible long-term trends, and the comparison of the retrieved HNO_3 profiles with model results and observations from other instruments. Chapters 5 and 6 concern the validation of data from space-based instruments. Chapter 5 examines the issue of comparing the variability of coincident measurements by two instruments, in an effort to improve upon the typical analysis of coincident comparisons. Chapter 6 examines the issue of validating the reported random errors of space-based instruments through quantification of measurement scatter. A summary and suggestions for future work are provided in Chapter 7.

1.5 Personal contribution

Work on the material of this thesis began with my involvement with the MANTRA balloon campaign of 2002. I inherited the emission radiometer instruments from Dr. Brendan Quine, and operated them on both the 2002 and 2004 campaigns. I also inherited from Dr. Quine a functioning suite of programs comprising a retrieval code for

the emission radiometer observations, which is described by Quine et al. (2005). Over the next few years I made substantial modifications to the retrieval code, including making improvements to the forward model based on laboratory tests, improving the optimization subroutines, and implementing a staged retrieval algorithm incorporating an onion-peeling retrieval which was built from scratch. Data from the emission radiometer measurements of 1990 was obtained from the Meteorological Service of Canada (MSC), and the codes to read in this raw data were produced by Dr. Quine based on Fortran codes from MSC. I have written all of the analysis algorithms used in the satellite validation work of Chapters 5 and 6.

Much of the material of Chapters 4 and 5 has been published previously by Toohey et al. (2007) and Toohey and Strong (2007), respectively. A manuscript for publication based on the results shown in Chapter 6 is currently in preparation.

Chapter 2

Concepts in stratospheric chemistry and dynamics

This chapter presents an overview of some important concepts used to describe and understand the abundances and distributions of stratospheric trace gases. O_3 plays an especially central role in the chemical and physical processes of the stratosphere. A basic description of O_3 chemistry is included within §2.1, in order to exemplify some of the important chemical concepts introduced there. The heating of the stratosphere produced by the absorption of UV radiation by O_3 shapes the vertical and horizontal thermal structure of the stratosphere, which sets the scene in §2.2 for an introduction to some important concepts in stratospheric dynamics.

2.1 Chemical concepts

2.1.1 Measures of atmospheric composition

The volume mixing ratio (VMR) of a gas X is the number of moles of X per mole of air. Because the VMR for trace gases is typically small, it is usually multiplied by 10^6, 10^9, or 10^{12} to obtain parts per million by volume (ppmv), parts per billion by volume (ppbv), or parts per trillion by volume (pptv), respectively. The VMR of a gas has the virtue of remaining constant when the air density changes (as happens

when the temperature or the pressure changes).

The number density of a gas X is defined as the number of molecules of X per unit volume of air. It is expressed commonly in units of molecules per cubic centimeter (molecules cm^{-3}). The number density of gas X is generally denoted [X].

2.1.2 Chemical production and loss

Chemical reactions in the stratosphere lead to the production and loss of chemical species. The chemical reactions underlying the presence of stratospheric O_3 are described by the Chapman mechanism, which is presented here based on the treatment by Jacob (1999, pp. 164–191). The production of ozone begins with the photolysis of O_2 yielding two O atoms:

$$O_2 + h\nu \longrightarrow O + O \quad (\lambda < 240\,\text{nm}). \tag{R1}$$

In order to overcome the bond energy of the O_2 molecule the photons must have wavelengths less than 240 nm. Such high energy ultraviolet photons are present in the solar spectrum at high altitudes. The highly reactive O atoms combine rapidly with O_2 to form ozone

$$O + O_2 + M \longrightarrow O_3 + M, \tag{R2}$$

where M is a third body, usually the abundant N_2 or O_2 molecule, required to conserve energy.

Photolysis of O_3 produces O atoms:

$$O_3 + h\nu \longrightarrow O_2 + O. \tag{R3}$$

This reaction is not a terminal sink for O_3 since the O atom produced may recombine with O_2 by (R2) to regenerate O_3. This null cycle effectively absorbs solar UV radiation, while conserving O_3. For O_3 to be lost, the O atom must undergo another

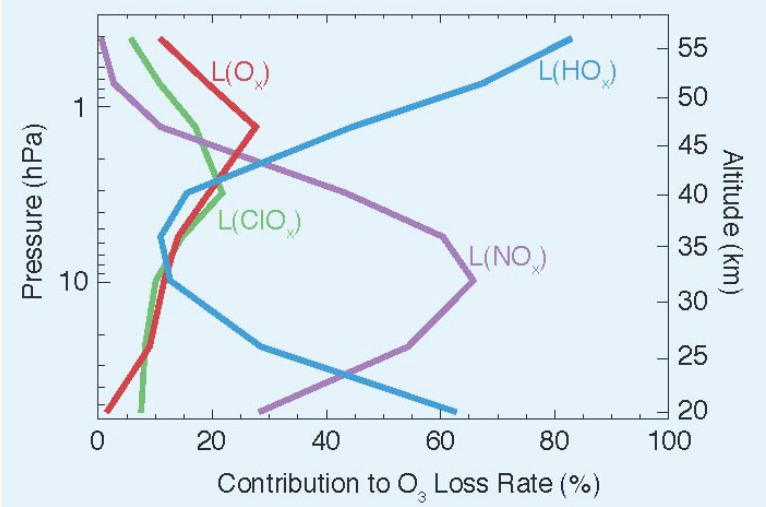

Figure 2.1: Percent contribution to total ozone loss in midlatitudes from catalytic cycles. Reproduced from IPCC/TEAP (2005).

reaction, which in the Chapman mechanism is

$$O_3 + O \longrightarrow 2\,O_2. \tag{R4}$$

In the years after the formulation of the Chapman mechanism, it was discovered that the loss of O_3 through (R4) is much less important than that resulting from reaction with other trace gases. Specifically, O_3 is depleted through a number of catalytic reactions, involving oxides of hydrogen, nitrogen, chlorine and bromine.

A general catalytic ozone loss cycle is given by the chemical reaction set

$$\begin{aligned} X + O_3 &\longrightarrow XO + O_2 \\ XO + O &\longrightarrow X + O_2 \\ \text{net} : O + O_3 &\longrightarrow 2\,O_2 \end{aligned}$$

where X acts as a placeholder for a catalytic agent such as H, OH, Cl, Br or NO.

Catalytic cycles can be very efficient, even if the concentration of the catalyst X is several orders of magnitude smaller than that of O_3. The efficiency of such a cycle regarding the destruction of O_3 is given by the number of times the cycle repeats

itself before the catalyst is eventually lost. Figure 2.1 shows the fraction of total O_3 loss at midlatitudes due to catalytic cycles involving the radical families shown. In the midlatitudes (and tropics), nitrogen oxide radical (NO_x) cycles are the dominant sink for O_3 in the middle stratosphere, whereas hydrogen oxide (HO_x) radical cycles dominate in the lower and upper stratosphere.

In the polar latitudes, different catalytic cycles can lead to large amounts of O_3 loss, such as that associated with the Antarctic "ozone hole". The cold temperatures of the polar vortex allow the formation of polar stratospheric clouds (PSCs), which provide surfaces for heterogenous chemical reactions which transform chlorine reservoir species (such as $ClONO_2$ and HCl) into active forms (such as Cl_2 and ClO). When sun returns to the polar region in spring, O_3 depletion occurs through catalytic cycles involving active chlorine, such as the following (Jacob, 1999):

$$\begin{aligned} \mathrm{ClO} + \mathrm{ClO} + \mathrm{M} &\longrightarrow (\mathrm{ClO})_2 + \mathrm{M} \\ (\mathrm{ClO})_2 + \mathrm{h}\nu &\longrightarrow \mathrm{Cl} + \mathrm{ClOO} \\ \mathrm{ClOO} + \mathrm{M} &\longrightarrow \mathrm{Cl} + \mathrm{O_2} + \mathrm{M} \\ 2\,\mathrm{Cl} + 2\,\mathrm{O_3} &\longrightarrow 2\,\mathrm{ClO} + 2\,\mathrm{O_2} \\ \text{net} : 2\,\mathrm{O_3} + \mathrm{h}\nu &\longrightarrow 3\,\mathrm{O_2}. \end{aligned}$$

2.1.3 Chemical rate kinetics and lifetimes

The relative importance of the various production and loss reactions for any particular trace gas is determined by the rates of the reactions. Chemical reaction rates are generally proportional to the abundance of reactants. For the hypothetical reaction

$$\mathrm{X} + \mathrm{Y} \rightarrow \mathrm{P} + \mathrm{Q}, \tag{2.1}$$

the loss rate of X is equal to the frequency of collisions between molecules X and Y multiplied by the probability that a collision will result in chemical reaction. The collision frequency is proportional to the product of the number densities [X][Y]. The

reaction rate can then be expressed as

$$\frac{d}{dt}[\mathrm{X}] = -k[\mathrm{X}][\mathrm{Y}] \tag{2.2}$$

where k is an empirically determined constant of proportionality, or rate constant. In a similar manner, the rate of reactions involving any number of reactants can be expressed as the product of the concentration of each of the reactants and a rate coefficient associated with the reaction.

Photochemical processes are initiated by the absorption of a photon by a molecule XY, leading to photolysis:

$$\mathrm{XY} + \mathrm{h}\nu \rightarrow \mathrm{X} + \mathrm{Y}. \tag{2.3}$$

The rate at which molecule XY is photolyzed is expressed as

$$\frac{d}{dt}[\mathrm{XY}] = -\mathrm{J}_{\mathrm{xy}}[\mathrm{XY}], \tag{2.4}$$

where J_{xy} is the rate constant for photolytic loss of XY, or the photolysis frequency. J_{xy} is determined by the number of photons available (which defines the solar actinic flux), the ability of the molecule to absorb these photons (the absorption cross-section), and the probability that the photon absorption leads to the decomposition of the molecule (quantum yield) (Brasseur and Solomon, 2005).

The loss rate for any species defines a typical time scale over which any individual molecule exists before destruction. The chemical lifetime describes the time over which a constituent decays to 1/e of its initial value in the absence of any production processes. The chemical lifetime (τ_x) of a species X can be calculated as the number density divided by the total chemical loss rate (L_x):

$$\tau_x = \frac{[\mathrm{X}]}{L_x}. \tag{2.5}$$

Chemical loss rates are determined by the chemical kinetics of the loss reactions, with the total chemical loss rate for X given by the sum of the rates of all reactions that

destroy X.

As a simple example, stratospheric nitrous oxide (N_2O) is destroyed primarily by photolysis:

$$N_2O + h\nu \xrightarrow{J_{N_2O}} N_2 + O \quad (\lambda < 200\,\text{nm}) \tag{R5}$$

and the lifetime of N_2O is thus given by

$$\tau_{N_2O} = \frac{[N_2O]}{J_{N_2O}[N_2O]} = \frac{1}{J_{N_2O}}, \tag{2.6}$$

i.e., the reciprocal of the photolysis frequency J_{N_2O}. The photolysis frequency itself depends on the solar actinic flux, so the lifetime of N_2O decreases with increasing altitude (as the overhead amount of O_3 decreases and so the UV actinic flux increases). The lifetime of N_2O is on the order of 1000 years in the lowermost stratosphere, decreasing to one year at ∼30 km, and is of the order of a month between 40 and 70 km (Brasseur and Solomon, 2005).

Chemical lifetimes for all stratospheric trace gases can be calculated with knowledge of the rates of their loss reactions, i.e., the concentrations of the species with which they react, the rate coefficients of the loss reactions, and knowledge of the atmospheric conditions upon which the rate constants depend, including actinic flux, temperature and pressure. Chemical lifetimes for many stratospheric trace gases are available for reference in Brasseur and Solomon (2005).

2.1.4 Chemical families

The concept of chemical families is often used to simplify the complex interrelationships of stratospheric chemistry. An examination of the Chapman mechanism exemplifies the usefulness of the family concept.

Laboratory tests show that the rates of the four reactions of the Chapman mechanism are of very different magnitude. Reactions (R2) and (R3) are fast, and the lifetimes of O and O_3 are consequently short, on the order of seconds to minutes throughout the stratosphere. The reactions (R1) and (R4) are relatively slow com-

pared to (R2) and (R3). Thus, the Chapman mechanism predicts a rapid cycling between O and O_3 through (R2) and (R3), and a slower cycle between O_2 and (O + O_3) by (R1) and (R4). Because of the rapid cycling between O and O_3, it is useful to define the chemical family odd oxygen $O_x = O_3 + O$, which is produced by (R1) and consumed by (R4). Reactions (R2) and (R3) are said to control the partitioning of O_x. The use of chemical families allows a clearer distinction to be made between reactions that represent net and gross production and loss terms over the time scale considered. For example, (R3) is a gross, but not a net loss term for O_3 over time scales longer than a few seconds, because nearly all the O atoms which are formed by this reaction rapidly reform O_3. On the other hand, (R4) (or any one of the catalytic O_3 loss cycles) produces a net loss of O_3 over an extended time scale (Jacob, 1999; Brasseur and Solomon, 2005).

All chemical families share some important characteristics. Chemical families originate from a particular source gas, whose destruction by photolysis or oxidation leads to the formation of a member of the chemical family. Members of a family cycle between themselves rapidly compared to the rate at which the family is created or destroyed. Finally, a chemical family is destroyed either through chemical loss of a member of the family that doesn't create another member of the family, or through transport out of the stratosphere.

The members of a chemical family can be generally divided into two groups, based on their respective lifetimes. Reservoir species refer to those family members with relatively long lifetimes, while radical species are short lived, and more highly active in chemical reactions. The nitrogen oxide ($NO_x = NO + NO_2$), hydrogen oxide ($HO_x = OH + HO_2$) and chlorine oxide ($ClO_x = Cl + ClO$) families responsible for catalytic ozone loss are examples of radical species. Ozone loss is slowed by the conversion of radical species into reservoir species such as HNO_3, H_2O and HCl.

2.2 Dynamical concepts

2.2.1 The vertical temperature profile

The absorption of solar UV radiation by O_3 (and to a lesser degree O_2) leads to significant heating of the middle atmosphere. The atmosphere is also heated from below by the surface of the Earth, which is heated itself through the absorption of solar radiation. The atmosphere loses heat through the emission of infrared radiation by a number of trace gases (including CO_2, H_2O and O_3) into space. The temperature profile that results from surface and O_3 heating has temperature maxima at the surface and at approximately 50 km altitude.

The vertical temperature profile is used to partition the atmosphere vertically into layers based on changing temperature gradients. The layers are called "spheres" and each boundary between spheres is a "pause". The lowest layer, the troposphere, exhibits generally decreasing temperature with increasing altitude up to the tropopause. The temperature and location of the tropopause vary with latitude and season. At the equator, its mean altitude is located near 18 km, and the corresponding temperature is about 190 K, while in polar regions its elevation is only about 8 km, and the temperature is roughly 220 K. Above the tropopause begins the stratosphere, which exhibits increasing temperature with altitude up to a maximum of about 270 K at the level of the stratopause at 50 km (1 hPa). Above the stratopause is the mesosphere, characterized by decreasing temperature up to the mesopause at 85 km. The stratosphere and mesosphere are collectively known as the middle atmosphere. The three lowest layers of the atmosphere, described above, constitute the homosphere, where the major constituents N_2 and O_2 are of relatively constant mixing ratio.

The thermal structure of the stratosphere, with temperature increasing with height, results in a region that is convectively stable. As a result, motion of air in the stratosphere is predominantly horizontal, along surfaces of constant entropy. Atmospheric entropy is commonly quantified by potential temperature (θ), the temperature an air parcel would have if expanded or compressed adiabatically to a reference pressure, usually taken to be 1000 hPa. Potential temperature is calculated

as

$$\theta = T\left(\frac{p_0}{p}\right)^{\kappa} \tag{2.7}$$

where p and T are the air parcel's pressure and temperature, p_0 is the reference pressure, and $\kappa = R/c_p$, where R is the gas constant for air, and c_p is the specific heat capacity at constant pressure for air. Surfaces of constant potential temperature (i.e., constant entropy) are known as isentropes.

2.2.2 Zonal winds

Heating of the stratosphere due to absorption of UV radiation by O_3 follows a seasonal and meridional pattern based on the total amount of sunlight absorbed, thus producing meridional temperature gradients with highest temperatures over the summer pole at solstice and in the tropics at equinox.

Horizontal temperature gradients imply horizontal pressure gradients, which would lead to horizontal flow along the gradient in the absence of other forces. All atmospheric flows are, however, subject to the Coriolis force, associated with the rotation of the Earth. The Coriolis force acts in a direction perpendicular to the velocity of the flow, thus a meridional pressure-gradient force can be balanced by the Coriolis force associated with a zonal flow. This relationship is known as "geostrophic balance".

Through geostrophic balance, stratospheric horizontal temperature gradients lead to eastward flow in the winter hemisphere, and westward in the summer hemisphere. Stratospheric winds thus change direction twice per year due to the change of temperature gradients with season (e.g., Wunch et al., 2005, and references therein).

2.2.3 Atmospheric waves

Disturbances to the zonally symmetric flow predicted by geostrophic balance are produced by the presence of atmospheric waves. Waves owe their existence to a restoring force, which can be produced in the atmosphere by the Coriolis force (leading to inertial waves) or by buoyancy in a stably stratified fluid (leading to gravity waves). In the first case, the force acts to oppose horizontal displacements, while in the second

case, it opposes vertical displacements.

Rossby waves

Rossby waves are the waves of largest scale and amplitude in the stratosphere, and are hence of the greatest importance for producing large-scale variability. Rossby waves owe their existence to meridional gradients in a derived quantity known as potential vorticity (PV). Following the treatment of Shepherd (2003), a simple picture of Rossby wave theory is formulated here based on the conservation of PV. In a simple model of two-dimensional incompressible motion on the surface of a sphere, PV, here denoted q, consists of the sum of vorticity associated with the motion (relative vorticity, ζ) and that associated with the Earth's rotation (planetary vorticity, f):

$$q = \zeta + f = \hat{\mathbf{z}} \cdot \nabla \times \mathbf{v} + 2\Omega \sin \phi \tag{2.8}$$

where $\hat{\mathbf{z}}$ is the unit vector in the direction of the local vertical, $\mathbf{v}$ is the three dimensional velocity vector, Ω is the angular frequency of the Earth's rotation and ϕ is latitude. In the absence of motion, $q = f$, and PV contours are zonally symmetric, describing a northward PV gradient at all latitudes.

A simple schematic of Rossby wave propagation is shown in Figure 2.2. Two PV contours (with constant and equal q) are shown, one horizontal, and one oscillatory. The horizontal line represents a PV contour for a resting flow, which lies along a constant latitude ϕ. The wavy line shows this q contour displaced meridionally by a sinusoidal wave. Where the q contour has moved northward, planetary vorticity f is increased. Conservation of q following the motion implies (through Equation 2.8) that the relative vorticity ζ is negative. In contrast, where the q contour has moved southward, ζ is positive. These disturbances to the ζ field induce, respectively, clockwise and counterclockwise velocities, with meridional velocities as shown by the arrows. The action of these arrows on the background q distribution is such as to cause the phase of the wave to move to negative latitude, i.e., westward. Thus Rossby waves, which are characterized by large-scale perturbations to the otherwise zonally

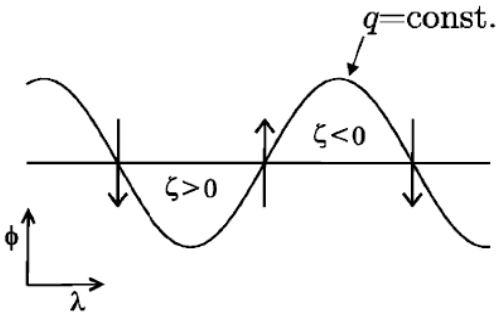

Figure 2.2: Schematic of Rossby wave propagation reproduced from Shepherd (2003).

symmetric PV field, always propagate to the west relative to any background wind.

The Rossby waves with the largest amplitudes are forced by the topography of the Earth's surface and by the thermal forcing associated with land-sea temperature contrasts, and tend therefore to be stationary with respect to the Earth's surface. Since Rossby waves propagate westward relative to the background wind, these stationary Rossby waves require an eastward wind in order to exist. In contrast, transient Rossby waves, forced primarily by tropospheric dynamical activity, travel with respect to the Earth's surface: both westward- and eastward-moving disturbances are found (Andrews et al., 1987).

In addition to horizontal propagation, Rossby waves can also propagate vertically. Charney and Drazin (1961) showed, however, that the vertical propagation of Rossby waves is dependent upon the zonal winds. The Charney-Drazin criterion states that in a zonal mean zonal wind $\bar{u}$, only waves that satisfy the relationship

$$0 < \bar{u} - c < \bar{u}_c \tag{2.9}$$

(where c is the phase speed of the wave) can propagate vertically. That is, the negative of the phase speed of the wave relative to the mean wind, $\bar{u} - c$, must be positive (eastward) but less than a critical value $\bar{u}_c$.

Zonal winds are predominantly eastward in the troposphere, and as such, Rossby waves propagate freely. Stationary Rossby waves can propagate into the stratosphere

only when the stratospheric winds are eastward, that is, in winter. The critical zonal wind speed ($\bar{u}_c$) beyond which stationary Rossby waves cannot exist depends on wavelength and is greater for longer waves. For typical wintertime stratospheric conditions, only the largest-scale waves, of zonal wavenumbers 1-3, can propagate into the stratosphere (Andrews et al., 1987). Thus, the wintertime stratosphere is disturbed continuously by large-scale stationary Rossby waves. The summer stratosphere is, in contrast, relatively undisturbed by stationary waves, which propagate only up to the zero-wind line, which normally lies between 15 and 20 km during summer (Wagner and Bowman, 2000).

The summer stratosphere is, however, subject to disturbance by another class of large-scale waves. Normal mode Rossby waves, or free modes, are a special class of unforced waves that correspond to a natural mode, or harmonic, of the atmosphere. Normal mode Rossby waves are traveling waves, the strongest of which have periods of 5-, 10- and 16-days (e.g., Salby, 1981). In an analysis of modelled chemical fields, Pendlebury et al. (2008) show that variations in chemical species are well-correlated with waves of 5- and 10-day periods between 30 and 60 km, implying that normal mode Rossby waves are a significant source of variability in the summer stratosphere.

2.2.4 Wave breaking

As atmospheric waves propagate upwards, wave amplitudes grow due to the decrease in atmospheric density. Amplitude growth cannot continue unabated. Eventually, the perturbation to the background state induced by the wave becomes so large that the atmosphere becomes unstable. When this happens, the wave "breaks", in a process analogous to the breaking of water waves on a beach (McIntyre and Palmer, 1983).

Wave breaking by gravity and Rossby waves is known to be very important in driving the large-scale mean meridional circulation of the atmosphere. Wave breaking transfers momentum from the wave to the atmosphere: this momentum transfer is called "wave drag". In order for this process to conserve angular momentum, wave breaking induces meridional motion: for example, a decelerated eastward flow can conserve angular momentum by shifting poleward, bringing it closer to the axis of

Earth's rotation. The result of Rossby wave breaking is thus poleward meridional flow, which drives the Brewer-Dobson circulation, to be discussed in §3.1.2.

Chapter 3

Stratospheric trace gas distributions and variability

The stratosphere is characterized by variations of trace gas abundances in both space and time. This chapter aims to describe how chemical and dynamical processes (many of which were introduced in Chapter 2) lead to climatological mean spatial distributions of trace gases (§3.1), and variations about the climatological mean state (§3.2).

3.1 Trace gas distributions

3.1.1 Transport

The combined effects of chemistry and dynamics on the mixing ratio χ of any chemical species can be expressed through the evolution equation, which for a fixed location can be written

$$\frac{\partial \chi}{\partial t} = -\mathbf{v} \cdot \nabla \chi + S \tag{3.1}$$

where t is time, $\mathbf{v}$ is velocity, ∇ is the three-dimensional spatial gradient operator, and S represents the net rate of chemical change, production minus loss (Brasseur and Solomon, 1984; Shepherd, 2003). Thus, the time rate of change of χ, $\partial\chi/\partial t$, can be nonzero due to transport, $-\mathbf{v} \cdot \nabla \chi$, or because of chemistry S. In a climatological

mean state, a steady-state balance can be assumed with $\partial\chi/\partial t = 0$. Thus, the steady state abundance of χ is produced by a balance between transport and chemistry.

The spatial gradients of χ are important in determining the effect of transport. When winds blow from higher to lower values of χ, $-\mathbf{v}\cdot\nabla\chi > 0$ and transport acts as a local source of χ. Likewise, when winds blow from lower to higher values of χ, $-\mathbf{v}\cdot\nabla\chi < 0$ and transport acts as a local sink for χ. When the local spatial gradient of χ is zero, winds have no effect on the local abundance of χ.

The relative importance of dynamics and chemistry in the balance that determines the abundance of χ depends on a comparison of the time scales associated with the processes. For the chemical processes, the chemical lifetime (τ_{chem}) introduced in §2.1.3 is used as an indicator of the rate at which the term S relaxes the balance toward "photochemical equilibrium". Comparable time scales of dynamical motions (τ_{dyn}), i.e., estimates of the time for advective processes (mean motions plus eddies) to transport χ by some significant amount[1] can be calculated with knowledge of the strength of the net transport in the zonal, meridional, and vertical directions. Three different cases (following Brasseur and Solomon, 2005) can be identified to characterize the effects of the competition between dynamics and chemistry in determining the distribution of a trace gas:

$\tau_{\text{chem}} \ll \tau_{\text{dyn}}$ Under these circumstances, the trace gas will be in photochemical equilibrium, and the effects of dynamics will not be directly important. Dynamics may still be important in an indirect fashion through its influence on temperature or on other species that are chemically linked to the species in question. For example, a very short-lived species such as NO_2 may be produced or destroyed by reaction with a longer-lived species that does depend on transport, such as N_2O.

$\tau_{\text{chem}} \gg \tau_{\text{dyn}}$ In this case, trace gases are passively advected and in the absence of localized sources or sinks will become well mixed due to the dispersive effects of

[1]Example criteria for the calculation of dynamical time scales include the time taken for transport through approximately a scale height in the vertical or from the equator to pole meridionally (Andrews et al., 1987).

transport (Andrews et al., 1987). The major species O_2 and N_2, for example, are so long lived in the middle atmosphere that they are thoroughly mixed by atmospheric motions, resulting in concentrations that are uniform throughout the homosphere.

$\tau_{\text{chem}} \approx \tau_{\text{dyn}}$ Under these circumstances the distribution of a trace gas depends critically on both dynamics and chemistry. The time constant for meridional transport is, for example, comparable to the photochemical lifetime of CH_4 and N_2O in the upper stratosphere, so that transport in the meridional plane is expected to be quite important in determining the distribution of these species. Such species are said to be "tracers" of the meridional motion.

3.1.2 The Brewer-Dobson circulation

Zonal winds, introduced in §2.2.2, are a product of the balance between the Coriolis force and meridional pressure gradients, and are thus "free" solutions of the governing equations in the sense that they do not require external forcing in order to exist. In contrast, persistent vertical or meridional motions require some type of forcing, and are as a result comparatively weak (Shepherd, 2003).

The presence of mean meridional and vertical transport in the stratosphere was first inferred by the distributions of stratospheric trace gases. While O_3 is produced most abundantly in the tropics, where incident radiation is largest throughout the year, it was observed to have largest total column abundances in polar regions, far from the region of its production. Dobson (1956) proposed a dynamical explanation for the observed ozone maximum at high latitudes, namely, that stratospheric air is advected poleward in the midlatitudes, and downwards in the polar region, supplying ozone to the polar lower stratosphere. This explanation was consistent with a conceptual model proposed earlier by Brewer (1949), wherein the observed low water-vapour mixing ratios of the stratosphere could be explained by "freeze-drying" of air by upward motion through the "cold trap" of the high cold tropical tropopause. The "Brewer-Dobson" circulation thus consists of a meridional cell in each hemisphere

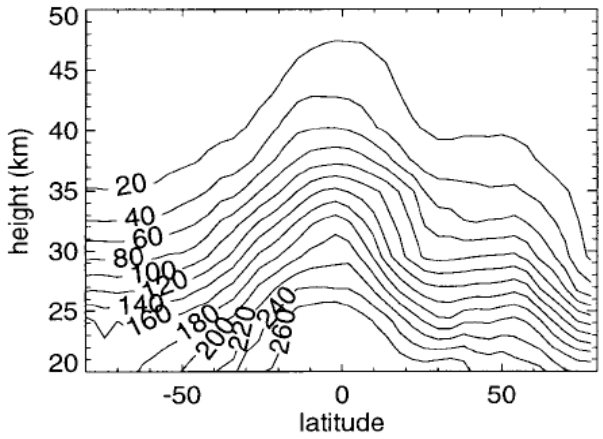

Figure 3.1: Zonal mean contour plot of Cryogenic Limb Array Spectrometer (CLAES) N_2O (ppbv) from measurements accumulated over January 1-18, 1993. Reproduced from Sparling (2000).

with rising motion across the tropical tropopause, poleward drift in the stratosphere, and by continuity of mass, a return flow into the troposphere in the extratropics.

The Brewer-Dobson circulation is especially useful for describing the meridional distribution of tracers such as N_2O (Figure 3.1). N_2O has a surface source, and a stratospheric sink, and therefore high concentrations indicate air that has only recently entered the stratosphere, while low values indicate air that has long resided there. Long-lived tracers (of tropospheric origin) display a distinct pattern of high values in the tropical latitudes, with concentrations falling off with altitude and latitude, indicative of the tropical upwelling, extratropical meridional motion, and polar downwelling produced by the Brewer-Dobson circulation (Plumb, 2002).

The Brewer-Dobson circulation is understood to be driven by the mean meridional motion produced by the breaking of Rossby waves in the middle atmosphere. As discussed in §2.2.4, Rossby wave breaking deposits momentum, and leads to poleward mean meridional flow. Stationary Rossby waves, which are able to propagate into the middle stratosphere of the winter hemisphere, drive a poleward flow there. Synoptic-scale Rossby waves dissipate in the lower stratosphere of both hemispheres in all seasons, driving a persistent poleward circulation in both hemispheres of the lower stratosphere (Shepherd, 2007). The poleward flows arising from the wave drag

induced by both synoptic- and planetary-scale Rossby waves leads, by mass balance, to upwelling in the tropics and downwelling in the extratropics (Haynes et al., 1991; Holton et al., 1995).

In addition to its effect on the mean meridional circulation, Rossby wave breaking leads to intense and large-scale two-way transport of air, or stirring. Eventually, mixing occurs as the breaking wave strings flow into long thin laminae, so that ultimately small-scale turbulence can mix the tracer throughout the domain (Andrews et al., 1987). In the stratosphere, wave motion is approximately adiabatic and so stirring and mixing is mainly quasi-horizontal, along isentropic surfaces. The region of the winter midlatitudes where the most intense Rossby wave breaking occurs has become known as the stratospheric "surf zone" (McIntyre and Palmer, 1983). The intense stirring and mixing produced by Rossby wave breaking in the surf zone acts to flatten the mean meridional distributions of chemical tracers, as seen in the winter (northern) hemisphere of Figure 3.1. In light of the fact that tracer distributions are shaped by both the mean meridional mass transport and the quasi-horizontal mixing produced by Rossby wave breaking, the Brewer-Dobson circulation, as the chemical transport circulation of the stratosphere, is understood to include both processes (Shepherd, 2002). A schematic of the meridional circulation of the atmosphere is depicted in Figure 3.2.

It is convenient (following Plumb, 2002) to regard the stratosphere as being separated into four regions: the summer hemisphere, the tropics, the wintertime midlatitude surf zone, and the winter polar vortex, as labeled in Figure 3.2. The surf zone is bounded by sharp gradients of PV and tracers (as can be seen, for example, in the N_2O contours of Figure 3.1) at the winter subtropical edge (separating it from the tropical region), and at the edge of the polar vortex. The sharp meridional gradients are often referred to as resulting from the presence of "mixing barriers" (Shepherd, 2007). The polar vortex mixing barrier has been long recognized for its importance in creating the chemical isolation needed to produce the Antarctic ozone hole. The presence of the subtropical mixing barrier has been confirmed based on the confinement of aerosols to the tropics after the eruption of Mt. Pinatubo (Trepte and Hitchman,

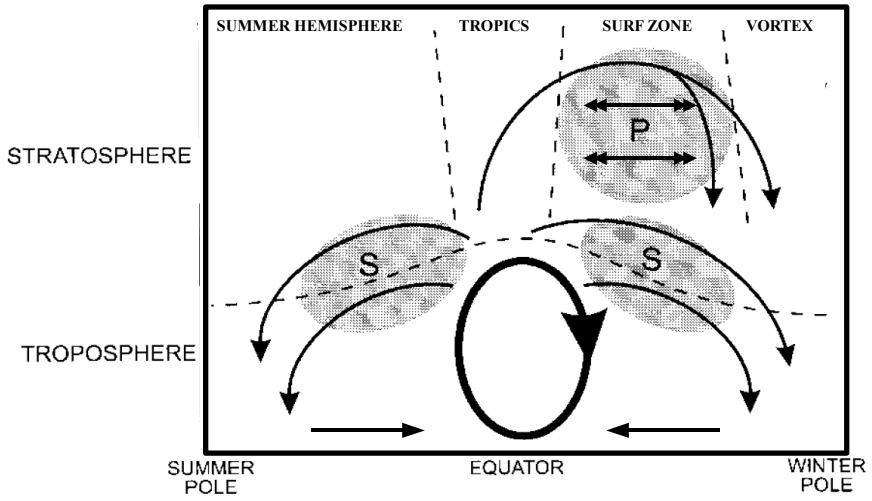

Figure 3.2: Schematic of the residual mean meridional circulation in the atmosphere, reproduced from Plumb (2002). The shaded regions (labeled "S" and "P") denote regions of breaking waves (synoptic- and planetary-scale waves, respectively), responsible for driving branches of the stratospheric circulation. The quasi-horizontal two-way transport of the surf-zone is indicated by horizontal double-headed arrows.

1992). Mixing barriers thus isolate, to some degree, the surf zone from the polar vortex and the tropics, which are characterized respectively by the downwelling and upwelling branches of the meridional circulation. The summer hemisphere is unique from the other regions in that it is relatively quiescent in terms of dynamic activity due to the blocking of Rossby wave propagation through the summer-time westward stratospheric winds.

3.2 Trace gas variability

In some of the first observational studies of stratospheric trace gas distributions, Dobson and Harrison (1926) noted that O_3 abundances over Oxford showed irregular variations on time scales of hours to days, superimposed on an annual cycle. These observations constitute some of the first evidence of short-term stratospheric trace gas variability. Variations of O_3 over much longer time periods - years to decades - have

motivated considerable scientific interest in stratospheric processes. The problem of calculating trends in O_3 amounts, and thus projecting future O_3 depletion (or recovery), is complicated by the fact that O_3 exhibits year-to-year variability, with magnitudes comparable to the long-term trend. This interannual variability acts as noise in the long-term series, and complicates the detection of long-term trends (e.g., Weatherhead et al., 1998; Weatherhead and Andersen, 2006).

This section introduces some basic concepts regarding stratospheric trace gas variability on long-term, interannual, and short-term time scales. As atmospheric variability is a vast area of research, the aim here is only to provide an introduction to some of the issues most relevant to the work of this thesis.

3.2.1 Long-term variability

On a global scale, long-term perturbations to the chemical balance of the stratosphere are produced by the various gases emitted as a result of anthropogenic activity. Figure 3.3 shows measured time series of atmospheric abundances for the trace gases CO_2, CH_4, and N_2O, and illustrates the drastic increases of these species over the last decades and centuries. All of these trace gases have lifetimes long enough that they are transported into the stratosphere, leading to changes in the stratospheric trace gas distributions.

The trace gases CO_2, CH_4 and N_2O are known as "greenhouse gases" (GHGs), since they are directly involved in the absorption and emission of infrared radiation in the atmosphere, and an increase in their atmospheric abundances leads to an increase in surface temperatures, and to a decrease in stratospheric temperatures. Through their effect on the radiative balance of the atmosphere, GHGs may lead to long-term changes in trace gas distributions, through a modification of temperature-dependent chemical reactions, or though an influence on the Brewer-Dobson circulation (e.g., Butchart et al., 2006).

The transport of CH_4 and N_2O to the stratosphere has a more direct effect on trace gases, since they act as source gases for the catalytic O_3-destroying species HO_x and NO_x,respectively (Brasseur and Solomon, 2005). In addition, the effect on O_3

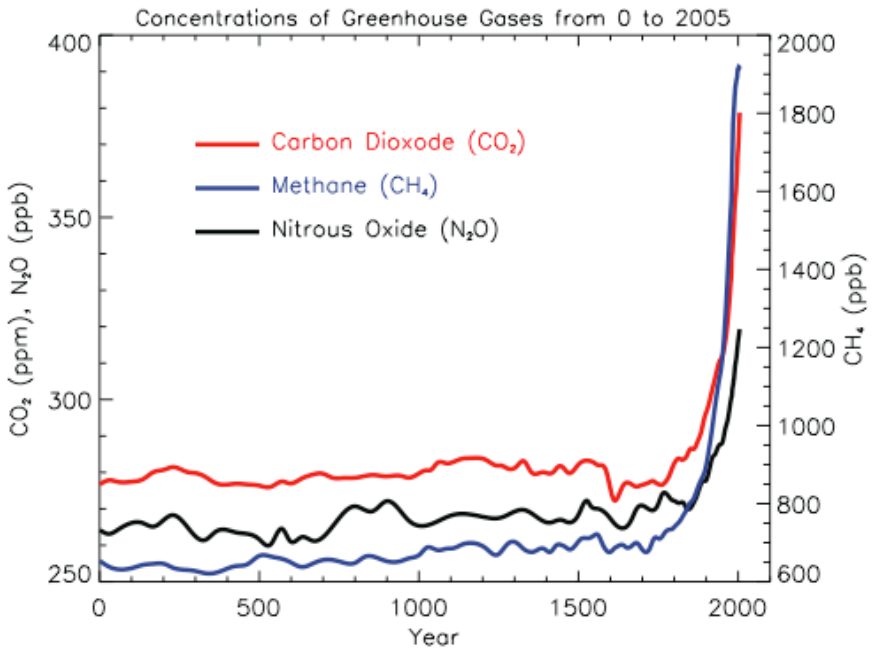

Figure 3.3: Observed increase in the atmospheric abundance of carbon dioxide (CO_2), nitrous oxide (N_2O), and methane (CH_4) at the Earth's surface, reproduced from IPCC (2007).

of ClO_x increases brought about by the emission of the chlorofluorocarbons (CFCs), especially in the special circumstances of the Antarctic polar vortex, is particularly well known (WMO, 2007).

3.2.2 Interannual variability

Interannual variability reflects deviations of the annual cycle from one year to the next. Such variability can be the result of specific forcing agents, or the unforced chaotic dynamical fluctuations of the atmosphere.

Unforced (or internal) variability in the atmosphere is due primarily to dynamics, since the dynamical equations that govern the atmosphere have solutions that do not reach steady-state, even under constant forcing (Shepherd, 2003). The quasi-biennial oscillation (QBO) of tropical stratospheric winds is an important source of internal atmospheric variability. Through its influence on stratospheric temperatures, and the Brewer-Dobson circulation, the QBO leads to interannual variability in a number of

trace gases (see Baldwin et al., 2001, and references therein).

Interannual variability of Rossby wave activity can have a large influence on trace gas distributions through its influence on the polar vortices (WMO, 2007). Stratospheric sudden warmings (SSWs) are dramatic dynamical events wherein polar temperatures undergo a rapid increase that reverses the climatological equator-to-pole gradient, causing eastward flow to become westward. SSWs play an important role in the photochemistry of O_3, especially in the Northern hemisphere. During winters without SSWs, the vortex is cold and stable, permitting the formation of PSCs which activate chlorine and lead to chemical O_3 loss in the presence of sunlight. Winters with SSWs, on the other hand, are characterized by a stratosphere that is too warm for PSCs to form, thus decreasing the potential for springtime O_3 loss.

Solar variability is known to exert a direct forcing on stratospheric O_3 abundances. The solar output of UV radiation exhibits an 11-year cycle, which corresponds to the appearance of sunspots. Solar maximum is characterized by increased sunspot activity and enhancement of UV radiation. Through the dependence of O_3 production on the photolysis of O_2, O_3 concentrations in the upper stratosphere are positively correlated with sunspot activity on the 11-year cycle, with percent changes estimated from observations to be $\sim$4% from solar minimum to maximum (Hood, 2004). A correlation between lower stratospheric O_3 and the solar cycle has also been observed. A dynamical explanation has been proposed, whereby the weak solar forcing in the upper stratosphere may influence the Brewer-Dobson circulation of the winter stratosphere (see Hood, 2004, and references therein).

The Sun also affects the Earth's atmosphere through more intermittent events like coronal mass ejections that enhance the precipitation of highly energetic particles (protons, electrons, and ions) into the mesosphere and upper stratosphere (e.g. Weeks et al., 1972; Crutzen et al., 1975). During an energetic particle precipitation (EPP) event, energetic particles ejected by the Sun are guided by the Earth's magnetic field towards the polar regions, where they ionize atmospheric constituents, leading eventually to increases in the oxidized species HO_x and NO_x. Through the creation of these catalytic O_3-destroying substances, EEP can have a direct effect upon O_3

in the upper stratosphere. When EEP-produced NO_y is transported downwards to the middle and lower stratosphere, EEP can influence middle and lower stratospheric O_3, and hence have a significant effect on total column O_3 (Randall et al., 2007). This "indirect effect" of EPP on O_3 gained considerable attention after the extraordinary period of solar activity in October–December 2003, which was followed by an unusual winter in the Northern Hemisphere in which the upper stratospheric vortex was the strongest on record during February and March (Manney et al., 2005). The combination of EPP and a stable upper stratospheric vortex led to unprecedented enhancements in stratospheric NO_y (López-Puertas et al., 2005b; Orsolini et al., 2005), and substantial reductions in O_3, during the NH late winter/spring of 2004 (e.g. Jackman et al., 2005; López-Puertas et al., 2005a; Randall et al., 2005a; Rohen et al., 2005).

Occasional major volcanic eruptions represent potentially important sources of SO_2 and HCl to the stratosphere, which directly modify the chemistry, create aerosols, and change the radiation and dynamics on interannual time scales (Coffey, 1996). SO_2 is important as a precursor to sulfate aerosol, which provides sites for heterogeneous reactions that would not otherwise occur. These heterogeneous reactions release catalytic O_3-destroying chlorine radicals from inert forms (such as HCl and $ClONO_2$), and convert reactive nitrogen species (NO, NO_2, N_2O_5, $ClONO_2$) into the more stable HNO_3. Both of these changes enhance the destruction of O_3. These direct perturbations are expected to last for a year or more after a volcanic eruption (Coffey, 1996).

3.2.3 Short-term variability

Short-term variability of trace gases occurs on time scales of hours to days. Short-lived chemical species display short-term variability primarily as a result of changing photochemical conditions. Many short-lived trace gas species display diurnal variability, cycling in response to the daily period of solar insolation. For example, diurnal variability is the major form of variability for O_3 above 40–50 km. Upper stratospheric O_3 has also been observed to be sensitive to variations in solar UV radiation

related to the 27-day rotation period of the sun (e.g., Hood, 1986; Hood and Zhou, 1998). Temperature fluctuations tied to atmospheric waves can also lead to variations in short-lived species. Long-lived species, on the other hand, react slowly to such changes in photochemical conditions, and instead show variability due to variations in transport.

Short-term variability present in global data sets (such as satellite-based measurements) of trace gases is often expressed in terms of deviations from zonal means. This concept of variability is subtly different from the idea of temporal variations at a particular location. Short-term variability expressed as deviations from a zonal mean is a function of both time and space, rather than just time. Of course, if the deviations from the zonal mean are zonally propagating, either by advection by the zonal winds or by wave propagation, the variability of a zonal band at one point in time may be comparable to the variability at a single point in that band over the time scale of the zonal propagation.

An illustrative example of the short-term zonal variability of trace gases, and its dependence on season and location, is shown by Kuell et al. (2004), in the form of zonal standard deviations of measurements made by the Cryogenic Infrared Spectrometers and Telescopes for the Atmosphere (CRISTA) instrument, which collected measurements of trace gas profiles over August 8–16, 1997, during a NASA space shuttle mission. Latitude-altitude slices of percent standard deviation were shown by Kuell et al. (2004) for the trace gas species O_3, HNO_3, N_2O, CH_4 and CFC-11. While the spatial pattern of short-term variability is different for each of the trace gas species studied, winter hemisphere variability is notably larger than that in the summer hemisphere in each case. As an example, standard deviations for N_2O are shown in Figure 3.4, which exhibit large variability in the Southern hemisphere, minimum variability in the tropics, and intermediate variability in the Northern hemisphere during August.

The meteorological conditions present in the winter hemisphere during the CRISTA observations, and the effect of these conditions on trace gas distributions, have been explored by Riese et al. (2002). Through an analysis of dynamical fields from the

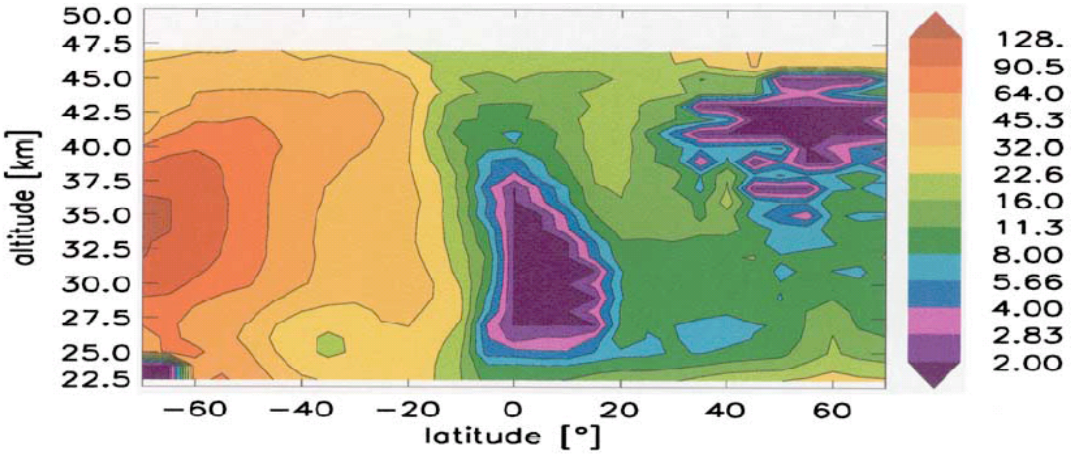

Figure 3.4: Standard deviation (in percent) from zonal means for N_2O as measured by CRISTA, August 8–16, 1997. Standard deviations were corrected by the reported statistical measurement error. Regions where the statistical measurement error dominates the measured standard deviation are colored in dark violet. Adapted from Kuell et al. (2004).

UK Met Office, Riese et al. (2002) found that the CRISTA measurements occurred during a period of exceptionally strong planetary scale (wave 1 and wave 2) Rossby wave activity in the Southern hemisphere. The wave activity described by Riese et al. (2002) can be understood to create the zonal variances shown in Kuell et al. (2004) through two related processes. Firstly, the Rossby wave activity was seen to displace the Southern hemisphere winter vortex from the pole, and elongate the shape of the vortex. This is illustrated in Figure 3.5, which shows a snapshot of the geopotential height fields, and N_2O VMRs at 10 hPa for 11 August 1997. The significant perturbation from zonal symmetry thus leads to variance in chemical fields in any zonal band, with the largest variances expected to be found at latitudes with the highest mean meridional gradients. Secondly, as a result of the displacement of the polar vortex, with the vortex edge reaching latitudes as low as 30°S, considerable trace gas flux from the tropics toward midlatitudes was observed, mainly in the form of a pronounced planetary-scale "streamer" advected out of the tropics and along the vortex edge (also seen in Figure 3.5). Through this process, tropical air with high N_2O VMR is seen to be drawn into midlatitudes. Together, the processes of vortex

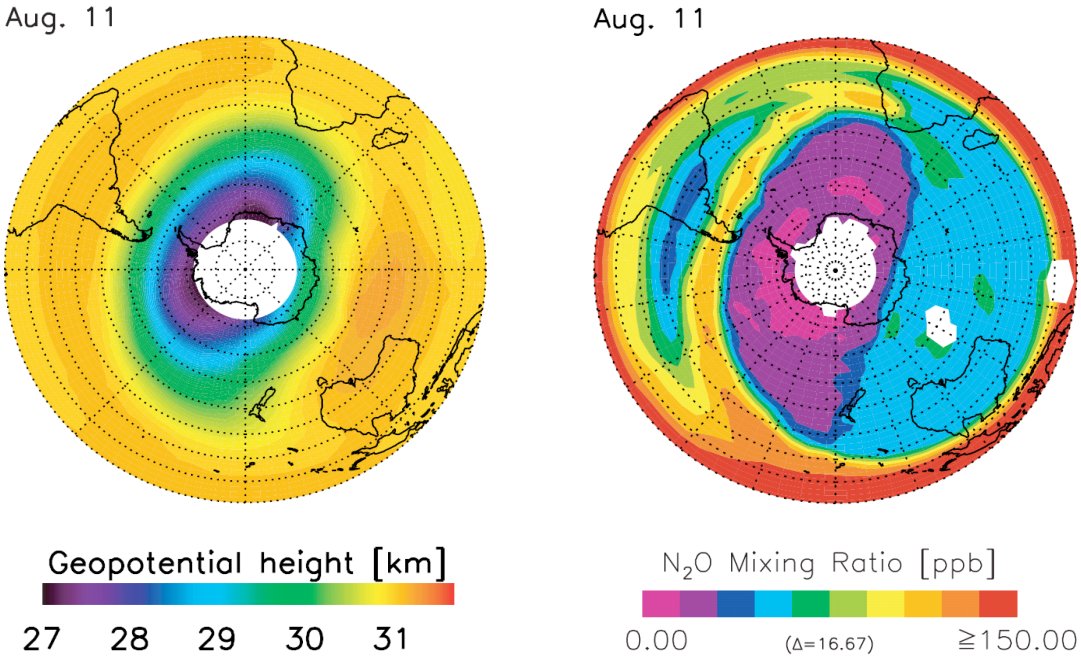

Figure 3.5: Dynamical and chemical conditions at 10 hPa in the Southern Hemisphere as measured on August 11, 1997 by CRISTA. Left: geopotential height fields derived from CRISTA measurements of pressure and geometric height. Right: interpolated N_2O VMRs. Adapted from Riese et al. (2002).

deformation and transport from tropical latitudes lead to extremely large variability about any midlatitude zonal band, as each latitude band contains both vortex air (with very low N_2O VMR) and tropical air (with very high VMR).

The enhanced variability in N_2O (and other trace gases) in the winter midlatitudes is consistent with the idea of the winter surf zone introduced in §3.1.2, where Rossby wave propagation and breaking leads to zonal asymmetries (reversible motions) and wave breaking (irreversible mixing), respectively. Another representation of the variability induced by Rossby waves in the surf zone is shown in Figure 3.6, which shows (in the left-hand panels) scatterplots of CLAES N_2O VMR χ at 30 km versus latitude. Right-hand panels show equal-area probability distribution functions (PDFs) $P(\chi)$ versus χ, turned sideways, which can be seen to conveniently partition the stratosphere into the regions of vortex, surf zone, tropics and summer extratrop-

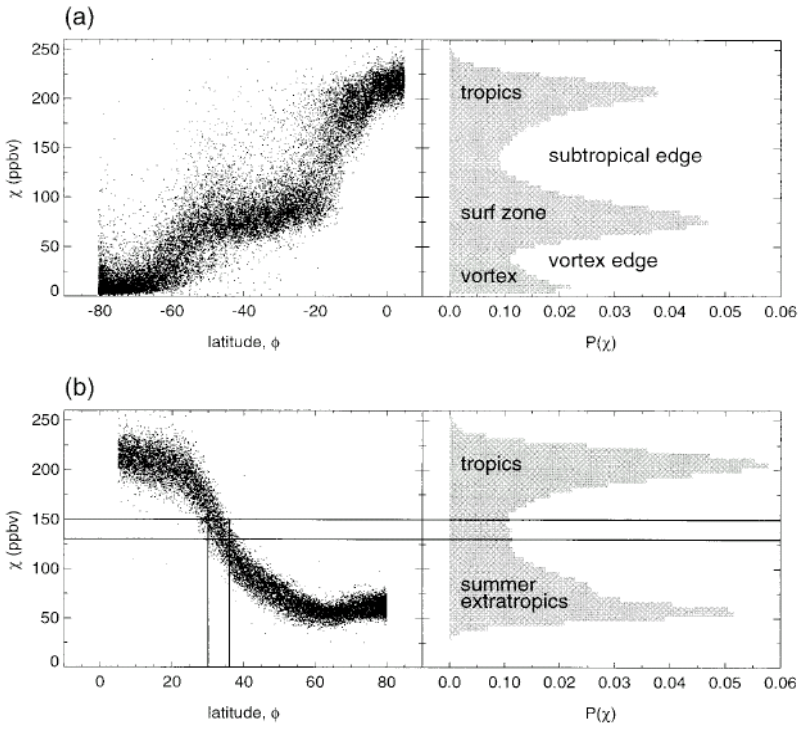

Figure 3.6: Left panels show scatterplots of CLAES N_2O mixing ratio χ at 30 km versus latitude; right panels show equal-area probability distribution functions (PDFs) $P(\chi)$ versus χ, turned sideways. (a) Southern Hemisphere winter, August 17 to September 16, 1992. (b) Northern Hemisphere summer, July 19 to August 10, 1992. Reproduced from Sparling (2000).

ics. The variability of the surf zone is quite apparent from the left-hand panel of Figure 3.6, where VMRs within the surf zone are seen to span the range of mixing ratios measured over the full Southern hemisphere, suggesting the surf zone contains air recently transported from both the tropics and the vortex regions. It is illustrative to compare the surf zone depicted in Figure 3.6 to that of Figure 3.1. While the zonal mean of N_2O VMRs leads to the flattened isopleths of N_2O in the surf zone of Figure 3.1, Figure 3.6 shows that the instantaneous variability about a latitude circle may be large.

As introduced in §2.2.3, the Rossby waves which drive the variability of the surf zone are characterized by large-scale meridional undulations in the derived quantity PV. Since PV is conserved, it behaves in a manner similar to passive chemical tracers such as N_2O. Thus, the variability seen in N_2O within latitude bands may be drasti-

cally reduced by segregating the chemical abundances not by geographical latitude, but by some "dynamical coordinate", based on the contours of PV. One such widely used coordinate is equivalent latitude, which is the latitude that would enclose the same area between it and the pole as a given PV contour (Butchart and Remsberg, 1986). One would expect that if the N_2O VMRs of Figure 3.6 were plotted versus equivalent latitude, the scatter of values in the winter hemisphere surf zone would be much reduced.

As would be expected based on the filtering of Rossby waves, the summer hemisphere shows less chemical zonal variability than the winter hemisphere, as illustrated in both Figures 3.4 and 3.6, although some variability remains. Summertime extratropical variability has been studied by Ehhalt et al. (1983), and characterized by "equivalent displacement height" (EDH). The EDH is the vertical distance that the time-mean vertical tracer profile would need to be displaced in order to produce the locally observed variance. The use of the EDH to characterize tracer variability should not be taken to imply that it is primarily vertical displacements that act to produce the variance. In fact, the magnitude of the EDH for a number of species (based on balloon-borne measurements during Northern hemisphere summer) is an order of magnitude greater than that for the observed potential temperature, indicating that horizontal (i.e., meridional) displacements are in fact more important for explaining the observed variance (Ehhalt et al., 1983).

Two mechanisms have been proposed to explain the variance observed by Ehhalt et al. (1983) and characterized using the EDH. Ehhalt et al. (1983) suggested that traveling normal-mode Rossby waves may be responsible. According to Andrews et al. (1987), the observed variances of the measured species could be explained only by meridional displacements of 1000 km. Hess and Holton (1985) argued that the magnitude of normal-mode Rossby waves is insufficient to produce the observed variance, and suggested that some of the tracer variability observed during summer is not due to wave activity, but rather represents variance generated by the large meridional and vertical parcel displacements associated with the springtime breakup of the polar vortex that become "frozen in" and advected by the symmetric easterly

circulation of the summer stratosphere.

Modern observational evidence has been collected in support of the frozen-in theory of summer variance. Observations collected by the Aura Microwave Limb Sounder (MLS) during Northern hemisphere spring and summer 2005 show a streamer of low-latitude (high N_2O, low H_2O) air drawn into high latitudes and confined in a tight anticyclone, then advected intact by the westward summer winds through late August (Manney et al., 2006). Similarly, Akiyoshi et al. (2002) report low-concentration N_2O regions in high latitudes of the lower stratosphere observed by the Improved Limb Atmospheric Spectrometer (ILAS) for one and a half months after the Arctic vortex breakdown in May 1997. The observed persistence of N_2O anomalies through spring and summer has been shown to be well simulated by chemical transport models (Orsolini, 2001; Akiyoshi et al., 2002).

Persistence of N_2O anomalies through the summer is understood to be due in large part to the long chemical lifetime of N_2O. Species with shorter lifetimes would be expected to relax photochemically after the breakup of the polar vortex. Indeed, Manney et al. (2006) report that observed O_3 anomalies at 30 km produced by the vortex breakup disappeared by April.

Satellite observations of ozone (e.g., Miles et al., 1994; Park and Russell III, 1994; Hoppel et al., 1999) have, however, revealed significantly enhanced summertime ozone variability (on the order of 5-20%) confined to a distinct region between 20 and 30 km, and 50° and 75° in latitude persisting throughout both Northern and Southern Hemisphere summers. Studies have concluded that this variability is due to the breaking of westward-propagating planetary waves Hoppel et al. (1999); Wagner and Bowman (2000). Luo et al. (1997) have shown that intense measured zonal asymmetries in O_3, HCl, NO and NO_2 can be understood to be a product of meridional motions due to Rossby wave activity acting on horizontal gradients along isentropic surfaces produced by the (near) constant sunlight conditions of the polar summer. In contrast, the variability of CH_4 and HF in the same region and time are shown to be small, due to the weak meridional gradients for these species.

In summary, the short-term variability of long-lived stratospheric tracers, ex-

pressed through the variance about zonal means, is a product of large-scale dynamical motions acting on time-mean horizontal gradients of chemical tracers, which are themselves a product of both photochemical and dynamical processes. Patterns of short-term variability are thus different for different trace gases, with different chemical lifetimes and hence different horizontal gradients.

Chapter 4

Balloon-borne emission radiometer measurements of HNO_3

4.1 Introduction

One of the central aims of the MANTRA mission was to evaluate observational evidence for changes in the components of stratospheric total reactive nitrogen (NO_y) (Strong et al., 2005). Towards this aim, two emission radiometer instruments played an important and unique role as part of the balloon payload. The instruments measure atmospheric thermal infrared emission from a number of trace gas species in in the 8–14 μm (715–1250 cm^{-1}) range, including emission at 11.3 μm due to nitric acid (HNO_3). The radiometers are closely related in design to instruments used to make some of the earliest measurements of HNO_3, and as such, represent a link to those pioneering measurements. Furthermore, instruments originally built as part of the same suite of instruments as those flown on the MANTRA flights were flown on balloon flights dating back to 1990. As a result, the emission radiometer raw data set spans a period of twelve years.

HNO_3 is the most abundant member of the reactive nitrogen family (known collectively as NO_y given by the sum of NO, NO_2, HNO_3, $2\times N_2O_5$, $ClONO_2$, $BrONO_2$ and HO_2NO_2) in the lower stratosphere, where it acts as a reservoir for NO_x. Since NO_x catalysis is the major source of O_3 loss in the midlatitude lower stratosphere

(see Figure 2.1), changes in HNO_3, due to changes in NO_y or its partitioning, have important implications for O_3 through related changes in NO_x.

There are relatively few long-term records of stratospheric HNO_3. The first long-term, near-global HNO_3 data set was produced from measurements made by the Microwave Limb Sounder (MLS) on board the Upper Atmosphere Research Satellite (UARS) (Santee et al., 1999). In an analysis of MLS HNO_3 observations between 1993 and 1997, Randel et al. (1999) reported decreases of approximately 2% per year in lower stratospheric, extra-tropical HNO_3. The analysis excluded from consideration the years 1991–1992 in an attempt to remove the effects of elevated aerosol loading due to the Mt. Pinatubo eruption of 1991, which is known to lead to elevated HNO_3 through heterogeneous reactions such as

$$N_2O_5 + H_2O_{(\ell)} \longrightarrow 2\,HNO_3, \tag{R1}$$

which converts N_2O_5 to HNO_3 in the presence of aerosols. Since aerosol loading returned to near background levels by 1993, it was therefore unclear whether the trend reported by Randel et al. (1999) was due to a slow relaxation of HNO_3 to pre-Pinatubo levels, or some underlying long-term trend in HNO_3 abundances.

The timing of the emission radiometer measurements is well suited to address this possible long-term trend, since measurements were taken before, and well after the Mt. Pinatubo eruption. While the sample size of the emission radiometer observations is small, the data set has the asset that measurements were taken during summer, when dynamical variability in the stratosphere is weak. In order to better understand the variability to be expected in HNO_3 profiles, simulated fields from the Canadian Middle Atmosphere Model (CMAM) have been investigated (see Appendix A for a description of the CMAM and the simulation results used in this thesis). Monthly mean CMAM HNO_3, at the vertical level of peak VMR ($\sim$24 km) and the model gridpoint closest to Vanscoy, Saskatchewan (52°N, 107°W), are shown in Figure 4.1 for ten years of the CMAM simulation. The seasonal cycle is apparent in the model output, with maximum HNO_3 in winter, and minimum in summer. The 2σ monthly

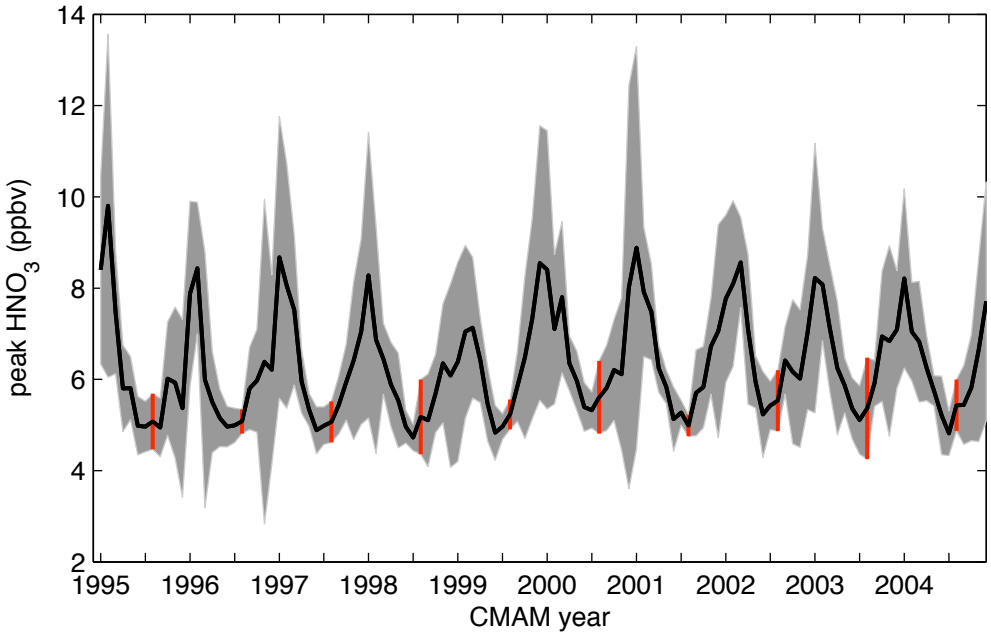

Figure 4.1: Time series of monthly mean CMAM HNO_3 VMR at ∼24 km and model gridpoint closest to geographical location of Vanscoy, Saskatchewan (52°N, 107°W). Monthly variability (2σ) is represented by shading, and August variability values are highlighted by red vertical lines.

variability of the 18-hourly fields is shown by the shaded region. The average standard deviation of August HNO_3 values (marked by red vertical lines) is approximately 6% of the mean, while in winter the same metric is 16%. Summer measurements are therefore the most useful measurements for assessing long-term changes in HNO_3.

The material of this chapter[1] represents the end result of work aiming to produce HNO_3 profiles from the emission radiometer raw data set, in order to assess long-term trends in midlatitude HNO_3. The history of the instruments and the instrument design are described in §4.2. The retrieval method used to produce vertical profiles from the radiance observations is described in §4.3. The quantity and quality of the raw data are outlined in §4.4. The retrieved profiles are presented in §4.5, and compared with retrievals from the ACE-FTS satellite instrument. The emission radiometer results are presented within the larger context of other satellite

[1]Much of the work of this chapter was published by Toohey et al. (2007).

measurements, including the post-Pinatubo MLS observations, in §4.6.

4.2 Instrument design and history

The emission radiometer, as a general instrument design, was first developed and used for atmospheric remote sensing by Pick and Houghton (1969). The basic design consists of a cryogenically cooled infrared detector, with a spectral band-pass filter attenuating the incoming radiation, allowing the measurement of the thermal emission features of atmospheric constituents.

The radiometer measures the cumulative emission along an upward-looking slant path. Vertical profile information is obtained by mounting the instrument on a balloon platform, and taking a series of measurements while the balloon ascends. At the beginning of the balloon ascent, the instrument views a slant path through the whole atmosphere, and the radiance measurements are at a maximum. As the instrument ascends, the atmosphere below the instrument is excluded from the slant path, and the collected radiance decreases.

Soon after the discovery of HNO_3 in the stratosphere (Murcray et al., 1968), cryogenically cooled radiometer instruments were built with the specific goal of measuring this stratospheric trace gas. The instruments were furnished with discrete band-pass filters to sample sections of the HNO_3 emission band at 11.3 μm, where there is minimal interference from other gas species (Murcray et al., 1973). Radiance estimates were derived from these band-pass measurements and a careful filter calibration. Balloon and aircraft borne emission radiometers of this design were used to begin to illuminate the vertical (Murcray et al., 1973) and horizontal (Murcray et al., 1975) structures of HNO_3.

The Atmospheric Environment Service (AES, later the Meteorological Service of Canada) built a suite of emission radiometers in the early 1970's. The AES emission radiometers were flown on a number of balloon flights, including the Stratoprobe series, that included measurements of NO_2 and HNO_3, leading to early estimates of the northern hemisphere midlatitude NO_y budget (Evans et al., 1976a, 1977, 1981,

1982a,b, 1985; Kerr and McElroy, 1976; Kerr et al., 1982; Ridley et al., 1976; Ridley et al., 1984).

Based on the success of the original AES emission radiometers, a second generation of radiometer was created in the 1980's. A number of these instruments were flown on a series of AES test flights from Vanscoy, Saskatchewan from 1989–1991 before being used in non-recovered Arctic flights. Two surviving instruments, code-named MX-31 and MX-36, were refurbished with minimal modifications and used in the biennial MANTRA flights of 1998–2004.

The main design features of the second generation emission radiometer include: a mechanical chopper at the entrance slit, a spectral band-pass filter, a mercury-cadmium-telluride detector, amplifying electronics, and an insulated liquid-nitrogen dewar surrounding the detector and optics, maintaining a temperature of approximately 77 K. The instrument design also includes a blackbody calibration flap that is automatically lowered to cover the field-of-view periodically. The flap, mounted externally to maintain a temperature above that of liquid nitrogen, has an embedded thermometer to provide temperature information necessary for performing in-flight radiometric calibration. The discrete spectral bandpass filters of the original design were replaced by a circular variable filter (CVF): a semicircular optical filter with a spectral bandpass dependent upon the angular position. With two filter segments mounted on a constantly turning wheel, the instruments are able to scan a wavelength region from 8–14 μm (715–1250 cm^{-1}) with a band-pass varying between 1% and 4% of the center wavelength.

4.3 Retrieval

Vertical profiles of HNO_3 were originally retrieved from data measured by the emission radiometers through an onion-peeling algorithm, wherein the amount of HNO_3 between two measurement altitudes was assumed to be proportional to the change in radiance between the measurements (Evans et al., 1976b). Instrument calibration parameters, necessary for converting the current output of the detector into radiances,

were determined through pre- and post-flight calibration tests.

Quine et al. (2005) introduced an updated retrieval algorithm which aimed to address many of the drawbacks of the onion-peeling method. This method centered around the use of a detailed forward model of the atmosphere and instrument: given atmospheric temperature and pressure information, expected trace gas abundances, and some instrument calibration parameters, the forward model simulates the spectral scans recorded by the instrument during flight. An optimization routine is then used to obtain a best fit between the simulated and measured spectra by adjusting the instrument parameters and trace gas amounts. The main advantage of the retrieval technique of Quine et al. (2005) is the incorporation of instrument parameters into the retrieved state vector. This makes possible the analysis of flight data which lacks pre- and post- calibration data, such is the case for the data from the 1990 flights. The main points of the retrieval algorithm are summarized below.

4.3.1 Atmospheric forward model

The atmosphere is modeled as a set of discrete 2-km-thick cells on a vertical grid. For each cell, a density-weighted effective mean temperature and pressure are determined based on sonde measurements. Spectral absorption coefficients for each cell and for eight principal emitting gas species (H_2O, CO_2, O_3, N_2O, CH_4, HNO_3, CFC-11 and CFC-12) are calculated using the GENSPECT line-by-line code (Quine and Drummond, 2002) with HITRAN 2004 spectral data (Rothman et al., 2005). HITRAN line parameter updates for HNO_3 are included (Flaud et al., 2006), based on work that went into the MIPAS (Michelson Interferometer for Passive Atmospheric Sounding) database. Radiative transfer code included in the GENSPECT package is used to calculate the radiance at each cell boundary based on the blackbody emission and transmission of each cell. This simulated spectral radiance profile is linearly interpolated to the radiometer measurement heights, and passed as the main input into the instrument forward model.

4.3.2 Instrument forward model

The aim of the instrument forward model is to accurately simulate the true mapping between input radiance and detector response. While it is assumed that the individual instruments are for the most part functionally identical, a few instrument parameters defining properties unique to each instrument are necessary, and are included in the forward model.

Two parameters define the mapping between the CVF angular position and wavenumber. A linear relationship between CVF position and wavelength is reported by the CVF manufacturer (Optical Coating Laboratory, Inc.) and assumed here. Initial estimates for the two parameters are produced manually by finding the angular position of the O_3 and HNO_3 peaks in the raw data. These instrument parameters are then included in the retrieval, and serve to shift and stretch the wavenumber axis of the measured spectra in order to match the simulated spectra.

A third instrument parameter is used in the construction of the instrument slit function. The approximate shape of the slit function is based upon laboratory Fourier transform spectrometer (FTS) measurements of the bandpass of a sample CVF. The slit function is roughly Gaussian in shape, although asymmetric, with a heavy tail to one side (see Figure 4.2). The same FTS measurements used to determine the slit function shape are also used to define the relationship between the width of the slit function and its center wavelength. The retrieved instrument parameter is used to specify the width of a boxcar function convolved with the experimentally determined slit function, accounting for the finite angular width of the focussed light passing through the CVF in each particular instrument. The resulting instrument slit function has a width of approximately 20 cm^{-1}.

While mercury-cadmium-telluride detectors are known to exhibit non-linear response to radiation (e.g., Borrello et al., 1977; Abrams et al., 1994), laboratory tests have shown that the emission radiometer instruments respond nearly linearly to blackbody radiation over much of their measurement range. Assuming such an idealized linear relationship, the instrument is modeled such that detector response S at a

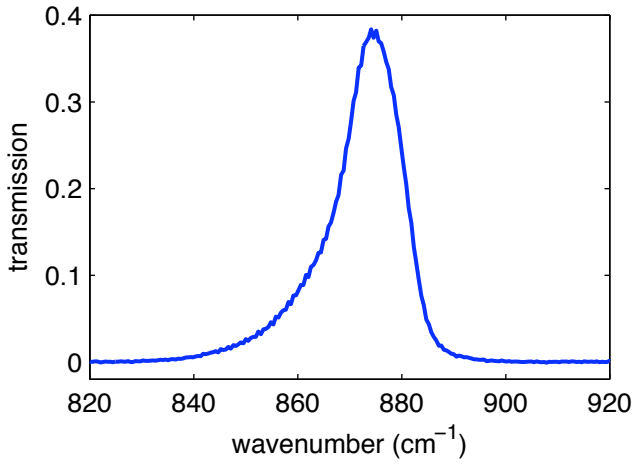

Figure 4.2: FTS-measured bandpass of the emission radiometer CVF, used to forward model the instrument's slit function.

given wavenumber ν', altitude z, and incident radiance spectrum I for a given effective viewing angle $\hat{\theta}$ is given by:

$$S(\nu', z) = R(\nu', z) \int_0^\infty I(\nu, \hat{\theta}, z) F(\nu - \nu') d\nu + \epsilon, \tag{4.1}$$

where R is the instrument responsivity, F is the slit function, and ϵ is a dark current level.

The average signal measured as an opaque section of the CVF attenuates the input radiation is used to define the dark current ϵ.

The effective viewing angle $\hat{\theta}$ is an approximate quantity describing the mean elevation angle of light collected by the instrument. It is a function of the mounting angle of the instrument on the balloon payload, the field-of-view (FOV) function of the instrument describing the angular dependence of the instrument's ability to collect radiation, and the radiation field as a function of angle. In order to define a constant $\hat{\theta}$ for any measurement set, we assume a plane parallel atmosphere with homogeneous emission. The radiation field is then given by the cosecant function describing the variation of atmospheric slant path with elevation angle, and $\hat{\theta}$ is given by the mean of the product of the FOV and cosecant functions. The FOV,

as determined by laboratory tests, is roughly toroidal, with sensitivity extending to $\pm 16^\circ$ and maximum sensitivity at $\pm 9^\circ$. The effective viewing angle for an instrument mounted at 20° elevation angle is approximately 17°.

In-flight blackbody calibration scans, performed every fifth scan at altitudes z', are used to define the instrument responsivity function R. The responsivity of the instrument changes as a function of altitude, most likely due to the fact that the detector does not typically reach liquid nitrogen temperature before launch, and continues to cool during the balloon ascent. The responsivity function is calculated as the ratio of detector output S (with dark current ϵ subtracted) to the theoretical blackbody radiance curve (based on the temperature measured by the flap thermometer, assuming the flap has a blackbody emissivity of unity) convolved with the instrument slit function:

$$R(\nu', z') = \frac{S(\nu', z') - \epsilon}{\int_0^\infty B_\nu(\nu, T) F(\nu - \nu') d\nu}. \tag{4.2}$$

The responsivity function is then interpolated to the atmospheric measurement heights (z) in order to calculate the simulated spectra via Eq. (4.1).

4.3.3 Optimization

The objective function is defined as the sum of squares of the difference between the simulated and measured spectral scans. The optimal retrieval is that which minimizes this objective function.

While the spectral range of the CVF spans 715–1250 cm^{-1}, the optimization is restricted to 825–945 cm^{-1}, a spectral window that includes strong emission features from HNO_3 as well as from the interfering species CFC-11 and CFC-12. Since there is significant overlap of the spectral features from these three species in the spectral window used due to the low resolution of the instruments, all three species are necessarily retrieved. The remaining species included in the atmospheric forward model for the full measurement range of the instrument (H_2O, CO_2, O_3, N_2O, and CH_4) are, for the present analysis, represented by constant a priori profiles. Interference by these species is small, and limited to the edges of the spectral window used.

Prior work (Quine et al., 2005) introduced the use of a non-linear optimization algorithm used to search for a global minimum of the objective function by iteratively perturbing the full state vector of instrument parameters and trace gas species on the full vertical grid. While this technique produced reasonable results, it required large amounts of computing resources, and time. In order to produce results on a faster time scale this approach has been modified.

In the current approach, the optimization routine is used to retrieve a reduced state vector composed of instrument parameters and scaling factors which multiply trace gas profiles. This reduces the state space of the optimization from the order of 30 (3 species by $\sim$10 altitudes) to six: 3 scaling factors plus the 3 instrument parameters (spectra shifting and stretching terms, plus the slit function width). We use the Direct routine (Jones et al., 1993) to perform the optimization, although in principle any global search algorithm could be employed.

The optimization routine is used in tandem with an onion-peeling routine which modifies the altitude structure of the profiles while keeping the instrument parameters fixed. Starting at the uppermost atmospheric cell and moving down, trace gas amounts are adjusted in order to minimize the measurement residual within that cell in the local spectral neighbourhood of the emission peak for each of the three species. The onion-peeling proceeds by the method of steepest descent, with the local gradient defined by first perturbing the cell VMR by 1%, and using the forward model to calculate the corresponding change in radiance. This process is repeated iteratively until the difference between the measured and simulated radiances reaches a preset convergence criterion.

Weighting functions and averaging kernels for the onion-peeling algorithm, produced by Toohey (2003), are shown in Figure 4.3 for the MX-31 instrument on the MANTRA 2002 flight. The weighting functions describe the sensitivity of the forward model to changes in the state vector (Rodgers, 2000), i.e., the derivative of the radiance for a given limb view with respect to the HNO_3 VMR as a function of height. At all measurement altitudes, the largest source of measured radiation is from the model layer directly above the measurement altitude. As the instrument

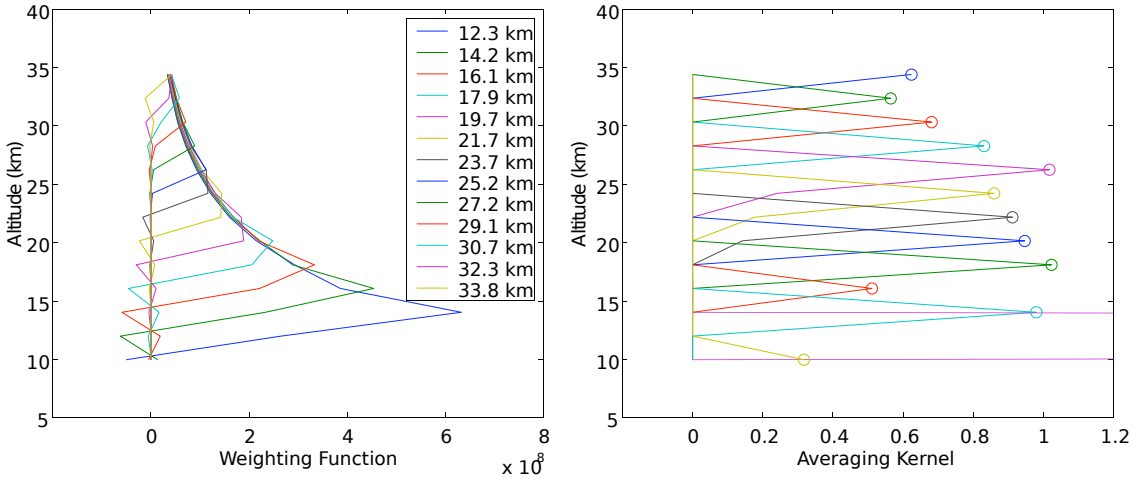

Figure 4.3: MANTRA 2002 weighting functions (left) and onion-peeling retrieval averaging kernels (right), where the nominal height of each kernel is marked by an open circle.

ascends, atmospheric layers below the instrument altitude are seen to contribute little to the collected radiance. The averaging kernels define the vertical resolution of the retrievals, and here show an almost ideal delta-function pattern, with most kernels displaying peak values of order one. The most egregious exception is at ∼12 km, where the extremely low VMR retrieved at that altitude has produced an exaggerated averaging kernel value. The width of most averaging kernels is one atmospheric model layer, implying that the retrieval method has a vertical resolution as good as or better than the resolution of the model atmosphere.

The profile-scaling and onion-peeling steps are iterated alternately as described above with one exception: the objective function to be minimized for the first optimization step is based solely on spectral scans within the lowermost atmospheric cell, since at low altitudes all three retrieved species show significant spectral features. The presence of three emission "peaks" allows the proper estimation of both the wavenumber-calibration-stretching and bandpass-filter-width instrument parameters. At higher altitudes, where only HNO_3 emission is significant, the effects of these two instrument parameters are indistinguishable. Experience has shown that this first-pass optimization is generally successful in identifying the correct "valley" of the full

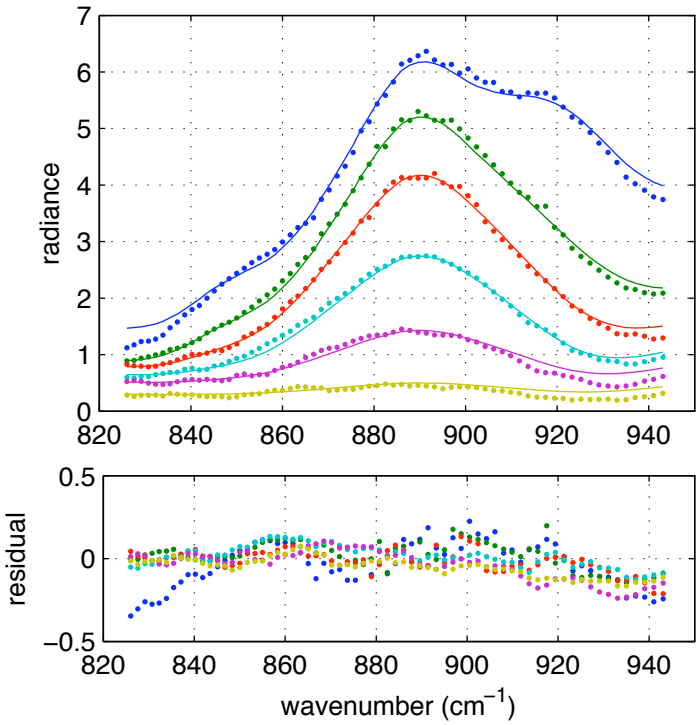

Figure 4.4: A selection of spectral fits and residuals, approximately equally spaced between 12 and 30 km height, from the emission radiometer fitting routine. Radiance and residuals are plotted in arbitrary instrument units. Results shown are for MANTRA 2000, MX-36. Plotted points represent the measurement scans, while lines show the simulated spectra.

objective function state space, which can have a great number of local minima.

Comparison of simulated and measured spectral scans for a sample instrument and year are shown in Figure 4.4.

4.3.4 Error analysis

The optimization discussed above aims to obtain a best fit between the measured and modelled spectra over the full spectral range. In order to estimate the retrieval error for each of the trace gases fit, a root-mean-square (RMS) residual is calculated between the simulated and measured spectra in the neighbourhood of the emission peak for each gas species fit (i.e., approximately 885 cm^{-1}, 850 cm^{-1} and 920 cm^{-1} for HNO_3, CFC-11 and CFC-12, respectively). This RMS value acts as a species specific

measure of the radiance noise. In order to estimate the retrieval error based on this noise, a Jacobian matrix describing the forward modeled change in radiance for a given change in species mixing ratio is created through a sensitivity study. Retrieval errors are then calculated based on the inverse of this Jacobian matrix, and we are therefore able to estimate the error in each retrieved species based on the radiance noise present in the spectral peaks of all three species.

The recorded temperature of the blackbody flaps is another major source of uncertainty. Laboratory tests have produced calibration coefficients for the two instruments used in the MANTRA flights. Based on these tests, we conservatively estimate the thermometer measurement error of the blackbody flaps for the instruments on the MANTRA flights as 2 K. Corresponding VMR errors have been calculated by simply performing the retrieval on raw data perturbed by the temperature uncertainty estimate. A 2 K error in temperature leads to a HNO_3 VMR error of approximately 10% at 24 km.

The uncertainty in the effective elevation angle is estimated to be $\pm 0.5^{\circ}$, which corresponds to an error of approximately 2.5% at 24 km.

Radiance error, blackbody temperature error, and viewing angle error are added in quadrature to obtain the total retrieval error as a function of altitude.

Finally, the retrieval altitudes themselves have errors as a result of uncertainty in the measurement of pressure by the sonde, and the propagation of this uncertainty through the hydrostatic equation used to calculate altitude. Richner and Viatte (1995) have simulated the effect of errors in measured pressure and temperature on calculated heights. Assuming an uncertainty of 1 hPa, the corresponding error in altitude is estimated from Figure 4 of Richner and Viatte (1995). Altitude error is on the order of a few meters below 20 km, but grows exponentially with height and reaches ± 2 km at 35 km.

This error analysis does not take into account errors due to the assumption of linear detector response. Blackbody tests have shown that while the response of the instrument is nearly linear over its full measurement range, non-linear response is apparent at the lowest measured radiances. Correction of non-linear effects would

Flight	Date	Instruments
AES	Aug 20, 1990	MX13
AES	Aug 30, 1990	MX19
MANTRA	Aug 24, 1998	MX31, MX36
MANTRA	Aug 29, 2000	MX31, MX36
MANTRA	Sep 03, 2002	MX31

Table 4.1: Balloon flights of the radiometer data set used in this work.

require characterization of the detectors used in each instrument. Since the retrieval method used here aims to be useful for historical flights for which the instruments are no longer available, we maintain our assumption of linearity. The effect of non-linear detector response would likely be systematic biases, especially at higher altitudes where measured radiances are at a minimum.

4.4 Data

In this work, we have collected data recorded by emission radiometer instruments during three MANTRA balloon flights (1998, 2000, and 2002), and during two flights in 1990. All data was collected through launches from a northern midlatitude site (Vanscoy, Saskatchewan, 52°N, 107°W) during late summer when dynamical variability is minimal. By analyzing the data with a consistent retrieval algorithm and spectroscopic line intensities, we connect modern measurements of HNO_3 with historical ones, and in so doing, develop a semi-quantitative picture of long-term HNO_3 profile changes.

Each MANTRA mission has included on the instrument payload two radiometers, MX-31 and MX-36. The two instruments were mounted with different elevation angles, with one at 20° and the other at 30–40°. Raw data of good quality was collected by both instruments during the flights of 1998 and 2000. One instrument failed in 2002, and both instruments failed in 2004, likely due to the effect of ice build-up on the mechanically rotating CVF.

The AES test flights of 1990 suffered from a variety of fatal and non-fatal failures,

making the raw data difficult to process. All flights suffered from poor data quality due to transponder drop-outs. Many flights also suffered from obvious mis-calibration of the blackbody flap temperatures.

Included in this work are results from two radiometer flights of August 1990. Of all data from the AES test flights, that from MX-19 on August 30 is of the highest quality. The ubiquitous transponder drop-outs and a higher-than-expected spectral noise are the data's only significant faults. Data from MX-13, flown on August 20, 1990, suffered some more critical non-fatal effects. Time synchronization between the radiometer clock and the radiosonde clock (in order to achieve proper altitude registration) was produced by defining a time offset that minimized the residual between the time derivatives of the sonde-measured air temperature and the instrument flap temperature. Furthermore, the flap temperature was bias corrected in order to bring the difference between flap temperature and air temperature into a range comparable to the other flights. Due to these necessary pre-processing steps, the temperature error of the blackbody flaps for this flight has been estimated to be 5 K.

The retrieved profiles of HNO_3, CFC-11 and CFC-12 are here compared with version 2.2 retrievals from ACE-FTS[2], launched in August 2003. We compare with ACE-FTS since it simultaneously retrieves all three species measured by the radiometers, and since ACE-FTS data has previously been used in a trend analysis of many species, including CFC-11 and CFC-12 (Rinsland et al., 2005). HNO_3 retrievals from ACE-FTS employ a set of microwindows near 900 cm^{-1} (11.3 μm), and another set near 1700 cm^{-1}, and use the HITRAN 2004 spectral database. ACE-FTS retrievals do not use the Flaud et al. (2006) HNO_3 update (but the emission radiometer retrieval does, as stated in §4.3.1), but the impact of the update to HITRAN 2004 in the 11.3-μm spectral range is small: the percent change in integrated line intensity over this band is on the order of 2% (Flaud et al., 2006).

Validation of the version 2.2 ACE-FTS HNO_3 data product has shown good agreement between ACE-FTS and other satellite data sets (Wolff et al., 2008), with mean

[2]The ACE-FTS instrument was introduced in Chapter 1. A more technical description of the ACE-FTS instrument and data set used here and throughout this thesis is included in Appendix B.

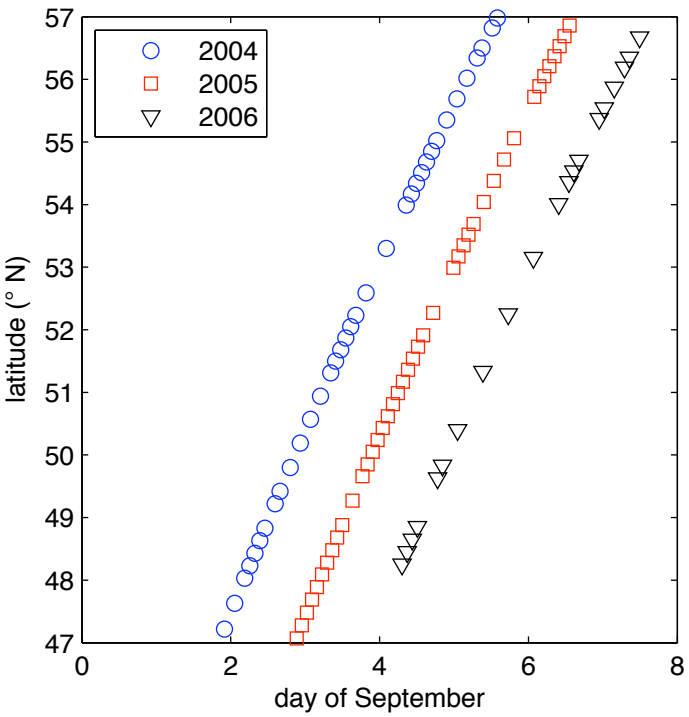

Figure 4.5: Time-latitude coordinates of Northern Hemisphere midlatitude ACE-FTS occultations in late summer, over three years.

absolute differences generally within ±1 ppbv (±20%) from 18 to 35 km. In comparisons with MIPAS and MLS in particular, mean relative differences lie within ±10% between 10 and 36 km.

Due to the orbit and solar occultation technique used by the ACE-FTS platform, the latitudinal distribution of measurements is highly dependent upon the time of year. ACE-FTS samples the northern midlatitudes in late summer briefly at the beginning of September (see Figure 4.5), repeating each year with a small offset in time between years. These late summer ACE-FTS observations are in close proximity to the window of past MANTRA flights (August 24 – September 3). For the comparison, we take averages of these late summer ACE-FTS measurements from 2004–2006 over the 10° latitude band centered on Vanscoy, Saskatchewan.

4.5 Results

The retrieved profiles of HNO_3 are shown in Figure 4.6. An overall MANTRA-era mean HNO_3 profile is calculated based on the weighted mean profiles for 1998 and 2000, and the single measured profile in 2002, and is shown on each panel for comparison. The retrieved profiles are quite consistent. The two retrieved profiles for the flights with simultaneous measurements (1998 and 2000) agree within error bars at almost all altitudes. Differences between the simultaneous measurements may be due to differences in horizontal sampling by the two instruments mounted with differing elevation angles, and the local horizontal gradient in the HNO_3 field, or due to different non-linear response of the instrument detectors as discussed in §4.3.4. The error in the 1990 profiles is large, due to uncertainties in the blackbody flap temperature and the measured radiance noise. Despite this, the profiles are not inconsistent with the MANTRA mean profile.

MANTRA yearly mean profiles of HNO_3, CFC-11 and CFC-12 are compared to mean Northern Hemisphere midlatitude late summer retrieved profiles from the ACE-FTS satellite instrument in Figure 4.7. The comparison shows good agreement for HNO_3: the MANTRA HNO_3 profiles lie within the ACE-FTS 2σ variability for most altitudes. The retrieved HNO_3 is slightly larger than the ACE-FTS profiles at the peak values around 24 km for MANTRA 1998 and 2002. Also, the MANTRA profiles are low compared to ACE-FTS at high altitudes: this may be due to the non-linear response of the detectors at low radiance levels, or due to altitude errors. The CFC profiles retrieved from the radiometer measurements as interfering species exhibit a large degree of scatter. Compared to ACE-FTS, the retrieved CFC-11 profiles are biased high, especially for the MANTRA 1998 flight. CFC-12 shows no obvious bias, but shows significant oscillations with height. The weaker emission of the CFCs leads to a more difficult retrieval: the estimated errors for the individual CFC profiles (not shown) range from approximately 20% at low altitudes to greater than 100% at high altitudes where the mixing ratios are small. Given the difficulty of the CFC retrieval, retrieval of values within a factor of two of the ACE-FTS mean profiles is considered

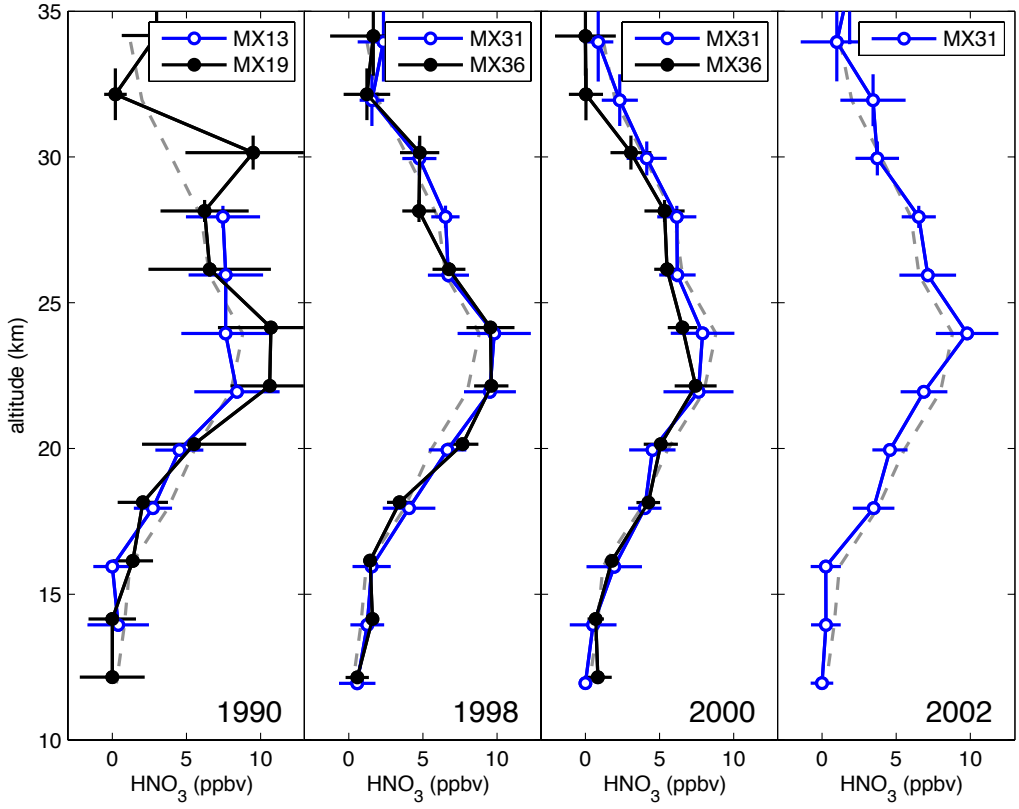

Figure 4.6: Retrieved HNO_3 profiles for the years shown. Instrument code names, specifying the source of the radiance data used in each retrieval, are given in the legend (see Table 4.1). Horizontal error bars show the estimated total uncertainty in the retrievals, vertical error bars show the estimated error in altitude. The mean profile over MANTRA missions (1998–2002) is shown by the dashed line.

acceptable.

Are the MANTRA HNO_3 measurements consistent with the variability of the CMAM? Figure 4.8 addresses this question by plotting the percent standard deviation of the three yearly mean MANTRA profiles and that from the 18-hourly model fields from the final week of August over 10 CMAM years. The two profiles show similar structure, with larger values in the lower stratosphere, and a broad minimum between 20 and 30 km. This comparison is somewhat unreliable as the MANTRA HNO_3 measurements represent a very sparse sampling of the true HNO_3 late-summer time series, with only three independent samples. We address the issue of sparse sampling

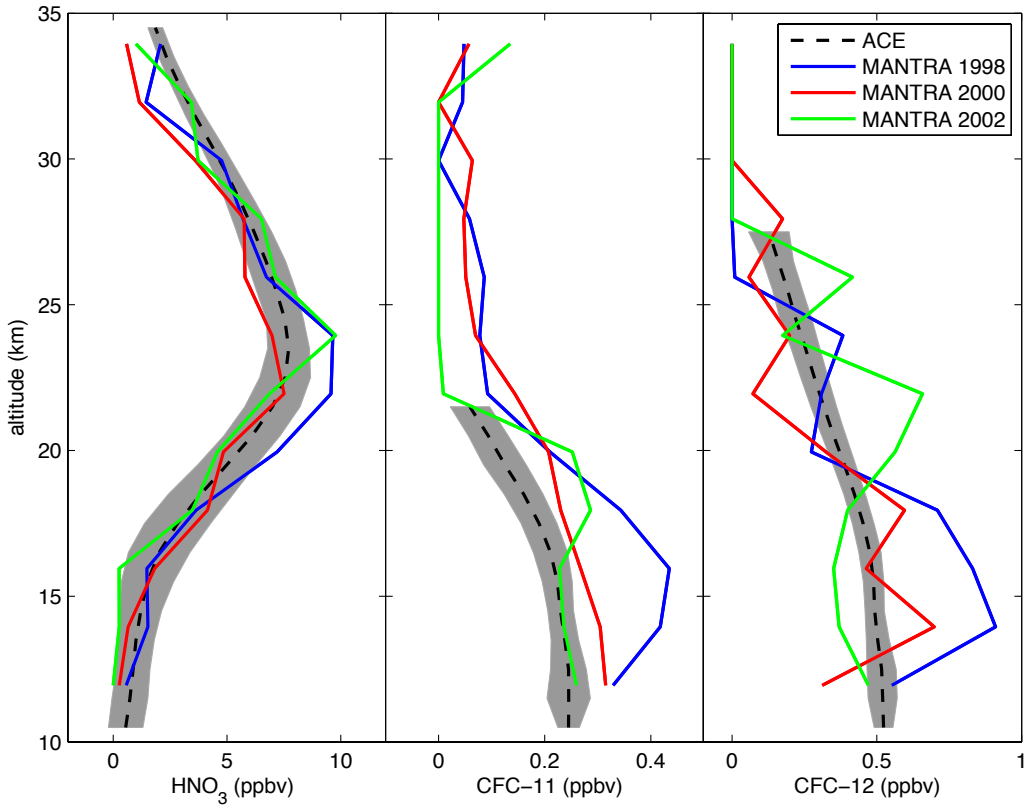

Figure 4.7: Profiles of HNO_3, CFC-11 and CFC-12 VMR from the MANTRA flights and zonal (47°– 57°) mean profiles from measurements by the ACE-FTS satellite instrument in early September, 2004–2006. Shading represents the 2σ variability of the ACE-FTS profile.

by MANTRA by calculating the standard deviation of the model profiles using only three random samples. By iterating this procedure a large number of times, we produce a probability distribution function of the calculated standard deviation for the given MANTRA sampling, and define a 99% confidence interval, shown in dark grey shading in Figure 4.8. By repeating the same procedure, but adding a 10% random error onto the sampled CMAM profiles, we get an even better sense of the variability expected in the MANTRA measurements (light grey). Between 18 and 30 km height, the variability of the MANTRA measurements is seen to be consistent with the variability of the model, given the sparsity and uncertainty of the measurements. Above 30 km and below 18 km, the variance of the MANTRA measurements is apparently adversely affected by instrument error.

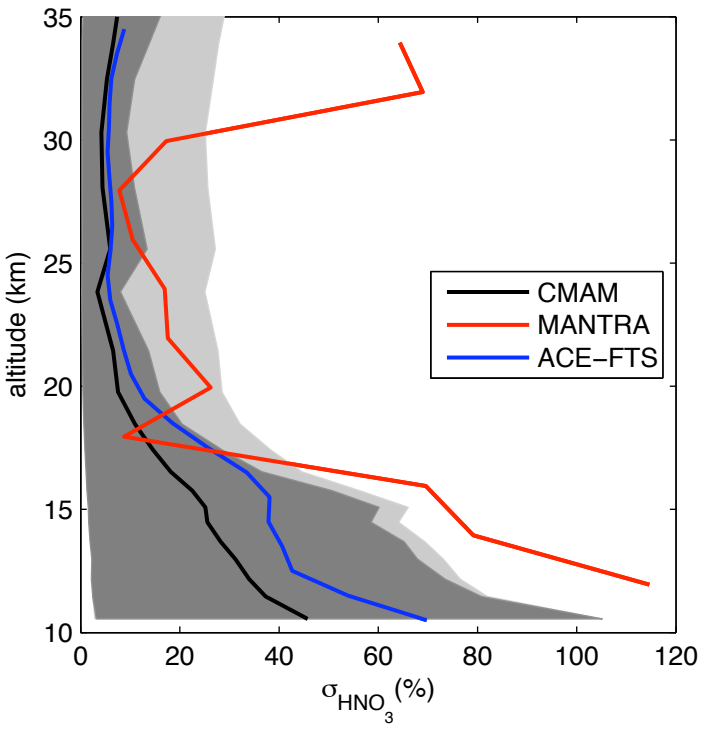

Figure 4.8: Percent standard deviations of HNO_3 VMR profiles from ACE-FTS measurements, MANTRA measurements, and simulated fields from CMAM. Grey shading represents the 99% confidence interval of values given by CMAM when only three independent samples are used (dark grey), and with an added 10% error on each sample (light grey).

The variance of ACE-FTS HNO_3 measurements, quantified by the standard deviation of the ACE-FTS measurements within the 47°–57°N latitude bin, is in excellent agreement with that of the CMAM (Figure 4.8). On the one hand, this close agreement validates our use of the model to explore the sampling issues inherent in the MANTRA measurements. On the other hand, under the assumption that the CMAM gives a good estimate of the variability of the true atmosphere, any difference between the CMAM and ACE-FTS variances should be due to the random error of the ACE-FTS retrievals. The close agreement between the measured and modeled variance is then evidence that ACE-FTS measurements display a high degree of precision, which is consistent with the small errors estimated for ACE-FTS (typically 1–5% for HNO_3 between 10 and 35 km) based on analysis of random spectroscopic fitting errors (Boone et al., 2005). The reported ACE-FTS random errors, and the issue of

comparing ACE-FTS measurement scatter to the trace gas variability simulated by the CMAM, will be revisited in Chapter 6.

4.6 Discussion

HNO_3 profiles retrieved from emission radiometer measurements from before and well after the Mt. Pinatubo eruption show no significant difference, however, the uncertainty of the retrievals is quite large compared to the possible trend. Figure 4.9 illustrates this point. Plotted are midlatitude HNO_3 mixing ratio anomalies at 24 km from the emission radiometer results (labeled MANTRA, averaged for years with simultaneous measurements) and from the UARS MLS, ACE-FTS and Aura MLS satellite instruments[3]. The satellite results are monthly means from the 10° zonal band centered on the latitude of Vanscoy, 52°N. Only months with more than 5 measurements within the latitude bin for each instrument are used. Anomalies are produced by subtracting climatological monthly means derived from four years of Aura MLS results from 2005 to 2008.

The trend in the UARS MLS timeseries is quite clear. Linear regression of the UARS MLS timeseries produces a slope estimate of -0.27 ± 0.06 ppbv per year or -3.2 ± 0.7 % per year, consistent with the 3% per year trend reported by Randel et al. (1999). Ideally, the emission radiometer HNO_3 measurements would shed light on both what HNO_3 levels were like just before the eruption of Mt. Pinatubo and the subsequent start of the UARS MLS measurement timeseries, and on the behaviour of HNO_3 levels in the years after the UARS MLS timeseries. Unfortunately, the estimated errors of the emission radiometer measurements are larger than the trend measured by UARS MLS, and as such, these measurements can neither confirm nor refute the presence of the trend observed by UARS MLS.

The closest approximation to a consistent satellite-based observational data set may be the combination of measurements by the UARS MLS discussed above with

[3]Descriptions of supplemental data sources, such as the UARS MLS and Aura MLS trace gas observations used here, are included in Appendix B.

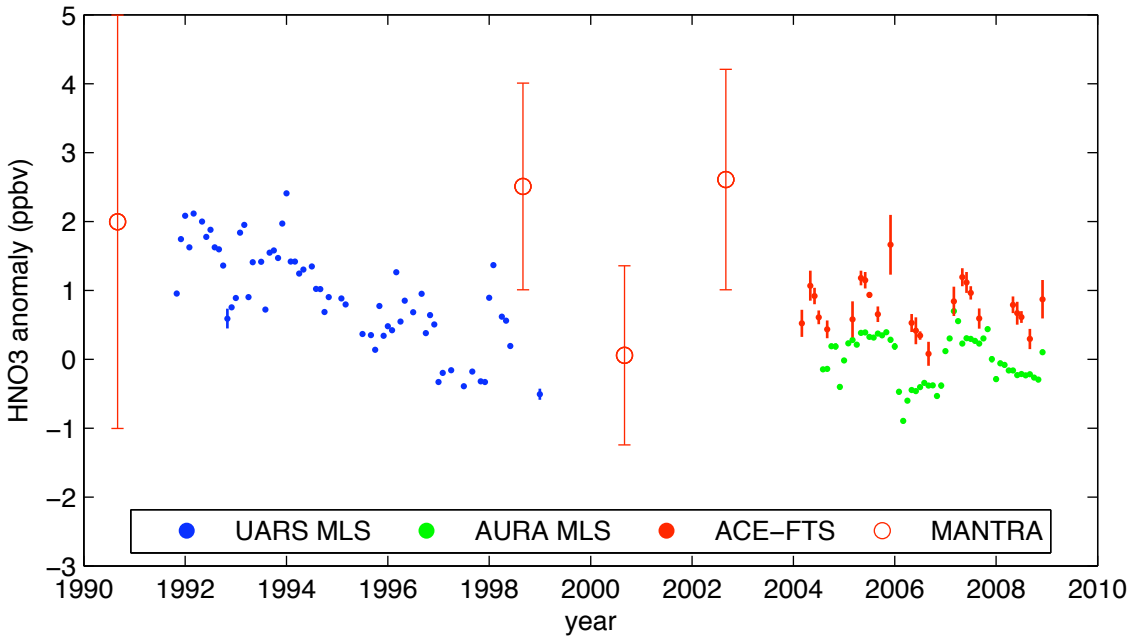

Figure 4.9: Multi-instrument timeseries of midlatitude monthly mean HNO_3 at 24 km. Values shown are anomalies from a monthly mean annual cycle calculated from Aura MLS monthly means from 2005–2008. Satellite results are shown as points, and are monthly zonal means from the latitude band 47–57°. The standard errors of the mean for each point are plotted as vertical bars. Emission radiometer results from MANTRA and the 1990 test flights are displayed as open circles with error bars.

those from the second generation MLS instrument onboard the Aura satellite platform, launched in 2004. In a validation study, Santee et al. (2007) show that version 2.2 Aura MLS results seem to be consistent with the lower bound of the UARS MLS results. Aura MLS retrievals of HNO_3, interpolated to 24 km, are shown in Figure 4.9 as anomalies from their mean seasonal cycle, and are indeed are comparable to the later years of the UARS MLS timeseries. This is consistent with the idea that the trend measured by UARS MLS was the beginning of a slow relaxation, and that HNO_3 values had approximately reached their pre-Pinatubo state by the end of the UARS MLS data set, and have been relatively constant since. Least squares fitting of an exponential decay function of the form $y = A\exp(-t/\tau)$ to the concatenated MLS timeseries (i.e., UARS and Aura) results in an estimated decay time constant (τ) of 3.6 years.

Examination of the Aura MLS HNO_3 anomaly timeseries reveals a significant amount of variability, with anomalies oscillating in time with a period of approximately two years. This is a clear indication of the influence of the QBO on these midlatitude HNO_3 abundances. Closer examination of the UARS MLS timeseries would suggest the presence of the QBO signal here as well. Elimination of this QBO signal, through a regression of the data to some proxy for QBO phase and amplitude could conceivably be used to perform a more rigorous analysis of the long-term changes in midlatitude HNO_3.

HNO_3 measurements from ACE-FTS are seen in Figure 4.9 to be positively biased compared to the Aura MLS result. This finding is consistent with comparisons between ACE-FTS and Aura MLS HNO_3 retrievals reported by Santee et al. (2007) and Wolff et al. (2008). Offsets between different instruments measuring HNO_3 have been often noted (e.g., Lary and Lait, 2006). One possible cause of HNO_3 measurement biases is uncertainty in the spectroscopic parameters used to retrieve profiles from radiance measurements. Laboratory measurements of the line intensities in the 11.3-μm band, in particular, have yielded results that differ by as much as 30% (Flaud et al., 2006). Retrieved volume mixing ratio magnitudes have been seen to be directly proportional to scaling of the line intensities (e.g., Irion et al., 2002), thus the uncertainty of spectroscopic parameters, especially for this spectral region, leads to a related uncertainty in the retrieved abundances.

In summary, accurate assessment of stratospheric nitrogen trends requires a consistent data set with considerably lower errors and/or denser spatiotemporal sampling than the emission radiometer measurements. The best available data set for assessing the long-term behaviour of HNO_3 after the Mt. Pinatubo eruption is that consisting of measurements made by the UARS and Aura MLS instruments. Since these two instruments share a common measurement technique, measuring in the same spectral region and performing retrievals based on constant spectroscopic parameterization, the assumption that there exists no systematic bias between the two instruments is well justified. Regression of an exponential decay function to the UARS+Aura MLS timeseries results in a fit with decay time constant of 3.6 years. Thus, it seems rea-

sonable to conclude that the linear trend seen in the UARS MLS HNO_3 data set is in fact the beginning of slow relaxation.

Chapter 5

Interpreting coincident measurement statistics

5.1 Introduction

Measurements are imperfect. Faults in the design, implementation or operation of instruments lead to measurements that differ from the truth. These errors can be decomposed into two parts: systematic error, which is relatively constant from one measurement to the next, and random error, which varies from one measurement to the next. Systematic error leads to data that is biased, while random error leads to data that exhibits scatter.

The process of obtaining systematic and random error estimates for a space-based remote sounding instrument and thereby assessing the quality of the measurement set produced by the instrument is known as satellite validation (e.g., von Clarmann, 2006). Validation studies often make use of situations wherein two instruments sample the atmosphere at nearly the same time and space. Differences between these coincident measurements are attributed to the systematic and random errors of each instrument. After compiling a set of coincident measurements, the central result of validation studies is usually a bias estimate based on the difference of means of the coincident sets. This result is useful for many reasons, a pragmatic one being that it is (often) a single number expressing the degree of agreement between the two

systems, e.g., "instruments X and Y agree to within 10% between 20 and 40 km". Mean values have the added advantage that random error "averages out" over a large enough sample, so that a difference of means is taken to be a good estimate of the systematic bias.

Validation studies often also include a comparison of the scatter of the two data sets, quantified through the variance or standard deviation (SD). Scatter comes from two sources, true atmospheric variability and random instrument error[1]. The material of this chapter was largely motivated by the desire to interpret the comparison of the SDs of coincident measurements, i.e., to extract from such comparisons information regarding both natural variability and random instrument error.

Interpretation of any statistical comparison is aided greatly by the explicit use of a comparison model. A simple model for coincident measurement comparisons is introduced in §5.3. This simple model will illustrate how random measurement error and atmospheric variability are combined into the variances and covariances of coincident measurements. Implications due to the presence of bias between measurements will be explored in detail in §5.4. Three techniques[2] for extracting underlying comparison model parameters out of the measurement statistics will be presented in §5.4.4 and applied to coincident data from the ACE-FTS and Aura MLS instruments in §5.5.

5.2 Terminology

"Measurements" are produced by "instruments" as an estimate of some physical quantity. In this work, measurements will refer to the final data produced by a satellite instrument, and provided to users. An instrument is thus comprised of all processes that lead to the measurement, including the physical instrument apparatus and retrieval algorithms.

[1]Scatter, spread, dispersion and variability are all terms used to describe the variance of values around their center. In this work, variability will refer to real variance about a mean atmospheric state, in the real atmosphere, or in a model. Scatter will be used in reference to the variance in measurements, a product of both atmospheric variability and random measurement error.

[2]Much of the material of this chapter was published by Toohey and Strong (2007).

Differences between any measurement (x_i) and the truth (τ_i) can be decomposed into systematic and random errors. A simple model relating measurement and truth is given by

$$x_i = B(\tau_i) + \epsilon_i \tag{5.1}$$

where systematic error is described by the bias function B and random error is described by the random, zero-mean variable ϵ_i. In practice, the bias of an instrument may be a function of any number of other variables. For simplicity, the bias function given here is a function of τ only, thus it can be seen as a calibration factor. An ideal instrument with no bias would have $B(\tau_i) = \tau_i$.

In the measurement model of Equation 5.1, the random errors are assumed to be independent of τ_i, i.e., they are absolute rather than relative errors. If one suspected that the errors would be better described as relative, the model could be modified to the form $x_i = B(\tau_i) + \epsilon_i\tau_i$. In this case, many of the equations in the following treatments (both here and in Chapter 6) would need to be modified, but a change to relative errors would not affect our ability to perform the estimation techniques which will be introduced[3].

The precision of an instrument is defined by the spread in possible values of ϵ, i.e., by the shape of its parent distribution. Under the common assumption that the errors are normally distributed, the distribution of ϵ is uniquely identified by the "random error variance", σ_ϵ^2. Given a large number of ϵ instances, the random error variance is equal to the mean of the square of those errors:

$$\sigma_\epsilon^2 = \frac{1}{n}\sum_{i=1}^{n}\epsilon_i^2. \tag{5.2}$$

In an idealized coincident measurement event, two instruments, X and Y, produce measurements at a particular time and place. For each coincidence (indexed by subscript i), the two instruments produce the measurements x_i and y_i. The following

[3]A key to much of the following statistical reasoning is that the random errors are uncorrelated with the truth, not necessarily independent. Thus a relative error model does not invalidate the logic of the following analysis, since, as long as ϵ_i is random, and uncorrelated with the truth, then $\mathrm{Cov}(\tau_i, \epsilon_i\tau_i) = 0$.

analysis treats measurements on each vertical level independently, thereby ignoring cross-correlation of errors between different heights. While this assumption is inappropriate for ground-based or nadir space-based measurements, for limb-sounding measuring systems, which have vertical resolutions on the order of the width of their vertical averaging kernels, it is a common and in many cases appropriate simplification.

5.3 Comparing unbiased coincident measurements

For two unbiased measurements with random error variances σ_δ^2 and σ_ϵ^2 respectively, coincident measurements by instruments X and Y can (as in Dunn, 1989) be modeled as:

$$x_i = \tau_i + \delta_i \tag{5.3}$$

$$y_i = \tau_i + \epsilon_i. \tag{5.4}$$

The variances and covariance of the measurements are then given by:

$$\sigma_x^2 = \sigma_\tau^2 + \sigma_\delta^2 \tag{5.5}$$

$$\sigma_y^2 = \sigma_\tau^2 + \sigma_\epsilon^2 \tag{5.6}$$

$$\sigma_{xy} = \sigma_\tau^2 \tag{5.7}$$

where the assumption that the errors are random (and thus uncorrelated with the truth or with each other) has been important in simplifying the relations. The variance of unbiased measurements is simply the sum of the natural variance σ_τ^2 (i.e., a quantification of the real atmospheric variability of the measurement set) and the random error variance σ_δ^2 or σ_ϵ^2. The covariance of unbiased coincident measurements is equal to the natural variance.

Given a sample of n coincident measurements, sample variances s_{xx} and s_{yy} and

covariance s_{xy} can be calculated from the data:

$$s_{xx} = \frac{1}{(n-1)} \sum_{i=1}^{n} (x_i - \bar{x})^2 \tag{5.8}$$

$$s_{yy} = \frac{1}{(n-1)} \sum_{i=1}^{n} (y_i - \bar{y})^2 \tag{5.9}$$

$$s_{xy} = \frac{1}{(n-1)} \sum_{i=1}^{n} (x_i - \bar{x})(y_i - \bar{y}), \tag{5.10}$$

where overbars denote arithmetic mean values, i.e.,

$$\bar{x} = \frac{1}{n} \sum_{i=1}^{n} x_i. \tag{5.11}$$

The expression for sample variance written here, with $(n-1)$ in the denominator, is an unbiased estimator of the population variance, compared to a similar expression with n in the denominator, which can be significantly positively biased for small sample sizes (Taylor, 1997).

The sample statistics s_{xx}, s_{yy}, and s_{xy} represent estimates (based on the collected data) of the expected statistics σ_x^2, σ_y^2, and σ_{xy} (based on the simple measurement model). Sample statistics can be substituted for the expected statistics and used to derive estimates of the error variances and natural variability:

$$\hat{\sigma}_\delta^2 = s_{xx} - s_{xy} \tag{5.12}$$

$$\hat{\sigma}_\epsilon^2 = s_{yy} - s_{xy} \tag{5.13}$$

$$\hat{\sigma}_\tau^2 = s_{xy}. \tag{5.14}$$

Estimated quantities derived from the data through the sample statistics will be denoted by a caret ($\hat{}$) above the variable.

So, for unbiased coincident measurements, three quantities of possible interest (the error variances of each instrument and the natural variability) can be estimated based on the three statistical measures of the data (the variances of each measurement set and the covariance). This simple example shows the potential of analysis of

the measurement variances, as a way of validating the random measurement errors reported for the two instruments, and as a way of producing estimates of the true variability of the atmosphere.

5.4 Comparing biased coincident measurements

The simple relationships between measurement error variances, natural variability and measurement (co)variances discussed above are significantly complicated by the presence of certain forms of bias.

Additive bias, representing a constant offset between measurement and truth, is modeled by adding constant terms to Equations 5.3 and 5.4. Since the calculation of variances and covariance is insensitive to constant terms, the analysis of the previous section would remain valid in the presence of additive bias between instruments.

Multiplicative bias represents a bias that is a linear function of the true state τ. In the presence of both additive and multiplicative bias, measurements x_i and y_i can be modeled as:

$$x_i = a_x + m_x\tau_i + \delta_i \tag{5.15}$$

$$y_i = a_y + m_y\tau_i + \epsilon_i. \tag{5.16}$$

where a and m represent the additive and multiplicative biases for the subscripted measurement sets x and y.

5.4.1 Comparing measurement means

A common method of comparison in satellite validation studies proceeds by calculating mean values, with the aim of reducing noise due to random errors. The difference of means is used to define inter-instrumental bias. In terms of the model including multiplicative bias (Equations 5.15 and 5.16), the difference of measurement means

is given by:

$$\bar{y} - \bar{x} \;=\; a_y - a_x + \bar{\tau}(m_y - m_x). \tag{5.17}$$

When there is no relative multiplicative bias, $m_y = m_x$, and the difference of means is equal to the difference of additive biases $(a_y - a_x)$, which defines the relative additive bias $(\alpha = a_y - a_x)$ between the two instruments.

On the other hand, if there is significant relative multiplicative bias, the difference of means will be affected by this relative multiplicative bias. Calculating the difference of means between different latitude bands, or different altitudes where $\bar{\tau}$ varies, will lead to different bias estimates. The inter-instrumental bias can in this case be difficult to describe accurately and succinctly, as it will depend on space and time. Furthermore, in a worst-case scenario, additive and multiplicative bias may cancel in a difference of means, and this validation technique may fail to detect bias when it does in fact exist.

5.4.2 Comparing measurement variances

From Equations 5.15 and 5.16, the following equations are obtained for the expected population variances for x and y, and the covariance of x and y:

$$\sigma_x^2 \;=\; m_x^2\sigma_\tau^2 + \sigma_\delta^2 \tag{5.18}$$

$$\sigma_y^2 \;=\; m_y^2\sigma_\tau^2 + \sigma_\epsilon^2 \tag{5.19}$$

$$\sigma_{xy} \;=\; m_x m_y \sigma_\tau^2. \tag{5.20}$$

As in the case of unbiased measurements (§5.3), some assumptions are required to derive the equations above. The measurement errors δ_i and ϵ_i are assumed to be independent, with zero mean. The two instruments' measurement errors are also assumed to be uncorrelated with each other (i.e., $\mathrm{Cov}(\delta, \epsilon) = 0$).

Eliminating σ_τ^2 in Equations 5.18 and 5.19 using Equation 5.20 leads to:

$$\sigma_x^2 = (m_x/m_y)\sigma_{xy} + \sigma_\delta^2 \quad (5.21)$$

$$\sigma_y^2 = (m_y/m_x)\sigma_{xy} + \sigma_\epsilon^2. \quad (5.22)$$

We can then define the relative multiplicative bias $\beta = m_y/m_x$. When the truth remains unknown, the absolute multiplicative bias (m) for any instrument cannot be found: comparisons of measurements can really only hope to find the relative multiplicative bias β. In the case that one instrument (say X) is taken to be the standard, then one can define $m_x = 1$, in which case $\beta = m_y$, and the comparison model equations can be rewritten in terms of β (as in Dunn, 1989; Hocking et al., 2001) rather than m_x and m_y:

$$x_i = \tau_i + \delta_i \quad (5.23)$$

$$y_i = \alpha + \beta\tau_i + \epsilon_i. \quad (5.24)$$

In terms of the relative multiplicative bias β, the expected measurement variances and covariance are then:

$$\sigma_x^2 = \sigma_\tau^2 + \sigma_\delta^2 \quad (5.25)$$

$$\sigma_y^2 = \beta^2\sigma_\tau^2 + \sigma_\epsilon^2 \quad (5.26)$$

$$\sigma_{xy} = \beta\sigma_\tau^2. \quad (5.27)$$

Equating the sample statistics with their expectations leads to two equations relating the error variances with the sample statistics:

$$\hat{\sigma}_\delta^2 = s_{xx} - \frac{1}{\hat{\beta}}s_{xy} \quad (5.28)$$

$$\hat{\sigma}_\epsilon^2 = s_{yy} - \hat{\beta}s_{xy}. \quad (5.29)$$

The only difference between the equations for $\hat{\sigma}_\delta^2$ and $\hat{\sigma}_\epsilon^2$ and those derived in the case of unbiased measurements (Equations 5.12 and 5.13) is the presence of the estimated relative multiplicative bias $\hat{\beta}$.

5.4.3 Parallels with linear regression

Equations 5.28 and 5.29 represent two equations for three unknowns. Hocking et al. (2001) presented equivalent expressions for the under-determined relationship between bias and instrument error variances in the analysis of radar-measured stratospheric winds. Hocking et al. (2001) went on to show that:

- assuming an error variance of zero for instrument X allows the estimation of $\hat{\beta}$ via Equation 5.28, and this $\hat{\beta}$ corresponds to the slope estimate derived from an ordinary least squares (OLS) linear regression of Y vs. X;
- similarly, assuming an error variance of zero for instrument Y leads to an estimate of β that corresponds to that produced by an OLS regression of X vs. Y;
- for comparisons of two instruments when both are subject to error, the true slope (β) should lie between the bounds specified by the two OLS fits.

Figure 5.1 shows a sample scatter plot of coincident measurements of O_3 by the ACE-FTS (X) and Aura MLS (Y) instruments. Coincidences are defined here as those occurring within $\pm 5^\circ$ longitude (λ), $\pm 1^\circ$ latitude (ϕ) and within ± 6 hours[4]. OLS fits of Y against X and X against Y are shown by red and blue lines, respectively, and correspond to β values of 0.83 and 0.98. In some cases, calculation of the OLS fits may be beneficial in itself to detect bias: in this example the OLS bounds exclude the value of $\beta = 1$, implying that bias exists.

[4]The selection of coincident criteria is based on striking a balance between reducing the variability due to differences in time and space between each pair of measurements, and collecting a large enough sample to produce statistics at high levels of significance. Since the measurement ground track of a satellite measurement series is approximately meridional, the latitude criterion used can usually be rather tight without significantly affecting the sample size. In addition, meridional gradients are typically stronger than zonal gradients, motivating the use of a tighter criterion for latitude

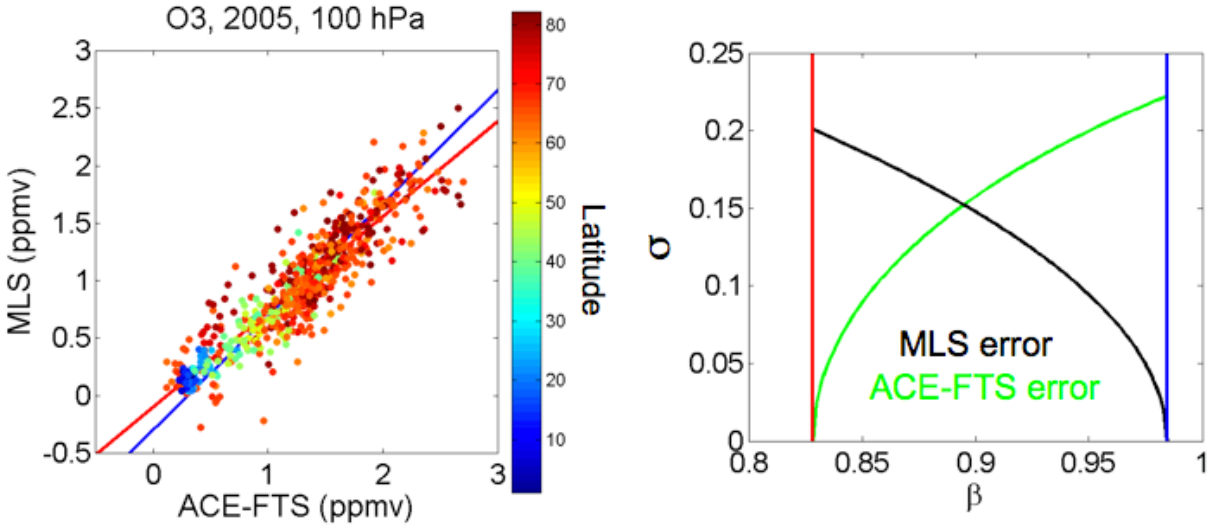

Figure 5.1: Least-squares fits of Aura MLS and ACE-FTS retrieved mixing ratios at 100 hPa over the year 2005. Left: Scatter plot of the coincident measurements, with points colour-coded according to absolute latitude. Least-squares fits of Aura MLS vs. ACE-FTS and ACE-FTS vs. Aura MLS are shown by red and blue lines respectively. Right: Red and blue vertical lines show the fit slope (β) values corresponding to the least-squares fits. Black and green lines show the theoretical Aura MLS and ACE-FTS random error SD corresponding to each point between the two least-squares fits.

With the knowledge that the correct answer lies between the two OLS fits, can the true bias β be determined? Isobe et al. (1990) present a useful overview of the multidisciplinary development and use of linear regression techniques, with a focus on astronomical uses. They identify five different techniques for producing a slope estimate between the two OLS limits, in the case that the source of the scatter in the data is not well described (i.e., the error variances are unknown). Each of these techniques produces unique results, and has various drawbacks. As is clear from Equations 5.28 and 5.29, exact estimation of β requires some knowledge of the measurement errors.

---

than longitude. The longitude and time criteria are chosen here rather arbitrarily, but are roughly comparable to those used in other validation studies comparing ACE-FTS and MLS retrievals of O_3, HNO_3 and N_2O (e.g., Froidevaux et al., 2008; Santee et al., 2007; Strong et al., 2008; Walker et al., 2005).

Hocking et al. (2001) illustrated the relationship described by Equations 5.28 and 5.29 by plotting the error SDs σ_δ and σ_ϵ given by iteratively assuming β values spanning the range between the two OLS bounds. Figure 5.1 shows a sample "Hocking plot" for the comparison of ACE-FTS and Aura MLS O_3 measurements at 100 hPa. The red and blue vertical lines highlight the values of β calculated via the two OLS methods, and correspond to the fit slopes in the adjacent plot. In the special case where the two instruments compared are believed to have equal error variances, the point on the Hocking plot where the two error variance estimates cross ($\beta \approx 0.89$ in this example) may be used to estimate β. Linear regression methods take error into account in a number of ways. When scatter due to error is the same for each point for each instrument (i.e. the data is homoscedastic) the ratio of the error measurement variances ($\sigma_\delta/\sigma_\epsilon$) can be used to produce an answer (e.g., Feigelson and Babu, 1992). The Hocking plot of Figure 5.1 is useful for understanding the power of the ratio of measurement error variances: $\sigma_\delta/\sigma_\epsilon$ spans -∞ to ∞ over the range of β values between the two OLS fits. Therefore, specifying $\sigma_\delta/\sigma_\epsilon$ uniquely specifies β. When the data is heteroscedastic (with errors varying from point to point) iterative techniques are required (e.g., Feigelson and Babu, 1992; York, 1966).

5.4.4 Variable estimation methods

As outlined above, estimation of the comparison variables in Equations 5.28 and 5.29 requires knowledge or assumptions regarding the measurement errors. Following the detailed treatment by Dunn (1989), three methods are described here for proceeding from Equations 5.28 and 5.29 towards values for $\hat{\beta}$, $\hat{\sigma}_\delta^2$, and $\hat{\sigma}_\epsilon^2$.

Method 1: Grubbs estimators

If there is no multiplicative bias between measurements ($\beta = 1$), then the instrument error estimates given by Equations 5.28 and 5.29 simplify to Equations 5.12 and 5.13,

repeated here with a subscript 1 denoting the method used:

$$\hat{\sigma}^2_{\delta 1} = s_{xx} - s_{xy} \tag{5.30}$$

$$\hat{\sigma}^2_{\epsilon 1} = s_{yy} - s_{xy}. \tag{5.31}$$

These relations were derived by Grubbs (1948, 1973) and are therefore frequently referred to as Grubbs estimators in the statistics literature (Dunn, 1989). Fioletov et al. (2006) implicitly assume $\beta = 1$ in deriving estimates of instrument error in comparisons of ozone profile measurements from satellite, ozonesonde and ground-based observations, although they used an equivalent formulation using the variance of measurement differences rather than measurement covariances.

Method 2: Use of reported error variances

If a reliable prediction of the error variance of one instrument is available, the multiplicative bias and the other instrument error variance can then be estimated. Let predictions of statistical parameters based on reported quantities be denoted by a tilde (˜) above the variable, in contrast to the estimates marked by a caret. Then, given a prediction of the error variance for instrument X, $\tilde{\sigma}^2_\delta$, Equations 5.28 and 5.29 can be solved for the estimates $\hat{\beta}$ and $\hat{\sigma}^2_\epsilon$:

$$\hat{\beta}_2|_{\tilde{\sigma}^2_\delta} = \frac{s_{xy}}{s_{xx} - \tilde{\sigma}^2_\delta} \tag{5.32}$$

$$\hat{\sigma}^2_{\epsilon 2} = s_{yy} - \frac{s^2_{xy}}{s_{xx} - \tilde{\sigma}^2_\delta}. \tag{5.33}$$

The relationship between the β estimation techniques introduced here and the least-squares fitting discussed in §5.4.3 is made clear by Equation 5.32. In the unrealistic (but often assumed) case of zero error in measurement x, $\tilde{\sigma}^2_\delta = 0$, and so $\hat{\beta}_2 = s_{xy}/s_{xx}$, which is equivalent to an ordinary least-squares fit of y to x. Equation 5.32 shows that ordinary least squares fitting of y to x underestimates the magnitude of the true slope when x is subject to error.

Retrieved quantities from satellite observations typically report an error derived

from the spectral fitting residuals, and the propagation of this fitting error through the retrieval algorithm (or some similar procedure). Each reported error for instrument X, d_i, is assumed to be a prediction of the absolute value of δ_i, the difference between the measurement and the truth (i.e. $d_i = |\delta_i|, \delta_i = x_i - \tau_i$). Recalling that the expected value of δ_i is zero, the reported instrument X error population variance can be written in terms of the reported error d_i:

$$\tilde{\sigma}_\delta^2 = \frac{1}{n}\sum_{i=1}^{n}(\delta_i - \bar{\delta})^2 = \frac{1}{n}\sum_{i=1}^{n} d_i^2. \tag{5.34}$$

The reported error variance is then simply the mean of the square of the individual measurement error predictions.

If the reported error variance for instrument Y is used instead, ($\tilde{\sigma}_\epsilon^2$, calculated as the mean of the square of the reported instrument Y error terms e_i) then estimates of the multiplicative bias and instrument X error variance are given by:

$$\hat{\beta}_2|_{\tilde{\sigma}_\epsilon^2} = \frac{s_{yy} - \tilde{\sigma}_\epsilon^2}{s_{xy}} \tag{5.35}$$

$$\hat{\sigma}_{\delta 2}^2 = s_{xx} - \frac{s_{xy}^2}{s_{yy} - \tilde{\sigma}_\epsilon^2}. \tag{5.36}$$

Equations 5.32 and 5.36 represent two different estimates of the multiplicative bias based on Method 2. Whether these two estimates agree with each other depends on the consistency between the reported errors and the measurement scatter. This issue will be explored in 5.5.2, through the application of Method 2 to actual data.

Other options for incorporating a priori error information into the analysis, including assuming $\tilde{\sigma}_\epsilon^2 = \tilde{\sigma}_\delta^2$, or taking the ratio $\tilde{\sigma}_\delta^2/\tilde{\sigma}_\epsilon^2$ as a known quantity, will not be explored here.

Method 3: Instrumental Variable Method

If some third variable, z, is measured in addition to x and y, and is correlated with τ, while uncorrelated with the measurement errors for x and y, then the measurement model parameters may be estimated directly from the data. This third measurement

is known as an instrumental variable[5] in the statistics literature (Dunn, 1989).

A second pseudo-coincident observation by instrument Y will be used in the following discussion as an instrumental variable. In this special case, the third coincident observation is modeled as:

$$z_i = \alpha + \beta(\tau_i + \eta_i) + \gamma_i \tag{5.37}$$

where the model parameters α and β are equal for the two observations by instrument Y. It is assumed that each measurement error γ_i is uncorrelated with δ_i or ϵ_i, and that $\sigma_\gamma^2 = \sigma_\epsilon^2$ (since both are the result of the random error of the same instrument). The term η represents a perturbation to the true state τ due to the fact that the second Y observation is not perfectly coincident in time and space with the X and first Y observation, i.e., it describes the representativeness error. The non-coincidence parameter η is assumed to be uncorrelated with the measurement errors. The variance of η over a set of measurements is given by σ_η^2.

The third coincident measurement leads to three additional Equations for the expected population variances and covariances:

$$\sigma_z^2 = \beta^2\sigma_\tau^2 + \beta^2\sigma_\eta^2 + \sigma_\epsilon^2 \tag{5.38}$$

$$\sigma_{xz} = \beta\sigma_\tau^2 \tag{5.39}$$

$$\sigma_{yz} = \beta^2\sigma_\tau^2. \tag{5.40}$$

Estimates of the right-hand-sides of Equations 5.38–5.40 (in terms of the estimates $\hat{\beta}, \hat{\sigma}_\tau^2, \hat{\sigma}_\eta^2$, and $\hat{\sigma}_\epsilon^2$) are obtained by replacing the expected variances σ_z^2, σ_{xz}, and σ_{yz} with the sample variances s_{zz}, s_{xz}, and s_{yz}. The ratio s_{yz}/s_{xz} is known as the instrumental variable estimate of β (Dunn, 1989), i.e.:

$$\hat{\beta}_3 = \frac{s_{yz}}{s_{xz}}. \tag{5.41}$$

[5] "Instrumental variable", as used here, is not to be confused with "instrument parameters" used in forward modeling, as in the modeling of emission radiometer radiances in Chapter 4.

The instrumental variable estimate of β is valid in general for any suitable choice of instrumental variable, not only for the special case of a secondary measurement by one instrument explored here.

With the multiplicative bias estimated directly from the data, it becomes possible to write Equations 5.28 and 5.29 strictly in terms of the measurement statistics:

$$\hat{\sigma}^2_{\delta 3} = s_{xx} - \frac{s_{xy}s_{xz}}{s_{yz}} \tag{5.42}$$

$$\hat{\sigma}^2_{\epsilon 3} = s_{yy} - \frac{s_{xy}s_{yz}}{s_{xz}}. \tag{5.43}$$

The total error variance for the secondary Y measurement is given by the second and third terms on the right-hand side of Equation 5.38: the sum of the non-coincidence error ($\beta^2\sigma^2_\eta$) and the Y measurement error variance (σ^2_ϵ). This quantity, σ^2_ν, can be estimated from the data:

$$\hat{\sigma}^2_{\nu 3} = s_{zz} - \frac{s_{xz}s_{yz}}{s_{xy}}. \tag{5.44}$$

Equation 5.44 is valid for any choice of instrumental variable. In the special case of an instrumental variable from a secondary measurement from instrument Y, $s_{xy} = s_{xz} = \beta\sigma^2_\tau$ (see Equations 5.27 and 5.39) and Equation 5.44 simplifies to $\hat{\sigma}^2_{\nu 3} = s_{zz} - s_{yz}$.

When the non-coincidence error variance σ^2_η is small, $\hat{\sigma}^2_{\nu 3} \approx \hat{\sigma}^2_{\epsilon 3}$. It should be noted that a non-coincidence parameter comparable to η could have been used in the original expression for y (Equation 5.16), explicitly describing the imperfect coincidence of the X and the primary Y observations. Under this model formulation, the non-coincidence error variance becomes confounded with the Y error variance σ^2_ϵ. As such, the estimated measurement error, $\hat{\sigma}^2_\epsilon$, calculated through all three methods, should be understood to be the sum of the true measurement error variance and an implicit non-coincidence error variance.

Estimating confidence intervals for estimates through bootstrapping

Without invoking any assumptions concerning the distribution of the measurement errors (normal or otherwise), the uncertainty in the estimated variables can itself be estimated through the use of bootstrapping (Efron and Tibshirani, 1994). This method allows for the determination of the standard error in each estimate through the sensitivity of the estimate to repeated random resampling of the data. In the following, bootstrapping is used to estimate the 95% confidence intervals of $\hat{\beta}_3$, $\hat{\sigma}^2_{\delta 3}$, and $\hat{\sigma}^2_{\epsilon 3}$ based on 1000 resamplings.

5.5 Comparing measurements by ACE-FTS and Aura MLS

The three statistical methods for measurement comparison introduced above will here be used to compare results from the satellite instruments Aura MLS and ACE-FTS over the full year of 2005. We focus on the relatively long-lived stratospheric species O_3, N_2O and HNO_3, in order to try to minimize the effect of non-coincidence error.

5.5.1 Data

Trace gas measurements made by the ACE-FTS, first introduced in Chapter 1, will be used in the following analysis. As in Chapter 4, results used here are from the version 2.2 data set. O_3 retrievals come from the version 2.2 ozone update.

Aura MLS (hereafter, for this chapter referred as MLS) version 1.51 retrieval results are used here, and are based on an optimal estimation method which includes use of a priori constraints (Livesey et al., 2006). MLS makes limb scans in the forward direction along the satellite orbit track, hence, consecutive scans cover significantly overlapping regions of the atmosphere. The MLS retrieval technique takes advantage of this fact by dividing the collected radiance data into "chunks" of about ten vertical scans, and simultaneously retrieving a similar number of profiles of atmospheric temperature and composition from each chunk. Therefore, the retrievals within each

chunk are not independent.

ACE-FTS and MLS observe the atmosphere with similar limb-viewing geometries. Consequently, the retrieved profiles from both instruments have similar resolutions in the horizontal (∼500 km) and vertical (3–6 km, depending on species and altitude and ACE-FTS occultation viewing geometry).

ACE-FTS spectra have a signal-to-noise ratio of greater than 300 over most of the spectral range. Uncertainties provided for the ACE-FTS mixing ratio results are 1σ statistical errors from its global-fitting retrieval algorithm, and do not include systematic contributions (Boone et al., 2005). MLS retrieval precisions are calculated as a function of the measurement error and the a priori error covariance matrix (Livesey et al., 2006).

At the time of the publication of the material of this chapter (Toohey and Strong, 2007), validation of ACE-FTS measurements had just recently begun (e.g., Walker et al., 2005; Froidevaux et al., 2006). Since that time, many works validating the ACE-FTS and MLS measurements have been published. Validation of ACE-FTS measurements of O_3 (Dupuy et al., 2009), N_2O (Strong et al., 2008) and HNO_3 (Wolff et al., 2008) have been published. These studies include statistical comparisons of ACE-FTS coincidences with other satellite instruments including MLS, and generally compare mean profiles for various latitude bins. No validation of the ACE-FTS random errors has yet been completed. Validation of MLS measurements (for O_3, N_2O and HNO_3, in Froidevaux et al., 2008; Lambert et al., 2007; Santee et al., 2007, respectively) has focussed on version 2.2 data, i.e., on data produced by a different retrieval algorithm than that used for the data (v.1.51) shown here. The main difference between data versions for the purposes of the present analysis is a large difference in HNO_3 retrievals, which will be discussed below. Each MLS validation study includes empirical estimates of the measurement precision, produced in an effort to validate the reported precisions. Empirical precision estimates are produced for O_3 and N_2O by taking the SD of differences between MLS measurements closely spaced in space and time (i.e. for "self-coincidences", as used in Method 3 here). For HNO_3, precision estimates are produced by Santee et al. (2007) by taking the SD of

MLS measurements in a small latitude band centered on the equator, where natural variability is believed to be minimal.

MLS data used in this work is screened based on the precision, status and quality fields of the MLS data files as described by Livesey et al. (2005). In addition, the MLS N_2O data is filtered according to flags provided by the MLS team, in order to remove a systematic high bias in low altitude polar vortex N_2O retrievals (as also discussed in Livesey et al. (2005)). In addition, in order to remove some suspicious ACE-FTS profiles, an ad hoc filter has been implemented, excluding from consideration any ACE-FTS O_3 profiles for which the error at any height exceeds 160 ppbv, and any N_2O profiles with retrieved VMR values outside the range -10 to 800 ppbv[6].

The implementation of the instrumental variable estimation methods described above requires a set of coincident measurements by ACE-FTS and MLS. In order to meet the data requirements of all three comparison methods, coincidences are found based on a two-stage coincidence criterion. All coincidences are defined herein as those occurring within $\pm 5^\circ$ longitude (λ) and $\pm 1^\circ$ latitude (ϕ). Primary coincidences are defined as those occurring within ± 6 hours. If multiple potential primary coincidences are found, the single coincidence for which the parameter $D = \Delta\phi(^\circ) + \Delta t(\text{hours})$ is minimized is chosen. Secondary coincidences (defining set z) are found within ± 12 hours of the primary coincidence. In order to ensure independence between the MLS measurement errors ϵ and γ, we require that the secondary coincidence be from a different retrieval "chunk" from the primary coincidence. Again, if multiple secondary coincidences are found, the parameter D defined above is minimized to find the best secondary coincidence. If no secondary coincidence is found, the primary coincidence is thrown out. Table 5.1 gives some statistics for the number of coincidences found and the space-time proximity of the coincidences.

ACE-FTS and MLS measurements must then be mapped onto a common vertical grid. An appropriate comparison should take into account the differing characteristics

[6]Negative values for retrievals should not necessarily be excluded outright, since absolute random error acting on small abundances can produce negative retrieval values. Exclusion of all negative values would lead to positive biases in this case. Here, a value of -10 ppbv is chosen as being larger than that due to typical random error, and is used to exclude a small number of extreme outliers.

species	O_3	N_2O	HNO_3
n	664	590	701
$\overline{\Delta\phi_1}(^\circ)$	0.04±0.47	0.03±0.47	0.03±0.47
$\overline{\Delta\phi_2}(^\circ)$	0.01±0.46	-0.01±0.47	-0.01±0.46
$\overline{\Delta\lambda_1}(^\circ)$	0.27±2.87	0.14±2.9	0.18±2.89
$\overline{\Delta\lambda_2}(^\circ)$	0.16±2.92	0.31±2.93	0.1±2.91
$\overline{\Delta t_1}$ (hr)	3.37±5.66	3.54±5.84	3.75±6.15
$\overline{\Delta t_2}$ (hr)	3.45±10.28	3.76±9.9	3.66±10.08

Table 5.1: Number of coincidences (n) for each species, and mean ($\pm 1\sigma$) differences between ACE-FTS and MLS primary (subscript 1) and secondary (subscript 2) measurement latitude (ϕ), longitude (λ) and local solar time (t).

of the observing systems. Comparison methods incorporating differences described by the instrument averaging kernels and error covariances have been described (e.g., Rodgers and Connor, 2003). Many comparisons (e.g., Froidevaux et al., 2006) proceed by the simpler route of interpolating to a common vertical grid. When the vertical resolutions of the two instruments are similar (as is the case for ACE-FTS and MLS), it is generally assumed that this simpler procedure will not adversely affect the statistics of the comparison. Also, this technique is applicable when averaging kernels and error covariances are unavailable for the data (as is the case here for ACE-FTS). In this comparison, retrieved ACE-FTS mixing ratio and uncertainty profiles, originally reported on a geometric height vertical grid of 1 km resolution, are linearly interpolated to the MLS pressure grid of six surfaces per decade change in pressure, using the ACE-FTS retrieved pressure profile for each measurement. This process adds an unquantifiable error to the resultant ACE-FTS mixing ratio profiles due to an unquantified error in the ACE-FTS retrieved pressure.

5.5.2 Results

O_3

Coincident measurements of O_3 by ACE-FTS and MLS over the full globe are compared in Figure 5.2. At each pressure surface, anomalies are calculated for both

measurement sets by subtracting the mean value over the full ensemble. The MLS O_3 anomalies are plotted in Figure 5.2 versus the coincident ACE-FTS anomalies for selected pressure surfaces. The points of the scatter plots are colour-coded based on the absolute value of latitude, with polar values in red, and equatorial values in blue.

In the absence of multiplicative bias (i.e., under the assumptions of Method 1), the measurements should lie along a 1:1 line with a slope of one (shown in black on the scatter plots), with scatter about the line due to the random error in each measurement. Any deviation of the slope from the 1:1 line is evidence of a multiplicative bias in the data. Lines with slopes corresponding to the multiplicative bias estimates $\hat{\beta}_2|_{\tilde{\sigma}^2_\delta}$, $\hat{\beta}_2|_{\tilde{\sigma}^2_\epsilon}$ and $\hat{\beta}_3$ (calculated by Equations 5.32, 5.35 and 5.41) are plotted on each scatter plot. The multiplicative bias estimate $\hat{\beta}_2|_{\tilde{\sigma}^2_\delta}$ is visibly less than one for all but one (10 hPa) pressure surface shown here, while $\hat{\beta}_2|_{\tilde{\sigma}^2_\epsilon}$ is quite close to one for all plots. The estimated multiplicative bias $\hat{\beta}_3$ is less than one at heights below 10 hPa, and quite close to one otherwise.

Multiplicative bias estimates $\hat{\beta}_2|_{\tilde{\sigma}^2_\delta}$, $\hat{\beta}_2|_{\tilde{\sigma}^2_\epsilon}$ and $\hat{\beta}_3$ are plotted in panel A of Figure 5.3 as a function of the MLS retrieval pressure surfaces. $\hat{\beta}_3$ is generally consistent with a slope of one, except above 1 hPa, and below 100 hPa, and at 14.68 and 23.54 hPa. Estimates calculated via Method 2 bracket $\hat{\beta}_3$, and generally follow its vertical structure.

Figure 5.3, panels B and C, show the measurement error estimates corresponding to each multiplicative bias profile in panel A. Error variance estimates calculated by Methods 1, 2 and 3 are converted to SD estimates by taking the square root. The error SD profiles are in units of parts per billion. Also shown are the reported error SD profiles, based on the reported errors for each instrument data set, calculated as the square root of Equation 5.34.

Method 1 leads to negative error variance estimates at a number of heights. The error SD is undefined for such cases, but is set to zero for the profile plots (e.g. at 10 hPa for ACE-FTS; 0.7, 14, 23, and 100 hPa for MLS). Negative variance estimates can result from a mis-specified model or an insufficient sample size (Dunn, 1989). In this case, comparing Panel A to Panel C shows that the negative variances calculated by

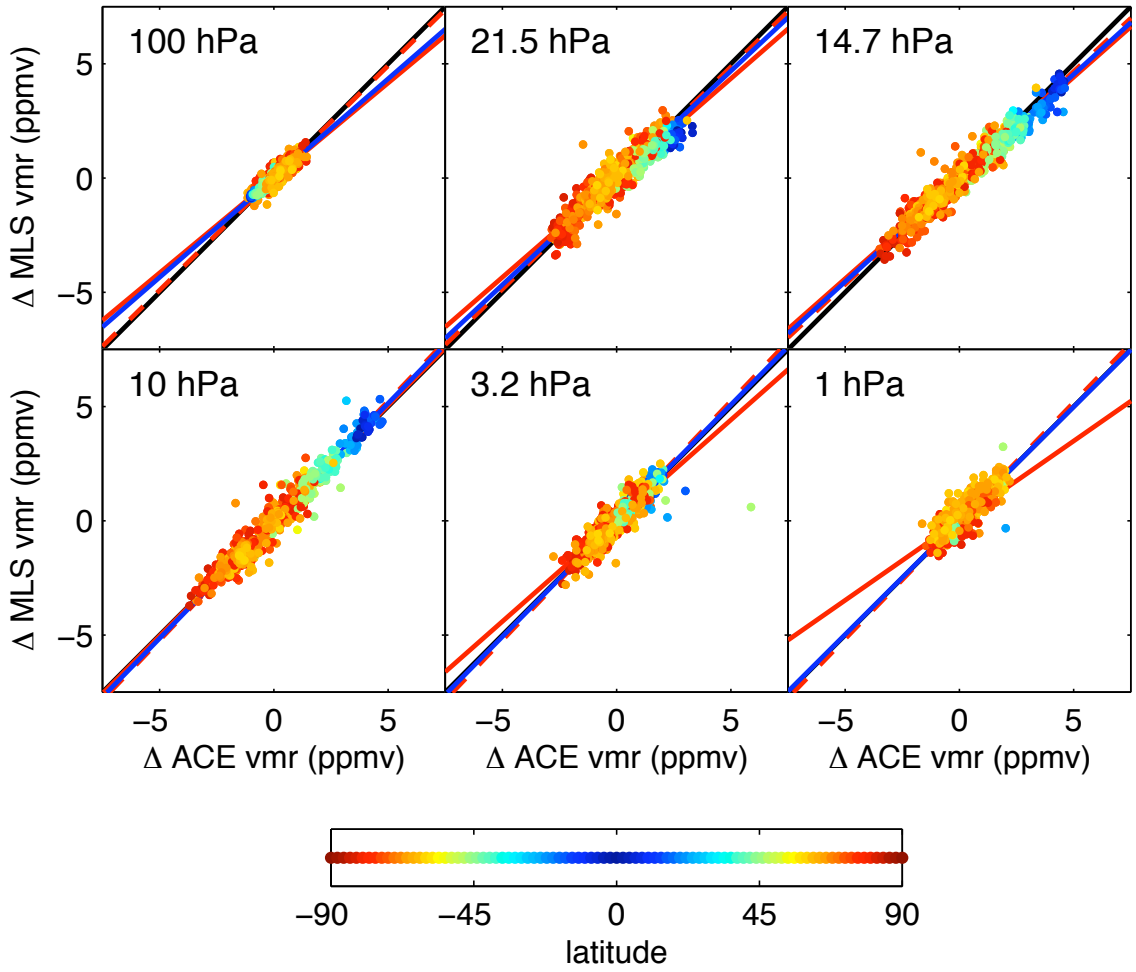

Figure 5.2: Scatter plots of MLS vs. ACE-FTS measured O_3 anomalies for selected pressure surfaces. Lines have slope: $\beta = 1$ (black), $\hat{\beta}_2|_{\tilde{\sigma}^2_\delta}$ (red), $\hat{\beta}_2|_{\tilde{\sigma}^2_\epsilon}$ (red dashed), and $\hat{\beta}_3$ (blue). Colour-coding of points is based on the absolute value of latitude, with equatorial measurements in blue, and polar measurements in red.

Method 1 for MLS occur at those heights for which the multiplicative bias estimates are furthest from a value of 1. Hence, at these heights the $\beta = 1$ assumption of Method 1 is most suspect, and the method leads to erroneous estimates. To be more specific, the form of Equation 5.29 shows that if the true multiplicative bias is less than 1, then assuming $\hat{\beta} = 1$ gives too much weight to the second term on the right-hand-side, leading to a negative value for the error variance. Similarly, the form of Equation 5.28 shows that if the true multiplicative bias is greater than 1, then assuming $\beta = 1$ will lead to negative values for the ACE-FTS error variance. This

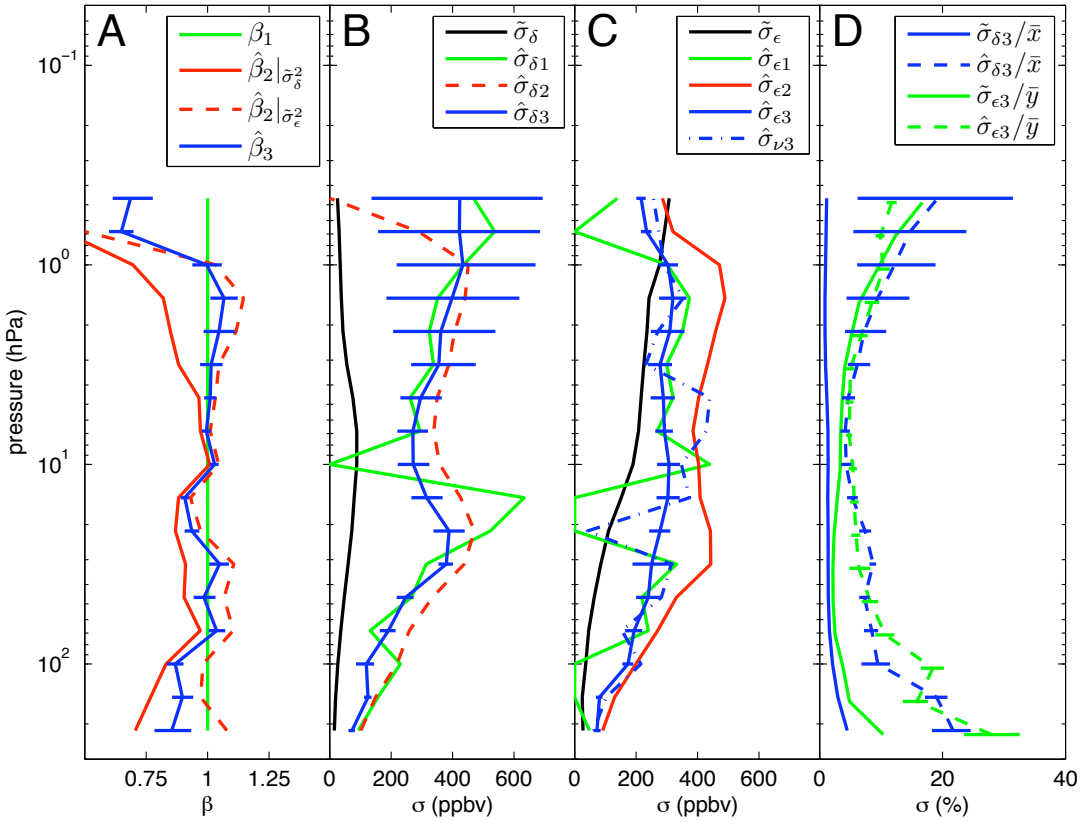

Figure 5.3: Estimated multiplicative bias and random error profiles from the comparison of ACE-FTS and MLS O_3 measurements. A: Three estimates of the multiplicative bias ($\hat{\beta}_2|_{\tilde{\sigma}_\delta^2}$ derived from reported ACE-FTS error variances, $\hat{\beta}_2|_{\tilde{\sigma}_\epsilon^2}$ derived from reported MLS error variances, and $\hat{\beta}_3$; Equations 5.32, 5.35 and 5.41) compared to $\beta_1 = 1$. B: Predicted and estimated ACE-FTS error SD profiles (Equations 5.34, 5.30, 5.36, and 5.42). C: Predicted and estimated MLS error SD profiles (Equations 5.34, 5.31, 5.33, 5.43, and 5.44). D: Predicted ($\tilde{\sigma}$) and Method 3 estimated ($\hat{\sigma}$) measurement error SD profiles for ACE-FTS (blue) and MLS (green), in percent of mean measurement. 95% confidence intervals for all quantities estimated through bootstrapping of Method 3 are shown with error bars.

occurs at 10 hPa (Panel B). Such unphysical estimates of the error variances motivate the use of Methods 2 and 3.

In Method 2, the error estimate for each instrument is calculated based on the reported error ($\tilde{\sigma}$) for the other instrument. It should be noted that by assuming that the reported error of one instrument is correct, we attribute any non-coincidence error to the other instrument. Thus, while the estimated MLS error variance contains

the implicit non-coincidence error like in the other Methods, the Method 2 estimated ACE-FTS error variance contains non-coincidence error as well. This fact explains why the Method 2 estimated ACE-FTS error variance is generally slightly larger than the estimate given by Methods 1 and 3. However, these estimates are all much larger than the reported ACE-FTS error variance, showing that there is more scatter in the data than can be explained by the reported errors for the two instruments. Thus, when the ACE-FTS error predictions are used to estimate the MLS error variance profile, the result is significantly larger than the other estimates and the prediction.

Method 3 error SD estimates generally agree with the results of Method 1, except for the problem heights discussed above, for which Method 3 leads to more realistic values. Comparison of $\hat{\sigma}_{\epsilon 3}$ and $\hat{\sigma}_{\nu 3}$ (recalling that $\hat{\sigma}_{\epsilon 3} = \hat{\sigma}_{\nu 3}$ when $\sigma_\eta = 0$) in Panel C shows that the effect of non-coincidence error is negligible below 30 hPa. The 95% confidence intervals calculated by bootstrapping the Method 3 analysis exclude the reported error SD profiles in all cases.

Panel D of Figure 5.3 shows reported error SDs and those estimated by Method 3 in terms of percent of the mean O_3 profile measured by each instrument. Below 100 hPa, the estimated errors for both instruments are up to 20% higher than the reported errors. Between 100 and 4 hPa, MLS estimated errors are ∼1–8% larger than the reported errors, while ACE-FTS estimates are ∼3–8% greater than the reported errors. Between 10 and 1 hPa, MLS estimated errors are within 2% of the reported errors, and ACE-FTS estimates are 3–10% greater than the reported errors, although the 95% confidence intervals at these higher altitudes imply the ACE-FTS error SD estimates could easily be within 5% of the reported errors.

N_2O

Figure 5.4 shows MLS versus ACE-FTS N_2O VMR anomalies for selected pressure surfaces. At high altitudes (e.g., 1 hPa), the absolute variance of MLS anomalies is much larger than that for ACE-FTS. This discrepancy in variance is likely due to noisy MLS retrievals (for which significant averaging is suggested for useful signals (Livesey et al., 2005)). As will be discussed below, the reported MLS N_2O error

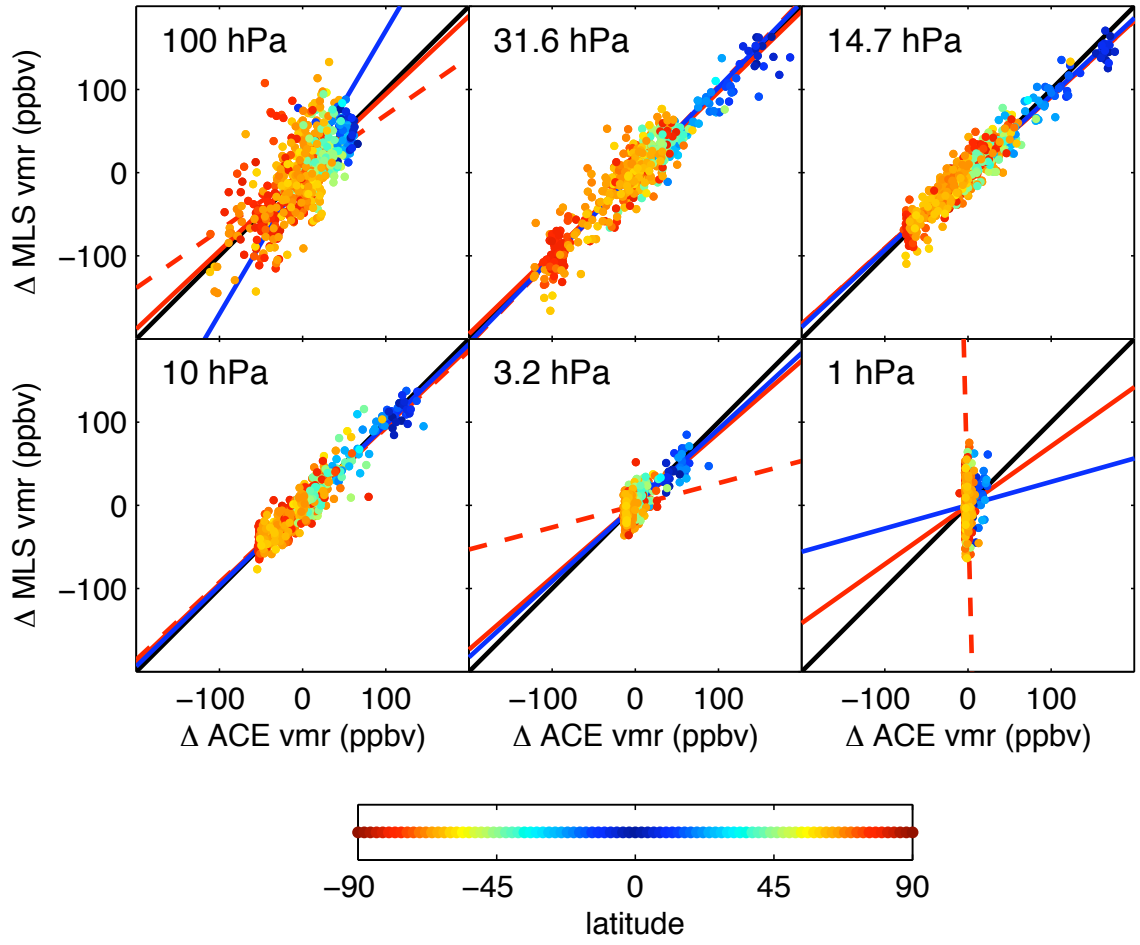

Figure 5.4: As Figure 5.2, but for N_2O.

variances, used as predictions of the error variance in Method 2, are larger than the true random error in the measurements. This leads to anomalously small (e.g., at 3.2 hPa) and even negative (at 1 hPa) estimates of multiplicative bias given by Method 2 when using these reported errors.

Throughout the middle stratosphere, the natural variance of N_2O is larger than at high altitudes, and the scatter lies roughly along the 1:1 line. At 100 hPa, the effects of the reported noisy retrievals at high latitudes is apparent, as the mask used to filter the N_2O data is seen to filter most, but not all, of the anomalously noisy data.

Figure 5.5, Panel A shows vertical profiles of the multiplicative bias estimates. At high altitudes, the small natural variability, and the large MLS variance, leads to

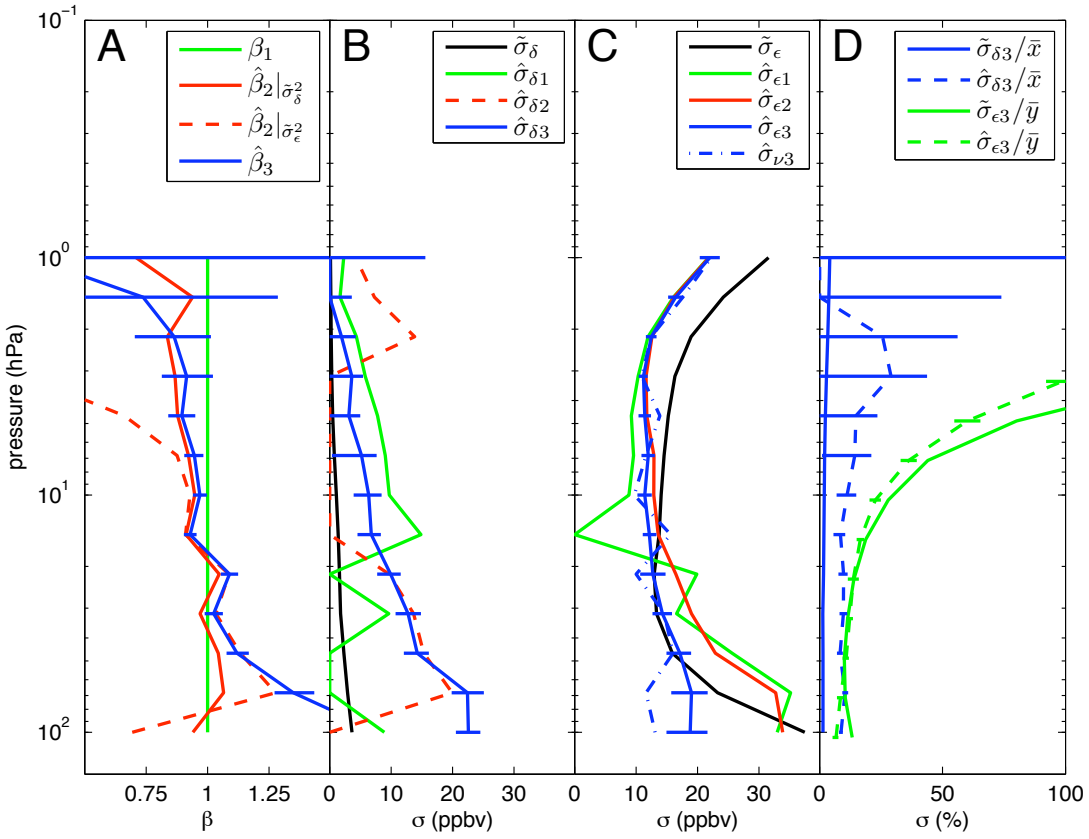

Figure 5.5: As Figure 5.3, but for N_2O.

large uncertainties in the slope estimates, and the uncertainties in the estimate from Method 3 include a slope of one. Moving down in altitude, between ∼4 and 14 hPa, the estimated bias is slightly, but significantly less than one. At heights below 30 hPa, $\hat{\beta}_2|_{\tilde{\sigma}_\epsilon^2}$ and $\hat{\beta}_3$ values are anomalously large due to the noisy MLS polar measurements, while $\hat{\beta}_2|_{\tilde{\sigma}_\delta^2}$ remains in close agreement with $\beta = 1$.

The error SD estimates corresponding to the bias estimates in Panel A are shown in Panels B and C. As was the case for O_3, Method 1 leads to a number of negative error variance estimates (as shown by plotted error SD values of zero) at heights where multiplicative bias is significant.

Focussing on the MLS error SD estimates of Panel C, at low altitudes the estimates based on Methods 1 and 2 agree closely and are ∼5–10 ppbv larger than the reported values, as might be expected due to the anomalous polar retrievals. Method

3 apparently underestimates the error SD at these low altitudes, as the slope estimate $\hat{\beta}_3$ is seen instead to compensate for the noisy polar measurements. Between 14 and 68 hPa, the 95% confidence intervals of $\hat{\sigma}_{\epsilon 3}$ include the reported error SD profile.

MLS error SD estimates from all three methods converge for heights above 10 hPa. At these heights, the large variance of the MLS measurements (s_{yy}) dominates all other terms in the expressions for MLS error variance. The MLS error SDs above 10 hPa are smaller than the reported values. This result is consistent with simulation results discussed by Livesey et al. (2005) in which the variance of quantities retrieved from simulated, noisy spectra was less than the error variance (or precision) reported by the retrieval algorithm. This occurs when the effect of "retrieval smoothing" artificially relaxes retrievals towards a mean or an a priori value.

Where the estimated MLS error SDs are smaller than the reported values, the estimation of ACE-FTS error SD based on the MLS predictions via Method 2 leads to anomalously negative variance values. Only Method 3 leads to physically acceptable error SD profiles for ACE-FTS over the vertical measurement range.

Panel D shows the estimated and reported error SD profiles in percent of the mean measured profile. The ACE-FTS error SD estimates are roughly 8% higher than the reported errors between 100 and 5 hPa. Above 5 hPa, the ACE-FTS error SD 95% confidence intervals become exceedingly large. The estimated MLS percent-error SD profile is in good agreement with the reported profile, although the estimate is smaller than the prediction at heights above 10 hPa, as discussed above.

HNO_3

HNO_3 provides an interesting test case as the scatter plots in Figure 5.6 are markedly sloped, signaling the definite presence of multiplicative bias. The slopes estimated via Methods 2 and 3 (Figure 5.7, Panel A) peak in magnitude at 21 hPa, where the absolute natural variance is largest.

By erroneously ignoring multiplicative bias, Method 1 leads to negative error variance estimates for ACE-FTS (Figure 5.7, Panel B) and, simultaneously, MLS error SD estimates (Figure 5.7, Panel C) up to five times the reported error SD

values over the vertical measurement range.

Method 2 leads to error SD estimates that are more realistic than those produced by Method 1. The Method 2 MLS error SD estimate (Panel C) is smaller than the reported error above 10 hPa, which is again consistent with the simulation results discussed by Livesey et al. (2005), wherein HNO_3 scatter was found to be smaller than the reported variance due to retrieval smoothing at these heights. Using reported error variances for MLS that are greater than the observed measurement variances in Method 2 again leads to anomalously small $\hat{\beta}_2|_{\tilde{\sigma}^2_\epsilon}$ estimates (Panel A) and correspondingly negative error variance estimates for ACE-FTS between 10 and 4 hPa.

Panel D shows the estimated and reported error SD profiles in percent of the mean measured profile. Between 10 and 70 hPa, Method 3 error SD estimates for ACE-FTS are roughly constant with height, and are approximately 10% (of the mean measurement value) larger than the predictions. The Method 3 error SD estimates for MLS are within ±5% of the reported values between 10 and 147 hPa. While the reported ACE-FTS error SD profile is a factor of 10–25% smaller than that for MLS, the estimated percent-error SD profiles are of comparable magnitude between 10 and 100 hPa. Only at high (above 10 hPa) and very low (146 hPa) heights does the comparison support the prediction of greater precision for the ACE-FTS measurements, as shown by the separation of the estimated percent-error profiles.

In contrast with the error SD estimates for N_2O and O_3 (below heights of 10 hPa), the secondary coincidence error SD ($\hat{\sigma}_{\nu 3}$) profile lies outside the 95% confidence interval of the $\hat{\sigma}_{\epsilon 3}$ estimate. This means that the non-coincidence variance σ^2_η is significant in this case, which suggests that a tighter coincidence criterion may be required for HNO_3 compared to the other species. This could be explored by repeating the analysis with a tighter coincident criterion, and comparing the results.

5.5.3 Discussion

The theoretical framework for the statistical comparison of coincident measurements used in this chapter introduced the use of multiplicative bias, which considerably com-

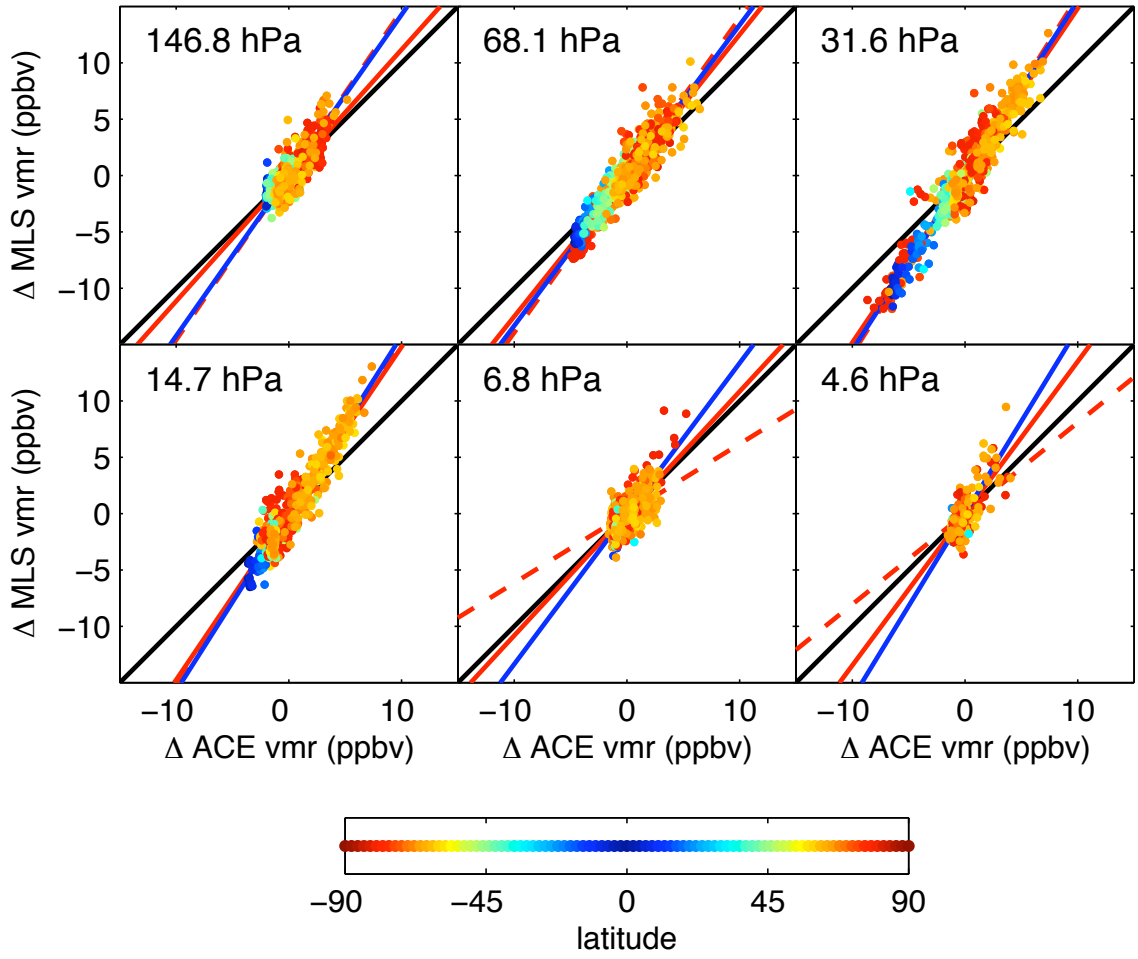

Figure 5.6: As Figure 5.2, but for HNO_3.

plicated the interpretation of the variances and covariance of the measurements. One might question whether multiplicative bias is significant in trace gas measurement comparisons. In the analysis of ACE-FTS and MLS coincident comparisons above, multiplicative bias was found to be significant in many cases, as evidenced either by β estimates (via Method 2 or 3) significantly different from 1, or by negative error variance estimates via Method 1. Multiplicative bias was found to be significant in the comparison of HNO_3 measurements at all heights. Santee et al. (2007) report that the version 1.51 MLS HNO_3 retrievals used here suffered from bias due to an error in a spectroscopic parameter used in the retrievals, and this error was subsequently corrected in version 2 MLS HNO_3 retrievals. Since there exist significant differences in spectroscopic parameters used in different retrievals (Flaud et al., 2006),

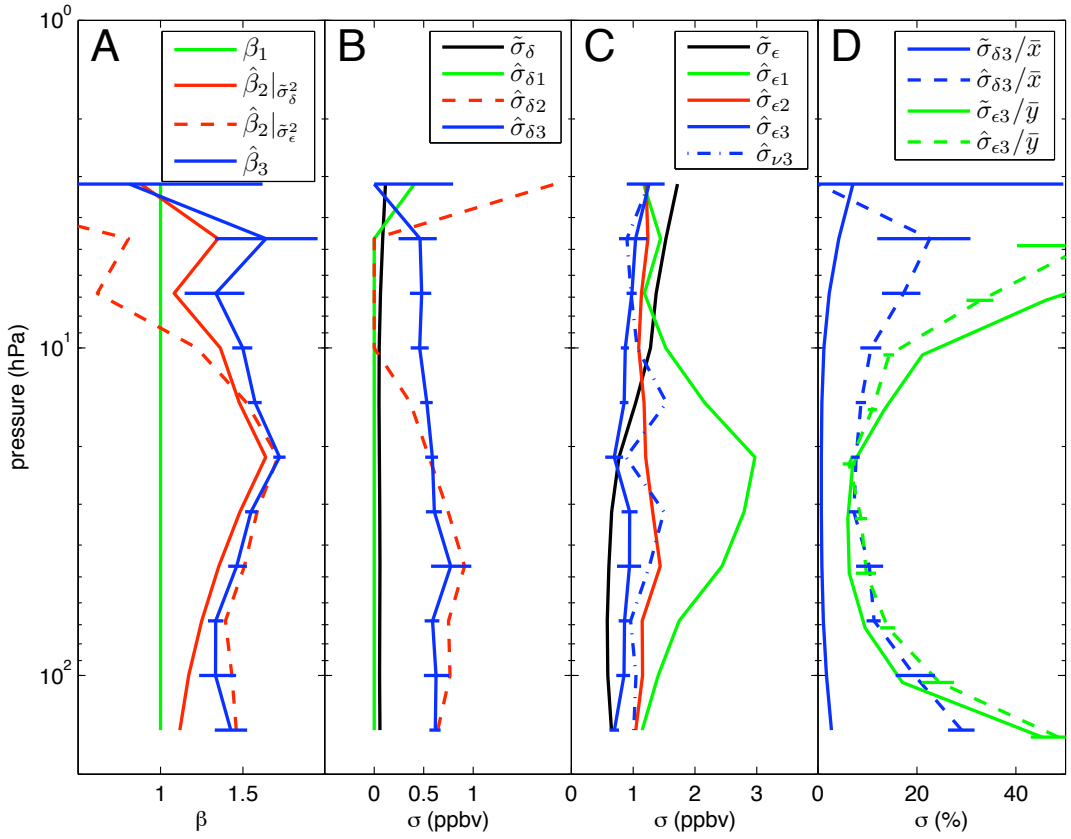

Figure 5.7: As Figure 5.3, but for HNO_3.

it seems likely that multiplicative bias is common between instruments (e.g., Lary and Lait, 2006). Including multiplicative bias in the analysis of coincident measurements promises to aid in the identification of biases created by retrieval differences such as spectroscopic inconsistencies.

Two methods for estimating the multiplicative bias have been proposed and tested. Method 2 produces β estimates based on the reported errors of one instrument. A serious drawback of this method is that the reported errors for many instruments are composed of more than just the random error, and may contain systematic errors. When the reported error variance for instrument A is larger than the scatter in its measurements, estimation of the random error for instrument B and the multiplicative bias will be significantly skewed (e.g., for N_2O above 10 hPa).

A better approach may be to estimate the random error of one instrument directly

from the data, and use this estimate in order to estimate β. This is in effect the approach of Method 3, wherein self-coincidences of MLS are used in the analysis. Such self-coincidences are used in MLS validation studies (e.g., Froidevaux et al., 2008; Lambert et al., 2007) in order to confirm the reported errors. There is no reason that random errors for MLS could not have been calculated first based on the analysis of the self-coincidences, and then used in an estimation of multiplicative bias. The instrumental variable technique of Method 3 in essence provides a framework for combining these two steps. Multiplicative bias estimates produced by Method 3 seem generally realistic, and instances where the technique reports significant multiplicative bias are supported by the calculation of negative error variances by Method 1. Two instances where Method 3 seems to be lacking are (1) for N_2O retrievals at low altitude, where noisy MLS retrievals lead to a skewed β estimate rather than a high error variance estimate, and (2) for HNO_3, where, based on the error in the v1.51 MLS retrievals reported by Santee et al. (2007), one would expect a constant multiplicative bias with height. The HNO_3 $\hat{\beta}_3$ estimate is seen to depend somewhat on height, with a maximum estimated value at 21.5 hPa, where the natural variance is largest.

Estimation of the multiplicative bias requires information about the random error of at least one instrument, and in the estimation methods used, produces an estimate of the other random instrument error. In theory, this random error estimate could be used to validate the reported instrument error - something that is not commonly produced from a comparison of coincident measurements. Method 2 produces a random error variance estimate based on the reported error of the other instrument. Method 3 produces a random error estimate for MLS based (effectively) on the variance of its self-coincidences, and uses this error variance to produce an estimate of the ACE-FTS error variance. Each error variance estimate produced is an upper limit, since each includes the effect of non-coincidence error.

The error variances for MLS estimated by Method 3 are in qualitative agreement with the estimates included in MLS validation studies by Froidevaux et al. (2008), Lambert et al. (2007), and Santee et al. (2007). These random error variances disagree somewhat with the reported MLS errors, in many cases since the reported errors

include systematic effects and smoothing error. For O_3, Froidevaux et al. (2008) cite atmospheric variability as the cause of the discrepancy between MLS reported errors and the calculated empirical estimates of precision.

The error variances for ACE-FTS estimated by Method 3 are significantly larger than the reported ACE-FTS error variances. While the reported errors for ACE-FTS are an order of magnitude or more smaller than those for MLS, only in a few cases are the error variances estimated by Method 3 smaller for ACE-FTS than for MLS (N_2O above 20 hPa, HNO_3 above 10 hPa and below 100). This could be a consequence of the true random error of ACE-FTS being larger than the reported error, the error in ACE-FTS pressures used to interpolate to the MLS pressure grid being significant, or non-coincidence error being significant. If the non-coincidence error is significant, i.e., the natural variance within the spatiotemporal coincidence criterion is comparable to or greater in magnitude than the measurement error variances, one would expect the derived error variances for each instrument to be similar in magnitude (since the coincidence statistics between ACE-FTS and MLS and between MLS measurements are very similar). This is in fact what is seen: the estimated error variances are almost equal for O_3 over the full comparison range, for N_2O below 20 hPa, and for HNO_3 between 10 and 100 hPa. It thus seems reasonable to hypothesize that non-coincidence error is having a significant effect on both estimated error variances. This effect is especially noticeable when comparing ACE-FTS estimated and reported errors, since the ACE-FTS reported errors are significantly smaller than the MLS errors. On the other hand, Froidevaux et al. (2008) showed empirical estimates of MLS O_3 precision that are slightly but significantly higher than the reported MLS errors: apparently, O_3 variability is significant compared to the MLS reported errors. The natural variability of the atmosphere thus complicates the estimation of the random errors of atmospheric measurements, and the validation of reported errors. This issue will be explored in more detail in Chapter 6.

5.6 Summary

This chapter has explored the relationship between the variance and covariance of coincident measurements, and random measurement errors and inter-instrumental biases. A measurement model was developed that differentiates between additive and multiplicative bias. In a comparison of coincident measurements by the ACE-FTS and Aura MLS instruments, multiplicative bias is found to be significant at certain heights for O_3 and N_2O, and at all heights for HNO_3. Future comparison studies that estimate both multiplicative and additive biases may produce validation metrics that better diagnose underlying differences in the data sets and instruments: for instance, identification of a multiplicative bias at all heights may be a clear signal of a difference in the spectroscopic parameters used in the retrievals.

Since the estimation of multiplicative bias is tied to the random errors in each data set, the methods used to estimate multiplicative bias also produce estimates of the random error. Comparing the random error estimates with the reported errors of ACE-FTS and Aura MLS leads to the conclusion that natural variability is likely a significant portion of the estimates. Nonetheless, the error variances estimated do represent upper limits on the random errors and therefore have scientific value. Chapter 6 will explore other techniques for the estimation of random errors, focussing on the ACE-FTS measurement data set.

Chapter 6

Validating the reported random errors of ACE-FTS measurements

6.1 Introduction

An important component of any measurement is the error reported with that measurement, which quantifies the expected difference between the measurement and the truth. Interpretation of a measurement may depend on the reported error: for instance, determining whether a measurement is consistent with an expected value, or with another measurement, depends on the reported error in the measurement.

Remote sounding measurements of atmospheric trace gas abundances are produced by retrieving trace gas profiles from spectra collected by an instrument. These measurements typically report an associated error, which can be said to be a bottom-up estimate, based on a priori knowledge of the measurement system and the uncertainties in the various assumptions and parameters used to produce the measurements. Propagating the various uncertainties through the measurement process and producing a corresponding error in the retrieved quantity is analogous to propagating errors in measured length through a calculation used to produce an estimate of some physical constant in the laboratory setting.

Assessment of the errors reported with trace gas measurements is understood to be an important part of the full satellite validation process (von Clarmann, 2006).

Since the true value of the quantity measured is rarely if ever known, validating the reported errors amounts to showing that the measurement data is consistent with the reported errors.

ACE-FTS measurements are generally believed to be of very high quality: the spectra that it collects are of high spectral resolution and show a high signal-to-noise ratio, the reported errors (as currently estimated) for many species are as low as 1%, and retrieved profiles show good agreement with other space-based missions. This chapter aims to quantify the quality of the ACE-FTS measurements through a validation of its reported random errors.

6.1.1 Measurement error

Measurement error can be decomposed into two parts, that which is consistent from one measurement to the next, and that which varies. The terms systematic error and random error, respectively, are used to describe these two types of uncertainty (e.g., Taylor, 1997).

Systematic errors are most readily identified through the comparison of measurement data sets produced by different instruments. Comparison of trace gas measurements by different instruments makes up the majority of most validation studies, and is used to determine the systematic difference between the data sets. Chapter 5 introduced a technique for extending the bias estimation process typical for satellite validation studies, by including multiplicative bias in the comparison model. In theory, once a systematic error has been identified through a comparison with a standard instrument or data set, it can be corrected for. This process is known as calibration (Dunn, 1989).

Random error is caused by inherently unpredictable fluctuations in the measurement system, and as such cannot be eliminated in the way that calibration aims to eliminate systematic error. The simplest measurement model describes a measurement x_i as being the sum of the true quantity being measured, τ_i, and some zero-mean random error ϵ_i:

$$x_i = \tau_i + \epsilon_i. \tag{6.1}$$

The most straightforward way to assess random error is to examine the results of repeated measurements on a single static quantity. Due to random error, the measured values will vary, and any quantification of the scatter in repeated measurements represents a quantification of the random error. For repeated measurements of the same quantity, $\tau_1 = \tau_2 = ... = \tau$. Given enough measurements, the true value τ is well estimated by the mean value of the measurements, $\bar{x}$. The differences between x_i and $\bar{x}$, i.e., the deviations about $\bar{x}$, then represent estimates of the errors ϵ_i:

$$x_i - \bar{x} \approx \epsilon_i. \tag{6.2}$$

In theory, any measure of the scatter of x about $\bar{x}$ represents a measure of the scatter in ϵ, and so can be used to quantify the random error. In practice, the scatter in a set of n measurements is most often quantified by the SD:

$$s_x = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2}. \tag{6.3}$$

Comparing Equation 6.3 with Equation 6.2, the SD of repeated measurements, used to quantify the overall uncertainty of the measurement process, is seen to be equivalent to the root-mean-square of the individual errors:

$$s_x = \sigma_\epsilon = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(\epsilon_i)^2}. \tag{6.4}$$

The use of the SD to quantify the random error is based at least partly on the general observation that random errors are often normally distributed (as, according to the Central Limit Theorem, would be the case if random errors are the result of some large number of independent random processes). In addition, the properties of the normal distribution lead to convenient interpretations of the errors. For instance, if the errors are normally distributed, 68% of repeated measurements should lie within σ_ϵ of the truth. Then, if a single measurement is made with the same instrument (with its known σ_ϵ error), one can infer that the truth has a 68% probability of being within

σ_ϵ of the measurement, and thus the error represents a 68% confidence interval for the measurement. Likewise, $2\sigma_\epsilon$ errors represent a 95% confidence interval for the measurement.

The statistical analysis of repeated measurements discussed above represents a top-down approach to random error assessment. As mentioned previously, random errors can also be estimated through a bottom-up approach, based on knowledge of the random fluctuations inherent to the measurement process. Bottom-up approaches themselves may take two forms. In the first, the resulting scatter in measurements due to uncertainties in independent measurement components may be estimated, and the overall random error calculated as the RMS of the independent relative random errors. This process, like the analysis of repeated measurements, gives a single-valued measure of the random error, which represents the width of the normal distribution that approximates the distribution of random errors.

A second method for bottom-up error estimation may examine the particular circumstances of each measurement, and produce an estimate of the random error for each measurement. Like the case described above, if the errors are assumed to be normally distributed, the single measurement error represents a 68% confidence interval for that particular measurement. The RMS of the individual errors represents (as in Equation 6.4) an overall measure of the random error, equivalent to the overall error which quantifies the scatter in repeated measurements.

6.1.2 ACE-FTS reported random errors

Associated with each individual VMR profile retrieved by ACE-FTS is a corresponding reported random error (RRE) profile, representing a statistical 1σ error, estimated through the retrieval spectral fitting process (Boone et al., 2005). Importantly, these RREs do not include any estimate of systematic error. In addition, it should be stressed that they are produced based on random errors in the spectral fitting process only, and so do not include possible random errors due to such factors as tangent height uncertainty, or interpolation to the 1-km vertical grid.

The RRE for any particular profile depends on the signal-to-noise ratio (SNR) of

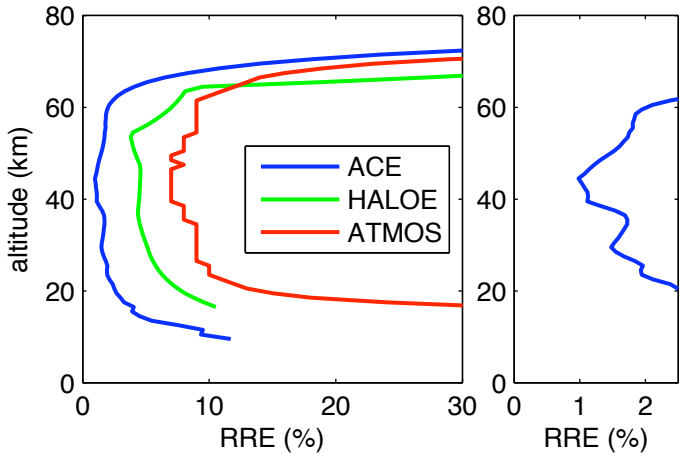

Figure 6.1: Comparison of the reported random error (RRE) in O_3 retrievals for the solar occultation instruments ACE-FTS, HALOE and ATMOS. The right hand panel shows the ACE-FTS RRE on an enlarged scale.

the measured spectra from which the profile is retrieved. Since the SNR can vary depending on such factors as ice build-up on the detector, or clouds in the instrument field-of-view, the RREs vary from retrieval to retrieval.

The spectra measured by ACE-FTS are typically of very good quality, with high SNR (Bernath et al., 2005). In theory, this should lead to small random errors in its retrieved VMRs. This is confirmed by comparing ACE-FTS RREs with "typical" random errors for other satellite instruments. Figure 6.1 shows the RRE for O_3 retrievals from ACE-FTS compared to those for the Halogen Occultation Experiment (HALOE) and Atmospheric Trace Molecule Spectroscopy Experiment (ATMOS) instruments. All three are solar occultation instruments, and so have similar measurement geometry and conditions. The ACE-FTS RRE profile shown is calculated as the RMS of the RREs for all ACE-FTS retrievals from the year 2005. The HALOE RRE profile shown represents the root-sum-square of the random error components reported in Table 1 of a HALOE O_3 validation study by Brühl et al. (1996). The ATMOS RRE profile represents the reported median random error for ATMOS filter 1 O_3 measurements, shown by Irion et al. (2002).

It is clear from Figure 6.1 that the RREs for ACE-FTS O_3 retrievals are signifi-

cantly smaller than the "typical" random errors reported for HALOE and ATMOS. A certain portion of this is very likely due to the technological superiority of the ACE-FTS instrument, resulting in a higher SNR, and correspondingly lower random errors in its retrievals. But, it is also possible that the ACE-FTS RREs are underestimating the true random error in the retrievals since they are not incorporating other sources of random error, such as pointing errors, which happen to be the primary reported source of error for the HALOE retrievals below approximately 40 km (Brühl et al., 1996).

A more complete analysis of the errors, random and systematic, present in the ACE-FTS retrievals is planned for the future (Boone et al., 2005). In the meantime, this chapter focusses on the problem of trying to validate the reported ACE-FTS random errors. Is it possible to prove or disprove that the ACE-FTS retrievals are as precise as is currently reported?

6.1.3 Random error validation methodology

As discussed above, the random error of a measurement system is defined by the scatter in repeated measurements: a quantification of this scatter describes the scatter produced by the random errors. In a well-controlled experimental set-up, random error can be assessed simply by taking a large number of repeated measurements, and quantifying the resulting scatter.

In terms of the simple measurement model of Equation 6.1, assuming that the errors ϵ_i are uncorrelated with the truth τ_i, the variance of any measurement set (repeated or not) is equal to the sum of the variances of the truth and the measurement error:

$$\sigma_x^2 = \sigma_\tau^2 + \sigma_\epsilon^2. \tag{6.5}$$

Under the special case of repeated measurements of the truth τ, the natural variance σ_τ^2 is zero, and the measurement variance is thus equivalent to the random error variance, $\sigma_x^2 = \sigma_\epsilon^2$. Taking the square root, the SDs are thus also equal, $\sigma_x = \sigma_\epsilon$, and this is really just a restatement of Equations 6.3 and 6.4.

The atmosphere is, however, not the static test object necessary for the repeated measurements outlined above. The atmosphere varies in space and time due to the chemical and dynamical processes that control trace gas distributions. Any set of trace gas measurements will therefore exhibit scatter due to a combination of both the natural variability of the true atmosphere, and scatter due to random measurement errors.

It is important to note that while random errors and natural variability are almost exclusively quoted in terms of standard deviation, it is the variances that add (via Equation 6.5) to give the measurement variance. In a practical sense, this means that natural variability and measurement error SD must be quite similar in magnitude to both be significant in the calculation of the measurement SD. For example, if the random error SD is 1/10 of the natural variability SD (i.e., $\sigma_\epsilon \sim \sigma_\tau/10$), it will contribute to only 1/100 of the measurement SD. Thus, it is reasonable to expect agreement between the reported random errors and the measurement scatter (i.e. $\sigma_x \approx \sigma_\epsilon$) in regions when and where the natural variability is significantly less than the random error.

Validation of random measurement errors using the method of repeated measurements described above thus depends on the identification of regions where natural variability is small compared to the measurement errors. A number of satellite validation studies, including those for the solar occultation instruments HALOE (Brühl et al., 1996) and ATMOS (Abrams et al., 1996), have shown good agreement between reported errors and measured variability in the tropical region, implying minimal natural variability there. This finding will be explored in §6.2, which will extend the discussion of short-term variability found in §3.2.3, and identify the tropics as a region of minimum variability. §6.3 will focus on the variability seen in the tropical ACE-FTS retrievals, and will assess how measured scatter is affected by the temporal sampling of the scatter statistics. The scatter in tropical measurements will then be compared to the ACE-FTS RREs. §6.4 will apply a similar analysis to simulated chemical fields from the CMAM, in an effort to resolve discrepancies between ACE-FTS measured scatter and RREs.

6.2 Identifying regions of minimum trace gas variability

As was discussed in detail in §3.2.3, short-term stratospheric trace gas variability is a product of both chemical and dynamical processes. Species with short chemical lifetimes vary due to changing photochemical conditions, which vary on a diurnal cycle. Long-lived species, with lifetimes on the order of transport time scales, will vary due to the effect of large-scale dynamical motions acting on time-mean spatial gradients of chemical tracers, which are themselves a product of both photochemical and dynamical processes. Patterns of short-term variability are therefore different for different trace gases, each with their own unique chemical lifetime and resulting spatial gradients.

Due to the strength of the zonal winds, long-lived trace gas species are expected to be, to first order, zonally symmetric, i.e., their VMR depends strongly on latitude and height, and to a much lesser extent on longitude. Short-term variability is therefore often expressed as variance about a zonal mean.

The discussion of §3.2.3 was based most heavily upon observed SDs about zonal means from the CRISTA experiment (Kuell et al., 2004). While these results are, to the author's knowledge, some of only a handful of published plots showing zonal variability for a number of trace gases, the data set is limited in that CRISTA made measurements for only eight days in 1997. A more long-term examination of natural trace gas variability within zonal bands can be extracted from a climatology of trace gas measurements from HALOE (Grooß and Russell III, 2005). In this climatology, monthly means and SDs of trace gas VMRs of O_3, H_2O, CH_4, NO_x, HCl, and HF measured by HALOE are reported for 5° latitude and equivalent latitude bins over the globe, calculated from all data spanning the years 1991-2002.

The statistics of this overall climatology are calculated from multiple years' worth of HALOE measurements for each spatial bin and calendar month. The SDs of the climatology thus contain the effects of both short-term and interannual variability (in addition to random measurement errors). For the present purpose of identifying

regions of minimum short-term variability, it would be advantageous to remove the interannual variability from the reported HALOE SD climatology. This is possible to a degree. In addition to the climatological monthly means and SDs, the HALOE climatology makes available (by download) means and standard deviations of the observations for the 5° equivalent latitude bins for each individual month of the full time series. It is possible to derive an estimate of the typical short-term variability which does not include interannual variability by averaging the SDs of the individual months of HALOE observations.

It should be kept in mind that since this HALOE variability climatology is derived in terms of equivalent latitude, the SDs may be significantly smaller than would be calculated for the same climatology in latitude space, especially in the winter hemisphere surf zone. It would be interesting to compare the magnitude of short-term variability expressed in terms of equivalent and traditional latitude; unfortunately this is not possible with the HALOE data as it is currently available.

Figure 6.2 shows equivalent latitude-height slices of the monthly SDs for O_3 and CH_4 averaged over the full 11-year HALOE data set. The trace gases O_3 and CH_4 are chosen here as two representative but contrasting examples of stratospheric variability. CH_4 is a long-lived species with a tropospheric source and horizontal gradients typical of other long-lived species, and is often used as a tracer of dynamical transport. O_3 is also relatively long-lived in the lower stratosphere, but is short-lived in the upper stratosphere. Its horizontal gradients are significantly different from those of CH_4, due to their different sources.

The minimum percent SD for any pressure surface is found (for both species) in a band centered on the equator. CH_4 variability is less than 10% through a large region of the lower tropical stratosphere. The measured variability increases with height, reaching a local maximum of about 12% at approximately 1 hPa, and decreases again above. For O_3, a band of variability less than 10% exists through all but the lower stratosphere.

The time dependence of the short-term variability of HALOE measurements is examined in Figure 6.3. The SDs for each individual month of the HALOE data set

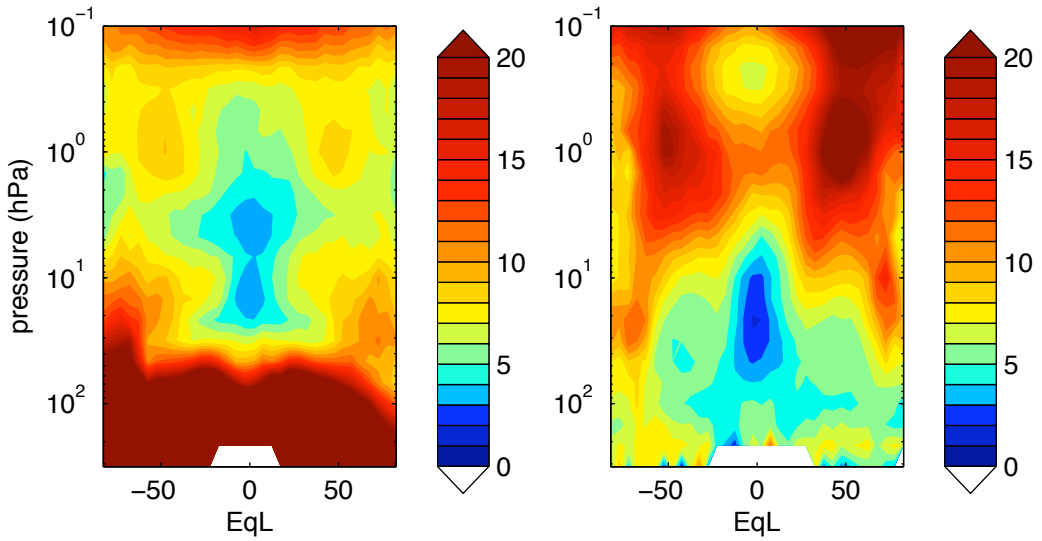

Figure 6.2: HALOE monthly zonal percent standard deviations for 5° equivalent latitude bins, averaged over the full 11 years of data, resulting in an annual mean short-term SD climatology for O_3 (left) and CH_4 (right).

have here been averaged for each calendar month. Figure 6.3 shows the resulting short-term SD climatology as a function of equivalent latitude and month at the pressure surfaces 31.6, 10 and 3.16 hPa, for CH_4 and O_3.

CH_4 variability in the lower stratosphere (31.6 hPa) is strongest in the polar spring region, with SD values exceeding 20% at latitudes indicative of the polar vortex edge. This variability is suggestive of variability produced by the displacement of the polar vortex from zonal symmetry, and of mixing across the vortex edge associated with the spring-time breakup of the vortex. CH_4 variability is minimal for the rest of the time-latitude space of the lower stratosphere, with the tropical 10°S–10°N band in particular showing the lowest SD values throughout the year.

In the middle stratosphere (10 hPa), the time-equivalent latitude distribution of CH_4 variability is more varied. Variability associated with the polar vortex edge is again strong at 10 hPa, and is seen to commence earlier in the winter than at 31.6 hPa. The wintertime variability also extends into the midlatitudes, which is indicative of the large-scale mixing of the surf zone. The midlatitude surf zone variability is seen to persist into the summer months, especially in the NH. This is quite possibly

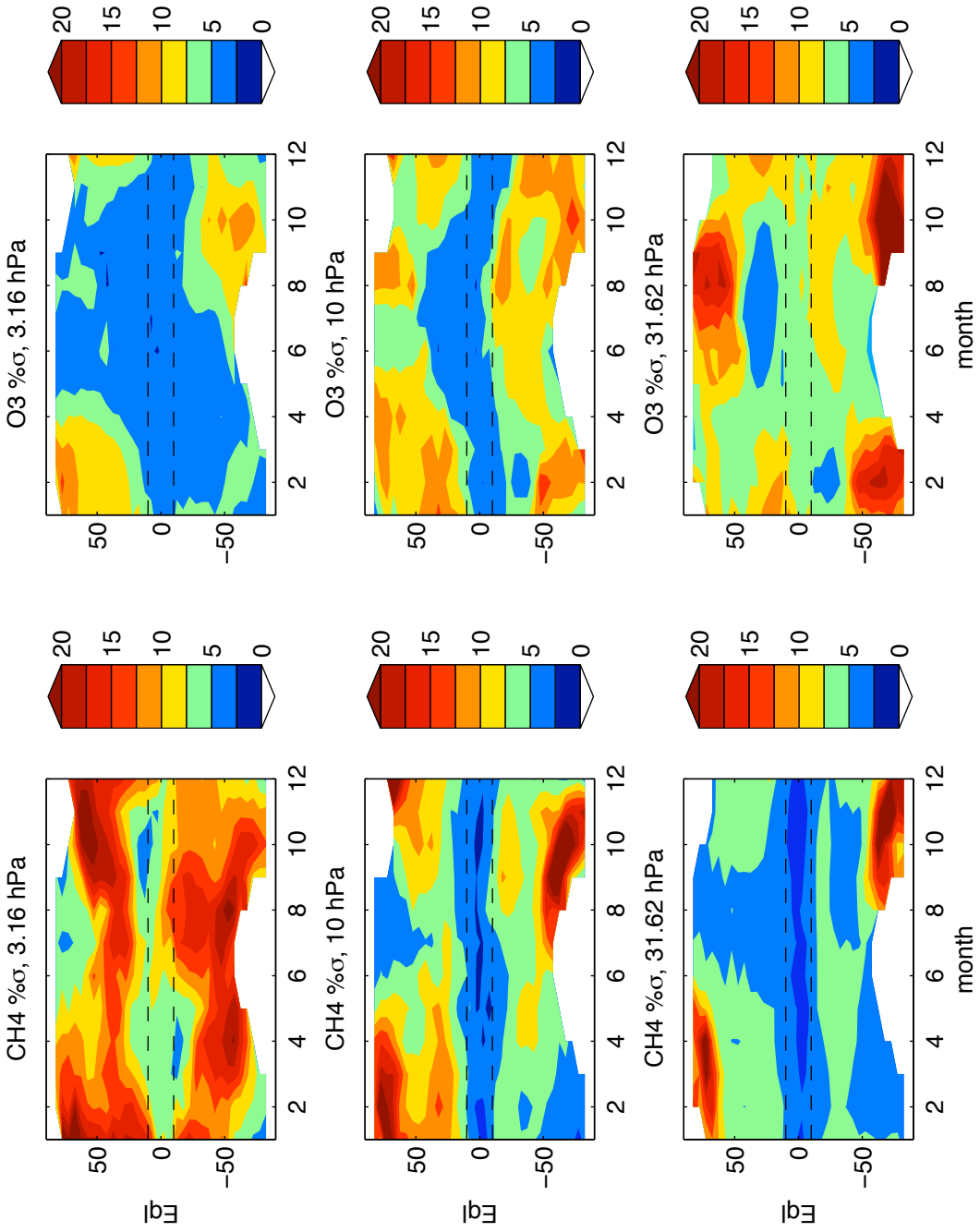

Figure 6.3: HALOE percent SDs plotted as function of month and equivalent latitude for CH_4 and O_3 at the pressure surfaces 31.6, 10 and 3.16 hPa. Dashed lines mark the 10°S to 10°N tropical region.

a signature of the frozen-in variance, produced by large-scale stirring and the long lifetime of CH_4, postulated by Hess and Holton (1985). A band of minimal variability is seen again in the tropics, with variability less than or equal to 5%. Variability is also minimal at high latitudes in summer.

The structure of CH_4 variability at 3.16 hPa is similar to that seen at 10 hPa, with a larger percent magnitude. Strong variability ($> 10\%$) is seen at 30°N through the full year. Again, minimum variability is found in the high-latitude summer, and in a band in the tropics, although the tropical band is seen to be "squeezed" by the midlatitude variability on either side of the equator during solstices.

The largest magnitudes of lower stratospheric (31.6 hPa) O_3 variability are found in the SH winter and spring. As was seen for CH_4, this variability is suggestive of variability produced by polar vortex asymmetry, and of mixing across the vortex edge associated with the spring-time breakup of the vortex. The relative strength of winter and spring variability for O_3 in the SH compared to the NH is likely due to the strong meridional gradients in O_3 across the vortex edge set up by the Antarctic O_3 hole. O_3 variability is also strong in the high-latitude summer of both hemispheres. This is consistent with observations and model results (e.g., Hoppel et al., 1999; Wagner and Bowman, 2000, , also see §3.2.3) which imply that meridional motions caused by Rossby wave breaking in the lower stratosphere acting on strong meridional gradients produced by the photochemical conditions of the polar summer can lead to significant O_3 variability, while leaving CH_4 and other longer-lived species (with weaker horizontal gradients) unaffected. Minimum O_3 variability is observed just equatorward of the summer maxima, although the tropical band is also relatively quiet, with variability less than 10% for most months.

At 10 hPa, signatures of high-latitude summer variability are still present (but not as strong as at 31.6 hPa). There is also wintertime surf zone variability apparent, concentrated at approximately 30° in each winter hemisphere. This wintertime surf-zone variability is seen to intrude into the 10°S–10°N tropical band. As a result, minimal variability is found in a band that oscillates seasonally about the equator. The intrusion of surf-zone mixing into tropical latitudes in the middle to upper strato-

sphere has been observed in an examination of the position of the "subtropical edge", which separates the relatively isolated tropical region from the winter extratropical surf-zone (Neu et al., 2003). Based on an analysis of the probability distribution functions of CH_4 measurements by HALOE and N_2O measurements by CLAES, Neu et al. (2003) observed the wintertime subtropical edge to remain poleward of 10° for heights below 10 hPa, and intrude into the equatorial region for heights above 10 hPa.

At 3.16 hPa, O_3 variability occurs most strongly at high latitudes in late winter/spring. The tropics show relatively weak variability (less than 5%) throughout the year. Summer midlatitudes are also rather quiet. In addition, signatures of the high-latitude late summer variability are still present.

In summary, the HALOE short-term SD climatology shows that the distribution of trace gas variability is strongly dependent on the species. For example, in summer midlatitudes O_3 shows minimal variability for all heights shown, while CH_4 shows quite large values of variability. Conversely, at high latitudes during summer, CH_4 shows minimal variability, while O_3 shows significantly large values of variability.

The tropics, on the other hand, show weak variability for both CH_4 and O_3 throughout the year. Many studies aiming to validate measurement errors have assumed the tropics to be a region of minimal variability, and shown good agreement between reported errors and tropical measurement scatter. For instance, Brühl et al. (1996) show sample SD profiles of HALOE O_3 measurements over two-day sample sizes, for summer 1992 and 1993, at low (10°, 14°N), middle (43°, 46°N) and high (76°N) latitudes, and show that O_3 SD profiles at low and midlatitudes are comparable in magnitude with each other, and with the ∼5% random errors reported for HALOE, while the high latitudes show considerably higher variance. Similarly, Abrams et al. (1996) quote SDs of ATMOS profiles for a number of species in tropical zonal bands (latitudes not specified), and show their general agreement with the reported random error estimates. Livesey et al. (2005) use SDs of measurements between 10°S and 10°N in order to assess the realism of the reported Aura MLS precisions for a number of species. Santee et al. (2007) show good agreement between

the SD of tropical (10°S–10°N) Aura MLS HNO_3 measurements (for single days) and reported precision estimates.

As a final note regarding the short-term SD climatology from HALOE, it is notable that the minimum short-term O_3 variability reported by HALOE in the tropics is comparable in magnitude to the reported random error of its measurements (∼5%). It is thus reasonable to question whether the minimum SDs are limited by natural variability, or by the scatter produced by random error. In other words, it is possible, based solely on the results of the HALOE climatology, that the natural variability of the tropics is less than 5% (for O_3) since the lower limit of the HALOE SDs, based on the reported random errors, should be ∼5%. This question will be addressed in the following sections.

6.3 Tropical ACE-FTS measurement scatter

Assured by the observational evidence from HALOE of minimal variability in the tropics, the following analysis of ACE-FTS data will focus on the measured variability of trace gas species in the 10°S–10°N tropical region. Issues related to the quantity (not quality) of ACE-FTS tropical measurements will first be introduced, followed by an examination of issues related to temporal partitioning of the data set. Finally, tropical measurement scatter will be compared to the RREs, in an attempt to validate those RREs.

6.3.1 Tropical sampling

The orbit geometry of the ACE-FTS satellite platform allows sampling of the tropical latitudes four times per year, in February, April, August and October (see Figure 6.4). The period of each tropical observation window is short: ACE-FTS samples latitudes between 10°S and 10°N during approximately eight days in each month of coverage. The sampling is split into periods of sunrise and sunset occultations, each of approximately four days length, with sunsets measured in the first half of each month, and sunrises in the second, with approximately 13 days separating the sunset

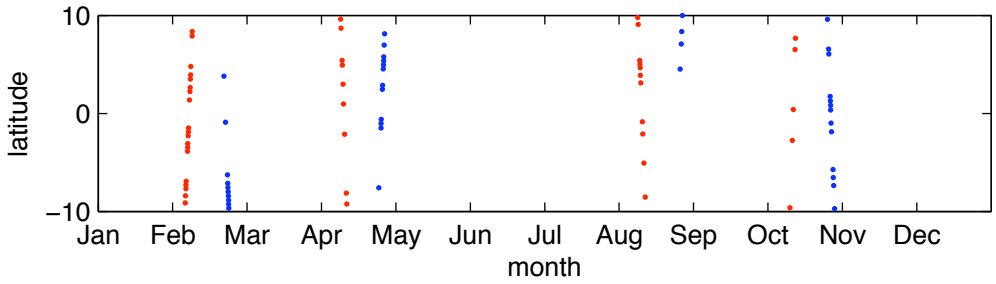

Figure 6.4: Latitude-time locations of retrieved profiles from ACE-FTS between 10°S and 10°N for the year 2006. Sunrises are shown in blue, sunsets in red.

and sunrise measurements.

The ACE-FTS satellite platform performs approximately 15 orbits per day, leading to 15 sunrise and 15 sunset occultations each. According to this occultation frequency, the ACE-FTS platform performs approximately 120 occultations between the latitudes of 10°S and 10°N, per month of coverage. This translates to approximately 480 tropical occultations per year.

Due to restrictions on data downlinking and measurement frequency, the number of retrieved profiles is, however, significantly less than the number of occultations performed by the satellite platform. Table 6.1 lists the number of retrievals produced as a function of month and year. The average number of retrievals per month is approximately 25, much less than the 120 occultations performed.

When examining the scatter of the tropical data set, time restriction is important in order to minimize any variance due to temporal variability (such as the seasonal cycle or interannual variability). Likewise, restricting the latitudes of the sample is important for reducing the variance due to latitudinal gradients. On the other hand, spatiotemporal restriction must be chosen so as to result in sample sizes that lead to statistically significant results. The sample standard deviation, as a measure of scatter, is an estimate of the true population standard deviation. Just as the sample mean has an associated error which depends on the sample size n, the sample standard deviation has an associated standard error (Taylor, 1997) given by:

	Feb	Apr	Aug	Oct
2004	0	27	33	18
2005	13	53	55	29
2006	29	21	15	18
2007	22	21	22	15
2008	29	8	19	0

Table 6.1: The number of retrieved ACE-FTS profiles between the latitudes 10°S and 10°N as a function of year and month.

$$SE = \frac{\sigma}{\sqrt{2n}}. \tag{6.6}$$

For small sample sizes, the relative uncertainty in the calculated SD can be quite large: e.g., for a sample size of 10, the percent error in the SD estimate is greater than 20%. Thus, the sample SD more accurately reflects the underlying population SD as sample sizes increase.

The calculation of empirical random error estimates from the ACE-FTS data must then strike a balance between reducing natural variability and reducing sampling error: tighter spatiotemporal sampling bounds decrease the natural variability, but reduce the sample size, thus increasing the sampling error.

In the following analysis, scatter statistics will be calculated for the tropical ACE-FTS measurements as a function of temporal bound. Examining the measurement scatter present in the data within different temporal sampling regimes will lead to a better understanding of the temporal variability of different time-scales. The statistics calculated with the tightest temporal bounds, and hence the smallest sample sizes, must be examined with the knowledge that uncertainty in the statistics may suffer from the small sample sizes.

6.3.2 Measurement scatter and temporal partitioning

The following analysis will make use of all ACE-FTS retrieved version 2.2 profiles, measured between 10°S and 10°N, for the years 2004–2008. The number of retrieved profiles comprising the full data set is listed as a function of month in Table 6.1. The

full data set is made up of 432 retrieved profiles.

Full O_3 data set

All retrieved O_3 profiles from the tropical set are plotted in panel A of Figure 6.5, with the mean of the full set overlaid in blue. The deviations of each profile from this mean are plotted in panel B. The scatter in O_3 deviations is generally less than 2 ppmv, and reaches a maximum between 30 and 40 km. The scatter is quantified by the SD, plotted in absolute and relative (percent of the mean) magnitude in black in panels C and D, respectively. While the absolute SD is maximum between 30 and 40 km, the relative SD is minimal, and roughly constant with altitude between 25 and 60 km, ranging between approximately 5 and 10%. The full tropical O_3 data set, which includes the effects of interannual, seasonal and diurnal natural variability, exhibits a SD which approaches 5% at 28 km, and is less than 10% between 24 and 60 km.

From visual inspection of the raw retrieved profiles or the deviations from the mean, one can see a number of outliers: points or profiles that differ significantly from the majority of the data. Below 20 km, a number of profiles display an outlying value at the lowest retrieved altitude. Above 20 km, a few profiles are consistently far from the rest of the data.

The presence of outliers in a data set poses a challenge for random error validation. On the one hand, unless data can be excluded on independent grounds, such as knowledge of anomalous conditions during the measurement, all data should be used in the precision validation, since the results of the validation should apply to the full data set. On the other hand, commonly used measures of scatter, like SD, are heavily influenced by outliers. A small number of outliers can drastically affect the SD, pulling it away from a value representative of the majority of the data.

Statistics that are relatively insensitive to the presence of outliers in data are known as robust. One of the most robust estimates of scale (or scatter) is the median absolute deviation (MAD) (Huber, 2004). The MAD was first promoted by Hampel (1974), who attributed it to Gauss. The following description is based on that given

by Maronna et al. (2006). Given a data set $\mathbf{x} = (x_1, ..., x_n)$, the MAD is defined as:

$$\text{MAD} = \text{med}\left(|\mathbf{x} - \text{med}(\mathbf{x})|\right), \tag{6.7}$$

where med($\mathbf{x}$) denotes the sample median[1] of $\mathbf{x}$. This estimator uses the sample median twice, first to get an estimate of the center of the data in order to form the set of absolute residuals about the sample median, ($|\mathbf{x} - \text{med}(\mathbf{x})|$), and then to compute the sample median of these absolute residuals.

The MAD represents the interval around the median that contains 50% of the data (Rousseeuw and Croux, 1993). As such, the MAD ignores the values of the 50% of the data that are outside this interval. The "robustness" of the MAD is defined by this property: up to 50% of the data can be composed of extreme outliers, and the MAD will still give a value representative of the scatter of the central 50% of the data distribution.

To make the MAD comparable to the SD, a normalized MAD ("MADN") can be defined as

$$\text{MADN}(\mathbf{x}) = \frac{\text{MAD}(\mathbf{x})}{0.6745}. \tag{6.8}$$

The reason for this definition is that 0.6745 is the MAD of a normally distributed set of random variables with $\sigma = 1$, and hence a normally distributed set with any standard deviation σ has a corresponding MADN = σ. It should be noted however that if the data is not normally distributed, the SD and MADN will not necessarily agree. In general, for data with outliers, the MADN will be less than the SD, due to its robustness. However, for distributions that do not peak at one location, and are for example "flat" (as in a uniform distribution) or "double-lobed", the MADN can exceed the SD.

The MADN for the full tropical ACE-FTS O_3 data set is shown in Figure 6.5

[1]Like the sample mean, the sample median is an estimator of the center of a data set. The median is conveniently defined in terms of the order statistics $(x_{(1)}, x_{(2)}, ..., x_{(n)})$ obtained by sorting the observations $\mathbf{x} = (x_1, ..., x_n)$ in increasing order. If n is odd, then $n = 2m - 1$ for some integer m, and in that case $\text{med}(\mathbf{x}) = x_{(m)}$. If n is even, then $n = 2m$ for some integer m, and it is customary to define $\text{med}(\mathbf{x}) = \frac{x_{(m)} + x_{(m+1)}}{2}$ (Maronna et al., 2006).

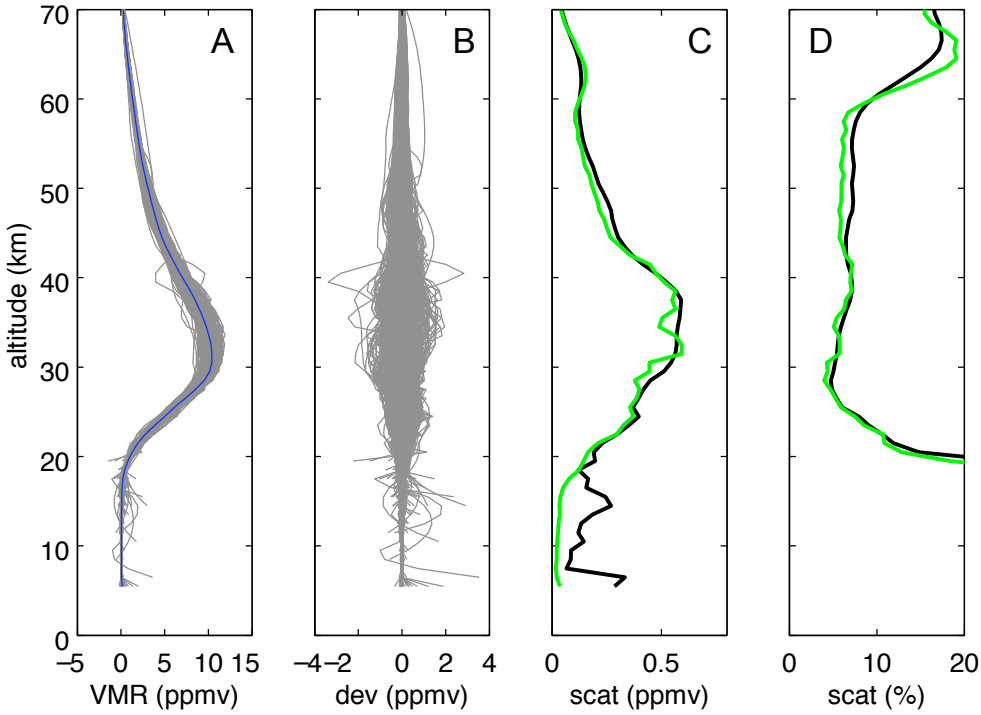

Figure 6.5: O_3 scatter statistics for the full tropical (10°S–10°N) ACE-FTS data set. Panel A: All measured O_3 profiles over years 2004–2008. The mean of all profiles is shown in blue. Panel B: Deviations from the overall mean. Panel C: Absolute scatter statistics, SD (black), and the MADN (green). Panel D: Relative scatter statistics, SD (black, shown as percent of mean) and MADN (green, shown as percent of median).

(green) in terms of absolute and relative (percent of median) magnitudes. The MADN is seen to be comparable in magnitude to the SD, but noticeably less sensitive to the outliers below 20 km, and somewhat less than the SD between 30 and 60 km due to the presence of a small number of outlying profiles. Above 60 km, the MADN is larger than the SD, due to the fact that the O_3 distribution at these altitudes is bimodal due to the effects of diurnal cycling. Throughout the presentation of results in this chapter, both SDs and MADs (or in a few select cases, MADNs) will be shown to help illuminate the behavior of the central portion (MAD), and the outlying values (SD) of the ACE-FTS measurement distributions.

Partitioned subsets

The scatter statistics calculated for the full ACE-FTS O_3 data set as shown in Figure 6.5 obviously contain contributions of interannual variations and the seasonal cycle. In order to minimize the effects of these forms of longer-term variability, it is desirable to calculate the scatter statistics for subsets of the data with small temporal bounds. However, since the statistical significance of any calculated statistics depends on the sample size, one wants to retain a sufficient number of observations within each subset. It is not easy to determine the ideal temporal bounding regime a priori. Instead, in the following, the full ACE-FTS data set will be partitioned based on a selection of temporal bounds, and the scatter statistics will be calculated for each of the resulting subsets. This will allow for an investigation of the effects of temporal partitioning on each of the trace gases analyzed.

The temporal partitions to be used in the following include:

All: data from all months (February, April, August and October) and years (2004–2008), as used in the analysis of O_3 above.

Months/All Years: the full data set partitioned by calendar month. The partitioning thus removes the mean annual cycle of variability from the full data set, but retains any interannual variability for each month. This level of partitioning is thus equivalent to that of the HALOE SD climatology produced by Grooß and Russell III (2005).

Months: data partitioned by year and month, thus removing both interannual and seasonal variability from the full data set. This level of partitioning is thus approximately equivalent to that produced by averaging the monthly SDs of HALOE measurements given by Grooß and Russell III (2005), as used as a climatology of short-term variability in §6.2.

SR/SS: data partitioned by occultation type within each month of each year. Given the tropical sampling pattern of the ACE-FTS occultations, this amounts to temporal partitions of approximately four days in length. Separating sunrise

	n_{tot}	m	n_{tot}/m
All	432	1	432
Month/All Years	432	4	108
Month	432	18	24
SR/SS	378	26	14.5
Day	77	8	9.6

Table 6.2: The effect of temporal partitioning on the sample sizes of the created subsets. At each level of partitioning, the total number of retrievals n_{tot} is partitioned into m subsets. n_{tot}/m then gives the average sample size of each subset. n_{tot} may change since any subsets with less than eight samples are excluded.

occultations from sunsets also removes the effect of any diurnal variability.

Day: data partitioned by year, month, and finally into single days. Since sunrise and sunset occultations do not occur on the same day in the tropics, Day partitioning also removes diurnal variability[2].

At each level of temporal partitioning, the sample sizes of each subset become progressively smaller. Table 6.2 details the relationship between the number of partitioned subsets (m), and the average sample size of each subset (n_{tot}/m). In order that the scatter statistics for each subset retain some level of significance, an arbitrary threshold of eight samples has been used as a minimum for each subset: when there are less than eight samples within a subset, it is excluded from the analysis. This sample size criterion leads to the exclusion of a handful of subsets for the SR/SS partitioning, taking the total number of measurements used from 432 to 378. A more drastic example is seen for Day partitioning, for which 82% of the retrieved profiles are excluded based on the fact that they come from days with less than 8 samples.

Scatter statistics can be calculated for each subset individually, resulting in an ensemble of scatter estimates. Rather than interpret an ensemble of different scatter estimates, a composite measure of the scatter can be computed based on the full set of deviations from the individual partition means (or medians). For example, when

[2]While not true of the global ACE-FTS occultation set, the local solar time (LST, see Equation 6.9) of tropical occultations corresponds with expectations of typical sunrise and sunset times, with sunrises occurring between approximately 5:00 and 6:15, and sunsets between 17:00 and 18:15.

partitioning by Month/All Years (row 1 of Figure 6.6), deviations can be produced based on the difference between each profile and its respective monthly mean (row 1, column 2). The RMS of the full set of deviations can then be used as a composite measure of the overall scatter. The composite RMS of the deviations of each partitioned subset is equivalent to the square root of a weighted mean of the subset variances, with weights equal to the sample size of each subset, relative to the total sample size. Based on this interpretation, and for the sake of simplicity, the RMS of the full set of deviations from the subset means will be referred to in the following as a composite standard deviation, denoted SD*.

The advantage of a composite measure of scatter, such as the SD* described above, is that it reflects the variability using all of the data, and thus avoids the philosophical issue of trying to validate the RREs of an instrument based on the scatter of only an extremely small subset of measurements. The SD* for the Month/All Years partitioned subsets is shown in black in absolute and percent magnitudes in row 1, columns 3 and 4 of Figure 6.6. A composite MADN, or MADN*, is produced analogously by calculating the median absolute value of the deviation of each profile from its respective monthly median, and is also shown (green) in columns 3 and 4 of Figure 6.6.

The process of partitioning the data into subsets based on temporal bounds, and calculating the SD* (and MADN*) based on deviations from the mean (median) of each subset, is performed for each of the partition bounds described above. Figure 6.6 shows four panel plots of calculated O_3 scatter statistics in the the style of Figure 6.5 for each of the levels of partitioning described above (except for All partitioning, shown previously in Figure 6.5), with each row corresponding to one level of partitioning.

The scatter in deviations steadily decreases with tighter temporal partitioning, as seen qualitatively in Figure 6.6 by the scatter in deviations in column 2, or by the quantification of scatter, SD* and MADN*, shown in columns 3 and 4. This reduction of scatter with increasing levels of temporal partitioning is shown more clearly in Figure 6.7. It should be noted that the SD* of the full data set, which includes

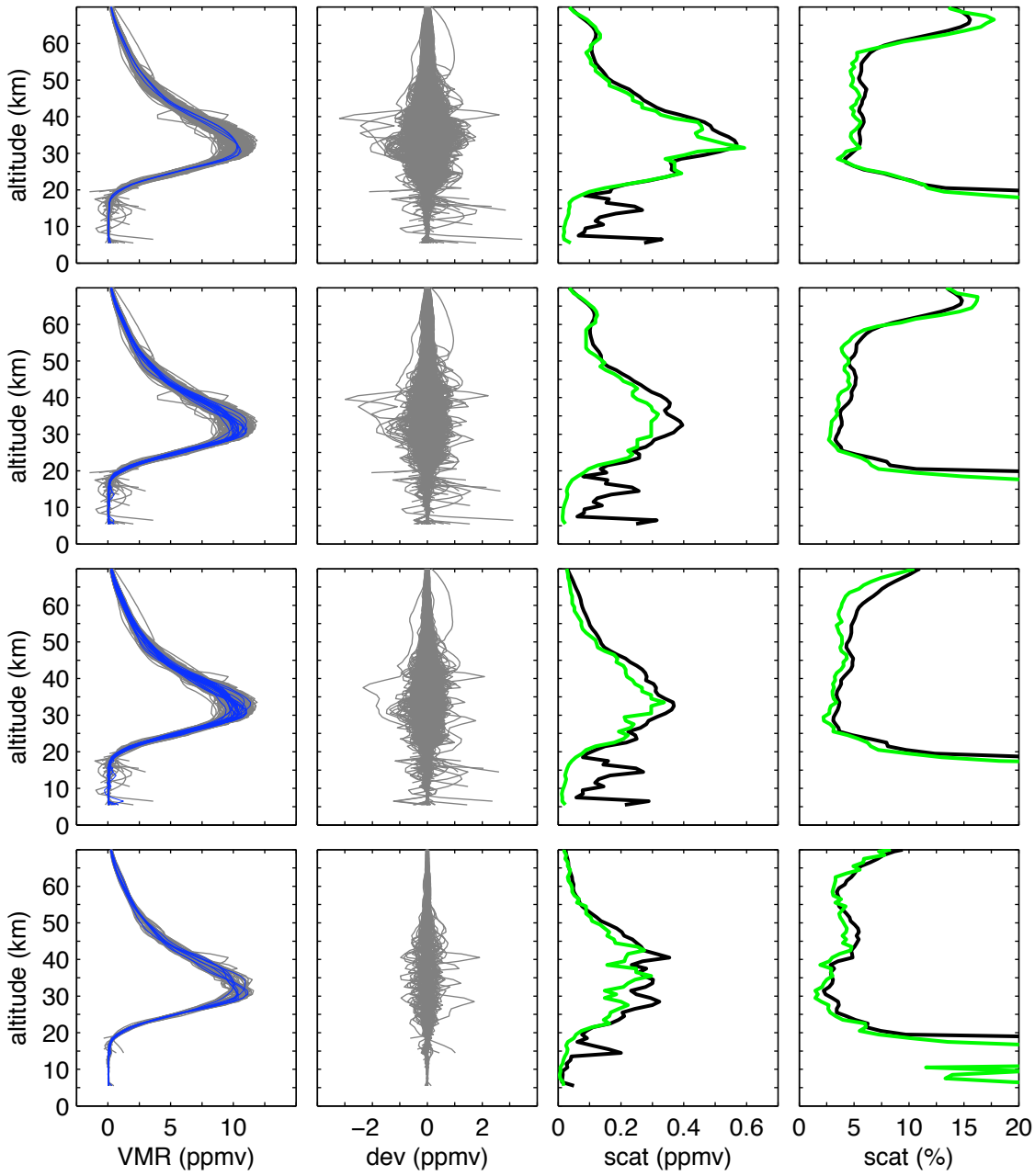

Figure 6.6: Scatter statistics for the tropical ACE-FTS O_3 data set, subject to different levels of temporal partitioning. Row 1: partitioned by Month/All Years, Row 2: partitioned by Month, Row 3: partitioned by SR/SS, Row 4: partitioned by Day. Column 1: raw measured profiles (grey), and means for each subset (blue). Column 2: deviation of each profile from the subset mean. Column 3: composite absolute SD (black) and MADN (green). Column 4: composite percent SD (black) and MADN (green).

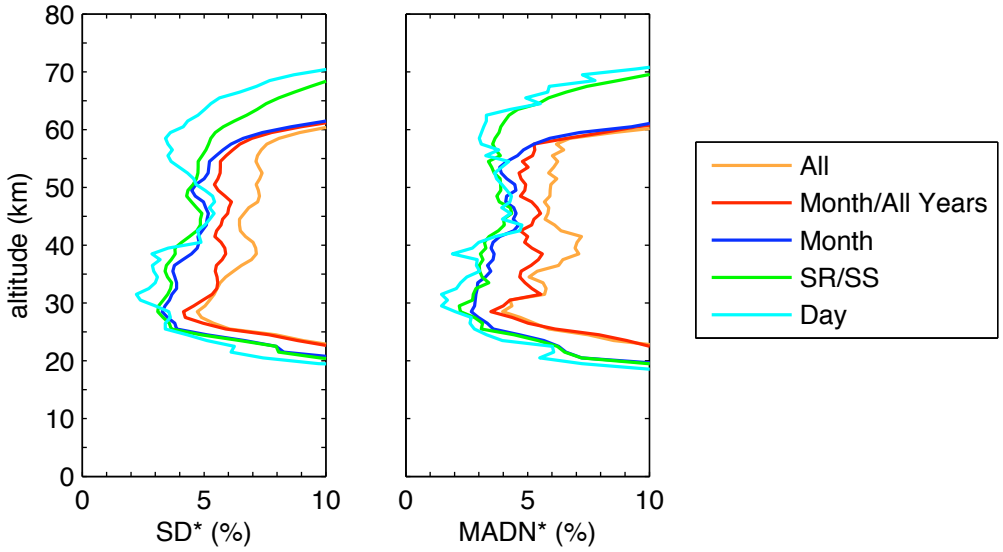

Figure 6.7: Composite SDs (left) and MADNs (right) from different levels of temporal partitioning of the tropical O_3 ACE-FTS measurement set. Scatter statistics shown for All (orange), Month/All Years (red), Month (blue), SR/SS (green) and Day (cyan) partitioning.

variability on time-scales of interannual, seasonal, and short-term, is not drastically different than that based on Day partitioning. All levels of partitioning give SD*s that are below 8% between the altitudes of roughly 25 and 55 km. Partitioning the data, and in so doing removing different scales of temporal variability, reduces the scatter by the order of a few percent. Notable reductions in scatter occur between 25 and 60 km due to the removal of seasonal and interannual variability and above 55 km due to the removal of diurnal variability. The reduction in scatter produced by partitioning by SR/SS subsets, in so doing removing the effect of diurnal variations, is displayed most clearly by the MADN* statistic. It should also be noted that the SR/SS MADN* profile is consistent with that produced by Day partitioning (with a much smaller sample size) for all but a few altitudes. The SR/SS MADN* profile is also quite close to the Month MADN* profile below 55 km. It would appear that there is only a small reduction in scatter produced by moving from Month to tighter levels of partitioning below 55 km, where diurnal variations are insignificant.

Confidence intervals

Confidence intervals for the scatter statistics SD* and MAD* (and MADN*) can be produced through the bootstrapping technique (Efron and Tibshirani, 1994), wherein the entire algorithm used to produce the statistics is repeated on a large number of random resamplings of the full data set. The width of the distribution of resulting estimates can then be used to define a confidence interval for the statistic.

Confidence intervals at the 95% level are shown for the O_3 statistics in Figure 6.8. Each column of Figure 6.8 shows the statistics (SD* row 1, MAD* row 2) at different levels of partitioning. For purposes of comparison, each panel also contains the All partition statistic.

The confidence intervals shown for O_3 are of relatively equal size at any altitude for All, Month/All Years, Month and SR/SS partitioning. For SD*s, the confidence intervals have widths ranging from approximately 1 to 3% between 22 and 60 km for these levels of partitioning. The Day SD* has somewhat larger confidence intervals, ranging in width from 2 to 5% between 22 and 60 km, due to the smaller total sample size resulting from the minimum subset size threshold. Confidence intervals for the MAD* are in general smaller than those for the SD*s, as would be expected given the robust nature of the MAD*, with widths ranging from 0.6 to 1% between 22 and 60 km for all but Day partitioning.

The confidence intervals show that while the Month/All Year SD* is not significantly different from the All SD*, the SD* for the other levels of partitioning are (at vertical levels shown).

6.3.3 Measurement scatter and reported random errors

The scatter in tropical ACE-FTS measurements, quantified by SD* and MAD* statistics, and presented as a function of temporal partitioning as in Figure 6.7, will presently be shown for a number of trace gases measured by ACE-FTS, and compared to the reported random errors for each specific species. The comparisons will each be briefly described, and an overall summary of the comparisons will follow.

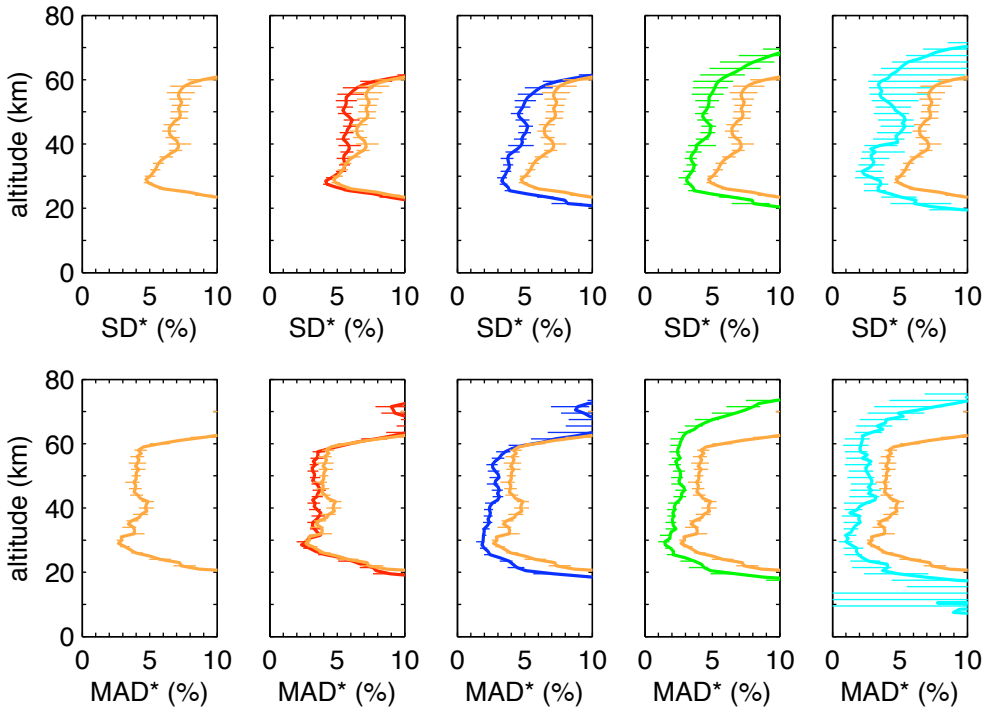

Figure 6.8: Tropical ACE-FTS O_3 scatter statistics SD* (top row) and MAD* (bottom row), shown for All (orange), Month/All Years (red), Month (blue), SR/SS (green) and Day (cyan) partitioning on separate panels, with 95% confidence intervals shown by horizontal lines.

The comparisons will be grouped based on the chemical lifetimes of the trace gas species and transport time scales shown by Brasseur and Solomon (2005). Following the treatment by Brasseur and Solomon (2005), "short-lived" species are defined to be those with lifetimes shorter than the time scale of zonal transport, and "long-lived" will refer to those species with lifetimes comparable to, or longer than, the meridional and vertical transport time scales. "Medium-lived" will be used here to describe those trace gases with lifetimes longer than the zonal transport time scale, but shorter than the vertical and meridional dynamical time scales. Since the lifetime of a trace gas depends on altitude, a species may be, for example, long-lived in the lower stratosphere but medium- or short-lived in the upper stratosphere. As such, each species examined below is a unique case, and the groupings are meant only to collect together species with roughly similar behavior.

Following the discussion of measurement errors in §6.1.1, the SD* of a set of repeated measurements defines the random error of the measurement process, and is equal to the RMS of the individual random errors. The scatter SD*s will therefore be compared to the RMS of the RREs in each of the following plots. Following the same logic, the MAD* of a set of repeated measurements should be equal to the median RRE, and the scatter MAD*s will therefore be compared to the median RREs.

Medium-lived species

O_x (whose lifetime defines an effective lifetime for O_3), HNO_3 and CO are all medium-lived (as defined above) through most of the stratosphere, with lifetimes of the order of a month in the lower stratosphere, decreasing with height. The lifetime of CO remains comparable to the dynamical time scales through the stratosphere, and actually increases with height above ∼40 km, becoming quite long-lived in the mesosphere and above. The lifetimes of O_x and HNO_3, in contrast, decrease with height, and both species should be considered short-lived in the upper stratosphere (above approximately 35-40 km), but are grouped here with CO based on their lifetimes in the middle and lower stratosphere.

O_3 measurement scatter and RREs are compared in Figure 6.9. Many of the features of the effect of temporal partitioning on the O_3 scatter have been discussed previously, such as the reduction in scatter above 50 km produced by partitioning sunrise and sunset measurements, and the smaller values of the MAD* compared to the SD* below 20 km, due to the different influence of outliers on the two measures of scatter. The short-term (SR/SS and Day) SD*s are on the order of 4-5% through most of the stratosphere, while the RREs are 1-2%. This discrepancy between the measured scatter and the RREs is apparent in both the absolute and relative SD* comparison plots.

The comparison of MAD*s with median RREs shows a much closer agreement, with the short-term MAD*s showing good agreement for all altitudes except 40-55 km.

HNO_3 measurement scatter and RREs are compared in Figure 6.10, top row. Note

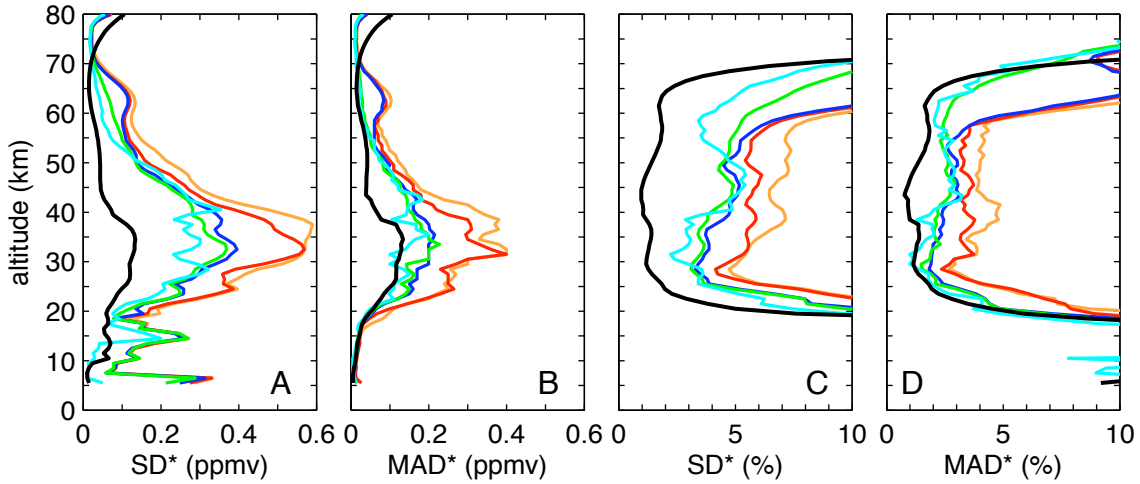

Figure 6.9: ACE-FTS O_3 measurement scatter, as a function of temporal partitioning (thin coloured lines), compared to reported random errors (RREs, thick black line). Colour-coding of temporally partitioned variability is as in Figure 6.7. A: Absolute composite standard deviations (SD*s) compared to the root-mean-square (RMS) RRE. B: Absolute composite median absolute deviations (MAD*s) compared to the median RRE. C: Relative SD*s compared to the RMS relative RRE. D: Relative MAD*s compared to the median relative RRE.

that here, and in many of the plots that follow, the scaling of the x axis is chosen so as to best compare the measurement scatter and RREs at altitudes where the RREs are relatively small, i.e., under 20% or so. For example, we are more concerned with comparing the measurement scatter and RREs for HNO_3 at altitudes above 20 km, where the relative RRE is on the order of 1%, than we are for altitudes below 20 km.

HNO_3 scatter quantified by the SD* suffers from the presence of a large number of outliers below 20 km, resulting in large and often oscillating SD*s in this region. MAD*s are, in contrast, small and well behaved in this region, and show good agreement between measured scatter and median RRE. As seen at high altitudes for O_3, HNO_3 shows a decrease in scatter associated with reducing the temporal partitioning to the SR/SS level, showing the presence of diurnal variability above 27 km. This effect is most readily apparent in the MAD* plots, which show close agreement between the SR/SS and Day partition statistics. Above 20 km, SD*s are significantly different from the RMS RREs, while the MAD*s are in relatively closer agreement

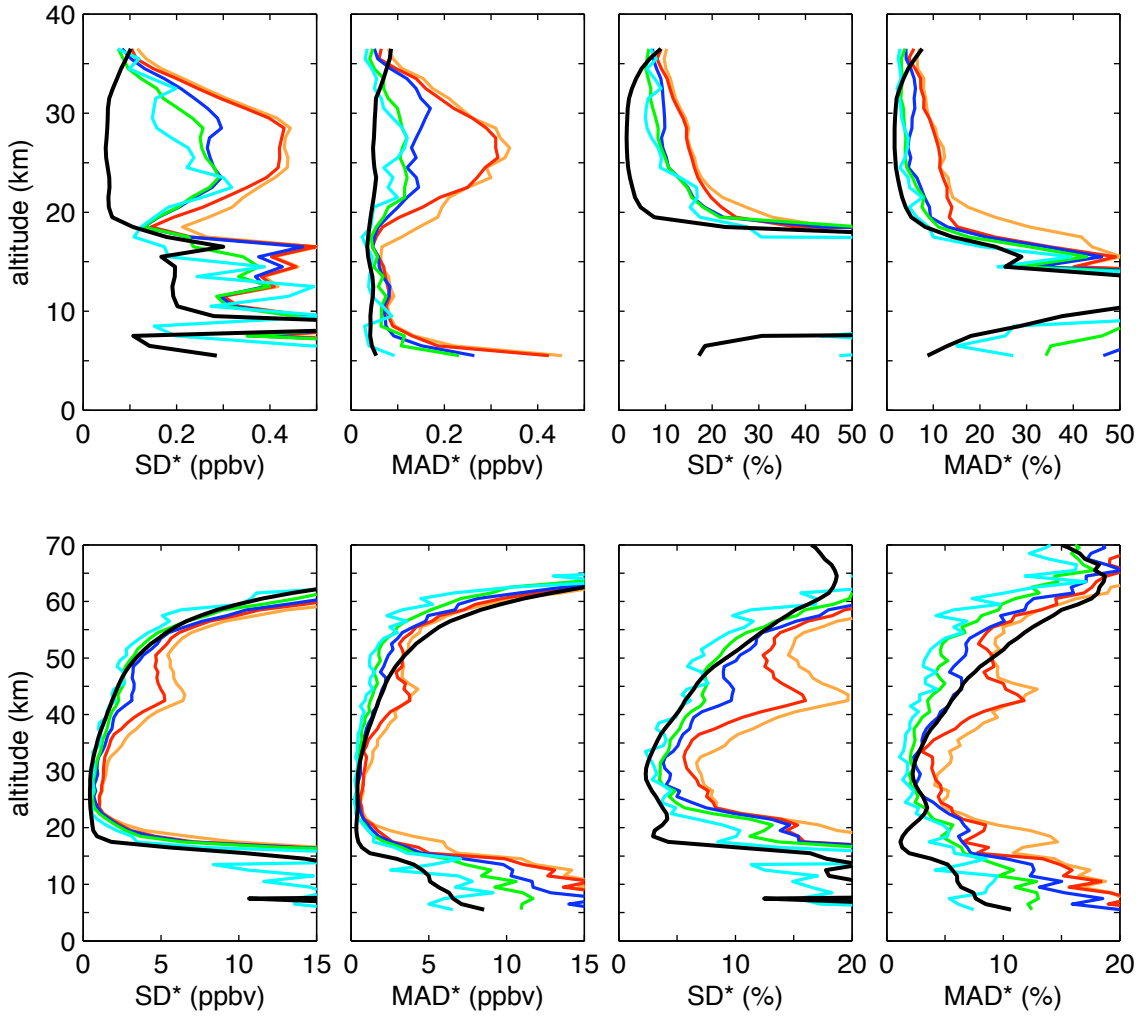

Figure 6.10: As Figure 6.9, for HNO_3 (top) and CO (bottom).

with the median errors.

CO scatter (Figure 6.10, bottom row) is only weakly influenced by temporal partitioning, with the most sizable reductions occurring between 40 and 50 km. Both short-term SD*s and MAD*s show good agreement with the RREs, with the short-term MAD*s actually displaying values less than the RREs between 35 and 60 km.

Short-lived species

The lifetimes of NO and NO_2 are extremely short in the stratosphere, on the order of minutes, and both have strong diurnal cycles. The sum of the two species, NO_x, is, on

the other hand, medium- to long-lived in the stratosphere. We therefore expect the variability of NO and NO_2 to be dominated by diurnal variations, but for the short-term variability of each species due to dynamical variability of NO_x to be potentially important as well.

Scatter statistics and RREs for the short-lived species NO and NO_2 are shown in Figure 6.11. The plots of Figure 6.11 are scaled so as to best display the magnitude of the short-term scatter, and as a result the long-term scatter statistics for NO_2 are larger than the x-axes of the plots. The peak value of the NO_2 All SD* is approximately 1.8 ppbv (or 40%) at 32 km, and the peak value of the All MAD* is approximately 1.7 ppbv (37%). The Month/All Year SD* is approximately equal to the All SD*, while the Month/All Year MAD*, with a peak value of 1.4 ppbv (35%) is only slightly smaller than the All MAD*. In contrast, at the Month level of partitioning, there is a sizeable difference between the SD* and the MAD*, with the SD* showing a value of 1.6 ppbv (38%) and the MAD* reduced down to 0.5 ppbv (20%). As might be expected, temporal partitioning to levels SR/SS and Day greatly reduces the measurement scatter for both NO and NO_2 - down to levels less than 10% - due to the elimination of the effects of diurnal cycling.

The short-term scatter statistics for NO are in good agreement with the RREs, with the MAD* showing particularly close agreement with the median RRE. Above 50 km, the short-term MAD*s for NO are less than the median RRE. Results for NO_2 show slightly worse agreement, especially for the SD* comparisons.

Long-lived species

The species N_2O, CH_4, H_2O and the CFC species CCl_2F_2 and CCl_3F are long-lived, with lifetimes comparable to or longer than the dynamical time scales over the vertical range of ACE-FTS measurements shown here.

N_2O (Figure 6.12, top row) shows a large degree of scatter in measurements and errors at altitudes below 20 km, as shown by the large SD*s in this altitude range, which exceed the horizontal scale of the absolute SD* plot as shown. Examination of the relative statistics shows the SD*s (for all but Day partitioning) are much

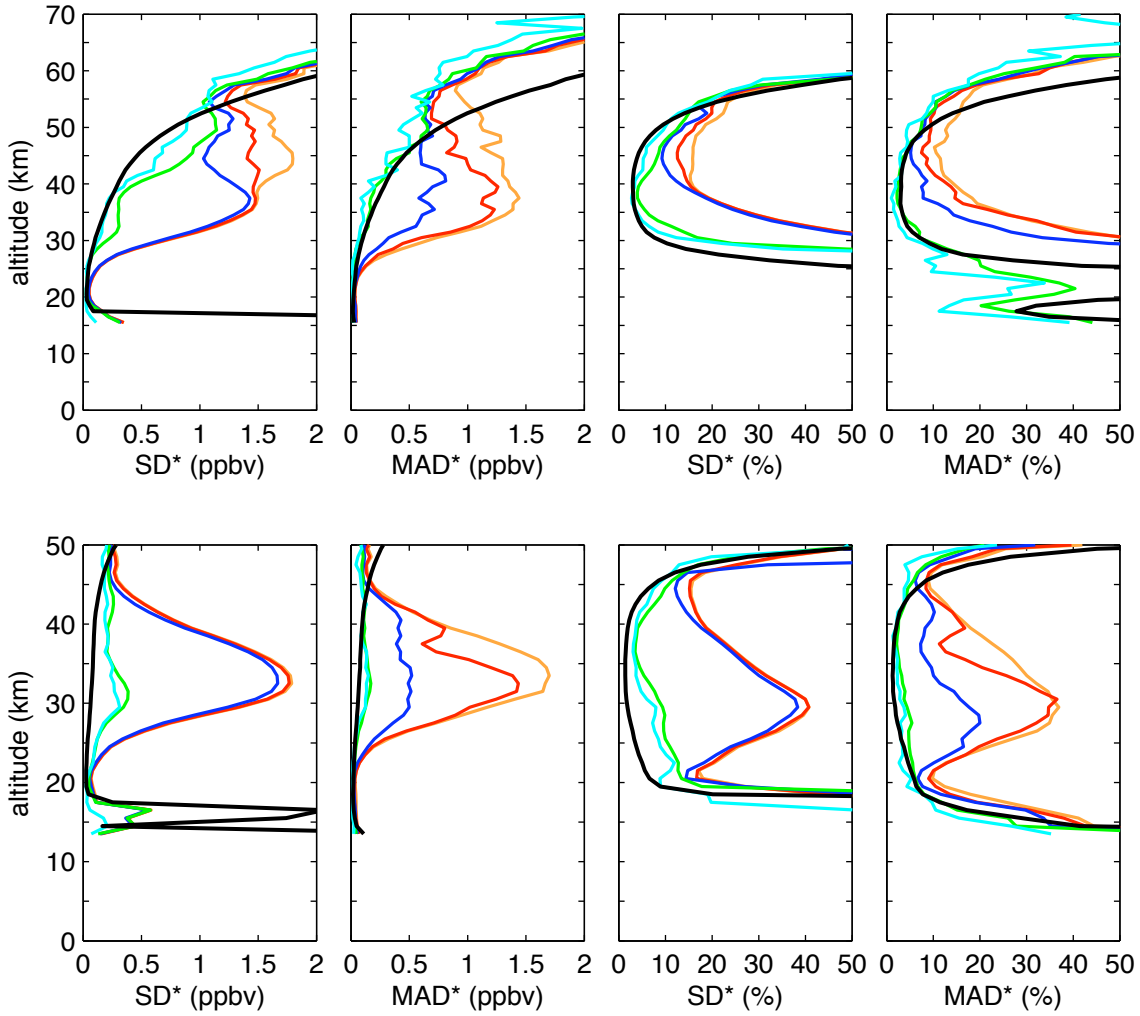

Figure 6.11: As Figure 6.9, for NO (top) and NO_2 (bottom).

larger than the RREs at these low altitudes. The effect of this scatter is removed in the MAD*s, which below 30 km are relatively constant in percent magnitude with altitude, and show reasonable agreement with the median RRE. Between 20 and 30 km, the relative SD*s (for all levels of partitioning) are approximately 5%. The RRE in this altitude region is approximately 1%: obviously there is a discrepancy between the RRE and the SD*s (to be explored in the following section). Above 30 km, the absolute SD*s and MAD*s are seen to be larger than the RREs, with a significant spread between the different levels of partitioning. Day statistics show the smallest magnitudes, and show a positive offset of ∼5 ppbv (2 ppbv) compared to the RRE

for the SD* (MAD*) comparison between 20 and 40 km. Because the VMR of N_2O becomes quite small in the middle to upper stratosphere, the discrepancies between the scatter statistics and the RRE are amplified in the plots of the relative statistics.

The comparisons for CH_4 (Figure 6.12, bottom row) show a very similar structure to those for N_2O, as might be expected for two species with similarly long lifetimes, both transported into the stratosphere from surface sources. CH_4 shows excellent agreement between MAD*s and RREs below 30 km. Above 30 km, as was seen for N_2O, there is a small offset between the measurement scatter at Day partitioning and the RREs, which is amplified by the small CH_4 VMR in the relative scatter comparisons.

H_2O short-term SD*s (Figure 6.13) show excellent agreement with the RREs. Day SD*s, in particular, show excellent agreement with the RRE in the lower stratosphere. Above 55 km, the SD*s (for SR/SS and Day partitioning) are significantly smaller than the RREs. The H_2O Day, SR/SS and Month MAD*s are all significantly smaller than the median RRE from 20-70 km.

The scatter in measurement and errors is extremely large for CCl_2F_2 and CCl_3F as measured by the SD* (Figure 6.14). These species are present at very low mixing ratios (pptv), and presumably the spectral fitting used in their retrieval is more sensitive to random errors than for other species, leading to a large number of outliers. Despite this, the MAD*s are generally close in magnitude to the median RREs, especially when one excludes from consideration the top and bottom of the retrieval vertical range. For instance, differences between the SR/SS MAD* and the median RRE are less than 1.4% for $CClF_2$ between 10 and 25 km, and are less than 2% for CCl_3F between 10 and 20 km. The statistical significance of these differences will be addressed in the following section.

Discussion

From inspection of the comparisons of scatter statistics for different temporal partitioning, a few general patterns emerge.

We will firstly examine the general results of temporal partitioning. For all but

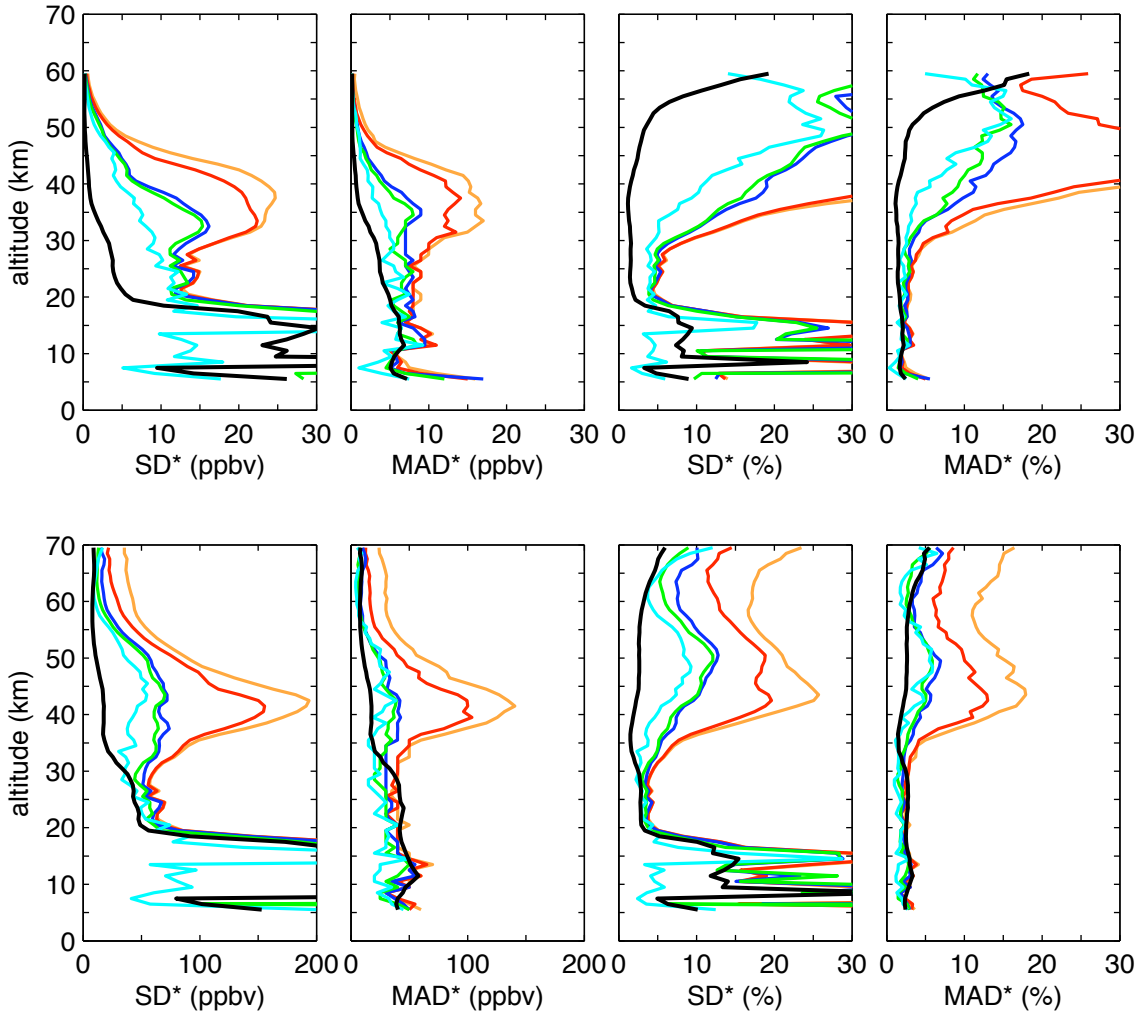

Figure 6.12: As Figure 6.9, for N_2O (top) and CH_4 (bottom).

the short-lived species NO and NO_2, the reduction in scatter produced by changing from All to Month/All Years is smaller than that for moving from Month/All Years to Month. In other words, for the tropical observations explored here, removing the seasonal cycle from this tropical data set is much less important than removing interannual variability in terms of reducing the scatter. The change from Month to SR/SS partitioning was, not surprisingly, important for removing variability tied to diurnal variations, which was important for the short-lived species NO and NO_2, for O_3 above 50 km, and for HNO_3 above 27 km.

SR/SS and Day statistics are often in close agreement with each other, as would

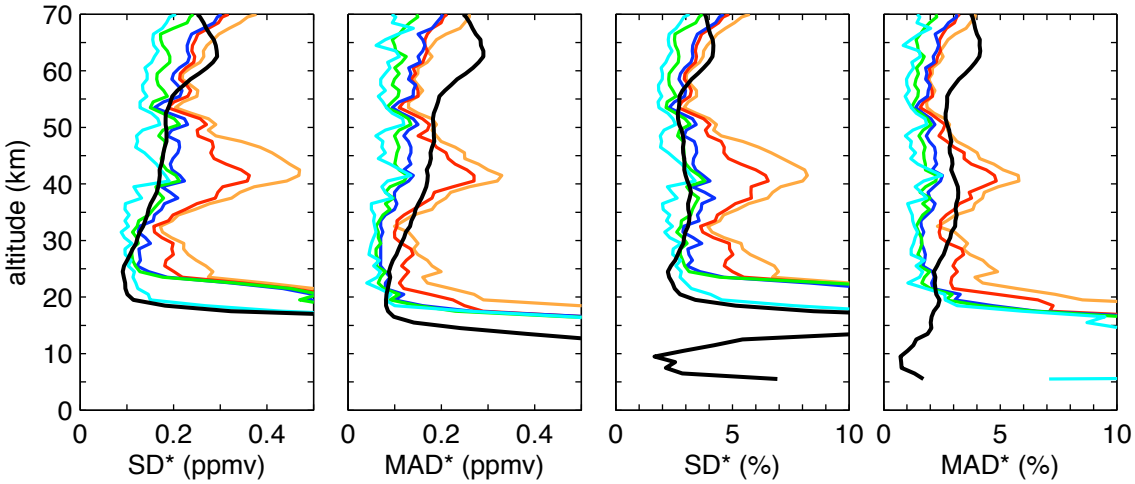

Figure 6.13: As Figure 6.9, for H_2O.

be expected, as the variability present over four days should not be much more than in a single day. However, there are cases where the Day statistics are significantly smaller than the SR/SS, e.g. for N_2O between 30 and 40 km, and to a lesser degree for CH_4 between 30 and 50 km (see Figure 6.12). Since the overall sample size of the Day partitioned set is much reduced from the full data set due to the sample size minimum threshold, it may be that Day partitioning excludes episodes of N_2O and CH_4 variability associated with surf-zone activity intruding into the tropics (as discussed in §6.2).

The short-term partitioned statistics, SR/SS and Day, show the smallest values, and hence generally the closest agreement with the RREs. The SR/SS and Day statistics are for the most part not significantly different from each other (i.e., their 95% confidence intervals overlap). Therefore, in order to summarize the comparisons of scatter and RRE, SR/SS statistics will be used as a measure of short-term scatter.

Agreement between SR/SS scatter statistics and RRE is examined in Figure 6.15, with each panel showing the difference between scatter and RRE (i.e., SD* - RMS(RRE) and MAD* - med(RRE)) for the ten trace gas species examined above. Confidence intervals at the 95% level, based on the confidence intervals calculated for the scatter statistics, are shown as horizontal lines.

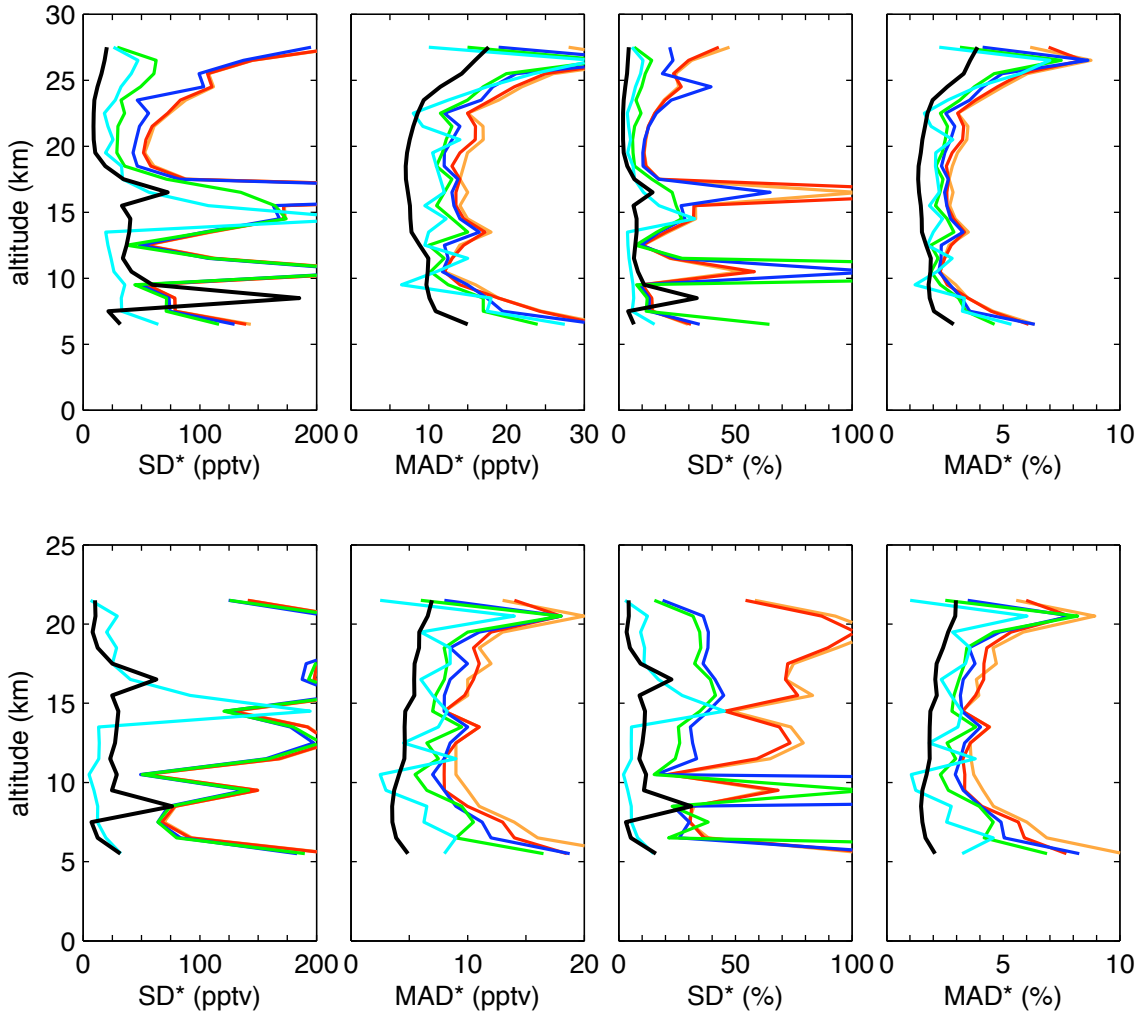

Figure 6.14: As Figure 6.9, for CCl_2F_2 (top) and CCl_3F (bottom).

Three cases present themselves in the comparisons of scatter with RREs. In "Case 1", differences between scatter and RREs are less than or approximately equal to zero (within error bars) for both the SD* and MAD* comparisons. Such is the case for H_2O at altitudes above 25 km, and for CH_4 between 20 and 30 km. For CO, differences between scatter and RREs above 25 km are less than 1%. For these species and altitude regions, this comparison amounts to a successful validation of the ACE-FTS RREs, in a manner similar to previous validation studies. The random measurement error can be no larger than the short-term variability observed, thus it can be concluded that the measurement error for these species and regions is equal

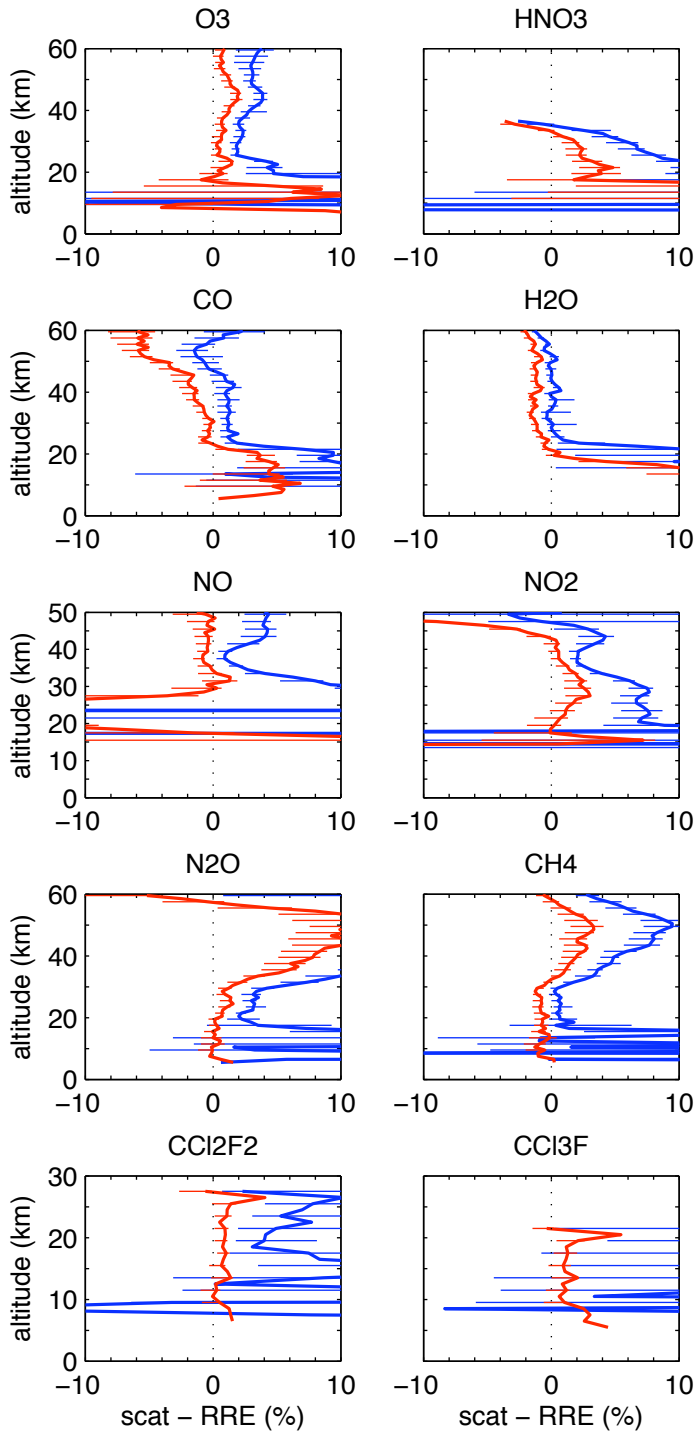

Figure 6.15: Differences between percent short-term scatter measured by ACE-FTS, and the percent RREs. Each panel shows the SR/SS partition SD* - RMS(RRE) (blue) and MAD* - med(RRE) (red). Confidence intervals, based on the confidence intervals calculated for the scatter statistics through bootstrapping, are shown as horizontal bars every 2 km.

to, or less than the RREs.

The second comparison result, "Case 2", describes instances where differences between the MAD* and median RRE are equivalent to or less than zero, while the SD* is significantly larger than the RMS RRE. Such is the case for O_3 between 20 and 40 km, NO between 30 and 50 km, NO_2 in two bands centered at 20 and 40 km, N_2O below 30 km, CH_4 below 20 km, and for most of the full range of measurements for the CFC species. Since the MAD* ignores the outer 50% of the input data, agreement between the data MAD*s and the median RREs shows that the width of the central 50% of the deviation distribution is in good agreement with the RREs. In other words, the RREs are representative of the scatter in 50% or more of the data. The larger differences between the SD*s and the RMS RREs are then understood to be a result of outliers in the data set, whose deviations are larger than their respective RREs.

The third and final comparison result, "Case 3", describes instances where both the SD* and MAD* are significantly larger than the RREs. Such is the case for O_3 between 40 and 50 km, for HNO_3 over all but the very highest retrieved altitudes, for CO below 25 km, for NO_2 between 25 and 35 km, for both N_2O and CH_4 above 30 km, and for H_2O below 25 km.

Discrepancies between scatter and RREs, in the form of either Case 2 or Case 3 described above, may be produced by some combination of (1) true random errors being larger than the RREs, or (2) short-term natural variability present in the measurement set. The present analysis has focussed on the variability of tropical measurements, since other measurements (e.g., those from HALOE) have shown the tropics to be a region of weak natural variability. However, the assumption that it is insignificant with respect to the ACE-FTS RREs has not to this point been directly addressed. It is impossible to address this issue with any currently available observational data set, since all observations are subject to random error, with most of them presumedly larger than the random errors of ACE-FTS measurements. Instead, the variability of chemical fields simulated in a chemistry-climate model will be used in the next section in order to investigate the possible amount of tropical short-term

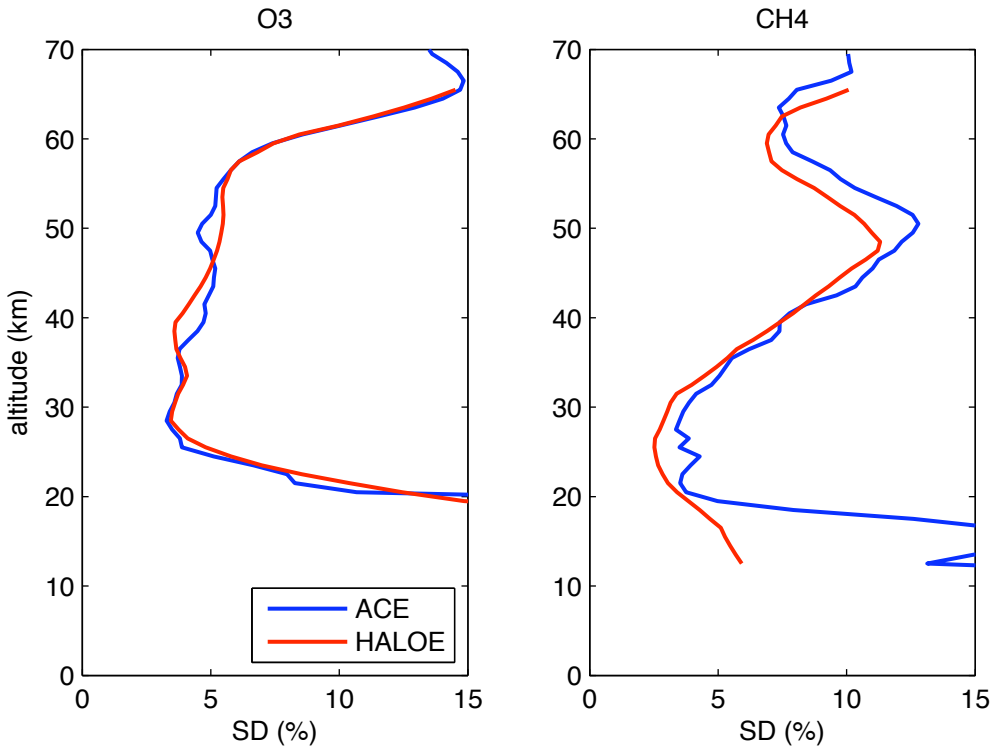

Figure 6.16: Variability in tropical ACE-FTS O_3 and CH_4 retrievals, compared to that from the HALOE climatology. For ACE-FTS, the profile shown represents the composite SD partitioned by month, while for HALOE the profile shows the mean monthly SD for the four calendar months of ACE-FTS tropical coverage.

variability.

Before doing so, however, it is instructive to compare the tropical scatter observed by ACE-FTS to that reported in the HALOE climatology. In §6.2, it was postulated that the minimum SD values of HALOE measurements in the tropics may have been representative of the random error in those measurements, and that the true natural variability may have been less. If so, and if the random errors of ACE-FTS measurements are smaller than those for HALOE (as reported), one would expect the scatter in ACE-FTS measurements to be smaller than that from HALOE. Figure 6.16 compares O_3 and CH_4 SD*s from ACE-FTS measurements, partitioned by Month, to the average SD profile from the HALOE time series of monthly SDs for the four months of ACE-FTS tropical coverage. There is excellent agreement between the scatter

statistics for both instruments. This close agreement supports the idea that natural variability, not random measurement error, is the dominant source of the scatter in both these measurements.

6.4 Tropical variability in the CMAM

It is impossible to determine from a single observational data set whether the scatter in the data is due to measurement error, or the real variability of the measured field. In order to address this question, one might compare the measured scatter with that expected based on an understanding of the measured field. CCMs represent working hypotheses of the interacting chemical and dynamical processes of the middle atmosphere. CCMs include representations, or parameterizations, of many of the processes leading to variability. Most studies focussing on the validation of CCM chemical fields have compared measured and modeled time-averaged fields, or variability on long-term, interannual, and seasonal time scales (e.g. Eyring et al., 2006; WMO, 2007). On short time scales, some studies comparing, e.g., monthly zonal SD fields of O_3 (de Grandpré et al., 2000) have found reasonable agreement between the short-term variability of CCM simulated fields and observations.

In this section, the scatter in chemical fields simulated by the Canadian Middle Atmosphere Model (CMAM) is diagnosed in the same manner as that used for the ACE-FTS measurements in the previous section. By making the assumption that the CMAM accurately reproduces, in a statistical sense, the variability of the true atmosphere, the variability of CMAM chemical fields can be used as a lower bound on the variability of any measurement set, itself subject also to random measurement error.

Chemical fields from ten years of a transient CMAM simulation of the 20th century (described in Appendix A) are used here. The chemical fields are available on every model gridpoint, in save intervals of 18 hours. The chemical fields are reported on pressure surfaces, and are interpolated onto the ACE-FTS 1-km altitude grid using the associated CMAM geopotential field.

In order to best reproduce the sampling of the tropical ACE-FTS measurements, the CMAM data is sampled based on the local solar time (LST), defined as:

$$\text{LST} = \text{UTC} + 24\left(\frac{\lambda}{360^\circ}\right) \tag{6.9}$$

where UTC represents the universal (Greenwich Mean) time, and λ the longitude. In the tropics, the LSTs of ACE-FTS occultations vary slowly with season, with sunrises occurring between approximately 5:00 and 6:15, and sunsets between 17:00 and 18:15. In the following analysis, only CMAM fields with LSTs within the bounds corresponding to ACE-FTS sunrises and sunsets are used. Furthermore, only the months February, April, August, and October, corresponding to the months of ACE-FTS tropical coverage, are used. With these constraints, and taking CMAM results for the four latitudes within the 10°S–10°N latitude bin used in the ACE-FTS analysis, the total CMAM sample size $n_{cmam} = 36800$.

6.4.1 CMAM variability in partitioned subsets

CMAM chemical fields are temporally partitioned as in the analysis of ACE-FTS measurements above, in order to explore the effect of temporal variability of different time scales on the overall variability of the simulated fields. The only difference between partitioning the CMAM and ACE-FTS fields is that CMAM SR/SS partitions are constrained to a length of four days, in order to match the length of ACE-FTS SR/SS partitions[3]. Table 6.3 details the relationship between the number of partitioned subsets (m), and the average sample size of each subset (n_{cmam}/m) for the CMAM analysis.

The scatter statistics for CMAM O_3 are shown in Figure 6.17. As in the ACE-FTS results shown in Figure 6.9, the scatter statistics steadily decrease with increasing temporal partitioning, with the most significant reductions occurring at high altitudes

[3]While there is only one sunrise and one sunset measurement period per month of ACE-FTS tropical coverage, partitioning a full month of CMAM fields into four-day periods results in seven sunrise and seven sunset periods.

	n_{cmam}	m	n_{cmam}/m
All	36800	1	36800
Month/All Years	36800	4	9200
Month	36800	40	920
SR/SS	36452	560	56.1
Day	35860	2252	15.9

Table 6.3: The effect of temporal partitioning on the sample sizes of the CMAM subsets. At each level of partitioning, the total number of retrievals n_{cmam} is partitioned into m subsets. n_{cmam}/m then gives the average sample size of each subset. n_{cmam} may change since any subsets with less than eight samples are excluded.

due to the elimination of the effects of diurnal variability. The strength of the diurnal variability above 55 km in the CMAM fields is much stronger than in the ACE-FTS measurements, due to differences in sampling. Whereas ACE-FTS samples the diurnally varying O_3 at exactly sunrise and sunset, i.e., at solar zenith angles (SZAs) of 90°, the CMAM has been sampled based on the window of local solar times determined by the ACE-FTS measurements. Therefore, in terms of SZA, the CMAM O_3 values are sampled from a finite window which includes 90°, and since O_3 VMRs change very rapidly as a function of SZA in the mesosphere (Allen et al., 1984), the CMAM fields display a larger variance. The same sampling issue may be behind the large increase in short-term variability seen in the CMAM O_3 above approximately 65 km, although other sources of short-term mesospheric variability may also contribute.

In contrast to the ACE-FTS results, the transition from All (orange lines) to Month/All Year (red) partitioning results in a much larger reduction in scatter than the transition from Month/All Year to Month (blue) and SR/SS (green), at least below 50 km. This suggests quite strongly that interannual variability is much weaker in the CMAM than in the true atmosphere, and is very likely due to the fact that this version of the CMAM has no QBO, which drives much of the interannual variability in the real tropical stratosphere (see §3.2.2).

Scatter statistics for more CMAM chemical fields are shown in Figure 6.18, focussing specifically on those species that showed discrepancies between measured scatter and reported random errors in §6.3.3. CMAM HNO_3 (top row) variability

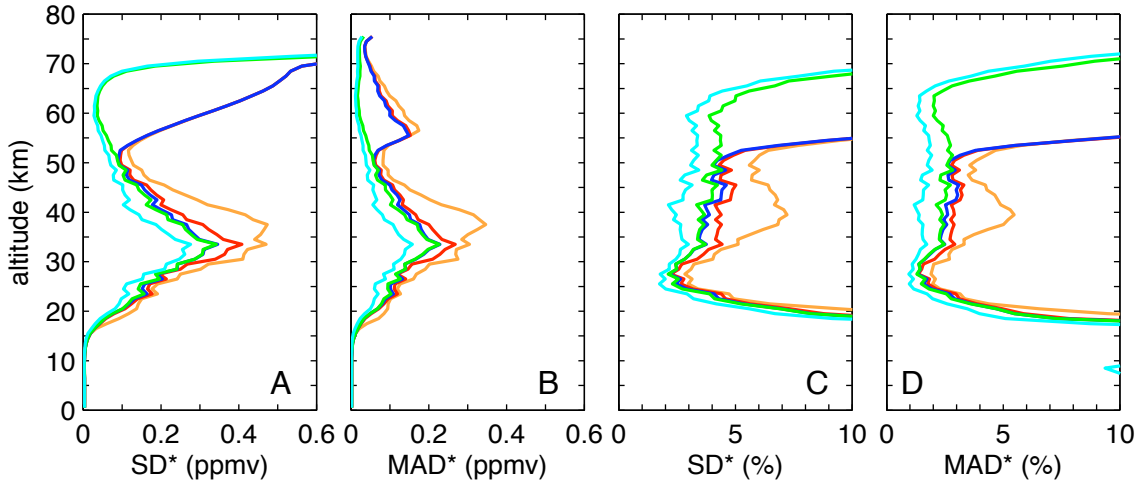

Figure 6.17: CMAM O_3 tropical scatter, as a function of temporal partitioning. Colour-coding of temporally partitioned variability is as in Figure 6.7. A: Absolute composite standard deviations (SD*s). B: Absolute composite median absolute deviations (MAD*s). C: Relative SD*s. D: Relative MAD*s.

peaks between 20 and 35 km, and is minimal at low altitudes, where CMAM VMRs are held fixed (although the relative variability is quite large due to the very low VMR of HNO_3 in the lowermost stratosphere and troposphere). As was seen for O_3, the effect of interannual variability (as shown by the difference between the Month[4] and Month/All Years scatter statistics), is much smaller in CMAM than was seen in the ACE-FTS observations. There is no evidence of diurnal variability in CMAM HNO_3, as was observed in the ACE-FTS observations above 27 km.

CMAM N_2O variability (Figure 6.18, middle row) is seen to peak between 30 and 40 km. Scatter is generally reduced by increasing temporal partitioning of the data. The non-zero variability of the All and Month/All Years statistics in the troposphere and lower stratosphere is due to a trend in N_2O surface emissions included in the CMAM run.

The scatter statistics for CMAM CH_4 are shown in Figure 6.18 (bottom), and show very similar behavior to that of N_2O.

[4]The Month scatter statistics (in blue) for HNO_3, and in fact for all species plotted in Figure 6.18, are in such close agreement with the SR/SS statistics (in green) that the Month statistics are often hidden by the SR/SS statistics.

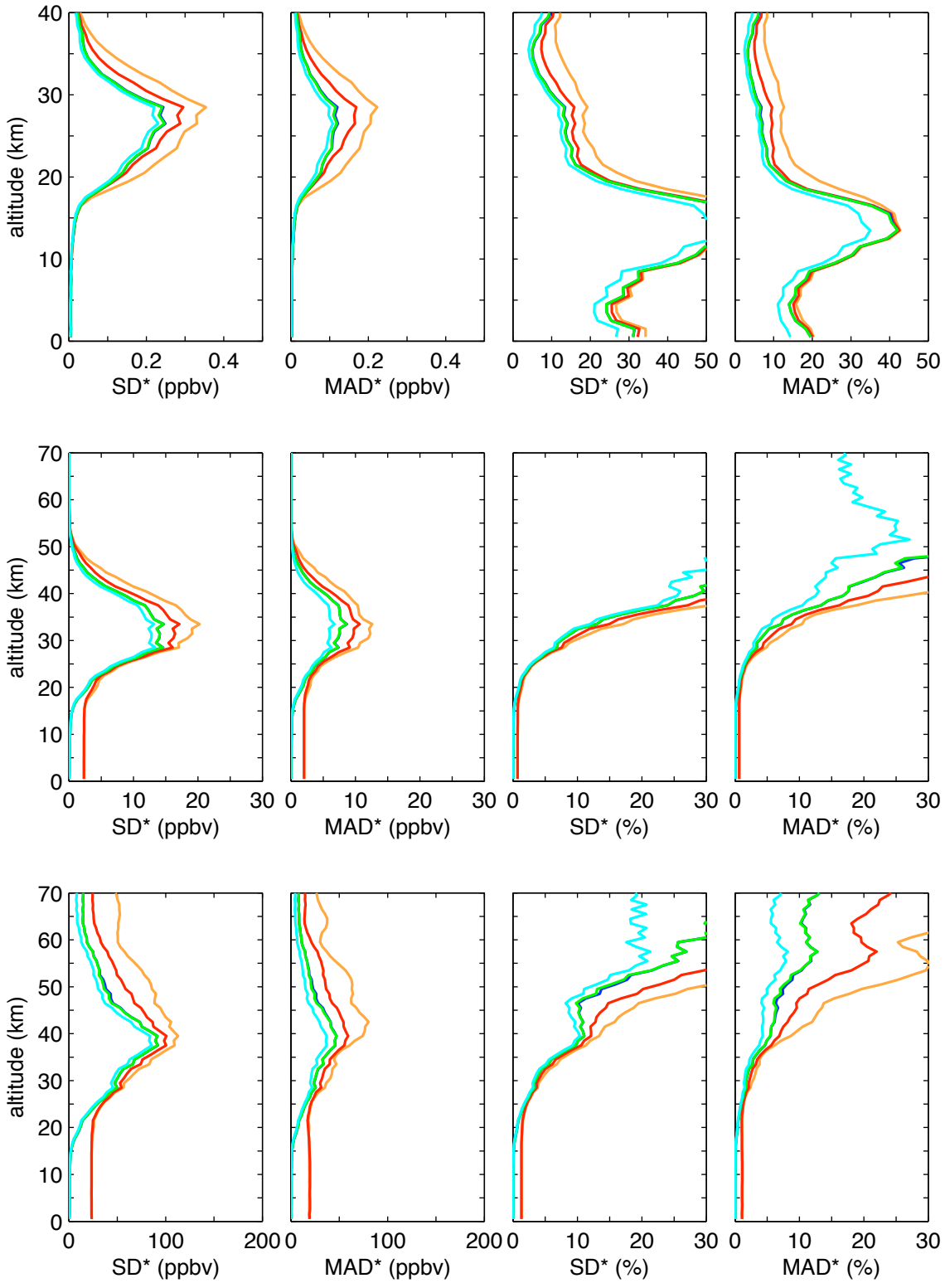

Figure 6.18: As Figure 6.17 for HNO_3 (top), N_2O (middle) and CH_4 (bottom).

6.4.2 CMAM variability at ACE-FTS sampling density

The effect of the sparse sampling density of the ACE-FTS measurements on the scatter statistics was quantified in §6.3.2 by calculating 95% confidence intervals for the statistics. Another way of understanding the effect of sparse sampling on the statistics can be achieved through analysis of model results. Specifically, by subsampling model fields at the density of the ACE-FTS measurements, and recalculating the scatter statistics, differences between the statistics calculated for the full set (as shown in the previous section) and the subsampled set should be comparable to differences between the ACE-FTS statistics and the true atmospheric variability for any temporal partitioning.

A subsampled CMAM set is produced by finding the closest CMAM space-time gridpoint for each individual ACE-FTS measured profile. Matches are found in terms of latitude, day of year, hour, and LST; i.e., for each ACE-FTS measurement, we find the closest CMAM latitude and time, and choose the longitude which gives the closest possible LST. For the five years of ACE-FTS data (2004–2008, inclusive) we use the first five years of CMAM data from the full 10-year CMAM set. With one CMAM profile for each ACE-FTS measured profile, the statistics of the subsampled CMAM data set are equivalent to those tabulated for ACE-FTS in Table 6.2.

Scatter statistics for the subsampled CMAM data set, subject to temporal partitioning, are shown in Figure 6.19 for O_3 and N_2O. The dominant difference between these scatter statistics and those for the full CMAM data set (cf., Figure 6.17 and Figure 6.18, middle panel) is the smaller values of the Day statistics (cyan lines). These results are consistent with the idea that the Day statistics grossly undersample the full variability present, based on the fact that the individual partitions are small in sample size, and many are thrown out since their sample sizes are too small. Thus, the small values of Day scatter seen for the ACE-FTS N_2O measurements (Figure 6.12) can be understood to be a product of this sampling issue.

Aside from the small values of the Day statistics, the other partition statistics are comparable in magnitude to those for the full CMAM sampling. This is seen for O_3

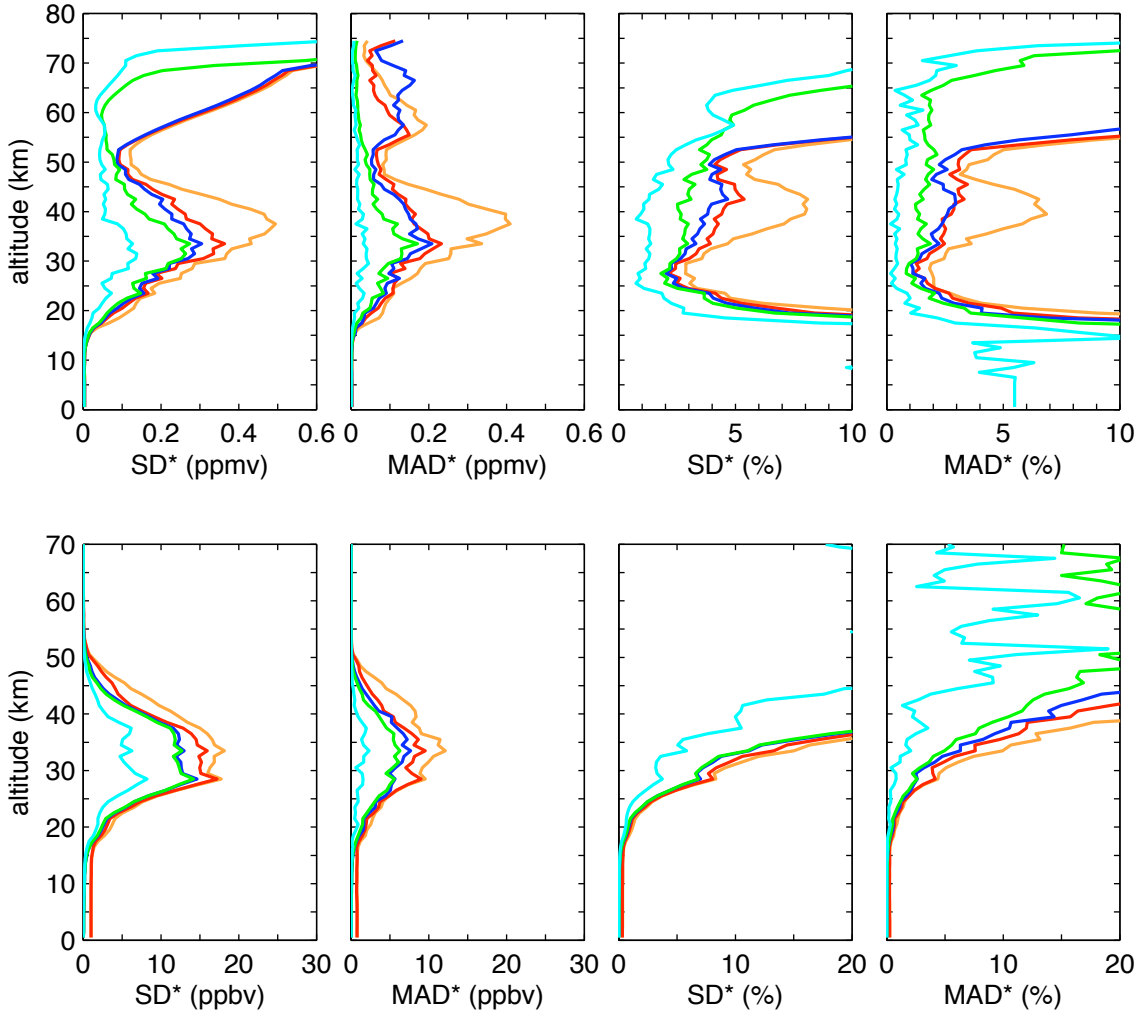

Figure 6.19: As Figure 6.17 but for CMAM sampled at ACE-FTS density, for O_3 (top), N_2O (bottom).

and N_2O shown in Figure 6.19, as well as other species (not shown) including CH_4 and HNO_3. This improves our confidence that the ACE-FTS statistics, at SR/SS and longer levels of partitioning, are representative of the true variability of the tropical stratosphere.

6.4.3 Comparing trace gas variability from CMAM with ACE-FTS measurement scatter

In this section, the variability of tropical chemical species simulated by the CMAM and estimated using the full CMAM data set will be compared to the scatter measured by ACE-FTS, and the reported random errors of the ACE-FTS measurements, focussing specifically on those species that showed discrepancies between measured scatter and reported random errors in §6.3.3. Modeled and measured scatter will be compared at the level of SR/SS partitioning, since this level of partitioning is the shortest time span for which the confidence intervals of the measurement scatter remain small.

Scatter statistics for tropical O_3 modeled by CMAM and measured by ACE-FTS are shown in Figure 6.20, along with the RMS RRE of the ACE-FTS measurements. Both the CMAM O_3 variability and the ACE-FTS measurement scatter are larger than the ACE-FTS RMS RRE in panel A throughout the full vertical range shown. The CMAM variability is comparable to the ACE-FTS measured scatter in vertical shape, with a maximum in absolute variability (panel A) at approximately 35 km, and percent variability (panel C) that increases relatively linearly between 30 and 60 km. There is very good agreement between the CMAM and ACE-FTS MAD*s, especially in terms of percent.

Comparisons of scatter statistics for modeled and measured HNO_3 are shown in Figure 6.21 (top). The CMAM variability between 20 and 35 km matches extremely well with the scatter measured by ACE-FTS, especially in terms of the MAD* statistic. Below 20 km, where CMAM variability is small, the ACE-FTS statistics are comparable in magnitude to the RRE. In terms of percent, the CMAM variability is

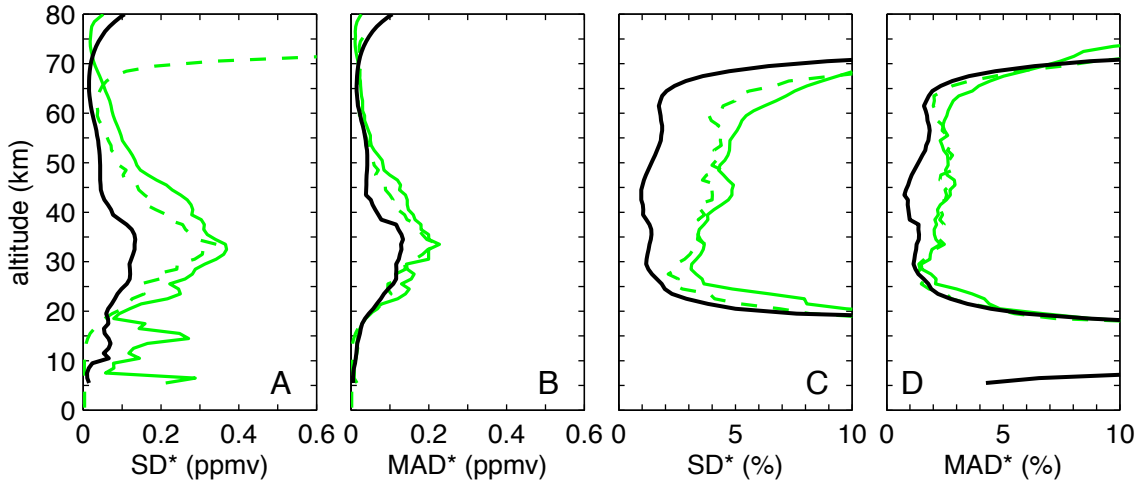

Figure 6.20: Tropical variability in CMAM O_3 (green dashed lines), compared to the scatter measured by ACE-FTS (green solid), both for SR/SS partitioning, and the ACE-FTS RREs (black) for absolute SD* (A), absolute MAD* (B), percent SD* (C) and percent MAD* (D).

larger than the ACE-FTS scatter: this is a result of a low bias in CMAM HNO_3.

Comparisons of scatter statistics for modeled and measured N_2O are shown in Figure 6.21 (middle). Again, as shown most clearly in the absolute magnitude comparisons, the variability of the CMAM simulated N_2O agrees very well with the scatter in ACE-FTS measurements between approximately 25 and 50 km. Furthermore, as was seen for HNO_3, below 25 km where the CMAM variability decreases to zero, the ACE-FTS scatter statistics show excellent agreement with the RREs. The results for the CH_4 comparison (Figure 6.21, bottom) are similar.

In summary, the variability in tropical chemical fields simulated by the CMAM is generally in good agreement with the measured scatter in tropical ACE-FTS measurements, especially at altitudes above 25 km. Based on the CMAM variability results, it would appear likely that differences between the short-term scatter in ACE-FTS measurements and the RREs, shown in Figures 6.9 through 6.14, are due to the presence of real short-term natural variability in the tropical middle atmosphere.

For sake of contrast, we turn now to examine the variability of CMAM H_2O, which showed excellent agreement between ACE-FTS scatter and RRE above 25 km.

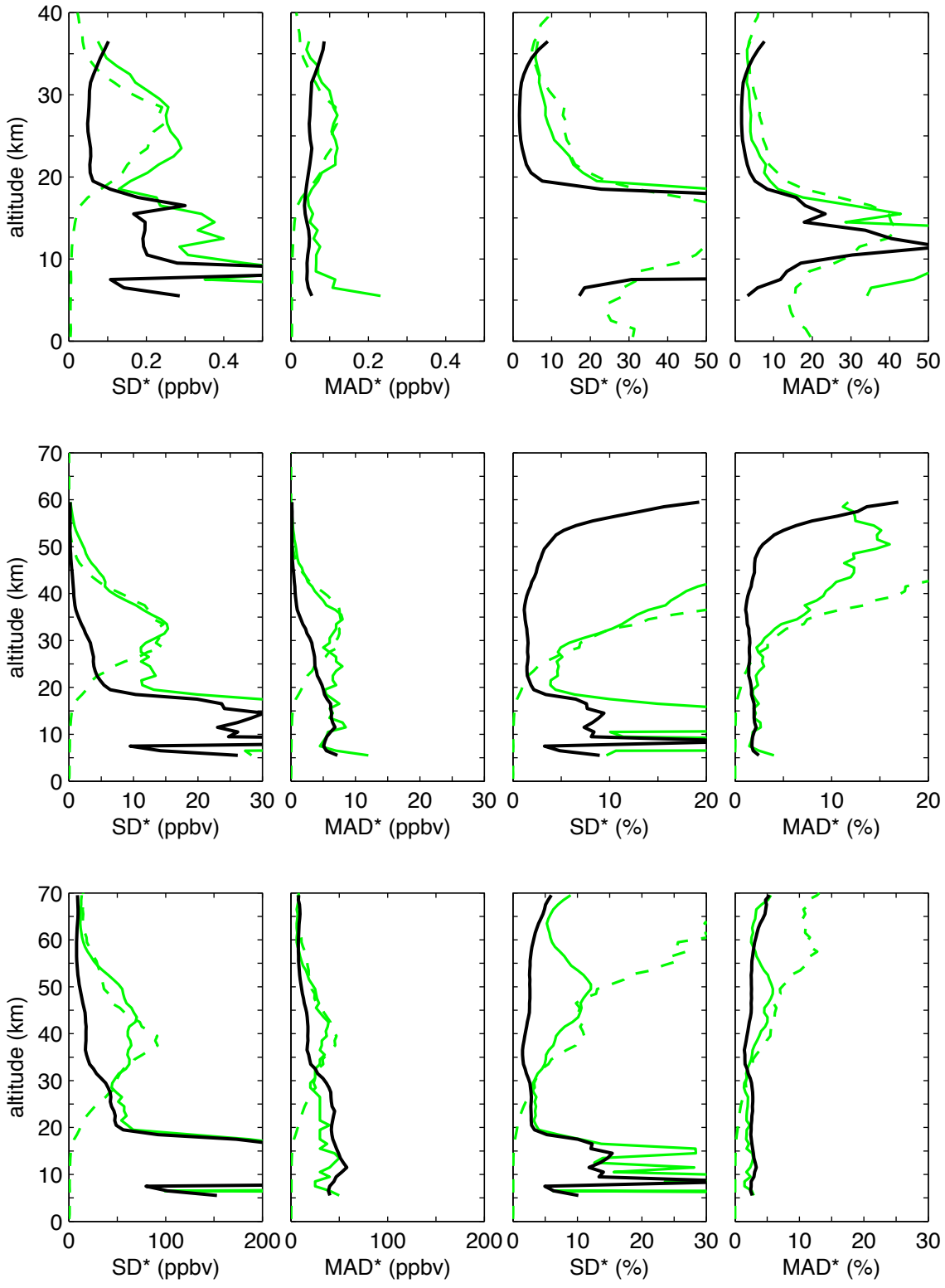

Figure 6.21: As Figure 6.20 for HNO_3 (top), N_2O (middle) and CH_4 (bottom).

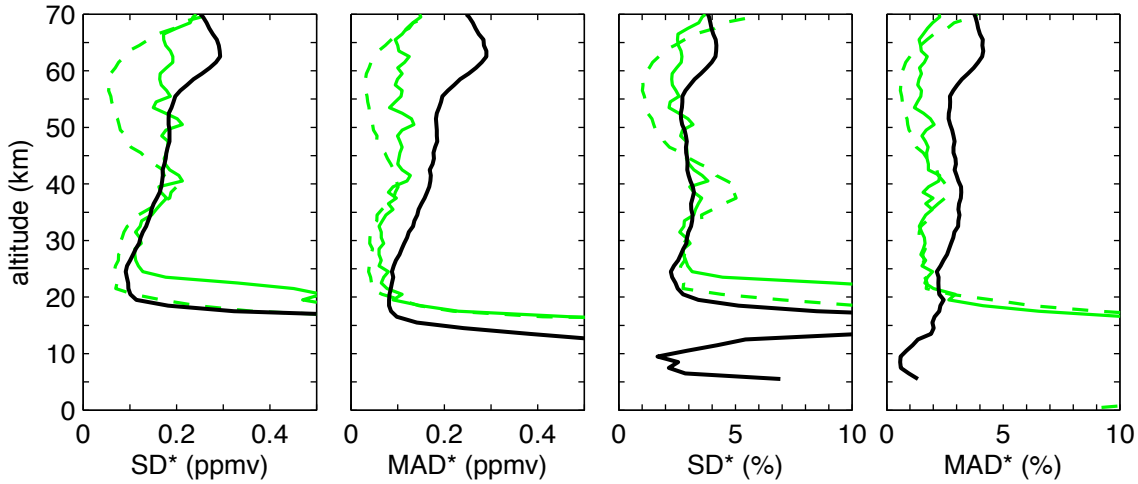

Figure 6.22: As Figure 6.22 for H_2O

CMAM variability, ACE-FTS scatter, and the ACE-FTS RRE for H_2O are shown in Figure 6.22. In terms of the SD* statistic, the CMAM variability is comparable to (20 – 40 km), or smaller than (above 40 km), the ACE-FTS RMS RRE. The ACE-FTS SD* is seen to agree well with the RMS RRE, except below 25 km. The ACE-FTS MAD*, on the other hand, is significantly smaller than the median RRE, and is in closer agreement with the CMAM variability than with the RRE. In other words, while the scatter of the full set of ACE-FTS measurements (i.e., the SD*) agrees well with the RREs, the scatter of the central portion of the measurements (i.e., the best 50%, quantified by the MAD*) is significantly less than the RREs for this portion of data. It would appear that the RREs of the "well behaved" retrievals may actually be overestimated.

6.5 Summary

This chapter has focussed on validating the reported random errors of the ACE-FTS trace gas measurements. Past studies have succeeded in validating the errors of other instrument measurements by comparing the scatter in measured profiles over time-space regions of small expected natural variability with the reported errors. For this

technique to succeed, one requires that the natural variability present in the data be small in comparison to the variability produced by the random instrument error.

Natural variability in the stratosphere, on the large scales resolved by remote sounding measurements, is often assumed to be minimal in the tropics compared to other latitudes. This is confirmed by an analysis of the HALOE climatology, which shows that the region bounded by the latitudes 10°S and 10°N remains relatively isolated from the strong variability associated with the winter hemisphere surf zone.

The scatter in tropical ACE-FTS measurements has been quantified based on a series of successively tighter temporal bounds. The scatter present in short-term subsets of the data (1-day and 4-day) is compared directly with the ACE-FTS reported random errors.

Good agreement between measured scatter and RREs is seen for CO and H_2O above 25 km. For these two species, this close agreement represents a validation of the random measurement errors in the manner of previous satellite validation studies.

For the other species analyzed, statistically significant differences were observed between the measured short-term scatter and the RREs. While these discrepancies may be a sign of larger than reported errors in the measurements, they may also be due to the presence of real natural variability. Examination of the scatter in chemical fields from the CMAM strongly supports the hypothesis that the scatter in the ACE-FTS measurements in these cases is due to real natural variability. Thus, while the RREs for the species other than CO and H_2O cannot be validated in the sense of showing agreement between scatter and RREs, it appears that the random errors of the ACE-FTS measurements are smaller than the short-term natural variability of the tropics.

This result has important implications for the use of tropical ACE-FTS data. This chapter has focussed on an effort to remove the "signal" from the tropical ACE-FTS data set in order to isolate the "noise". Usually though, the signal is what is important for scientific use. From this perspective, Figures 6.9 – 6.14 can be interpreted as quantifying the magnitude of signal over different time scales compared to the predicted noise. For H_2O and CO we have, at short time scales, measurement

scatter which is dominated by random error, which diminishes its scientific usefulness. The results of this chapter thus imply that for the other species examined in this work, the tropical ACE-FTS retrievals are scientifically useful at many altitudes right down to the shortest time scales.

Chapter 7

Conclusions

7.1 Summary of results

This thesis has examined measurements of stratospheric trace gases from a number of remote sounding instruments, with the purpose of using measurements to study ways in which the composition of the stratosphere varies in time and space, and ways by which the quality of the measurements can be assessed.

Chapter 4 focused on the analysis of spectral data recorded by low-resolution infrared radiometers from balloon flights in 1990, 1998, 2000, and 2002, and the retrieval of vertical profiles of HNO_3 from this data. The measurements were taken over a midlatitude Northern Hemisphere site in late summer, when stratospheric dynamical variability is minimal. The retrieved HNO_3 profiles show good agreement, in profile shape and magnitude, with measurements made by ACE-FTS at a similar season and latitude region over the years 2004–2006. The variance of the HNO_3 profiles measured over the MANTRA era (1998–2002) is in good agreement with the variability estimated by the CMAM, when the limited sampling and random error of the measurements are taken into account. The radiometer measurements represent a consistent data set with samples before and after the Mt. Pinatubo eruption of 1991 which perturbed NO_y partitioning, and hence HNO_3 levels, through the injection of aerosols into the stratosphere. The UARS MLS instrument, which began observation soon after the Mt. Pinatubo eruption, found a significant trend in HNO_3 over the

time span 1993–1997 (Randel et al., 1999) after aerosol levels had decreased to near background levels. Modeling results suggest that this measured trend may have been due to a slow relaxation to pre-Mt. Pinatubo eruption conditions, rather than any underlying long-term trend (Rinsland et al., 2003). Comparison of the HNO_3 profiles retrieved from the emission radiometer measurements of before and well after the Mt. Pinatubo eruption has found no significant difference; however, the uncertainty of the measurements precludes conclusive confirmation of the slow relaxation hypothesis.

Chapter 5 explored the complex relationship between the variance and covariance of coincident measurements by two instruments, the random errors of the instruments, and inter-instrumental biases. A statistical comparison model was developed that differentiates between additive and multiplicative bias. In a comparison of coincident measurements by the ACE-FTS and Aura MLS instruments, multiplicative bias was found to be significant at certain heights for O_3 and N_2O, and at all heights for HNO_3. The multiplicative bias found in the HNO_3 comparison is now known to be related to a systematic error in the version 1.5 MLS retrievals, related to spectroscopic line parameterization (Santee et al., 2007). Based on these results, it is suggested that future satellite validation studies that estimate both multiplicative and additive biases may be able to identify systematic errors, such as those due to spectroscopic errors, and produce accurate calibration factors between instruments.

Since the estimation of multiplicative bias in the comparison of coincident measurements is dependent on the random measurement errors of each data set, the methods used to estimate multiplicative bias in Chapter 5 also produce estimates of (or assume values for) the random errors. The random error estimates produced through the "instrumental variable" method introduced in Chapter 5 are significantly larger than the reported random errors for ACE-FTS and MLS. Whether this is a result of real differences in the atmosphere between the two measurements closely spaced (but not truly coincident) in time and space, or an underestimation of the reported errors, can not be determined solely from the instrumental variable analysis. This highlights a weakness of the instrumental variable method, and motivates the use of a more straight forward method for estimating random error.

The validity of the reported random errors of the ACE-FTS retrievals was explored in Chapter 6. Following the methodology of past validation studies, ACE-FTS errors were compared with the statistical scatter of measurements in the tropics, where natural variability was confirmed to be weak. For two trace gas species measured by ACE-FTS, CO and H_2O, the short-term scatter, quantified by a composite SD of measurements between 10°S and 10°N over four-day periods, was seen to be consistent with, or smaller than, the RMS RRE of the measurements. For these two species, this result represents a validation of the RREs in the manner of previous satellite validation studies.

For O_3, HNO_3, N_2O, and CH_4, the measured short-term scatter quantified by the composite SD was seen to be at certain altitudes significantly larger than the RREs. Thus, the RREs for these species cannot be validated in the sense of showing agreement between scatter statistics and RREs. While the discrepancies between scatter and the RREs may signal an underestimation of the RREs, they may also be due to the presence of natural variability. Examination of the scatter in chemical fields from the CMAM strongly supports the hypothesis that the scatter in the ACE-FTS measurements for these species above ∼20 km is due to real natural variability. This suggests that the true random errors of the ACE-FTS measurements are smaller than the short-term natural variability of the tropics. In other words, as long as the natural variability of the tropics is truly minimal compared to other regions, then this result implies that the scatter seen in any set of ACE-FTS measurements at altitudes above ∼20 km reflects real atmospheric variability.

7.2 Suggestions for future work

The results of Chapter 5 show the descriptive power of estimating multiplicative bias between instruments. Such bias may be relatively common between instruments, due to systematic differences in spectroscopic parameters, or other scaling errors. The question may then be asked, are the methods for estimating multiplicative bias explored in Chapter 5 advisable for future validation work?

Method 2, in which reported errors for one instrument were used to estimate the multiplicative bias, leads to anomalous results when the reported errors differed largely from the scatter in the measurements. Conceivably, it would be possible to improve these results by first obtaining estimates of the random error in each measurement set through independent means, such as an analysis of the scatter in tropical measurements. However, it was seen in Chapter 6 that, for ACE-FTS, the random error estimates based on the short-term scatter in tropical measurements are largely dominated by natural variability. Thus, estimates of multiplicative bias using random error estimates based on tropical scatter analysis would likely lead to skewed results.

The instrumental variable method aims to avoid direct dependence on random error information by incorporating a third piece of information. In Chapter 5, this method was applied by taking a secondary coincidence by Aura MLS as the third measurement. It can be argued that this methodology would be equivalent to estimating the Aura MLS random error independently through the variance of self-coincidences, and then using these random error estimates in the calculation of multiplicative bias through Method 2. Thus, in the special case used here, the instrumental variable method can be seen to be simply combining the two-step procedure discussed above into one streamlined calculation. While it may be elegant, it doesn't appear to add much power to the analysis.

It should be noted also that the estimation of multiplicative bias can be performed by "beating down" the random error rather than trying to include it in the analysis. Retrievals from the Sub-Millimeter Radiometer (Frisk et al., 2003; Murtagh et al., 2002; Olberg et al., 2003) have been compared with those from MIPAS (Urban et al., 2006) and Aura MLS (Urban et al., 2008) by performing linear fits on zonal means of the data from the various instruments. This simple technique produces estimates of multiplicative and additive bias, and likely avoids the deleterious effects of random errors in the regression.

There are however instances where the instrumental variable method would have unique and potentially powerful applications. The ACE satellite platform carries

two instruments: the ACE-FTS, which has been a focus of much of this thesis, and the MAESTRO instrument, which has not been mentioned hitherto. MAESTRO is a dual-grating diode array spectrophotometer operating in the UV to near infrared spectral range (McElroy et al., 2007). ACE-FTS and MAESTRO share a common light beam on the ACE platform, therefore their measurements are truly coincident. Comparisons of ACE-FTS and MAESTRO retrievals thus offer an ideal test case for comparison methodology techniques. Early attempts at comparing ACE-FTS and MAESTRO retrievals have, however, uncovered random and systematic errors in the MAESTRO O_3 retrievals that suggest the presence of an altitude registration problem (Kar et al., 2007). The MAESTRO errors produce anomalies that appear pseudo-random but heavily weighted to either positive (for sunset occultations) or negative (for sunrise occultations) values, and are not well handled by the techniques introduced in Chapter 5. As MAESTRO retrieval algorithms improve in future, it may be useful to use the truly coincident measurements of ACE-FTS and MAESTRO in the instrumental variable method described in Chapter 5.

Another scenario that appears to lend itself to the instrumental variable method is that of incorporating model analysis data in the comparison of coincident measurements. For example, in the comparison of measurements of a long-lived tracer, such as N_2O or O_3 in the lower stratosphere, a dynamical tracer such as PV obtained from a global reanalysis could be used as the third "measurement". This type of analysis would be in some ways similar to the use of PV to produce a "proxy" O_3 for use in validation studies (as in Randall et al., 2002, 2005b). The use of the instrumental variable method may prove useful in the incorporation of analysis data into satellite validation studies.

In Chapter 3, measurements made by the CRISTA instrument over just a few days in 1997 were used to explore the phenomenon of short-term, large-scale variability in the atmosphere, quantified in terms of SDs of measurements from zonal means. In Chapter 6, zonal SDs provided with a climatology of HALOE measurements were used to identify regions of minimum short-term variability. In the analysis of ACE-FTS scatter in Chapter 6, it was found that interannual variability, at least in the

tropics, is a large source of scatter in the measurements. This was also seen in the comparison of SDs of HALOE measurements from all the years of observations (i.e., a SD containing interannual variability) with an average SD, produced by averaging the SD maps for each month of HALOE observation. To the author's knowledge, there has not yet been a comprehensive study of the short-term variability of the atmosphere, measured in terms of the SDs about zonal means of satellite measurements, over the time span of a few years. Such a study could potentially clarify the effects of interannual and short-term variability on climatological SDs, as well as the effects of calculating zonal statistics based on latitude versus equivalent latitude coordinates. Since the short-term variability of trace gases is tied directly to the dynamical activity of the atmosphere, comparison of short-term variability measurements with that of CCMs may also be used to help diagnose the realism of the models (as in Erbertseder et al., 2005).

From a complementary perspective, investigating the short-term variability of trace gases in CCMs may be helpful for future observational studies. Satellite validation studies often choose coincidence criteria somewhat arbitrarily: it could thus be very useful to produce estimates of the non-coincidence error for given coincidence criteria on a species-by-species basis from a model, or perhaps produce coincidence criteria suggestions for each species based on a maximum threshold for non-coincidence error. Pendlebury et al. (2008) have investigated the potential effects of short-term variability on trace gas measurements in an analysis of normal-mode Rossby waves in midlatitude summer. It might be interesting to extend this analyses to other regions, and consider other sources of variability. For instance, the issue of frozen-in variance in long-lived tracers such as N_2O and CH_4, first postulated by Hess and Holton (1985), has not to the author's knowledge been investigated in a CCM.

The validation of ACE-FTS reported random errors in Chapter 6 focussed on the scatter of measurements in the tropical regions, since this method is consistent with prior work, and since an analysis of HALOE measurements confirmed that variability is small in the tropics in an annual mean sense, independent of species. It is possible, however, that regions and periods of even smaller natural variability may exist for

each individual trace gas. For instance, measured O_3 variability from HALOE was seen to be quite small in the summer midlatitudes. The work of Chapter 6 might thus be improved by analyzing the scatter of measurements in species-specific locations, in order to identify regions of minimal variability, and obtain closer agreement between measured scatter and RREs.

Since trace gas isopleths are sloped in the extratropics, variability within any extratropical zonal band may be reduced by removing the latitudinal gradient. This can be easily accomplished through linear regression, which can produce residuals from a simple fit of the measured data versus explanatory parameters such as latitude, equivalent latitude, and/or potential temperature. The fit residuals then take the place of deviations (from the mean or median) used in the statistical analysis of Chapter 6. Preliminary work on this type of analysis has shown promise, and it is hoped that future work may strengthen the validation of ACE-FTS RREs found in Chapter 6.

Appendix A

CMAM

The Canadian Middle Atmosphere Model (CMAM) is an extended version of the Canadian Centre for Climate Modelling and Analysis spectral General Circulation Model (GCM). The dynamical core and chemistry scheme are described by Beagley et al. (1997) and de Grandpré et al. (1997) respectively. The distributions of chemical species in the CMAM have been seen to generally compare well with observations (e.g., de Grandpré et al., 2000; Farahani et al., 2007; Hegglin and Shepherd, 2007; Jin et al., 2005; Jin et al., 2009; Melo et al., 2008).

The particular version of the CMAM used in this work is discussed by Eyring et al. (2006). The model resolution is 3.75°×3.75° (T32) in the horizontal, with 71 levels in the vertical, with a model lid at 0.0006 hPa (∼100 km). Small-scale orographic gravity waves are parametrized according to Scinocca and McFarlane (2000); non-orographic gravity waves follow the parametrization of Scinocca (2003). Chemical species are transported in the model based on a spectral tracer advection scheme in the horizontal, and finite elements in the vertical. Model chemistry and dynamics are coupled through the radiative heating of O_3 and H_2O. The model includes both gas-phase chemistry and heterogeneous chemistry on aerosols, and on polar stratospheric clouds (PSCs), although formation of nitric acid trihydrate (NAT) particles and dehydration/denitrification by gravitational settling of PSC particles are not included, based on the rationale described by Hitchcock et al. (2009).

CMAM results shown in this work are from a transient run of the years 1960-

2004 described by Eyring et al. (2006). The specifications of the simulations follow or are similar to the "reference simulation 1" (REF1) of CCMVal (Eyring et al., 2005) and include anthropogenic and natural forcings based on changes in sea surface temperatures, trace gases, and aerosol effects from major volcanic eruptions.

In the CMAM REF1 simulation:

1. Sea surface temperatures are prescribed as monthly means following a global sea ice and sea surface temperature data set, based on blended satellite and in situ observations, provided by the UK Met Office.

2. The surface concentrations of GHGs are based on IPCC (2001).

3. The surface halogens are from WMO (2003) and are extended through 2004.

4. Both chemical and direct radiative effects of enhanced stratospheric aerosol abundance from large volcanic eruptions are considered by prescribing observed sulfate aerosol surface area densities specified from a monthly climatology based on satellite data.

Simulated chemical fields from the last ten years (1995–2004) of the CMAM REF1 simulation have been used in this thesis, for comparison with measurements. The chemical fields are available for every model gridpoint, in save intervals of 18 hours, with this high sampling frequency (i.e., compared to monthly means) allowing the calculation of short-term variability. In order to best reproduce the sampling of measurements, CMAM chemical fields have been interpolated from pressure coordinates to approximate altitude using the geopotential field.

Mean profiles from the equatorial 10°S–10°N latitude band for four species, where CMAM means are calculated over the 10 years of CMAM data, and ACE-FTS means are calculated over four years (2004-2007), are shown in Figure A.1. Compared to ACE-FTS, CMAM O_3 is seen to be biased low above 30 km, which is in agreement with comparisons shown by Eyring et al. (2006) and Hegglin and Shepherd (2007). Hegglin and Shepherd (2007) have noted that O_3 from the CMAM run in data assimilation mode, i.e., with temperatures constrained by observations, show better

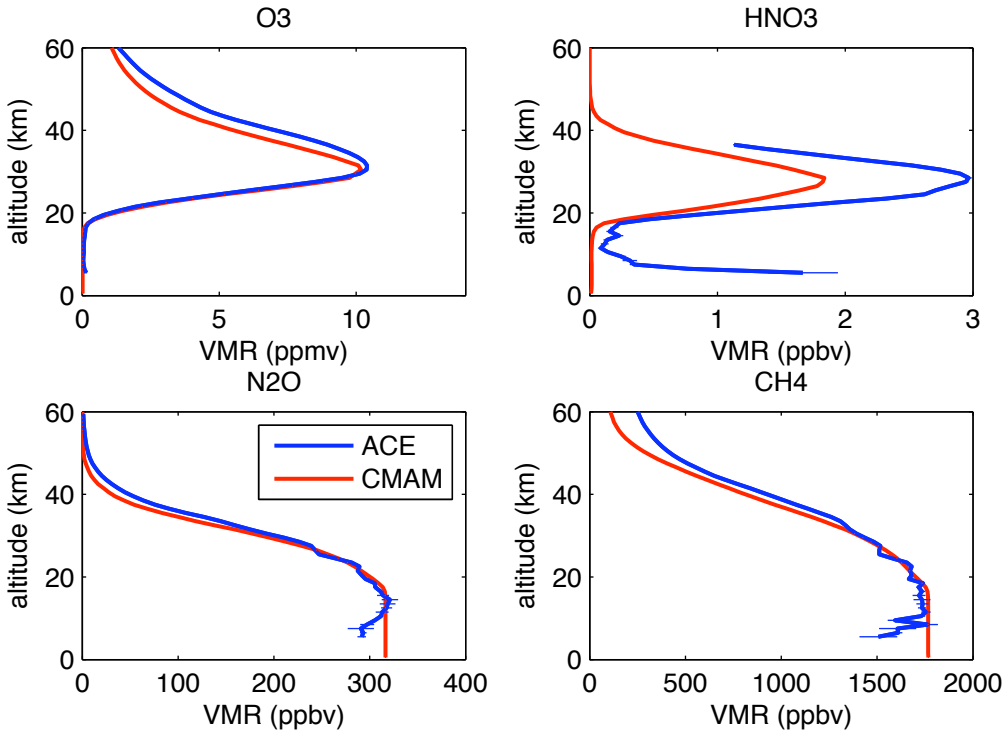

Figure A.1: Equatorial (10°S – 10°N) CMAM 10-year mean profiles of O_3, HNO_3, N_2O and CH_4 compared to 4-year mean ACE-FTS measurements (from 2004–2007) from the same region. The standard errors of the means calculated for both the model and measurements are shown as horizontal lines, and are smaller than the width of the line for the plotted profile for all but the lowest altitude ACE-FTS results.

agreement with the ACE-FTS O_3 measurements. This implies the presence of an upper stratospheric warm bias in the free-running CMAM, which leads to the low bias in CMAM upper stratospheric O_3. Slight low biases for N_2O and CH_4 in the middle to upper stratosphere are known to be related to the vertical diffusion coefficient in the model (A. I. Jonsson, personal communication). CMAM HNO_3 is significantly biased low at all altitudes. The cause of this HNO_3 bias is currently unknown.

Appendix B

Observational data source descriptions

Observational data sets used in this thesis, and the instruments that have produced them, are described below, listed by order of appearance in the main text.

B.1 ACE-FTS

The Atmospheric Chemistry Experiment - Fourier transform spectrometer (ACE-FTS) onboard the SciSat-1 satellite, launched 12 August 2003 into a low-Earth circular orbit (altitude 650 km, inclination 74°), collects high resolution (0.02 cm^{-1}) infrared (2.2 – 13.3 μm, 750 – 4400 cm^{-1}) spectra, measuring atmospheric extinction by solar occultation (Bernath et al., 2005). ACE-FTS performs approximately 15 sunrise and 15 sunset occultations per day, with a latitudinal coverage that depends strongly on time of year. Over a full year, the latitudinal coverage of ACE-FTS covers approximately 85°N to 85°S (Bernath et al., 2005).

Profiles as a function of altitude for pressure, temperature, and over 30 trace gases are retrieved from these spectra. The details of ACE-FTS processing are described in Boone et al. (2005). Briefly, a non-linear least squares global fitting technique is employed to analyze selected microwindows (0.3–30 cm^{-1} wide portions of the spectrum containing spectral features for the target molecule). Prior to performing VMR

retrievals, pressure and temperature as a function of altitude are determined through the analysis of CO_2 lines in the spectra. A crucial aspect of the pressure and temperature retrieval process is pointing knowledge. No information from sensors onboard the satellite (other than a clock) is used in the computation of tangent altitudes; instead the approach treats tangent altitudes as unknown parameters in the pressure/temperature retrievals, i.e., the tangent heights for the measurement scans are themselves retrieved. For the VMR retrievals, forward model calculations employ the spectroscopic constants and cross-section measurements from the HITRAN 2004 line list (Rothman et al., 2005). First-guess profiles are based on ATMOS measurements, but the retrievals are not sensitive to this a priori information.

The altitude spacing of the FTS measurements, controlled by the scan time and the orbit of the satellite, is typically 3–4 km. It varies with the beta angle, the angle between the satellite velocity vector and a vector from the Earth to the Sun, where a beta angle of zero features the sun setting (or rising) exactly perpendicular to the Earth horizon. The altitude spacing ranges from 2 km for long occultations with high beta angle ($\sim$55°) to 6 km for occultations with beta angle zero. The vertical resolution of the profile retrievals is related to the field of view of the instrument, which is circular with a diameter of 1.25 mrad, and gives a maximum altitude resolution between 3 and 4 km for a satellite 2700 km from the tangent point (Boone et al., 2005).

Each ACE-FTS retrieved profile is reported on both the vertical grid of the tangent heights of the recorded spectra, and on a standard grid of 1-km resolution. Piecewise quadratic interpolation is used to cast information from the retrieval grid onto a 1-km grid (Boone et al., 2005). In order to calculate statistics on vertical layers, it is convenient to have the measurements on standard levels. Therefore, this thesis makes use of the results reported on the 1-km grid.

ACE-FTS results shown in this thesis are from the version 2.2 data set, with O_3 from the v2.2 O_3 update. The results of a number of V2.2 validation studies are collected in a special issue of Atmospheric Chemistry and Physics `http://www.atmos-chem-phys.net/special_issue114.html`, including articles focussing on O_3

(Dupuy et al., 2009), N_2O (Strong et al., 2008), HNO_3 (Wolff et al., 2008), NO and NO_2 (Kerzenmacher et al., 2008).

B.2 UARS MLS

The Microwave Limb Sounder (MLS) onboard the Upper Atmosphere Research Satellite (UARS) measured millimeter- and submillimeter-wavelength thermal emission as the instrument field of view was scanned through the atmospheric limb (Barath et al., 1993; Waters et al., 1999).

UARS was launched on September 12, 1991, and the first full day of MLS data was obtained on September 21. Latitudinal coverage of the MLS measurements extended from 80° on one side of the equator to 34° on the other. Approximately 10 times per year, UARS performed a 180° "yaw maneuver" such that MLS alternated between viewing northern and southern high latitudes. A calendar of MLS daily data coverage and a detailed chronology of MLS operations are provided by Livesey et al. (2003). After several years in orbit, degradation in the performance of the MLS antenna scan mechanism, together with a reduction in power available from the UARS spacecraft, resulted in markedly reduced data sampling (Santee et al., 2004). MLS was put in standby mode in July 1999 to conserve its remaining lifetime for possible future measurements (Livesey et al., 2003). It was operated again for two brief periods in February and March 2000 to obtain Arctic observations (Santee et al., 2000), and briefly in mid-August 2001 to obtain correlative observations for the Odin satellite.

The MLS retrieval algorithms, described in detail by Livesey et al. (2003), are based on the standard optimal estimation method (Rodgers, 2000). MLS retrievals, for version 5 and later data sets, are reported on a vertical pressure grid, with six surfaces per decade change in pressure (corresponding to ∼2.5 km spacing). The vertical resolution of the retrievals depend on the species retrieved, and may be coarser than the reporting grid.

Version 6 UARS MLS HNO_3 data is used in Chapter 4. Version 6 HNO_3 retrievals are of significantly higher quality than previous versions, as discussed by Santee et al.

(2004). The vertical resolution of the v6 HNO_3 data varies with altitude from 4.5 km at 100 hPa to 10.5 km at 4.6 hPa. Single-profile precision for the HNO_3 retrievals is 1.0–1.5 ppbv throughout the vertical range of measurement (Santee et al., 2004).

Version 6 UARS MLS HNO_3 data was obtained from the Constituent Observational Database (found at `http://www.autochem.info/`) maintained by Dr. David Lary. This data source is especially convenient since the MLS retrievals, originally reported only on pressure surfaces, are here given along with corresponding geometric altitudes taken from the UKMO reanalysis meteorological fields. UARS MLS data is also available through the NASA Goddard Space Flight Center's Distributed Active Archive Center (DAAC), `http://daac.gsfc.nasa.gov/`.

B.3 Aura MLS

The Microwave Limb Sounder (MLS) onboard the Aura satellite, launched on 15 July 2004, measures millimeter- and submillimeter-wavelength thermal emission from the Earth's limb with seven radiometers covering five broad spectral regions (Waters et al., 2006). The Aura MLS fields of view point forward in the direction of orbital motion and vertically scan the limb in the orbit plane, leading to data coverage from 82°S to 82°N latitude on every orbit. Thus Aura MLS obtains continuous daily sampling of both polar regions, with none of the temporal gaps from yaw maneuvers that occurred with UARS MLS. The MLS limb scans are synchronized to the Aura orbit, with 240 scans per orbit at essentially fixed latitudes. This results in ∼3500 scans per day, with an along-track separation between adjacent retrieved profiles of 1.5° great circle angle (∼165 km). The longitudinal separation of MLS measurements, set by the Aura orbit, is 10°– 20° over low and middle latitudes, with much finer sampling in the polar regions.

The MLS Level 2 data (retrieved geophysical parameters and diagnostics at the measurement locations along the suborbital track) are generated from input Level 1 data (calibrated radiances and engineering information) by the MLS data processing software. The MLS retrieval algorithms, described in detail by Livesey et al. (2006),

are based on the standard optimal estimation method; they employ a two-dimensional approach that takes into account the fact that limb observations from consecutive scans cover significantly overlapping regions of the atmosphere. The data are divided into overlapping "chunks" consisting of the measurements in a 15° span of great circle angle (typically about 10 vertical profiles); retrievals are performed for each of these chunks independently and then joined together to produce a complete set of output (Livesey et al., 2006). Most MLS data products are reported on a fixed vertical pressure grid with six levels per decade change in pressure in the troposphere and stratosphere.

Aura MLS retrievals of O_3, HNO_3 and N_2O used in Chapter 5 are from version 1.5, and were downloaded from the NASA Goddard Space Flight Center's Distributed Active Archive Center, `http://daac.gsfc.nasa.gov/`.

B.4 ATMOS

The Atmospheric Trace Molecule Spectroscopy (ATMOS) experiment is a Fourier-transform interferometer that measured infrared solar absorption at a spectral resolution of $\sim$0.01 cm^{-1}. ATMOS performed measurements during four Space Shuttle flights: Spacelab 3 and the Atmospheric Laboratory for Applications and Science (ATLAS)-1, -2, and -3 missions, in 1985, 1992, 1993, and 1994, respectively (Irion et al., 2002).

ATMOS version 3 retrievals use a robust method of simultaneously fitting multiple gases within each spectral window. This is combined with a global-fit algorithm to retrieve a vertical VMR profile simultaneously at all altitudes within an occultation. The ATMOS O_3 random error profile shown in Chapter 6 is from the version 3 retrievals, reported on by Irion et al. (2002) and available at `http://atmos.jpl.nasa.gov/atmos`.

B.5 HALOE

The Halogen Occultation Experiment (HALOE) onboard the UARS spacecraft, launched September 12, 1991, measured infrared radiation in a number of channels covering selected portions of the spectral range from 2.45 to 10.04 μm, measuring atmospheric extinction by solar occultation (Russell III et al., 1993). HALOE performed approximately 15 sunrise and 15 sunset occultations per day, with a latitudinal coverage that depends strongly on time of year. Over a full year, the latitudinal coverage of HALOE covers 80°N to 80°S.

HALOE results shown in Chapter 6 are taken from the HALOE climatology of Grooß and Russell III (2005), which includes (1) means and (2) standard deviations of all HALOE measurements from 1991–2002, in 5°-wide latitude and equivalent latitude bands on a vertical pressure grid, and (3) means and (4) standard deviations of HALOE measurements for equivalent latitude bins (i.e., not for latitude bins) for each individual month of the HALOE mission.

References

Abrams, M. C., Toon, G. C., and Schindler, R. A.: Practical example of the correction of Fourier-transform spectra for detector nonlinearity, Applied Optics, 33, 6307–6314, 1994.

Abrams, M. C., Chang, A. Y., Gunson, M. R., Abbas, M. M., Goldman, A., Irion, F. W., Michelsen, H. A., Newchurch, M. J., Rinsland, C. P., Stiller, G. P., and Zander, R.: On the assessment and uncertainty of atmospheric trace gas burden measurements with high resolution infrared solar occultation spectra from space by the ATMOS experiment, Geophysical Research Letters, 23, 2337–2340, doi: 10.1029/96GL01794, 1996.

Akiyoshi, H., Sugata, S., Sugita, T., Nakajima, H., Hayashi, H., Kurokawa, J., and Takahashi, M.: Low-N_2O air masses after the breakdown of the Arctic polar vortex in 1997 simulated by the CCSR/NIES nudging CTM, Journal of the Meteorological Society of Japan, 80, 451–463, 2002.

Allen, M., Lunine, J. I., and Yung, Y. L.: The vertical distribution of ozone in the mesosphere and lower thermosphere, Journal of Geophysical Research, 89, 4841–4872, doi:10.1029/JD089iD03p04841, 1984.

Andrews, D., Holton, J., and Leovy, C.: Middle Atmosphere Dynamics, Academic Press, 1987.

Baldwin, M. P., Gray, L. J., Dunkerton, T. J., Hamilton, K., Haynes, P. H., Randel, W. J., Holton, J. R., Alexander, M. J., Hirota, I., Horinouchi, T., Jones, D. B. A.,

Kinnersley, J. S., Marquardt, C., Sato, K., and Takahashi, M.: The quasi-biennial oscillation, Reviews of Geophysics, 39, 179–230, doi:10.1029/1999RG000073, 2001.

Barath, F. T., Chavez, M. C., Cofield, R. E., Flower, D. A., Frerking, M. A., Gram, M. B., Harris, W. M., Holden, J. R., Jarnot, R. F., and Kloezeman, W. G.: The Upper Atmosphere Research Satellite microwave limb sounder instrument, Journal of Geophysical Research, 98, 10 751, doi:10.1029/93JD00798, 1993.

Beagley, S. R., de Grandpré, J., Koshyk, J. N., McFarlane, N. A., and Shepherd, T. G.: Radiative-dynamical climatology of the first-generation Canadian Middle Atmosphere Model, Atmosphere–Ocean, 35, 293–331, 1997.

Bernath, P. F., McElroy, C. T., Abrams, M. C., Boone, C. D., Butler, M., Camy-Peyret, C., Carleer, M., Clerbaux, C., Coheur, P.-F., Colin, R., DeCola, P., DeMazière, M., Drummond, J. R., Dufour, D., Evans, W. F. J., Fast, H., Fussen, D., Gilbert, K., Jennings, D. E., Llewellyn, E. J., Lowe, R. P., Mahieu, E., McConnell, J. C., McHugh, M., McLeod, S. D., Michaud, R., Midwinter, C., Nassar, R., Nichitiu, F., Nowlan, C., Rinsland, C. P., Rochon, Y. J., Rowlands, N., Semeniuk, K., Simon, P., Skelton, R., Sloan, J. J., Soucy, M.-A., Strong, K., Tremblay, P., Turnbull, D., Walker, K. A., Walkty, I., Wardle, D. A., Wehrle, V., Zander, R., and Zou, J.: Atmospheric Chemistry Experiment (ACE): Mission overview, Geophysical Research Letters, 32, L15S01, doi:10.1029/2005GL022386, 2005.

Boone, C. D., Nassar, R., Walker, K. A., Rochon, Y., McLeod, S. D., Rinsland, C. P., and Bernath, P. F.: Retrievals for the atmospheric chemistry experiment Fourier-transform spectrometer, Applied Optics, 44, 7218–7231, 2005.

Borrello, S., Kinch, M., and LaMont, D.: Photoconductive HgCdTe detector performance with background variations, Infrared Physics, 17, 121–125, 1977.

Brasseur, G. and Solomon, S.: Aeronomy of the Middle Atmosphere, D. Reidel Publishing Company, 1984.

Brasseur, G. and Solomon, S.: Aeronomy of the Middle Atmosphere, Springer, 2005.

Brewer, A. W.: Evidence for a world circulation provided by the measurements of helium and water vapor distribution in the stratosphere, Quarterly Journal of the Royal Meteorological Society, 75, 351–363, 1949.

Broderick, A. J., Oliver, R. C., Jesson, J. P., and Delwiche, C. C.: Section 1: Does Man Influence Stratospheric Ozone?, in: Stratospheric Ozone and Man, edited by Bower, F. A. and Ward, R. B., CRC Press Inc., 1982.

Brühl, C., Drayson, S. R., Russell III, J. M., Crutzen, P. J., McInerney, J. M., Purcell, P. N., Claude, H., Gernandt, H., McGee, T. J., McDermid, I. S., and Gunson, M. R.: Halogen Occultation Experiment ozone channel validation, Journal of Geophysical Research, 101, 10 217–10 240, doi:10.1029/95JD02031, 1996.

Butchart, N. and Remsberg, E. E.: The area of the stratospheric polar vortex as a diagnostic for tracer transport on an isentropic surface, Journal of the Atmospheric Sciences, 43, 1319–1339, doi:10.1175/1520-0469(1986)043, 1986.

Butchart, N., Scaife, A., Bourqui, M., de Grandpre, J., Hare, S., Kettleborough, J., Langematz, U., Manzini, E., Sassi, F., Shibata, K., et al.: Simulations of anthropogenic change in the strength of the Brewer–Dobson circulation, Climate Dynamics, 27, 727–741, 2006.

Charney, J. G. and Drazin, P. G.: Propagation of planetary-scale disturbances from the lower into the upper atmosphere, Journal of Geophysical Research, 66, 83–109, doi:10.1029/JZ066i001p00083, 1961.

Coffey, M. T.: Observations of the impact of volcanic activity on stratospheric chemistry, Journal of Geophysical Research, 101, 6767–6780, doi:10.1029/95JD03763, 1996.

Crutzen, P. J., Isaksen, I. S. A., and Reid, G. C.: Solar proton events - Stratospheric sources of nitric oxide, Science, 189, 457–459, doi:10.1126/science.189.4201.457, 1975.

de Grandpré, J., Sandilands, J. W., Mcconnell, J. C., Beagley, S. R., Croteau, P. C., and Danilin, M. Y.: Canadian Middle Atmosphere Model: Preliminary results from the chemical transport module, Atmosphere–Ocean, 35, 385–431, 1997.

de Grandpré, J., Beagley, S. R., Fomichev, V. I., Griffioen, E., McConnell, J. C., Medvedev, A. S., and Shepherd, T. G.: Ozone climatology using interactive chemistry: Results from the Canadian Middle Atmosphere Model, Journal of Geophysical Research, 105, 26 475–26 492, doi:10.1029/2000JD900427, 2000.

Dobson, G. M. B.: Origin and distribution of the polyatomic molecules in the atmosphere, Proceedings of the Royal Society of London. Series A, Mathematical and Physical Sciences, 236, 187–193, 1956.

Dobson, G. M. B. and Harrison, D. N.: Measurements of the amount of ozone in the Earth's atmosphere and its relation to other geophysical conditions, Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character, 110, 660–693, `http://www.jstor.org/stable/94460`, 1926.

Dunn, G.: Design and Analysis of Reliability Studies: The Statistical Evaluation of Measurement Errors, Oxford University Press, 1989.

Dupuy, E., Walker, K. A., Kar, J., Boone, C. D., McElroy, C. T., Bernath, P. F., Drummond, J. R., Skelton, R., McLeod, S. D., Hughes, R. C., Nowlan, C. R., Dufour, D. G., Zou, J., Nichitiu, F., Strong, K., Baron, P., Bevilacqua, R. M., Blumenstock, T., Bodeker, G. E., Borsdorff, T., Bourassa, A. E., Bovensmann, H., Boyd, I. S., Bracher, A., Brogniez, C., Burrows, J. P., Catoire, V., Ceccherini, S., Chabrillat, S., Christensen, T., Coffey, M. T., Cortesi, U., Davies, J., De Clercq, C., Degenstein, D. A., De Mazière, M., Demoulin, P., Dodion, J., Firanski, B., Fischer, H., Forbes, G., Froidevaux, L., Fussen, D., Gerard, P., Godin-Beekmann, S., Goutail, F., Granville, J., Griffith, D., Haley, C. S., Hannigan, J. W., Höpfner, M., Jin, J. J., Jones, A., Jones, N. B., Jucks, K., Kagawa, A., Kasai, Y., Kerzenmacher, T. E., Kleinböhl, A., Klekociuk, A. R., Kramer, I., Kllmann, H., Kuttippurath, J., Kyrölä, E., Lambert, J.-C., Livesey, N. J., Llewellyn, E. J., Lloyd,

N. D., Mahieu, E., Manney, G. L., Marshall, B. T., McConnell, J. C., McCormick, M. P., McDermid, I. S., McHugh, M., McLinden, C. A., Mellqvist, J., Mizutani, K., Murayama, Y., Murtagh, D. P., Oelhaf, H., Parrish, A., Petelina, S. V., Piccolo, C., Pommereau, J.-P., Randall, C. E., Robert, C., Roth, C., Schneider, M., Senten, C., Steck, T., Strandberg, A., Strawbridge, K. B., Sussmann, R., Swart, D. P. J., Tarasick, D. W., Taylor, J. R., Tétard, C., Thomason, L. W., Thompson, A. M., Tully, M. B., Urban, J., Vanhellemont, F., Vigouroux, C., von Clarmann, T., von der Gathen, P., von Savigny, C., Waters, J. W., Witte, J. C., Wolff, M., and Zawodny, J. M.: Validation of ozone measurements from the Atmospheric Chemistry Experiment (ACE), Atmospheric Chemistry and Physics, 9, 287–343, `http://www.atmos-chem-phys.net/9/287/2009/`, 2009.

Efron, B. and Tibshirani, R. J.: An Introduction to the Bootstrap, Chapman & Hall/CRC, 1994.

Ehhalt, D. H., Röth, E. P., and Schmidt, U.: On the temporal variance of stratospheric trace gas concentrations, Journal of Atmospheric Chemistry, 1, 27–51, 1983.

Erbertseder, T., Eyring, V., Dameris, M., Bittner, M., and Loyola, D.: Comparison of zonal variability in total ozone derived from ERS-2 GOME and coupled chemistry-climate model, in: Envisat & ERS Symposium, vol. 572 of *ESA Special Publication*, 2005.

Erbertseder, T., Eyring, V., Bittner, M., Dameris, M., and Grewe, V.: Hemispheric ozone variability indices derived from satellite observations and comparison to a coupled chemistry-climate model, Atmospheric Chemistry and Physics, 6, 5105–5120, `http://www.atmos-chem-phys.net/6/5105/2006/`, 2006.

Evans, W. F. J., Kerr, J. B., Wardle, D. I., McConnell, J. C., Ridley, B. A., and Schiff, H. I.: Intercomparison of NO, NO_2, and HNO_3 measurements with photochemical theory, Atmosphere, 14, 189–198, 1976a.

Evans, W. F. J., Lin, C. I., and Midwinter, C. L.: The altitude distribution of nitric acid at Churchill, Atmosphere, 14, 172–179, 1976b.

Evans, W. F. J., Kerr, J. B., McElroy, C. T., O'Brien, R. S., Wardle, D. I., and Ridley, B. A.: The odd nitrogen mixing ratio in the stratosphere, Geophysical Research Letters, 4, 235–238, doi:10.1029/GL004i006p00235, 1977.

Evans, W. F. J., Kerr, J. B., McElroy, C. T., O'Brien, R. S., and McConnell, J. C.: Simulation of nitrogen constituent measurements from the August 28, 1976, Stratoprobe III Flight, Journal of Geophysical Research, 86, 12 066–12 070, doi:10.1029/JC086iC12p12066, 1981.

Evans, W. F. J., Kerr, J. V., McElroy, C. T., O'Brien, R. S., and McConnell, J. C.: Measurements of NO_2 and HNO_3 during a stratospheric warming at 54° in February 1979, Geophysical Research Letters, 9, 493–496, doi:10.1029/GL009i004p00493, 1982a.

Evans, W. F. J., McElroy, C. T., Kerr, J. V., and McConnell, J. C.: Simulation of the October 23, 1980 Stratoprobe flight, Geophysical Research Letters, 9, 223–226, doi:10.1029/GL009i003p00223, 1982b.

Evans, W. F. J., McElroy, C. T., and Galbally, I. E.: The conversion of N_2O_5 to HNO_3 at high latitudes in winter, Geophysical Research Letters, 12, 825–828, doi:10.1029/GL012i012p00825, 1985.

Eyring, V., Kinnison, D. E., and Shepherd, T. G.: Overview of planned coupled chemistry-climate simulations to support upcoming ozone and climate assessments, SPARC Newsletter, 25, 11–17, 2005.

Eyring, V., Butchart, N., Waugh, D. W., Akiyoshi, H., Austin, J., Bekki, S., Bodeker, G. E., Boville, B. A., Brühl, C., Chipperfield, M. P., Cordero, E., Dameris, M., Deushi, M., Fioletov, V. E., Frith, S. M., Garcia, R. R., Gettelman, A., Giorgetta, M. A., Grewe, V., Jourdain, L., Kinnison, D. E., Mancini, E., Manzini, E., Marchand, M., Marsh, D. R., Nagashima, T., Newman, P. A., Nielsen, J. E., Pawson, S., Pitari, G., Plummer, D. A., Rozanov, E., Schraner, M., Shepherd, T. G., Shibata, K., Stolarski, R. S., Struthers, H., Tian, W., and Yoshiki, M.: Assessment

of temperature, trace species, and ozone in chemistry-climate model simulations of the recent past, Journal of Geophysical Research, 111, doi:10.1029/2006JD007327, 2006.

Farahani, E., Fast, H., Mittermeier, R. L., Makino, Y., Strong, K., McLandress, C., Shepherd, T. G., Chipperfield, M. P., Hannigan, J. W., Coffey, M. T., Mikuteit, S., Hase, F., Blumenstock, T., and Raffalski, U.: Nitric acid measurements at Eureka obtained in winter 2001–2002 using solar and lunar Fourier transform infrared absorption spectroscopy: Comparisons with observations at Thule and Kiruna and with results from three-dimensional models, Journal of Geophysical Research, 112, D01 305, 2007.

Feigelson, E. D. and Babu, G. J.: Linear regression in astronomy II, The Astrophysical Journal, 397, 55, 1992.

Fioletov, V. E., Tarasick, D. W., and Petropavlovskikh, I.: Estimating ozone variability and instrument uncertainties from SBUV(/2), ozonesonde, Umkehr, and SAGE II measurements: Short-term variations, Journal of Geophysical Research, 111, 2305, doi:10.1029/2005JD006340, 2006.

Flaud, J.-M., Brizzi, G., Carlotti, M., Perrin, A., and Ridolfi, M.: MIPAS database: Validation of HNO_3 line parameters using MIPAS satellite measurements, Atmospheric Chemistry and Physics, 6, 5037–5048, `http://www.atmos-chem-phys.net/6/5037/2006/`, 2006.

Frisk, U., Hagström, M., Ala-Laurinaho, J., Andersson, S., Berges, J.-C., Chabaud, J.-P., Dahlgren, M., Emrich, A., Florén, H.-G., Florin, G., Fredrixon, M., Gaier, T., Haas, R., Hirvonen, T., Hjalmarsson, Å., Jakobsson, B., Jukkala, P., Kildal, P. S., Kollberg, E., Lassing, J., Lecacheux, A., Lehikoinen, P., Lehto, A., Mallat, J., Marty, C., Michet, D., Narbonne, J., Nexon, M., Olberg, M., Olofsson, A. O. H., Olofsson, G., Origné, A., Petersson, M., Piironen, P., Pons, R., Pouliquen, D., Ristorcelli, I., Rosolen, C., Rouaix, G., Räisänen, A. V., Serra, G., Sjöberg, F., Stenmark, L., Torchinsky, S., Tuovinen, J., Ullberg, C., Vinterhav, E., Wadefalk,

N., Zirath, H., Zimmermann, P., and Zimmermann, R.: The Odin satellite. I. Radiometer design and test, Astronomy and Astrophysics, 402, L27–L34, doi:10.1051/0004-6361:20030335, 2003.

Froidevaux, L., Livesey, N. J., Read, W. G., Jiang, Y. B., Jimenez, C., Filipiak, M. J., Schwartz, M. J., Santee, M. L., Pumphrey, H. C., Jiang, J. H., Wu, D. L., Manney, G. L., Drouin, B. J., Waters, J. W., Fetzer, E. J., Bernath, P. F., Boone, C. D., Walker, K. A., Jucks, K. W., Toon, G. C., Margitan, J. J., Sen, B., Webster, C. R., Christensen, L. E., Elkins, J. W., Atlas, E., Lueb, R. A., and Hendershot, R.: Early validation analyses of atmospheric profiles from EOS MLS on the Aura satellite, Geoscience and Remote Sensing, IEEE Transactions on, 44, 1106–1121, doi:10.1109/TGRS.2006.864366, 2006.

Froidevaux, L., Jiang, Y. B., Lambert, A., Livesey, N. J., Read, W. G., Waters, J. W., Browell, E. V., Hair, J. W., Avery, M. A., McGee, T. J., Twigg, L. W., Sumnicht, G. K., Jucks, K. W., Margitan, J. J., Sen, B., Stachnik, R. A., Toon, G. C., Bernath, P. F., Boone, C. D., Walker, K. A., Filipiak, M. J., Harwood, R. S., Fuller, R. A., Manney, G. L., Schwartz, M. J., Daffer, W. H., Drouin, B. J., Cofield, R. E., Cuddy, D. T., Jarnot, R. F., Knosp, B. W., Perun, V. S., Snyder, W. V., Stek, P. C., Thurstans, R. P., and Wagner, P. A.: Validation of Aura Microwave Limb Sounder stratospheric ozone measurements, Journal of Geophysical Research, 113, 15, doi:10.1029/2007JD008771, 2008.

Grooß J.-U. and Russell III, J. M.: Technical note: A stratospheric climatology for O_3, H_2O, CH_4, NO_x, HCl and HF derived from HALOE measurements, Atmospheric Chemistry and Physics, 5, 2797–2807, `http://www.atmos-chem-phys.net/5/2797/2005/`, 2005.

Grubbs, F. E.: On estimating precision of measuring instruments and product variability, Journal of the American Statistical Association, 43, 243–264, 1948.

Grubbs, F. E.: Errors of measurement, precision, accuracy and the statistical comparison of measuring instruments, Technometrics, 15, 53–66, 1973.

Hampel, F. R.: The influence curve and its role in robust estimation, Journal of the American Statistical Association, 69, 383–393, 1974.

Haynes, P. H., McIntyre, M. E., Shepherd, T. G., Marks, C. J., and Shine, K. P.: On the 'downward control' of extratropical diabatic circulations by eddy-induced mean zonal forces., Journal of the Atmospheric Sciences, 48, 651–680, doi:10.1175/1520-0469(1991)048, 1991.

Heard, D. E.: Analytical Techniques for Atmospheric Measurement, Blackwell Publishing Ltd., 2006.

Hegglin, M. I. and Shepherd, T. G.: O_3-N_2O correlations from the Atmospheric Chemistry Experiment: Revisiting a diagnostic of transport and chemistry in the stratosphere, Journal of Geophysical Research, 112, 19 301, doi:10.1029/2006JD008281, 2007.

Hess, P. G. and Holton, J. R.: The origin of temporal variance in long-lived trace constituents in the summer stratosphere, Journal of the Atmospheric Sciences, 42, 1455–1463, 1985.

Hitchcock, P., Shepherd, T. G., and McLandress, C.: Past and future conditions for polar stratospheric cloud formation simulated by the Canadian Middle Atmosphere Model, Atmospheric Chemistry and Physics, 9, 483–495, `http://www.atmos-chem-phys.net/9/483/2009/`, 2009.

Hocking, W. K., Thayaparan, T., and Franke, S. J.: Method for statistical comparison of geophysical data by multiple instruments which have differing accuracies, Advances in Space Research, 27, 1089–1098, doi:10.1016/S0273-1177(01)00143-0, 2001.

Holton, J. R., Haynes, P. H., McIntyre, M. E., Douglass, A. R., Rood, R. B., and Pfister, L.: Stratosphere-troposphere exchange, Reviews of Geophysics, 33, 403–439, doi:10.1029/95RG02097, 1995.

Hood, L. L.: Coupled stratospheric ozone and temperature responses to short-term changes in solar ultraviolet flux - An analysis of Nimbus 7 SBUV and SAMS data, Journal of Geophysical Research, 91, 5264–5276, doi:10.1029/JD091iD04p05264, 1986.

Hood, L. L.: Effects of solar UV variability on the stratosphere, in: Solar Variability and its Effects on Climate. Geophysical Monograph 141, edited by Pap, J. M., Fox, P., Frohlich, C., Hudson, H. S., Kuhn, J., McCormack, J., North, G., Sprigg, W., and Wu, S. T., vol. 141 of *Washington DC American Geophysical Union Geophysical Monograph Series*, p. 283, 2004.

Hood, L. L. and Zhou, S.: Stratospheric effects of 27-day solar ultraviolet variations: an analysis of UARS MLS ozone and temperature data, Journal of Geophysical Research, 103, 3629–3638, doi:10.1029/97JD02849, 1998.

Hoppel, K. W., Bowman, K. P., and Bevilacqua, R. M.: Northern hemisphere summer ozone variability observed by POAM II, Geophysical Research Letters, 26, 827–830, doi:10.1029/1999GL900118, 1999.

Huber, P.: Robust Statistics, Wiley-Interscience, 2004.

IPCC: Climate Change 2001: The Scientific Basis, Contribution of Working Group I to the Third Assessment Report of the Intergovernmental Panel on Climate Change, (Eds.) Houghton, J. T., Ding, Y., Griggs, D. J., Noguer, M., van der Linden, P. J., Dai, X., Maskell, K., and Johnson, C. A., Cambridge University Press, 2001.

IPCC: Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change, (Eds.) Solomon, S., Qin, D., Manning, M., Chen, Z., Marquis, M., Averyt, K. B., Tignor, M., and Miller, H.L., Cambridge University Press, 2007.

IPCC/TEAP (Intergovernmental Panel on Climate Change/Technology and Economic Assessment Panel): IPCC/TEAP Special Report on Safeguarding the Ozone Layer and the Global Climate System: Issues Related to Hyd

Irion, F. W., Gunson, M. R., Toon, G. C., Chang, A. Y., Eldering, A., Mahieu, E., Manney, G. L., Michelsen, H. A., Moyer, E. J., Newchurch, M. J., et al.: Atmospheric Trace Molecule Spectroscopy (ATMOS) experiment version 3 data retrievals, Applied Optics, 41, 6968–6979, 2002.

Isobe, T., Feigelson, E. D., Akritas, M. G., and Babu, G. J.: Linear regression in astronomy, The Astrophysical Journal, 364, 104–113, doi:10.1086/169390, 1990.

Jackman, C. H., DeLand, M. T., Labow, G. J., Fleming, E. L., Weisenstein, D. K., Ko, M. K. W., Sinnhuber, M., and Russell, J. M.: Neutral atmospheric influences of the solar proton events in October-November 2003, Journal of Geophysical Research, 110, 9, doi:10.1029/2004JA010888, 2005.

Jacob, D. J.: Introduction to Atmospheric Chemistry, Princeton, Univ. Press, 1999.

Jin, J. J., Semeniuk, K., Jonsson, A. I., Beagley, S. R., McConnell, J. C., Boone, C. D., Walker, K. A., Bernath, P. F., Rinsland, C. P., Dupuy, E., Ricaud, P., De La Noë, J., Urban, J., and Murtagh, D.: Co-located ACE-FTS and Odin/SMR stratospheric-mesospheric CO 2004 measurements and comparison with a GCM, Geophysical Research Letters, 32, 15, doi:10.1029/2005GL022433, 2005.

Jin, J. J., Semeniuk, K., Beagley, S. R., Fomichev, V. I., Jonsson, A. I., McConnell, J. C., Urban, J., Murtagh, D., Manney, G. L., Boone, C. D., Bernath, P. F., Walker, K. A., Barret, B., Ricaud, P., and Dupuy, E.: Comparison of CMAM simulations of carbon monoxide (CO), nitrous oxide (N_2O), and methane (CH_4) with observations from Odin/SMR, ACE-FTS, and Aura/MLS, Atmospheric Chemistry and Physics, 9, 3233–3252, `http://www.atmos-chem-phys.net/9/3233/2009/`, 2009.

Jones, D. R., Perttunen, C. D., and Stuckman, B. E.: Lipschitzian optimization without the Lipschitz constant, Journal of Optimization Theory and Applications, 79, 157–181, 1993.

Kar, J., McElroy, C. T., Drummond, J. R., Zou, J., Nichitiu, F., Walker, K. A., Randall, C. E., Nowlan, C. R., Dufour, D. G., Boone, C. D., Bernath, P. F.,

Trepte, C. R., Thomason, L. W., and McLinden, C.: Initial comparison of ozone and NO_2 profiles from ACE-MAESTRO with balloon and satellite data, Journal of Geophysical Research, 112, 16 301, doi:10.1029/2006JD008242, 2007.

Kerr, J. B. and McElroy, C. T.: Measurement of atmospheric nitrogen dioxide from the MSC stratospheric balloon program, Atmosphere, 14, 166–171, 1976.

Kerr, J. B., McElroy, C. T., and Evans, W. F. J.: Midlatitude summertime measurements of stratospheric NO_2, Canadian Journal of Physics, 60, 196–200, 1982.

Kerzenmacher, T., Wolff, M. A., Strong, K., Dupuy, E., Walker, K. A., Amekudzi, L. K., Batchelor, R. L., Bernath, P. F., Berthet, G., Blumenstock, T., Boone, C. D., Bramstedt, K., Brogniez, C., Brohede, S., Burrows, J. P., Catoire, V., Dodion, J., Drummond, J. R., Dufour, D. G., Funke, B., Fussen, D., Goutail, F., Griffith, D. W. T., Haley, C. S., Hendrick, F., Höpfner, M., Huret, N., Jones, N., Kar, J., Kramer, I., Llewellyn, E. J., López-Puertas, M., Manney, G., McElroy, C. T., McLinden, C. A., Melo, S., Mikuteit, S., Murtagh, D., Nichitiu, F., Notholt, J., Nowlan, C., Piccolo, C., Pommereau, J.-P., Randall, C., Raspollini, P., Ridolfi, M., Richter, A., Schneider, M., Schrems, O., Silicani, M., Stiller, G. P., Taylor, J., Tétard, C., Toohey, M., Vanhellemont, F., Warneke, T., Zawodny, J. M., and Zou, J.: Validation of NO_2 and NO from the Atmospheric Chemistry Experiment (ACE), Atmospheric Chemistry and Physics, 8, 5801–5841, `http://www.atmos-chem-phys.net/8/5801/2008/`, 2008.

Kondratyev, K. Y. and Varotsos, C. A.: Atmospheric Ozone Variability: Implications for Climate Change, Human Health, and Ecosystems, Springer, 2000.

Kuell, V., Olschewski, F., Jarisch, M., Offermann, D., and Grossmann, K. U.: Trace gas variability in the stratosphere, Advances in Space Research, 34, 1722–1730, doi:10.1016/j.asr.2003.05.041, 2004.

Lambert, A., Read, W. G., Livesey, N. J., Santee, M. L., Manney, G. L., Froidevaux, L., Wu, D. L., Schwartz, M. J., Pumphrey, H. C., Jimenez, C., Nedoluha,

G. E., Cofield, R. E., Cuddy, D. T., Daffer, W. H., Drouin, B. J., Fuller, R. A., Jarnot, R. F., Knosp, B. W., Pickett, H. M., Perun, V. S., Snyder, W. V., Stek, P. C., Thurstans, R. P., Wagner, P. A., Waters, J. W., Jucks, K. W., Toon, G. C., Stachnik, R. A., Bernath, P. F., Boone, C. D., Walker, K. A., Urban, J., Murtagh, D., Elkins, J. W., and Atlas, E.: Validation of the Aura Microwave Limb Sounder middle atmosphere water vapor and nitrous oxide measurements, Journal of Geophysical Research, 112, 24, doi:10.1029/2007JD008724, 2007.

Lary, D. J. and Lait, L.: Using probability distribution functions for satellite validation, IEEE Transactions on Geoscience and Remote Sensing, 44, 1359–1366, doi:10.1109/TGRS.2005.860662, 2006.

Livesey, N. J., Read, W. G., Froidevaux, L., Waters, J. W., Santee, M. L., Pumphrey, H. C., Wu, D. L., Shippony, Z., and Jarnot, R. F.: The UARS Microwave Limb Sounder version 5 data set: Theory, characterization, and validation, Journal of Geophysical Research, 108, 4378, doi:10.1029/2002JD002273, 2003.

Livesey, N. J., Read, W. G. Filipiak, M. J., Froidevaux, L., Harwood, R. S., Jiang, J. H. Jimenez, C., Pickett, H. M., Pumphrey, H. C., Santee, M. L., Schwartz, M. J., Waters, J. W., and Wu, D. L.: Version 1.5 Level 2 data quality and description document, Tech. Rep. D-32381, JPL, 2005.

Livesey, N. J., Snyder, W. V., Read, W. G., and Wagner, P. A.: Retrieval algorithms for the EOS Microwave Limb Sounder (MLS), IEEE Transactions on Geosciences and Remote Sensing: The EOS Aura Mission, pp. 1144 – 1155, doi:10.1109/TGRS.2006.872327, 2006.

López-Puertas, M., Funke, B., Gil-López, S., von Clarmann, T., Stiller, G. P., Höpfner, M., Kellmann, S., Fischer, H., and Jackman, C. H.: Observation of NO_x enhancement and ozone depletion in the Northern and Southern Hemispheres after the October-November 2003 solar proton events, Journal of Geophysical Research, 110, 9, doi:10.1029/2005JA011050, 2005a.

López-Puertas, M., Funke, B., Gil-López, S., von Clarmann, T., Stiller, G. P., Höpfner, M., Kellmann, S., Mengistu Tsidu, G., Fischer, H., and Jackman, C. H.: HNO_3, N_2O_5, and $ClONO_2$ enhancements after the October-November 2003 solar proton events, Journal of Geophysical Research, 110, 9, doi:10.1029/2005JA011051, 2005b.

Luo, M., Park, J. H., Lee, K. M., Russell III, J. M., and Bruehl, C.: An analysis of HALOE observations in summer high latitudes using airmass trajectory and photochemical model calculations, Journal of Geophysical Research, 102, 16 145, doi:10.1029/97JD00694, 1997.

Mankin, W., Atlas, E., Canrell, C., Eisele, F., and Fried, A.: Chapter 11: Observational Methods: Instruments and Platforms, in: Atmospheric Chemistry and Global Change, edited by Brasseur, G. P., Orlando, J. J., and Tyndall, G. S., Oxford University Press, 1999.

Manney, G. L., Krüger, K., Sabutis, J. L., Sena, S. A., and Pawson, S.: The remarkable 2003-2004 winter and other recent warm winters in the Arctic stratosphere since the late 1990s, Journal of Geophysical Research, 110, 4107, doi: 10.1029/2004JD005367, 2005.

Manney, G. L., Livesey, N. J., Jimenez, C. J., Pumphrey, H. C., Santee, M. L., MacKenzie, I. A., and Waters, J. W.: EOS Microwave Limb Sounder observations of "frozen-in" anticyclonic air in Arctic summer, Geophysical Research Letters, 33, 6810, doi:10.1029/2005GL025418, 2006.

Maronna, R. A., Martin, R. D., and Yohai, V. J.: Robust Statistics, J. Wiley, 2006.

McElroy, C. T., Nowlan, C. R., Drummond, J. R., Bernath, P. F., Barton, D. V., Dufour, D. G., Midwinter, C., Hall, R. B., Ogyu, A., Ullberg, A., Wardle, D. I., Kar, J., Zou, J., Nichitiu, F., Boone, C. D., Walker, K. A., and Rowlands, N.: The ACE-MAESTRO instrument on SCISAT: description, performance, and preliminary results, Applied Optics, 46, 4341–4356, doi:10.1364/AO.46.004341, 2007.

McIntyre, M. E. and Palmer, T. N.: Breaking planetary waves in the stratosphere, Nature, 305, 593–600, doi:10.1038/305593a0, 1983.

Melo, S. M. L., Blatherwick, R., Davies, J., Fogal, P., de Grandpré, J., McConnell, J., McElroy, C. T., McLandress, C., Murcray, F. J., Olson, J. R., Semeniuk, K., Shepherd, T. G., Strong, K., Tarasick, D., and Williams-Rioux, B. J.: Summertime stratospheric processes at northern mid-latitudes: comparisons between MANTRA balloon measurements and the Canadian Middle Atmosphere Model, Atmospheric Chemistry and Physics, 8, 2057–2071, `http://www.atmos-chem-phys.net/8/2057/2008/`, 2008.

Miles, T., Grose, W. L., Remsberg, E. E., and Lingenfelser, G.: Evolution of the Southern hemisphere subpolar middle atmosphere during summer and autumn, Journal of the Atmospheric Sciences, 51, 677–693, doi:10.1175/1520-0469(1994)051, 1994.

Murcray, D. G., Kyle, T. G., Murcray, F. H., and Williams, W. J.: Nitric Acid and Nitric Oxide in the Lower Stratosphere, Nature, 218, 78–79, doi:10.1038/218078a0, 1968.

Murcray, D. G., Goldman, A., Csoeke-Poeckh, A., Murcray, F. H., Williams, W. J., and Stocker, R. N.: Nitric Acid Distribution in the Stratosphere, Journal of Geophysical Research, 78, 7033–7038, doi:10.1029/JC078i030p07033, 1973.

Murcray, D. G., Barker, D. B., Brooks, J. N., Goldman, A., and Williams, W. J.: Seasonal and latitudinal variation of the stratospheric concentration of HNO_3, Geophysical Research Letters, 2, 223–225, doi:10.1029/GL002i006p00223, 1975.

Murtagh, D., Frisk, U., Merino, F., Ridal, M., Jonsson, A., Stegman, J., Witt, G., Eriksson, P., Jiménez, C., Megie, G., de La Noë, J., Ricaud, P., Baron, P., Pardo, J. R., Hauchcorne, A., Llewellyn, E. J., Degenstein, D. A., Gattinger, R. L., Lloyd, N. D., Evans, W. F. J., McDade, I. C., Haley, C. S., Sioris, C., von Savigny, C., Solheim, B. H., McConnell, J. C., Strong, K., Richardson, E. H., Leppelmeier, G. W.,

Kyrölä, E., Auvinen, H., and Oikarinen, L.: Review: An overview of the Odin atmospheric mission, Canadian Journal of Physics, 80, 309, doi:10.1139/p01-157, 2002.

Neu, J. L., Sparling, L. C., and Plumb, R. A.: Variability of the subtropical "edges" in the stratosphere, Journal of Geophysical Research, 108, 4482, doi: 10.1029/2002JD002706, 2003.

Olberg, M., Frisk, U., Lecacheux, A., Olofsson, A. O. H., Baron, P., Bergman, P., Florin, G., Hjalmarson, Å. ., Larsson, B., Murtagh, D., Olofsson, G., Pagani, L., Sandqvist, A., Teyssier, D., Torchinsky, S. A., and Volk, K.: The Odin satellite. II. Radiometer data processing and calibration, Astronomy and Astrophysics, 402, L35–L38, doi:10.1051/0004-6361:20030336, 2003.

Orsolini, Y. J.: Long-lived tracer patterns in the summer polar stratosphere, Geophysical Research Letters, 28, 3855–3858, doi:10.1029/2001GL013103, 2001.

Orsolini, Y. J., Manney, G. L., Santee, M. L., and Randall, C. E.: An upper stratospheric layer of enhanced HNO_3 following exceptional solar storms, Geophysical Research Letters, 32, 12, doi:10.1029/2004GL021588, 2005.

Park, J. H. and Russell III, J. M.: Summer polar chemistry observations in the stratosphere made by HALOE, Journal of the Atmospheric Sciences, 51, 2903–2913, doi:10.1175/1520-0469(1994)051, 1994.

Pendlebury, D., Shepherd, T. G., Pritchard, M., and McLandress, C.: Normal mode Rossby waves and their effects on chemical composition in the late summer stratosphere, Atmospheric Chemistry and Physics, 8, 1925–1935, `http://www.atmos-chem-phys.net/8/1925/2008/`, 2008.

Pick, D. R. and Houghton, J. T.: Measurements of atmospheric infrared emission with a balloon-borne multifilter radiometer, Quarterly Journal of the Royal Meteorological Society, 95, 535, 1969.

Plumb, R.: Stratospheric transport, Journal of the Meteorological Society of Japan, 80, 793–809, 2002.

Quine, B. M. and Drummond, J. R.: GENSPECT: A new generation line-by-line code with a bounded interpolation accuracy, Journal of Quantitative Spectroscopy and Radiative Transfer, 74, 147–165, 2002.

Quine, B. M., Toohey, M., Drummond, J. R., Strong, K., Wunch, D., Midwinter, C., and McElroy, C. T.: The concentration profile of nitric acid and other species over Saskatchewan in August 1998: Retrieval from data recorded by thermal-emission radiometry, Atmosphere–Ocean, 43, 361–376, 2005.

Randall, C. E., Lumpe, J. D., Bevilacqua, R. M., Hoppel, K. W., Fromm, M. D., Salawitch, R. J., Swartz, W. H., Lloyd, S. A., Kyro, E., von der Gathen, P., Claude, H., Davies, J., DeBacker, H., Dier, H., Molyneux, M. J., and Sancho, J.: Reconstruction of three-dimensional ozone fields using POAM III during SOLVE, Journal of Geophysical Research, 107, 8299, doi:10.1029/2001JD000471, 2002.

Randall, C. E., Harvey, V. L., Manney, G. L., Orsolini, Y., Codrescu, M., Sioris, C., Brohede, S., Haley, C. S., Gordley, L. L., Zawodny, J. M., and Russell, J. M.: Stratospheric effects of energetic particle precipitation in 2003-2004, Geophysical Research Letters, 32, 5802, doi:10.1029/2004GL022003, 2005a.

Randall, C. E., Manney, G. L., Allen, D. R., Bevilacqua, R. M., Hornstein, J., Trepte, C., Lahoz, W., Ajtic, J., and Bodeker, G.: Reconstruction and simulation of stratospheric ozone distributions during the 2002 Austral winter, Journal of the Atmospheric Sciences, 62, 748–764, doi:10.1175/JAS-3336.1, 2005b.

Randall, C. E., Harvey, V. L., Singleton, C. S., Bailey, S. M., Bernath, P. F., Codrescu, M., Nakajima, H., and Russell III, J. M.: Energetic particle precipitation effects on the Southern Hemisphere stratosphere in 1992-2005, Journal of Geophysical Research (Atmospheres), 112, 8308, doi:10.1029/2006JD007696, 2007.

Randel, W. J., Wu, F., Russell III, J. M., and Waters, J.: Space-time patterns of trends in stratospheric constituents derived from UARS measurements, Journal of Geophysical Research, 104, 3711–3728, doi:10.1029/1998JD100044, 1999.

Richner, H. and Viatte, P.: The hydrostatic equation in the evaluation algorithm for radiosonde data, Journal of Atmospheric and Oceanic Technology, 12, 649–656, 1995.

Ridley, B. A., Bruin, J. T., Schiff, H. I., and McConnell, J. C.: Altitude profile and sunset decay measurements of stratospheric nitric oxide, Atmosphere, 14, 180–188, 1976.

Ridley, B. A., Luu, S. H., Hastie, D. R., Schiff, H. I., McConnell, J. C., Evans, W. F. J., McElroy, C. T., Kerr, J. B., Fast, H., and Obrien, R.: Stratospheric odd nitrogen - Measurements of HNO_3, NO, NO_2, and O_3 near 54 deg N in winter, Journal of Geophysical Research, 89, 4797–4820, doi:10.1029/JD089iD03p04797, 1984.

Riese, M., Manney, G. L., Oberheide, J., Tie, X., Spang, R., and Küll, V.: Stratospheric transport by planetary wave mixing as observed during CRISTA-2, Journal of Geophysical Research, 107, 8179, doi:10.1029/2001JD000629, 2002.

Rinsland, C. P., Weisenstein, D. K., Ko, M. K. W., Scott, C. J., Chiou, L. S., Mahieu, E., Zander, R., and Demoulin, P.: Post-Mount Pinatubo eruption ground-based infrared stratospheric column measurements of HNO_3, NO and NO_2 and their comparison with model calculations, Journal of Geophysical Research, 108, 4437, doi:10.1029/2002JD002965, 2003.

Rinsland, C. P., Boone, C., Nassar, R., Walker, K., Bernath, P., Mahieu, E., Zander, R., McConnell, J. C., and Chiou, L.: Trends of HF, HCl, CCl_2 F_2, CCl_3F, $CHClF_2$ (HCFC-22), and SF_6 in the lower stratosphere from Atmospheric Chemistry Experiment (ACE) and Atmospheric Trace Molecule Spectroscopy (ATMOS) measurements near 30°N latitude, Geophysical Research Letters, 32, doi: 10.1029/2005GL022415, 2005.

Rodgers, C. D.: Inverse Methods for Atmospheric Sounding: Theory and Practice, World Scientific, 2000.

Rodgers, C. D. and Connor, B. J.: Intercomparison of remote sounding instruments, Journal of Geophysical Research, 108, doi:10.1029/2002JD002299, 2003.

Rohen, G., von Savigny, C., Sinnhuber, M., Llewellyn, E. J., Kaiser, J. W., Jackman, C. H., Kallenrode, M.-B., Schröter, J., Eichmann, K.-U., Bovensmann, H., and Burrows, J. P.: Ozone depletion during the solar proton events of October/November 2003 as seen by SCIAMACHY, Journal of Geophysical Research (Space Physics), 110, 9, doi:10.1029/2004JA010984, 2005.

Rothman, L. S., Jacquemart, D., Barbe, A., Chris Benner, D., Birk, M., Brown, L. R., Carleer, M. R., Chackerian, C., Chance, K., Coudert, L. H., et al.: The HITRAN 2004 molecular spectroscopic database, Journal of Quantitative Spectroscopy and Radiative Transfer, 96, 139–204, 2005.

Rousseeuw, P. J. and Croux, C.: Alternatives to the median absolute deviation, Journal of the American Statistical Association, 88, 1273–1283, 1993.

Russell III, J. M., Gordley, L. L., Park, J. H., Drayson, S. R., Hesketh, W. D., Cicerone, R. J., Tuck, A. F., Frederick, J. E., Harries, J. E., and Crutzen, P. J.: The Halogen Occultation Experiment, Journal of Geophysical Research, 98, 10 777, doi:10.1029/93JD00799, 1993.

Salby, M. L.: Rossby normal modes in nonuniform background configurations. Part II. Equinox and solstice conditions, Journal of the Atmospheric Sciences, 38, 1827–1840, doi:10.1175/1520-0469(1981)038, 1981.

Santee, M. L., Manney, G. L., Froidevaux, L., Read, W. G., and Waters, J. W.: Six years of UARS Microwave Limb Sounder HNO_3 observations: Seasonal, interhemispheric, and interannual variations in the lower stratosphere, Journal of Geophysical Research, 104, 8225–8246, doi:10.1029/1998JD100089, 1999.

Santee, M. L., Manney, G. L., Livesey, N. J., and Waters, J. W.: UARS Microwave Limb Sounder observations of denitrification and ozone loss in the 2000 Arctic late winter, Geophysical Research Letters, 27, 3213–3216, doi:10.1029/2000GL011738, 2000.

Santee, M. L., Manney, G. L., Livesey, N. J., and Read, W. G.: Three-dimensional structure and evolution of stratospheric HNO_3 based on UARS Microwave Limb Sounder measurements, Journal of Geophysical Research, 109, doi:10.1029/2004JD004578, 2004.

Santee, M. L., Lambert, A., Read, W. G., Livesey, N. J., Cofield, R. E., Cuddy, D. T., Daffer, W. H., Drouin, B. J., Froidevaux, L., Fuller, R. A., Jarnot, R. F., Knosp, B. W., Manney, G. L., Perun, V. S., Snyder, W. V., Stek, P. C., Thurstans, R. P., Wagner, P. A., Waters, J. W., Muscari, G., de Zafra, R. L., Dibb, J. E., Fahey, D. W., Popp, P. J., Marcy, T. P., Jucks, K. W., Toon, G. C., Stachnik, R. A., Bernath, P. F., Boone, C. D., Walker, K. A., Urban, J., and Murtagh, D.: Validation of the Aura Microwave Limb Sounder HNO_3 measurements, Journal of Geophysical Research, 112, 24, doi:10.1029/2007JD008721, 2007.

Scinocca, J. F.: An accurate spectral nonorographic gravity wave drag parameterization for general circulation models, Journal of the Atmospheric Sciences, 60, 667–682, doi:10.1175/1520-0469(2003)060, 2003.

Scinocca, J. F. and McFarlane, N. A.: The parametrization of drag induced by stratified flow over anisotropic orography, Quarterly Journal of the Royal Meteorological Society, 126, 2353–2394, doi:10.1256/smsqj.56801, 2000.

Shepherd, T. G.: Issues in stratosphere-troposphere coupling, Journal of the Meteorological Society of Japan, 80, 769–792, 2002.

Shepherd, T. G.: Large-scale atmospheric dynamics for atmospheric chemists, Chemical Reviews, 103, 4509–4532, 2003.

Shepherd, T. G.: Transport in the middle atmosphere, Journal of the Meteorological Society of Japan, 85B, 165–191, 2007.

Sparling, L. C.: Statistical perspectives on stratospheric transport, Reviews of Geophysics, 38, 417–436, doi:10.1029/1999RG000070, 2000.

Strong, K., Bailak, G., Barton, D., Bassford, M. R., Blatherwick, R. D., Brown, S., Chartrand, D., Davies, J., Drummond, J. R., Fogal, P. F., Forsberg, E., Hall, R., Jofre, A., Kaminski, J., Kosters, J., Laurin, C., McConnell, J. C., McElroy, C. T., McLinden, C. A., Melo, S. M. L., Menzies, K., Midwinter, C., Murcray, F. J., Nowlan, C., Olson, J. R., Quine, B. M., Rochon, Y., Savastiouk, V., Solheim, B., Sommerfeldt, D., Ullberg, A., Werchohlad, S., Wu, H., and Wunch, D.: Mantra-A balloon mission to study the odd-nitrogen budget of the stratosphere, Atmosphere–Ocean, 43, 283–299, doi:10.3137/ao.430401, 2005.

Strong, K., Wolff, M. A., Kerzenmacher, T. E., Walker, K. A., Bernath, P. F., Blumenstock, T., Boone, C., Catoire, V., Coffey, M., De Mazière, M., Demoulin, P., Duchatelet, P., Dupuy, E., Hannigan, J., Höpfner, M., Glatthor, N., Griffith, D. W. T., Jin, J. J., Jones, N., Jucks, K., Kuellmann, H., Kuttippurath, J., Lambert, A., Mahieu, E., McConnell, J. C., Mellqvist, J., Mikuteit, S., Murtagh, D. P., Notholt, J., Piccolo, C., Raspollini, P., Ridolfi, M., Robert, C., Schneider, M., Schrems, O., Semeniuk, K., Senten, C., Stiller, G. P., Strandberg, A., Taylor, J., Tétard, C., Toohey, M., Urban, J., Warneke, T., and Wood, S.: Validation of ACE-FTS N_2O measurements, Atmospheric Chemistry and Physics, 8, 4759–4786, `http://www.atmos-chem-phys.net/8/4759/2008/`, 2008.

Taylor, J. R.: An Introduction to Error Analysis, University Science Books, 1997.

Toohey, M.: Retrieval of atmospheric trace gas profiles from low-spectral-resolution emission radiometers, Master's report, University of Toronto, Department of Physics, 2003.

Toohey, M. and Strong, K.: Estimating biases and error variances through the com-

parison of coincident satellite measurements, Journal of Geophysical Research, 112, doi:10.1029/2006JD008192, 2007.

Toohey, M., Quine, B. M., Strong, K., Bernath, P. F., Boone, C. D., Jonsson, A. I., McElroy, C. T., Walker, K. A., and Wunch, D.: Balloon-borne radiometer measurements of Northern Hemisphere mid-latitude stratospheric HNO_3 profiles spanning 12 years, Atmospheric Chemistry and Physics, 7, 6075–6084, `http://www.atmos-chem-phys.net/7/6075/2007/`, 2007.

Trepte, C. R. and Hitchman, M. H.: Tropical stratospheric circulation deduced from satellite aerosol data, Nature, 355, 626–628, doi:10.1038/355626a0, 1992.

Urban, J., Murtagh, D., Lautié, N., Barret, B., Dupuy, É., , de La Noë, J., Eriksson, P., Frisk, U., Jones, A., Flochmoën, É. L., Olberg, M., Piccolo, C., Ricaud, P., , and Rösevall, J.: Odin/SMR limb observations of trace gases in the polar lower stratosphere during 2004-2005, in: Atmospheric Science Conference, vol. 628 of *ESA Special Publication*, 2006.

Urban, J., Pommier, M., Murtagh, D. P., Santee, M. L., and Orsolini, Y. J.: Nitric acid in the stratosphere based on Odin observations from 2001 to 2007 – Part 1: A global climatology, Atmospheric Chemistry and Physics Discussions, 8, 9569–9590, `http://www.atmos-chem-phys-discuss.net/8/9569/2008/`, 2008.

von Clarmann, T.: Validation of remotely sensed profiles of atmospheric state variables: strategies and terminology, Atmospheric Chemistry and Physics, 6, 4311–4320, `http://www.atmos-chem-phys.net/6/4311/2006/`, 2006.

Wagner, R. E. and Bowman, K. P.: Wavebreaking and mixing in the Northern Hemisphere summer stratosphere, Journal of Geophysical Research, 105, 24 799–24 808, doi:10.1029/2000JD900320, 2000.

Walker, K. A., Randall, C. E., Trepte, C. R., Boone, C. D., and Bernath, P. F.: Initial validation comparisons for the Atmospheric Chemistry Experiment (ACE-FTS), Geophysical Research Letters, 32, L16S04, doi:10.1029/2005GL022388, 2005.

Waters, J. W., Read, W. G., Froidevaux, L., Jarnot, R. F., Cofield, R. E., Flower, D. A., Lau, G. K., Pickett, H. M., Santee, M. L., Wu, D. L., Boyles, M. A., Burke, J. R., Lay, R. R., Loo, M. S., Livesey, N. J., Lungu, T. A., Manney, G. L., Nakamura, L. L., Perun, V. S., Ridenoure, B. P., Shippony, Z., Siegel, P. H., Thurstans, R. P., Harwood, R. S., Pumphrey, H. C., and Filipiak, M. J.: The UARS and EOS Microwave Limb Sounder (MLS) experiments., Journal of the Atmospheric Sciences, 56, 194–218, doi:10.1175/1520-0469(1999)056, 1999.

Waters, J. W., Froidevaux, L., Harwood, R. S., Jarnot, R. F., Pickett, H. M., Read, W. G., Siegel, P. H., Cofield, R. E., Filipiak, M. J., Flower, D. A., Holden, J. R., Lau, G. K., Livesey, N. J., Manney, G. L., Pumphrey, H. C., Santee, M. L., Wu, D. L., Cuddy, D. T., Lay, R. R., Loo, M. S., Perun, V. S., Schwartz, M. J., Stek, P. C., Thurstans, R. P., Boyles, M. A., Chandra, K. M., Chavez, M. C., Chen, G.-S., Chudasama, B. V., Dodge, R., Fuller, R. A., Girard, M. A., Jiang, J. H., Jiang, Y., Knosp, B. W., LaBelle, R. C., Lam, J. C., Lee, K. A., D. Miller, J. E. O., Patel, N. C., Pukala, D. M., Quintero, O., Scaff, D. M., Snyder, W. V., Tope, M. C., Wagner, P. A., and Walch, M. J.: The Earth Observing System Microwave Limb Sounder (EOS MLS) on the Aura satellite, IEEE Transactions on Geoscience and Remote Sensing, 44, 1075–1092, doi:10.1109/TGRS.2006.873771, 2006.

Weatherhead, E. C. and Andersen, S. B.: The search for signs of recovery of the ozone layer, Nature, 441, 39–45, doi:10.1038/nature04746, 2006.

Weatherhead, E. C., Reinsel, G. C., Tiao, G. C., Meng, X.-L., Choi, D., Cheang, W.-K., Keller, T., DeLuisi, J., Wuebbles, D. J., Kerr, J. B., Miller, A. J., Oltmans, S. J., and Frederick, J. E.: Factors affecting the detection of trends: Statistical considerations and applications to environmental data, Journal of Geophysical Research, 103, 17 149–17 162, doi:10.1029/98JD00995, 1998.

Weeks, L. H., Cuikay, R. S., and Corbin, J. R.: Ozone Measurements in the Mesosphere During The Solar Proton Event of 2 November 1969., Journal of the Atmospheric Sciences, 29, 1138–1142, doi:10.1175/1520-0469(1972)029, 1972.

WMO: Scientific Assessment of Ozone Depletion: 2002, World Meteorological Organization Global Ozone Research and Monitoring Project–Report No. 47, 2003.

WMO: Scientific Assessment of Ozone Depletion: 2006, World Meteorological Organization Global Ozone Research and Monitoring Project–Report No. 50, 2007.

Wolff, M. A., Kerzenmacher, T., Strong, K., Walker, K. A., Toohey, M., Dupuy, E., Bernath, P. F., Boone, C. D., Brohede, S., Catoire, V., von Clarmann, T., Coffey, M., Daffer, W. H., De Mazière, M., Duchatelet, P., Glatthor, N., Griffith, D. W. T., Hannigan, J., Hase, F., Höpfner, M., Huret, N., Jones, N., Jucks, K., Kagawa, A., Kasai, Y., Kramer, I., Küllmann, H., Kuttippurath, J., Mahieu, E., Manney, G., McElroy, C. T., McLinden, C., Mébarki, Y., Mikuteit, S., Murtagh, D., Piccolo, C., Raspollini, P., Ridolfi, M., Ruhnke, R., Santee, M., Senten, C., Smale, D., Tétard, C., Urban, J., and Wood, S.: Validation of HNO_3, $ClONO_2$, and N_2O_5 from the Atmospheric Chemistry Experiment Fourier Transform Spectrometer (ACE-FTS), Atmospheric Chemistry and Physics, 8, 3529–3562, `http://www.atmos-chem-phys.net/8/3529/2008/`, 2008.

Wunch, D., Tingley, M. P., Shepherd, T. G., Drummond, J. R., Moore, G. W. K., and Strong, K.: Climatology and predictability of the late summer stratospheric zonal wind turnaround over Vanscoy, Saskatchewan, Atmosphere-Ocean, 43, 301–313, 2005.

York, D.: Least-squares fitting of a straight line, Canadian Journal of Physics, 44, 1079, 1966.