DIGITAL ACCESS TO SCHOLARSHIP AT HARVARD
DASH.HARVARD.EDU

HARVARD LIBRARY
Office for Scholarly Communication

The Sources and Significance of Stratospheric Water Vapor: Mechanistic Studies from Equator to Pole

Citation

Smith, Jessica Birte. 2012. The Sources and Significance of Stratospheric Water Vapor: Mechanistic Studies from Equator to Pole. Doctoral dissertation, Harvard University.

Permanent link

http://nrs.harvard.edu/urn-3:HUL.InstRepos:10121968

Terms of Use

This article was downloaded from Harvard University's DASH repository, and is made available under the terms and conditions applicable to Other Posted Material, as set forth at http://nrs.harvard.edu/urn-3:HUL.InstRepos:dash.current.terms-of-use#LAA

Share Your Story

The Harvard community has made this article openly available.
Please share how this access benefits you. Submit a story .

Accessibility

©2011 Jessica Birte Smith

All rights reserved.

Professor James G. Anderson Jessica Birte Smith

The Sources and Significance of Stratospheric Water Vapor: Mechanistic Studies from Equator to Pole

Abstract

It is the future of the stratospheric ozone layer, which protects life at Earth's surface from harmful ultraviolet (UV) radiation, that is the focus of the present work. Fundamental changes in the composition and structure of the stratosphere in response to anthropogenic climate forcing may lead to catastrophic ozone loss under current, and even reduced, stratospheric halogen loading.

In particular, the evolution toward a colder, wetter stratosphere, threatens to enhance the heterogeneous conversion of inorganic halogen from its reservoir species to its catalytically active forms, and thus promote in situ ozone loss. Water vapor concentrations control the availability of reactive surface area, which facilitates heterogeneous chemistry. Furthermore, the rates of the key heterogeneous processes are tightly controlled by the ambient humidity. Thus, credible predictions of UV dosage require a quantitative understanding of both the

sensitivity of these chemical mechanisms to water vapor concentrations, and an elucidation of the processes controlling stratospheric water vapor concentrations.

Toward this end, we present a set of four case studies utilizing high resolution in situ data acquired aboard NASA aircraft during upper atmospheric research missions over the past two decades.

1) We examine the broad scale humidity structure of the upper troposphere and lower stratosphere from the midlatitudes to the tropics, focusing on cirrus formation and dehydration at the cold-point tropical tropopause. The data show evidence for frequent supersaturation in clear air, and sustained supersaturation in the presence of cirrus. These results challenge the strict thermal control of the tropical tropopause.

2) We investigate the likelihood of cirrus-initiated activation of chlorine in the midlatitude lower stratosphere. At midlatitudes the transition from conditions near saturation below the local tropopause to undersaturated air above greatly reduces the probability of heterogeneous activation and in situ ozone loss in this region.

3) We probe the details of heterogeneous processing in the wintertime Arctic vortex, and find that in situ measurements of OH provide incontrovertible evidence for the heterogeneous reaction of HOCl with HCl. This reaction is critical to sustaining catalytically active chlorine and prolonging ozone loss in the springtime vortex.

4) We revisit the topic of midlatitude ozone loss with an emphasis upon the response of ozone in this region to changes in the chemical composition and

thermal structure of the lower stratosphere induced by anthropogenic climate change. Specifically, we show evidence for episodic moisture plumes in the overworld stratosphere generated by the rapid evaporation of ice injected into this region by deep convection, and find that these high water vapor plumes have the potential to alter the humidity of the lower stratosphere, and drastically increase the rate of heterogeneous chemistry and in situ ozone loss, given sufficient reactive surface.

Table of Contents

Abstract **iii**

Table of Contents **vi**

List of Figures xiii

List of Table xviii

Dedication **xix**

1 Introduction **1**

1.1 Population, Energy Demand, and Climate 2

1.1.1 Population Growth and Energy Demand 2

1.1.2 Energy Demand & Carbon Dioxide Emissions 5

1.1.3 Carbon Dioxide Forcing 8

1.1.4 Feedbacks in Earth's Climate 9

1.2 Science Motivated by Societal Objectives 11

1.2.1 UV Flux & Stratospheric Ozone 12

1.2.2 Stratospheric Ozone Loss 13

1.2.3 Ozone Recovery in a Changing Climate 14

1.2.4 The Importance of Stratospheric Water Vapor 17

1.2.5 How Water Vapor Enters the Stratosphere 21

1.2.6 The Future of Stratospheric Water Vapor 25

1.2.7 Water Vapor Measurement Uncertainties....26
1.3 Outline of Dissertation....29
References....33

2 The HWV Instrument: Principle of Detection 40

2.1 Introduction....40
2.2 Photo-fragment Fluorescence Detection....41
2.3 HWV Instrument Description....43
2.3.1 Detection Axis....43
2.3.2 Aircraft Interface....50
2.4 HWV Measurement & Performance....52
2.4.1 Raw Data to Measured Mixing Ratio....52
2.4.2 Instrument Performance Characteristics....55
2.4.3 Recent Instrument Improvements – Lyman-α Filer....56
References....58

3 The HWV Instrument: Calibration & Validation 59

3.1 Introduction....59
3.2 Calibration/Validation Methodology....59
3.3 Laboratory Calibration....62
3.3.1 Selection and Verification of a Calibration Standard....63
3.3.2 Laboratory Calibrations at In Situ [H_2O], P and T....71
3.4 In-Flight Diagnostic Tests....81
3.4.1 Calibration Verification: Absorption vs. Fluorescence....81

3.4.2 In Situ Contamination Assessment: Velocity Scan 84

3.4.3 Accuracy of In-Cloud Data: Sublimation Analysis 86

3.5 In Situ Instrument Comparisons 93

3.5.1 HWV vs. HTW 94

3.5.2 HWV vs. HOxotope and ICOS 96

3.5.3 HWV vs. JLH 99

3.6 Use of a priori Knowledge as a Constraint 100

3.7 Conclusions 102

References 103

4 Investigation of Measurement Offsets 105

4.1 Introduction 105

4.2 Evaluation of Water Contamination 106

4.3 Evaluation of Instrumental Artifact 107

4.3.1 Offsets Associated with Lamp Scatter ($\Delta SCT_{out\text{-in}}$) 110

4.3.1.1 Theory 110

4.3.1.2 Laboratory Quantification 111

4.3.1.3 Effect on Measurement 114

4.3.2 Spurious Fluorescence (FLR_X) 116

4.3.2.1 Theory 116

4.3.2.2 Laboratory Quantification 126

4.3.2.3 Summary 137

4.4 Offset-Free Axis 138

4.5 Evidence For/Against an Offset in the Flight Data 139

4.5.1 Evidence Supporting an Instrumental Offset in Flight 141
4.5.2 Evidence Opposing an Instrumental Offset in Flight 142
References 148

5 Aqua Validation and Instrument Tests (AquaVIT) 149

5.1 Introduction 149
5.2 AquaVIT Campaign 150
5.2.1 AquaVIT Structure 150
5.2.2 The AIDA Chamber as a Reference Standard 153
5.2.3 The AIDA Chamber as a Common Source 156
5.3 Performance of HWV During AquaVIT 159
5.3.1 Evaluation of the HWV Calibration 162
5.3.2 Quantifying the Contributions due to Leaks and Outgassing 171
5.3.3 Evaluation Instrument Offsets 182
5.4 AquaVIT Mission Summary and Conclusions 186
5.5 Thoughts for Future Laboratory Campaigns 190
References 193

6 The Supersaturation Puzzle: An In Situ Measurement Perspective 194

6.1 Introduction 194
6.2 Relative Humidity 195
6.3 The Data Set 196
6.4 RH_i Data Accuracy 197
6.4.1 Accuracy of In Situ Water Vapor Measurements: Clear Air 199

6.4.2 Accuracy of In Situ Water Vapor Measurements: In Cloud.....199

6.4.3 Accuracy of Airborne Temperature Measurements.....199

6.5 Clear Air / In-Cloud Discrimination.....202

6.6 RH_i Results Overview.....210

6.7 Clear Air RH_i.....215

6.7.1 Clear Air RH_i: Ice Supersaturated Regions (ISSRs).....215

6.7.2 Clear Air RH_i: Ice Nucleation Mechanisms.....222

6.7.3 Clear Air RH_i: Dependence of Temperature & H_2O.....228

6.8 In-Cloud RH_i.....234

6.8.1 In-Cloud RH_i: Persistent Ice Supersaturation.....234

6.8.2 In-Cloud RH_i: Dependence of Temperature & H_2O.....238

6.9 Discussion & Implications.....240

6.9.1 Clear Air RH_i.....241

6.9.2 In-Cloud RH_i.....248

6.10 Conclusions.....251

References.....254

7 Mechanisms of Midlatitude Ozone Loss: Heterogeneous Chemistry in the Lowermost Stratosphere? 259

7.1 Introduction.....259

7.2 Chemical Processes of Rate-Limiting Radicals.....262

7.3 Susceptibility of Midlatitude Lower Stratospheric Free Radical Concentrations to Cirrus Clouds.....267

7.4 Evidence From In Situ Aircraft Observations.....269

7.4.1 Cloud Occurrence Frequency 270
7.4.2 Water Vapor and Temperature in the Lower Stratosphere 274
7.4.3 Discussion of the *Solomon et al.* [1997] Analysis 279
7.4.4 Particle Reactive Surface Area 284
7.4.5 Reaction Probability on Cold Aerosols.....285
7.4.6 Observations of ClO in the Vicinity of the Tropopause 289
7.5 Summary and Conclusions 292
References.....293

8 In Situ Observations of OH: The Signature of Heterogeneous Processing in the Polar Vortex 297

8.1 Introduction.....297
8.2 The Chemical Context 300
8.2.1 Heterogeneous Chlorine Activation and O_3 Loss 300
8.2.2 HO_x and Cl_x.....301
8.3 The Signature of Heterogeneous Chemistry.....303
8.4 The Proposed Chemical Mechanism 305
8.5 Heterogeneous Rates in the Late Winter Vortex 309
8.5.1 Reactive Surface Area.....310
8.5.2 Heterogeneous Rates on Solid NAT 312
8.5.3 Heterogeneous Rates for Liquid Binary & Ternary Aerosol 313
8.6 Model Versus Measurement 314
8.6.1 Model 1: Heterogeneous Loss Along the Flight Track.....314
8.6.2 Model 2: Heterogeneous Loss the Night Prior 320

8.6.3 Additional Considerations 322
8.7 Conclusions 327
References 328

9 Revisiting Ozone Chemistry in the Midlatitude Lower Stratosphere 332

9.1 Introduction 332
9.2 Halogen Photochemistry 334
9.3 Stratospheric Moisture 342
9.3.1 Moist Plumes 342
9.3.2 Inefficient Dehydration in the TTL 349
9.3.3 Water Vapor Accuracy 351
9.4 Heterogeneous Chemistry 353
9.4.1 Response to High Water Vapor & Cold Temperature 353
9.4.2 Probing Heterogeneous Chemistry with HO_x 355
9.5 Reactive Surface Area 357
9.6 Summary 359
References 361

List of Figures

1.1 Historical growth of population and global energy consumption.....4

1.2 Rise in atmospheric CO_2 concentrations.....6

1.3 Observed and projected CO_2 emissions.....7

1.4 Inter-annual variability in Arctic ozone depletion and stratospheric temperatures.....15

1.5 Comparison of water vapor measurements in the upper troposphere and lower stratosphere.....27

2.1 Diagram of the Lyman-α detection axis.....43

2.2 Bandpass of the PMT collection optics.....47

2.3 HWV instrument integrated into the WB-57 aircraft.....51

2.4 Progression from raw PMT counts to measured mixing ratio.....54

3.1 Chain linking the in situ measurement to a reference standard.....60

3.2 Schematic illustrating the use of independent reference standards.....65

3.3 Schematic of HWV calibration bench.....66

3.4 Data from a representative calibration run.....73

3.5 Linearity of calibration data.....76

3.6 Laboratory calibration data acquired at different temperatures.....80

3.7 Simultaneous fluorescence and radial absorption measurements.....83

3.8 Data from a representative in situ diagnostic velocity scan.....85

3.9 Representative particle size distributions for thick cirrus.....89

3.10 Cloud particle sublimation model results 89

3.11 Comparison of HWV and JLH clear air and in-cloud data.....93

3.12 Comparison of HWV and HTW clear air data.....95

3.13 Comparison of HWV and HOXotope and ICOS data 99

3.14 Measurements of RH_i in a wave cloud.....101

4.1 Raw PMT counts versus time 109

4.2 Data from measurements of the chopper window reflectivity.....112

4.3 Low water vapor calibration of HWV-E 119

4.4 Normalized counts per ppmv water vapor in air and nitrogen.....124

4.5 Photograph of laboratory calibration setup.....126

4.6 Simultaneous low water calibration of HWV and HOXotope.....127

4.7 Zero-intercept data from two diagnostic tests.....129

4.8 Zero-intercept test results in the context of a simple model 130

4.9 Offset model results 133

4.10 Zero-intercept measurements for two different axis configurations....135

4.11 Zero-intercept differences for two different axis configurations.....136

4.12 Ambient mixing ratio and absolute difference between HWV and HTW in clear air 143

4.13 Ambient mixing ratio, percent difference and absolute difference between a suite of water vapor instruments aboard the WB-57 146

5.1 AIDA chamber sampling configuration during AquaVIT.....151

5.2 Representative AquaVIT daily comparison run 152
5.3 Examination of AquaVIT sampling homogeneity.....157
5.4 Summary of HWV calibration during AquaVIT 163
5.5 Fractional differences among the core instruments at high water 165
5.6 Offset diagnostic data and model results at varying flows 173
5.7 Offset diagnostic data and model results at varying pressures 175
5.8 Diagnostic data acquired while sampling from the AIDA chamber....176
5.9 Estimated offset correction and uncertainty vs. chamber pressure.....178
5.10 Differences between HWV and AIDA TDL measurements raw and corrected.....179
5.11 Residual differences between HWV and the other core instruments ..184
6.1 Comparison of in situ cloud property measurements 206
6.2 Temperature, water vapor and RH_i dataset overview 213
6.3 Empirical distribution functions of RH_i in clear air and cloud for CRYSTAL-FACE.....214
6.4 Case-study of contrail-seeded cirrus.....219
6.5 Normalized frequency distribution of clear air RH_i by mission 225
6.6 Clear air RH_i as a function of temperature.....229
6.7 Case-study of water vapor and temperature dependence of clear air RH_i distribution.....233
6.8 Empirical distribution functions of RH_i in-cloud by mission 235
6.9 In-Cloud RH_i as a function of temperature 236
6.10 In-Cloud RH_i as a function of temperature and cloud IWC.....237

6.11 Case-study of in-cloud RH_i variability....239

7.1 Seasonal mean trend in midlatitude ozone from 1970 – 1996....261

7.2 Mean ozone trend and aircraft sampling area superimposed on a schematic of the structure and dynamics of the UT/LS....262

7.3 Modeled ClO as a function of day of year and altitude....269

7.4 Histogram of ice saturation frequency in percent....271

7.5 Representative vertical profiles of water vapor, saturation mixing ratio, temperature, and saturation temperature in the midlatitude UT/LS....275

7.6 Normalized histograms of degrees from saturation in the tropopause region....276

7.7 Near tropopause temperature and mixing ratio data plotted on a water vapor/ice phase diagram....278

7.8 Convolution of ER-2 humidity data with a function capturing both natural tropopause variability and the SAGE II vertical resolution....282

7.9 Particle surface area density as a function of altitude....285

7.10 Heterogeneous reaction probabilities and sulfuric acid weight percent as a function of altitude....287

7.11 First-order time constants for heterogeneous processing as a function of altitude....288

7.12 Stratospheric ClO observations plotted as a function of altitude....290

8.1 Chemical and physical properties of sunrise flight segment....304

8.2 OH as a function of solar zenith angle....305

8.3 Schematic of HO_x chemistry in normal and perturbed airmasses....307

8.4 Surface area density and temperature along flight track.....311

8.5 Reaction probability for HCl + HOCl as a function of temperature....313

8.6 Measured and modeled OH at sunrise 316

8.7 First-order lifetime of OH with respect to gas phase and heterogeneous reactions 319

8.8 Back-trajectory solar zenith angle and temperature data.....320

8.9 Measured and modeled OH at sunrise 322

9.1 First-order lifetime of ozone as a function of ClO mixing ratio 335

9.2 Modeled fractional odd-oxygen loss by different catalytic cycles 341

9.3 Stratospheric water vapor mixing ratio as a function of potential temperature at midlatitudes.....343

9.4 Vertical profiles of water vapor, carbon dioxide and ozone.....344

9.5 Vertical profile of water vapor color-coded by δD.....345

9.6 Hydration event in midlatitude lower stratosphere color-coded by fractional convective influence.....348

9.7 First-order lifetime of $ClONO_2$ as a function of temperature and water vapor mixing ratio.....355

List of Tables

2.1 Performance characteristics for the WB-57 HWV instrument 56

3.1 Calibration/Validation methodology for HWV 61

3.1 Uncertainty associated with the bubbler standard 68

3.2 Uncertainty associated with the VUV absorption standard 71

4.1 Pressure, flow, and temperature dependence of offset terms.....125

4.2 Versions of HWV by mission 141

5.1 Summary of core instrument measurement comparisons at high water during AquaVIT 167

5.2 Performance characteristics of HWV during AquaVIT.....189

6.1 RH_i analysis mission names and dates 197

6.2 Performance characteristics of MMS and P/T instruments 200

6.3 Cloud property instruments and criteria for airmass discrimination ...203

6.4 Altitude zone and frequency of ISSRs and clouds by mission 216

To Joseph N. Demusz

A man of integrity, good humor, and passion, who always believed in the value of his life and work.

&

To Gabrielle Giffords

She embodies the best of our representative democracy.

Chapter 1

Introduction

"You and I, my dear friend, have been sent into life at a time when the greatest lawgivers of antiquity would have wished to live. How few of the human race have ever enjoyed an opportunity of making an election of government, more than of air, soil, or climate, for themselves or their children!"

– John Adams, The Foundation of Government

[*McCullough*, 2001]

In the two centuries since the founding of the United States and John Adams penned these words, we have come to recognize that the air, soil and climate are not beyond the influence of the human race, but rather, are exceptionally vulnerable to human activity. Indeed, we live at a time, unprecedented in human history, where the scale of the human endeavor has the power to propel Earth's climate into a vastly different state than the one we comfortably inhabit at present, at a rate orders of magnitude faster than the dominant geologic and solar processes that have been shown to control climate over most of Earth's history. Thus, we live in a time when we must make an election of policy in order to preserve and protect the quality of the air, soil and

climate for future generations. We must work toward a rigorous understanding of how our current choices will drive the evolution of the climate system. Though achieving this understanding is a truly multi-faceted and global endeavor, a detailed knowledge of the coupling between atmospheric chemistry and climate, is a necessary and critical step in the process.

1.1 Population, Energy Demand, and Climate

1.1.1 Population Growth and Energy Demand

In order to quantify the present, and forecast the future impact of human activity on the Earth's climate, we begin with a modified version of a formula introduced by Ehrlich and Holdren [*Ehrlich and Holdren*, 1971; *Harrison et al.*, 2000]. Broadly, the formula describes the net global "Impact" as the product of three terms: global Population, Affluence (i.e. per person consumption or appetite), and Technology, which determines how and what resources are used per unit of consumption (IPAT). Written this way, we realize that all three terms contribute to, and have the potential to amplify, the resultant impact.

We focus first upon assessing the scale of current global energy demand. Within the IPAT framework, the rate of worldwide energy consumption [Joules/year] may be represented by the multiplication of three primary components: world population [#] × globally averaged gross domestic product (GDP) per capita per year [$\$\cdot\#^{-1}\cdot year^{-1}$] × energy consumption per dollar [$J\cdot\$^{-1}$] [*Lewis and Nocera*, 2006], *and references therein*. In the year 2000, for example, the global population was just over six billion, globally averaged GDP was

~$5000 per person per year, and global energy consumption per dollar, also referred to as energy intensity, was ~1.5 x 10^7 (J/$) [*Lewis and Nocera*, 2006], *and references therein*; [http://data.worldbank.org/indicator/NY.GDP.PCAP.CD (Retrieved November 3, 2010)]. These present day values yield a rate of energy consumption equal to (6×10^9 people) × ($5000 per person per year globally averaged) × (1.5×10^7 J/$) = 0.4×10^{21} J/year , or 0.4 Zeta Joules (ZJ)/year to fuel the global human enterprise.

Next we examine how we expect these numbers to evolve over the next half century. The growth in global population ("P") has been staggering, as is evident in Figure 1. From the emergence of modern humans ~160,000 years agoto the beginning of the Common Era the world's population grew to ~2.5 million, and it was not until the time of the American Revolution, i.e., another ~1800 years, that it reached its first billion. Over the next 200 years the population growth rate exploded, and by 1975 the world's population had grown to ~4 billion, achieving another fourfold increase in less than ~$1/10^{th}$ the time. Despite declines in the population growth rate over the past few decades, we are currently adding approximately one billion people to the planet every 15 years, and a moderate future projection estimates that the world's population will be ~9×10^9 people in 2050 [*Harrison et al.*, 2000], *and references therein*. Thus, the rate of energy consumption by mid-century due to population growth alone is expected to grow to ~0.6 ZJ/year.

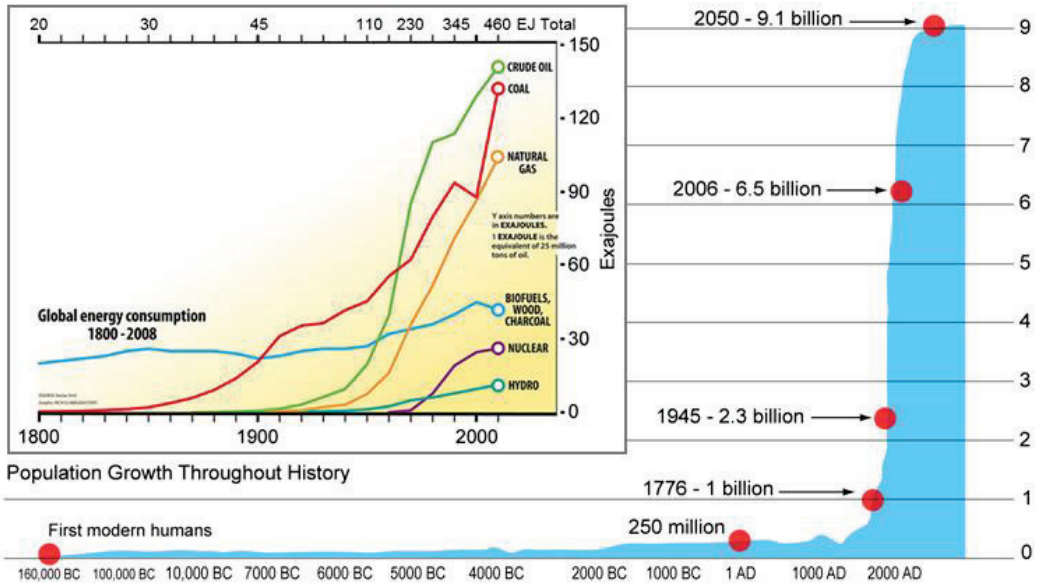


Figure 1.1. Global population and global energy consumption throughout history. The primary figure shows the world's population (billions) as a function of time from the emergence of modern humans in 160,000 BCE through to an estimated 9.1 billion in 2050. The inset shows a plot of global energy consumption in Exajoules (10^{18} J) by energy source from 1800 to 2000. Selected annual totals are indicated in the bar along the top. [Figure adapted from: [*Gore*, 2006] and [*Kurczy*, 2010], Retrieved September 15, 2011, from http://www.csmonitor.com/var/ezflow_site/storage/images/media/global-energy-consumption-1800-2008/8985558-1-eng-US/Global-energy-consumption-1800-2008_full_600.jpg]

Furthermore, it is generally assumed that per capita income ("A") will also grow over the next few decades, particularly with the increased technological and economic development of China and India. As a result, even with anticipated declines in average global energy intensity, T, resulting from technological advancements in energy generation and efficiency, the per person multiplier ("AT") is expected to increase from the current level of [($5000 per person per year globally averaged) × (1.5×10^7 J/$)] = 2.4 Terra Watts (TW) per billion people to ~4 TW per billion people in 2050, pushing global energy demand to ~1.0 ZJ/year by mid-century [*Lewis and Nocera*, 2006]. Thus, despite

uncertainties in the estimates of these terms, and the fact that the terms cannot truly be decoupled from one another, the scale of global energy demand can be expected to roughly double in the next 50 years.

1.1.2 Energy Demand & Carbon Dioxide Emissions

A factor not explicitly included in the formula as it is written above is the connection between the staggering growth of the world's rate of energy consumption, and the impact on Earth's climate. Carbon dioxide (CO_2) emissions associated with primary energy generation provide this link. The focus upon CO_2 emissions arises because 1) ~80% of global energy demand is currently met by the combustion of fossil fuels[1, 2], 2) the dominance of fossil fuels in global energy generation is projected to continue for decades into the future, 3) CO_2 is a greenhouse gas, i.e., it plays a key role in moderating the amount of energy available to the Earth's climate system, and 4) anthropogenically elevated CO_2 concentrations are likely to last for centuries, e.g., [*Archer et al.*, 2009; *Tans*, 1997].

Thus, while global population increased ~sixfold between the beginning of the industrial revolution (~1800) and the year 2000, and total global energy consumption grew by a factor of 15 to 20 (Figure 1.1 inset), CO_2 emissions, generated primarily from fossil fuel combustion, rose 800-fold over the same

[1] The combustion of natural gas, i.e., methane (CH_4), a representative fossil fuel, yields energy and CO_2: $CH_4 + 2 \cdot O_2 \rightarrow CO_2 + H_2O + \textit{energy}$.

[2] This dependence is evident in the inset of Figure 1.1, which shows the contribution of different energy sources to the global total. Crude oil, coal, and natural gas, all fossil fuel sources, dominate the energy mix.

period [*Harrison et al.*, 2000], *and references therein*, leading to the well-documented increase of atmospheric concentrations over pre-industrial levels, e.g., [*Keeling et al.*, 1976]. Figure 1.2 shows the variation in measured atmospheric CO_2 concentrations over geologic and human history. The dramatic rise beginning around ~1800, and continuing at an increasing rate of growth through to the present, is evident in this figure.

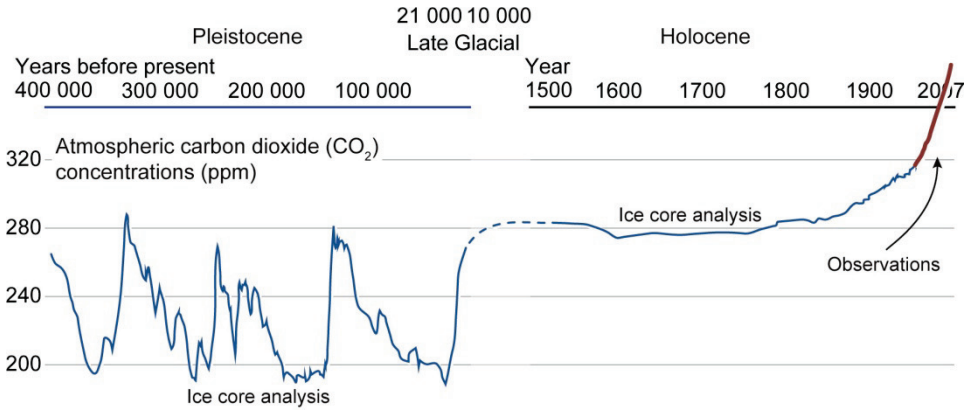


Figure 1.2. Rise in atmospheric concentrations of CO_2 over geologic and human history. Current atmospheric CO_2 concentrations are higher than the Earth has seen in over 400,000 years. Furthermore, the growth rate of CO_2 concentrations is also increasing. [Figure adapted from: Historical trends in carbon dioxide concentrations and temperature, on a geological and recent time scale. (2007). In *UNEP/GRID-Arendal Maps and Graphics Library*. Retrieved November 3, 2010, from http://maps.grida.no/go/graphic/historical-trends-in-carbon-dioxide-concentrations-and-temperature-on-a-geological-and-recent-time-scale.]

In the year 2000 atmospheric CO_2 concentrations were ~369 parts per million (ppmv), or ~31% greater than pre-industrial levels (~280 ppmv), and far beyond the natural range exhibited over the last 650,000 years [*Solomon et al.*, 2007]. This year, 2011, they have exceeded 390 ppmv, and the International Panel on Climate Change (IPCC) report of 2007 predicts that CO_2 concentrations will reach between ~450 to 550 ppmv, nearly double pre-industrial levels, by 2050.

The variability in the projected concentrations arises largely from the uncertainty in the carbon intensity of future development scenarios. However, emissions data acquired during the years since the report was released show that we have already surpassed the most fossil fuel intensive emissions forecasts, e.g., [*Allison et al.*, 2009; *Raupach et al.*, 2007].

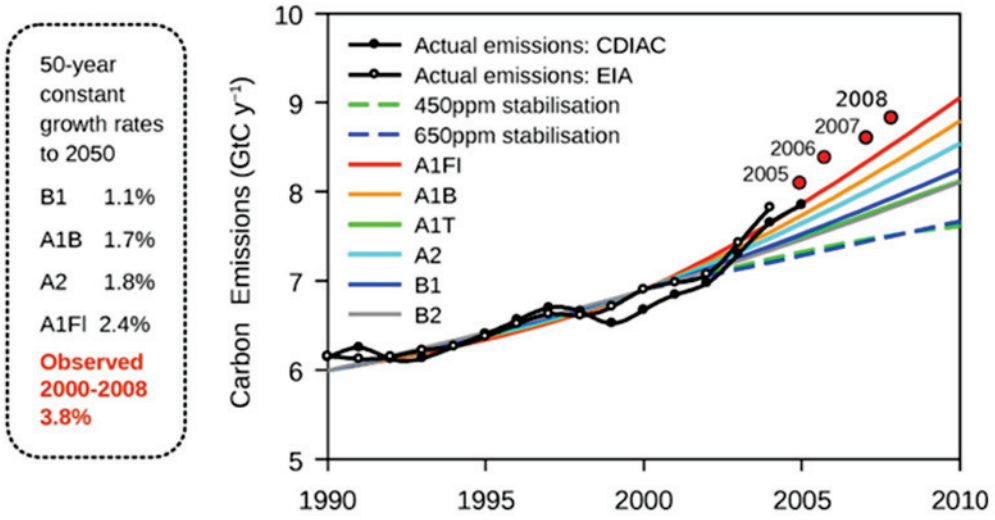


Figure 1.3. Observed and projected anthropogenic CO_2 emissions from fossil fuel sources versus time. Recent emissions data show that we have already surpassed the most fossil-fuel intensive development projections (A1F1[3]), and are far from the projected stabilization trajectories.

Figure 1.3 shows observed global CO_2 emissions compared with projected emissions adapted from a set of IPCC scenarios (represented by the colored lines)[4] [*Nakicenovic et al.*, 2000]. The figure also shows the emissions growth

[3] "In this world, people pursue personal wealth rather than environmental quality."
– IPCC SRES A1FI Emission Scenario Storyline
[http://www.ipcc-sata.org/sres/ddc_sres_emissions.html (Retrieved November 8, 2010)]

[4] Broadly, the scenarios explore different projections of greenhouse gas and aerosol emissions associated with different plausible and coherent social, economic, technological, and environmental development storylines. For detailed descriptions of each of the scenarios, refer to the IPCC Third Assessment Report (TAR) Special Report on Emissions Scenarios.

rates for two different stabilization targets of 450 ppmv and 650 ppmv respectively. Observations in recent years, derived from data provided by the US Department of Energy Carbon Dioxide Information Center (CDIAC), exceed even the most fossil-fuel intensive development scenario, A1FI. Furthermore, the results of two carbon cycle models run with the A1FI scenario (not shown) suggest that we are on target to achieve atmospheric concentrations of ~1000 ppmv, nearly quadrupling pre-industrial levels, by the end of this century [*Houghton and Intergovernmental Panel on Climate Change. Working Group I.*, 2001]. This is a sobering result, and as we shall see in subsequent sections, it has profound implications for the future of Earth's climate.

1.1.3 Carbon Dioxide Forcing

The Earth intercepts ~5500 ZJ from the Sun each year (equivalent to 340 W/m^2). Of this, ~4000 ZJ/year (~70%) are absorbed by the Earth system, and the remaining ~30% is scattered or reflected back to space by the air, aerosols, and clouds, as well as bright areas (e.g., ice) on the Earth's surface. This incoming, predominantly visible, radiation from the Sun is balanced by the return of ~4000 ZJ/year (~240 W/m^2) from the Earth to space as infrared (IR) radiation.

Trace gases in the Earth's atmosphere such as water vapor (H_2O), carbon dioxide (CO_2), and methane (CH_4), absorb and emit in the infrared, and are capable of intercepting a significant fraction of the outgoing long-wave radiation (OLR) and reradiating it back to surface. In this way, these IR absorbing "greenhouse" gases (GHGs) dictate the amount of the energy cycled within the

Earth's climate system. As GHG concentrations rise, the amount of energy retained within the climate system increases, and in order to reestablish the energy balance between incoming and outgoing radiation the Earth's surface temperature must rise. Indeed, without the absorption and reemission of IR radiation by the natural abundance of H_2O, CO_2, and other trace species in the Earth's atmosphere, the mean surface temperature would be well below freezing. Therefore, in the absence of other factors, the direct consequence of increasing CO_2 concentrations is an increase in global mean surface temperatures.

In the simplest analysis, a doubling of the atmospheric concentration of CO_2 reduces the Earth's OLR by ~4 W/m^2, and in order to restore energy balance the planet must warm by ~1.2 °C [*Ad Hoc Study Group on Carbon*, 1979; *Watson et al.*, 2001]. This effect can be quantified using the following equation: $\Delta T_s = \lambda \Delta F$, which relates changes in the Earth's surface temperature at radiative equilibrium, ΔT_s, to changes in the Earth's radiative energy budget associated with an external perturbation, ΔF, e.g., increasing concentrations of GHGs. The term ΔF is referred to as a "radiative forcing" and it is usually measured in W/m^2 at the top of the atmosphere. The proportionality factor, λ, captures the sensitivity of the Earth's climate to perturbations and has units of K/(W/m^2). In the "no feedback" case described above, the value of λ is equal to 0.3 K/(W/m^2).

1.1.4 Feedbacks in Earth's Climate

Feedbacks within the Earth's climate system complicate the determination of the Earth's climate sensitivity to increasing CO_2 concentrations, as well as the

projected rate of climate change. Feedbacks act to either amplify or diminish an initial forcing and its associated temperature increase. Models of the Earth's climate show that the response of water vapor to increasing CO_2 and other GHGs constitutes one of the dominant positive feedbacks [*Held and Soden*, 2000]. The Clausius-Clapeyron equation describes the relationship between the equilibrium vapor pressure of water over liquid as a function of temperature. Specifically, the partial pressure of water increases nearly exponentially with increases in temperature. Thus, a warmer atmosphere can accommodate higher water vapor concentrations, and because water vapor acts as a greenhouse gas, higher water vapor concentrations lead to further warming. General circulation models (GCMs) that include this and other feedbacks within the climate system estimate λ's equal to ~0.8 K/(W/m^2), and generate ~3 °C of warming for a doubling of CO_2 from 280 ppmv, with much of the additional warming attributable to the positive climate feedback associated with increasing water vapor concentrations [*IPCC*, 2001].

Feedbacks also complicate and amplify the range of forecast impacts. The concern is not simply over the rise of global mean surface temperatures, but a spans wide range of more immediate and potentially extreme outcomes, e.g., 1) the evolution of regional-scale temperatures and atmospheric circulation patterns, which are coupled to precipitation and drought patterns, 2) the frequency and magnitude of extreme weather events, 3) ocean acidification, and 4) sea level rise resulting from a combination of thermal expansion, and irreversible ice loss from the Earth's glacial systems and polar ice sheets. Furthermore, changes induced in

the large scale stratospheric circulation and/or increases in convective activity that can penetrate into the stratosphere can lead to radical changes in stratospheric chemistry and dynamics. These changes have severe implications for the future of stratospheric ozone and the flux of harmful ultraviolet (UV) radiation reaching the Earth's surface.

Thus, where we were once at the mercy of shifts in climate due to natural forces beyond our control, over the past two centuries we have become the dominant drivers of climate change. The explosive growth in population, made possible by advances in technology and ultimately propelled by fossil fuel combustion has been achieved at the price of altering, perhaps irrevocably, the air, soil, ocean and climate, with significant consequences for present and future generations.

1.2 Science Motivated by Societal Objectives

A primary motive for Earth science research is to seek to understand the Earth system in support of clear societal objectives. In order to make informed decisions for the preservation of human health and prosperity for future generations, we must be able to predict climatic and environmental parameters of critical societal significance, e.g., regional temperatures and precipitation patterns, fresh water resources, sea level, air quality, and UV flux with confidence and reliability.

1.2.1 UV Flux & Stratospheric Ozone

Within this context, the research presented here is focused upon elucidating the processes that dictate the flux of UV radiation reaching Earth's surface. UV-B (280 nm < λ < 320 nm) absorption by living organisms causes damage at the molecular level by splitting DNA base pairs, and increases in UV-B flux at the surface have been shown to increase the incidence of human skin cancer, as well as to adversely affect terrestrial plants and aquatic ecosystems, e.g., [*WMO*, 2003]. Because stratospheric ozone protects life on Earth by absorbing this harmful UV radiation, accurate forecasts of UV dosage depend upon a detailed understanding of the chemical and dynamical processes, which dictate ozone abundance and spatial distribution, and of the response of these processes to changing climate conditions. It is imperative, therefore, that we understand 1) the large scale processes that drive the stratospheric circulation and consequently determine the distribution of critical long-lived species, e.g., ozone, water vapor, and ozone depleting substances, e.g., the chlorine and bromine containing chlorofluorocarbons (CFCs) and hydrochlorofluorocarbons (HCFCs), 2) the dynamical mechanisms that result in the irreversible transport of both long- and short-lived chemical species between the troposphere and the stratosphere, 3) the chemical mechanisms, including both the gas phase and surface catalyzed heterogeneous components of ozone production and loss as a function of altitude, latitude and season, and 4) the combined chemical and dynamical processes that control the availability of reactive surface area.

1.2.2 Stratospheric Ozone Loss

Stratospheric ozone is constantly being produced, primarily through the UV initiated photolysis of O_2: $UV + O_2 \rightarrow O + O$, followed by the reaction of O with O_2: $2{\cdot}O + 2{\cdot}O_2 \rightarrow 2{\cdot}O_3$. Its abundance is determined by the balance between its photochemical production and photochemical loss processes. Catalytic loss cycles involving various radical families, e.g., HO_x (OH, HO_2) and NO_x (NO, NO_2), constitute some of the most important natural loss mechanisms[5].

Halogen radicals, e.g., Cl_x (Cl, ClO) and Br_x (Br, BrO), resulting primarily from the UV photolysis of halogen source gases (e.g. CFCs), also contribute significantly to the photochemical loss of ozone. Increases in the halogen loading of the stratosphere, as a result of anthropogenic emissions of ozone depleting source gases, have led to the observed global decrease in stratospheric ozone over the past several decades. Accordingly, surface UV-B radiation has increased above naturally occurring amounts over the same period [*WMO*, 1999].

Though a global phenomenon, halogen catalyzed loss is particularly dramatic over the poles during winter and spring, where radical concentrations are elevated by chemical processes occurring on the surfaces of cold aerosol and polar stratospheric clouds (PSCs)[6]. The rapid heterogeneous conversion of inorganic chlorine from its reservoir species (e.g. HCl and $ClONO_2$) to its

---

[5] For example: $OH + O_3 \rightarrow HO_2 + O_2$
$\underline{HO_2 + O_3 \rightarrow OH + 2O_2}$
Net: $2{\cdot}O_3 \rightarrow 3{\cdot}O_2$

[6] There are three subtypes of PSC: type 1a PSCs consist of solid nitric acid tri-hydrate (NAT) particles ($3{\cdot}\ HNO_3{:}H_2O$); type 1b exist as liquid supercooled ternary solution (STS) droplets, i.e. mixtures of nitric acid, sulfuric acid and water ($HNO_3{\cdot}H_2SO_4{\cdot}H_2O$); and type 2 consist of solid water ice particles.

catalytically active forms, greatly amplifies ozone depletion in these airmasses. The sensitivity of Antarctic ozone to anthropogenic emissions of chemical compounds containing chlorine and bromine is well documented, e.g., [*WMO*, 1994; 1999; 2003]. Indeed, the seasonal destruction of ozone over the South Pole during the past four decades, resulting in the "Ozone Hole," is one of the most vivid demonstrations of the impact of human activity on the natural environment.

1.2.3 Ozone Recovery in a Changing Climate

The internationally agreed upon Montreal Protocol, initiated in 1987, and its subsequent amendments, has led to the progressive phase out of ozone depleting substances. As a result, the forecast, in the absence of other factors, is for ozone recovery over the next fifty years [*WMO*, 2003], limited only by the removal of stratospheric halogen compounds via the slow return of air to the troposphere and elimination through precipitation and deposition at the Earth's surface. However, the atmospheric halogen burden is not the only factor. Ozone levels are also sensitive to 1) changes in stratospheric dynamics, e.g., stronger and colder winter vortices, and/or variability in the strength of the stratospheric overturning circulation, which transports ozone from its production region in the tropics to mid and high latitudes [*Hegglin and Shepherd*, 2009], and 2) changes in stratospheric chemistry, via increases in stratospheric water vapor concentrations, which can significantly enhance the probability of surface catalyzed chlorine activation, [*Kirk-Davidoff et al.*, 1999], and/or increases in stratospheric concentrations of short-lived bromine compounds which can amplify gas phase

loss rates [*Salawitch et al.*, 2005]. Furthermore, if, in an effort to counteract global warming, we pursue the proposed geoengineering scheme of injecting sulfur into the stratosphere, we can expect not only a delayed recovery, but also increased loss, particularly in the Arctic, over the next century [*Tilmes et al.*, 2008].

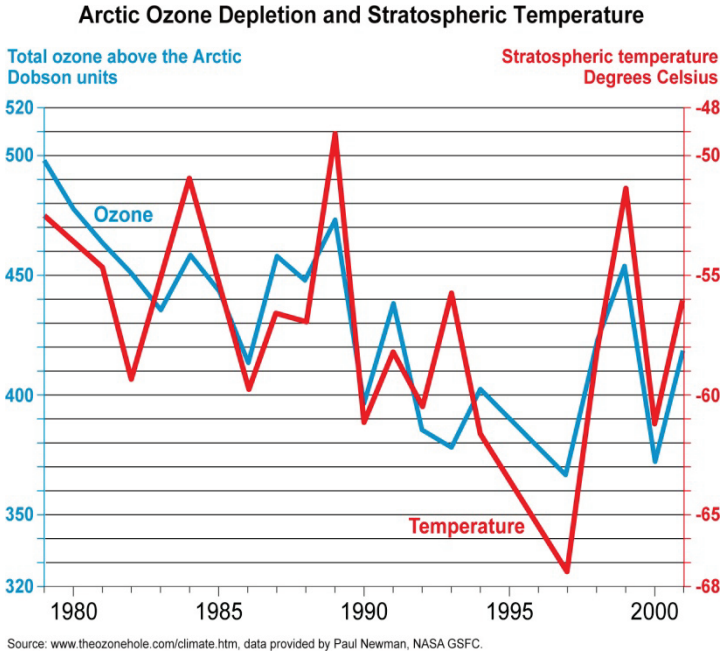


Figure 1.4. Coupling between Arctic ozone depletion and stratospheric temperatures. [Figure adapted from: Arctic ozone depletion and stratospheric temperature. (2007). In *UNEP/GRID-Arendal Maps and Graphics Library*. Retrieved November 3, 2010, from http://maps.grida.no/go/graphic/arctic-ozone-depletion-and-stratospheric-temperature.]

The Arctic stratosphere, in fact, provides an especially instructive (and sobering) model for the tight coupling between ozone chemistry and climate. Ozone losses in the Arctic polar vortex exhibit extreme interannual variability. This variability is highly correlated with vortex temperatures, as is shown in Figure 1.4. The temperature dependence arises because the frequency and

persistence of cold vortex temperatures dictates the availability of PSCs and/or cold liquid aerosol, which facilitate the heterogeneous reactions that amplify the concentrations of catalytically active chlorine and bromine free-radical species. Thus, the magnitude of stratospheric ozone loss in any given Arctic winter is directly related to the frequency of occurrence of temperatures below a critical threshold for PSC formation, T_{PSC}.

The sustained low temperatures required for PSC formation depend upon 1) the ambient humidity, which determines T_{PSC}[7], e.g. [*Hanson and Mauersberger*, 1988; *Tabazadeh et al.*, 1997], and 2) the strength and duration of the vortex, which depends upon stratospheric dynamics. At present, the observed variability in Arctic ozone loss is largely due to interannual changes in vortex dynamics. However, this example highlights the extreme sensitivity of ozone loss to conditions that favor PSC formation, and demonstrates that subtle changes in stratospheric temperatures, humidity and/or aerosol loading can have a profound impact on future ozone concentrations.

Therefore, while ozone recovery is anticipated over the next several decades in accordance with the projected decrease in the stratospheric burden of chlorine and bromine, the rate and extent of recovery depends upon the evolution and impact of a number of additional factors. In particular, the sensitivity of ozone loss rates to stratospheric temperatures and humidity, makes the future of stratospheric water vapor pivotal to any discussion of ozone chemistry [*Kirk-Davidoff et al.*, 1999].

[7] Higher water vapor partial pressures correspond to higher nucleation threshold temperatures.

1.2.4 Importance of Stratospheric Water Vapor

In the present work we are motivated by the role of stratospheric water vapor in facilitating the catastrophic loss of stratospheric ozone at present and in the future.

There are three primary ways in which stratospheric water vapor can impact ozone chemistry:

1. Stratospheric water vapor, via its reaction with O^1D, is a significant source of the hydroxyl (OH) radical in both the troposphere and the stratosphere[8]. As mentioned earlier, stratospheric OH and HO_2 radicals are involved in one of the primary gas-phase catalytic loss cycles of ozone. Model studies have suggested that increases in stratospheric water vapor can enhance the efficiency of the HO_x cycle and thus amplify ozone loss throughout the stratosphere, e.g., [*Dvortsov and Solomon*, 2001; *Stenke and Grewe*, 2005].
2. Ambient water vapor concentrations determine the threshold temperature for rapid aerosol growth and PSC formation via the nucleation of ice and other solid hydrates, which, as we have discussed, facilitate the rapid conversion of chlorine from its reservoir species to its radical forms, e.g., [*Kirk-Davidoff et al.*, 1999]. Changes by as little as +1 ppmv to current stratospheric mixing ratios, which measure ~5 ppmv, can raise the ice formation threshold by ~+1.0 K, and similarly raise the nitric acid trihydrate (NAT) threshold by +0.5 to +1.0 K [*Hanson and Mauersberger*,

[8] OH production: $O^1D + H_2O \rightarrow 2 \cdot OH$

1988; *Tabazadeh et al.*, 1997]. A recent model study examining ozone loss in the arctic vortex shows that the formation of solid NAT and ice particles is enhanced, and the length of time favoring PSC formation is extended, in response to increases in stratospheric water vapor mixing ratios [*Vogel et al.*, 2011]. Thus, increases in stratospheric water concentrations can intensify ozone loss by enhancing the availability of PSCs and cold aerosol surface area.

3. The heterogeneous reaction probabilities (γ's) are kinetic values that quantify the probability of the irreversible reactive uptake of a trace gas species on a condensed surface in the event of a collision. The reactions of chlorine and bromine reservoir species, $ClONO_2$, HOCl and $BrONO_2$, on sulfuric acid aerosol in the stratosphere are highly dependent on the ambient humidity. This dependence arises because of the sensitivity of γ to aerosol composition. The reaction probability increases as the ambient humidity increases and the acid weight percent of the aerosol decreases. Therefore, increases in stratospheric water vapor can substantially increase the probability of heterogeneous activation independent of aerosol availability.
4. Finally, stratospheric water vapor, like other GHGs, radiatively cools the stratosphere while warming the troposphere, e.g., [*Forster and Shine*, 2002; *Solomon et al.*, 2010] *and references therein*. As a result, increases in stratospheric water vapor concentrations have the potential to lower

stratospheric temperatures, further promoting PSC formation, aerosol activation and associated ozone loss.

The humidity response of the upper tropical troposphere and lower stratosphere (UT/LS) is also significant within the broader context of atmospheric chemistry and climate. As mentioned earlier, the response of water vapor to an initial warming induced by anthropogenic forcing constitutes one of the most important positive feedbacks in the climate system. In particular, changes in the flux of outgoing longwave radiation have been shown to be especially sensitive to water vapor concentrations in the middle and upper troposphere [*Held and Soden*, 2000]. Despite its low concentrations in this region, the radiative impact of water vapor in the cold, upper troposphere is disproportionately large because the thermal emission at ambient temperatures is significantly less than that from the warmer atmosphere and surface below. Higher than expected humidities in the tropical upper troposphere reduce the outgoing longwave radiative flux at the top of the atmosphere, relative to a standard, sub-saturated relative humidity profile. One model study shows, for example, that ~10% changes in the relative humidity of upper troposphere yield changes in OLR on the order of 1 W/m^2, [*Jensen et al.*, 1999] *and references therein*.

Furthermore, the potential for temperature driven increases in water vapor concentrations in the upper troposphere is amplified by the fact that for a uniform temperature change, the fractional change in water vapor partial pressure is larger at colder temperatures. Thus, an increase of ~1 °C at the surface (~20 °C) yields a

~+6% change in the ambient saturation mixing ratio, but a ~+16% change at temperatures representative of the UT (~−70 °C). In addition, there is reason to expect that the warming in the UT will be larger than that in the lower troposphere because of the forecast evolution of the moist adiabatic lapse rate [*Held and Soden*, 2000] *and references therein*.

Water vapor in the tropical upper troposphere is also tightly coupled to the growth of supercooled liquid aerosol, and the formation and coverage of thin (often sub-visible) cirrus clouds, which play an important role in the chemistry, dynamics and radiative budget of this region, e.g., [*Solomon et al.*, 2010] *and references therein*. The role of clouds in the climate system is complicated because clouds are capable of both reflecting a portion of the short-wave incoming solar radiation (albedo effect) and absorbing a portion of the long-wave outgoing infrared radiation from Earth's surface and atmosphere (greenhouse effect). The net effect of clouds depends upon the sum of these two components and varies greatly as a function of their altitude and their visible optical depth. It is generally accepted that for thin, often sub-visible cirrus, the absorption in the IR outweighs the scattering and reflectance of the sun's incident radiation [*Jensen et al.*, 1999; *Jensen et al.*, 2001b; *Lee et al.*, 2009]. Thus, cirrus clouds, present either as a result of in situ formation or from convective outflow, can have a significant impact on the radiative properties of this region, leading to enhanced reductions in OLR, and increases in the local radiative heating rates [*Jensen et al.*, 1999; *Stephens*, 2005].Changes in cirrus cloud coverage in the upper troposphere

and lower stratosphere (UT/LS) have the potential to amplify the effects of a changing climate.

Finally, water vapor and its isotopologues are key tracers for the mechanism(s) of stratosphere/troposphere exchange (STE) because they contain unique information about the trajectories (specifically the thermal histories) of air parcels prior to their irreversible entry into the lower stratosphere [*Brewer*, 1949; *Hanisco et al.*, 2007; *Keith*, 2000; *Mote et al.*, 1996; *Moyer et al.*, 1996; *Sayres et al.*, 2010]. Given that the majority of greenhouse gases and ozone loss radical precursors have their sources in the troposphere, identifying the primary mechanisms of STE is essential to understanding the present and predicting the future of stratospheric chemistry.

1.2.5 How Water Vapor Enters the Stratosphere

Atmospheric water vapor concentrations range from a few percent at the Earth's surface to a few parts per million in the stratosphere. Vertical transport in the troposphere is dominated by convection. In the tropics convection can rapidly move airmasses from the planetary boundary layer to near the level of neutral buoyancy (LNB), typically between 10 and 14 km, and is dehydrated via rapid condensation and precipitation. The ~14 km level (~150 hPa, ~355 K), which is just above the region of main convective outflow in the tropics, defines the base of the Tropical Transition Layer, or Tropical Tropopause layer (TTL). The TTL is a several kilometer thick layer bearing the characteristics of both the troposphere and the stratosphere, where the frequency of convection falls off dramatically, and

where vertical motions begin to be dominated by slow ascent driven by the mean stratospheric circulation, known as the Brewer-Dobson circulation, e.g., [*Brewer*, 1949; *Fueglistaler et al.*, 2009; *Holton et al.*, 1995] and references therein. The top of the TTL is set at ~18.5 km (~70 hPa, ~425 K), which is, in general, above the mean tropical cold point tropopause at ~380 K. Air detraining near the base of the TTL typically contains several tens of ppmv of water as both vapor and condensate. Thus, dehydration processes operating within the TTL, often associated with the formation of in situ cirrus, further reduce water vapor mixing ratios to the few parts per million level (mean ~3.5 – 4.0 ppmv, e.g., [*Dessler and Kim*, 1999]) observed in air that has recently entered the lower stratosphere. Once the dehydrated air has irreversibly entered the stratosphere, it is carried poleward by the mean circulation before it descends back into the troposphere at higher latitudes.

Thermal Control

[*Brewer*, 1949], first postulated that the aridity of the stratosphere is due to the dehydration of air slowly rising through the "cold trap" at the tropical tropopause. Broadly referred to as "thermal control," airmasses carried by the mean transport circulation have their water vapor mixing ratios dehydrated to the ice saturation mixing ratio (SMR_i) determined by the tropical tropopause temperature at the point of irreversible entry into the stratosphere, or some colder temperature encountered along the air parcel's trajectory. Water vapor in excess of the SMR_i condenses and precipitates out. The strong correlation observed

between the seasonal cycles in the tropical tropopause temperatures and stratospheric water vapor concentrations, termed the "tape recorder," [*Mote et al.*, 1996], provides compelling evidence for the dominant role of thermal control in setting entry water vapor concentrations, i.e., $[H_2O]_e = SMR_i(T_{min})$ [*Fueglistaler and Haynes*, 2005; *Fueglistaler et al.*, 2005]. However, there remains considerable speculation regarding the details and efficiency of the dehydration and transport processes, e.g., [*Gettelman et al.*, 2002; *Jensen and Pfister*, 2005; *Jensen et al.*, 2001a; *Weinstock et al.*, 2001], as well as the role of deep convection in delivering both dehydrated and hydrated air to the TTL and lower stratosphere [*Danielsen*, 1982; 1993; *Dessler*, 2002; *Dessler et al.*, 2007; *Hanisco et al.*, 2007; *Sherwood and Dessler*, 2001]. Frequent observations of supersaturation in both clear air and in clouds in the tropical upper troposphere, as well as indications of direct convective injection of water as ice, suggest that the mechanism of dehydration and entry into the stratosphere is more complex than the temperature controlled scenario outlined above.

Long-term records of stratospheric mixing ratios suggest that there has been an increase of ~2 ppmv in water vapor concentrations since the middle 1950s [*Kley et al.*, 2000; *Oltmans et al.*, 2000; *Rosenlof et al.*, 2001]. Though the oxidation of methane in the stratosphere is a source of water vapor, with one molecule of methane yielding two molecules of water, $CH_4 + 3/2 \cdot O_2 \rightarrow CO_2 + 2 \cdot H_2O$, increases in methane concentrations over the same period of ~0.55 ppmv can only account for half of the observed trend in water [*Kley et al.*, 2000]. Curiously, the positive decadal trend in stratospheric water vapor coincides with a

negative decadal trend in tropical tropopause temperatures [*Randel et al.*, 2009; *Seidel and Randel*, 2006]. Thus, the observation of these trends, if true, also challenges the validity of strict thermal control at the tropical tropopause, and they have led to a number of unconstrained hypotheses positing additional mechanisms for water vapor entry into the stratosphere, e.g., [*Rosenlof*, 2002; 2003; *Rosenlof and Reid*, 2008; *Rosenlof et al.*, 2001].

Convection

A growing body of both in situ and satellite datasets show that convection can penetrate into the TTL and lower stratosphere, leading to episodic enhancements in water vapor mixing ratios due to the deposition and rapid evaporation of convective ice, e.g., [*Corti et al.*, 2008; *Hanisco et al.*, 2007; *Iwasaki et al.*, 2010]. Water, and in particular, its heavy isotopologue, HDO, are uniquely sensitive tracers of convective air. Because HDO is heavier, it is preferentially removed relative to H_2O during condensation, and the ratio of HDO/H_2O decreases as air ascends from the boundary layer through the tropopause. Thus, the enrichment in HDO observed in both the extra-tropical lower stratosphere, and the tropical UT/LS provides incontrovertible evidence for the influence of convection, i.e., the rapid vertical mixing of air from near the boundary layer to high altitudes, in these regions [*Dessler and Sherwood*, 2003; *Dessler et al.*, 2007; *Hanisco et al.*, 2007; *Keith*, 2000; *Sayres et al.*, 2010]. In their quantitative analysis of in situ measurements of H_2O and HDO, [*Hanisco et al.*, 2007] conclude that deep convective events penetrating the overworld

stratosphere can contribute significantly to the water vapor budget over the mid-North American continent in summer. As a result, changes in the frequency deep of convective events in the tropics and/or midlatitudes, in response to GHG forcing, must be considered in forecasts of stratospheric water vapor.

1.2.6 Future of Stratospheric Water Vapor

Two primary scenarios that have the potential to cause increases in stratospheric water vapor concentrations include 1) a warming of the tropical cold-point tropopause and/or tropopause region, and 2) an increase in the convective injection of ice directly into the sub-saturated lower stratosphere.

The tropical cold-point tropopause may be expected to warm 1) in response to surface warming in the tropics, which is amplified in the UT by a corresponding increase in latent heat release [*Austin and Reichler*, 2008], and/or 2) as a result of a GHG induced slowing of the stratospheric overturning circulation [*Kirk-Davidoff et al.*, 2002]. The first cause of tropopause warming results from a very simple model of the evolution of the thermal lapse rate in response to increasing surface temperatures. The second cause of tropopause warming follows a chain of linked responses [*Kirk-Davidoff et al.*, 2002], beginning with amplified surface warming at the poles, which leads to a decrease in the surface equator-to-pole temperature gradient and a consequent decrease in the mean eddy kinetic energy in the midlatitude troposphere. The reduction in eddy kinetic energy causes a decrease in momentum deposition in the stratosphere associated with wave breaking, and thus a weaker mean overturning circulation.

Because tropical tropopause temperatures are inversely correlated with the strength of the circulation, e.g., [*Randel et al.*, 2006], a slower circulation leads to warmer temperatures and higher water vapor mixing ratios.

Convection that deposits ice directly into the overworld stratosphere represents a conduit for water to bypass the cold-point tropopause. Increases in convective strength associated with warmer sea surface temperatures may increase the frequency of events capable of penetrating the tropical tropopause. Similarly, increased tropospheric moisture in the extra-tropics has the potential to lead to stronger continental convection and penetration of the midlatitude tropopause. In either case, it is evident that convection provides the means for circumventing the strict thermal control of the cold-point tropical tropopause. Additionally, an increase in convective injection may enhance the delivery of short-lived ozone depleting species to the stratosphere, e.g., bromine-rich biogenic compounds associated with coastal emissions, e.g., [*Gettelman et al.*, 2009; *Hoskings et al.*, 2010; *Salawitch et al.*, 2005].

1.2.7 Water Vapor Measurement Uncertainties

Accurate and precise water vapor measurements are required 1) to dissect the mechanisms that determine stratospheric water vapor concentrations, 2) to determine the humidity conditions that are required for ice nucleation, and to sustain cloud formation, and persistence, 3) to understand the radiative and chemical impacts of UT/LS water vapor and clouds, and 4) to detect both short- and long-term variability and/or trends in UT/LS humidity. The current suite of

instruments, both in situ and remote, has not met the requirements necessitated by these scientific goals.

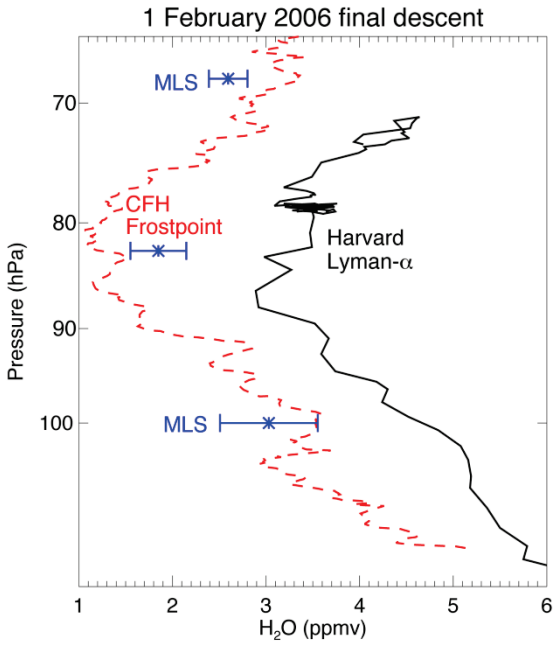


Figure 1.5. Comparison of vertical profiles of water vapor in the tropical upper troposphere. Water vapor mixing ratios measured by MLS (blue symbols), CFH (dashed red line), and HWV (solid black line), three instruments with a long and documented measurement history, are plotted as a function of pressure. The data were acquired on February 1, 2006, over San Jose, Costa Rica, with near spatial and temporal simultaneity to reduce the impact of atmospheric heterogeneity. The absolute difference between HWV and either CFH or MLS is ~1.5 ppmv ± 0.5 ppmv in the tropical UT/LS.

Figure 1.5 captures the level of discrepancy in tropical UT/LS water vapor among three well established instruments, including the Harvard Water Vapor instrument (HWV) flown aboard the WB-57, the balloon-borne National Oceanic and Atmospheric Administration (NOAA) Climate Monitoring and Diagnostics Laboratory (CMDL) Cryogenic Frostpoint Hygrometer (CFH), and the Microwave Limb Sounder (MLS) aboard the Aura satellite. In order to isolate instrumental differences from atmospheric heterogeneity, a special effort was

made to coordinate the spatial and temporal proximity of these measurements. The largest difference between HWV and CFH is approximately equal to 1.5 ppmv ±0.5 ppmv (equivalently ~75%), outside their combined uncertainty.

[*Weinstock et al.*, 2009] demonstrate that differences of this order, ~1 – 2 ppmv in the tropical upper troposphere, provide no constraint on mechanistic models for water vapor transport across the tropical tropopause. Similarly, a recent reanalysis of two long-term data sets, acquired by the NOAA CMDL Frostpoint Hygrometer (FPH) and Halogen Occultation Experiment (HALOE) aboard the Upper Atmosphere Research Satellite (UARS), shows there is little agreement about the magnitude, and even the sign, of long-term trends in stratospheric water vapor concentrations [*Scherer et al.*, 2008]. The inability to detect long-term trends, and attribute them, through model validation, severely limits our understanding of the factors controlling stratospheric water vapor in the present climate state, and prohibits any reliable forecast of stratospheric water vapor in a changing climate.

Despite the low mixing ratios of water vapor in the UT/LS (< 10 ppmv), it has been suggested that water in this region is a significant driver of climate [*Solomon et al.*, 2010]. The persistent discrepancy between water vapor measurements perpetuates uncertainty in estimates of the direct radiative forcing from stratospheric water vapor, thereby limiting the utility of such calculations until the measurement differences are resolved. Similarly, the differences provide no constraint on microphysical models of cloud/ice nucleation [*Kramer et al.*, 2009]. To paraphrase [*Kramer et al.*, 2009], a lively debate persists on whether

our understanding of ice cloud microphysics is incomplete or whether the measurements are beset by large errors. The uncertainty in ice cloud nucleation is especially deplorable given the importance of cirrus clouds to both radiative calculations and chemical models.

Thus, the current measurement discrepancies have far reaching implications for both our understanding of the present day atmosphere and for our ability to model future climate scenarios.

1.3 Outline of Dissertation

The importance of achieving accurate and precise measurements of water vapor provides the motivation for an overview of the performance of HWV. In Chapter 2, we review the Lyman-α photo-fragment fluorescence detection technique and operational details of the HWV instrument, and in Chapter 3, we provide a detailed discussion of the calibration and in-flight validation strategy. The discussion in Chapter 3 focuses on the results of recent laboratory work that targeted achieving high accuracy at low (<5 ppmv) water vapor mixing ratios. The details of the HWV calibration and validation are presented within a framework that is broadly applicable to all in situ instruments. In Chapter 4, we focus on the identification and quantification of potential systematic instrumental offsets/biases, in an attempt to address the enduring differences in flight between HWV and other water vapor sensors evident in Figure 1.5. Finally, in Chapter 5, the full calibration and validation protocol of HWV is evaluated within the context of an international laboratory comparison campaign hosted by the Aerosol

Interaction and Dynamics in the Atmosphere (AIDA) facility in Karlsruhe, Germany. The agreement between HWV and a core subset of other sensors during this campaign is typically consistent with our experimental uncertainty, and cannot explain the in situ disagreement.

In Chapter 6, we utilize HWV measurements acquired in situ during six aircraft campaigns spanning the tropics to the midlatitudes, to examine the humidity structure of the UT/LS and TTL. Observations of water vapor and relative humidity with respect to ice (RH_i) in the tropical and sub-tropical upper troposphere show that supersaturation with respect to ice is common in both clear air and in the presence of cirrus. The combination of frequent clear air supersaturation in the UT and sustained supersaturation in the presence of near tropopause cirrus challenges the theory of strict thermal control by the tropical cold-point temperature. Both results indicate that stratospheric entry level water vapor concentrations may be higher than $SMR_i(T_{min})$. A review of recent laboratory and in situ studies, that examine the effect of chemical composition on aerosol nucleation properties, offers a reasonable explanation for our in situ humidity observations. Furthermore, these studies suggest that anthropogenically induced changes in aerosol composition may be capable of affecting stratospheric water vapor concentrations.

In Chapter 7, we analyze the humidity structure of the midlatitude lower stratosphere, in combination with halogen measurements in the same region, to examine the potential for chlorine activation on cold aerosol/and or cirrus clouds. We find that the midlatitude lower stratosphere is very dry, with no evidence of

ice saturation above the first 0.5 km over the local tropopause. Cirrus lifetimes in this region are expected to be exceedingly short, and the particle data show no evidence of clouds in this region. Accordingly, the in situ ClO radical measurements are very low, except during a single encounter with an airmass with unusually high aerosol loading from the Mt. Pinatubo eruption. This study demonstrates the sensitivity of heterogeneous activation to ambient humidity and surface area availability, and provides a benchmark for future studies examining lower stratospheric humidity, aerosol loading, and halogen activation.

In Chapter 8, we analyze high resolution in situ OH data acquired during a sunrise flight aboard NASA's ER-2 aircraft in the wintertime Arctic vortex over Kiruna, Sweden in January 2000. The data provides the most direct evidence to date for the occurrence of HOCl + HCl + PSC/aerosol is $\rightarrow H_2O + Cl_2 +$ PSC/aerosol. Just prior to the aircraft encounter, the airmass in question experienced over 36 hours of darkness and coincident cooling, allowing for the formation of PSCs and the complete loss of HOCl through heterogeneous processing. Given sufficient reactive surface area and cold temperatures, this reaction prolongs chlorine activation and sustains chemical ozone loss by providing a reaction partner for HCl under the depleted $ClONO_2$ conditions characteristic of the springtime vortex. Indeed, this reaction no doubt contributed to the large losses observed during the winter of 2000, and was central to the unprecedented losses observed during the winter of 2011. The data acquired on this flight provide incontrovertible proof of this chemical loss mechanism, and

again highlight the sensitivity of ozone loss to stratospheric temperature and humidity conditions.

Finally, in Chapter 9, we revisit the topic of midlatitude ozone loss with an emphasis upon the response of ozone in this region to changes in the chemical composition and thermal structure of the lower stratosphere induced by anthropogenic climate change. We revisit the conclusions of Chapters 6, 7, and 8, and examine the potential for higher BrO concentrations in the lower stratosphere to partially resolve the observed midlatitude ozone trends. We also show evidence for the direct injection of water vapor as ice into the lower stratosphere via deep convective systems, and find that these high water vapor plumes have the potential to alter the humidity of the lower stratosphere, and drastically increase the rate of heterogeneous chemistry, given sufficient reactive surface. We discuss the sensitivity of the chemistry in this region to future perturbations arising from the positive climate forcing due to rising GHG concentrations, e.g., increases in the frequency and/or magnitude of these deep convective intrusions, as wells as from efforts to mitigate it through the intentional injection of sulfur into the lower stratosphere.

References

Ad Hoc Study Group on Carbon, D. a. C. (1979), *Carbon Dioxide and Climate: A Scientific Assessment*, National Academy of Sciences; National Research Council, Washington D.C.

Allison, I., et al. (2009), The Copenhagen Diagnosis, 2009: Updating the wrold on the Latest Climate Science*Rep.*, THe University of New South Walkes Climate Change Research Centre (CCRC), Sydney, Australia.

Archer, D., et al. (2009), Atmospheric Lifetime of Fossil Fuel Carbon Dioxide, *Annu Rev Earth Pl Sc*, *37*, 117-134.

Austin, J., and T. J. Reichler (2008), Long-term evolution of the cold point tropical tropopause: Simulation results and attribution analysis, *J Geophys Res-Atmos*, *113*.

Brewer, A. W. (1949), Evidence for a World Circulation Provided by the Measurements of Helium and Water Vapour Distribution in the Stratosphere, *Q J Roy Meteor Soc*, *75*(326), 351-363.

Corti, T., et al. (2008), Unprecedented evidence for deep convection hydrating the tropical stratosphere, *Geophys Res Lett*, *35*(10).

Danielsen, E. F. (1982), A Dehydration Mechanism for the Stratosphere, *Geophys Res Lett*, *9*(6), 605-608.

Danielsen, E. F. (1993), Insitu Evidence of Rapid, Vertical, Irreversible Transport of Lower Tropospheric Air into the Lower Tropical Stratosphere by Convective Cloud Turrets and by Larger-Scale Upwelling in Tropical Cyclones, *J Geophys Res-Atmos*, *98*(D5), 8665-8681.

Dessler, A. E. (2002), The effect of deep, tropical convection on the tropical tropopause layer, *J Geophys Res-Atmos*, *107*(D3).

Dessler, A. E., and H. Kim (1999), Determination of the amount of water vapor entering the stratosphere based on Halogen Occultation Experiment (HALOE) data, *J Geophys Res-Atmos*, *104*(D23), 30605-30607.

Dessler, A. E., and S. C. Sherwood (2003), A model of HDO in the tropical tropopause layer, *Atmos Chem Phys*, *3*, 2173-2181.

Dessler, A. E., T. F. Hanisco, and S. Fueglistaler (2007), Effects of convective ice lofting on H2O and HDO in the tropical tropopause layer, *J Geophys Res-Atmos*, *112*(D18).

Dvortsov, V. L., and S. Solomon (2001), Response of the stratospheric temperatures and ozone to past and future increases in stratospheric humidity, *J Geophys Res-Atmos*, *106*(D7), 7505-7514.

Ehrlich, P. R., and J. P. Holdren (1971), Impact of Population Growth, *Science*, *171*(3977).

Forster, P. M. D., and K. P. Shine (2002), Assessing the climate impact of trends in stratospheric water vapor, *Geophys Res Lett*, *29*(6).

Fueglistaler, S., and P. H. Haynes (2005), Control of interannual and longer-term variability of stratospheric water vapor, *J Geophys Res-Atmos*, *110*(D24).

Fueglistaler, S., M. Bonazzola, P. H. Haynes, and T. Peter (2005), Stratospheric water vapor predicted from the Lagrangian temperature history of air entering the stratosphere in the tropics, *J Geophys Res-Atmos*, *110*(D8).

Fueglistaler, S., A. E. Dessler, T. J. Dunkerton, I. Folkins, Q. Fu, and P. W. Mote (2009), Tropical Tropopause Layer, *Rev Geophys*, *47*.

Gettelman, A., W. J. Randel, F. Wu, and S. T. Massie (2002), Transport of water vapor in the tropical tropopause layer, *Geophys Res Lett*, *29*(1).

Gettelman, A., P. H. Lauritzen, M. Park, and J. E. Kay (2009), Processes regulating short-lived species in the tropical tropopause layer, *J Geophys Res-Atmos*, *114*.

Gore, A. (2006), *An inconvenient truth : the planetary emergency of global warming and what we can do about it*, 325 p. pp., Rodale Press, Emmaus, Pa.

Hanisco, T. F., et al. (2007), Observations of deep convective influence on stratospheric water vapor and its isotopic composition, *Geophys Res Lett*, *34*(4).

Hanson, D., and K. Mauersberger (1988), Vapor-Pressures of Hno3/H2o Solutions at Low-Temperatures, *J Phys Chem-Us*, *92*(21), 6167-6170.

Harrison, P., F. Pearce, and American Association for the Advancement of Science (2000), *AAAS atlas of population & environment*, xi, 204 p. pp., University of California Press, Berkeley, CA.

Hegglin, M. I., and T. G. Shepherd (2009), Large climate-induced changes in ultraviolet index and stratosphere-to-troposphere ozone flux, *Nat Geosci*, *2*(10), 687-691.

Held, I. M., and B. J. Soden (2000), Water vapor feedback and global warming, *Annu Rev Energ Env*, *25*, 441-475.

Holton, J. R., P. H. Haynes, M. E. Mcintyre, A. R. Douglass, R. B. Rood, and L. Pfister (1995), Stratosphere-Troposphere Exchange, *Rev Geophys*, *33*(4), 403-439.

Hoskings, J. S., M. R. Russo, P. Braesicke, and J. A. Pyle (2010), Modelling deep convection and its impacts on the tropical tropopause layer, *submitted to Atmos. Chem. Phys.*

Houghton, J. T., and Intergovernmental Panel on Climate Change. Working Group I. (2001), *Climate change 2001 : the scientific basis : contribution of Working Group I to the third assessment report of the Intergovernmental Panel on Climate Change*, x, 881 p. pp., Cambridge University Press, Cambridge ; New York.

IPCC (2001), *Climate Change 2001, Third Assessment Report: The Scientific Basis, 7.2.1 Physics of the Water Vapour and Cloud Feedbacks*, 1 CD-ROM pp., IPCC Secretariat, Geneva.

Iwasaki, S., T. Shibata, J. Nakamoto, H. Okamoto, H. Ishimoto, and H. Kubota (2010), Characteristics of deep convection measured by using the A-train constellation, *J Geophys Res-Atmos*, *115*.

Jensen, E. J., and L. Pfister (2005), Implications of persistent ice supersaturation in cold cirrus for stratospheric water vapor, *Geophys Res Lett*, *32*(1).

Jensen, E. J., L. Pfister, A. S. Ackerman, A. Tabazadeh, and O. B. Toon (2001a), A conceptual model of the dehydration of air due to freeze-drying by optically thin, laminar cirrus rising slowly across the tropical tropopause, *J Geophys Res-Atmos*, *106*(D15), 17237-17252.

Jensen, E. J., W. G. Read, J. Mergenthaler, B. J. Sandor, L. Pfister, and A. Tabazadeh (1999), High humidities and subvisible cirrus near the tropical tropopause, *Geophys Res Lett*, *26*(15), 2347-2350.

Jensen, E. J., O. B. Toon, S. A. Vay, J. Ovarlez, R. May, T. P. Bui, C. H. Twohy, B. W. Gandrud, R. F. Pueschel, and U. Schumann (2001b), Prevalence of ice-supersaturated regions in the upper troposphere: Implications for optically thin ice cloud formation, *J Geophys Res-Atmos*, *106*(D15), 17253-17266.

Keeling, C. D., R. B. Bacastow, A. E. Bainbridge, C. A. Ekdahl, P. R. Guenther, L. S. Waterman, and J. F. S. Chin (1976), Atmospheric Carbon-Dioxide Variations at Mauna-Loa Observatory, Hawaii, *Tellus*, *28*(6), 538-551.

Keith, D. W. (2000), Stratosphere-troposphere exchange: Inferences from the isotopic composition of water vapor, *J Geophys Res-Atmos*, *105*(D12), 15167-15173.

Kirk-Davidoff, D. B., D. P. Schrag, and J. G. Anderson (2002), On the feedback of stratospheric clouds on polar climate, *Geophys Res Lett*, *29*(11).

Kirk-Davidoff, D. B., E. J. Hintsa, J. G. Anderson, and D. W. Keith (1999), The effect of climate change on ozone depletion through changes in stratospheric water vapour, *Nature*, *402*(6760), 399-401.

Kley, D., J. M. Russell, and C. Phillips (2000), SPARC Assessment of Upper Tropospheric and Stratospheric Water Vapour*Rep.*, WCRP, WMO, Paris.

Kramer, M., et al. (2009), Ice supersaturations and cirrus cloud crystal numbers, *Atmos Chem Phys*, *9*(11), 3505-3522.

Kurczy, S. (2010), China to mold future world energy use: IEA, in *The Christian Science Monitor*, edited, The Christian Science Monitor.

Lee, J., P. Yang, A. E. Dessler, B. C. Gao, and S. Platnick (2009), Distribution and Radiative Forcing of Tropical Thin Cirrus Clouds, *J Atmos Sci*, *66*(12), 3721-3731.

Lewis, N. S., and D. G. Nocera (2006), Powering the planet: Chemical challenges in solar energy utilization, *P Natl Acad Sci USA*, *103*(43), 15729-15735.

McCullough, D. G. (2001), *John Adams*, 751 p., 740 p. of plates pp., Simon & Schuster, New York.

Mote, P. W., K. H. Rosenlof, M. E. McIntyre, E. S. Carr, J. C. Gille, J. R. Holton, J. S. Kinnersley, H. C. Pumphrey, J. M. Russell, and J. W. Waters (1996), An atmospheric tape recorder: The imprint of tropical tropopause temperatures on stratospheric water vapor, *J Geophys Res-Atmos*, *101*(D2), 3989-4006.

Moyer, E. J., F. W. Irion, Y. L. Yung, and M. R. Gunson (1996), ATMOS stratospheric deuterated water and implications for troposphere-stratosphere transport, *Geophys Res Lett*, *23*(17), 2385-2388.

Nakicenovic, N., I. P. o. C. Change., and I. P. o. C. C. W. G. III. (2000), *Special report on emissions scenarios: a special report of Working Group III of the Intergovernmental Panel on Climate Change*, 599 p. pp., Cambridge University Press, Cambridge ; New York.

Oltmans, S. J., H. Vomel, D. J. Hofmann, K. H. Rosenlof, and D. Kley (2000), The increase in stratospheric water vapor from balloonborne, frostpoint hygrometer measurements at Washington, DC, and Boulder, Colorado, *Geophys Res Lett*, *27*(21), 3453-3456.

Randel, W. J., F. Wu, H. Vomel, G. E. Nedoluha, and P. Forster (2006), Decreases in stratospheric water vapor after 2001: Links to changes in the tropical tropopause and the Brewer-Dobson circulation, *J Geophys Res-Atmos*, *111*(D12).

Randel, W. J., et al. (2009), An update of observed stratospheric temperature trends, *J Geophys Res-Atmos*, *114*.

Raupach, M. R., G. Marland, P. Ciais, C. Le Quere, J. G. Canadell, G. Klepper, and C. B. Field (2007), Global and regional drivers of accelerating CO2 emissions, *P Natl Acad Sci USA*, *104*(24), 10288-10293.

Rosenlof, K. H. (2002), Transport changes inferred from HALOE water and methane measurements, *J Meteorol Soc Jpn*, *80*(4B), 831-848.

Rosenlof, K. H. (2003), Atmospheric science - How water enters the stratosphere, *Science*, *302*(5651), 1691-1692.

Rosenlof, K. H., and G. C. Reid (2008), Trends in the temperature and water vapor content of the tropical lower stratosphere: Sea surface connection, *J Geophys Res-Atmos*, *113*(D6).

Rosenlof, K. H., et al. (2001), Stratospheric water vapor increases over the past half-century, *Geophys Res Lett*, *28*(7), 1195-1198.

Salawitch, R. J., D. K. Weisenstein, L. J. Kovalenko, C. E. Sioris, P. O. Wennberg, K. Chance, M. K. W. Ko, and C. A. McLinden (2005), Sensitivity of ozone to bromine in the lower stratosphere, *Geophys Res Lett*, *32*(5).

Sayres, D. S., L. Pfister, T. F. Hanisco, E. J. Moyer, J. B. Smith, J. M. St Clair, A. S. O'Brien, M. F. Witinski, M. Legg, and J. G. Anderson (2010), Influence of convection on the water isotopic composition of the tropical tropopause layer and tropical stratosphere, *J Geophys Res-Atmos*, *115*.

Scherer, M., H. Vomel, S. Fueglistaler, S. J. Oltmans, and J. Staehelin (2008), Trends and variability of midlatitude stratospheric water vapour deduced from the re-evaluated Boulder balloon series and HALOE, *Atmos Chem Phys*, *8*(5), 1391-1402.

Seidel, D. J., and W. J. Randel (2006), Variability and trends in the global tropopause estimated from radiosonde data, *J Geophys Res-Atmos*, *111*(D21).

Sherwood, S. C., and A. E. Dessler (2001), A model for transport across the tropical tropopause, *J Atmos Sci*, *58*(7), 765-779.

Solomon, S., I. P. o. C. Change., and I. P. o. C. C. W. G. I. (2007), *Climate change 2007: the physical science basis: contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*, viii, 996 p. pp., Cambridge University Press, Cambridge ; New York.

Solomon, S., K. H. Rosenlof, R. W. Portmann, J. S. Daniel, S. M. Davis, T. J. Sanford, and G. K. Plattner (2010), Contributions of Stratospheric Water Vapor to Decadal Changes in the Rate of Global Warming, *Science*, *327*(5970), 1219-1223.

Stenke, A., and V. Grewe (2005), Simulation of stratospheric water vapor trends: impact on stratospheric ozone chemistry, *Atmos Chem Phys*, *5*, 1257-1272.

Stephens, G. L. (2005), Cloud Feedbacks in the Climate System: A Critical Review, *J Climate*, *18*(2), 237-273.

Tabazadeh, A., O. B. Toon, and E. J. Jensen (1997), Formation and implications of ice particle nucleation in the stratosphere, *Geophys Res Lett*, *24*(16), 2007-2010.

Tans, P. P. (1997), The CO2 lifetime concept should be banished - Comment, *Climatic Change*, *37*(3), 487-490.

Tilmes, S., R. Muller, and R. Salawitch (2008), The sensitivity of polar ozone depletion to proposed geoengineering schemes, *Science*, *320*(5880), 1201-1204.

Vogel, B., T. Feck, and J. U. Grooß (2011), Impact of stratospheric water vapor enhancements caused by CH4 and H2O increase on polar ozone loss, *J. Geophys. Res.*, *116*(D5), D05301.

Watson, R. T., D. L. Albritton, D. J. Dokken, C. Intergovernmental Panel on Climate, P. United Nations Environment, and O. World Meteorological (2001), *Climate change 2001 : synthesis report*, Stand-alone ed., viii, 184 p. pp., WMO ;UNEP, Geneva, Switzerland?] :[Nairobi, Kenya?].

Weinstock, E. M., E. J. Hintsa, D. B. Kirk-Davidoff, J. G. Anderson, A. E. Andrews, R. L. Herman, C. R. Webster, M. Loewenstein, J. R. Podolske,

and T. P. Bui (2001), Constraints on the seasonal cycle of stratospheric water vapor using in situ measurements from the ER-2 and a CO photochemical clock, *J. Geophys. Res.*, *106*(D19), 22707-22724.

Weinstock, E. M., et al. (2009), Validation of the Harvard Lyman-alpha in situ water vapor instrument: Implications for the mechanisms that control stratospheric water vapor, *J Geophys Res-Atmos*, *114*.

WMO (1994), *Scientific assessment of ozone depletion, 1994*, World Meteorological Organization.

WMO (1999), *Scientific assessment of ozone depletion, 1998*, World Meteorological Organization.

WMO (2003), *Scientific assessment of ozone depletion, 2002*, World Meteorological Organization.

Chapter 2

The HWV Instrument:

Principle of Detection

2.1 Introduction

Harvard has had over twenty years experience measuring water vapor in situ in the upper troposphere and lower stratosphere. The original in situ instrument was designed for a stratospheric balloon payload, and it flew 4 flights from 1987 to 1989 [*Schwab et al.*, 1990]. In 1992, a new version was developed for integration into the nosecone of the NASA's ER-2 high-altitude research aircraft [*Weinstock et al.*, 1994]. This instrument, referred to in this document as HWV-E, regularly participated in airborne campaigns from the fall of 1992 to the winter/spring of 2000. In 1998, the most recent version of the instrument (HWV) was developed with a smaller, 2" versus 3", square duct and a light-weight design. A companion total water (vapor + condensed phase) instrument (HTW) was developed in parallel, and both instruments have been flying missions on NASA's WB-57 aircraft since 2001 [*Weinstock et al.*, 2006a; *Weinstock et al.*, 2006b]. All versions of Harvard water vapor utilize the Lyman-α (121.6 nm) initiated photo-fragment fluorescence detection method. In this chapter, we provide an

explanation of the measurement technique and a detailed description of the current generation instrument.

2.2 Photo-fragment Fluorescence Detection

The Lyman-α photo-fragment fluorescence detection method was developed for the in situ measurement of stratospheric water vapor via balloon-borne instrumentation because of its molecular specificity and high sensitivity [*Kley and Stone*, 1978] and [*Bertaux and Delannoy*, 1978]. Lyman-α radiation, generated by a small amount of hydrogen gas in a radiofrequency plasma discharge lamp, photo-dissociates water vapor in the in situ sample duct. A fraction of the resulting OH fragments are formed in their first excited electronic state ($A^2\Sigma^+$) denoted OH*. These excited state fragments either fluoresce, or are quenched by collisions with nitrogen and oxygen. The detected fluorescence signal is directly related to the number density of water vapor in the sample duct via the following series of equations.

OH Production*: $H_2O + h\nu\ (121.6\ \text{nm}) \rightarrow OH(A^2\Sigma^+) + H;\ \psi, \sigma, \varphi$ (2.1)

Fluorescence: $OH(A^2\Sigma^+) \rightarrow OH(X^2\Pi) + h\nu\ (305 - 325\ \text{nm});\ \tau \sim 10^{-6}\text{s}$ (2.2)

Quenching: $OH(A^2\Sigma^+) + M \rightarrow OH(X^2\Pi);\ k_{qair}$ (2.3)

Fraction of OH that Fluoresce*: $f_{flr} = \dfrac{\tau^{-1}}{(\tau^{-1} + k_{qair}[M])}$ (2.4)

Signal: $S_{OH} = [H_2O]\ C_{flr} = [H_2O] \left\{ C_0 \times \dfrac{1}{1 + q_{air}[M]} \right\} \sim \dfrac{C_0}{q_{air}} \times \dfrac{[H_2O]}{[M]}$ (2.5)

Equation 2.1 describes the production of OH* via the photo-dissociation of H_2O. The terms ψ, σ, and φ refer to the 121.6 nm lamp flux at the detection volume, the absorption cross section of H_2O at 121.6 nm, and the quantum yield for OH* (~10%), respectively [*Kley and Stone*, 1978]. The excited OH* can then lose its excess energy either through the loss of a photon, i.e., fluorescence at ~315 nm, or through collisional quenching with air. Equations 2.2 and 2.3 describe these loss processes, where τ is the radiative lifetime of OH*, and k_{qair} [cm^3/s] is the quenching rate coefficient for air, determined as a weighted sum of quenching due to both N_2 and O_2 [*Kley and Stone*, 1978]. Equation 2.4 combines these two loss channels to determine the fraction of OH* that fluoresces, f_{flr}.

The fluorescence signal is detected by a photomultiplier tube (PMT) positioned at right angles to the Lyman-α lamp, and Equation 2.5 describes the relationship between the number of fluorescence photons detected, S_{OH} [counts], and [H_2O], the water vapor number density in the detection volume. The proportionality constant, C_{flr} [(counts/s)/(molecules/cm^3)] implicitly includes the photo-dissociative terms described above, the factor defining the branching ratio of fluorescence, in addition to the factors that determine the collection efficiency of the detection axis such as the transmissivity of the filter assembly in front of the PMT, and the PMT quantum efficiency. A rearrangement of terms yields the two empirically determined calibration constants of the detection axis: C_0, with units of [(counts/s)/(molecules/cm^3)], defines the sensitivity of the detection axis to [H_2O] at zero air density, and $q_{air} = \tau \times k_{qair}$ [cm^3/molecule], is the empirically determined quenching factor. Equation 2.5 also shows that at the altitudes of the

upper troposphere and lower stratosphere, where the quenching term q_{air}[M] >> 1, the observed fluorescence signal is essentially proportional to the water vapor volume mixing ratio (*vmr* = [H_2O]/[M]).

2.3 HWV Instrument Description

2.3.1 Detection Axis

The mechanical and optical configurations of the HWV detection axis are shown in Figure 2.1. The PMT and collection optics are positioned at right angles to both the lamp and the flow tube, i.e., above the plane of the page. The region imaged onto the photocathode of the PMT is indicated by the circle in the center of the duct.

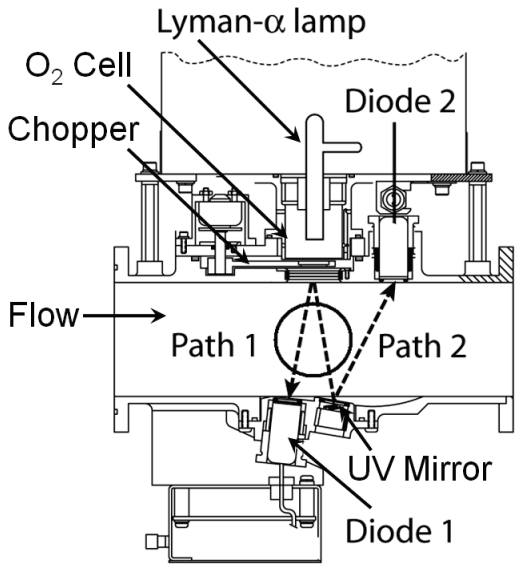


Figure 2.1. Diagram of the Harvard Lyman-α detection axis. The flow duct, Lyman-α lamp, chopper assembly, O_2 cell, vacuum ultraviolet diodes, and mirror are labeled. In this configuration, the photomultiplier and its associated collection optics are located above the plane of the page. The circle indicates the detection region imaged onto the photocathode of the PMT.

Lamp

The source of Lyman-α radiation is a sealed, low pressure, radiofrequency plasma discharge lamp in which a small amount of hydrogen gas is maintained in equilibrium with hot uranium hydride in a balance of neon. The uranium hydride, held in the heated lamp sidearm, serves as both the source of hydrogen as well as an efficient chemical getter for impurities in the lamp. The lamp body is made out of high-purity quartz tube and is sealed at the transmission end with a MgF_2 window. Positioned at the other end is a silicon photodiode, which selectively monitors the 486.1 nm hydrogen Balmer line through a combination of a narrow band (~1 nm) interference filter and 10 nm bandpass filter. The signal from the silicon photodiode is fed into a circuit controlling the heat on the uranium hydride source. The feedback circuitry is capable of maintaining the Lyman-α output stable to within ~1%.

Air Filter Cell

A flowing air cell, referred to as the "O_2 cell", is mounted between the transmission end of the lamp and the instrument duct. It functions as an optical filter for 121.6 nm light. While the spectral purity of the hydrogen lamps is generally good, with >90% of the flux between 115 nm and 150 nm resulting from 121.6 nm, there are slight impurities primarily due to the oxygen triplet at 130.4 nm and the Lyman bands of molecular hydrogen around 150 nm. N_2 is transparent in the region of interest (<150 nm), however, O_2 has an appreciable absorption cross section throughout the vacuum ultraviolet (VUV), with a notable

and fortuitous exception in a narrow band, the “oxygen window”, around 121.6 nm. The O_2 absorption cross section drops by more than an order of magnitude between 121.0 nm and 122.5 nm. Thus, with the pressure inside the cell maintained at ~1 atmosphere above ambient and a path length of ~2.5 cm between the end of the lamp and the MgF_2 window, which provides the seal to the instrument duct, the air filter cell effectively ensures a monochromatic light source [*Pittman*, 2005], *and references therein*.

VUV Flux Measurements

The VUV photodiodes serve two primary functions. The pair is used to monitor the VUV lamp flux, and also provides a complementary method for measuring water vapor in situ. The primary photodiode is positioned across the flow duct from the lamp and is used to normalize the fluorescence signal. The combination of the VUV coated mirror and second photodiode can also be used to account for fluctuations in the lamp flux, providing valuable redundancy. Furthermore, when there are short term changes in water vapor concentrations of sufficient magnitude, an independent measurement of water vapor utilizing Beer’s law and the absorption path between the lamp and the second photodiode is possible. This secondary measurement provides a means of verifying the robustness of the fluorescence calibration in situ. Examples utilizing flight data are discussed in more detail in following sections.

Though most aspects of the optical design were preserved from one generation of Harvard instruments to the next, the optical arrangement of the first

photodiode and mirror differs between the ER-2 and WB-57 detection axes in a subtle, yet potentially important way. The VUV mirror in HWV-E is 38 mm in diameter and has a 12 mm hole in its center. This allows the mirror to be fitted concentrically around the first photodiode. In this configuration the photodiode is positioned directly across the duct from the lamp and receives the majority of the collimated 121.6 nm lamp flux. In the reconfigured optical design of the WB-57 detection axis, shown in Figure 2.1, the first photodiode and mirror are mounted adjacent to one another, such that the majority of the lamp flux hits the portion of the aluminum mount between the optics. The surface of the mount is painted with a flat black paint infused with carbon black to reduce scattered light in the duct. The potential significance of this minor optical rearrangement is discussed further in Chapters 3 and 4.

PMT and Collection Optics

Photo-fragment fluorescence is a "two-color" detection technique. The generation of the excited OH* fragments is achieved with 121.6 nm light, and the red-shifted fluorescence at ~315 nm is detected. Though rotationally excited OH* is predominantly formed in the $v' = 1$ vibrational level of the $A^2\Sigma^+$ excited electronic state, it rapidly relaxes to $v' = 0$. Thus, the strongest fluorescence is from the highly excited rotational levels ($N' = 20$–22), of the ($0 \rightarrow 0$) band, at wavelengths from ~305 nm to 325 nm [*Weinstock et al.*, 1994].

The filter array in front of the PMT is designed to measure the signal within a narrow band, ~10 – 20 nm at full-width half-maximum, around 312 nm, and

minimize the transmission of light outside the bandpass, i.e., from ~200 to 650 nm. Optimally, the fluorescence signal emitted from a small detection volume (~1 cm^3) in the core of the flow is 1) collimated by a anti-reflection coated quartz condensing lens, 2) filtered through a ~10 – 20 nm bandpass interference filter centered at 313 nm, with a peak transmission of >40% and a transmission of $<10^{-5}$ of peak from 200 to 420 nm outside the region of interest, 3) passed through an additional antireflection coated UG11 visible filter that blocks light from 410 to 650 nm, 4) redirected at right angles by a mirror coated for 99% reflectivity from 300 to 320 nm, and for >95% transmission in the red and near infrared, 5) passed through a dielectric filter which blocks radiation from 650 to 750 nm, and 6) focused by a lens onto the photocathode of the PMT.

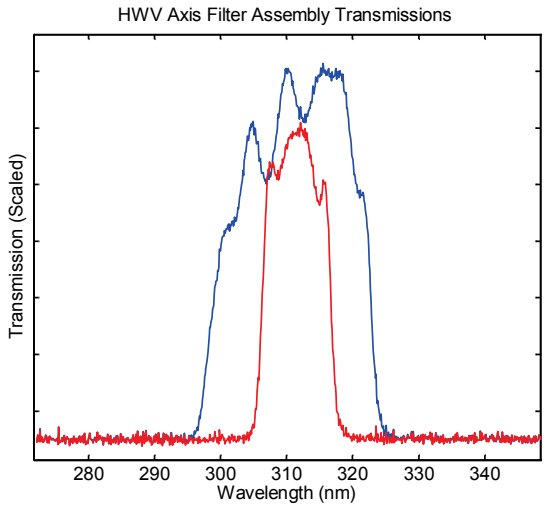


Figure 2.2. Bandpass of PMT collection optics. This figure shows the effective transmission of the optical stack in front of the PMT as a function of wavelength (nm). The two curves correspond to the two different assemblies, utilizing two narrowband filters with different transmission specifications.

Figure 2.2 shows a plot of the effective transmission as a function of wavelength for the optical filter stack. The difference between the two curves results from the choice of the narrowband filter. Light in this wavelength range has two known and easily identified sources, OH* fluorescence and light from the UV lamp which is scattered off of the detection axis surfaces and into the collection optics. Thus, the raw signal detected by the PMT is predominantly a sum of the desired fluorescence signal and UV lamp scatter.

Quartz Window Chopper

There are two sources of light in the detection band pass (~315 nm) that are not a result of VUV stimulated fluorescence and that must be quantified so that an appropriate correction to the detected PMT signal can be applied: 1) the hydrogen lamp, which has a non-negligible output in the near UV caused by trace impurities in the discharge[1], and 2) light from the sun that is scattered down the instrument duct and into the collection optics. Rayleigh scatter and scattering off of the surfaces of instrument duct from both these sources contribute to a non-zero background signal. Strategically positioned baffles in front of the lamp and in the PMT collection optics, in combination with the black coated surfaces of the duct reduce, but cannot eliminate, the contributions due to lamp and solar scatter.

The magnitude of the background signal is measured by periodically inserting an anti-reflection coated quartz window in front of the lamp. The quartz window is mounted on a mechanical actuator ("chopper") and blocks the beam for

[1] Oxygen impurities in the lamp combine with hydrogen to form OH, which can then be excited in the discharge and fluoresce.

2 – 4 seconds out of every 10 second sampling period. The window transmits ~98% of the near UV from the lamp, but effectively blocks all of the VUV radiation, thus eliminating the VUV-stimulated fluorescence signal. With the window "out" of the lamp path the PMT signal is a combination of the fluorescence signal + (lamp background + solar background). With the window "in", the signal is solely due to the non-fluorescent background, i.e., (lamp background + solar background). In theory, a simple subtraction yields the actual fluorescence signal. In practice, however, because the quartz window is not 100% transmissive of the near UV light (each surface reflects ~1%), another small correction must be applied to account for the slight reduction in the lamp's contribution to the measured background signal with the chopper "in". The quartz window does not attenuate the solar scatter contribution. Thus, in order to apply the appropriate background correction it is necessary to separately quantify the lamp and solar scatter terms.

Solar Background

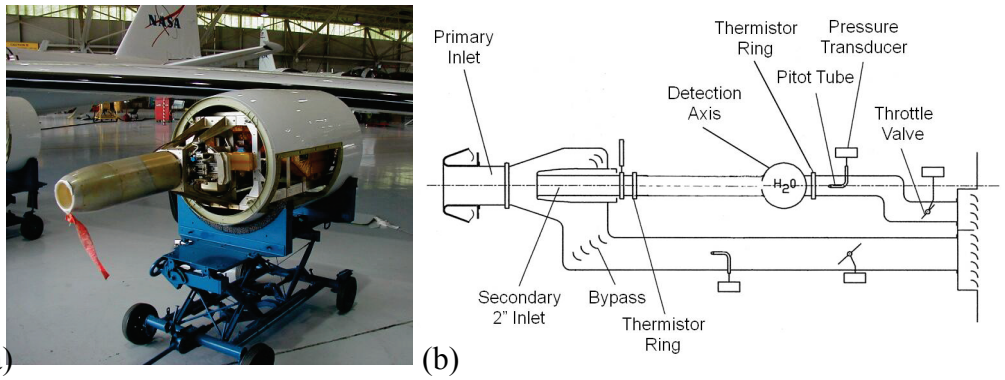
In the version of the instrument flown aboard the ER-2, HWV-E, the primary valve at the inlet to the instrument was closed for ~30 seconds, roughly every hour or two of flight. Closure of the primary valve provided a measure of the background contribution from the lamp alone, i.e., in the absence of any solar contribution. Roughly 90% of the total background signal was found to be due to solar radiation reflected off the walls and optics of detection axis, and the remaining 10% due to the lamp.

In the current generation of Harvard instruments, HWV and HTW, both ducts are configured with sections that serve as light traps, such that the solar contribution to the background in both axes is zero. The elimination of solar scatter in the 2" axes was essential because the compact design greatly increased the degree of scattered light detected by the PMT. With no solar contribution, the background signal measured with the chopper window "in" is simply equal to the contribution from the lamp.

2.3.2 Aircraft Interface

NASA's WB-57 is a mid-wing, high-altitude aircraft capable of flying for approximately 6 hours and reaching altitudes in excess of 60,000 feet (~18 km) while carrying a ~6,000 pound instrument payload. The aircraft is piloted by a crew of two, a pilot and an instrument operator. The instrument operator has on/off controls for the operation of the suite of sensors that comprise the instrument payload. The HWV instrument has participated in ~65 flights in 8 missions over the past ten years aboard the WB-57.

The HWV detection axis and ducting fit into a spear pod, which is mounted below the left wing of the WB-57 aircraft. Figure 2.3a shows a picture of the instrument integrated into the pod, and Figure 2.3b is a mechanical schematic showing the flight configuration of the sampling duct, pressure and temperature gauges, and detection axis.



(a) (b)

Figure 2.3. Photograph and diagram of HWV in its aircraft configuration. Panel (a) shows a photograph of the water vapor instrument mounted in the WB-57 spear pod. Panel (b) shows a schematic of the instrument, illustrating the subsystems that control the flow of air through the instrument as well as provide measurements of temperature, pressure, and velocity in the duct.

Air is ram-fed through a 4.5" diameter external inlet into the instrument's primary duct, and the flow is throttled down from the typical aircraft velocity of ~200 m/s to ~100 m/s. A secondary 2" square internal duct, which interfaces with the detection axis described above, samples the core of the flow. A specially shaped portion of the 2" ductwork upstream of the detection axis serves as a light trap to eliminate solar scatter from the detection region. A pitot tube mounted in the secondary duct downstream of the detection axis provides a measure of both the pressure in the detection region and the mass flow through the instrument, and the temperature is measured with a thermistor ring assembly mounted in the flow. The secondary and primary flows are recombined before being exhausted. A throttle valve in the secondary duct, controlled by a stepper motor, allows for the periodic variation of the flow velocity from ~20 to 100 m/s during flight. The ability to vary the flow is used as a diagnostic tool to test for contamination due to water vapor outgassing from the instrument walls. Typically, the fast flow

through the duct allows for a fast time response measurement, and eliminates wall effects.

2.4 HWV Measurement & Performance

In the following section, the process of converting raw PMT counts to ambient water vapor mixing ratios is described, and the instrument performance characteristics are summarized.

2.4.1 Raw Data to Measured Mixing Ratio

Equation 2.6, which is simply a rearrangement of the terms in Equation 2.5, shows the relationship between the normalized PMT counts, S_{norm}, the calibration constants, and water vapor concentrations. This equation captures the fundamental simplicity of the photo-fragment fluorescence technique, i.e., detected counts are directly proportional to water vapor concentration. Furthermore, as was stated above, throughout much of the UT/LS $1/C_0$ is <5% of q_{air}/C_0, and thus, S_{norm} is almost directly proportional to the ambient water vapor mixing ratio.

$$[\mathrm{H_2O}] = S_{norm}\left(\frac{1}{C_0} + \frac{q_{air}}{C_0}[\mathrm{M}]\right) \tag{2.6}$$

Equations 2.7 and 2.8 describe how S_{norm} is derived from the raw PMT signal. S_{out} and S_{in} in Equation 2.7 correspond to the PMT signal where "in" and "out" refer to chopper position. The signal with the chopper out is shown to be a combination of water vapor fluorescence, FLR_{H2O}, and near UV scatter from the

lamp, SCT_{out}. The signal with the chopper in is simply due to the near UV scatter from the lamp, SCT_{in}, and the factor, f_{chop}, in Equation 2.7 accounts for the slight attenuation of the near UV lamp light due to surface reflections off of the chopper window. This factor, referred to as the chopper reflectivity correction, is empirically determined and is typically ~1–2%.

$$\Delta S = S_{out} - S_{in} \times (1 + f_{chop}) = (FLR_{H2O} + SCT_{out}) - (SCT_{in}) \times (1 + f_{chop}) \quad (2.7)$$

$$S_{norm} = \frac{\Delta S}{VUV \cdot e^{\sum_X \sigma_X [X]_\ell}} \times 10^4 \quad (2.8)$$

Equation 2.8 describes the normalization of the fluorescence signal, S_{norm}. The term VUV represents the raw signal from the VUV photodiode monitoring the lamp flux, and the exponential corrects for the attenuation of the VUV flux that occurs along the path, ℓ, from the fluorescence region in the center of the detection axis to the front of photodiode. The sum over the species X in the exponential accounts for absorption due to O_2, CO_2, and CH_4 along ℓ. In the UT/LS the total correction is ~<5%. The attenuation correction for ambient H_2O is determined iteratively, and in the UT/LS it is typically <3%. The normalized signal is then multiplied by the factor containing the two empirically determined calibration constants, C_o and q_{air}, to yield the measured water vapor concentration.

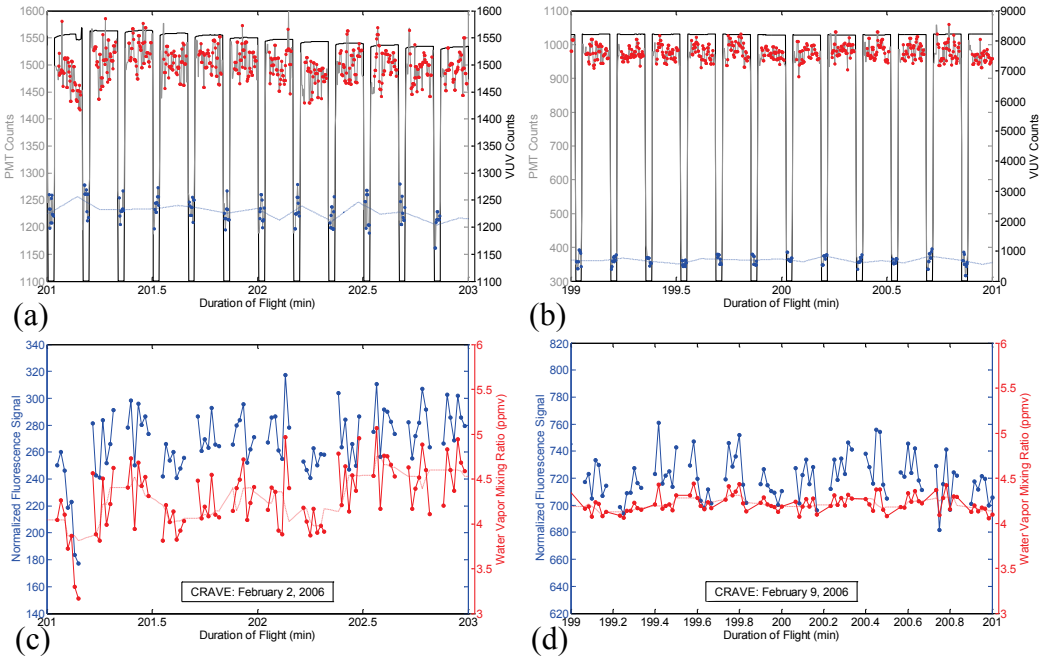


Figure 2.4. Progression from raw counts to measured mixing ratio. The two vertical columns correspond to data acquired by two different detection axes on two different flights during the CRAVE mission. Panels (a) and (b) show raw counts from the PMT, as well as raw data from the VUV diode used to normalized the fluorescence signal. Panels (c) and (d) show the corresponding normalized fluorescence signal (blue) and measured mixing ratio (red).

Figure 2.4 panels (a) and (c) show raw data from both the VUV diode (black line) used to normalize the fluorescence signal, and thus account for variability in the output of the lamp, and the PMT (gray line). The data were acquired on the flights of February 2, 2006, and February 9, 2006, respectively, during the Costa Rice Aura Validation Experiment (CRAVE), with two different detection axes. The chopper cycles are readily evident in the diode signal. The combined fluorescence + background signal measured by the PMT with the chopper "out" is highlighted in red, and the background signal measured with the chopper "in" is highlighted in blue. With the chopper "in" the VUV signal is effectively zero. Panels (b) and (d) of Figure 2.4 show the corresponding

normalized PMT signals, as well as the final water vapor mixing ratios. The measured mixing ratio in both cases is ~4.2 ppmv.

2.4.2 Instrument Performance Characteristics

The HWV instrument performance characteristics are summarized in Table 2.1. The instrument was designed with the goal of making accurate and precise measurements with a fast time response over the full range of mixing ratios encountered in the mid- to upper troposphere and lower stratosphere, i.e., from ~1 – 500 ppmv. The fast flow through the instrument duct translates into a residence time of a few hundredths of a second, and allows for a rapid response measurement. The accuracy of the measurement is primarily determined by laboratory calibrations traceable to fundamental physical properties of water (see Chapter 3). Uncertainties in the laboratory calibration are typically ≤3%. The typically high signal-to-noise ratio[2] (S/N), represented by the data acquired on the flight of February 9, 2006, shown in Figure 2.4, panels (b) and (d), allows for a precision of better than 0.1 ppmv in 1 second. Thus, we estimate the total uncertainty of the in situ measurement to be within ±5%, except where we are precision limited, i.e., at the lowest mixing ratios.

---

[2] The following describes the calculation of the signal-to-noise (S/N) ratio: $S/N = \frac{S_{out} - S_{in}}{\sqrt{S_{out} + S_{in}}}$

Feb. 2: $\frac{1500 - 1225}{\sqrt{1500 + 1225}} = 5.3 @ 8\,Hz\ for\ 4.2\ ppmv \rightarrow a\ precision\ of \pm 0.28\,ppmv @ 1Hz$

Feb. 9: $\frac{975 - 360}{\sqrt{975 + 360}} = 16.8 @ 8\,Hz\ for\ 4.2\ ppmv \rightarrow a\ precision\ of \pm 0.09\,ppmv @ 1Hz$

	1 - 10 ppmv	10 - 100 ppmv	>100 ppmv
Sampling Period	1 second	1 second	1 second
Pressure Range	<500 hPa	<500 hPa	<500 hPa
Precision (typical)	±<0.2 ppmv	±<1 – 5%	±<1%
Accuracy	±5 – 10%	±5%	±5%
Artifacts/Offset	<0.3 ppmv	<0.3 ppmv	<0.3 ppmv

Table 2.1. The instrument performance characteristics for the 2" HWV instrument that flies aboard NASA's WB-57 aircraft.

The data plotted in Figure 2.4 provide an excellent example of the range of instrument performance as dictated by instrument precision. We contrast an axis with poor S/N, (February 2, 2006), with an axis with a S/N ratio ~3x higher (February 9, 2006). Note that the calibrated accuracy of both axes is identical. In the first case, the combination of high background scatter and low lamp flux compromised the instrument's performance. The gain in the VUV electronics of this detection axis was set too high, and the signal saturated at the optimal lamp flux. The backup detection axes, installed mid-mission, achieved much improved performance for the flight of February 9, 2006, and subsequent flights. The reduced background signal and improved sensitivity is readily evident in the raw PMT data of the second axis shown in Figure 2.4 panel (c). The same vertical scale was used for measured mixing ratio in Figure 2.4 panels (b) and (d) in order to highlight the improvement in measurement precision of HWV.

2.4.3 Recent Instrument Improvements – Lyman-α Filter

The newest version of the HWV instrument utilizes a 121.6 nm filter in place of the standard MgF_2 O_2 cell window. The filter effectively ensures that all

the light entering the duct is 121.6 nm ± 15 nm (full-width half-maximum), obviating the need for the chopper. Tests in the laboratory demonstrate that the background counts are reduced to ~10 counts at 8 Hz, from ~360 counts for an optimal axis or >1000 for a poor axis (see Figure 2.4 panels (c) and (a) respectively). Because the 121.6 nm signal is also attenuated, improvements in the S/N ratio are not as remarkable, however, several other uncertainties associated with high and variable background scatter are resolved with the filter. In particular, uncertainties associated with the chopper reflectivity correction (discussed in greater detail in Chapter 4) are eliminated. Furthermore, with the filter in place, changes in the scattering properties of the axis during a mission no longer lead to degradation in performance. Transient changes in the background scatter associated with the presence of cloud particles in the duct also no longer compromise the S/N during cloud encounters. Finally, without the chopper, a continuous measurement is possible, allowing for a higher frequency data rate.

References

Bertaux, J. L., and A. Delannoy (1978), Vertical Distribution of H2O in Stratosphere as Determined by UV Fluorescence In Situ Measurements, *Geophys Res Lett*, *5*(12), 1017-1020.

Kley, D., and E. J. Stone (1978), Measurement of Water Vapor in Stratosphere by Photo-Dissociation with Ly-Alpha (1216 Å) Light, *Rev Sci Instrum*, *49*(6), 691-697.

Pittman, J. (2005), *Transport in the tropical and subtropical lower stratosphere: insights from in situ measurements of chemical tracers*.

Schwab, J. J., E. M. Weinstock, J. B. Nee, and J. G. Anderson (1990), In Situ Measurement of Water Vapor in the Stratosphere with a Cryogenically Cooled Lyman-Alpha Hygrometer, *J Geophys Res-Atmos*, *95*(D9), 13781-13796.

Weinstock, E. M., E. J. Hintsa, A. E. Dessler, J. F. Oliver, N. L. Hazen, J. N. Demusz, N. T. Allen, L. B. Lapson, and J. G. Anderson (1994), New Fast-Response Photofragment Fluorescence Hygrometer for Use on the NASA ER-2 and the Perseus Remotely Piloted Aircraft, *Rev Sci Instrum*, *65*(11), 3544-3554.

Weinstock, E. M., et al. (2006b), Measurements of the total water content of cirrus clouds. Part I: Instrument details and calibration, *J Atmos Ocean Tech*, *23*(11), 1397-1409.

Weinstock, E. M., J. B. Smith, D. Sayres, J. V. Pittman, N. Allen, and J. G. Anderson (2006a), Measurements of the total water content of cirrus clouds. Part II: Instrument performance and validation, *J Atmos Ocean Tech*, *23*(11), 1410-1421.

Chapter 3

The HWV Instrument:

Calibration & Validation

3.1 Introduction

As we emphasized in Chapter 1, accurate water vapor measurements at all levels of the atmospheric column, and in the upper troposphere and lower stratosphere in particular, are essential to understanding the factors that control the chemical composition of Earth's atmosphere, and to answering fundamental questions about the Earth's climate sensitivity. We argued that the current level of disagreement and uncertainty among water vapor measurements is hindering scientific progress, as trend detection and model validation, from the cloud scale to the climate scale, are dataset dependent. The severity of this predicament provides the impulse for the detailed discussion of the calibration and validation of the Harvard water vapor instrument (HWV) presented here.

3.2 Calibration/Validation Methodology

We begin with the assertion that the stated accuracy of a measurement is comparable to a falsifiable hypothesis. Therefore, for an accuracy claim to be

credible, the methods that led to its estimation must be documented, and must be capable of being examined, tested and confirmed by independent investigators.

The diagram in Figure 3.1 illustrates the accumulation of uncertainty in the chain linking a hypothetical in situ measurement to an accepted laboratory standard. The green ovals form the outline of a successful calibration/validation strategy. Adhering to the stated premise, they highlight the importance of end-to-end identification and documentation of the sources of uncertainty and emphasize the need for transparency at every step. The red boxes depict methodological defects that jeopardize the credibility of the final measurement.

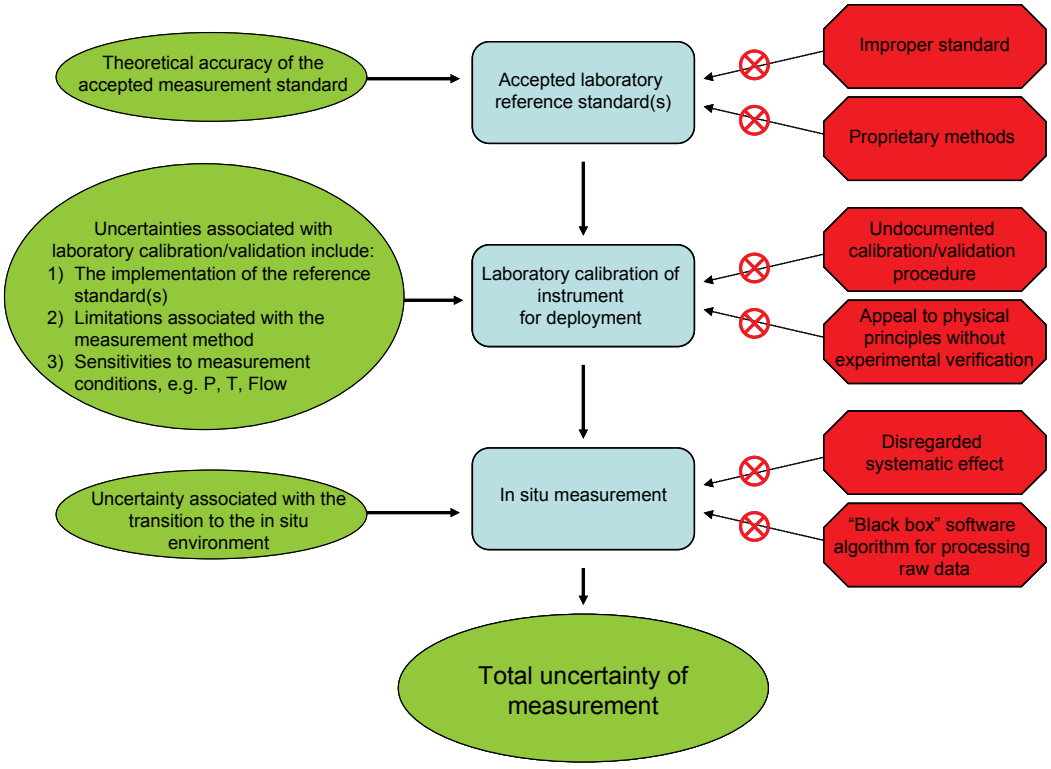


Figure 3.1. Schematic for establishing unassailable measurement accuracy. The flow of this graphic illustrates the accumulation of uncertainty in linking the in situ measurement to an accepted laboratory standard. The green ovals map out the core of a successful methodology. The red boxes depict methodological defects that jeopardize the credibility of the final uncertainty estimate.

Table 3.1 provides a practicable version of the calibration/validation methodology outlined in Figure 3.1. The left-hand column specifies the basic steps of a generalizable protocol for by any in situ instrument. The right-hand column indicates how each step has been realized in the calibration and validation process of HWV. The details of each step are discussed in subsequent sections.

	Calibration/Validation Step	**HWV Example**
0.	Selection of measurement method & Understanding the basic physical equations linking the method to the desired measurement	Photo-fragment fluorescence Calib. constants: C_0 and q_{air}
1.	Selection of a primary standard	i) SMR_{liq} @ Room T & P ii) Beer's law, $\sigma_{H2O\ @\ 121.6\ nm}$
2.	Laboratory calibration under flight conditions	~1-300 ppmv H_2O 20-200 mbar Low T tests: 230-300 K
3.	Integration onto aircraft platform & Evaluation of in-flight performance with prescribed diagnostic tests	i) Absorption vs. Fluorescence ii) Duct velocity scans
4.	Comparison with other instruments	HWV vs. HTW HWV vs. ICOS & HOXotope HWV vs. JLH
5.	Use of a priori knowledge to constrain observable	Expected RH_i in wave cloud: RH_i = 100 %

Table 3.1. Calibration and validation procedure for the HWV instrument. This table provides a generally applicable methodology for the calibration and validation of in situ instrumentation. Examples specific to HWV are listed for each step and discussed in detail in the following sections.

3.3 Laboratory Calibration

The accurate measurement of ambient water vapor mixing ratios with the photo-fragment fluorescence technique depends upon the accurate determination of the optical sensitivity of the HWV detection axis, defined by the calibration constants, C_0 and q_{air}, cf. Equations 2.5 and 2.6. These factors are determined empirically in the laboratory as a function of known water vapor concentration, $[H_2O]_{add}$, pressure and temperature. A rearrangement of Equation 2.5 leads to the equations that define the two-step calibration procedure, i.e., Equations 3.1 and 3.2. Equation 3.1b is identical to Equation 3.1a, except that it explicitly includes a constant non-zero intercept.

$$S_{norm} = [H_2O]_{add} \bullet C_{flr}; \; with \; C_{flr} = C_0 \bullet \frac{1}{1 + q_{air}[M]} \; at \; fixed \; [M] \tag{3.1a}$$

$$S_{norm} = ([H_2O]_{add} + [H_2O]_{bkgd}) \bullet C_{flr} = [H_2O]_{add} \bullet C_{flr} + \text{constant}; \quad at \; fixed \; [M] \; and \; fixed \; [H_2O]_{bkgd} \tag{3.1b}$$

$$\frac{1}{C_{flr}} = \frac{1}{C_0} + \frac{q_{air}}{C_0}[M] \tag{3.2}$$

To determine the calibration constant C_{flr} in Equation 3.1, we hold the duct pressure (equivalently [M]) constant and vary the concentration of water vapor, $[H_2O]_{add}$, at the detection axis. A plot of the normalized fluorescence signal, S_{norm}, versus added water vapor, $[H_2O]_{add}$, yields a line with slope C_{flr} and an intercept. The non-zero intercept is attributed to the presence of residual (background) water vapor from a combination of sources upstream of the detection axis: 1) the dry air supplied by the gas cylinder; 2) the gas addition lines upstream of the calibration duct; and 3) the walls of the duct itself upstream of the

detection axis. A non-zero intercept does not impact the accurate determination of the quantity C_{flr} as long as it remains constant over the duration of the water vapor titration at constant [M].

In order to obtain the pressure dependence of C_{flr}, we execute water vapor titrations at several different duct pressures. Examination of Equation 3.2 shows that a plot of $1/C_{flr}$ versus [M] yields a line with a slope proportional to the quenching coefficient, q_{air}/C_0, and an intercept equal to the inverse of the axis sensitivity in the absence of quenching, i.e., at zero pressure, $1/C_0$. Thus, the accuracy of the calibration critically depends upon how accurately we know the quantity of water vapor, $[H_2O]_{add}$, at the detection axis (i.e., our reference standard), and upon how accurately we measure the pressure and temperature of the duct flow.

3.3.1 Selection and Verification of a Calibration Standard

While the theoretical accuracy of any given standard may be exceptional, its practical implementation in the laboratory introduces sources of random error and sources of systematic error that limit its realizable accuracy and reliability. This is true even for recognized reference standards. The International Organization for Standardization (ISO) document on "Accuracy (trueness and precision) of measurement methods and results" provides a guide to assessing and establishing practical estimations of accuracy through experiment. The introductory section outlines the general principles that must be understood when assessing accuracy, and lists several factors that can contribute to variability or

uncertainty: 1) the operator; 2) the equipment used; 3) the calibration of the equipment; 4) the ambient environmental condition; and 5) the time elapsed between measurements [*ISO-5725-1*, 1994]. Great care must be taken in order to minimize the effects of these potential error sources, and diagnostic tests devised to identify and quantify the chosen standard's sensitivity to each of these factors.

In order to verify the accuracy of the chosen reference standard within the operational constraints of the laboratory calibration system, it is advantageous to implement a second, largely independent, reference. Accordingly, our laboratory calibration system simultaneously utilizes two independent reference methods, tied to two very different physical properties of water, for evaluating $[H_2O]_{add}$. Figure 3.2 provides a conceptual illustration of how the use of two independent reference techniques establishes the accuracy of the laboratory calibration. The two standards are represented by the dashed Gaussian lineshapes, drawn on orthogonal axes to signify the fact that their errors are uncorrelated. The positions of the peaks represent the effect of systematic errors or biases associated with the implementation of each calibration technique, and the widths represent the uncertainty associated with the precision of each technique. The solid Gaussian lineshape represents the minimization of systematic errors that is established by the required overlap of the independent reference techniques.

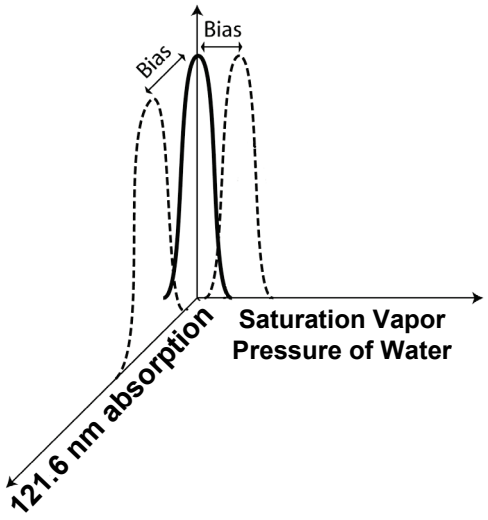


Figure 3.2. Schematic representation of how the use of independent reference techniques serves to minimize systematic errors that arise in the course of laboratory implementation.

Figure 3.3 shows a schematic of the HWV calibration system, illustrating the integration of the flight detection axis with two independent reference standards: 1) a dual-stage air saturation device, the "bubbler", and 2) a direct (Beer's law) absorption measurement. The dual-stage bubbler functions both as the water vapor source to the calibration flow system, as well as a water vapor reference standard, and is tied to the equilibrium partial pressure of vapor over liquid water at a well defined temperature and pressure. The second reference, which relies on the absorption of 121.6 nm radiation by water vapor in the calibration duct, provides a straightforward means of verifying the accuracy of the calibrated water vapor addition between the point of delivery and the point of detection. The axial absorption measurement (green line) detects the attenuation of 121.6 nm light due over a distance of ~1 meter along the axis of the duct.

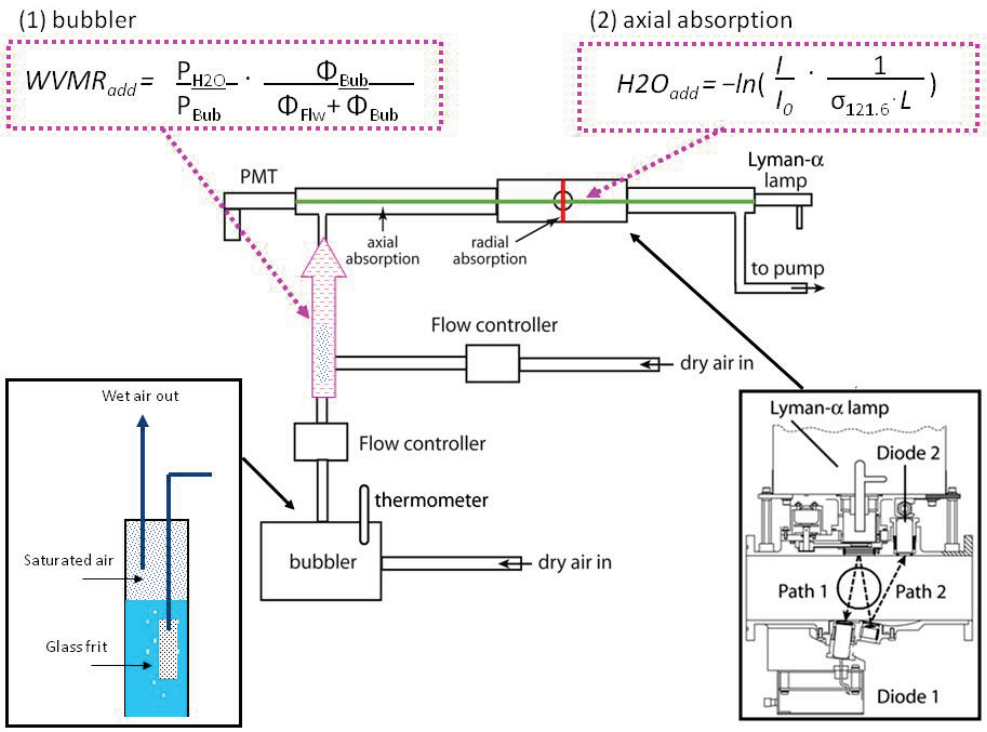


Figure 3.3. Schematic of HWV calibration bench. The flight detection axis is integrated into the laboratory flow system. Controlled quantities of humid air are added via the bubbler. The mixing ratio is set by the equilibrium partial pressure of vapor over liquid in the bubbler and the ratio of the saturated and dry air flows. The water concentration is confirmed by a Beer's law absorption measurement along the primary axis of calibration duct.

Bubbler

Carefully quantified amounts of humid air are added to the calibration flow system using a dual-stage bubbler. Dry air is fed through a glass frit submerged in a water bath such that the headspace above the bath maintains the equilibrium vapor pressure set by the bath temperature. In a dual-stage bubbler the saturated air is then fed through a second frit and water bath, guaranteeing air within 99.9% of saturation. The fully saturated air is drawn off, fed through a calibrated flow controller and mixed with the primary dry air flow.

Equation 3.3a shows how the water vapor mixing ratio of the flow, $WVMR_{add}$, is determined.

$$WVMR_{add} = \frac{P_{H2O}(T_{Bub})}{P_{Bub}} \bullet \frac{\phi_{Bub}}{\phi_{Bub} + \phi_{Dry}} \tag{3.3a}$$

The equilibrium partial pressure for water vapor over liquid water, P_{H2O}, at the bubbler is determined using the parameterization provided by [*Murphy and Koop*, 2005], and an accurate measurement of the bubbler temperature, T_{Bub}. P_{Bub} is the total pressure measured at the bubbler, and the ratio of pressures, P_{H2O}/P_{Bub}, is equal to the saturation mixing ratio, SMR_{liq} exiting the bubbler. The saturated air is then added to a steady dry air stream. The flow rates of saturated and dry air are denoted by ϕ_{Bub} and ϕ_{Dry} respectively, and the ratio of the saturated flow to the total flow establishes the mixing ratio of the sample. The saturated air flow is typically 1 – 2% of the dry air flow.

Each calibration run explores a range of water vapor mixing ratios (~1 – 500 ppmv) and pressures (20 – 300 mbar) in order to determine the calibration constants and validate the linearity of the detection axis throughout the full range of mixing ratios encountered in situ.

Though mixing ratio is the conserved quantity in the flow system, the sensitivity of the detection axis depends upon the number density of water. Equation 3.3b shows the conversion between the water vapor mixing ratio and water vapor number density at the detection axis, $[H_2O]_{add}$, using pressure and temperature measurements near the detection region.

$$[H_2O]_{add} = WVMR_{add} \bullet [M(P_{flow}, T_{flow})] \tag{3.3b}$$

Therefore, the accuracy of the bubbler reference standard is limited by uncertainties in the measurements of pressure, P_{Bub} and P_{flow}, temperature, T_{Bub} and T_{flow}, the mass flow rates, ϕ_{Bub} and ϕ_{Dry}, and by the parameterization for determining the saturation mixing ratio [*Hintsa et al.*, 1999; *Weinstock et al.*, 2006a; *Weinstock et al.*, 2006b]. There are numerous parameterizations for determining P_{H2O} over a wide range of temperatures. At room temperature, however, they all agree to within 0.2% [*Murphy and Koop*, 2005]. Table 3.2 summarizes the error budget for this method. The total uncertainty estimate associated with this method is ±5.0%.

Method 1: Bubbler			
	est. unc.	**est. unc. (%)**	**max. est. unc. (%)**
$\boldsymbol{WVMR_{add}}$			
P_{Bub}	±0.5 torr	~0.1%	~1% max
T_{Bub}	±0.2 K	~1% in K	~2% in SMR_{liq}
P_{H2O}		~0.2%	~0.2%
[M]			
P_{flow}	±0.5 torr	over 30 to 300 torr	~0.1 to 1%
T_{flow}	±2 K max	~1%	~1%
Flow			
ϕ_{Dry}	3%	3%	3%
ϕ_{H2O}	3%	3%	3%
		TOTAL:	**~5%**

Table 3.2. Uncertainty estimate for the bubbler reference standard.

Systematic errors associated with delivering a known mixing ratio of water vapor unperturbed to a sensor downstream plague every such system regardless of the details and inherent accuracy of the humidity source, and are not

included in Table 3.2. Adsorption onto and desorption from surfaces between the humidity source and the point of detection, or between two independent points of detection, must always be considered.

Though we mention both adsorption and desorption, the primary concern is with the introduction of additional water to the sample flow due to the net outgassing of water from interior surfaces of the addition lines and/or calibration ducting. The net outgassing rate may depend upon the temperature, pressure, flow rate, dead volumes, the duct material and/or treatment of the duct surface, as well as exposure history. Furthermore, leaks into the calibration system upstream of the detection axis constitute an additional source of uncertainty. HWV, like many of the in situ instruments, does not operate with a significant pressure differential in situ, however, in order to test the instruments at atmospherically relevant pressures in the laboratory the detection axis must be leak tight. Even a small leak of room air into the ducting can lead to a non-negligible contribution of water vapor to the sample flow.

As demonstrated in Equation 3.1b, if the background contribution of outgassing and leaks is constant over a water vapor titration at constant [M], the accuracy of C_{flr} will not be compromised. However, it is imperative that the stability of the background be verified, and where possible, the contributions due to these spurious sources of water vapor quantified and minimized. This is particularly true for titrations at low mixing ratios.

Axial Absorption

The second reference method measures the water vapor concentration along the central axis of the calibration flow tube via the direct absorption of 121.6 nm radiation. The Lyman-α lamp and vacuum ultra-violet PMT used for the axial absorption measurement are independent of the detection axis under calibration. The absorption cross section of H_2O at 121.6 nm, and Beer's law, are used to determine H_2O_{add}, in accordance with Equation 3.4.

$$[H_2O]_{add} = \frac{-1}{l \cdot \sigma_{H2O}} \ell n(\frac{I_{add}}{I_0}); \quad with\ Beer's\ law: I = I_0\, e^{-l \cdot \sigma_X \cdot [X]} \tag{3.4}$$

The signals I_0 and I_{add} refer to the Lyman-α signal when no water is added to the flow and the signal during the titration, respectively. Note that the absorption method cannot be used to measure the absolute water vapor concentration in the duct. It can only measure relative changes in the mixing ratio. Thus, the presence of a non-zero water vapor background has no impact on the determination of $[H_2O]_{add}$ as long as the background is constant during each titration sequence. Similarly, as long as the duct pressure remains constant we do not need to explicitly account for absorption due to oxygen. This method has the advantage of measuring the water vapor concentration directly in the flow duct, and thus provides an excellent independent verification of 1) the stability of the background, and 2) the accuracy of $[H_2O]_{add}$ as determined by the bubbler.

The accuracy of the absorption method is limited by the uncertainty in the cross section, σ_{H2O}, as well as by the measurement of the path length, l. We use the water vapor cross section at Lyman-α measured by [*Lewis et al.*, 1983], which has an estimated uncertainty of ~3% [*Lewis et al.*, 1983]. The summary of errors

associated with the absorption method is given in Table 3.3. The total uncertainty is estimated to be ~4%.

Method 2: Axial Absorption			
	est. unc.	**est. unc. (%)**	**max. est. unc. (%)**
L	±2.0 cm	2%	2%
σ_{H2O}	3%	3%	3%
		TOTAL:	**~4%**

Table 3.3. Uncertainty estimate for the axial absorption reference standard.

The demonstrated agreement between the two methods over numerous calibration runs, exploring a wide range of duct pressures and flow velocities, is within ±3%, well within their combined uncertainty. This level of agreement provides conclusive evidence that we have minimized systematic errors in the implementation of each calibration technique, and that the water vapor measured by the fluorescence detection axis under calibration is known, within the limits of its precision, to better than ±5%. A more thorough discussion of the results from representative calibration runs is provided in the following section.

3.3.2 Laboratory Calibrations at In Situ [H_2O], P and T

Calibrations of the detection axes are routinely executed in the laboratory prior to and after each field mission. A portable version of the calibration system has also been utilized to monitor the stability of the calibration during field campaigns. In general, however, detection axis performance has proven to be robust, and has not necessitated frequent adjustment or re-calibration.

Derivation of Calibration Constants

Figure 3.4 shows data from a representative calibration run performed on November 5, 2006. Figure 3.4a shows the series of five water vapor titrations, each performed at a constant duct pressure. Normalized fluorescence counts, S_{norm}, and the corresponding amount of water vapor added to the calibration system, $WVMR_{add}$, are both plotted as a function of time. As expected, the normalized fluorescence signal is effectively proportional to the water vapor mixing ratio over a pressure range representative of the UT/LS.

Figure 3.4b shows S_{norm} plotted as a function of the added water vapor number density, $[H_2O]_{add}$. Each line in Figure 4.3b corresponds to a titration in Figure 4.3a, where the first and last titrations were executed at the same pressure, 65 hPa. The slope of each line is equal to the sensitivity, C_{flr}, of the detection axis at a fixed pressure, or equivalently, a constant duct number density, [M], see Equation 3.1a.

Figure 3.4c demonstrates the derivation of the pressure dependent and pressure independent components of C_{flr}, by plotting $1/C_{flr}$ as a function of [M]. Recall the pressure dependence arises because of collisional quenching by air. As shown in Equation 3.2, the slope of the line is equal to q_{air}/C_0 and the intercept is equal to $1/C_0$. The calibration constants are derived from robust linear least squares fits to the data. This fitting algorithm uses an iteratively re-weighted least squares analysis and is more robust than a simple linear least squares fit in the presence of a few outliers. Inspection of the fit residuals (not shown) demonstrates the advantage of this fit method over a simple linear least squares

method, where a couple of points at high water vapor mixing ratios leverage the fit.

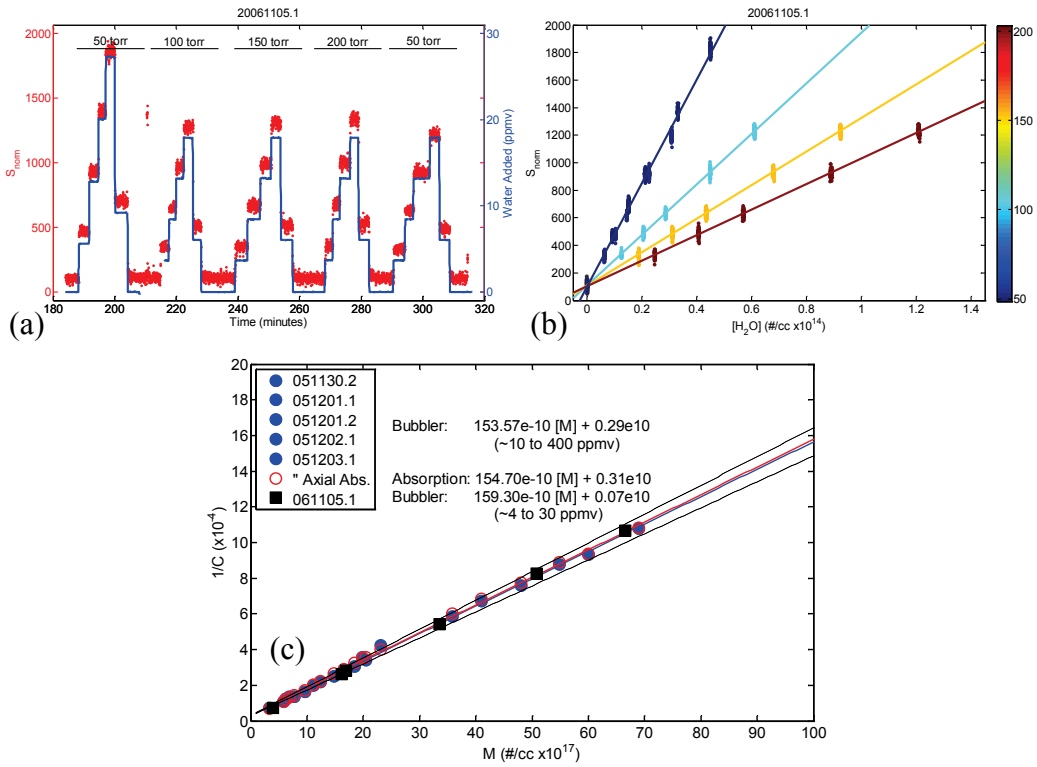


Figure 3.4. Data from a representative calibration run. Panel (a) shows the normalized fluorescence signal (red), S_{norm}, and the added water vapor mixing ratio (blue) as a function of time for the calibration executed on November 5, 2006. This run included five water vapor titrations at four pressure settings. In panel (b), the sensitivity of the detection axis, C_{flr}, at fixed pressure is determined by the slope of each of the four lines. The colorbar denotes the pressure (in hPa) in the duct. In panel (c), the pressure dependence of C_{flr} is determined by plotting $1/C_{flr}$ as a function of [M]. The slope and intercept of the line yield the final calibration constants, q_{air}/C_0 and $1/C_0$, respectively. The details of this figure are described in the text.

In Figure 3.4c, the data for the November 5, 2006, run are plotted together with the results of five calibration runs of the same detection axis executed a year earlier. The data from the prior runs are plotted in the solid blue circles and in the open red circles, where the different colors correspond to the independent results

obtained with the bubbler and the absorption reference standards for $[H_2O]_{add}$, respectively. The blue and red lines show the results of fits to the data from all five calibration runs executed in November and December of 2005, where the colors again specify the underlying water vapor reference standard. This plot demonstrates 1) the outstanding agreement, well within ±3%, between the results of the two independent calibration standards, 2) the remarkable stability of the detection axis over the course of a year, within ±5%, despite being flown during the CRAVE mission in January and February of 2006, and 3) the linearity of the calibration over an order of magnitude in pressure (~10 hPa to ~300 hPa) and two orders of magnitude in water vapor mixing ratio (~4 to 400 ppmv).

Low Water Vapor Calibrations

The concern over the accuracy of the in situ measurements, particularly at the lowest water vapor mixing ratios, ~1 to 5 ppmv, prompted us to extend the range of our calibration runs to ~<1 ppmv to confirm the accuracy of the photo-fragment fluorescence detection technique in this range.

The focus on low water vapor calibrations revealed new challenges, in particular, achieving a truly dry or near zero background in the calibration duct. In theory the presence of a non-zero background should have no impact on the accurate determination of the quantity C_{flr}. In practice, however, small fluctuations of a high water vapor background have the potential to substantially impact calibrations at low water vapor mixing ratios.

Minimizing the background water vapor requires careful conditioning of the entire calibration system. Because water is a "sticky" molecule, every component of the calibration system from the gas cylinder regulators to the walls of the instrument duct can act as a water vapor source. As mentioned earlier, the background is expected to be a combination of contributions from the walls of the water vapor addition plumbing and walls of the flow system, leaks into the duct, and any residual contribution from the dry air source itself. Thus, in order to ensure the accuracy of the low water calibrations, we regularly pre-conditioned the calibration system by pumping it out overnight in the presence of a steady flow of dry N_2 or air to achieve values of $[H_2O]_{bkgd}$ of a few ppmv or less. In addition, we improved the vacuum integrity of the ducting, and used stainless steel fittings and tubing wherever possible to minimize the water vapor "memory" of the gas addition lines.

Data from mid and low water vapor calibration runs executed during the Aqua Validation and Instrument Tests (AquaVIT) campaign are shown in Figure 3.5. The AquaVIT campaign was a focused water vapor instrument comparison effort hosted by the AIDA facility at the Forschungszentrum Karlsruhe in Karlsruhe, Germany, during October 2007 (see Chapter 5). Figure 3.5a shows selected titrations from calibrations executed on October 16, 2007, and October 19, 2007. The first calibration run explored mixing ratios from ~1 ppmv to 21 ppmv in steps of 1 to 5 ppmv, and had a background mixing ratio of ~1.1 ppmv. The second explored mixing ratios from ~0.5 to 3 ppmv in steps of ~0.5 ppmv, and had a stable background of ~0.6 ppmv when no water was added. The bubbler

data, $WVMR_{add}$, are plotted in black, and the colored lines show the bubbler data corrected for the small non-zero background. Mixing ratios measured by the fluorescence detection axis are shown in the colored dots. The fluorescence measurements for both runs shown here were acquired using the calibration constants from a composite of four calibrations (including the two shown) performed during the campaign.

Figure 3.5a, shows the excellent agreement between the mixing ratios measured at the detection axis and the corrected bubbler data. These data also highlight the fast response and outstanding precision of the water vapor addition system (bubbler), the comparably fast response and excellent precision of the detection axis, and finally the low water vapor background achieved during both calibration runs.

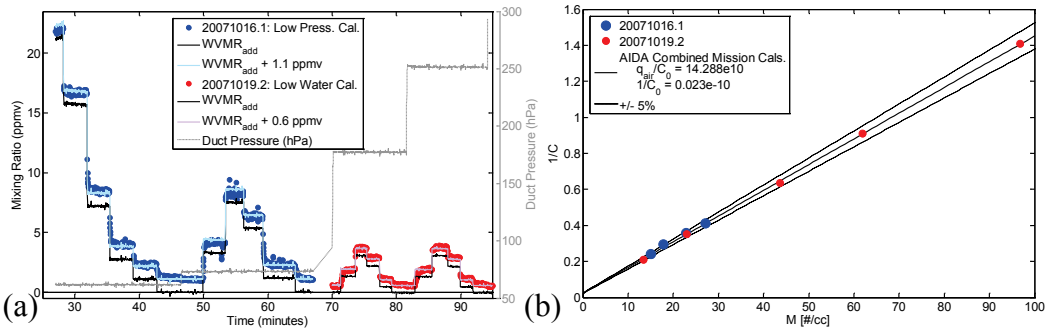


Figure 3.5. Linearity of Lyman-α detection method over a wide range of pressures and mixing ratios. Panel (a) shows selected titrations from a calibration executed at low pressures, ~60 to 120 hPa (blue data), and at low water vapor mixing ratios, 0 to 5 ppmv (red data). The fluorescence data, evaluated using the derived calibration constants, are plotted in the colored dots, added water vapor, $WVMR_{add}$, utilizing the bubbler reference standard, is plotted in the black line, and the colored lines show the bubbler plus a constant offset representative of the background mixing ratio in the carrier flow. Panel (b) shows the calibration figure, i.e., $1/C_{flr}$ vs. [M], with the data from both calibration runs. The solid black line represents the best fit to the full set of calibrations performed at the AIDA facility, and the dashed lines show the ±5% uncertainty range.

The inverse of the sensitivities, C_{flr}, derived from the titration data in Figure 3.5a are plotted as a function of [M], in the calibration plot shown in Figure 3.5b. The solid black line represents the results of a linear fit to the complete set of calibration data acquired during the AquaVIT campaign, and the dashed lines denote the ±5% range. The data from these two runs fall well within the ±5% uncertainty envelope. Furthermore, the individual slopes for these two runs (not shown) agree with the final calibration fit to within 5%. These results again demonstrate the linearity and robustness of the calibration over an order of magnitude in pressure and from a few tenths of a ppmv to tens of ppmv.

Low Temperature Calibrations

Consistent with the fourth factor of the ISO 5725-1 report [*ISO-5725-1*, 1994], the success of the HWV instrument depends not only upon the accuracy of its laboratory calibration, but also on its ability to achieve that accuracy in the flight environment. Differences, including temperature, pressure, flow rate, mixing ratio and even airmass composition, between the laboratory and flight environments must be carefully considered for every in situ instrument. Where possible, a combination of laboratory diagnostic tests and/or calibrations must be carried out under conditions closely representative of those in flight. The preceding paragraphs have demonstrated our ability to calibrate the HWV detection axis in the laboratory with an accuracy of ±5% under pressure and mixing ratio conditions representative of the UT/LS. In this section we discuss the

results of laboratory tests exploring the temperature sensitivity of the detection axis.

Laboratory tests of the early balloon instrument revealed the potential for a systematic temperature effect. Cold optics caused an increase in the detection axis sensitivity. This sensitivity change was due almost entirely to an increase in the transmission of the interference filter at –40 ºC [*Schwab et al.*, 1990]. Accordingly, the PMT filter assembly housing is now continuously purged with dry air during flight and kept warm with a thermostated heater to prevent temperature induced changes in the transmission and bandpass of the filters [*Weinstock et al.*, 1994]. Furthermore, all the main optical and electrical components of the detection axis are packaged with resistive heaters and insulated. Diagnostic thermistors are mounted on key components of the instrument in order to track temperature changes during flight. The thermistor mounted on the filter assembly housing, for example, shows that its temperature is well regulated and reliably held at ~+15ºC. Thus, we do not expect the HWV axis sensitivity to change in the flight environment.

The temperature dependence of the quenching term, which dictates the branching between quenching and fluorescence, has also been investigated. OH* fluorescence from low rotational states demonstrates a measurable increase in quenching efficiency with a decrease in temperature, i.e., a change of ~10% from 298 K to 200 K. An increase in quenching leads to a reduction in fluorescence, and could cause our measurements at cold temperatures to be erroneously low. However, at the very high rotational levels produced by the photo-dissociation of

H_2O, the temperature dependence of the quenching rate is negligible [*Hintsa et al.*, 1999]. [*Hintsa et al.*, 1999], calibrated HWV-E over a range of temperatures from room temperature down to −30°C and found no change in axis sensitivity.

Similarly, the temperature dependence of the absorption cross section of water vapor at 121.6 nm has been considered. Results from laboratory studies by [*Lewis et al.*, 1983] have demonstrated, however, that the cross section does not exhibit a significant temperature dependence, within ±4%, from room temperature to temperatures representative of the UT/LS. In fact, their experimental uncertainty is larger than the magnitude of any potential temperature dependence.

Despite the indirect evidence supporting the temperature insensitivity of the technique, we felt it was necessary to directly and unambiguously test the temperature response of the current generation of the HWV detection axis. Replicating the tests of [*Hintsa et al.*, 1999], we calibrated the HWV detection axis over a range of temperatures in order to empirically verify its temperature independence, and explicitly demonstrate its accuracy under the thermal conditions of flight.

We built an insulated box around the entire calibration flow system. Liquid nitrogen from a 240 liter tank was fed into a heat exchanger and fan assembly located inside the box to provide cooling. The coldest temperatures achieved were ~–55ºC, similar to temperatures measured inside the instrument pod during flight. Though we did not actively cool the flow, by the time it reached the detection axis, flow temperatures were typically only ~15ºC warmer than the ambient box temperature. During cold temperature scans and cold soaks of a

couple hours duration, the diagnostic thermistor on the PMT filter assembly showed that the temperature was stable at ~+20ºC, consistent with in-flight diagnostic data.

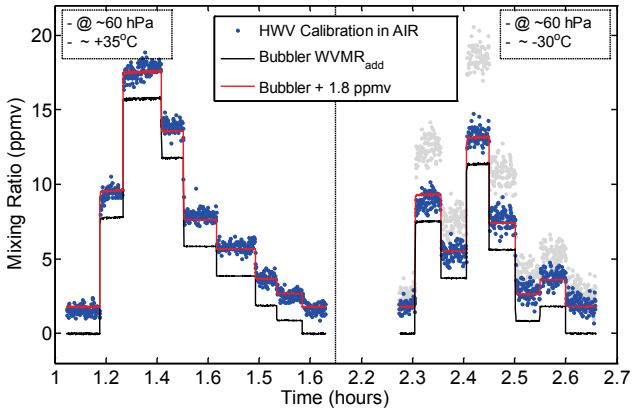


Figure 3.6. Titrations from calibrations run over a ~70°C temperature range. The fluorescence data (blue) in both panels were analyzed with the same calibration constants, and demonstrate the temperature independence of the HWV detection axis. The light grey data show the effect of a hypothetical 40% increase in sensitivity caused by a systematic temperature dependence of the PMT collection optics.

Figure 3.6 shows the results of one temperature sensitivity test, with calibration data acquired prior to and during a cold soak. The mixing ratio and pressure conditions are nearly identical, but the temperatures are radically different. The temperature differential between the data sets is ~70ºC, with box temperatures at +35ºC prior to cooling and approaching –35ºC during the cold soak. The raw bubbler data, $WVMR_{add}$, are plotted in black, and the bubbler data adjusted by a nominally constant 1.8 ppmv are in red. The blue data correspond to the measured mixing ratios, determined using the set of calibration constants. For comparison, the grey data represent mixing ratios resulting from a hypothetical 40% increase in the sensitivity of the PMT collection optics. These data

demonstrate the insensitivity of the detection axis calibration over a temperature range of nearly 70ºC, with the low temperature end approaching the environmental conditions of flight.

3.4 In-Flight Diagnostic Tests

In the following section we evaluate the performance of the HWV instrument in situ. Specifically, we discuss the results of two diagnostic tests that we regularly perform in flight. These tests were designed to 1) evaluate the applicability of the laboratory calibration in flight, and 2) examine the impact of outgassing from the instrument ducting on the measured mixing ratio. The first test verifies the stability of the detection axis calibration in the transition from the laboratory to the atmospheric environment, and the second test verifies that the ambient mixing ratio is delivered unperturbed to the detection region.

3.4.1 Calibration Verification: Absorption vs. Fluorescence

Validation of the HWV calibration in flight is achieved by comparing the photo-fragment fluorescence measurements of water vapor with simultaneous measurements made by direct absorption. Natural perturbations to water vapor, of sufficient magnitude above the background, provide opportunities for synchronous measurements by both methods [*Hintsa et al.*, 1999; *Weinstock et al.*, 2006a; *Weinstock et al.*, 2006b]. This approach to verifying instrument accuracy in flight is analogous to our choice of dual calibration standards in the

laboratory. We are comparing measurements of water vapor made by two independent methods, each subject to its own set of systematic errors.

The absorption measurement utilizes the pair of VUV photodiodes, which monitor the fluorescence lamp flux, see Figure 3.3. We measure the cross duct (radial) absorption between the two diodes. As shown in the inset in Figure 3.3, the first diode is mounted across the flow from the lamp and adjacent to the VUV mirror, and the second diode is positioned flush with the opposite wall downstream of the lamp. Equation 3.5a is a modified version of Beer's law that uses the ratio of the two diode signals. Equation 3.5b is simply a rearrangement of 3.5a used to derive the ambient water concentration. The effective path length is then equal to the difference between Path 2 and Path 1 in Figure 3.3, and the use of the diode ratio provides a means of accounting for changes in the VUV lamp flux.

$$\frac{I_{VUV2}}{I_{VUV1}} = \frac{I_{VUV2,0}\, e^{-l_2 \cdot \sigma_{H2O} \cdot [\mathrm{H_2O}]}}{I_{VUV1,0}\, e^{-l_1 \cdot \sigma_{H2O} \cdot [\mathrm{H_2O}]}} = \frac{I_{VUV2,0}}{I_{VUV1,0}} e^{-(l_2 - l_1) \cdot \sigma_{H2O} \cdot [\mathrm{H_2O}]} \tag{3.5a}$$

$$[\mathrm{H_2O}] = \frac{-1}{\Delta l \cdot \sigma_{H2O}} \ell n(\frac{R_{VUV}}{R_{VUV,0}}); \; with \; \Delta l = l_2 - l_1 \; \& \; R_{VUV} = \frac{I_{VUV2}}{I_{VUV1}}, etc. \tag{3.5b}$$

The radial absorption method relies on finding features along the flight track where 1) there is a change in the ambient water vapor mixing ratio of sufficient magnitude to be detectable over the short, ~6 cm, path length, and where 2) there is little to no change in the ambient pressure, i.e., where we can assume the change in the concentration of other absorbers such as O_2, O_3, CO_2 and CH_4, is minimal. The ratio of the diode signals, R_{VUV}, during an excursion in ambient water defines our transmitted intensity over the path difference between

the diodes, and a linear interpolation between the short regions extending before and after the excursion defines our incident intensity, or baseline reference, $R_{VUV,0}$.

Data from such a feature, encountered on the February 9, 2006, flight during the CRAVE mission, are plotted in Figure 3.7. The results of the radial absorption measurement are plotted (red dots) in Figure 3.7a as a function of time. Also shown are the results of the measurements made by the calibrated fluorescence detection axis (blue dots). Because the absorption measurement can only measure the change in water vapor over the feature, or ΔH_2O, the background water vapor measured by the fluorescence technique has been added back in for the comparison. Figure 3.7b shows ΔH_2O measured by the absorption technique plotted as a function of ΔH_2O measured by the fluorescence technique. Also shown are the results of a linear least squares fit, and the 1:1 line.

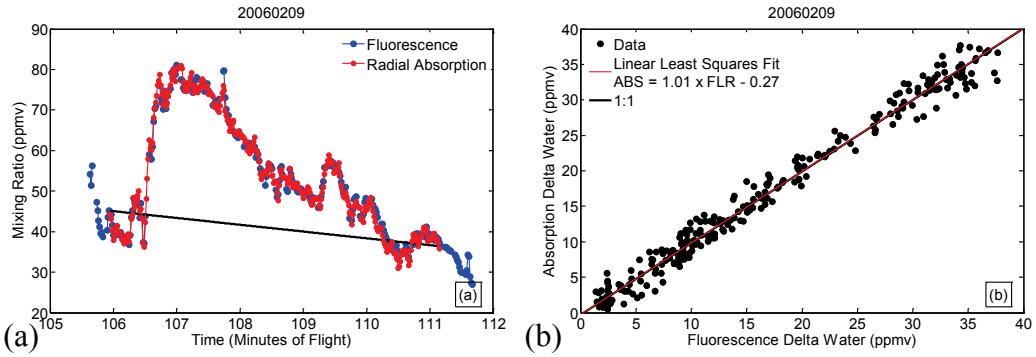


Figure 3.7. Comparison of fluorescence and radial absorption measurements of water vapor from the flight of February 9, 2006 during the CRAVE mission. Panel (a) shows water vapor measured by the direct absorption of Lyman-α radiation across the instrument duct (red), and by the calibrated fluorescence detection axis (blue). Panel (b) quantifies the comparison. Water measured by absorption is plotted versus water measured by fluorescence. The solid line shows the results of a linear least squares fit to the data, demonstrating agreement to within ~1%. The 1:1 line is shown for comparison.

The level of agreement shown between these two independent measurement methods is within ~1%, demonstrating that the calibration determined under laboratory conditions is valid in the flight environment, and furthermore, that the accuracy achieved in the laboratory (±5%) is achieved in flight. Comparisons, such as the one shown in Figure 3.7, from the CRAVE mission, also provide a means of monitoring the stability of the calibration over a mission.

3.4.2 In Situ Contamination Assessment: Velocity Scans

One of the dominant concerns facing every in situ water vapor instrument, balloon and aircraft alike, is the effect of water vapor outgassing or shedding off of surfaces and into the detection volume. In the HWV instrument, the velocity in the secondary duct, determined from pitot tube measurements of differential pressure, is typically ~70 m/s ± 10 m/s. This is equivalent to mass flows on the order of ~1000 to ~3000 slm depending upon the ambient pressure. At these fast flow rates the impact of outgassing is expected to be minimal. However, in order to confirm that this potential contamination source is truly insignificant, we periodically vary the flow rate in flight using a throttle valve located in the detection axis duct. If we assume that an outgassing signal is due to a constant small flux of water vapor, reducing the velocity, and thus the mass flow in the duct, should increase the contamination contribution to the measured mixing ratio.

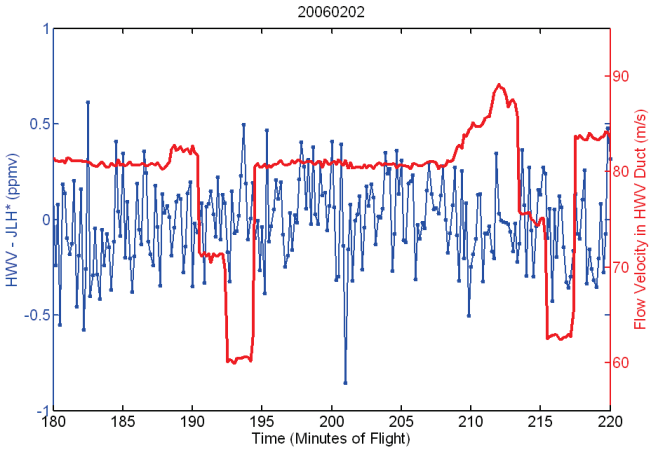


Figure 3.8. Results of in situ diagnostic velocity scans. The velocity in the HWV duct (red line) is plotted as a function of time during the flight of February 2, 2006. The difference between HWV, which samples via an internal duct, and JLH, which is an open path instrument located under the wing of the WB-57, is also plotted as a function of time. The results show that the HWV measurement is not sensitive to velocity in the duct. There is no evidence, therefore, for contamination of the ambient mixing ratio in flight.

Figure 3.8 shows the results of two representative velocity scans executed during the flight of February 2, 2006, during the CRAVE mission. The velocity in the duct was varied at roughly 25 minute intervals. The red line shows the stepwise changes in duct velocity as a function of time during the flight. In order to look for corresponding changes in the measured mixing ratio we compare the HWV measurements with measurements made simultaneously by the independent Jet Propulsion Laboratory Laser Hygrometer (JLH), an open path tunable diode laser instrument located under the wing of the WB-57. The JLH data have been linearly adjusted for ease of comparison and interpretation so that the mean difference is zero, with JLH* = JLH × 0.78 – 1.5. The difference between HWV and JLH* plotted in Figure 3.8 (blue line) shows no sign of a velocity dependence. Though noisy, there is no indication of the difference between the

instruments increasing as the flow velocity in HWV is decreased. These tests are performed routinely during every flight and have never shown evidence for water vapor contamination in the HWV instrument.

3.4.3 Accuracy of In-Cloud Data: Sublimation Analysis

Model

The HWV instrument measures ambient water vapor mixing ratios in an internal duct. Ram heating of air inside the duct has the potential to partially sublimate ice particles by the time they reach the detection axis, leading to a systematic high bias of in-cloud measurements resulting solely from sampling conditions. The magnitude of this bias depends directly upon the extent of sublimation, which is expected to be a function of the cloud particle size distribution, particle surface to volume ratio, residence time in the duct, and the difference between the ambient and duct temperatures. Unfortunately, assumptions regarding particle shape and density, combined with uncertainties in particle size distribution measurements, render exact calculations impossible. Therefore, we developed a simple model to investigate the potential for sublimation in the duct and to estimate its maximal effect.

The model examines the sublimation of spherical ice particles during the time it takes them to travel from the entrance of the ram-heated duct to the detection axis, approximately one meter downstream. Inputs to the model include the velocity, pressure, temperature and water vapor mixing ratio inside the duct, as well as ambient particle number densities determined from the combination of

the Nucleation-Mode Aerosol Size Spectrometer (NMASS), Focused Cavity Aerosol Spectrometer (FCAS) and Cloud, Aerosol Precipitation Spectrometer (CAPS) datasets, covering particle sizes from approximately 4 nm to 1.6 mm in diameter [*Baumgardner et al.*, 2001; *Jonsson et al.*, 1995]. Sublimation occurs in the model via surface area loss. At each time step, the new (reduced) particle size is determined, and the corresponding water vapor contribution to the flow calculated.

Duct flow velocities of 40 – 80 m/s ensure that particles entering the duct will travel to the detection axis in ~0.01 to 0.03 seconds. In this time, the temperature of the airmass typically increases to +15 ºC above ambient[1], leading to sublimation.

Several factors in the model favor sublimation. The use of the temperatures recorded downstream of the detection axis enhances the impact of sublimation, since in reality there is a positive temperature gradient from the duct entrance to the thermistor ring. Temperatures upstream of the detection axis will be colder than those used in the model. Also, an ice density of 0.9 g/cm^3 is assumed, which is a conservative estimate that again maximizes the effect of sublimation. Finally, the model was initially run assuming spherical particles, and then modified to account for variable shapes. Model runs with variable shapes tended to reduce the net effect of sublimation. Though non-spherical particles fully sublimate more quickly than spheres, the difference in rate is generally not large enough to matter on the extremely short time scales inside the instrument

[1] The difference between the ambient and duct temperatures is determined from Meteorological Measurement System (MMS) measurements of ambient temperatures [Scott et al., 1990], and measurements from an array of thermistors in the duct downstream of the detection axis.

duct. In the model, only the smallest particles, those with diameters <1.0 μm, sublimate completely by the time the air sample reaches the detection axis. Additionally, for an equivalent surface area lost to sublimation, there is a larger volume/mass of water contributed from spherical particles than from non-spheres.

We modeled two cirrus cloud encounters from the Cirrus Regional Study of Tropical Anvils and Cirrus Layers – Florida Area Cirrus Experiment (CRYSTAL-FACE) mission. The first is representative of moderately thick cirrus. The cloud encounter occurred from ~20:26 to 20:29 UT hours on the flight of July 23, 2002. The measured water vapor mixing ratio in the cloud was ~28 ± 2 ppmv, and the peak cloud ice water content (IWC) was ~210 ppmv. The cloud IWC content is determined from the pair of Harvard water instruments, where IWC = total water – water vapor [*Weinstock et al.*, 2006a; *Weinstock et al.*, 2006b].

The model results from this particular example are shown in Figures 3.9 and 3.10. Figure 3.9 shows the particle size distribution for this cloud encounter (open black circles) based on a composite of the particle measurements at the peak of the cloud encounter, and Figure 3.10 shows the mixing ratio contribution due to sublimation from each particle bin as a function of time, as well as the total contribution from sublimation and the measured water vapor mixing ratio in the cloud. The temperatures and pressures of the ambient air and duct are noted in the figure. The duct was ~15ºC warmer than ambient and the flow velocity was ~50 m/s.

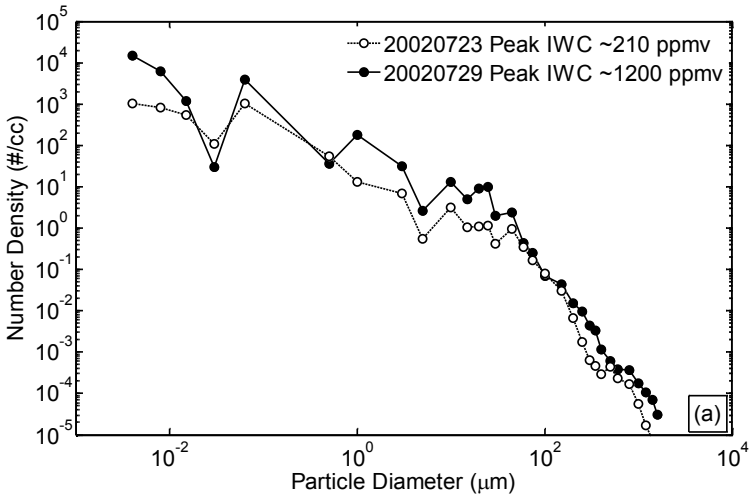


Figure 3.9. Cirrus particle size distributions. This figure shows size distributions (number density vs. particle size) for two thick cirrus clouds encountered during the CRYSTAL-FACE mission.

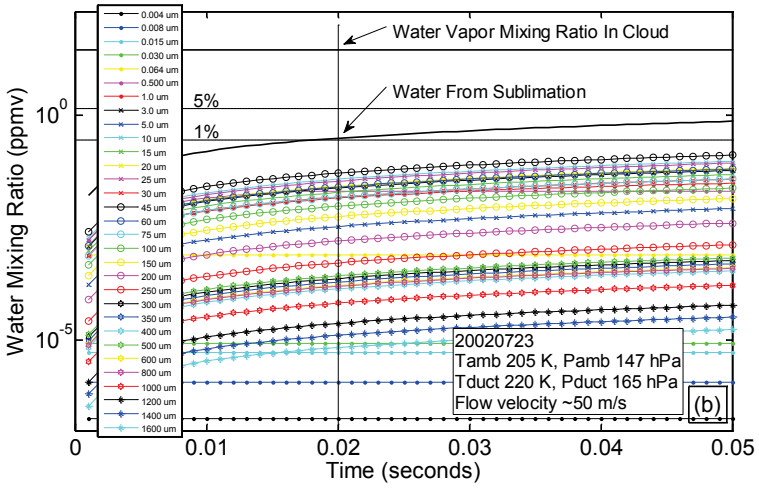


Figure 3.10. Sublimation model results. Mixing ratio contributions for each particle size bin (colored symbols) are plotted as a function of time for the cloud encounter on July 23, 2000. The vertical line at 0.02 seconds represents the nominal position of the HWV detection axis. The horizontal solid black line shows the ambient water vapor mixing ratio measured in the cloud. The dashed and dotted lines correspond to 5% and 1% of the ambient mixing ratio respectively. The total contribution due to sublimation is shown in the solid black curved line.

For the second example, we chose an extremely thick cirrus cloud encounter from the CRYSTAL-FACE flight of July 29, 2002. The peak IWC in the cloud was ~1200 ppmv (19:21 to 19:33 UT hours). To put this in context, the ice water content of this cloud, and its underlying particle number density, was significantly greater than the vast majority of the cirrus sampled in all of the recent missions aboard the WB-57. During CRYSTAL-FACE only ~0.5% of the clouds sampled had ice water contents >1000 ppmv. This cloud encounter, therefore, provides an upper bound on the degree of contamination attributable to sublimation in the HWV duct.

During this cloud encounter the duct velocity was ~55 m/s, and the duct temperature was at most 12ºC warmer than ambient. The particle size distribution for this cloud is shown in Figure 3.9 (black circles). The model results (not shown) indicate the net contribution from sublimation after 0.02 seconds is on the order of 2.2 ppmv, which amounts to a positive bias of ~2.5% given a measurement of ~90 ppmv water vapor in the cloud. Thus, according to the model, the most extreme cloud encounter yields a maximum measurement bias $\leq$+3%.

A significant degree of uncertainty arises from the particle size and number densities used in the model. There is an increasing body of evidence showing that many of the particle instruments are subject to errors in their particle number densities and size distributions due to the shattering of large particles off of inlets, as well as off the body of the aircraft [*Field et al.*, 2006; *Jensen et al.*, 2009; *Korolev and Isaac*, 2005]. Shattering leads to erroneously high number

densities of particles with diameters ~10 μm. In our model, the 10 μm to 25 μm diameter particles contribute significantly to the sublimation signal in both of the cloud encounters examined. If the real number densities of these size particles are reduced, the net effect of sublimation will be greatly diminished. Confounding the issue, however, we have to expect that large particles will shatter in our instrument duct. Model sensitivity tests show that our measurements are impacted by sublimation at the ~5% level only when the ~10 μm particle number densities are increased (above the already inflated measured values) by an additional factor of five. As a result, we do not expect these uncertainties to substantially impact our estimates of the sublimation contribution, and feel confident that the measurement bias introduced by ice particle sublimation in our instrument duct is <+5%.

Measurements

Initially it was hoped that direct comparisons with the open path JLH instrument would provide an empirical method for quantifying the net effect of contamination on an event-by-event basis. The JLH instrument is mounted in the free stream under the aircraft wing, and is not impacted by ice particle sublimation. However, signal-to-noise limitations of both instruments, combined with flight-to-flight and mission-to-mission variability in their agreement, rendered a rigorous quantitative analysis impossible. Furthermore, during the CRYSTAL-FACE mission, JLH was unable to measure water vapor during moderate to thick cloud events because of enhanced noise, possibly due to charge

separation on cloud particles [Robert Herman, *personal communication*, 2002]. Despite these limitations, judicious comparisons of our measurements inside and outside of clouds with simultaneous measurements made by JLH provide an additional means of constraining an in-cloud bias.

Figure 3.11 shows plots of JLH clear air (blue circles) and in-cloud (red stars) data versus HWV from the CRYSTAL-FACE and Middle latitude Cirrus Experiment (MidCIX) missions respectively, with corresponding linear least squares fits. Significant sublimation in HWV should lead to a separation between the clear air and in-cloud data in each plot, with a corresponding difference in the slopes of the least squares fits. With JLH on the y-axis and HWV on the x-axis, the magnitude of the in-cloud slope is expected to decline as a result of contamination in HWV from sublimation. Linear fits to the clear air and in-cloud data from CRYSTAL-FACE mission indicate a possible high bias in the HWV in-cloud data. The slope of the fit to the in-cloud data is lower than the slope of the clear air data by ~3%. However, a more detailed analysis of the differences between the in-cloud HWV and JLH data shows no dependence upon the magnitude of the cloud ice water content, or on cloud particle number density, surface area or volume. If the systematic bias were due to sublimation in HWV, we would expect it to demonstrate a dependence on these factors. Furthermore, the comparable comparison of MidCIX data, shown in Figure 8.11b, yields an apparent bias of ~2%, in the opposite direction. In the end, we conclude that the combined uncertainty in the JLH and HWV clear air and in-cloud measurements is probably larger than any systematic bias in the HWV in-cloud data, and that the

net effect of sublimation suggested by the model and confirmed by the in situ comparisons is well within +5%.

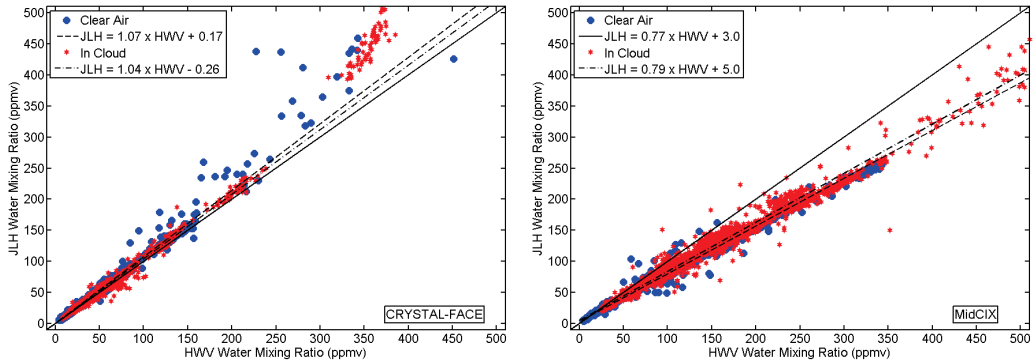


Figure 3.11. Comparison of HWV and JLH clear air and in-cloud data. Clear air (blue circles) and in-cloud (red stars) data measured by JLH are plotted versus HWV. Panel (a) shows data from the CRYSTAL-FACE mission, and panel (b) shows data from the MidCIX mission. The JLH instrument is an open path instrument mounted under the wing of the aircraft and is not subject to contamination from particle sublimation. Significant sublimation in HWV would cause a measureable decline in the in-cloud slope relative to the clear air slope. No such signal is evident in either comparison.

3.5 In Situ Instrument Comparisons

While much of the burden of in flight validation rests on carefully designed diagnostic tests, comparisons with other instruments on board the aircraft, like those with JLH in the preceding sections, may also be used to evaluate instrument performance. In this section HWV flight data are compared with measurements made simultaneously by a select set of instruments, each utilizing a unique and independent measurement method and/or sampling strategy.

3.5.1 HWV vs. HTW

We begin by comparing measurements made simultaneously by HWV and the Harvard Total Water (HTW) instrument [*Weinstock et al.*, 2006a; *Weinstock et al.*, 2006b], which measures both the vapor and condensed phase water content of the ambient air. HWV and HTW have identical detection axes, which are often calibrated simultaneously in the laboratory under identical conditions. However, fundamental differences in their in situ sampling strategies lead to extremely different operational conditions in flight. Comparisons between HTW and HWV, therefore, provide complementary data on the in situ performance of the fluorescence detection method.

The HTW instrument includes an isokinetic inlet for accurate cloud particle sampling, and a 600 watt heater that fully evaporates water in the condensed phase prior to detection. The combination of heater elements positioned in the flow and additional heat supplied to the walls of the sampling duct raises the temperature of the air sample to ~+15 ºC, or ~90 degrees above ambient by the time it reaches the detection region. In contrast, flow temperatures in HWV are elevated to a maximum of 20 degrees above ambient due to ram heating, and generally measure from −70 ºC to −40 ºC. Because temperatures within HTW remain very close to those of the laboratory, comparisons with HTW provide another excellent metric for empirically evaluating the stability of the HWV calibration as temperatures in the duct and detection axis plummet to near ambient conditions in flight.

Percent differences between HTW and HWV clear air data from the Pre-Aura Validation Experiment (PreAVE) mission are plotted as function of ambient temperature in Figure 3.12a. Also shown is the average difference in each 4 K temperature bin over the full range, i.e., from 186 K to 276 K. Figure 3.12a shows that the two detection axes, which were calibrated with respect to the same standards in the laboratory, agree to within their combined uncertainty (~7%) in the flight environment[2], thus confirming the temperature independence of the fluorescence measurement.

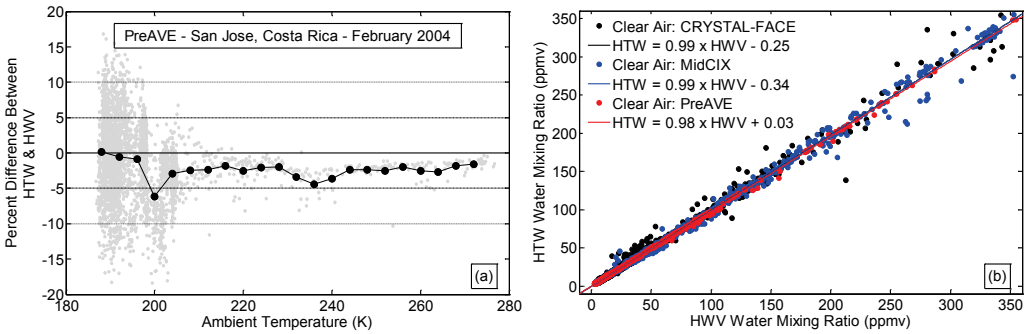


Figure 3.12. Comparisons of HWV and HTW clear air data. Panel (a) shows the percent difference between HTW and HWV clear air data (gray circles) from the PreAVE mission as a function of ambient temperature. The average value for each 4 K temperature bin is also shown (black circles). Panel (b) shows HTW clear air data plotted as function of HWV clear air data from the CRYSTAL-FACE (black circles), MidCIX (blue circles) and PreAVE (red circles) missions. The colored lines correspond to robust linear least squares fits to the data, again demonstrating the excellent agreement between this pair of instruments.

Figure 3.12b shows direct comparisons of HTW and HWV clear air data from the CRYSTAL-FACE, MidCIX, and PreAVE missions. The results of

[2] We attribute much of the short-term variability in their agreement to systematic errors in HTW, where the instrument walls can act as both a temporary source of, and sink for water, because of the combination of 1) the high surface area of the long heated inlet and 2) the comparatively low mass flow rate through the HTW instrument.

robust linear least squares fits to the data are also plotted, and they quantify the excellent agreement (~<2%) between HWV and HTW from mission to mission over two orders of magnitude in water vapor, and the full range of UT/LS pressures and temperatures.

3.5.2 HWV vs. HOXotope and ICOS

We turn now to a discussion of results from the Aura Validation Experiment – Water Isotope Inter-comparison Flights (AVE-WIIF) campaign flown out of Houston, TX in July of 2005. This mission provided an unprecedented opportunity for the in-flight comparison of four laboratory-calibrated Harvard University hygrometers. In three flights during the campaign the four Harvard hygrometers made simultaneous measurements of water vapor in the UT/LS aboard the WB-57 aircraft. Just as the choice of dual calibration standards serves to establish accuracy and minimize bias in the laboratory calibration, the comparison of water vapor instruments employing distinct detection methods and sampling strategies provides the means to substantiate measurement accuracy in situ. The calibration of each instrument to the same laboratory standard, or pair of standards, does not compromise the independence of the measurements in flight. Rather, establishing agreement in the laboratory is the first step toward achieving accurate in situ measurements as it provides a firm basis for diagnosing potential systematic errors that arise in the implementation of each detection method under ambient conditions.

The four Harvard instruments included HWV and HTW with the addition of the "HOXotope" and the Integrated Cavity Output Spectroscopy, ICOS, instruments. HOXotope and ICOS utilize two fundamentally different detection methods to measure water vapor and its isotopologues [*Sayres et al.*, 2009; *St Clair et al.*, 2008]. In brief, the HOXotope instrument photolyzes water with a Xenon excimer lamp at 172 nm and measures the OH (or OD) fragment by laser-induced fluorescence, and the ICOS instrument measures water vapor via cavity-enhanced absorption with a quantum cascade laser tuned to a water absorption line in the mid-infrared.

Following the methodology of HWV and HTW, these instruments calibrate in the laboratory with respect to two independent traceable standards. Calibrated water addition for HOXotope and ICOS is achieved using 1) the bubbler apparatus, described above, which is tied to the vapor pressure of water over liquid at room temperature, and 2) a micro-droplet injector, which avoids isotopic fractionation during the calibration procedure by introducing bulk water to the flow in drops of a fixed size (~60 μm diameter) at a set rate. The injector calibration is determined gravimetrically by collecting the liquid dispensed at a known rate over the period of an hour and weighing the sample to determine the mean droplet water mass [*Sayres et al.*, 2009; *St Clair et al.*, 2008]. Additionally, the temperatures and pressures of the ICOS optical bench and absorption cell and the HOXotope lasers and detection region are under active control in the laboratory as well as in flight. As a result, both instruments are calibrated in the laboratory under pressure, temperature and flow conditions nearly identical to

those of flight, thus reducing the uncertainties associated with the transition to the aircraft platform.

During AVE-WIIF, the accuracy of HOXotope was estimated to be ~7%, based upon the ~5% accuracy of the calibration, plus additional uncertainties arising from low laser power and poor precision. Similarly, the accuracy of ICOS was estimated to be ~7%, arising from a combination of the accuracy of the calibration, and uncertainties due to noise in the raw spectra. Non-white noise structures in the baseline can produce additive errors that chiefly impact the accuracy at low water vapor concentrations.

Figure 3.13 shows the results of comparisons between HOXotope (blue) and HWV, and ICOS (red) and HWV from the flight of August 3, 2005 during the AVE-WIIF mission. Data from the first hour and a half of each flight were excluded because of the potential for contamination on ascent in both ICOS and HOXotope[3]. The plot also shows the results of robust linear regressions applied to the two data sets. The plots show the better than 5% agreement among the suite of Harvard instruments. Because both HOXotope and ICOS operate at nominally the same pressure and temperature in flight as in the laboratory, the agreement demonstrated in Figure 3.13 once again demonstrates that 1) the pressure dependence, i.e., quenching, of the HWV calibration is correct, and that 2) the HWV detection axis is insensitive to the difference between the laboratory and in situ temperatures. Furthermore, this figure shows the linearity of all three

[3] Contamination in these instruments on ascent was a recognized problem. Water shedding off of the aircraft fuselage was considered to be a dominant source of this contamination [*St Clair et al.*, 2008]. The inclusion of data from the first hour and a half impacts the intercept by ~+0.50 ppmv, consistent with possible contamination, but does not significantly impact determination of the slope.

techniques over two orders of magnitude in water vapor mixing ratio. This level of agreement provides confirmation that the full complement of Harvard hygrometers is capable of making in situ measurements of water vapor within their stated accuracies.

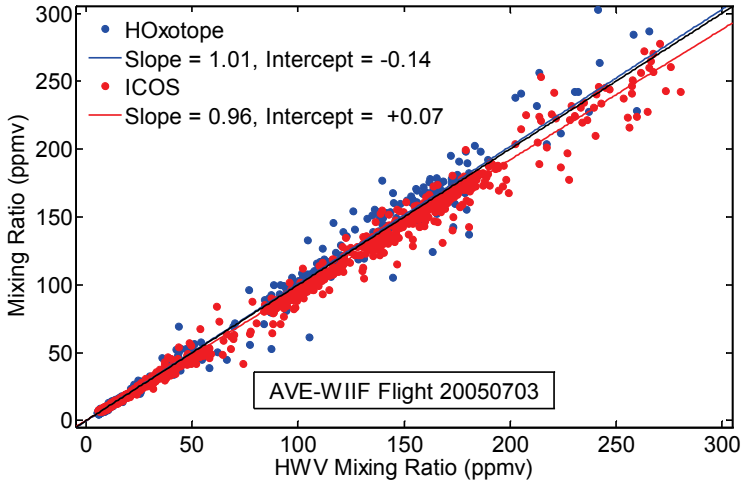


Figure 3.13. Comparison of HWV, HOXotope and ICOS water vapor data from the AVE-WIIF mission. HOXotope (blue) and ICOS (red) water vapor mixing ratios are plotted as a function of HWV. Also shown are the results of robust linear regressions, which quantify the excellent agreement between these instrument pairs over two orders of magnitude in water vapor mixing ratio.

3.5.3 HWV vs. JLH

Unfortunately, direct comparisons between HWV and JLH, two instruments that utilize radically different measurement and sampling methods, have not been possible since the CRYSTAL-FACE mission in 2002. In recent missions, JLH has experienced a host of undocumented issues associated with laser failures, varying spectroscopic water lines, improperly characterized laser linewidths, noise in the laser power supply, and uncertainties associated with the laser modulation necessary for the second harmonic detection. All JLH data

archived from 2004 to the summer of 2007 are preliminary, invalidating any direct comparison with HWV over this period [Robert Herman and Robert Troy, *personal communication*].

3.6 Use of a priori Knowledge as a Constraint

Validation of instrument performance based primarily upon agreement with a priori expectations of atmospheric phenomena is untenable, however, examples such as the one presented here do increase confidence in instrument performance[4]. During the MidCIX mission, flown out of Houston, TX during April and May of 2004, the WB-57 flew through an orographic wave cloud over north central Colorado. Wave clouds provide an opportunity to examine ice nucleation and cloud formation in the atmosphere under well-constrained physical conditions [*Heymsfield and Miloshevich*, 1995]. Based on the high surface area densities within a wave cloud, it is expected, that water vapor within the cloud should be in near equilibrium with ice, i.e., the relative humidity with respect to ice, RH_i, within the cloud should be equal to 100%. Using this knowledge, deviations away from 100% may be attributed to measurement errors.

Figure 3.14 shows RH_i results from the wave cloud encounter on the flight of May 5, 2004, using HWV, and temperature and pressure from the Meteorological Measurement System (MMS) [*Scott et al.*, 1990]. RH_i is plotted as a function of ice particle number density as measured by the Forward Scattering

[4] Other examples using a prior knowledge as a constraint on the accuracy of water vapor measurements include the hydrogen budget analysis (H_2O + 2 x CH_4) of *Hurst et al.*, 1999, the ice crystal growth analysis of *Jensen et al.*, 2008, and the RH_i analysis of *Kramer et al.*, 2009.

Spectrometer Probe (FSSP). The peak of the particle size distribution is centered at 10 to 20 μm, and the number density of 10 μm particles is comparable to that observed in thick cirrus. The data show that as the number density, and thus surface area, increases in the cloud, RH_i converges toward ~104%. This 4% deviation from 100% is well within the combined uncertainties of HWV, and the ambient temperature and pressure measurements. Though this observation helps to verify the accuracy of HWV in situ, it does not take the place of carefully executed laboratory calibrations, with accuracies defined by accepted traceable standards. It simply provides further evidence of the applicability of our laboratory calibration to flight conditions, and supports the conclusion that the net effect of sublimation in the duct is likely negligible, and reliably <5%.

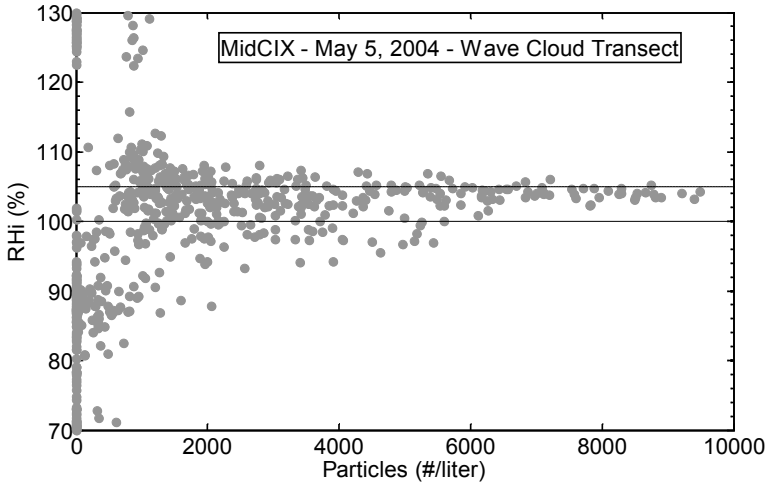


Figure 3.14. RH_i in a wave cloud. RH_i, determined from HWV measurements and ambient temperature and pressure converges to ~100% as the particle number density reaches its peak, consistent with theoretical expectations of the humidity within these clouds.

3.7 Conclusions

In this chapter, we outlined the foundations of our laboratory calibration and in-flight validation strategy, and demonstrated the practical realization of this strategy through a step-by-step description of the process applied to the HWV detection axis. Calibrations of HWV referenced to two independent standards, i.e., 1) the saturation vapor pressure of water over liquid at room temperature and pressure, and 2) the direct absorption of 121.6 nm light, demonstrated agreement to ~3%, well within their combined uncertainty, thereby validating the accuracy of both the HWV detection axis and the calibration system. The results of a representative sample of laboratory calibrations and sensitivity tests of HWV further established that we can achieve this accuracy and precision under the pressure, temperature and water vapor mixing ratio regimes encountered in flight. An examination of actual flight data, and the results of focused in-flight diagnostics confirmed, 1) the validity of the laboratory calibration in situ, and 2) the absence of wall contamination during in situ sampling. Additionally, comparisons with a select set of independent airborne instruments, utilizing radically different detection methods and sampling techniques, were used to confirm 1) the stability of the laboratory calibration, 2) the temperature, pressure, flow, and ice particle insensitivity of the HWV detection axis, and 3) the linearity of the photo-fragment fluorescence detection method from a few to several hundred ppmv water vapor. Finally, we showed data that were consistent with a priori knowledge of the evolution of RH_i in a wave cloud, further corroborating the robust in situ performance of HWV.

References

Baumgardner, D., H. Jonsson, W. Dawson, D. O'Connor, and R. Newton (2001), The cloud, aerosol and precipitation spectrometer: a new instrument for cloud investigations, *Atmos Res*, *59*, 251-264.

Field, P. R., A. J. Heymsfield, and A. Bansemer (2006), Shattering and particle interarrival times measured by optical array probes in ice clouds, *J Atmos Ocean Tech*, *23*(10), 1357-1371.

Heymsfield, A. J., and L. M. Miloshevich (1995), Relative humidity and temperature influences on cirrus formation and evolution: Observations from wave clouds and FIRE II, *J Atmos Sci*, *52*(23), 4302-4326.

Hintsa, E. J., E. M. Weinstock, J. G. Anderson, R. D. May, and D. F. Hurst (1999), On the accuracy of in situ water vapor measurements in the troposphere and lower stratosphere with the Harvard Lyman-alpha hygrometer, *J Geophys Res-Atmos*, *104*(D7), 8183-8189.

Hurst, D. F., et al. (1999), Closure of the total hydrogen budget of the northern extratropical lower stratosphere, *J Geophys Res-Atmos*, *104*(D7), 8191-8200.

ISO-5725-1 (1994), *Accuracy (trueness and precision) of measurement methods and results : part 1 : general principles and definitions*, vi, 17 p. pp., International Organization for Standardization, Genève, Switzerland.

Jensen, E. J., et al. (2005), Ice supersaturations exceeding 100% at the cold tropical tropopause: implications for cirrus formation and dehydration, *Atmos Chem Phys*, *5*, 851-862.

Jensen, E. J., et al. (2009), On the importance of small ice crystals in tropical anvil cirrus, *Atmos Chem Phys*, *9*(15), 5519-5537.

Jonsson, H. H., et al. (1995), Performance of a Focused Cavity Aerosol Spectrometer for Measurements in the Stratosphere of Particle Size in the 0.06–2.0-μm-Diameter Range, *J Atmos Ocean Tech*, *12*(1), 115-129.

Korolev, A., and G. A. Isaac (2005), Shattering during sampling by OAPs and HVPS. Part I: Snow particles, *J Atmos Ocean Tech*, *22*(5), 528-542.

Kramer, M., et al. (2009), Ice supersaturations and cirrus cloud crystal numbers, *Atmos Chem Phys*, *9*(11), 3505-3522.

Lewis, B. R., I. M. Vardavas, and J. H. Carver (1983), The Aeronomic Dissociation of Water-Vapor by Solar-H Lyman Alpha-Radiation, *J Geophys Res-Space*, *88*(Na6), 4935-4940.

Murphy, D. M., and T. Koop (2005), Review of the vapour pressures of ice and supercooled water for atmospheric applications, *Q J Roy Meteor Soc*, *131*(608), 1539-1565.

Sayres, D. S., et al. (2009), A new cavity based absorption instrument for detection of water isotopologues in the upper troposphere and lower stratosphere, *Rev Sci Instrum*, *80*(4).

Schwab, J. J., E. M. Weinstock, J. B. Nee, and J. G. Anderson (1990), Insitu Measurement of Water-Vapor in the Stratosphere with a Cryogenically Cooled Lyman-Alpha Hygrometer, *J Geophys Res-Atmos*, *95*(D9), 13781-13796.

Scott, S. G., T. P. Bui, K. R. Chan, and S. W. Bowen (1990), The Meteorological Measurement System on the Nasa Er-2 Aircraft, *J Atmos Ocean Tech*, *7*(4), 525-540.

St Clair, J. M., et al. (2008), A new photolysis laser-induced fluorescence instrument for the detection of H2O and HDO in the lower stratosphere, *Rev Sci Instrum*, *79*(6).

Weinstock, E. M., J. B. Smith, D. Sayres, J. V. Pittman, N. Allen, and J. G. Anderson (2006a), Measurements of the total water content of cirrus clouds. Part II: Instrument performance and validation, *J Atmos Ocean Tech*, *23*(11), 1410-1421.

Weinstock, E. M., E. J. Hintsa, A. E. Dessler, J. F. Oliver, N. L. Hazen, J. N. Demusz, N. T. Allen, L. B. Lapson, and J. G. Anderson (1994), New Fast-Response Photofragment Fluorescence Hygrometer for Use on the Nasa Er-2 and the Perseus Remotely Piloted Aircraft, *Rev Sci Instrum*, *65*(11), 3544-3554.

Weinstock, E. M., et al. (2006b), Measurements of the total water content of cirrus clouds. Part I: Instrument details and calibration, *J Atmos Ocean Tech*, *23*(11), 1397-1409.

Chapter 4

Investigation of Measurement Offsets

4.1 Introduction

In view of the long standing discrepancies among the water measurements, discussed in Chapter 1, and in direct response to the results of the CRAVE mission, which led to what has become the canonical figure highlighting these differences (Figure 1.5), we turn now to the question of measurement offsets in HWV.

Our calibrations, the subject of Chapter 3, establish an accuracy of ~±5% for the measurement sensitivity of the HWV detection axis, and a judicious combination of laboratory tests and in-flight diagnostics show that the accuracy achieved in the laboratory is reproducible in flight. A limitation of our current calibration system, however, is that neither reference standard is capable of establishing the absolute mixing ratio in the duct. As a result, we do not have a straightforward means of identifying and quantifying the non-zero background mixing ratio and/or any instrumental offset or measurement bias.

The bubbler standard accurately determines the water vapor added to the duct flow, but it cannot account for water vapor contamination in the duct from other sources. Similarly, the axial absorption standard is only capable of

measuring the *change* in the water concentration during a titration, i.e., $\Delta[H_2O] = ([H_2O]_{add} + [H_2O]_{bkgd}) - [H_2O]_{bkgd}$. not the absolute concentration. This is because this method relies on the periods prior to and following the titration to obtain a baseline.

In Equation 3.1b, we attributed all of the non-zero intercept to background water vapor in the duct. However, there is also the possibility that we have a small, but non-zero instrumental artifact/offset. The impact of a measurement offset on the accuracy of our measurements depends upon both the magnitude of the potential offset and the ambient water vapor mixing ratio, with the greatest effect expected at the lowest mixing rations, e.g., below ~5 ppmv. This chapter provides a detailed discussion of the evidence for and against an instrumental offset based on the results of extensive laboratory tests and a re-examination of flight data.

4.2 Evaluation of Water Contamination

In the simplest terms, a measurement offset results from PMT counts that are not explicitly due to the desired water vapor fluorescence. These additional counts may arise from either a spurious source of water vapor, or an instrumental artifact.

In the flight environment, an offset arising from water vapor contamination is considered to be a sampling problem, e.g., due to hysteresis or sublimation, rather than a detection axis bias. In-flight diagnostics, discussed at length in Chapter 3, have been devised to explicitly test for these sources of

parasitic water, and the results of these tests have repeatedly demonstrated that the ambient sample is not contaminated by either the instrument walls, or the sublimation of ice particles within the duct. In fact, the fast flow through the instrument duct (~80 m/s), makes the measurement essentially insensitive to these sources, and qualifies HWV as an in-flight comparison standard for instruments that do require a short (typically <1 hour) "dry out" after takeoff, e.g., ICOS, HOxotope, and HTW [*Sayres et al.*, 2009; *St Clair et al.*, 2008; *Weinstock et al.*, 2006; *Weinstock et al.*, 2009].

In the laboratory, a non-zero water vapor background may be due to contamination from the calibration source gas, and/or leaks or outgassing from the addition lines and instrument walls. However, as we showed in Chapter 3, the accurate determination of the calibration constants is effectively insensitive to a nominally constant water vapor background. In fact, following calibration, the fluorescence data can be analyzed to obtain the absolute mixing ratio in the duct, consequently providing a measure of the background water vapor mixing ratio. The greatest impact, therefore, of a non-zero background due to water vapor in the laboratory is that it obscures the detection and quantification of a true instrumental artifact.

4.3 Evaluation of Instrumental Artifact

Two distinct mechanisms can lead to systematic measurement biases in the fluorescence detection method: 1) 315 nm radiation from the lamp that is scattered in the duct and not properly quantified, and 2) VUV stimulated

fluorescence from a non-water vapor source. In order to understand and identify these potential sources of error, we once again review the core of HWV detection technique.

The filter assembly in front of the PMT limits the detection of radiation to a relatively narrow band centered at 315 nm, i.e., near the peak of OH* fluorescence (see Chapter 2, Figure 2.2). An anti-reflection coated quartz window mounted on a mechanical actuator, "chopper", is used to measure the 315 nm background light from the lamp in the absence of fluorescence. When the quartz window is positioned in front of the lamp it blocks (fully absorbs) the 121.6 nm light, thus preventing water photolysis and OH fluorescence, and the measured background is attributed entirely to lamp scatter. The chopper is cycled in and out of the lamp path, allowing for continuous monitoring of the 315 nm background. Figure 4.1 shows a one-minute segment of raw PMT data (black dots) acquired in the laboratory. The "OH Signal + Background" data are highlighted in red circles, and the blue line represents the interpolated "Background" signal. The difference, equal to "OH Signal", is essentially proportional to the water vapor mixing ratio in the duct, see Equation 2.5.

Errors in the determination of the "OH Signal" can lead to systematic measurement offsets. Positive offsets, leading to erroneously high measurements of water vapor, can occur if the signal measured with the chopper out is too high, or conversely, if the background measurement made with the chopper in is too low. Similarly, negative offsets, leading to erroneously low measurements of

water vapor, can occur if the chopper out signal is too low or the chopper in signal is too high.

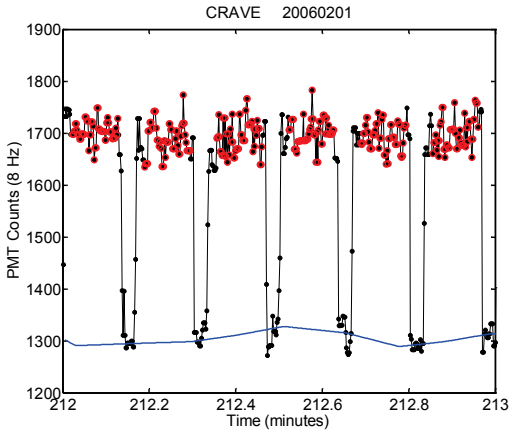


Figure 4.1. Raw PMT counts as a function of time. This figure shows one minute of raw PMT signal (black dots and line) acquired in situ during the CRAVE mission. The red circles highlight the ~315 nm signal with the chopper out, OH Signal + Background, and the blue line shows the interpolated 315 nm background signal measured with the chopper in, Background.

Equation 4.1 provides the framework for understanding the sources of these offsets.

$$\Delta S = S_{out} - S_{in} = \{ SCT_{out} + [FLR_{H2O} + FLR_X] \} - \{ (1 + f_{chop}) \times SCT_{in} \}$$

$$= \Delta SCT_{out\text{-}in} + [FLR_{H2O} + FLR_X] \sim [H_2O] @ M \qquad (4.1)$$

ΔS is the difference between the measured PMT counts with the chopper out, S_{out}, and the chopper in, S_{in}. The PMT counts with the chopper out are fully represented by the sum of the lamp scatter, SCT_{out}, the fluorescence signal from OH* due to water vapor photolysis in the duct, FLR_{H2O}, and contributions from any non-water vapor source of VUV stimulated fluorescence, FLR_X. The PMT counts with the chopper in are simply attributed to background lamp scatter, SCT_{in}, and the pre-factor $(1 + f_{chop})$ is included to account for the attenuation of

the lamp scatter by the chopper window. Recall f_{chop} is the chopper reflectivity factor, see Chapter 2, Equation 2.7. Note there is no contribution from VUV stimulated fluorescence from water vapor or any other source with the quartz window in front of the lamp. In order to achieve an accurate bias-free water vapor measurement, we want the terms, $\Delta SCT_{out\text{-}in}$ and FLR_X to be equal to zero.

4.3.1 Offsets Associated with Lamp Scatter ($\Delta SCT_{out\text{-}in}$)

4.3.1.1 Theory

The potential for lamp scatter to contribute to positive or negative offsets is well understood, and we correct for it. However, uncertainties associated with the correction, specifically in the determination of f_{chop}, can lead to small systematic biases. Therefore, in the interest of thoroughness, we discuss the process of determining the scatter correction, and the uncertainties associated with it here.

Scattered 315 nm light from the lamp can generate a positive offset if the background scatter is lower with the chopper in than with the chopper out, i.e., $\Delta SCT_{out\text{-}in} > 0$. Indeed, under normal circumstances, the ~315 nm photon flux entering the duct is lower with the chopper window positioned in front of the lamp because ~1% of the photon flux is reflected per surface of the anti-reflective coated quartz window. Thus, if the measured background signal, SCT_{in}, is low by ~2% the difference $\Delta SCT_{out\text{-}in}$ will equal $0.02 \times SCT_{out}$, resulting in a small positive offset to the measurement. As was discussed briefly in Chapter 2, we account for this potential offset by multiplying SCT_{in} by the factor, $1 + f_{chop}$, with

$f_{chop} \sim 0.02$, such that the term $\Delta SCT_{out\text{-}in} = 0$. However, if the chopper reflectivity factor, f_{chop}, is too low/high the resulting correction will be too low/high leading to measurements that are systematically too high/low. Uncertainties in the empirical determination of f_{chop} dictate the uncertainty associated with this correction, and may result in small positive or negative biases to the measurement.

4.3.1.2 Laboratory Quantification

We have traditionally utilized two methods for determining the chopper reflectivity. The first method measures the chopper reflectivity in the absence of a fluorescence signal. At room pressure (1 atm) and humidity it is assumed that virtually no VUV makes it into the detection volume in the center of the duct, and consequently there is no fluorescence signal at the PMT. Furthermore at one atmosphere the quenching rate is very high. Thus, the 315 nm signal measured by the PMT with the chopper out is assumed to be solely due to lamp scatter, and the difference between the scattering signal with the chopper out and in can be measured directly.

Figure 4.2a shows data from a diagnostic run executed on February 4, 2006, during the CRAVE mission. The figure shows the empirically determined fraction f_{chop}, i.e., $\{SCT_{out} - SCT_{in}\} / SCT_{in}$ as a function of time. The data are noisy, but the results are generally consistent with a value of ~1.5% ± 1.0%. During the final 4 minutes of the run the lamp heater set point was increased leading to a small fluorescence signal. The data show that the contribution of a non-zero fluorescence signal leads to a systematic overestimate of f_{chop}. This

method of determining f_{chop}, therefore, provides an upper limit to the correction factor, i.e., it is unlikely that the true value for the chopper reflectivity will be higher than the lowest average value (~1.5%) measured by this method.

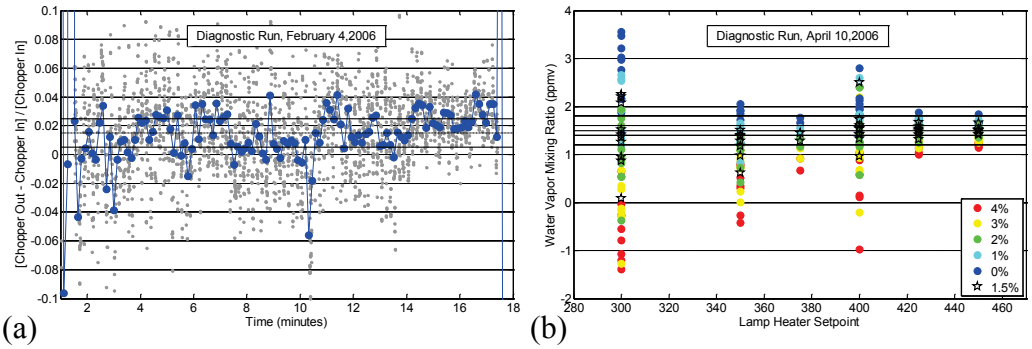


Figure 4.2. Chopper window reflectivity measurements. Panel (a) shows the fractional difference between the measured PMT signal with the chopper "out" and with the chopper "in" under conditions where the fluorescence contribution is effectively zero. The signal with the chopper "out" is ~1.5% higher than with the chopper "in". Panel (b) shows water vapor measurements as a function of the lamp heater setpoint where different values for the chopper window reflectivity correction have been applied (color code). The true mixing ratio is constant, and the measurements converge to the true value, ~1.5 ppmv, at the highest lamp heater setpoint. Thus, the correct chopper reflectivity value (~1.5%) is that which yields a constant mixing ratio over the full range of lamp heater setpoints.

The second method of quantifying f_{chop} involves measuring a constant water vapor mixing ratio at a constant pressure while varying the VUV flux by changing the lamp heater set point. This method relies on the theory that changing the lamp heater set point should have no effect on the water vapor measurement. Recall that the calibration determines the sensitivity of the detection axis to water vapor in terms of fluorescence counts normalized to the VUV flux as measured by the diode across the duct from the lamp. If we have not properly accounted for the chopper reflectivity, however, the scatter dependent offset term, $SCT_{out} - (1 + f_{chop}) \times SCT_{in} = \Delta SCT_{out\text{-}in}$, will be non-zero, and its proportional contribution to

the measured mixing ratio will vary as the lamp flux is varied. At low lamp flux, i.e., comparatively high background scatter and low fluorescence signal, if f_{chop} is too low we will measure a significantly higher value than the true mixing ratio. Similarly, if f_{chop} is too high under these conditions, we will overcorrect and our measured mixing ratio will be lower than the true value. As the lamp flux increases, however, the fluorescence signal rapidly increases, and the net contribution resulting from errors in the background scattering term, $\Delta SCT_{out\text{-}in}$, decreases. Thus, at high VUV flux we expect the measurement to converge toward the true value. Only when the correction for the chopper reflectivity, f_{chop}, is correct, i.e., when $SCT_{out} - (1 + f_{chop}) \times SCT_{in} = \Delta SCT_{out\text{-}in} = 0$, will we achieve a constant water vapor measurement over the full range of the VUV lamp flux.[1]

Figure 4.2b shows data from a chopper reflectivity diagnostic run performed on April 10, 2006, after the CRAVE mission. The lamp heater set point was varied while sampling a constant low water vapor mixing ratio, ~1.5 ppmv. The figure shows the measured mixing ratio at six different heater set points, assuming five different chopper reflectivity factors. Examination of the figure shows that a value of $f_{chop} = 1.5\% \pm 0.5\%$ (the black dots in the figure) yields the most consistent measurement of the mixing ratio across the full range of the VUV flux settings, thus we conclude that this is the best estimate for the chopper

[1] This method is subject to two underlying assumptions: 1) first we assume that raising/lowering the lamp heat will increase/decrease the magnitude of both the fluorescence signal and the background scatter. However, we expect the response of the background scatter to be far less sensitive to changes in the lamp heater set point than that of the VUV flux. If the VUV flux and the 315 nm background signal responded in a similar fashion we would not observe a change in the offset, 2) we further assume that the fractional correction for the difference between SCT_{out} and SCT_{in} is a constant, i.e., that the proportion of scattered light, f_{chop}, that gets blocked by the chopper window does not change with lamp flux.

reflectivity correction. While neither method for the determination of f_{chop} is free from error, the consistency between the results of both methods increases our confidence in the accuracy of this subtle correction factor.

4.3.1.3 Effect on Measurement

Uncertainty in the determination of f_{chop} and the resulting non-zero contributions from $\Delta SCT_{out\text{-}in}$, have the potential to lead to offsets in the measurement. The magnitude of this offset, however, depends upon 1) the magnitude of the chopper window reflectivity, 2) our ability to accurately quantify the reflectivity, and 3) the relative contribution of the background scatter signal to the total ("OH Signal" + "Background") signal measured. Under conditions of high background scatter and a low sensitivity to water vapor, we expect errors in the scattering correction to lead to larger magnitude offsets. However, if the background scatter is low and the sensitivity to water vapor is high we expect errors associated with the background subtraction to be comparatively small, and any resulting offset to be below the instrument precision. Note that in flight the VUV flux from the lamp is typically set to achieve a high signal-to-noise ratio for the measurement, and simultaneously minimize errors associated with the background scatter and chopper reflectivity correction.

Data acquired by HWV during the CRAVE mission provide the means of quantifying the net impact of offsets resulting from the lamp scatter correction under far from optimal conditions. The combination of extremely low ambient

mixing ratios (~< 2 ppmv), and poor instrument performance offer an opportunity for quantifying the maximum effect uncertainties in f_{chop} are likely to have on the measurement.

During the first four flights of the second half of the CRAVE mission two factors compromised the performance of the HWV detection axis, 1) unusually low Lyman-α lamp flux[2], and 2) higher than expected background scatter in the duct. The resulting low signal to noise ratio during these flights severely reduced the measurement precision (see Chapter 2), and made the measurement more sensitive to systematic errors associated with the background scatter. During these flights, varying f_{chop} from ~0.01 to 0.02, yielded deviations on the order of ±0.2 ppmv, resulting in ~10% uncertainties in the measurement at the lowest mixing ratios (~2.0 ppmv). Though the accuracy of the HWV calibration was shown to be robust over the course of the mission, as corroborated by pre- and post-mission laboratory calibrations and in-flight absorption-fluorescence comparisons, the accuracy of the measurement was compromised on four flights by the low signal and high background of the detection axis. In contrast, during the last two flights of the CRAVE mission, the HWV detection axis was replaced with a backup axis. In the backup axis, the combination of high VUV lamp flux and a low background signal decreased the sensitivity of the measurement to uncertainties in f_{chop} by more than a factor of 5, effectively rendering these uncertainties inconsequential.

[2] The gain of the VUV diode located across from the lamp was set too high, and at reasonable lamp heater set points the diode signal saturated. In order to keep the diode signal on scale, we were forced to run the lamp heater at set point much lower than its optimal value.

In conclusion, the uncertainties associated with scattering correction are known, can be quantified and, most importantly, can be minimized, such that they do not (typically) impact the performance of the detection axis. Even in the most unfavorable circumstances, the offsets associated with the scattering correction are small (~<0.2 ppmv), and have an almost equal likelihood of being either positive or negative. There is no way that these uncertainties can explain the long-standing systematic difference between the in situ measurements, e.g., between HWV and CFH.

4.3.2 Spurious Fluorescence (FLR_X)

4.3.2.1 Theory

Equation 4.1 shows that positive offsets may also be associated with a VUV stimulated fluorescence signal that mimics a water vapor signal, denoted FLR_X. Possible sources include a non-water vapor gas contaminant, or certain optics, and/or painted surfaces in the instrument duct that fluoresce in the bandpass of the PMT collection optics when irradiated with VUV light. Identifying and accurately quantifying offsets of this type, which we expect to be ~< 1 ppmv, however, is challenging.

In order to better understand the difficulty of detecting and measuring an instrumental bias, FLR_X, we enumerate the terms contributing to the measured fluorescence signal, FLR_{TOT}, in Equation 4.2.

$$FLR_{TOT} = FLR_{H2O} + FLR_X \qquad (4.2)$$
$$= [FLR_{H2Oadded} + FLR_{H2Oflowbkgd} + FLR_{H2Oleak} + FLR_{H2Owalls}] + FLR_X$$

In this equation, the term, FLR_X, is the only term that is a true instrumental offset, and the only term that is expected to impact measurements in situ. The rest of the terms, which are unique to the laboratory environment, are 1) water vapor intentionally added to the flow, $FLR_{H2Oadded}$, 2) residual water vapor present in the carrier flow, $FLR_{H2Oflowbkgd}$, 3) water vapor due to leaks into the flow tube, $FLR_{H2Oleak}$, and 4) water vapor due to outgassing from the flow tube walls and plumbing upstream of the detection axis, $FLR_{H2Owalls}$. During calibrations, when the bubbler flow is zero, i.e., when $FLR_{H2Oadded} = 0$, it has generally been assumed that the majority of the residual non-zero signal is due to water vapor fluorescence from one or more of the remaining three water vapor sources, i.e., not due to an instrumental bias.

One approach to quantifying the magnitude of real fluorescence artifact, FLR_X, therefore, is to accurately quantify the sum of the contribution of the residual water vapor terms. This approach is challenging because it requires an independent reference for the water vapor mixing ratio present in the instrument duct that can measure accurately at and below ~1 ppmv. A second, less satisfactory but more practicable approach is to bound the magnitude of FLR_X by minimizing the contribution of each of the water vapor terms, thus providing an upper limit to the magnitude of any potential instrumental offset.

Hintsa et al., [1999], ran diagnostic tests on the 3" Lyman-α water vapor detection axis, which flew on NASA's ER-2 aircraft from 1993 through 2000 (HWV-E), utilizing the second approach [*Hintsa et al.*, 1999; *Wennberg et al.*, 1994]. The ER-2 detection axis was checked for spurious offsets by flowing large

amounts of dry nitrogen, generated from the boil-off of a liquid nitrogen (LN_2) Dewar, through the flow tube. The results yielded a minimum fluorescence signal equivalent to ~0.3 ppmv H_2O in air. Furthermore, in keeping with Equation 4.2, *Hintsa et al.*, [1999], conclude that most of this signal was actually due to residual water in the liquid nitrogen tank, gas lines, and warm flow tube walls, and estimated that any non-water vapor fluorescence signal in the HWV-E detection axis would likely be at the level of ~0.1 ppmv or less. Furthermore, in flight comparisons of HWV-E with both the NOAA Lyman-α hygrometer and JLH showed that instrumental differences (offsets) did not exceed ~0.2 ppmv.

We repeated the tests of *Hinsta et al.*, [1999], on the 3" detection axis, HWV-E in February, 2007. We drew dry nitrogen gas from a LN_2 Dewar and fed it directly into the calibration flow system. The results of one of these low water vapor calibrations and background tests of the HWV-E axis are shown in Figure 4.3. The titration provides the calibration constants used to derive the measured mixing ratios in the figure. The difference between the measured mixing ratio and the water added through the bubbler is ~0.1 ppmv, providing a constraint on the sum of the contributions from the remaining water sources, i.e., residual water in the nitrogen flow, leaks and outgassing, plus any instrumental artifact. Therefore, the maximum possible offset is ~0.1 ppmv in nitrogen, which is roughly equivalent to ~0.2 ppmv in air, consistent with the results of *Hinsta et al.*, [1999].

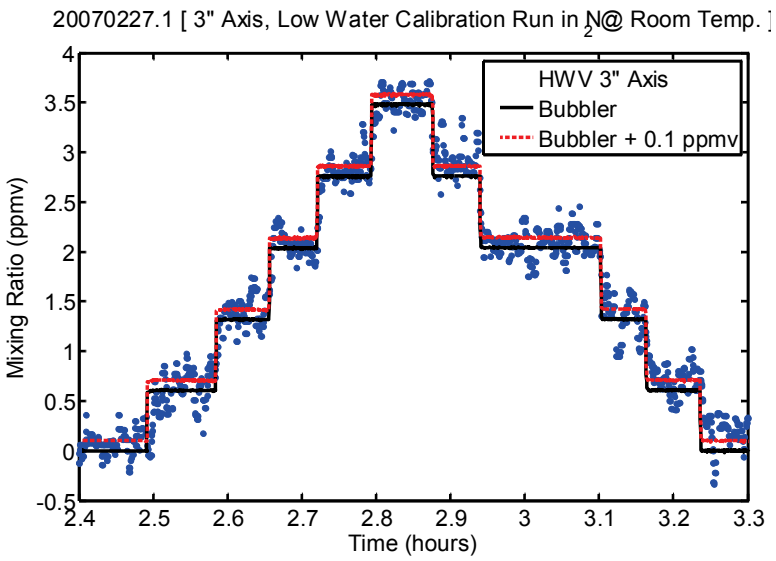


Figure 4.3. Low water vapor calibration and background test of HWV-E. The fluorescence measurements (blue dots), bubbler (black line), i.e., the quantity of water vapor intentionally added to the flow, and bubbler with an additional 0.1 ppmv offset (red line) attributed to the sum of the remaining terms in Equation 4.2, are plotted as a function of time. These data provide a nominal upper bound equivalent to ~<0.2 ppmv in air for any instrumental bias.

In contrast, our initial tests of the 2" detection axis, HWV, flown on NASA WB-57 missions from 2001 to the present, revealed evidence of a measurement offset equivalent to ~1.0 to ~1.5 ppmv H_2O in air. As part of a concerted effort to eliminate sources of residual water vapor, we pared down plumbing upstream of the detection axis, removed flow controllers and valves, measured and minimized the leak rate, and purged the system for up to six hours continuously at high flow rates. Despite every attempt to reduce the impact of residual water vapor sources, we were left with a significant non-zero signal. Thus, in order to provide a bound on the magnitude of an instrumental bias, we were forced to characterize the contributions of each of the remaining terms, i.e., $FLR_{H2Oflowbkgd}$, $FLR_{H2Oleak}$, $FLR_{H2Owalls}$, and FLR_X. Because each term is expected

to behave differently in response to variations in flow conditions, we explored the theoretical mass flow and pressure dependence of each of the terms, and analyzed the behavior of the measured offset.

Background Water

The contribution associated with $FLR_{H2Oflowbkgd}$, for example, is unlikely to respond to changes in either the pressure or the flow rate, since we expect the mixing ratio to be nominally constant in the carrier flow over a single run. In other words, the mixing ratio background of the primary flow is a conserved quantity in our calibration system.

Leak

The contribution of a leak upstream of the detection axis, $FLR_{H2Oleak}$, is expected to depend primarily upon the mass flow in the duct, and to have little or no pressure dependence over the typical range of operating pressures[3]. The flow associated with a constant leak rate, measured in hPa per minute, is given by, Leak Flow [slm] = (Volume of Calibration System [L]) × ($T_{standard}$ / T_{duct}) × (Leak Rate [hPa/min]) / ($P_{standard}$ [hPa]), with $T_{standard}$ = 0 °C and $P_{standard}$ = 1 atm. Thus, the contribution of a leak is determined by the ratio of the leak flow to the main flow scaled by the mixing ratio associated with a leak of room air into the duct, i.e., $FLR_{H2Oleak}$ (ppmv) = ~0.40 × SMR_{liq} × (Leak Flow [slm]) / (MainFlow

[3] The leak rate can be measured as a function of pressure within the duct. Repeated tests have shown that in practice the leak rate is effectively constant over the typical range of operating pressures, ~25 to 400 hPa.

[slm])[4], where SMR_{liq} is the saturation mixing ratio of water vapor over liquid at room temperature and pressure, and the relative humidity of the laboratory is estimated to be ~40%. Given this relationship, the mixing ratio contribution of $FLR_{H2Oleak}$, for a constant leak rate, will vary inversely with mass flow in the duct.

Outgassing

Next we consider the behavior of water vapor outgassing from the walls of the flow plumbing and instrument duct, $FLR_{H2Owalls}$. The literature on water as a contaminant in vacuum systems and/or gas delivery systems for both scientific research and industrial applications is vast. The cause of this contamination, even in the most rigorously designed systems, results from atmospheric exposure. Even brief exposure can lead to monolayer coverage of internal surfaces, with ~10^{15} to 10^{16} water molecules per square centimeter, and typical surface residence times of approximately 30 hours [*Hinkle*, 2004]. Unfortunately, neither our calibration system nor our detection axis was designed to minimize water absorption. In fact, all the internal surfaces of the detection axis are coated with carbon black paint intentionally designed to have high surface area in order to reduce scattered light. Therefore, despite our efforts to regularly dry the system by flowing a dry nitrogen purge under modest vacuum for several hours (~4 to 12 hours) prior to running diagnostic tests, non-zero contributions of $FLR_{H2Owalls}$ must be considered.

[4] A back of the envelope calculation shows that for a leak rate of ~0.5 hPa/min into a volume of 2 liters, the mass flow of the leak is ~1 sccm. If room air is 40% saturated, and the main flow rate is 40 slm, the leak contribution, assuming it is all upstream of the detection axis, is estimated to be ~0.20 ppmv. Leak rates of twice this magnitude into larger volumes can lead to contributions of ~0.80 ppmv.

The contribution due to the desorption of water from the walls of the flow tube and flow plumbing, $FLR_{H2Owalls}$, is expected to be a more complex function of both the pressure and mass flow in the duct, as well as temperature and time[5]. In the simplest case, however, we assume the molar number density of water sourced from the walls, $n_{H2Owalls}$ [#/cc], varies linearly with the diffusion coefficient of water vapor in air, D, and inversely with the volumetric flow in the duct, Q_V, i.e., $n_{H2Owalls} \sim D(T,P)/Q_V$. Because the diffusion coefficient varies nearly linearly with temperature and inversely with pressure, we can express it as $D = C_D \times (T_{duct}/P_{duct})$, with the proportionality constant C_D. Rewriting the equation to determine the mixing ratio contribution yields $FLR_{H2Owalls}$ (ppmv) = $n_{H2Owalls}/M_{duct} \times 10^6 = (R \cdot C_D) \times 10^6 \times (T_{duct}/P_{duct}) \cdot (T_{duct}/P_{duct}) / Q_V$, with $M_{duct} = P_{duct}/(R \cdot T_{duct})$, and the universal gas constant, R. Finally, converting from volumetric flow in the duct to standard mass flow yields $FLR_{H2Owalls} = C \times (T_{duct}/P_{duct})/Q_M$, with $Q_V = Q_M \times (P_{std}/T_{std}) \times (T_{duct}/P_{duct})$, and all the constant terms grouped together in C. Thus, we expect any mixing ratio signal due to $FLR_{H2Owalls}$ to vary inversely with pressure at constant mass flow, and inversely with mass flow at constant pressure. Furthermore, we expect the signal to increase in response to heating the duct, and decrease with time.

Instrumental Artifact

Finally, we examine the dependence of FLR_X on the various flow parameters. The response of FLR_X will depend upon the nature of its source. If it

[5] Note that the "stickiness" of water to the walls is significantly higher than that of water to itself, hence removal of the first several monolayers is expected to be orders of magnitude faster than the removal of the final monolayer.

is due to a constant gaseous component/impurity in the carrier flow it will behave like $FLR_{H2Oflow}$, i.e., it will show no dependence upon pressure or mass flow rate. Likewise, if it is due to a gaseous contaminant from the flow tube walls it will behave similarly to $FLR_{H2Owalls}$. However, if the contribution of FLR_X is due to VUV stimulated emission of ~315 nm light from optics or surfaces in the duct, we expect it to depend primarily upon the magnitude of the VUV flux, and thus, the signal of FLR_X normalized by the VUV flux will be a constant number of counts. In this case it should show no dependence upon the mass flow, however, it will bear a unique pressure dependent signature.

Recall the expression for the normalized OH* fluorescence counts, S_{OH}:

$S_{OH} = [\mathrm{H_2O}]_{Total} \times C; \; with \; C = C_0 \times \frac{1}{1 + q_{air}[\mathrm{M}]}$. Thus, for q_{air}[M] >> 1, i.e., for duct pressures >~100 hPa, S_{OH} is simply proportional to the water vapor mixing ratio in the duct, and is nominally independent of the duct pressure (once the VUV flux has been corrected for absorption due to O_2). However, at low pressures the relationship between S_{OH} and water vapor mixing ratio becomes highly dependent upon pressure. As the pressure goes to zero we expect S_{OH} to go to zero. The calibration accounts for this [M] dependence.

Figure 4.4 shows how the normalized fluorescence signal, S_{OH}, is expected to vary with pressure for a constant 1 ppmv mixing ratio in the duct based on one set of calibration constants. The blue line shows the number of normalized fluorescence counts per ppmv of water vapor as a function of pressure in air, and the red line shows the corresponding number of counts as a function of pressure in nitrogen. The difference, nearly equal to a factor of 2, is due to the fact that

oxygen quenches OH* more effectively. The black line in the figure corresponds to a theoretical constant background signal, S_X, due to spurious fluorescence. For this set of calibration constants the effective offset in ppmv due to a constant normalized signal of ~120 counts is equivalent to ~1 ppmv water vapor in nitrogren and ~2 ppmv water vapor in air at pressures > ~100 hPa. Note also that the equivalent mixing ratio increases dramatically as the pressure decreases below ~50 hPa, reaching ~3 ppmv in air at 15 hPa. As a result, an offset due to FLR_X will have no mass flow dependence at constant pressure, but we expect its effective contribution in ppmv to increase dramatically as the duct pressure is decreased.

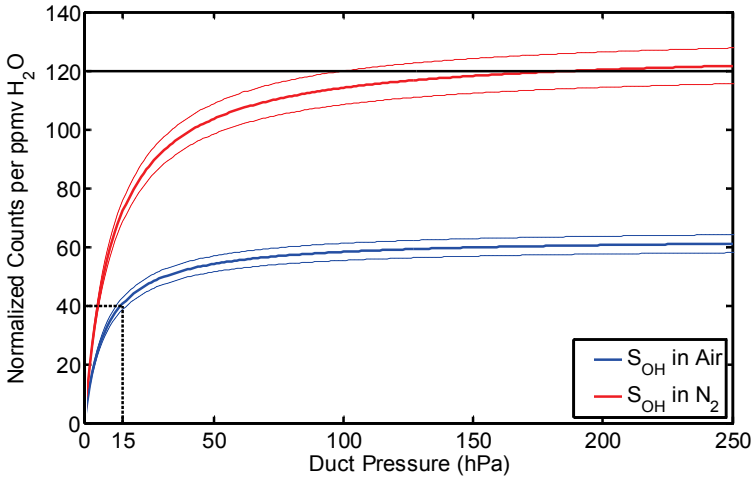


Figure 4.4. Normalized counts per ppmv of water vapor as a function of pressure. At high pressures ~120 normalized counts is roughly equivalent to ~1 ppmv in nitrogen and ~2 ppmv in air. As the pressure decreases below 100 hPa the number of counts per ppmv decreases rapidly, such that for a fixed number of counts the equivalent mixing ratio increases. Note that this relationship is defined by the calibration constants, and will change as the calibration changes.

Table 4.1 summarizes the expected behavior for each of the terms, and provides the basis for interpreting the results of the diagnostic tests we ran during the fall of 2007.

	Pressure Dependence	**Mass Flow Dependence**	**Temperature Dependence**	**Constant**
$FLR_{H2Oflowbkgd}$	No	No	No	Mixing Ratio
$FLR_{H2Oleak}$	(Weak)	Yes Inverse	No	Leak Rate
$FLR_{H2Owalls}$	Yes Inverse	Yes Inverse	Yes	(N/A)
FLR_X	Yes Inverse	No	No(?)	Normalized Counts

Table 4.1. Table summarizing the pressure, flow and temperature dependencies of the four potential offset terms. The last column indicates the conserved, or constant, quantity for each.

We feel it is necessary to reiterate that the primary motivation for characterizing and quantifying each of these terms is to provide a rigorous bound on the magnitude of an instrumental bias, FLR_X. This is the only term that has the potential to lead to errors in flight. Terms due to unintentional water vapor sources do not impact the in situ measurements. The pressure inside the duct in flight is higher than ambient due to the ram fed flow, thus there is no possibility of a leak. Furthermore, due to the high mass flow rates in flight, a factor of ~30 greater than those achieved in the laboratory, the contribution from any residual outgassing is undetectable. This is confirmed by in-flight diagnostic tests, see Chapter 3. Finally, as was discussed in Chapter 3, the presence of a constant (at fixed flow and pressure) non-zero intercept does not impact the accuracy or validity of the calibration procedure and the determination of the calibration constants.

4.3.2.2 Laboratory Quantification

Figure 4.5 shows a photograph of the laboratory setup with three of the Harvard water vapor instruments, ICOS [*Sayres et al.*, 2009], HWV, and HOXotope [*St Clair et al.*, 2008], arranged in parallel and plumbed to a common water vapor addition source. The system allows us to operate all three instruments either simultaneously or independently. The pink insulated box around two HWV detection axes allows us to run calibrations at simulated flight temperatures.

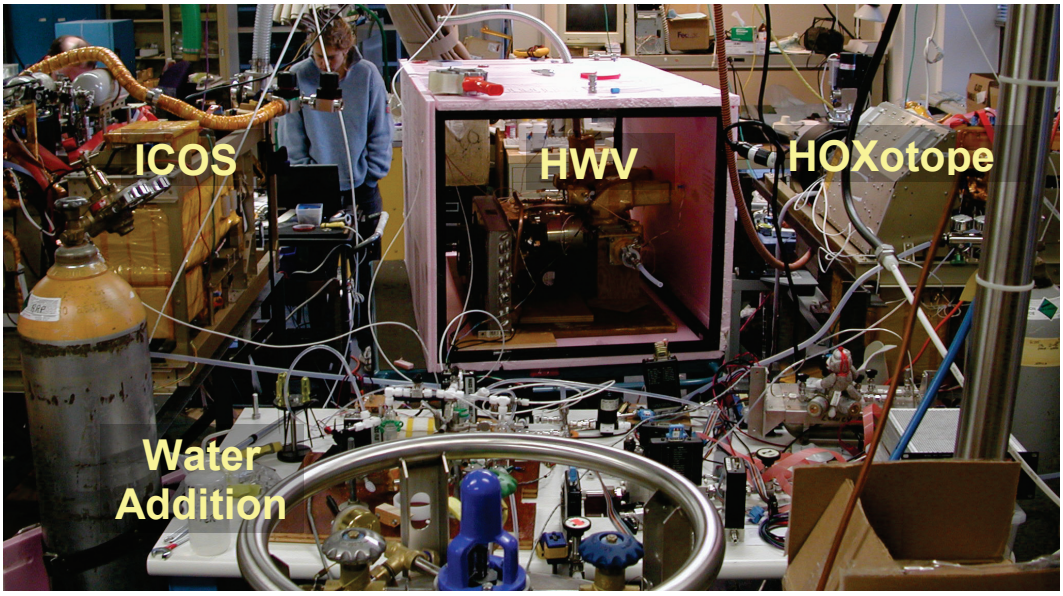


Figure 4.5. Photograph of the laboratory calibration setup. The HWV detection axis (center) is located in the pink insulated box, which allows for runs at temperatures near those encountered in flight. Also pictured are ICOS (left) and HOXotope (right), and the water vapor addition system.

Figure 4.6 shows the results of a low water vapor calibration and zero intercept diagnostic test of HWV in dry nitrogen. The figure shows measurements acquired simultaneously by HWV and HOXotope, where the titration serves as a nominal calibration. Also shown is the calculated water vapor mixing ratio added to the duct via the bubbler apparatus. If the dominant source of the non-zero offset

is to due residual water vapor in the carrier flow and/or addition lines feeding both instruments, i.e., $FLR_{H2Oflowbkgd}$, we expect the measured offsets to be equal in magnitude. The HOXotope instrument measures ~0.5 ± 0.1 ppmv when the bubbler flow is zero, providing an upper limit on the magnitude of the contribution from $FLR_{H2Oflowbkgd}$[6]. The measured offset in HWV is ~1.3 ± 0.1 ppmv, suggesting that there are significant additional biases unique to HWV. The difference between HOXotope and HWV, ~0.8 ± 0.1 ppmv, provides a nominal lower bound to the sum of the three remaining terms in Equation 4.2.

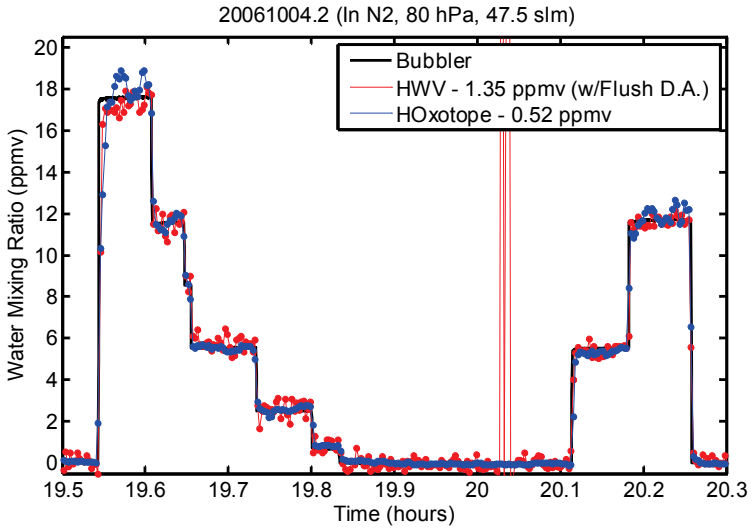


Figure 4.6. Simultaneous calibration and low water vapor intercept run with HOxotope and HWV. The titration serves as the calibration for the HOxotope and HWV measurements. The bubbler data, representing $FLR_{H2Oadded}$, and the HWV and HOxotope measurements are plotted as a function of time. The non-zero offsets measured by each instrument have been subtracted. The net difference between HWV and HOxotope is ~0.8 ± 0.1 ppmv, with HWV measuring systematically higher.

[6] The HOXotope instrument is effectively "wall less" [*St Clair et al.*, 2008]. Once the sample flow has passed through the initial ~1 meter of inlet plumbing it is subjected to ~0.5 meter of VUV illumination from a bank of excimer lamps. The OH and OD fragments from the photolysis region are then detected downstream via laser-induced fluorescence. Water sourced from the walls downstream of the photolysis zone will not be photolyzed and thus will not be detected. Thus it is expected that wall effects in HOXotope in both the laboratory and in flight will be insignificant.

Leak

In order to quantify the contribution of the remaining terms, $FLR_{H2Oleak,}$ $FLR_{H2Owalls,}$ and FLR_X, we ran a series of diagnostic tests in which we systematically varied the flow parameters in Table 4.1. Figure 4.7a shows the results of a diagnostic test during which the duct pressure was held constant (at 60 hPa), and the mass flow was varied by a factor of four. In this run, the measured offset, ~1.1 ± 0.2 ppmv, shows a small dependence upon the mass flow, i.e., bearing the signature of a leak, $FLR_{H2Oleak}$, and/or outgassing from the walls of the duct, $FLR_{H2Owalls}$. We expect that a factor of four decrease in the mass flow will yield a factor of four increase in the sum of the signal from both sources. The measured change in the offset from the highest mass flow (94 slm) to the lowest mass flow (~23 slm) was ~0.4 ppmv. Simple mathematical accounting suggests that the maximum contribution from the sum of the leak and outgassing terms at 94 slm is ~0.1 ppmv, increasing proportionately to ~0.5 ppmv at 23 slm, with an additional constant offset of ~0.7 ± 0.1 ppmv due to other sources. This remaining ~0.7 ppmv is not susceptible to changes in mass flow and thus, according to our analysis, it is likely due to some combination of FLR_X and/or $FLR_{H2Oflowbkgd}$. The fact that HWV previously measured a background ~0.8 ppmv higher than HOXotope suggests that much of this remaining offset may be due to a spurious fluorescence signal, FLR_X, rather than residual water vapor in the carrier flow.

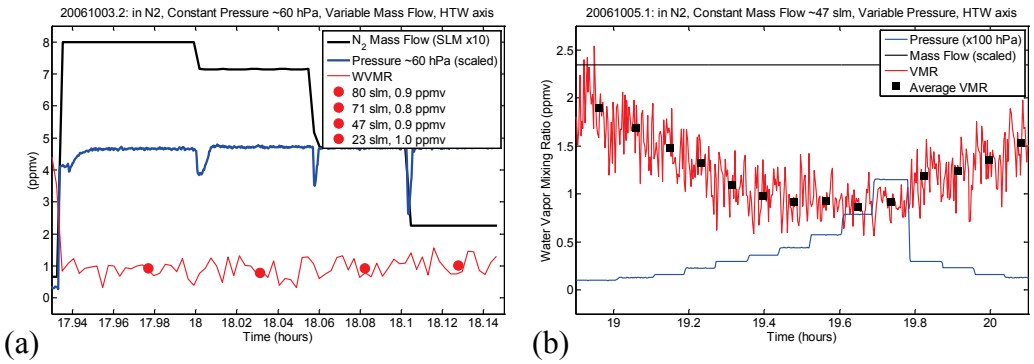


Figure 4.7. Data from two independent diagnostic tests. Panel (a) shows the results of a constant pressure, varying mass flow diagnostic run in nitrogen. Mass flow in the duct (black line), and the duct pressure (blue line) are plotted as a function of time. The corresponding measurements of the residual water vapor mixing ratio are plotted in the red line, and the solid red circles represent the mixing ratio mean at each of the four mass flow settings. Panel (b) shows the results of constant mass flow, varying pressure diagnostic run in nitrogen. Again, mass flow (black line), duct pressure (blue line), and the measured mixing ratio (red line, solid black squares) are plotted as a function of time.

Walls

Figure 4.7b shows the results of a comparable diagnostic test in which the mass flow was held constant (at 47 slm) and the duct pressure was varied. These data show a dramatic response with respect to variations in pressure, with the measured mixing ratio decreasing by approximately a factor of two as pressure increases in the duct. A plot of the average mixing ratio as a function of the average pressure at each step, as shown in Figure 4.8a, exhibits this dependence more plainly. The measurements are represented by the black squares.

Figure 4.8a also shows the results of two simple models that attribute the behavior 1) solely to pressure dependent contributions from outgassing, $FLR_{H2Owalls}$ (open blue circles), and 2) solely to a constant fluorescence signal, FLR_X equal to 110 counts (filled red circles). The general behavior of the measurements is well represented by both models. However, the pressure

dependent outgassing model, cannot explain the residual underlying offset measuring ~0.6 ppmv. The outgassing model with an additional constant 0.6 ppmv is represented by the filled blue circles. Because the outgassing signal asymptotes to near zero with increasing pressure, and the contributions of both the outgassing and leak terms asymptote to zero with increasing mass flow, there is no way to explain the residual offset without invoking either a non-zero mixing ratio background in the carrier flow or an instrumental bias.

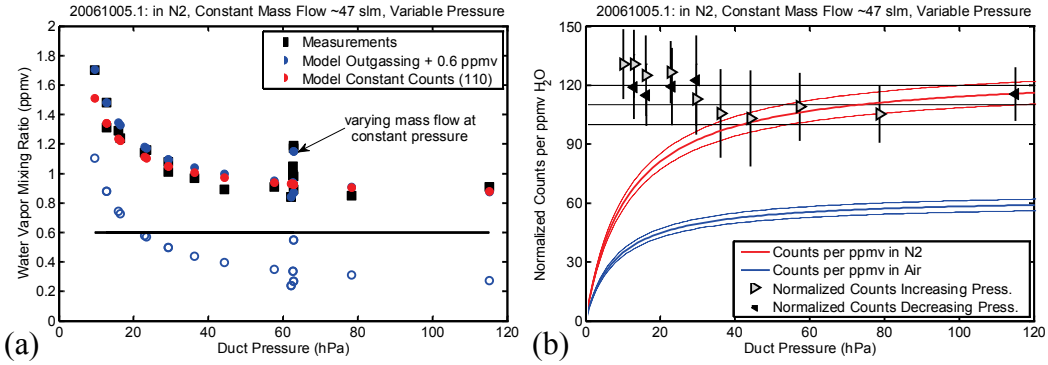


Figure 4.8. Diagnostic test results in the context of a simple model. Panel (a) shows the mean mixing ratio measurements (solid black squares) from the diagnostic tests of Figure 4.7a and b plotted as a function of duct pressure (hPa). The results of a model attributing all of the residual water vapor to outgassing are plotted in the open blue circles. The solid blue circles represent the outgassing model results plus an arbitrary constant 0.6 ppmv (black line). The solid red circles correspond to the results of a model attributing all of the residual water vapor to a spurious, i.e., non-water vapor fluorescence signal equivalent to a constant 110 normalized counts. Panel (b) provides the context for the constant counts model. The solid red and blue lines show the sensitivity of the detection axis, in normalized counts per ppmv of water vapor, as a function of pressure for nitrogen and air respsectively. The triangles show the normalized counts at each pressure level of the varying pressure diagnostic run, with the symbols representing the mean and the vertical line capturing the standard deviation of the measurements. The solid black line corresponds to a constant 110 counts, a signal which is equivalent to ~1.0 ppmv at ~70 hPa, and ~1.8 ppmv at ~10 hPa in nitrogen.

Spurious Fluorescence

Invoking a constant fluorescence signal due to non-water vapor sources, equal to ~110 normalized counts, also approximates the behavior of the measured signal. Furthermore, a constant signal of ~110 counts naturally asymptotes to ~0.8 ppmv as the pressure increases. This source simultaneously accounts for the constant offset measured in the mass flow diagnostic test, as well as much of the residual difference between the measurements of HWV and HOXotope.

Figure 4.8b shows a plot of the measured normalized fluorescence counts (black open and solid triangles for increasing and decreasing pressure) as a function of pressure within the framework of Figure 4.4. This figure makes explicit the pressure dependent contribution of a constant, non-water vapor fluorescence signal, FLR_X, of ~110 ± 10 counts (black line), and shows that the measured data are broadly consistent with a model assuming constant counts. The discrepancy between the measured mixing ratio and the model, evident at the lowest pressures in both Figures 4.8a and 4.8b, can be reconciled with a small outgassing contribution.

Model

In reality, it is exceedingly difficult to accurately quantify the contribution of each component. During most of the runs the signal to noise ratio of the detection axis was low, severely limiting our ability to distinguish the different terms, the magnitude of each being on the order of tenths of ppmv. Furthermore, the low signal to noise ratio exacerbated errors due to the chopper reflectivity

correction. In addition, the calibration of the axis under investigation was not determined with the usual level of rigor and uncertainties in the calibration from one run to the next, particularly at low pressures, partially compromised the validity of the tests. Finally, leak rates, outgassing rates, and even the water vapor in the carrier flow were not necessarily constant between runs, further confounding the interpretation of the results.

Despite these limitations, we attempted to synthesize the results of the various tests with a simple model. Figure 4.9 shows the results of this effort. The measured mixing ratio is again plotted in the solid black squares. The total modeled offset (green circles) is a sum of four terms: $FLR_{H2Obackground}$ = 0.1 ppmv (black line); $FLR_{H2Owalls}$, which ranges from ~0.1 at the highest pressures to 0.5 ppmv at the lowest pressures (dark blue dots); $FLR_{H2Oleak}$, which is effectively a constant ~0.1 ppmv with respect the pressure, but varies in response to changing mass flow (cyan circles); and FLR_X, due to a constant 70 counts, which is ~1.0 ppmv at the lowest pressures and asymptotes to ~0.6 ppmv as the pressure increases. The results of the constant pressure, varying mass flow run are evident as the vertical distribution of points at ~60 hPa in the figure, and these data serve to bound the total contribution of $FLR_{H2Oleak} + FLR_{H2Owalls}$. The results of this model are by no means conclusive, however, they provide reasonable estimates for the contributions from each of the terms, and provide a nominal "worst case" for the contribution do to an instrumental offset, equivalent to ~0.6 ppmv in nitrogen and ~1.2 ppmv in air.

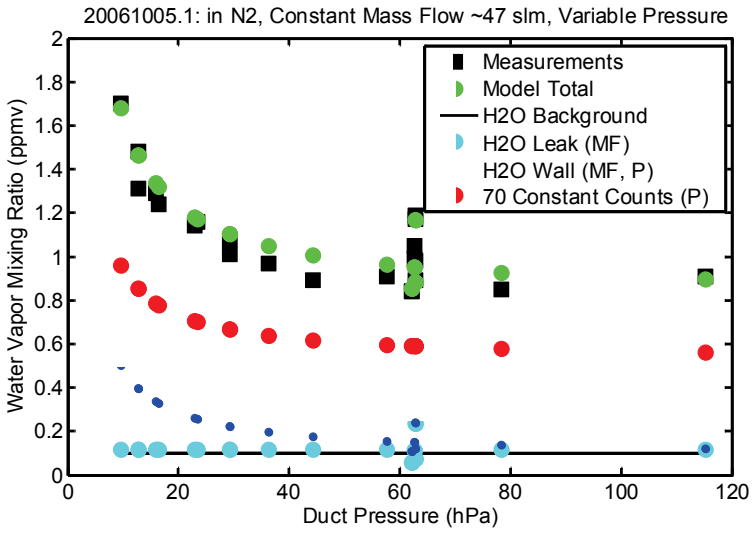


Figure 4.9. Results of a model capturing the contribution of each of the four potential offset terms. This figure shows the mean mixing ratio measurements (solid black squares) from the diagnostic tests of Figure 4.7a and b plotted as a function of duct pressure (hPa). The model results show the contribution of each of the four potential offset terms in Equation 4.2 as a function of pressure. The absolute values for each term are constrained by the results of the diagnostic tests described in the text. The constant contribution from residual water vapor in the carrier flow is plotted in the solid black line. The contributions due to 1) outgassing from the walls (dark blue dots), 2) a leak into the duct (solid light blue circles), and 3) a constant spurious signal equivalent to 70 normalized fluorescence counts (solid red dots) each demonstrate their characteristic mass flow and/or pressure dependencies. The total offset is plotted in the solid green circles.

Additional Tests

Subsequent diagnostic tests examined the possibility of fluorescence by a gaseous component in the O_2 cell. The residual signal showed no dependence upon the O_2 cell pressure. We also explored the effect of changing the chemical composition of the carrier flow by adding ozone and CO_2 to better simulate the composition of the UT/LS. None of these tests substantially altered the results.

Optical Configuration

The most dramatic and potentially illuminating response of the diagnostic runs was achieved when we replaced a flush diode/mirror assembly, which holds both the VUV diode and mirror directly across the duct from the lamp, with a recessed piece. The recessed piece was designed to position the diode ~4 cm further from the lamp in order to reduce the contribution of scattered light from the optics and the wall, and to simultaneously increase the path length for radial absorption measurements. It was used during the Clouds and Water Vapor in the Climate System (CWVCS) mission in the HTW instrument, but was replaced with the standard flush mount for all subsequent missions. With the recessed diode mount in place the offset plaguing the HWV measurements in the laboratory was dramatically reduced. A series of back-to-back runs on October 16, 2006, and subsequent days showed that replacing the flush diode assembly with the recessed diode mount repeatedly and reliably eliminated the apparent measurement bias.

Data acquired during a diagnostic run on January 25, 2007, and plotted in Figure 4.10, neatly summarize these results. During this run two Lyman-α detection axes were situated in series. The first of the two Lyman-α detection axes was configured with the recessed diode mount (Ax1), and the second axis held the standard flush diode assembly (Ax2). In addition, in Ax2 the MgF_2 window forming the seal between the O_2 cell and the duct was replaced with a 121.6 nm filter. This optic has a transmission peak at 121.6 nm with a full-width half-maximum of ~15 nm. Because the filter effectively eliminates all wavelengths emitted by the lamp except for 121.6 nm, it eliminates the need for the chopper,

and thus any uncertainties associated with the chopper reflectivity correction. The ICOS instrument was operated in parallel in order to provide a nominal upper bound to the residual water vapor in the carrier flow, and titrations were executed in both dry N_2 and in air in order to ascertain the magnitude of the residual offset in both. Figure 4.10 shows the measurements as a function of time with the legend providing the key to the Ax1 (recess), Ax2 (flush + filter), and the ICOS high sensitivity absorption line data, as well as water added via the bubbler.

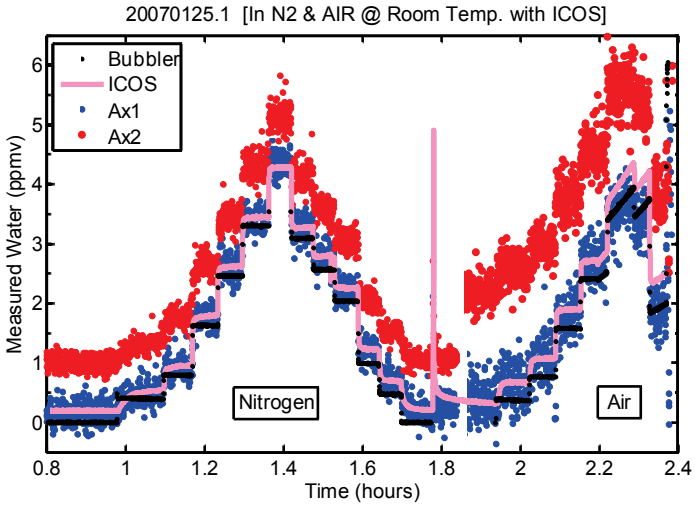


Figure 4.10. Offset comparison of two HWV axes with different optical configurations – measurements. Water vapor measured by HWV Ax1 (blue dots), with the recessed diode assembly, HWV Ax2 (red dots), with the flush diode assembly, and ICOS (pink line) are plotted as a function of time. The water vapor added by the bubbler (black line) is shown for comparison. The colored dots correspond to the fluorescence measurements, where the data have been analyzed using the results of the titrations in nitrogen and air, respectively. The excellent agreement between HWV Ax1 and ICOS is evident. The difference between these measurements and the bubbler, ~0.2 ppmv in nitrogen, and ~0.4 ppmv in air, is attributed to residual water vapor in the carrier flows. HWV Ax2 shows a substantial offset relative to HWV Ax1 and ICOS.

Figure 4.11 shows the corresponding differences between the Ax1 and Ax2 measurements and those of ICOS. This figure shows that Ax1 yields zero-

intercept measurements in both N_2 and air that are in remarkable agreement with ICOS. The agreement between these two independent instruments suggests that the non-zero offset due to residual water vapor in the carrier flow, i.e., the contribution of $FLR_{H2Obackground}$ is at most ~0.2 ppmv in N_2, and ~0.4 ppmv in air. Furthermore, these data indicate that additional contributions associated with leaks, outgassing or spurious fluorescence sources in Ax1 are likely negligible. In contrast, the offset in Ax2 measures ~0.8 ppmv higher in N_2 and ~1.8 ppmv higher in air than ICOS, suggesting a bias in the Lyman-α system uniquely associated with the presence of the flush diode assembly and most likely due to a spurious fluorescence source.

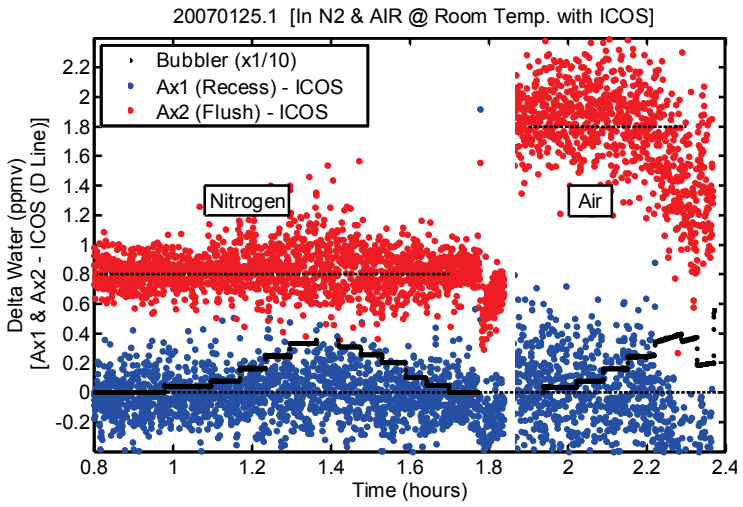


Figure 4.11. Offset comparison of two HWV axes with different optical configurations – absolute differences. The absolute differences between water vapor measured by HWV Ax1 and ICOS (blue dots), and between HWV Ax2 and ICOS (red dots) are plotted as a function of time. The water vapor added by the bubbler, scaled by a factor of 10 (solid black line), is shown for reference. The difference between HWV Ax1 and ICOS is centered at ~0.0 ± 0.2 ppmv in both nitrogen and air, and is constant over the full run. The difference between HWV Ax2 and ICOS is centered at ~0.8 ±0.2 ppmv in nitrogen, and ~1.8 ±0.3 ppmv in air. The change in the magnitude of the offset between nitrogen and air is consistent with the presence of spurious fluorescence counts in HWV Ax2, which has the flush diode assembly.

Though the mechanism is still not fully understood, subsequent laboratory tests suggested that the painted surface of the flush diode assembly was a potential source of VUV stimulated fluorescence. The mechanical design of the flush diode assembly (see Figure 2.1) positions both the mirror and the diode ~0.5 cm from the centerline of the lamp, leaving a ~1.0 cm^2 area between the optics that receives the majority of the collimated VUV photon flux. This portion of the flush aluminum mount is coated with the same non-reflective carbon black paint that covers all the interior surfaces of the detection axis. Furthermore, crude but adequate tests with a penlight mounted on a translating stage showed that this region of the wall and mount was visible to the PMT. Thus, the leading hypothesis is that VUV stimulated fluorescence off of the painted surface of the flush diode/mirror mount assembly contributes to a systematic measurement bias in HWV, and that this bias is eliminated by replacing the flush diode assembly with a recessed mount[7].

4.3.2.3 Summary

In summary, a long series of diagnostic tests executed after the CRAVE mission designed to characterize the behavior of the 2" HWV detection axis at low water mixing ratios, ~<5.0 ppmv, revealed a persistent instrumental measurement bias largely consistent with a constant number of normalized VUV

[7] The optical design of the 3" axis (HWV-E) is significantly different than that of the 2" axis (HWV). In the 3" axis the diode is positioned directly across from the lamp, in a hole cut from the center of the VUV mirror. Therefore, the majority of the collimated light from the lamp is directly incident on the diode. Furthermore, because of the larger dimension of the 3" axis the walls of the detection axis are completely shielded from the field of view of the PMT with modest baffling. Not only does this reduce the potential of a spurious fluorescence signal, but it also significantly reduces the scattering component.

stimulated fluorescence counts. The residual normalized fluorescence signal, FLR_X, was found to be insensitive to duct flow, pressure and temperature, and carrier flow composition, and the magnitude of the resulting offset was roughly equivalent to ~1.0 ± 0.5 ppmv H_2O in air at 100 hPa. Every laboratory examination of the 2" HWV axis showed that the offset was effectively eliminated when the flush diode mount was replaced with a recessed assembly. In general, this was true regardless of which of the two individual detection axes (HWV or HTW) was used. However, the exact magnitude of the offset may be subtly axis dependent. Furthermore, subsequent tests performed by Maryann Sargent during the spring of 2008 showed that the apparent measurement bias associated with the flush diode assembly was also affected by changing the bandwidth of filters in the PMT collection optics [*Maryann Sargent, personal communication*].

4.4 Offset-free Axis

In response to these laboratory results, a nominally "offset-free" axis was assembled. A new recessed mount for the mirror and the diode was fabricated, with the positioning of both optics carefully determined to be out of the field of view of the PMT. Subsequent diagnostic runs and calibrations showed that it performed reliably, with a total offset including contributions from every source in Table 4.1, measuring <0.50 ppmv in both air and nitrogen flows. Figure 4.10 provides a representative example of a low water vapor calibration and the performance of the offset-free HWV axis. These data show the total offset of

HWV to be ~<0.4 ppmv in a dry air flow, with the residual difference between ICOS and HWV measuring ~0 ±0.1 ppmv.

This offset-free design of the 2" HWV detection axis is now the standard, and it was first deployed during the TC4 mission in the summer of 2008, and again in the AIDA comparison campaign in the fall of 2008. Results from the AIDA chamber laboratory comparison, discussed in detail in Chapter 5, offered additional confirmation that residual biases in the 2" HWV axis with the recess diode assembly were no greater than ~0.45 ppmv equivalent in air with respect to a select set of core instruments.

4.5 Evidence For/Against an Offset in the Flight Data

The results from the low water and zero-intercept laboratory diagnostic tests of 2006 – 2007 showed that the 2" detection axis with the flush diode assembly exhibited a systematic measurement bias of ~1.0 ± 0.5 ppmv equivalent in air, and that the behavior of this offset was consistent with a nominally constant source of fluorescence counts from a non-water vapor source. The robustness of this result motivated a redesign of the 2" detection axis. Thus, the fundamental question becomes: What evidence exists for a measurement offset in the flight data? Furthermore, if the flight data show evidence of a systematic bias, what is its approximate magnitude? and when did it arise? In this section we review the flight data in an attempt to answer these questions.

Identifying and quantifying an offset in the laboratory is laborious and relies on both independent reference measurements and a series of rigorously

designed diagnostic tests. In contrast, the process of identifying an offset in flight relies almost entirely upon comparisons with other instruments[8]. Examining measurement differences between instrument pairs is both tedious and inexact, because differences may be due to calibration, i.e., scale factor, errors, as well as to absolute instrumental offsets. Additionally, the measurements of one or both instruments may be influenced by ambient pressure, temperature, flow, airmass composition and time (either during a single flight, over the course of a mission or from one mission to the next), further complicating the quantification and attribution of systematic errors[9]. Finally, no single instrument is currently capable of providing an accepted in situ reference measurement for comparison. With these caveats in mind, and noting that the interpretations of the results are often contradictory, we present a series of instrument comparisons.

For reference, Table 4.2 provides an overview of the versions of HWV by mission, date, aircraft and location. Other instruments that flew on the same platform and the contemporaneous satellite or balloon-borne instruments are also listed. The shading indicates the transition from HWV 3" to the HWV 2" version with the flush diode assembly after 2000, and the transition from the HWV 2" version with the flush diode assembly to the HWV 2" version with the recessed assembly after 2006.

[8] Recall that the HVW in-flight radial absorption measurements, similar to the axial absorption measurements in the laboratory, can only measure changes in the mixing ratio relative to a baseline. Thus, while they provide an excellent validation of the calibration, i.e., the slope, under flight conditions they offer no information about the magnitude of the intercept.

[9] We argue (see Chapter 3) that in the case of HWV we have carefully examined the instrument's sensitivity to each of these parameters in the laboratory, and demonstrated that the detection axis calibration is insensitive to the transition from the laboratory to the in situ environment.

Mission	Dates	Aircraft	Location	HWV Version	Aircraft Comparison Instruments	Other Comparison Instruments
-	1993 - 2000	ER-2	-	HWV 3”	JLH	FPH, HALOE
CWVCS	Aug. 2001	WB-57	San Jose, Costa Rica	HWV 2” *Flush*	HTW (*Recess*)	-
CRYSTAL FACE	Jul. 2002	WB-57	Key West, FL	HWV 2” *Flush*	HTW (*Flush*), JLH	
Pre-AVE	Jan. 2004	WB-57	San Jose, Costa Rica	HWV 2” *Flush*	HTW (*Flush*), JLH	CFH, HALOE
AVE-WIIF	Jul. 2005	WB-57	Houston, TX	HWV 2” *Flush*	HTW (*Flush*), ICOS, HOx	MLS
CRAVE	Feb. 2006	WB-57	San Jose, Costa Rica	HWV 2” *Flush*	ICOS, JLH	CFH, MLS
TC4	Aug. 2007	WB-57	San Jose, Costa Rica	HWV 2” *Recess*	ICOS, HOx, JLH	CFH, MLS
AquaVIT	Oct. 2007	Lab	AIDA Chamber Karlsruhe	HWV 2” *Recess*	JLH, FISH	CFH

Table 4.2. Version of HWV by mission, date, aircraft and location.

4.5.1 Evidence Supporting an Instrumental Offset in Flight

The most convincing evidence for a measurement bias in flight relies on measurement results acquired during the Tropical Composition, Cloud and Climate Coupling (TC4) flight mission. During this mission we configured the HWV instrument to fly with two detection axes in series, the first axis flew with the old style flush diode assembly and the second axis flew with the newly designed recessed mount. Though both axes were plagued by chopper mechanism

failures, and their performance and reliability compromised by large calibration changes mid-mission, on average the measurements showed differences on the order of ~1.5 ppmv, with the flush diode axis measuring systematically higher. These data suggest that the effect observed in the laboratory is reproducible in flight. The question remains, however, whether this issue, which was discovered in the laboratory after the CRAVE mission and demonstrated to exist in flight during TC4, was present during CRAVE and earlier missions.

Comparisons with the NOAA CMDL frost-point hygrometer, CFH, have consistently shown differences with 2" HWV. Moreover, analyses of data taken during the CRAVE mission specifically show that differences between HWV and CFH are generally consistent with a measurement offset of the order of ~1.5 ppmv [*Vomel et al.*, 2007a; *Weinstock et al.*, 2009]. However, because of 1) the limited number of successful comparisons during the CRAVE mission, one for HWV/CFH and five for ICOS/CFH, 2) the variability in the performance of all of the instruments from flight to flight, which was at or near the level of their absolute differences, and 3) real atmospheric variability, a full characterization of the discrepancy, and its possible dependence upon pressure, temperature, and mixing ratio was impossible.

4.5.2 Evidence Opposing an Instrumental Offset in Flight

Compelling evidence against a measurement bias in the HWV 2" axis is provided by measurement comparisons between HWV and HTW acquired during

the CWVCS mission in 2001, and by measurement comparisons between HWV, HTW, HOXotope and ICOS during the AVE-WIIF mission in 2005.

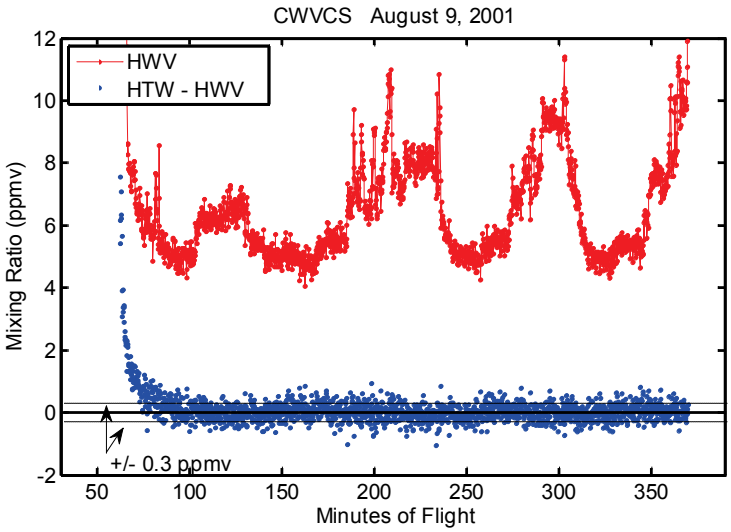


Figure 4.12. Ambient mixing ratio and absolute difference between HWV and HTW in clear air. This figure shows in situ measurements made by HWV on the August 9, 2001, flight during the CWVCS mission as a function of time (red dots). The absolute difference in clear air between HWV, flown with the flush diode assembly, and HTW, flown with the recessed assembly, are plotted in the blue dots. After the short "dry out" period at the beginning of the flight, where HTW > HWV, the differences are centered at 0.0 ±0.3 ppmv. This agreement is well within the 7% combined calibration uncertainty of the two instruments, and there is no indication of a measurement bias in HWV.

During the CWVCS mission, the HWV instrument flew with the flush diode assembly and the HTW instrument flew with the recessed assembly. Clear air comparisons at low water vapor mixing ratios in the tropical UT/LS show excellent agreement between the two instruments during select flights, with typical differences well within ~±0.3 ppmv. Figure 4.12 shows a comparison from the flight of August 9, 2001. Water vapor mixing ratios measured by both instruments are plotted as a function of time. The absolute difference between the two measurements is also plotted. HTW measures higher than HWV during the

first ~1 hour because the walls of the inlet require a finite "dry out" period[10]. After this period the measurements converge to better than ±0.3 ppmv, and show remarkable consistency over a range of ambient mixing ratios, temperatures and pressures.

A linear fit to the data (HTW vs. HWV), excluding the first two hours, shows agreement between the two measurements to within ~1% (slope), with an intercept of ~−0.1 ppmv. These data show the level of agreement that is possible between two calibrated and independently operated instruments in flight, and convincingly demonstrate that the use of the flush diode mount did not lead to a measurement offset in the HWV instrument during the CWVCS mission. Furthermore, HWV and HTW were both configured with identical flush diode/mirror mounts during the CRYSTAL-FACE, PreAVE and AVE-WIIF missions, and no change was observed in their level of agreement.

Comparisons from the AVE-WIIF campaign staged out of Houston during the summer of 2005 provide further evidence against the existence of a measurement bias in HWV. AVE-WIFF was designed as a comparison mission for a suite of instruments newly developed for the in situ measurement of water vapor and its isotopologues. During the mission, data were acquired and archived by each instrument team using a single "blind" approach, i.e., no instrument team had access to data from the other participant instruments. HOxotope and ICOS were tested along with the ALIAS instrument from the Jet Propulsion Laboratory. HWV, HTW and JLH were chosen to provide nominal reference measurements.

[10] The comparatively slow flow in HTW (~10 slm) and long inlet duct exacerbate the effects of hysteresis, however, proper pre-flight conditioning can significantly reduce the hysteresis lifetime [*Weinstock et al*., 2006a; *Weinstock et al.*, 2006b].

Figure 4.13a shows mixing ratio data from the suite of Harvard hygrometers as a function of time during the flight of July 7, 2005. Percentage differences and absolute differences between HWV and each of the other instruments during this flight are shown in Figures 4.13 b & c, respectively. Note that the percentage differences mostly represent calibration (scale) errors, and that the absolute differences capture systematic offsets (biases). The "dry out" period required for the HTW, ICOS and HOxotope instruments is evident during the first ~2 hours of flight. For data after this period, the agreement between HWV and the other Harvard instruments is typically better than ±5%, with absolute differences <±0.3 ppmv over nearly two orders of magnitude in ambient mixing ratio.

All four Harvard instruments are calibrated in the laboratory with respect to two independent references standards, and HTW, ICOS, and HOXotope operate under virtually identical temperature, pressure and flow conditions in the laboratory and in flight [*Sayres et al.*, 2009; *St Clair et al.*, 2008; *Weinstock et al.*, 2009]. Furthermore, the four Harvard instruments utilize three different detection techniques and four different in situ sampling strategies. As a result, we expect that each instrument will be subject to largely independent sources of systematic error in both the laboratory and flight environments, and further, that agreement among them provides a stringent test that these errors have been minimized. The agreement demonstrated in Figure 4.13, well within the combined uncertainty of the suite of instruments, serves as a validation of the Harvard calibration and validation methodology, and provides the best empirical proof of the absence of

any systematic offset in the HWV detection axis with the flush diode assembly during AVE-WIIF.

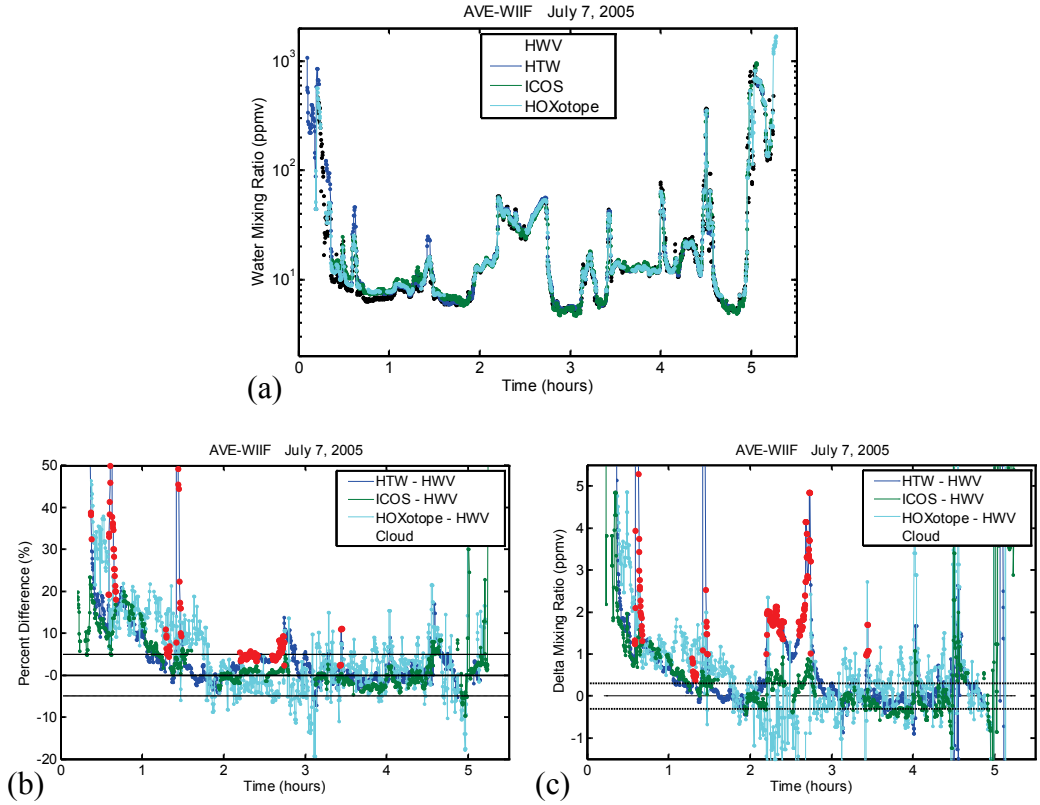


Figure 4.13. In situ comparison of the suite of Harvard hygrometers. Panel (a) shows in situ water vapor measurements as a function of time acquired on the July 7, 2005, flight during the AVE-WIIF mission by the full suite of Harvard instruments described in the text. Panel (b) shows the percent difference between each of the instruments and HWV. The "dry out" period at the beginning of the flight is evident. After the first 2 hours of flight, all the instruments converge to measure within ~±5%, i.e., well within their combined uncertainties. This figure demonstrates the robustness of the calibration in the in situ environment, and the excellent agreement among this suite of instruments over two orders of magnitude in ambient water vapor. Panel (c) shows the absolute difference between each of the instruments relative to HWV. Again, after the "dry out" period, the data fall within ~±0.3 ppmv (the dashed black lines). HWV, flown with the flush diode assembly, shows no evidence of a measurement bias in this dataset. During this flight segment most of the data were acquired within clear air. The in-cloud data are highlight by the red circles. Note the departure of HTW during these intervals.

Finally, the flight comparisons among Aura MLS and CFH and the 2" HWV axis with the *flush* diode assembly during CRAVE, discussed above, show results that are virtually identical to those between MLS and CFH and the 2" axis with the *recessed* assembly during TC4. Both sets of comparisons show HWV measuring consistently higher by ~1.5 ppmv. Assuming that the performance of the MLS and CFH instruments was stable during the period between CRAVE and TC4, these data indicate that the HWV axis with the flush diode assembly during CRAVE performed in a manner identical to the HWV axis with the recess during TC4, i.e., that there was no additional offset or bias present in CRAVE. Similarly, comparisons of data from HOXotope and the HWV axis with the recess diode assembly acquired during TC4 exhibit agreement comparable to that shown in Figure 4.13.

References

Hinkle, L. D. (2004), Effect of purge pressure on desorbing water removal rate, *J Vac Sci Technol A*, *22*(4), 1799-1803.

Hintsa, E. J., E. M. Weinstock, J. G. Anderson, R. D. May, and D. F. Hurst (1999), On the accuracy of in situ water vapor measurements in the troposphere and lower stratosphere with the Harvard Lyman-alpha hygrometer, *J Geophys Res-Atmos*, *104*(D7), 8183-8189.

Read, W. G., et al. (2007), Aura Microwave Limb Sounder upper tropospheric and lower stratospheric H2O and relative humidity with respect to ice validation, *J Geophys Res-Atmos*, *112*(D24).

Sayres, D. S., et al. (2009), A new cavity based absorption instrument for detection of water isotopologues in the upper troposphere and lower stratosphere, *Rev Sci Instrum*, *80*(4).

St Clair, J. M., et al. (2008), A new photolysis laser-induced fluorescence instrument for the detection of H2O and HDO in the lower stratosphere, *Rev Sci Instrum*, *79*(6).

Vomel, H., D. E. David, and K. Smith (2007a), Accuracy of tropospheric and stratospheric water vapor measurements by the cryogenic frost point hygrometer: Instrumental details and observations, *J Geophys Res-Atmos*, *112*(D8).

Vomel, H., et al. (2007b), Validation of Aura Microwave Limb Sounder water vapor by balloon-borne Cryogenic Frost point Hygrometer measurements, *J Geophys Res-Atmos*, *112*(D24).

Weinstock, E. M., J. B. Smith, D. Sayres, J. V. Pittman, N. Allen, and J. G. Anderson (2006), Measurements of the total water content of cirrus clouds. Part II: Instrument performance and validation, *J Atmos Ocean Tech*, *23*(11), 1410-1421.

Weinstock, E. M., et al. (2009), Validation of the Harvard Lyman-alpha in situ water vapor instrument: Implications for the mechanisms that control stratospheric water vapor, *J Geophys Res-Atmos*, *114*.

Wennberg, P. O., R. C. Cohen, N. L. Hazen, L. B. Lapson, N. T. Allen, T. F. Hanisco, J. F. Oliver, N. W. Lanham, J. N. Demusz, and J. G. Anderson (1994), Aircraft-Borne, Laser-Induced Fluorescence Instrument for the in-Situ Detection of Hydroxyl and Hydroperoxyl Radicals, *Rev Sci Instrum*, *65*(6), 1858-1876.

Chapter 5

Aqua Validation and Instrument Tests (AquaVIT)

5.1 Introduction

In June 2007, the international UT/LS water vapor community organized a meeting, the International Workshop on Upper Tropospheric Humidity, focused upon measurements and modeling of upper tropospheric humidity. The core question of the meeting was: Is the current understanding of ice microphysics complete, or do the recent humidity measurements (showing substantial supersaturation) require new physical mechanisms to explain ice nucleation and cirrus cloud formation in the upper troposphere? It was concluded that differences in humidity measurements across instruments and platforms were, and continue to be, too great to answer this fundamental question, and it was suggested that a collaborative international laboratory effort be initiated to address the outstanding measurement discrepancies. The Aqua Validation and Instrument Tests (AquaVIT) campaign, planned on the final day of the workshop, and hosted by the Aerosol Interaction and Dynamics in the Atmosphere (AIDA) facility at the Forschungzentrum Karlsruhe in October 2007, provided the first international

opportunity for the simultaneous comparison of a suite of water vapor measurement techniques in a well controlled laboratory setting.

In the following chapter we summarize the structure of the AquaVIT comparison campaign, evaluate the performance of the AIDA chamber as a common calibration source, assess the performance of HWV during the campaign, identify some of the limitations of the AquaVIT, and discuss strategies for resolving the in situ humidity measurement discrepancies in future campaigns.

5.2 AquaVIT Campaign

5.2.1 AquaVIT Structure

Figure 5.1 shows a schematic of the AIDA chamber with the participating instrument sampling positions indicated. The AIDA facility instruments, which were expected to provide reference measurements during the campaign, included a long path tunable diode laser (TDL) mounted inside the chamber and a commercial MBW 373LX frost point hygrometer (FP), which sampled externally through a heated tube and was advertised for use down to low pressures and frost point temperatures of –95˚C. A subset of instruments that have an extensive history of making atmospheric measurements in the UT/LS were termed “core instruments” and they were the focus of the comparison during the first week of the campaign. The core instruments consisted of the HWV Lyman-α hygrometer, CFH, the Fast In situ Stratospheric Lyman-α Hygrometer (Geophysica) (FISH2) and the Fluorescent Advanced Stratospheric Lyman-α Hygrometer for Balloon (FLASH-B2), which all shared an external heated sampling line exiting the

chamber near the base on Level 1. The core instruments also included the FISH1 (LearJet) Lyman-α hygrometer sampling off an outlet port located near the top of the chamber on Level 3, and JLH and FLASH-B1 Lyman-α hygrometers, which were mounted inside the AIDA chamber. HWV participated only during the first week of the campaign where static water vapor, temperature, and pressure chamber conditions were tested. The primary goal was to simultaneously provide the same water mixing ratio to the suite of participant instruments and compare their performance over a range of pressures and mixing ratios ranging from ~200 ppmv to <0.5 ppmv.

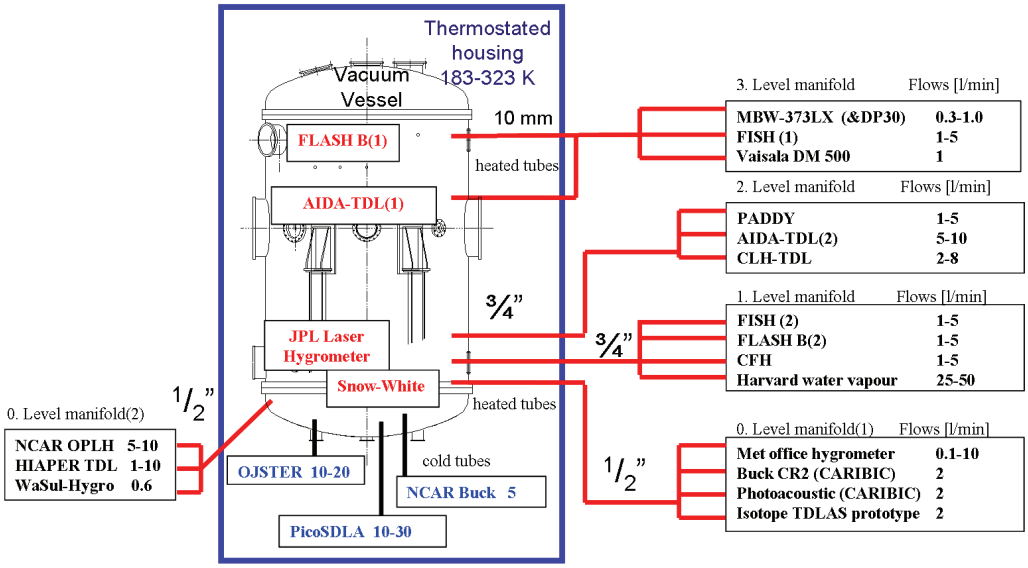


Figure 5.1. Schematic representation of the AIDA chamber and the instrument configuration used during the AquaVIT campaign.

The duration of a typical daily run was roughly eight hours. The chamber was primed, i.e., fully evacuated and dosed with a known quantity of water, and then back filled to the first stable pressure level with dry air. Formal sampling began once the starting pressure was stabilized. Over the course of each static run

the pressure was increased in a stepwise fashion with the addition of dry air, and decreased in a stepwise fashion by pumping. During the fill stages the mixing ratio decreased as the pressure increased. Figure 5.2 shows the structure of the run of October 15, 2007. The chamber pressure (blue line), temperature (green line), and mixing ratio (red line) as measured by the AIDA TDL are shown as a function of time. During the static intervals the chamber pressure and temperature are nominally constant.

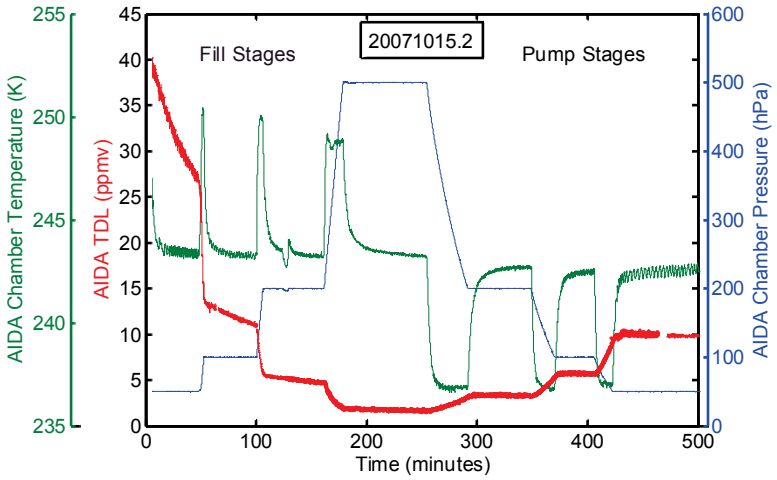


Figure 5.2. Representative structure for a daily comparison run. The AIDA chamber temperature (green) and pressure (blue) are plotted as a function of time. Note that during the static portions (~1 hour each) the temperatures and pressures of the chamber are nominally constant. The mixing ratio of the chamber, as measured by the AIDA TDL facility instrument, is also shown. The mixing ratio decreases progressively primarily through dilution as the pressure is increased during the first half of the run. It increases again as the pressure is stepped down, due to desorption of water from the chamber walls.

5.2.2 The AIDA Chamber as a Reference Standard

The organizers of the campaign initially assumed that the two AIDA facility instruments would provide an unassailable reference standard for the water vapor comparison. In this section, therefore, we evaluate the performance of the AIDA chamber facility instruments and discuss their utility as absolute standards. Specifically, did the AIDA instruments, 1) an extractive chilled mirror hygrometer, i.e., Frost Point device (FP), and 2) a cross chamber long-path Tunable Diode Laser (TDL), provide reliable measurements of the water vapor mixing ratio within the chamber that were consistent with each other within their stated accuracies?

A cursory examination of the results from the two AIDA reference standards (not shown) revealed that the AIDA FP reference was frequently operating under conditions where it was incapable of reliably making a measurement. During the static tests of the AquaVIT campaign, the level of disagreement between the AIDA TDL and the AIDA FP was routinely outside their combined uncertainty of ~5%[1].

During the first three runs on October 15, 16 and 17, 2007, the disagreement between the two AIDA standards showed a significant nonlinear dependence upon the chamber pressure. At the lowest pressures, typically less than 100 hPa, the discrepancy exceeded 15% with the AIDA FP exhibiting a systematic “high bias”. Furthermore, despite convergence toward nominal agreement between the techniques at chamber pressures >100 hPa, the

[1] The stated uncertainty of the AIDA TDL is ±4%, and the uncertainty of the AIDA FP reference standard, is expected to be within ±0.1ºC of the frostpoint, which is equivalent to or better than ±2% in ppmv.

performance from one run to the next was not consistent. In the best case, the run of October 16, 2007, the two reference standards converged to within ±4% at the higher pressure levels. However, during a similar run on October 15, 2007, the agreement even at the highest pressures was at best 15%. Finally, the results from the last two runs of the static tests, on October 18 and 19, were exceedingly poor. During these runs, chamber mixing ratios were routinely low (less than ~5 ppmv), and the differences at *all* pressure levels exceeded 100%, with the AIDA FP reporting values more than ten times those of the AIDA TDL at the lowest pressures.

With any measurement method, systematic errors associated with implementation or exposure to new conditions must be considered in the assessment of measurement accuracy. It would appear that the AIDA FP is incapable of making accurate measurements at low pressure (<150 hPa) and/or at low mixing ratios (<20 ppmv). The nature of the "high bias" does not appear to be consistent with a single error source, i.e., its character does not have the signature of a well-defined pressure or mixing ratio dependence. A quick consultation of the operational manual reveals that the optimal sample gas pressure is 2 bar, with a sample gas flow rate of 0.5 to 1.0 liter/min. During AquaVIT, the pressure range in the chamber was at most 500 hPa, and the flow rates were often a factor of 2 to 5 times greater than the recommended rate. Attempting to make measurements outside the specified ranges will almost certainly impact the quality of the measurement. Though fully characterizing the source and behavior of the error is complicated because of the coupling between flow, pressure, and mixing ratio, a

thorough investigation of the sensitivity of the AIDA FP measurement to the operating parameters needs to be executed before this instrument can be considered a reference standard.

Cross-calibrations between the AIDA FP and the Physikalisch Technische Bundesanstalt (PTB) water permeation source were carried out directly after the AquaVIT campaign. In this arrangement, the PTB functioned as the calibrated reference source. The estimated accuracy of the PTB is 2% for mixing ratios between 0.5 and 5 ppmv [*Fahey et al.*, 2009]. Therefore, the agreement between the two should be within 3%. Though the available range of the permeation source is narrow and the time to equilibration is long, the data acquired at six mixing ratio setpoints from 1.6 ppmv to 13.6 ppmv over a 24 hour period showed that the AIDA FP measured systematically lower than the PTB by ~4% ±3% [*Fahey et al.*, 2009]. This level of disagreement is outside their stated uncertainties despite optimal operational conditions for the AIDA FP (1 atm, 1 SLM). It also provides an excellent example of the difference between the theory of an internationally accepted traceable standard and the realizable accuracy associated with implementation. Again, the requisite sensitivity tests were not executed, so identifying the sources of the systematic error and attributing them to one (or both) of the instruments is not possible.

Furthermore, the poor performance of the AIDA FP during the static runs of the AquaVIT campaign, combined with the systematic disagreement with the PTB permeation source under more favorable measurement conditions casts doubt on the stated accuracy of the other AIDA reference instrument, the AIDA

TDL. Though the AIDA TDL uses the spectral line parameters published in HITRAN, its validation, i.e., the detection and minimization of potential systematic error sources, requires an independent reference-quality measurement of the water vapor in the AIDA chamber. The AIDA FP, which is permanently installed at the chamber, has been used as that reference standard. Assessments of the AIDA TDL were executed at chamber pressures ranging from 0.01 to 1.0 bar, again operating well outside the recommended range of the AIDA FP, and into the region where it demonstrated exceedingly poor performance during AquaVIT. In summary, the methods of calibration and/or validation of the two AIDA reference instruments (FP and TDL) were not consistent with the requirements for an internationally accepted calibration facility. As a result, consistent with the consensus decision among the campaign referees, we treat the AIDA facility instruments as participants in the comparison effort rather than as reference standards.

5.2.3 The AIDA Chamber as a Common Source

In the following section we evaluate the reliability of the chamber as a common water vapor source to a suite of instruments. Specifically, were the mixing ratios homogeneous within the chamber, and did the chamber deliver these mixing ratios unperturbed to the extractive instruments?

In order to assess the performance of the AIDA chamber we use data from the two FISH Lyman-α photo-fragment fluorescence hygrometers developed and operated by the Forschungszentrum Jülich (Germany) [*Zoger et al.*, 1999]. Both

were developed for airborne campaigns. During AquaVIT, FISH1 (LearJet) was located on platform 3 and sampled through a shared 10 mm sampling line also serving the MBW 373LX. FISH2 (Geophysica) was located on platform 1, ~5 m lower down the chamber, and shared a ¾" sampling line with HWV, CFH and FLASH. This pair of instruments provides insight into both the chamber homogeneity and the integrity of the external sampling lines. The results from the two FISH instruments show excellent agreement, generally within their combined uncertainty (~8.5% +0.25 ppmv), for all 5 static experiments over the full range of mixing ratios and chamber pressures.

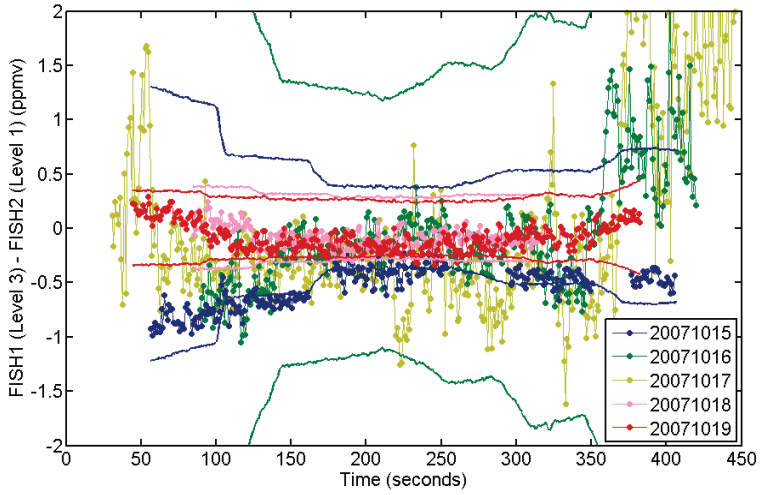


Figure 5.3. Examination of AIDA chamber sampling homogeneity. This figure shows the absolute difference (colored dots) between the pair of FISH Lyman-α hygrometers as a function of time for each of the five formal comparison runs. FISH1 and FISH2 sampled from shared extractive lines on Levels 3 and 1, respectively. The colored lines designate the combined uncertainty bounds of the measurements. The bounds for October 17, 2007, are off the scale of the figure.

Figure 5.3 shows the absolute difference between the FISH instrument pair (colored dots) for each of the runs and the associated error envelope (colored

lines). The larger differences on October 16 and 17, 2007, correspond to the higher mixing ratio data at the lowest pressure levels on these days, and are well within the uncertainty bounds.

In addition to the pair of FISH instruments, the FLASH-B team from the Central Aerological Observatory of Roshydromet, Dolgoprudny (Russia) operated a pair of instruments. These instruments again utilize the Lyman-α photofragment fluorescence technique, and were developed for balloon-borne water vapor measurements [*Kretova et al.*, 1991; *Yushkov et al.*, 2001]. In the atmosphere, the total uncertainty of the FLASH-B measurements is expected to be less than 10% at mixing ratios greater than 3 ppmv, increasing to about 20% at mixing ratios less than 3 ppmv. During the AquaVIT campaign, FLASH-B1 was mounted inside the AIDA vessel at the height of the Level 3 manifold and was capable of making direct in situ measurements. FLASH-B2 was an extractive instrument set up outside the chamber, and connected through a common heated sampling line with FISH2, HWV, and CFH. In theory, this instrument pair offered a complimentary check upon chamber homogeneity as well as a check upon potential differences between in situ and extractive sampling.

Unfortunately the performance of the FLASH-B instruments during the AquaVIT campaign was complicated by several sources of systematic error that do not impact the measurements in situ, but rendered the pair ineffective for diagnosing chamber performance in the laboratory. For the FLASH-B instruments, scattered light from the hydrogen discharge lamp off of the walls of both the AIDA chamber and the small extractive chamber produced a non-

fluorescence artifact or offset, which severely impacted the measurements. The offsets due to stray light backscattering were estimated to be equivalent to 7 ± 2.5 ppmv and 230 ppmv for FLASH-B1 and FLASH-B2 respectively. In addition, the FLASH-B2 instrument was significantly impacted by contamination, despite careful preconditioning, due to water outgassing from the walls of the small extractive chamber. Outgassing, rather than a leak, was understood as the cause because the signal diminished as function of time during each run. The contamination precluded an independent determination of the backscatter offset in FLASH-B2, as FLASH-B1 data were used as a diagnostic. This compromised the independence of the pair and further limited their utility in characterizing AIDA chamber performance.

Despite the limitations of the diagnostic data, we conclude that the chamber homogeneity was robust. The FISH instrument data do not show evidence of significant biases or systematic errors associated with sampling at different levels. Furthermore, the results of comparisons among the full suite of "core instruments", discussed in subsequent sections, did not show evidence for systematic errors associated with in situ or extractive sampling.

5.3 Performance of HWV During AquaVIT

Having established that the AIDA chamber is capable of functioning as a reliable water vapor source to a suite of instruments measuring simultaneously, we turn to an evaluation of the performance of HWV. Consistent with the methodology described in Chapter 3, and similar to the performance assessments

carried out prior to and during any field campaign, we 1) performed frequent laboratory calibrations, tied to a primary standard, over a range of conditions representative of the campaign, 2) analyzed the uncertainties associated with the calibration, and assessed the stability of the calibration over the course of the campaign, 3) identified potential sources of systematic error associated with integration into campaign platform, i.e., onto the aircraft, or at the AIDA facility, 4) performed diagnostic tests to constrain these additional uncertainties, and 5) compared our measurements with those of other instruments.

We use Equation 5.1 to guide the discussion of uncertainty in the HWV measurement. The equation explicitly identifies potential error sources, associated both with the calibration and with other sources of systematic error, and captures the fundamental separability among them. The equation also provides the basis for interpreting the AquaVIT comparison results. Note that this version of the equation is specifically tailored to the AIDA laboratory environment. The desired quantity for comparison, i.e., the common or "true" value, is the sum of water vapor intentionally added to the chamber, vmr_{added}, and any residual water vapor in chamber flow, vmr_{flow}. The final HWV measurement, $vmr_{measured}$, however, may be affected by uncertainty in the calibration, $F_{unc.cal,}$ plus undesired contributions from leaks, outgassing from the walls, and any instrumental measurement bias.

$$vmr_{measured} = (1 + F_{unc.\ cal.}) \times \{ [vmr_{added} + vmr_{flow\ bkgd.}] + vmr_{leak} + vmr_{walls} + vmr_{instrument} \}. \quad (5.1)$$

The accuracy of the measurement depends first and foremost upon accuracy of the calibration. As discussed in Chapter 3, the calibration quantifies

the detection sensitivity of the instrument. Uncertainties in the calibration are represented by a scale or gain factor, denoted by $F_{unc.\ cal.}$ in Equation 5.1. These uncertainties are typically within ~±5% (see Chapter 3), and thus we expect $F_{unc.\ cal.}$ to be within ±0.05.

Uncertainties and/or errors in the HWV measurement may also arise from the combination of $vmr_{leak} + vmr_{walls} + vmr_{instrument}$. vmr_{leak} captures the contribution of leaks from the ambient environment into the HWV instrument upstream of the detection axis. The magnitude of this term will depend upon the pressure and flow rate in the instrument duct, and on humidity of the ambient environment. vmr_{walls} represents the contribution due to the desorption of water from the walls of the instrument and ducting upstream of the detection axis. Again, the magnitude of this contribution will depend upon pressure and flow in the duct, and it may have a time-dependent signature. Finally, $vmr_{instrument}$ represents the contribution to the measured signal from non-water vapor sources. This term accounts for potential systematic errors associated with spurious, i.e., non-water vapor, fluorescence or other stray light sources (see Chapter 4). In the following analysis, we focus upon the quantifying each one of these terms.

A summary of the expected performance of HWV is given in Chapter 2, Table 2.1, and the details describing and justifying each entry in the table are discussed in Chapters 2, 3, and 4.

5.3.1 Evaluation of the HWV Calibration

Field Calibrations

We calibrated the HWV axis in the Harvard laboratory prior to departing for AquaVIT, and performed calibrations utilizing the same bubbler apparatus prior to, or immediately following, each daylong static run during the mission. Figure 5.4 shows the results of these calibrations. The purple dots and solid purple line correspond to the pre-mission calibration data and the results of a robust linear regression, respectively. The calibration data spanned pressures from 40 to 200 hPa, and water vapor mixing ratios from ~2 to 100 ppmv. Also shown are the results of the five calibration runs performed during the AIDA campaign. The data from the first run (blue stars) on October 13, 2007, show outstanding agreement with the results of the pre-mission calibrations. A fit to the calibration data of the October 13, 2007, run yields results within 2% of the combined pre-mission values. This level of agreement is well within the expected uncertainty of the calibration, and demonstrates our ability to perform calibrations with reproducible accuracy at a remote location.

The results from the four subsequent calibrations performed at the AIDA facility reveal a measurable shift in the calibration. Figure 5.4 shows the data from these runs in addition to a least squares fit (black line) to the data bracketed by dotted lines representing the estimated ±5% uncertainty. The difference between the pre-mission calibrations, including the first run done at the AIDA facility, and the subsequent four calibration runs is ~5.5%, just outside the uncertainty bounds. Calibration changes, though infrequent, can occur during a

mission and are generally attributed to changes in the sensitivity of the UV diode, which is used to monitor and normalize the UV lamp flux. As discussed in Chapter 3, we have numerous diagnostics to evaluate such calibration changes. During AquaVIT switching between chamber sampling and the calibration system was executed by opening and shutting two valves, making it possible to both quickly identify and accurately quantify calibration change. In the analysis of the AIDA static runs, we chose to use the pre-mission calibration for the first run on October 15, 2007, and applied the revised calibration numbers to the remainder of the static runs.

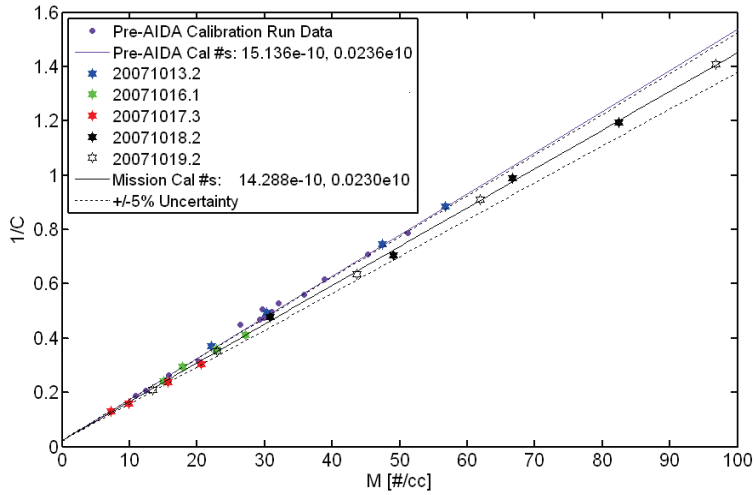


Figure 5.4. Calibration summary of the primary HWV axis during the AquaVIT campaign. The inverse of the detection axis sensitivity is plotted as a function of the number density in the duct. The reproducibility and linearity of the calibration over a wide range of pressures and mixing ratios is evident.

AIDA Instrument Comparisons

Next we evaluate the calibration of HWV through AquaVIT instrument comparisons. The consensus among the referees and participants of the AquaVIT campaign was that there was no single measurement that represented the "real" mixing ratio in the chamber [*Fahey et al.*, 2009]. As discussed earlier, the AIDA FP instrument was subject to unidentified and unquantified errors that rendered it ineffective as a reference standard. Though the performance of the AIDA TDL during AquaVIT was rational, we do not feel that it is sufficiently vetted, particularly with respect to calibration uncertainties (see above), to serve as the sole standard for judging the performance of HWV or any other instrument participating in the comparison. Thus we compare HWV with a selection of the core instruments including FISH2 and CFH, which shared a common sampling line with HWV, and the AIDA TDL and JLH, which measured in situ.

As described in Equation 5.1, an error associated with the HWV calibration will be evident as a scale factor. For example, in the absence of other large sources of systematic error, we expect the slope of a linear fit between data from HWV and Instrument X to capture differences primarily associated with their respective calibrations. Similarly, we expect plots of the fractional difference between mixing ratio data from HWV and Instrument X, to yield a constant value less than or equal to the combined calibration uncertainty of the instrument pair.

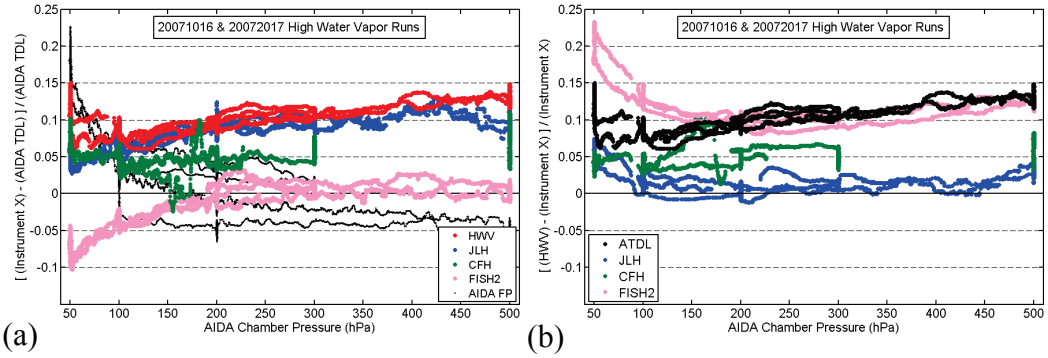


Figure 5.5. Fractional differences between core instruments. Panel (a) shows the fractional difference between a subset of the core instruments and the AIDA TDL as a function of chamber pressure for the two high water vapor comparison runs. Panel (b) is a comparable plot using HWV as the reference instrument. The fractional differences at high water and higher chamber pressures capture calibration, or scale, errors.

Accordingly, Figure 5.5a shows the fractional difference between the AIDA TDL and four of the core participant instruments, HWV (±5% accuracy at pressures >100 hPa), JLH (±10%), CFH (±2%), and FISH2 (±6%) plotted as a function of pressure, using the data from the high water static runs on October 16 and 17, 2007. Only data where all five instruments archived data were included, and a median filter was applied. The data from these runs were chosen because the impact of artifacts, offsets, and other systematic errors on the measurements is expected to be minimal relative to the impact on measurements from lower mixing ratio runs. Additionally, chamber temperatures on these days were warmer than 210 K. Figure 5.5b highlights the calibration differences between HWV and the same set of instruments. The fundamental disagreements among the calibrations of these core instruments are immediately apparent. Systematic errors, possibly associated with pressure effects, are also evident.

As is evident in both panels of Figure 5.5, the subset of instruments examined here is essentially divisible into three camps: 1) HWV and JLH measure systematically high with respect to the AIDA TDL but agree well with each other, 2) FISH2 agrees well with the AIDA TDL, particularly for chamber pressures >200 hPa, and 3) CFH bridges the gap. Figure 5.5b emphasizes the excellent agreement between JLH and HWV, which is within 3% ±3%, well within their combined uncertainty. The agreement between CFH (green dots) and the AIDA TDL, shown in Figure 5.5a, is ~+5% ±5% and is just at the edge of their combined uncertainty envelope, with CFH measuring systematically higher. Similarly, the agreement between HWV and CFH, shown in Figure 5.5b, is ~+5 ±5%, with CFH measure systematically lower. Finally, the relationship between FISH2 and the AIDA TDL is excellent, well within the ±3% at pressures greater than 200 hPa. At lower chamber pressures, however, the agreement rapidly degrades, with FISH2 measuring systematically lower by as much as ~8% at the lowest pressure level, ~50 hPa. The differences between JLH (blue dots) and the AIDA TDL are ~+8% ±5%, falling just within the combined uncertainty of the pair due the conservative uncertainty estimate of JLH, while the differences between HWV (red dots) and the AIDA TDL are slightly larger, ~+10% ±5%, and are outside their combined uncertainty. Overall these differences, in some cases exceeding 5% between instrument pairs, are larger than expected and are indicative of unidentified systematic errors affecting more than one of the core instruments.

Instr. X = HWV × *Slope* + *Intcpt.*	Slope	Intercept
ATDL	**0.92**	−0.39
CFH	**0.95**	−0.01
JLH	**0.99**	−0.08
FISH2	**0.89**	+0.15

Table 5.1. Results of robust linear least squares fits to instrument comparisons. The slope of the linear fit captures scale, or gain, errors, which are generally attributed to errors in instrument calibration.

Table 5.1 quantifies the relationship between HWV and each of the selected instruments, showing the results of robust linear regressions. As in Figure 5.5, we used the data from the high water static runs of October 16 and 17, 2007. Furthermore, the data were selected for high, i.e., >20 slm, flow rates through the HWV instrument, in order to reduce the impact of uncertainties associated with low pressure and low flow data. The numerical results support the visual analysis of Figure 5.5, and provide the scale factors that are used henceforth when comparing the results of HWV with those of the other instruments.

It is worth noting that the agreements/disagreements do not correspond with measurement method or measurement location, i.e., in situ or extractive. For example, the agreement between the two TDL instruments (ATDL and JLH) is just barely within their combined uncertainty envelope (~10%), despite the fact that they both utilized the same spectral line, with the same line parameters from the HITRAN database, and operated under the same temperature and pressure conditions inside the AIDA chamber. In addition, JLH was configured to take direct absorption measurements every 20 seconds, periodically freeing it from uncertainties associated with the 2f detection method. Furthermore, the greatest differences were observed between the pair of Lyman-α hygrometers, FISH2 and

HWV, which shared a sampling line. Nor does the choice of calibration reference standard explain the differences. At pressures >200 hPa neither FISH2, which calibrates with a frost point reference standard, nor HWV, which is tied to the liquid water saturation vapor pressure through the bubbler, showed appreciably better agreement with the participant frost point instrument, CFH. Both measure within ~5% ±3% of CFH, with HWV measuring systematically higher than CFH and FISH2 measuring systematically lower.

Discussion

In theory, comparisons of this type, which use different detection techniques to make simultaneous measurements over a range of mixing ratios and pressures, can facilitate an identification of the sources of systematic error that lead to instrument inaccuracies of the magnitudes observed during AquaVIT. However, the formal structure of AquaVIT, with the blind data protocol, and the sluggish nature of chamber equilibration precluded this type of investigative and responsive comparison. Therefore, in the absence of more rigorous tests to either substantiate the accuracy of the AIDA TDL, or discover sources of systematic error affecting both JLH and HWV, and to a lesser degree CFH, we feel confident in the both the choice of the bubbler as our reference standard, and in the stated ±5% uncertainty of the HWV calibration. Comparisons with both the JLH and CFH instruments demonstrate a level of agreement consistent with this uncertainty estimate. Furthermore, there is some evidence, albeit indirect, to support the existence of a ~–4% systematic gain error affecting one or both of the

AIDA facility instruments. Based upon a very limited cross-calibration dataset, it was shown that the AIDA FP measured ~4% lower than the PTB permeation source. Because the AIDA TDL has routinely evaluated its accuracy using the AIDA FP as its measurement reference standard, it is possible that its measurements are subject to errors of a similar magnitude.

So far we have ignored the obvious secondary effects evident in Figure 5.5, e.g., the systematic pressure dependence between the core instruments. When compared with the AIDA TDL, the evidence for pressure related errors among the different instruments, ranked from least to the greatest is: CFH, HWV, JLH and FISH2. Similarly, in the comparison utilizing HWV as the reference instrument, the order, again from least to greatest is: CFH, JLH, AIDA TDL and FISH2. In all cases, FISH2 is the outlier, with a large non-linear deviation at pressures below ~200 hPa. In general, the linearity exhibited between HWV, CFH and JLH is excellent. Once the data have been corrected for the ~5% gain errors, the remaining systematic errors associated with pressure are well within the uncertainty estimates, providing further confirmation of the sound performance of all three.

Because collisional quenching competes with fluorescence, the sensitivity of any Lyman-α detection axis is a function of pressure. The final calibration numbers are in fact derived from a linear fit of $1/C$, where C is the sensitivity at a given pressure, versus number density, M, see Figure 5.4, and Chapter 3. As discussed above, errors in the slope of the calibration line manifest as scale factor or gain errors. In Figure 5.5 these errors will move the fractional difference up or

down, but will not result in a pressure dependence. However, errors in the determination of the calibration intercept, i.e., at $M = 0$ [#/cc], can yield a nonlinear pressure dependence, with the low pressure data revealing the largest sensitivity to intercept errors. The AIDA comparisons show that HWV performed well across the full pressure range, however, the FISH2 data bear the characteristic signature of an intercept error.

In summary, the calibration of HWV during AquaVIT was tied to the saturation mixing ratio of water vapor over liquid at room temperature, as provided by the bubbler water vapor addition system that is routinely used in Harvard laboratory calibrations. We believe that we can achieve an accuracy of ~5% or better with the bubbler (see Chapter 3). Cross-validation with the long-path absorption of 121.6 nm light in the Harvard laboratory provides verification of this. The excellent agreement between the first calibration run at the AIDA facility and the results of calibration runs prior to the mission, executed at Harvard, serves as a confirmation that no significant systematic errors were introduced in the transition, and the stability of the calibrations at the AIDA facility is evident in Figure 5.4. Instrument comparisons between HWV and CFH and HWV and JLH serve to further corroborate the accuracy of HWV.

Comparisons between the AIDA TDL and the German Lyman-α instrument, FISH2, yielded disagreements that are just outside the combined uncertainties. Unfortunately, however, the structure of the mission did not allow time for a close examination of these differences.

5.3.2 Quantifying the Contributions due to Leaks and Outgassing

Now that we have established confidence in the stated ±5% of the HWV calibration, i.e., in our determination of $F_{unc.\ cal.}$, we will attempt to quantify the remaining terms in Equation 5.1. Recall that during the AIDA comparisons the combination of $vmr_{added} + vmr_{flow\ bkgd.}$ is the common quantity measured by the participant instruments. Thus, the value that we need to quantify and remove, or at least bound for the purposes of the comparison, is the sum of the remaining terms: vmr_{leak}, vmr_{walls}, and $vmr_{instrument}$. Data from targeted diagnostic runs and the AquaVIT comparisons will be used for this purpose.

We begin by evaluating the contribution due to leaks into the HWV instrument duct. The common line serving HWV, CFH, FISH2 and FLASH2 was determined to be leak tight. However, the leak rate into our ducting, with the valve between HWV and the common line shut, was ~0.8 hPa/min. A back-of-the-envelope calculation shows that this is equivalent to a ~1.6 sccm flow rate of room air into the HWV instrument, assuming a ~2 liter volume for our instrument and ducting. At room temperature and pressure, the saturation mixing ratio is ~2.5%. If we assume an ambient relative humidity of 50% and a flow rate through the HWV instrument equal to 20 slm, the offset due to a leak yields vmr_{leak} ~1.0 ppmv. For a constant leak rate, the magnitude of the offset is inversely proportional to flow. It is also possible that the leak rate will depend upon the pressure differential between the room and the duct, i.e., it will be higher at lower pressures duct pressures. Furthermore, at the lowest chamber pressures it was impossible to maintain the requested flow rates, exacerbating the flow and

pressure effects on vmr_{leak}[2]. At the lowest pressure level in the chamber (~50 hPa) the fastest flow rate we were able to maintain was ~18 slm through a high throughput flow controller and pump. Under these conditions the vmr_{leak} term can become a significant fraction of $vmr_{measured}$.

The contribution to due outgassing from the ducting upstream of the detection axis, vmr_{walls}, behaves like a "virtual leak". As is discussed in detail in Chapter 4, the outgassing signal, similar to a leak, is expected to vary inversely with both the flow in the duct and with pressure.

Diagnostic Tests and Development of Leak and Outgassing Correction

We performed two types of diagnostic runs in order to test our sensitivity to leaks and outgassing. In the first case, we sampled from the AIDA chamber during the nominally constant pressure and mixing ratio stages, and varied the pressure and flow in our duct by throttling the exit valve located downstream of the detection axis. In the second case we utilized the facility's dry air supply and operated independent of the chamber. These independent tests gave us more control over the pressures and flows in our duct and as a result provided better diagnostic data. Our tests became more sophisticated as the campaign progressed, and we as realized the importance of quantifying these contributions for the very low water vapor chamber runs.

[2] At 500 hPa in the chamber, for example, a 50 slm flow rate is equivalent to 100 liters/min, but at 50 hPa it equals 1000 liters/min.

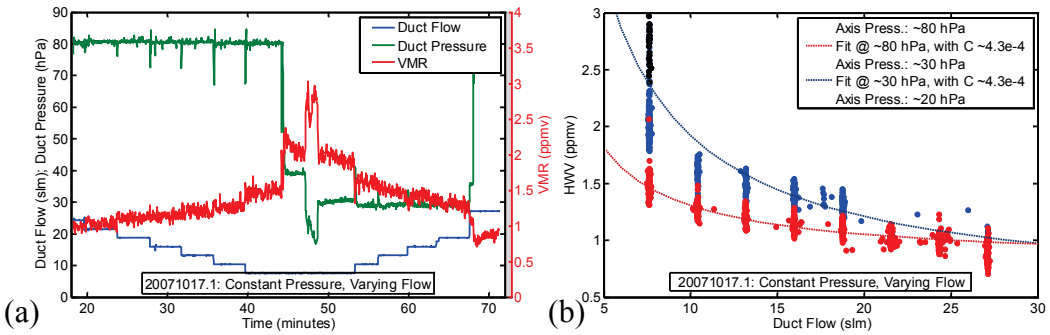


Figure 5.6. Offset diagnostic data and model results. Panel (a) shows water vapor mixing ratio measurements by HWV (red line) as a function of time during a diagnostic test where the duct pressure (green line) was nominally held constant and the mass flow (blue line) was varied. The background mixing ratio, $vmr_{flow\ bkgd.}$, and any instrumental offset, $vmr_{instrument}$, are expected to be constant at constant pressure. Therefore, the changes in the measured mixing ratio are attributed to the leak and outgassing terms. Panel (b) shows the HWV measurements plotted as a function of pressure (color dots). Also shown are the results of an empirically tuned model (dashed lines), which was designed to quantitatively account for a pressure and flow dependent sampling offset.

The results of a test designed to explore the sensitivity to flow at constant pressure are shown in Figure 5.6. Figure 5.6a shows the duct pressures and flows as well as the measured mixing ratio as a function of time. We ran the facility's dry air through a 50 slm flow controller and into our instrument. At each flow set point we adjusted the exit valve to maintain constant pressure in the duct, thereby allowing flow to be the controlling variable. We explored the effect of varying the flow at two duct pressures. 80 hPa was the lowest pressure we could achieve at 30 slm when we were plumbed through the smaller of two AIDA facility flow controllers. Switching our exit line to the larger flow meter reduced the pressure to ~30 hPa. Figure 5.6b shows the measured mixing ratio as a function of flow, where the colors denote the two duct pressures. The mixing ratio of the dry air source, $vmr_{flow\ bkgd.}$, is expected to be a constant for the duration of the test. We

also expect the contribution of any instrumental artifact to be constant, $vmr_{instrument}$, thus variability in the measurement is attributed to the combination of the leak and outgassing terms.

These data provided the foundation for the development of a simple model to correct the HWV measurement for the contributions due to leaks and outgassing. The model is represented by Equation 5.2, where the offset due to leaks and outgassing, *Offset*, is measured in ppmv, and *C* is an empirically determined scale factor, the value of which is dependent upon the humidity of the room, the size of the leak, and the amount of water coating the surfaces of the ducting.

$$Offset = C \times [\ (P_{room} - P_{duct}) / P_{duct}\ ] \times (1 / Flow) \qquad (5.2)$$

The ratio of the difference in pressure, between the room and the duct, to the duct pressure captures the potential pressure dependent portion of both the leak and outgassing terms. The inverse dependence of both terms on the mass flow in the duct is also represented. Figure 5.6b shows the results of the model with values for *C* (~4.3×10^{-4} [ppmv·sccm]) chosen to match the measurements. The model results have also been adjusted with an additive constant to match the measured values at the highest flow setting. The model reproduces the shape of the flow dependence and captures the difference between the two pressure settings. The full series of diagnostic tests showed that the value of *C* was typically ~8.5×10^{-4} [sccm · ppmv] ± 4.0×10^{-4} [sccm · ppmv].

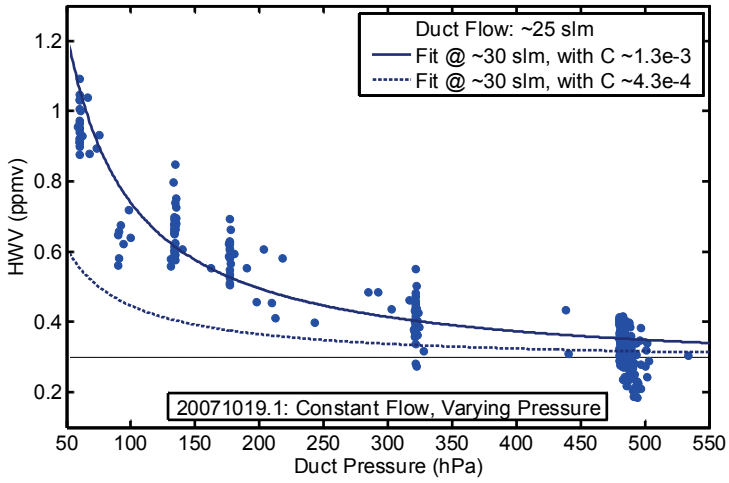


Figure 5.7. Offset diagnostic data and model results. This figure shows measurements (blue dots) from constant pressure, varying mass flow diagnostic run as a function of duct pressure. The results of an empirically tuned model, designed to quantitatively account for a pressure and flow dependent sampling offset, are plotted in the solid blue line. The dashed blue line demonstrates the sensitivity of the model to the choice of the scale factor, *C*.

Figure 5.7 shows the results of a comparable test designed to explore the sensitivity of the measurement to changing pressure at a constant flow. We ran the facility's dry air through a 50 slm flow controller and controlled the pressure in the duct with the exit valve. The measurements and the scaled model results are plotted as a function of duct pressure. Here, good agreement is shown with a value of $C \sim 1.3 \times 10^{-3}$. The model result using $C = 4.3 \times 10^{-4}$ is shown for comparison. Unfortunately, given the limited number of tests we were able to run during the campaign, we were not able to fully explore the instrument's behavior over the full span of relevant pressures and flows. It is expected that the value of *C* will vary from day to day, however, the observed factor of three change between these two tests is most likely due to factors not accounted for in our simple model. The degree of uncertainty in determining the value of *C*, largest at

the lowest flows and pressures, is included in the uncertainty of the final water vapor measurement.

Correcting HWV Data and Estimating Correction Uncertainty

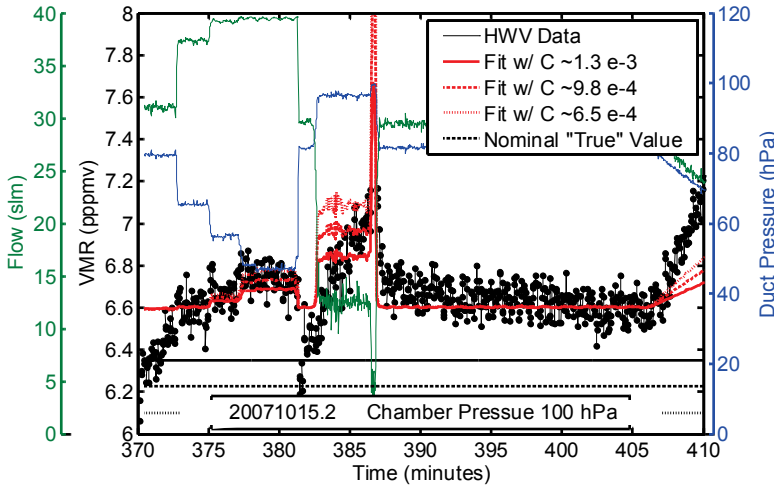


Figure 5.8. HWV measurements during a constant pressure leg of the October 15, 2007, AIDA comparison run. HWV measurements (black dots) are plotted as a function of time. The chamber pressure was held constant at 100 hPa. The pressure (blue line) and mass flow (dark green line) in the HWV duct were varied during the first part of this leg for diagnostic purposes. In theory, the mixing ratio delivered to the HWV instrument is nominally constant over this interval. The red lines show the results of the offset model, which quantitatively accounts for pressure and flow dependent sampling offsets, due to leaks and outgassing in HWV. The different model results (solid, dashed and dotted lines) utilize different values for the empirically determined scale factor *C*, and assume different values for the constant mixing ratio delivered by the chamber.

Figure 5.8 shows data taken during a pressure and flow diagnostic test run while sampling from the AIDA chamber. The plot shows the HWV duct pressure (blue line) and flow (dark green line) as a function of time. The tight coupling between pressure and flow is evident during this diagnostic test, where slower flows correspond to higher pressures and vice versa. The pressure/flow scan

shown in Figure 5.8 was executed at a constant 100 hPa pressure level in the AIDA chamber, where the delivered mixing ratio is expected to be relatively constant. During this pressure/flow diagnostic period we sampled at each set point for ~2 minutes. The uncorrected HWV mixing ratio data (black dots), i.e., the measurement of the actual delivered mixing ratio plus the contributions due to vmr_{leak}, vmr_{walls}, and $vmr_{instrument}$, show evidence of time constants longer than the time expected for equilibration at each set point. These time constants suggest the presence of uptake and desorption from the walls of the AIDA chamber itself, processes not included in the simple leak model.

Also shown in Figure 5.8 are model results for three different values of *C*. The red lines are the sum of the calculated model offset and the nominal "true" mixing ratio delivered by the AIDA chamber. The "true" values (black lines) are lower than the actual HWV measurements by ~0.50 to ~0.25 ppmv depending upon the value of *C* chosen. Higher values of *C* correspond to larger offset corrections and thus lower "true" values. The data from this test show good qualitative agreement with the model, but on their own are not sufficient to constrain *C* to the desired level of accuracy.

Because of the design of the AIDA campaign, the experimenters did not have access to the facility data or to data from any of the participant instruments. As a result, we were forced to rely on the data from our limited set of diagnostic tests to determine the optimal correction factor. Our original estimate for *C* was $9.8\times10^{-4} \pm 3.3\times10^{-4}$. This was the value we applied to the data we archived initially, and the uncertainty of the reported measurement includes the large

uncertainty in the offset determination. Figure 5.9 shows the calculated offset attributed to leaks and outgassing in ppmv as a function of the AIDA chamber pressure using a value of C ~9.8×10^{-4}. Also shown is our estimate of the uncertainty in this term. As expected, the largest correction factors (~2 ppmv) with the largest uncertainty (±0.7 ppmv) are associated with the slowest flows and lowest pressures. According to our model, the contribution from leaks and outgassing rapidly exceeds ~0.1 ppmv for chamber pressures less than ~200 hPa, reaching ~1.0 ppmv ± 0.3 ppmv at ~80 hPa. At higher chamber pressures the impact of these terms is reduced, and thus the impact of other sources of uncertainty becomes more significant.

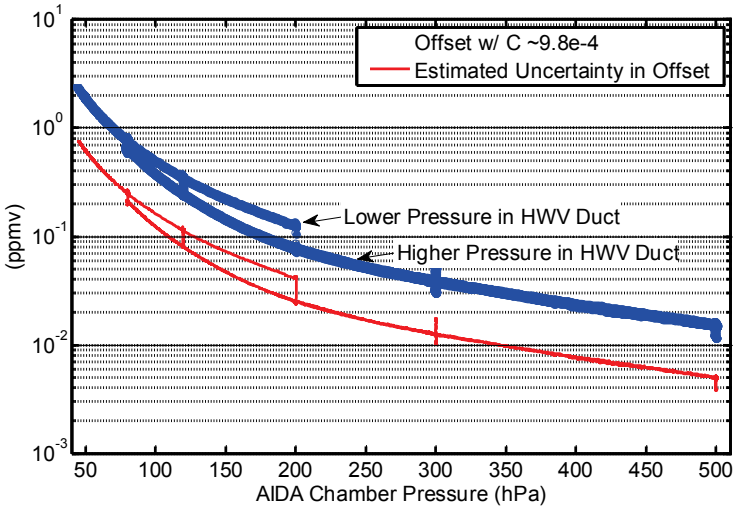


Figure 5.9. Estimated offset correction to HWV and uncertainty in correction factor. This blue dots show the estimated magnitude of the mixing ratio correction to the HWV measurements for sampling errors, e.g., leaks and outgassing in the HWV instrument and sample lines, as a function of the AIDA chamber pressure. The red line shows the uncertainty associated with the offset correction.

Figures 5.10a & b show the pressure in the AIDA chamber, the AIDA TDL mixing ratio, and the absolute difference between HWV and the AIDA TDL

before and after the HWV data have had the offset correction applied. In this analysis, the AIDA TDL data have been increased by ~8% in order to account for the calibration difference identified earlier. Figure 5.10a shows a plot of the difference between the HWV measurements and those of the AIDA TDL for a portion of the October 15, 2007, AquaVIT run. The AIDA chamber pressure settings range from 50 hPa to 500 hPa, with mixing ratios ranging from ~15 ppmv to 2 ppmv respectively. The figure shows the difference between the two instruments without correcting HWV, as well as the difference with HWV corrected using $C = 9.8\text{x}10^{-4} \pm 3.3\text{x}10^{-4}$.

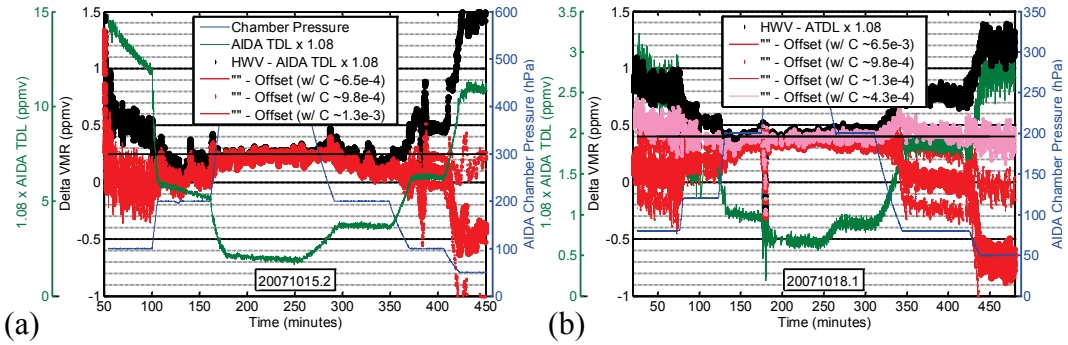


Figure 5.10. Difference between HWV and the AIDA TDL – raw and corrected. Panels (a) and (b) show mixing ratios measured by the AIDA TDL facility instrument (dark green line) located within the chamber, and the AIDA chamber pressure (blue line) as a function of time, for two different comparison days. Panel (a) shows data from the run of October 15, 2007, with mixing ratios ranging from ~2 ppmv to 15 ppmv and pressures from ~50 hPa to 500 hPa. Panel (b) shows data from the run of October 18, 2007, with lower mixing ratios, ~<3 ppmv, and pressures. The differences between HWV and the AIDA TDL are plotted in the black dots, where the AIDA TDL data have been scaled by a factor of 1.08 to account for the calibration difference between the two instruments. The red dots show the differences between the corrected HWV data and the AIDA TDL, where different model scale factors have been applied.

Figure 5.10b shows a similar plot for data from the low water vapor run on October 18, 2007. Pressures during this run varied from 50 to 300 hPa, with the

corresponding mixing ratios ranging from ~3 ppmv to 0.7 ppmv. The difference between the corrected HWV data and the AIDA TDL is evaluated with four *C* values. On this day the best, i.e., most consistent result at all pressure and mixing ratio levels, is achieved with $C \sim 4.3 \times 10^{-4}$ (pink line).

Taken together these plots reveal four primary conclusions: (1) The estimate for *C* applied to the archived data appears to be too large and thus the model overcorrects the measurements at the lowest pressures, i.e., below ~150 hPa; (2) A value of $C \sim 5.4 \times 10^{-4} \pm 1.0 \times 10^{-4}$ yields the best agreement with the AIDA TDL, i.e., the absolute difference is nominally constant ~0.30 ppmv ± 0.10 at all pressure levels and mixing ratios; (3) Even when the data have been corrected for leaks and outgassing there remains some variability in the difference that appears to be associated with changes in pressure and/or flow, suggesting the presence of time constants in the AIDA chamber, delivery plumbing, instrument duct and/or sampling line that are not included in the simple leak model; (4) Finally, the remaining difference of ~0.30 ppmv ±0.10 ppmv between the AIDA TDL and HWV at the highest pressures and flows is not consistent with a leak or outgassing.

Summary of Leak and Outgassing Correction

We reiterate that in the laboratory the sources of systematic error are inevitably and predictably different from those in flight. As was stated earlier, Equation 5.1 was tailored for evaluating the performance of HWV in the AIDA laboratory environment. Neither outgassing nor leaks are expected to impact the

in situ HWV measurements. In flight there is little pressure differential across the instrument duct, in fact pressures inside the duct are greater than ambient, often by ~10 hPa. In addition, mass flows in the instrument duct are a factor of ~30 greater in flight than those achieved in the laboratory, minimizing wall effects in situ. (The results of diagnostics designed to detect and quantify the impact of outgassing in flight are discussed in more detail in Chapter 3.) Thus, the systematic bias and large uncertainties associated with leaks and outgassing that degraded the performance of HWV at the AIDA facility do not affect the flight measurements.

In short, the evaluation of potential error sources must be made relevant to each specific instrument implementation, and the performance of the instrument judged accordingly. There is an understandable temptation to extrapolate instrument performance during the AquaVIT campaign to performance in flight. However, because the environmental and sampling conditions during the AIDA campaign for most of the instruments were not representative of the flight environment, the consensus of the organizers of the comparison, the referees, and the participating PIs was to approach such extrapolations with caution.

5.3.3 Evaluating Instrumental Offsets

Chopper Reflectivity Correction

HWV was subject to an additional source of uncertainty during the AquaVIT mission. The chopper arm was mounted askew, such that in the closed (chopper in) position light was scattered off of the quartz window mount directly

into the PMT. This caused the scattered light to be higher with the chopper in than with the chopper out, leading to an erroneously high background subtraction and a negative measurement offset (see discussion of this correction in Chapter 4). Diagnostic data acquired during the mission, in combination with focused tests performed in the Harvard laboratory after the mission, were used to determine the appropriate correction for the background scatter. Additive uncertainties associated with the determination of the chopper correction factor were of the order of ~±0.20 ppmv.

Instrumental Artifact

Measurement differences between any one of the participant instruments and the "true" water vapor value, or between any pair of participant instruments, can be attributed to multiple factors including, calibration errors and sampling errors. We have attempted to identify, quantify, and correct for calibration and sampling errors in the previous two sections, and focus here upon the remaining systematic differences that may be due to a HWV instrumental artifact. Comparisons during AquaVIT provided a means of quantifying this term relative to a selection of the core instruments, including the AIDA TDL, JLH, CFH and FISH2.

Recall that in Chapter 4, prior to the TC4 and AquaVIT campaigns we assembled a new "offset-free" HWV detection axis (the one used in the AIDA campaign). More accurately stated, we assembled an axis with a measured offset of ~0.2 ppmv in nitrogen, which is equivalent to ~0.4 ppmv in air (Figure 4.10 in

Chapter 4). Because this residual non-zero measurement represents the sum of $vmr_{flow\ bkgd.} + vmr_{instrument}$, a value of ~0.4 ppmv provides the maximum possible estimate of an instrumental artifact in air. Furthermore, if we make the reasonable assumption that ~0.1 ppmv of the measured offset is due to residual water in the dry nitrogen source, we achieve an upper bound to an instrumental bias equal to ~0.1ppmv in nitrogen and ~0.2 ± 0.1 ppmv in air.

The residual offset between the AIDA TDL and HWV of ~0.30 ppmv ± 0.10 ppmv, evident in both panels of Figure 5.10, cannot reasonably be attributed to a leak and/or outgassing. Though we've demonstrated the limitations and large uncertainties of the offset correction model at low chamber pressures and flows, the uncertainty in the model results at the high pressures and flow rates is far less than the magnitude of the observed difference. For example, we expect an outgassing offset, associated with desorption from the walls of the instrument duct, to diminish at the higher pressures and flows, and we expect such an offset to decrease with time at a given pressure and mixing ratio set point. There is some evidence of an outgassing signature in the first sampling leg of both the runs shown in Figure 5.10. However, the stability of the residual difference, particularly evident in Figure 5.10b over a four hour period during the fastest flows and highest pressures, has the quality of a systematic instrumental offset, i.e., $vmr_{instrument}$.

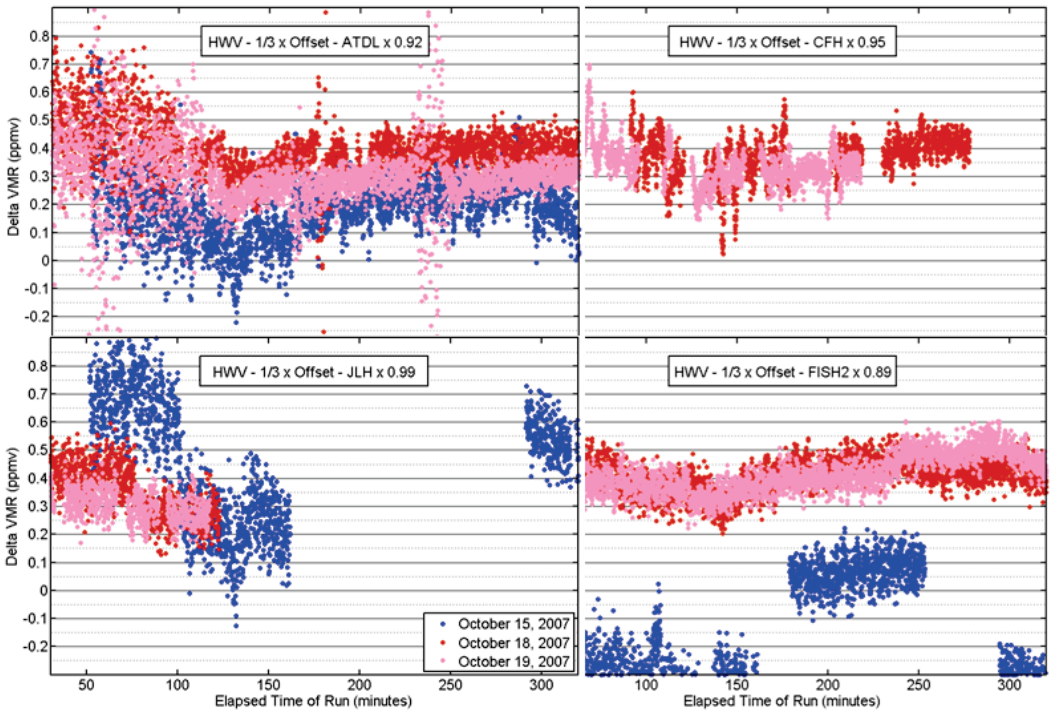


Figure 5.11. Differences between four of the core instruments and HWV. Plotted in each of the four panels are the differences between the HWV measurements, which have been corrected for offsets due to outgassing and leaks, and measurements from the AIDA TDL, CFH, JLH and FISH2, respectively, as a function of time. The colors denote the date of the run. Each instrument has been scaled to account for calibration, i.e., gain, differences between it and HWV. Thus, the remaining differences (~0.35 ppmv) between HWV and the other core instruments are attributed to an instrumental bias in HWV and/or systematic errors in the comparison instrument.

In Figure 5.11, the absolute difference between HWV and four of the core instruments is plotted as a function of elapsed time during each run, where calibration differences have been accounted for with the appropriate scale factor, and the leak and outgassing correction has been applied to HWV. Only data from the fill stages of the low water vapor runs of October 15, 18, and 19, 2007 are shown. Furthermore, for reasons discussed above, we are most confident in the data at the higher pressure levels and flow rates, and thus focus the discussion on the data acquired between 100 and 300 minutes. Indeed, there is a greater degree

of variability evident among all the instruments in the first 100 minutes of these runs. JLH did not record data at chamber pressures >200 hPa for leaving very little data for comparison, and CFH did not participate on October 15, 2007.

The difference between HWV and the AIDA TDL, shown in Figure 5.11, is reasonably consistent from run to run, and is roughly ~0.30 ± 0.10 ppmv. In contrast, the differences between HWV and JLH and HWV and FISH2 show erratic behavior during the run of October 15, 2007, suggesting the presence of errors not associated with HWV. The consistency of the differences during the final two low water vapor runs, however, allows for estimates of the offset of HWV relative to JLH, ~0.30 ± 0.10 ppmv, of HWV relative to FISH2, ~0.40 ± 0.10 ppmv, and of HWV relative to CFH, ~0.35 ± 0.10 ppmv. In summary, the HWV measurement exhibits a relatively constant positive offset relative to a selection of the core instruments that is not representative of a sampling error, but bears the signature of an instrumental measurement artifact of the order of ~0.30 ppmv. Though indicative of a HWV measurement bias, these differences are not of sufficient magnitude to explain or resolve the outstanding discrepancies among HWV and CFH in situ.

5.4 AquaVIT Mission Summary and Conclusions

The AquaVIT campaign, hosted by the AIDA research facility, was the first-ever formal comparison campaign that simultaneously tested the primary in situ water vapor measurement methods in a controlled laboratory setting. The underlying goal of the first week of the campaign was to examine the

performance of a select suite of instruments, which share a long history of making measurements in the UT/LS, in order to address the outstanding discrepancies among the measurements, and to provide a concrete foundation for interpreting the extreme supersaturations (see Chapter 6). These core instruments included the Harvard and Forschungszentrum Jülich (Germany) Lyman-α photofragment fluorescence hygrometers (HWV and FISH), and the NOAA cryogenic frost point hygrometer (CFH), which shared a common heated external sampling line. The selection of core instruments also included the JPL tunable diode laser hygrometer (JLH), and the AIDA facility TDL (ATDL), which were both mounted inside the chamber. Results from the first week showed that the AIDA chamber successfully functioned as a common water vapor source to the full complement of core instruments. It is currently the only facility capable of providing a stable mixing ratio to such a large suite of instruments simultaneously. Unfortunately, however, systematic differences between the AIDA facility instruments, the AIDA TDL and a cryogenic frost point device (AIDA FP) on an external sampling line, demonstrated that there was no single instrument or pair of instruments capable of providing a measure of the "true" water vapor mixing ratio supplied by the chamber. Despite this, the excellent performance of the AIDA chamber and the core instruments during the first week yielded an unprecedented laboratory comparison dataset.

Though in some respects, testing the instruments under the well-controlled conditions of the laboratory, i.e., removed from the vagaries of flight, provides a "best-case" scenario for comparison, the differences between the laboratory and

in situ environments fundamentally limits an extrapolation of the laboratory results to flight. The nature of the differences between the laboratory and flight sampling conditions was different for each instrument. In some cases, the sources of error or uncertainty associated with the laboratory implementation degraded instrument performance, while in other cases it is possible that the laboratory environment improved instrument performance. Though the pressures and temperatures were representative of atmospheric conditions for those instruments mounted inside the chamber, such as JLH, the flows were a fraction of those experienced on an aircraft or balloon sonde. For those instruments mounted externally and sampling through heated lines, FISH1 and FISH2, FLASHB-2, HWV, and CFH, the pressure differential across the instrument ducting, the temperatures and the flow rates were considerably different than those encountered in flight. We have shown that both the pressure difference and the low flow rates during AquaVIT negatively impacted the performance of HWV. However, it is possible that laboratory flow rates through CFH, which were roughly an order of magnitude slower than in flight, improved frost layer formation and instrument performance. The formality of the blind comparison precluded a thorough evaluation of the differences between the laboratory and flight implementations, and their impacts for each instrument.

Overall, HWV performed well during AquaVIT. Table 5.2 summarizes the performance characteristics of HWV during the campaign. (Note the mission specific differences from the expected in situ performance characteristics enumerated in Table 2.1.) Our precision was excellent, within ~0.03 ppmv for 5

second data for the lowest mixing ratios in the chamber, <1.0 ppmv, and was within 0.5% for chamber mixing ratios greater than 10 ppmv. In fact, HWV frequently outperformed the AIDA TDL. The first laboratory calibration performed at the AIDA facility was consistent with calibrations performed at Harvard prior to the campaign, agreeing to better than 2%. The consistency among subsequent calibrations performed during the campaign, over a representative range of pressures and water vapor mixing ratios, further demonstrated the reproducibility and robustness of the HWV calibration. Uncertainty in pinning the intercept translated into the larger errors bars for measurements at the lowest pressures. While we did not have the opportunity to cross-calibrate with the MBW-373 LX during the campaign, comparisons of HWV with both JLH and CFH showed agreement within the combined uncertainty of each instrument pair. Though the agreement between HWV and both the AIDA TDL and FISH2 fell outside their combined uncertainty envelopes, in the absence of an accepted reference standard, we feel that the agreement with JLH and CFH corroborates the HWV calibration and its associated uncertainty. At a minimum, it suggests that calibration issues alone cannot explain the large differences evident in the in situ data. Again, the structure of the campaign did not allow for an investigation of these calibration discrepancies.

	1 - 10 ppmv	10 - 100 ppmv	>100 ppmv
Sampling Period	1 second	1 second	1 second
Pressure Range	<100 hPa >100 hPa	<100 hPa >100 hPa	<100 hPa >100 hPa
Precision (typical)	±<0.1 ppmv	±<1%	±<1%
Accuracy	±10% ±5%	±10% ±5%	±10% ±5%
Artifacts/Offset	–0.20/+0.50 ppmv	–0.20/+0.50 ppmv	–0.20/+0.50 ppmv

Table 5.2. The instrument performance characteristics for the 2" HWV instrument integrated at the AIDA laboratory facility.

Differences between the laboratory and flight implementations, in addition to some atypical instrument performance issues led to larger measurement uncertainties than expected for HWV. The large pressure differential across the HWV instrument duct led to leaks. The low flow rates in the laboratory also increased the impact of hysteresis, i.e., the adsorption and desorption of water vapor from the walls of the instrument duct. At the highest pressures and faster flows the combined contribution from leaks and outgassing was minimized, amounting to less than ~0.05 ppmv. However, at lower pressures and/or slower flows these contributions exceeded ~1.0 ppmv, and became a significant fraction of the water vapor signal measured by HWV. We attempted to correct for these offsets using data from a limited set of diagnostic tests to tune a simple model. The large uncertainties associated with the correction, however, increased the uncertainty of the final measurement. The uncertainty in the determination of the leak/outgassing correction, which ranged from ~0.01 ppmv at the highest pressures and flows to ~1.0 ppmv at the lowest pressures and flows, is included in the estimated ~0.50 ppmv uncertainty attributed to artifacts and offsets. Additional uncertainties associated with the chopper reflectivity correction were

of the order of ~±0.20 ppmv, and are also included in the final uncertainty estimate.

Finally, HWV exhibited a constant positive measurement offset with respect to a selection of the core instruments. The behavior of this offset is not consistent with any of the sources of error previously described. Our estimate for the magnitude of this artifact, based on laboratory studies prior to the AquaVIT campaign, was between +0.0 and +0.3 ppmv. Results of comparisons during AquaVIT showed differences within ~0.35 ± 0.10 ppmv. Although this value is larger than our initial estimate, it is within the bounds of our total estimated uncertainty. There is also potential for negative biases in one or more of the comparison instruments. Finally, a positive offset of this magnitude, ~0.30 ppmv, cannot resolve the in situ measurement discrepancies of ~1.5 ppmv.

5.5 Thoughts for Future Laboratory Campaigns

The AquaVIT campaign provided a well-thought out, well-planned, and well-intentioned first step in addressing the accuracy of the in situ relative humidity data, and in resolving the enduring in situ measurement discrepancies, via a formal instrument comparison. However, there was little acknowledgement of the difference, for each instrument, between the laboratory setting at the AquaVIT facility and the in situ sampling environment, and no prescribed means of addressing how these differences might impact the measurements and the applicability of the comparison. While every in situ instrument must be capable of being tied to and tested against a laboratory standard, laboratory calibration and

validation is only a portion of the full flight performance assessment (see Chapter 3). The sources of systematic error for a given instrument may differ substantially between the laboratory and the atmosphere. Consequently, the applicability of an instrument's performance in the laboratory to the atmospheric environment may not be straightforward. Furthermore, agreement (or disagreement) among instruments in the laboratory may not accurately represent in-flight comparisons. In the case of HWV, for example, the instrument is subject to substantial artifacts in the laboratory from a combination of leaks into the duct and outgassing from the instrument walls. Neither of these artifacts is of concern in situ. However, in the laboratory, these issues, and the applied correction, introduce uncertainty in HWV measurements. Thus, future campaigns should make it a priority, 1) to provide a sampling environment that simulates flight conditions as closely as possible for all the participant instruments, 2) to understand the remaining sampling and measurement differences between the laboratory and in situ environments, and 3) to devise a methodology for quantifying the impact of these differences on the measurements, and on the comparison.

Comparisons over a wide range of mixing ratios and sampling conditions with a variety of instruments employing different detection techniques have the potential to reveal the sources of systematic error that lead to measurement differences. However, the formal nature of the AquaVIT campaign precluded the type of focused diagnostic tests that make possible the identification and quantification of errors. In order to make future laboratory campaigns more effective in addressing the systematic differences evident in the in situ

comparisons, it is necessary to first identify the potential error sources for each instrument, accounting for the differences between the atmospheric and laboratory environments, and to develop a campaign structure that permits an interactive approach. Future efforts should work toward a methodology that facilitates instrument diagnosis by addressing instrument differences and sensitivities in real-time.

References

Fahey, D. W., R. S. Gao, and O. Mohler (2009), Summary of the AquaVIT Water Vapor Intercomparison: Static Experiments.

Kretova, M. V., M. G. Khaplanov, and V. A. Yushkov (1991), Resonance Fluorescence Hygrometer for Humidity Measurement in the Stratosphere, *Instrum Exp Tech+*, *34*(3), 689-692.

Yushkov, V. A., N. M. Sitnikov, A. E. Ulanovskii, F. Ravegnani, and G. Redaelli (2001), Measurements of the ozone and water vapor contents in the stratospheric Antarctic cyclone from the high-altitude M55 Geophysica Aircraft, *Izv Atmos Ocean Phy+*, *37*(3), 275-280.

Zoger, M., et al. (1999), Fast in situ stratospheric hygrometers: A new family of balloon-borne and airborne Lyman alpha photofragment fluorescence hygrometers, *J Geophys Res-Atmos*, *104*(D1), 1807-1816.

Chapter 6

The Supersaturation Puzzle:

An In Situ Measurement Perspective

6.1 Introduction

We present in situ observations of upper tropospheric and lower stratospheric water vapor and relative humidity with respect to ice, RH_i. These data were acquired by the Harvard water vapor (HWV) instrument which flew aboard NASA's WB-57 high-altitude research aircraft during six missions sampling air from the tropics to midlatitudes. The tropical dataset consists of measurements made during two field campaigns flown out of San Jose, Costa Rica: the CWVCS mission, in August 2001, and the PreAVE, in January 2004; Sub-tropical data were obtained during the CRYSTAL-FACE mission in July 2002 on flights out of Key West, Florida; and midlatitude data were acquired during the MidCIX mission, and the AVE-0506 and AVE-WIIF missions flown out of Houston, TX, in April/May 2004 and June/July 2005 respectively. The full complement of data allows us to explore the behavior of relative humidity in the upper troposphere/lower stratosphere over a fifty-degree range in temperature, and three orders of magnitude in both water vapor and cloud ice water content.

During the six field campaigns examined here we frequently observed ice-supersaturation in clear air as well as in the presence of cirrus. Our clear air RH_i data at temperatures warmer than ~200 K are consistent with the established theory of homogeneous nucleation, however at temperatures representative of the tropical upper troposphere (<200 K), our observations challenge the accepted model for ice nucleation and cloud formation. Similarly, our in-cloud RH_i data at temperatures >200 K follow the accepted convergence of RH_i toward thermodynamic equilibrium. However, at the coldest temperatures, often within thin, near-tropopause cirrus, our data show evidence of prevalent and sustained supersaturation. If representative, our cold temperature clear air and in-cloud RH_i results may substantially alter the expected frequency and coverage of UT cirrus, and significantly diminish dehydration within the tropical tropopause layer (TTL), leading to higher than expected stratospheric entry-level water vapor mixing ratios, $[H_2O]_e$.

6.2 Relative Humidity

Relative humidity provides a framework for addressing the processes involved in the exchange of water between its gaseous and condensed phases in the atmosphere. Relative humidity with respect to ice, RH_i, is defined as the ratio of the ambient water vapor mixing ratio to the saturation mixing ratio with respect to ice, SMR_i, and is typically expressed as a percentage: $RH_i = VMR_{amb}/SMR_i \times 100$, with $SMR_i = fn(T,P)$. SMR_i is the thermodynamic function of state describing the water vapor mixing ratio in equilibrium with plane hexagonal ice at ambient

temperatures and pressures. The ratio, RH_i, can, therefore, be viewed as a measure of the departure from the thermodynamic equilibrium condition. As such, RH_i supplies the basic context for understanding the factors controlling the phase changes and phase distributions of atmospheric water. In the present analysis, in situ measurements of RH_i in clear air and within clouds provide the foundation for exploring the relationship between water vapor and cirrus in the UT/LS.

6.3 The Dataset

We obtain RH_i using ambient measurements of water vapor, temperature and pressure acquired in situ aboard NASA's WB-57 high-altitude research. In the present analysis we use data from the Harvard Water Vapor instrument (HWV), which utilizes the established Lyman-α photo-fragment fluorescence detection technique to make accurate and precise measurements of ambient water vapor [*Weinstock et al.*, 1994; *Weinstock et al.*, 2009], and depending upon the payload configuration, we use either the NASA Ames Meterological Measurement System (MMS) [*Scott et al.*, 1990], or the NOAA Pressure and Temperature (P/T) data for accurate temperature and pressure measurements. The complete RH_i dataset covers more than fifty degrees in temperature, three orders of magnitude in both water vapor and cloud ice water content, and a geographic range from the midlatitudes to the tropics. Table 6.1 lists the full complement of missions, including date and location.

Mission	Acronym	Location	Dates
Clouds and Water Vapor in the Climate System	CWVCS	San Jose, Costa Rica	Aug. 2001
Cirrus Regional Study of Tropical Anvils and Cirrus Layers – Florida Area Convection Experiment	CRYSTAL-FACE (CF)	Key West, FL	Jul. 2002
Pre AURA Validation Experiment	PreAVE	San Jose, Costa Rica	Jan. 2004
Midlatitude Cirrus Cloud Experiment	MidCIX	Houston, TX	Apr./May 2004
Aura Validation Experiment	AVE-0506	Houston, TX	Jun. 2005
Water vapor Isotope Inter-comparison Flights	AVE-WIIF	Houston, TX	Jul. 2005

Table 6.1. NASA WB-57 missions during which we used HWV to derive relative humidity data. The locations and dates of each mission are noted.

6.4 RH_i Data Accuracy

Before we begin a thorough dissection of the relative humidity dataset, we provide a brief, but necessary, discussion of the accuracy of the observations. An accuracy of at least ±10% is desirable in RH_i. Uncertainties larger than this begin to impact the strength of our conclusions. Thus, we dedicate the next section to a discussion of the accuracy and the uncertainties of the independent in situ water vapor and temperature measurements.

The accuracy of RH_i depends upon the accuracy of its principal components, i.e., the measurements of ambient water vapor and temperature, as well as the validity of the equation used to calculate SMR_i.

We calculate SMR_i using Equation 7 of [*Murphy and Koop*, 2005], which describes the partial pressure of water vapor over plane hexagonal ice as a function of temperature. Their parameterization, derived from the Clapeyron

equation and constrained by measurements of the vapor pressure at the triple-point, is both efficient and applicable over a wide temperature range, from the triple-point to ~110 K. They estimate the uncertainty in their parameterization is less than ±0.5% in calculated vapor pressure over plane hexagonal ice at 200 K.

The Clasius-Clapeyron equation shows that the vapor pressure of water over ice increases nearly exponentially with temperature. Therefore, measurements of both SMR_i and RH_i are highly sensitive to uncertainties in measured temperature. Under conditions representative of the UT, e.g., a temperature of 200 K and pressure of 115 mbar, an error of 0.5 K in the measured temperature (0.25%), yields an ~8% difference in derived RH_i. At temperatures <200 K, the sensitivity to temperature uncertainties is even greater. In contrast, the sensitivity of RH_i to errors in the measurement of water are directly proportional, i.e., a 5% error in H_2O yields a 5% error in derived RH_i.

In the present analysis, the combination of the stated uncertainties of the independent water vapor and temperature measurements, ~±5% and ~±0.5 K respectively, yields an uncertainty of ~±10% in measured RH_i. Additional sources of uncertainty associated with the P/T temperature measurements and aircraft roll angle [*Murphy*, 2005], and with the in-cloud water vapor measurements are discussed in more detail below.

6.4.1 Accuracy of In Situ Water Vapor Measurements: Clear Air

Table 2.1 summarizes the expected in-flight performance of HWV. For a thorough discussion of the calibration and validation of HWV, and the evaluation of each of the terms in Table 2.1 refer to Chapters 2, 3 and 4.

6.4.2 Accuracy of In Situ Water Vapor Measurements: In Cloud

For a detailed discussion of the accuracy of HWV in clouds refer to section 3.4.3 in Chapter 3.

6.4.3 Accuracy of Airborne Temperature Measurements

Accurate, high resolution in situ measurements of temperature and pressure necessitate instruments which have undergone a rigorous process of calibration and validation analogous to that of HWV, i.e., instruments that are calibrated to a laboratory standard, carefully integrated onto the aircraft platform, and verified under the unique conditions of flight. During the suite of missions listed in Table 6.1 the WB-57 aircraft was equipped with one or both of the following instruments for measuring temperature and pressure, 1) the Meteorological Measurement System (MMS) instrument and, 2) the NOAA Pressure-Temperature (P/T) instrument. The following are common to both the MMS and P/T instruments: 1) temperatures are measured with Rosemount platinum resistance temperature detectors (RTD), which are calibrated at the NIST-certified facility at Rosemount, Inc., and exhibit long-term stability, 2) the temperature and pressure probes are positioned near the nose of the WB-57 and

are strategically located around the fuselage to sample out of the aircraft boundary layer and avoid interference from other probes, and 3) the raw temperature measurements are corrected for the effects of compression heating and airflow through the probe housing during flight. Furthermore, the stated accuracies for both instruments are comparable, with the pressure measurements accurate to within 0.1 mbar, and temperatures accurate to within 0.5 K or better, under atmospheric conditions. Table 6.2 summarizes the performance characteristics of both sets of probes.

Temperature & Pressure	**Typical Value**	**Precision**	**Accuracy**
MMS @ ~1 Hz			
Pressure	~60 hPa	±0.02 hPa	±0.1 hPa, ~0.2%
Temperature	~180 K	±0.02 K	±0.3 K, ~0.2%
NOAA P/T @ 1 Hz			
Pressure		±0.03 hPa	±0.1 hPa
Temperature		±0.03 K	±0.5 K

Table 6.2. Summary of the accuracy and precision for each of the two pressure and temperature instruments aboard the NASA WB-57 aircraft: 1) the NASA Ames Meteorological Measurement System (MMS); and 2) the NOAA Pressure & Temperature instrument (P/T).

A complete description of the MMS probe, including its calibration and validation, is provided in the paper by [*Scott et al.*, 1990]. To summarize, the MMS validation process entails repeated laboratory calibration, followed by in-flight validation and performance assessment, and concludes with comparison with other established measurement methods. The laboratory work involves not only the calibration of the individual sensors – the probes are sent back to Rosemount for calibration after each major deployment – but also necessitates

calibrating the sensors within their specific MMS configuration. In the case of the temperature probes, for example, the "recovery correction," which accounts for air flow through the probe during flight, is confirmed by wind tunnel testing, and by consistency checks in flight utilizing wind field measurements [T.P. Bui, *personal communication*]. Finally, [*Gaines et al.*, 1992], published an analysis comparing MMS measurements aboard the NASA ER-2 aircraft with radiosonde temperatures, and concluded that the comparisons showed agreement to within the expected measurement accuracy and the natural variability of the meteorological parameters.

The CRYSTAL-FACE mission provided an opportunity for a direct comparison of MMS and the NOAA P/T temperature data. The data demonstrate good agreement between the two instruments, well within their combined uncertainty (±0.8 K) over a broad temperature range[1]. Furthermore, the results of the comparison do not show any detectable sensitivity to the presence of clouds. Within the UT, the region of interest here, the agreement is within ~0.2 ±0.2 K, with P/T measuring systematically lower and yielding RH_i values ~3% higher on average.

In the present analysis, we preferentially use MMS temperatures because of the documented calibration and validation procedure and the long flight history of the MMS probes.

[1] MMS data prior to the flight of July 9, 2002, were excluded, because the primary temperature sensor was destroyed by a lightning strike on the flight of July 7, 2002, and its calibration could not be verified after the mission.

6.5 Clear Air / In-Cloud Discrimination

Distinguishing clear air from cloudy air is also critical to the RH_i analysis, as we expect different factors to govern the relative humidity distributions and relative humidity extrema in clear and cloudy environments. We devote the following section to a discussion of the instruments used for clear air/in-cloud discrimination, the detection thresholds and errors associated with the different instruments, and how the final selection criteria were defined.

Because of variability in the WB-57 payload configuration, the instrument(s) used for cloud detection, and thus, the criteria for discriminating clear from cloudy air changed from mission to mission. Table 6.3 lists each mission and the suite of instruments considered for cloud detection and characterization. During the CRYSTAL-FACE, MidCIX and AVE-WIIF missions the plane was equipped with both optical particle counters and direct cloud ice water content instruments. During the two tropical missions (CWVCS and PreAVE), however, payload and/or instrument limitations forced us to rely solely upon the Harvard ice water content measurements, and during the AVE-0506 mission out of Houston, TX, the WB-57 did not carry any instruments capable of direct cloud detection. Even with a full suite of cloud sensors, however, the identification of clear and cloudy air is often difficult, particularly at cloud boundaries or in regions of patchy, thin, sub-visible cirrus, because of instrument detection thresholds, response times, and data rates.

Mission	Ice Water Content	Particle Instruments	Clear Air	In Cloud
CWVCS	HTW	*N/A*	IWC < 0.5 ppmv	IWC ≥ 2
CF	HTW	CAPS, SPP-100	SPP < 1 L^{-1}	SPP ≥ 3
PreAVE	HTW	*N/A*	IWC < 0.5 ppmv	IWC ≥ 1
MidCIX	HTW	CAPS, SPP-100	SPP < 1 L^{-1}	SPP ≥ 3
AVE-0506	*N/A*	*N/A*	*N/A*	*N/A*
AVE-WIIF	HTW	CAPS	CAPS < 3 $\mu m^2/cm^3$	CAPS ≥ 10

Table 6.3. Cloud property instruments and criteria in each mission used to differentiate clear and cloudy air. Instrument acronyms and descriptions and are given in the text.

Cloud Particle Instrumentation

There are two primary types of instrument designed to sample and study cloud properties in situ, those that measure the bulk ice water content of clouds, and those that probe the physical properties (size, shape and number) of the cloud aerosol and ice crystals. In brief, the total water instruments like the Harvard Total Water instrument (HTW) [*Weinstock et al.*, 2006a; *Weinstock et al.*, 2006b], measure the total water content of an airmass, i.e., water in both its vapor and condensed phases, by ensuring the complete evaporation of the condensed phase fraction by the time the air sample reaches the detection region of the instrument. HTW is designed to measure isokinetically, i.e., the airmass velocity at the instrument inlet is equal to the free stream velocity. As such, ambient particle number densities are unperturbed in the sample. No additional processing is required, and the departure between the results of the total water measurements and those of an independent vapor instrument, e.g., HWV, indicate the presence of a cloud, with the difference equal to the cloud ice water content.

In contrast, the optical instruments are sensitive to the presence of aerosol particles or cloud crystals via the scattering of laser light or direct imaging. The SPP-100/FSSP-100 from Stratton Park Engineering Company, Inc., (SPEC), is a Forward Scattering Spectrometer Probe (FSSP) with a Droplet Measurement Technologies (DMT) Signal Processing Package (SPP), which generates size distributions by collecting and analyzing light scattered in the forward direction by particles passing through a focused laser beam. This instrument has demonstrated high sensitivity to particle detection, for particles from 2 to ~60 μm in diameter, and has an operational detection threshold of 10 particles/L at 1 Hz (effectively equivalent to ~3 particle/L for 10-second average data). As a result, we preferentially use the SPP-100 total number density data for clear air/in-cloud discrimination in the present analysis. We also use the SPP-100 data to assess the cloud detection performance of the other instruments.

The DMT Cloud, Aerosol and Precipitation Spectrometer, CAPS, instrument characterizes aerosol particles and small to medium-sized cloud crystals ranging from 0.5 to 1600 μm in diameter [*Baumgardner et al.*, 2001]. CAPS is comprised of two optical probes: 1) the cloud aerosol spectrometer (CAS) which collects light scattered in the forward and backward directions, and retrieves particle size and shape information in the range from 0.5 to 50 μm; and 2) the cloud imaging probe (CIP) which records the images of particles that pass through its collimated laser beam. These images yield particle size and habit data with a resolution of 25 μm for particle diameters from 50 to 1600 μm. Note that we use the CAPS total surface area data, a derived data product, because CAPS

does not provide merged particle number density data. CAPS records number densities only for individualized particle size bins, which, while useful for generating size distributions, is not effective for cloud discrimination.

Analysis of Particle Probe Data

The CRYSTAL-FACE and MidCIX missions had payload configurations that provided the opportunity to compare SPP-100 particle number density data with both the CAPS surface area density data, and the cloud IWC measurements made by the pair of Harvard instruments. We use the results of these comparisons to explore and define the cloud discrimination criteria for each mission.

The agreement between the 10 Hz particle probe data from MidCIX is shown in Figure 6.1a. CAPS total particle surface area density data ($\mu m^2/cm^3$) is plotted as a function of total particle number density data from the SPP-100 (#/L). An examination of the plot reveals four primary conclusions about the performance of these two particle instruments: 1) the correlation between the data sets is generally good over nearly four orders of magnitude; 2) the detection threshold for the SPP-100 data is effectively ~1 L^{-1} – these are the lowest values reported by the SPP-100 in both the MidCIX and CRYSTAL-FACE data sets; 3) the SPP-100 reports zeros for nominally clear air data; and 4) the CAPS total surface area density data appear to have a positive offset on the order of 1 – 10 $\mu m^2/cm^3$.

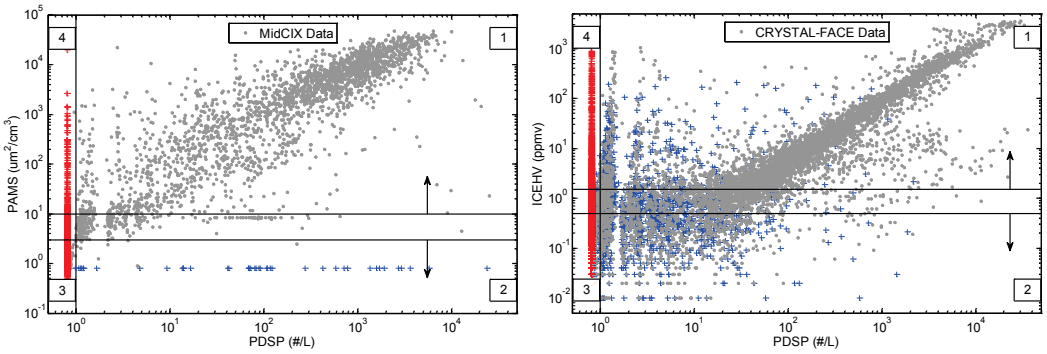


Figure 6.1. Comparison of in situ cloud property measurements. Panel (a) shows CAPS particle surface area density data ($\mu m^2/cm^3$) plotted as a function of SPP-100 particle number densities (#/L) from the MidCIX mission. Panel (b) shows cloud ice water content (ppmv) as determined from the pair of Harvard hygrometers plotted as a function SPP-100 data (#/L) from the CRYSTAL-FACE mission. The detection thresholds used for distinguishing clear air from cloud are shown in the solid black lines. The colored crosses in each figure and the numbered "Zones" are described in the text.

Both CAPS and SPP-100 report zero in nominally clear air, however, the proportion of zero data differs greatly between the two data sets. According to the SPP-100 data alone, ~50% of the UT/LS airmasses sampled by the WB-57 during MidCIX were cloud free. In contrast, simultaneous measurements by CAPS recorded "cloud free data," i.e. 0 $\mu m^2/cm^3$, only 5% of the time. CAPS frequently (95%) records non-zero surface area density when the SPP-100 number densities equal zero. In order to be shown on the log/log scale in Figure 6.1a, the SPP-100 zero data (red crosses) and CAPS zero data (blue crosses) were set equal to 0.8. A cursory examination of the relative humidity data (not shown) reveals that >95% of these non-zero CAPS data occur in sub-saturated air, supporting the SPP-100 classification and suggesting that the CAPS data have a small positive offset. Further inspection of the CAPS measurements, when the SPP-100 data are equal zero, shows that ~90% of the CAPS values are ~<10 $\mu m^2/cm^3$, i.e., much of the

CAPS data measuring <10 $\mu m^2/cm^3$ is falsely indicative of cloud. Comparison of the CAPS data with the Harvard IWC data (not shown) also supports this conclusion.

We evaluate the CAPS clear air/in-cloud criteria with SPP-100 data representing "truth". The difficulty of airmass discrimination is captured in Figure 6.1a, which is subdivided into four Zones. Zone 1, where CAPS measures values ≥ 10 $\mu m^2/cm^3$ and SPP-100 measures >3 particle/L, delineates the Zone where cloudy air is identified with confidence. 95% of the data identified by CAPS as being in cloud (≥ 10 $\mu m^2/cm^3$) are confirmed by the SPP-100. Similarly, Zone 3 where the SPP-100 measures < 1 particle/L, and where CAPS measures <3 $\mu m^2/cm^3$, delineates data where we are very likely in clear air. Zones 2 and 4, in contrast, denote regions of ambiguity, where sole reliance upon the CAPS criteria leads to a high probability of erroneous identification. In general, with CAPS as the only metric, the criteria shown in Figure 6.1 and summarized in Table 6.3 lead to ~5% of the airmasses being misidentified. Furthermore, in the absence of other cloud data, the data in the "inter-zone" region, i.e., CAPS surface area densities >10 $\mu m^2/cm^3$ and <3 $\mu m^2/cm^3$, are eliminated from the analysis because of the high probability of misidentification. Unfortunately, this removes much of the data in thin cirrus and near cloud boundaries.

The Harvard IWC measurements are plotted as a function of the SPP-100 data in Figure 6.1b. Again, these data show that the correlation between the datasets is good over four orders of magnitude in both ice particle number density and cloud ice water content. As in Figure 6.1a, the SPP-100 zero data (red

crosses) were set equal to 0.8 L^{-1}, and the absolute value of the negative IWC data were plotted in blue crosses. The IWC data are noisy at and below ~1.0 ppmv, or equivalently, where particle number densities fall below ~10 L^{-1}. This essentially defines the detection threshold of the IWC content data. The distribution of IWC data, selected for SPP-100 <10 particles/L, is centered at zero with >80% of the data between ±1.0 ppmv.

The combination of HTW and HWV yields cloud ice water content, providing a direct measure of this quantity in ppmv. The efficacy of this measurement as an airmass discriminator, however, is poor relative to the SPP-100 particle counter. The Harvard instrument pair has difficulty distinguishing between the thinnest clouds and clear air, especially for clouds with ice water contents below about ~1.0 ppmv. Uncertainties due to hysteresis in the HTW instrument, calibration uncertainties between the two instruments, and limitations in instrument precision render the detection of clouds with IWC <0.5 ppmv unrealistic. As with the CAPS measurements, the risk of misidentifying airmasses operates in both directions as the detection limit of the instrument is approached. We may erroneously identify truly cloudy air as clear, or vice versa.

Identifying cloudy air with certainty is the easier of the two problems because the detection criteria may be set well above the detection threshold of the measurement, i.e., we can confidently identify air with ≥10 ppmv IWC as being in cloud. The difficulty lies, therefore, in the fact that much of in-cloud data are excluded. However, if we set the limit for cloud detection too low, i.e., ≥0.5 ppmv, we run the risk of erroneously including a large fraction of clear air data. In

the present analysis we choose a low end cutoff of between 1.0 and 2.0 ppmv (mission dependent) for the identification cloudy air.

Similarly, identifying air as clear which in reality has a detectable (though small) cloud ice water content, is also likely when relying solely upon the Harvard instrument pair. Given an upper limit IWC of 0.5 ppmv, the risk of misidentifying truly cloudy air as clear air is ~15%. However, reducing the clear air upper limit by a factor of four simply reduces number of clear air data points, without appreciably improving the fraction of air that is incorrectly identified. Thus, relying on Harvard IWC as the sole cloud detection method can lead to the misidentification of airmasses as much as ~15% of the time.

The optimal clear air/in-cloud selection criteria are given in Table 6.3. As discussed, we chose to use the SPP-100 particle number density when available because of its high sensitivity to cloud detection and reliable performance mission-to-mission. However, during the CWVCS and PreAVE missions, due to payload constraints or poor instrument performance, we were forced to rely on the IWC measurements provided by the combination of HTW and HWV, and during AVE-WIIF we utilized the particle surface area density measurements of the CAPS instrument.

The inclusion of truly clear air in the cloudy air dataset may have the effect of expanding both tails of the in-cloud RH_i distribution, as clear air data are more likely to be both under-saturated and highly super-saturated than in-cloud data. Indeed this effect is enhanced because this error is more likely to occur at cloud boundaries. However, it's unlikely that these misidentifications will alter in

any significant way the mean in-cloud results, or even the fundamental shape of the in-cloud RH_i distribution. Misidentifying truly cloudy air as clear, i.e., erroneously including in-cloud data in the clear air dataset, is even less likely to impact the conclusions. The most important aspect of the clear air distribution is the supersaturated tail, as the maximum clear air RH_i provides information about cloud nucleation thresholds. Basic physical arguments essentially require in-cloud RH_i to be equal to or less than the maximum clear air value, thus the presence of a few in-cloud data points will not impact the determination of the RH_i maximum.

6.6 RH_i Results Overview

The combined dataset provides a unique opportunity to compare and contrast upper tropospheric relative humidity measurements obtained over a range of climatological conditions. In the following analysis and discussion we make use of both ensemble data as well as specific high resolution case studies to investigate the behavior and implications of the RH_i observations.

We sampled air in several distinct temporal and geographic settings. The data span a geographic range from the tropics to the midlatitudes, and were acquired during three of the four seasons. Specifically, the CWVCS data, acquired on flights out of San Jose, Costa Rica, provide numerous vertical profiles through the summertime tropical upper troposphere, and tropical tropopause layer (TTL). The PreAVE data, also acquired on flights out of San Jose, Costa Rica, provide comparable profiles through the TTL during the drier, colder, and comparatively quiescent winter season. Data obtained during the CRYSTAL-FACE mission are

representative of the sub-tropical upper troposphere during summer. The sampling region during CRYSTAL-FACE was centered over southern Florida and was subject to the influence of frequent local summertime convection. Finally, during the MidCIX and AVE-0506/AVE-WIIF missions, staged out of Houston, Texas, the aircraft sampled the midlatitude UT/LS in the spring and summer seasons respectively. During the latter mission, the aircraft frequently encountered air in the lower stratosphere bearing the signature of summertime deep continental convection over the Midwestern United States [*Hanisco et al.*, 2007].

Figure 6.2 provides the context for the relative humidity datasets during each of the six missions. The panels on the left show vertical profiles of temperature, the middle panels show the corresponding vertical profiles of the ambient water vapor mixing ratio, and those on the right show the resulting RH_i. Potential temperature is chosen as the vertical coordinate. Supersaturated clear air data, i.e., where RH_i>100%, are plotted in red. In-cloud data are plotted in blue.

This mission climatology figure reveals three features that are more fully developed in subsequent sections. The first feature, evident in every data set, is that air may be supersaturated without cloud formation. This observation is consistent with a growing body of in situ evidence for the existence of prevalent and sustained regions of near tropopause supersaturation in clear air [*Gierens and Spichtinger*, 2000; *Gierens et al.*, 1999; *Heymsfield et al.*, 1998; *Jensen et al.*, 2001a; *Jensen et al.*, 2001b; *Kramer et al.*, 2009; *Ovarlez et al.*, 2002]. The current data set extends the geographic and thermal range of these observations. These ice-supersaturated regions (ISSRs) occur up to, but rarely above the local

tropopause, which serves as a nominal upper bound for the ice-supersaturation as well as cirrus, consistent with the results of [*Gierens et al.*, 1999], and [*Smith et al.*, 2001]. We will expand on the significance of these clear air ISSRs, in particular, what they reveal about modes of ice nucleation and cirrus formation, in greater detail below.

The second shared feature is that the supersaturated airmasses (red dots) occur preferentially at the lowest temperatures, though they appear to span the full range of observed water vapor mixing ratios. The distribution of ice-supersaturated data is skewed toward the coldest temperatures particularly in the vicinity of the local tropopause. The comparatively warmer tropopause regions are undersaturated. This is not unexpected. A number of studies have shown that much of the structure and variability in clear air RH_i in the UT is preferentially driven by variability in ambient temperature [*Gierens et al.*, 1999; *Karcher and Haag*, 2004]. We explore the role of temperature in driving the observed clear air RH_i distribution in subsequent sections.

The third feature is that much of the data acquired in the presence of clouds (blue dots) also show significant supersaturation. The RH_i distributions in clouds are nominally symmetric in the sub-tropical and midlatitude datasets, the lower three rows of the figure, with nearly equal fractions of the in-cloud data occurring in both saturated and unsaturated airmasses. However, in the two tropical datasets, the in-cloud RH_i data are biased toward a much higher frequency of substantial supersaturation. These results and their implications will also be explored more thoroughly in subsequent sections.

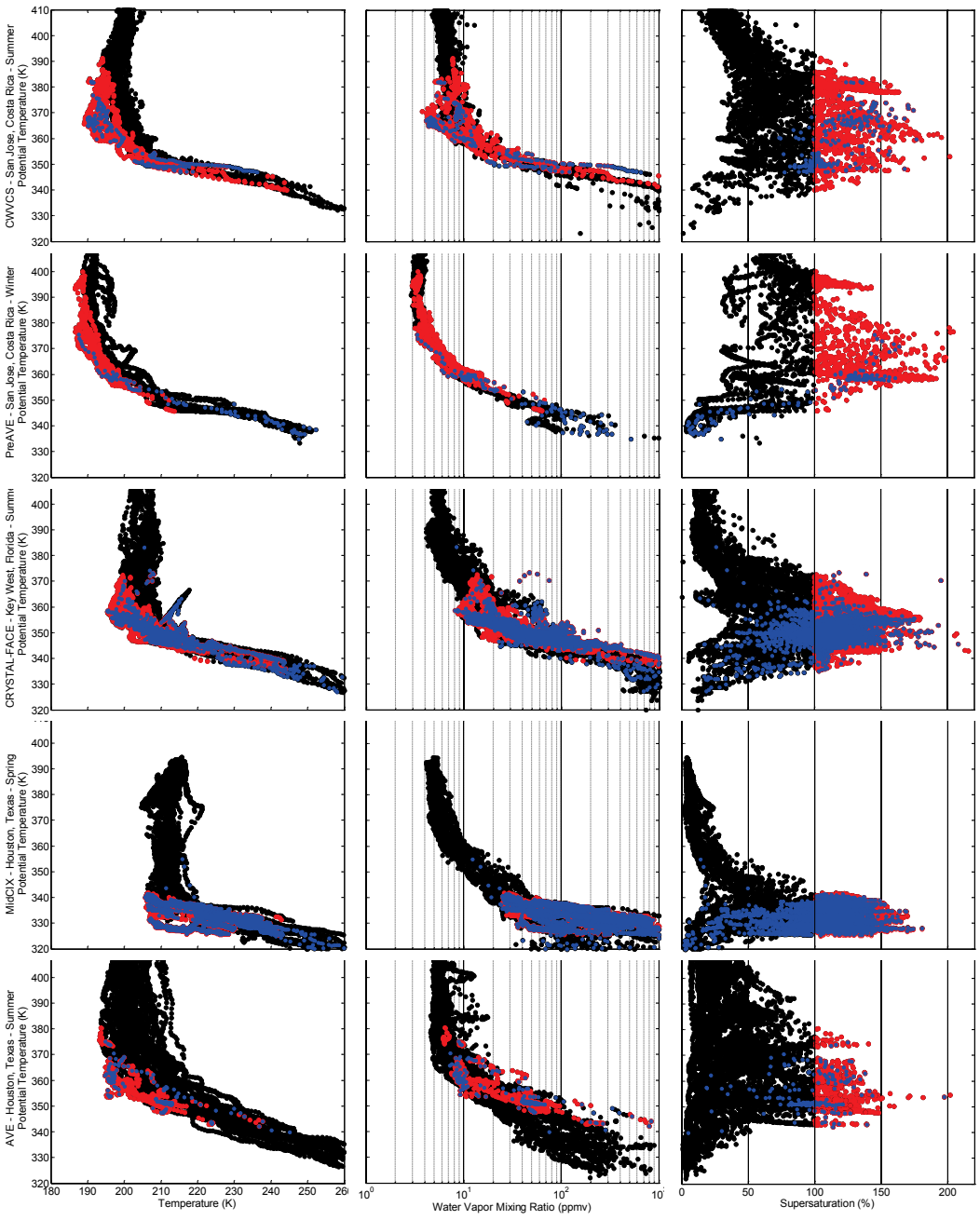


Figure 6.2. Temperature, water vapor, and humidity profiles for each mission. The red dots indicate supersaturated air, and the blue dots indicate cloud presence.

Figure 6.3 offers a more quantitative portrayal of the features qualitatively evident in Figure 6.2. In this figure, the clear air (red) and in-cloud (blue)

empirical distribution functions of *RHi* data from the CRYSTAL-FACE mission are shown.

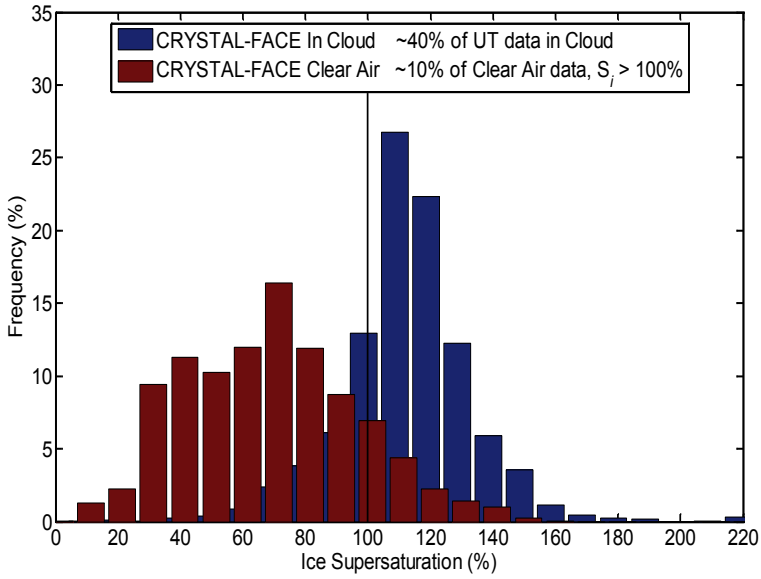


Figure 6.3. Empirical distribution functions of RH_i in % for clear air (red) and in-cloud data (blue) from the CRYSTAL-FACE mission. Data were selected to be in the UT/LS, i.e., between 325 and 405 K.

The data in Figure 6.3 were selected to be within a range covering the local upper troposphere and lowermost stratosphere (between ~325 K and 405 K). These selection criteria effectively delineate the zone of upper tropospheric and near tropospause cirrus cloud occurrence for CRYSTAL-FACE, and include cirrus resulting from convective outflow as well as those formed in situ. Accordingly, the in-cloud RH_i data in Figure 6.3 capture nearly 100% of the clouds sampled during the CRYSTAL-FACE mission, excluding only warm wet clouds encountered on ascent, and the clear air data are representative of the 'cloud-free' condition in the same region. The CRYSTAL-FACE data again show that, 1) regions of supersaturated clear air were not uncommon (~10% frequency

of occurrence), 2) air with a substantial degree of supersaturation (~160%) can occur without cloud formation, and 3) a broad range of RH_i, with substantial supersaturation, is observed within clouds.

6.7 Clear Air RH_i

6.7.1 Clear Air RH_i: Ice Supersaturated Regions (ISSRs)

As mentioned above, the existence of ISSRs in the UT is now generally accepted. However, the frequency of occurrence of ISSRs is driven by local meteorological patterns and is highly dependent upon geographic region and season. The data from the six missions evaluated here reveal some of this variability.

The effect of season and geographic location on the vertical temperature and humidity structure for each mission is visually evident in Figure 6.2, and Table 6.4 provides a quantitative summary of ISSR frequency by mission. Note, we calculated the ISSR frequency independently for each mission in the region bounded by the upper and lower potential temperature levels at which clear air supersaturation was observed. We chose these bounds to eliminate data from the largely undersaturated lower stratosphere, which would simply lower the percentages in all cases. The expansion and contraction of the ISSR zone from one geographic region and season to the next is evident.

Table 6.4 also records the zone in which clouds were observed for each mission, and the frequency of cirrus occurrence in a fixed region between 325 K and 405 K. The frequency of cloud encounters is determined for this more broadly

defined UT/LS region, because it is important to include airmasses that are sub-saturated as well as supersaturated. It is worth noting that in the tropical datasets the clear air supersaturations extend to higher theta levels than the upper limits of observed cloud occurrence.

Mission	Season Region	Zone of Clear Air $RH_i > 100\%$	Clear Air ISSR Frequency	Zone of Cloud Occurrence	Cloud Frequency in UT/LS
CWVCS	Summer Tropics	350 – 390 K	~35%	345 – 380 K	~5%
PreAVE	Winter Tropics	345 – 400 K	~35%	335 – 375 K	~3%
CF	Summer Sub-trop.	340 – 370 K	~15%	335 – 370 K	~25%
MidCIX	Spring Midlat.	325 – 340 K	~15%	325 – 340 K	~50%
AVE-0506	Summer Midlat.	350 – 380 K	~5%	*N/A*	*N/A*
AVE-WIIF	Summer Midlat.	345 – 375 K	~2%	340 – 375 K	~15%

Table 6.4. The region and season of each mission, the upper and lower limits to observations of clear air ice-supersaturation, and the upper and lower limits to cloud observations. The frequency of ISSRs is determined as a percentage of supersaturated clear iar data to total clear air data in the ISSR zone for each mission. The frequency of cloud observations is determined as a percentage of the cloudy data to total data in the region between 325 K and 405 K for each mission.

Tropical ISSRs and $[H_2O]_e$

In both tropical datasets the frequency of clear air ice-supersaturations substantially exceeds that of the CRYSTAL-FACE mission. Approximately 35% of the clear air data are supersaturated during the summertime CWVCS mission, and a similar fraction are supersaturated during the wintertime PreAVE mission. Inspection of Figure 6.2 shows that despite the noticeably lower water vapor mixing ratios during PreAVE, which are generally consistent with the winter-

phase of the seasonal cycle of water vapor irreversibly entering the stratosphere, $[H_2O]_e$, the colder wintertime UT temperatures manage to sustain high RH_i throughout the region and preserve the probability of encountering supersaturated air. Furthermore, in both missions, air with RH_i >100% is present throughout much of the TTL [*Fueglistaler et al.*, 2009], extending from a lower boundary near 350 K, which is just above the level of main convective outflow, through the level of the local tropopause (~380 K), and up to ~400 K.

The presence and frequency of these clear air supersaturations, particularly in the tropical datasets, has the potential to impact our understanding of the dehydration efficiency of tropical cold-point temperatures, and thus the strict thermal control of $[H_2O]_e$, e.g., [*Jensen and Pfister*, 2005; *Jensen et al.*, 2001a; *Weinstock et al.*, 2009], as well as estimates of outgoing long-wave radiative fluxes at the top of the atmosphere, e.g., [*Jensen et al.*, 2001b]. [*Jensen et al.*, 2001a] used a cloud model to explore the effect of ISSRs as well as inefficient dehydration by optically thin, laminar cirrus occurring near the tropical tropopause on $[H_2O]_e$. Air crossing the tropical tropopause through ISSRs, i.e., in the absence of ice nucleation and cirrus formation, can lead to significantly elevated values of $[H_2O]_e$, relative to a model that simply assumes the mixing ratio is equal to the SMR_i set by the local cold-point temperature (RH_i = 100%). To fully quantify the impact of ISSRs on $[H_2O]_e$, however, their prevalence and extent, as well as their spatial overlap with regions of irreversible transport into the tropical stratosphere, and degree of supersaturation, needs to be known. [*Jensen et al.*, 2001a] admit that their analysis is speculative because of 1) the

scarcity and 2) the level of uncertainty of water vapor measurements in the tropical tropopause region.

Extra-tropical ISSRs and Contrail-seeded Cirrus

Though the frequent observation of ISSRs (~15%) observed during the extra-tropical CRYSTAL-FACE and MidCIX missions are not expected to be important in the context of determining $[H_2O]_e$, they may still be significant from a local chemical and radiative perspective. In particular, these regions of clear air supersaturation can extend the lifetime of convective cirrus, and also promote the development of cirrus seeded by aircraft contrails [*Immler et al.*, 2007; *Jensen et al.*, 1998].

The flight of May 3, 2004, from the MidCIX mission, provides a spectacular example of mesoscale meteorology producing the conditions – a higher and consequently cooler tropopause over north-central Oklahoma – for anthropogenically seeded cirrus. Relative to flights executed in the same region two weeks earlier, the tropopause on this flight was elevated by ~200 m, and was ~1 K cooler. During this flight, the WB-57 sampled thin, wispy, cirrus along a nominally constant altitude leg within about 50 m of the tropopause (~13 km) over the Texas and Oklahoma panhandles. The pilot log at the time of the encounter states, "There was a good mix of visible and sub-visible cirrus on this leg."

Figure 6.4a locates the WB-57 flight track in the larger geographical context, and the time of interest here corresponds to the final minutes of the green

segment. Figure 6.4b shows NOAA-17 satellite images of the same region, bracketing the WB-57 transect in time. These images show that the cirrus sampled on this flight leg were almost certainly associated with aircraft contrails. Finally, Figure 6.4c shows the results of the in situ instrumentation as a function of flight time.

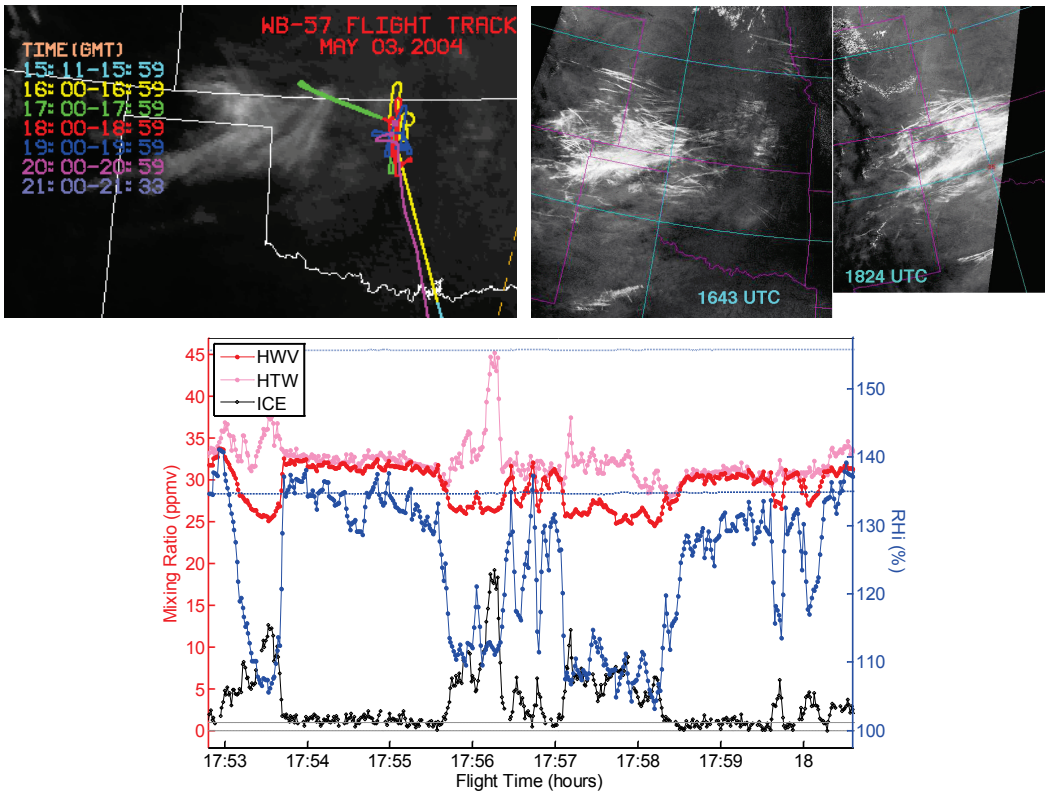


Figure 6.4. Panel (a) shows the WB-57 flight track on the flight of May 3, 2004, color-coded by time. The in situ data correspond to the green portion intersecting the cirrus over the border between Oklahoma-Kansas border. Panel (b) shows satellite imagery over this same region bracketing the time of the WB-57 cirrus encounter. Panel (c) shows the in situ water vapor (red), total water (pink), IWC (black), and RH_i (blue) data. Note the rapid transitions from supersaturated clear air, to near equilibrium within the cloud, at the cloud boundaries. The dotted blue line at 157% RH_i indicates the threshold required for homogeneous nucleation in this airmass. The dashed blue line at ~135% is the level found to be sufficient for heterogeneous nucleation in midlatitude data analyzed by [*Heymsfield et al.*, 1998].

In Figure 6.4c, the observed 1 Hz relative humidity data are plotted in blue, the ambient total water and water vapor mixing ratios are in pink and red respectively, and the derived ice water product, which is simply equal to the difference between the Harvard instrument pair is in black. Clear air corresponds to areas in which HTW and HWV are effectively equal, i.e., the measured IWC is ~<1.0 ppmv. These clear air regions were verified with particle counter data (not shown). The cirrus sampled during this leg of the flight shows the characteristic signature of having been formed in situ, rather than having originated from convective outflow. Total water is nearly constant, with the exception of a few excursions, indicating that most of the cloud water content results from the scavenging of ambient water vapor. Furthermore, the active scavenging of vapor by the cloud creates the abrupt transition in relative humidity evident at the edge of each cloud pass, from supersaturated clear air to near equilibrium in the sampled cirrus.

The in situ RH_i measurements, shown in Figure 6.4c, indicate broad clear air regions in which the relative humidity is ~135% ± 5%. These sustained ISSRs likely result from the fact that the ambient RH_i is not high enough for homogeneous ice nucleation. In addition, we conclude that there are no efficient heterogeneous condensation nuclei leading to ice nucleation at these RH_i levels. Figure 6.4b reveals, however, that the narrow bands that have been perturbed by the passage of jet aircraft show persistent contrail formation, with the steady evolution (over at least 2 hours) of in situ cirrus. In this case, the aircraft probably contributed a negligible net amount of water to the ambient air, as indicated by

the relatively constant total water values. However, the contribution of particulate matter from the aircraft exhaust clearly initiated ice nucleation, contrail formation and eventually led to the cirrus we sampled. This series of figures provides a simple introduction to the interplay between the natural variability of the atmosphere and anthropogenic activity. Unfortunately, the aircraft was not equipped with instrumentation to explore the more complex, but no less important, chemical and radiative effects associated with this type of event.

Finally, we contrast the ISSR frequency observed during MidCIX (~15%) with that observed during the AVE missions (~<5%), which sampled the same broad geographic region. The low frequency of supersaturated clear air data from the AVE missions is initially surprising, since the seasonal cycle of water vapor in the tropics is at a maximum during the summer, and this is also the season of intense midlatitude continental convection. Indeed, the water vapor profiles in Figure 6.2 reveal that the summertime midlatitude clear air data, acquired during the AVE missions, have mixing ratios ~20% to ~100% greater than the comparable springtime values measured during MidCIX throughout the UT/LS. Furthermore, there is strong evidence of moist convective plumes in the summertime data reaching the UT, and extending well above the local tropopause level (~360 K) into the LS. Figure 6.2 reveals, however, that the temperature structure of the upper troposphere is the dominant factor driving ISSR frequency, not water vapor mixing ratio. The temperatures of the UT, from ~320K to 350 K, are ~20 ± 10°C cooler during MidCIX than during AVE, leading to a significantly higher frequency of ISSRs during the spring season. The leverage of ambient

temperatures upon the humidity structure of the UT/LS is explored further in subsequent sections.

6.7.2 Clear Air RH_i: Ice Nucleation Mechanisms

In this section, we follow the example of [*Heymsfield et al.*, 1998], and [*Haag et al.*, 2003], and examine the distribution of clear air ice-saturation for each mission in order to extract information about the mechanisms of ice nucleation in the atmosphere. Specifically, we focus on observations of the maximum clear air supersaturation, seeking information about the relative humidity required to nucleate ice, RH_{nuc}, in situ. We begin with a brief summary of the two modes of ice nucleation, and then describe how we can identify these modes in the present dataset.

There are two distinct modes of ice particle nucleation leading to the formation of cirrus clouds in the upper troposphere, e.g., [*DeMott et al.*, 2003a], *and references therein*. Homogeneous nucleation refers to the autonomous freezing of the ubiquitous aqueous aerosol. The composition of this aerosol is expected to be sulfuric acid and water (H_2SO_4/H_2O) in various states of dilution, though the presence of nitric acid may also lead to the formation of ternary ($HNO_3/H_2SO_4/H_2O$) mixtures [*Carslaw et al.*, 1995]. Laboratory studies on the freezing of submicrometer-sized supercooled sulfuric acid aerosols at the Aerosol Interactions and Dynamics in the Atmosphere (AIDA) facility at the Forschungszentrum, Karlsruhe, Germany, have shown that substantial supersaturation is required to initiate homogeneous nucleation under conditions

representative of the upper troposphere [*Mohler et al.*, 2003]. Over the range of typical UT temperatures, i.e., from ~230 – 190 K, the laboratory ice saturation ratios, measured at the onset of ice formation, varied from ~1.40 to 1.75 respectively. Furthermore, experiments conducted on ambient aerosol under well-controlled temperature and humidity conditions showed that homogeneous freezing required ice supersaturations of a similar magnitude (~1.55 to 1.65) [*DeMott et al.*, 2003a]. Conceptually this process may be understood to occur in the following manner. As the liquid aerosol cools it takes up water vapor and its sulfuric acid weight percent decreases. At a critical point, shown to be largely determined by the water activity of the solution [*Koop et al.*, 2000], stable ice germs form and freezing commences.

The second mode of nucleation, known as heterogeneous nucleation, describes the process by which freezing is initiated with the aid of a foreign body, or "efficient" ice nuclei. These ice nuclei may be incorporated into the bulk of the liquid aerosol, may initiate freezing upon surface contact with the liquid aerosol, or may initiate ice formation via the deposition of vapor directly onto their surface [*DeMott et al.*, 2003a]. In general, heterogeneous ice nuclei constitute a small fraction of the background aerosol. However, they may be quite diverse and include both partially soluble and insoluble species. Examples include mineral dust (found to be very efficient) [*DeMott et al.*, 2003b], metals, soot, solid organics, and ammonia/sulfate mixtures, e.g., $(NH_4)_2SO_4$. Regardless of type, the degree of saturation required for heterogeneous nucleation to commence is necessarily lower than that for homogeneous nucleation.

Accurately measuring RH_{nuc} in the laboratory and in the atmosphere is challenging because it requires both highly specific environmental conditions as well as instrumentation with a high sensitivity to particle detection, e.g., [*Heymsfield et al.*, 1998] (atmosphere), and [*Abbatt et al.*, 2006] (laboratory). The difficulty arises in part because of the competition between the time it takes for the particle density to become detectable and the simultaneous evolution of relative humidity at the onset of ice nucleation. RH_i often continues to increase during the initial phases of cirrus development because of competition between the excess 'vapor supply rate' due to continued cooling, and the 'vapor depletion rate', which is governed by the rate at which vapor is scavenged by the growing ice crystals. Thus, the measured RH_{nuc} is generally expected to be biased slightly high relative to the "true" RH_{nuc} at which the first few ice crystals nucleate.

Furthermore, measurements of the clear air RH_i at or near most cloud boundaries will rarely coincide with a region of active nucleation, and thus seldom provide an actual measure of RH_{nuc}. The accepted understanding of the structure of in situ cirrus is that nucleation occurs preferentially in the supersaturated upper regions of the cloud, particle growth occurs through the bulk, and sedimentation followed by sublimation occurs in subsaturated air near the cloud base. As a result, certain in situ and remote studies targeted cirrus cloud tops in order to constrain RH_{nuc}, while others targeted and sampled wave clouds because of the well defined conditions at the upwind edge where the phase transition occurs, e.g., [*Comstock et al.*, 2004; *Heymsfield and Miloshevich*, 1995]. Our RH_i dataset, acquired predominantly in and around cirrus generated

from convective anvils and detrained into either supersaturated or undersaturated ambient air, is significantly more complex than such controlled studies. Thus, it is with these limitations in mind, that we examine the distribution of clear air RH_i for each mission.

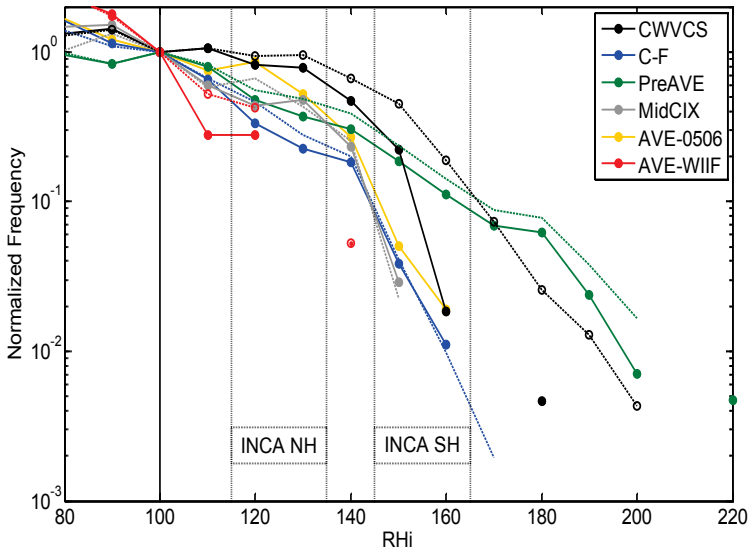


Figure 6.5. This figure shows the observed frequency distribution of clear air relative humidity observations normalized by the number of observations at saturation (RH_i = 100%) for each mission. Data were selected to be within the UT/LS.

Figure 6.5 shows a logarithmic plot of the relative frequency of clear air UT/LS data within each 10% RH_i bin color-coded by mission. The number of points in each RH_i bin are normalized to the total in the bin centered at RH_i = 100%. The difference between the solid lines with symbols and the dashed lines results from the choice of clear air selection criteria. The dashed lines include data that fall in the ambiguous regions between the designated clear and cloudy air requirements. Except in the case of the CWVCS mission, the inclusion of these data does not significantly change the results. The datasets show two distinct

groupings. The midlatitude and sub-tropical data exhibit a "cutoff" RH_{nuc} value between ~140-160%, while the tropical data show clear air RH_i data in excess of 180%.

Of the extra-tropical data, the results from the AVE-WIIF mission are outliers, as the apparent RH_{nuc} "cutoff" value is below the generally accepted range required for homogeneous nucleation. In fact, these data provide an excellent example of the ambiguity of interpreting the in situ results. For example, we might conclude 1) that heterogeneous nucleation, initiated around RH_i ~140%, is the dominant mode in this dataset, or 2) that the air has simply not reached the homogeneous threshold, and the distribution is dictated by the coldest temperatures reached, rather than the onset of ice nucleation. A detailed examination of the cloud encounters and the surrounding clear air reveals that in every case the ambient environment was substantially undersaturated, consistent with the low frequency of ISSRs observed in this mission. The small fraction of clear air data that show supersaturation, and drive the tail of the distribution in Figure 6.5, are not in the vicinity of cirrus, and thus provide no definitive information about the modes of ice nucleation in this dataset. The clouds sampled were most likely convective remnants, and none appeared to be in an ice-generating or growth phase. This is an important result as it illustrates one of the limitations in determining RH_{nuc} from a data set that was not specifically designed to probe cirrus nucleation mechanisms.

The rest of extra-tropical datasets exhibit a "cutoff" RH_{nuc} value of ~160%. The clear air frequency distribution of RH_i in these datasets steadily

decreases above the thermodynamic equilibrium value of 100%, and there are no instances of RH_i exceeding ~166%. This is consistent with the laboratory studies cited above, which showed that the homogeneous nucleation of sulfuric acid aerosols required substantial supersaturation, ~165%, at temperatures below 200 K [*Mohler et al.*, 2003]. We conclude, therefore, that homogeneous nucleation was an active, though perhaps not dominant, mode of ice/cloud nucleation during the extra-tropical missions.

These results do not preclude the possibility that heterogeneous nucleation was also active during the extra-tropical missions. Indeed, it is evident from the cirrus encounter during the MidCIX flight on May 3, 2004, shown in Figure 6.4, that aircraft contrails promoted cirrus formation and growth below the homogeneous threshold (dotted blue line at 157%). Clear air humidity in the sample region was ~135%, consistent with levels found to be adequate for heterogeneous nucleation given sufficient nuclei. In this case, the direct injection of efficient heterogeneous ice nuclei in the aircraft exhaust seeded the cirrus sampled on this flight. In fact, these RH_{nuc} values are comparable to the thresholds measured in continental midlatitude data by [*Heymsfield et al.*, 1998] (dashed blue line at ~135%). [*Heymsfield et al.*, 1998] found that the RH_{nuc} values measured over the continental U.S. were consistently lower than those over maritime environments, and speculated that the number and diversity of aerosol types over the continent, including mineral materials from the surface and contributions from jet aircraft exhaust, facilitated ice nucleation at lower ice-supersaturations. We have evidence, therefore, that both modes of ice nucleation

are active in the extra-tropical dataset. Furthermore, the extra-tropical data are generally in agreement with ice nucleation thresholds measured in earlier aircraft missions as well as laboratory studies.

The RH_{nuc} values for the tropical data sets, in contrast, exceed the accepted limit for homogeneous nucleation. Clear air RH_i in these datasets extends to unprecedented values. In CWVCS, clear air RH_i data extend beyond 180% even with the more restrictive clear air criteria. Data from PreAVE exceed ~200%, approaching the equilibrium value for vapor over supercooled liquid water. These extreme clear air supersaturations, particularly evident in the wintertime PreAVE dataset, challenge the current understanding of homogeneous ice nucleation, and suggest that there may be mechanisms that inhibit ice nucleation in the real atmosphere. If prevalent, these high nucleation thresholds have the potential to increase modeled estimates of the frequency and extent of ISSRs in the TTL, and decrease estimates of the frequency and coverage of thin, laminar near-tropopause cirrus. Consequently, they will impact calculations of $[H_2O]_e$, as well as the chemical and radiative budgets of the tropical UT/LS.

6.7.3. Clear Air RH_i: Dependence on Temperature & H_2O

While the ice nucleation threshold, RH_{nuc}, is essentially defined by the maximum clear air RH_i observed, the larger structure and variability of clear air RH_i in the UT is, by definition, driven by variability in ambient temperatures and water vapor mixing ratios. In this section we examine the relative role of each in

determining the clear air RH_i distributions using a combination of the composite dataset and specific case studies.

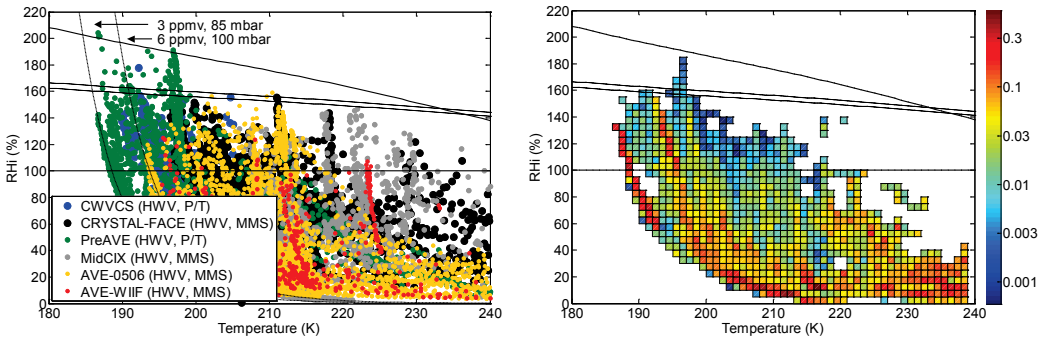


Figure 6.6. Clear air RH_i data plotted as a function of temperature. Panel (a) shows the data color-coded by mission. Panel (b) shows the data color-coded by the fractional frequency of occurrence. In this figure, the frequencies are determined as a function of the total number of points in each 1 K temperature slice.

Figure 6.6a shows the clear air relative humidity data for the combined dataset plotted as a function of temperature and color-coded by mission. The solid black line at RH_i = 100% indicates the equilibrium value for water vapor over plane hexagonal ice, and the solid black line extending from 140% to 210% designates the equilibrium value for water vapor over supercooled liquid. In both cases we have used the parameterizations derived by [*Murphy and Koop*, 2005]. The black dotted lines represent parameterized fits to the AIDA chamber homogeneous nucleation data, where the relative humidity nucleation threshold, RH_{nuc}, was measured as a function of temperature for micrometer sized sulfuric acid aerosol [*Mohler et al.*, 2003]. Finally, the dashed black lines follow the evolution of RH_i as the ambient temperature changes for constant mixing ratios of 3 and 6 ppmv, respectively.

Figure 6.6a captures some of the subtlety of the clear air RH_i distributions, and exposes the response of RH_i to changing vapor mixing ratios and temperatures. Specifically, data that are distributed vertically in Figure 6.6.a, i.e., along lines of constant temperature, demonstrate the variability in RH_i associated with changes in ambient water vapor, and data that follow the trend of the black dashed lines of constant mixing ratio, represent the response of RH_i to changing temperature.

Varying Water Vapor Mixing Ratio

An examination of the vertically distributed data at constant temperature, acquired while sampling along level flight paths in clear air, shows that the water vapor mixing ratio can vary and drive the ambient RH_i up to the level required for homogeneous nucleation. For data with temperatures >200 K, it is clear that this level is an effective threshold, i.e., there are no clear air RH_i data that exceed the homogeneous threshold at temperatures warmer than 200 K. This is consistent with the distribution function analysis of Figure 6.5, and Figure 6.6 reveals that these threshold data are associated with a few discrete sampling events in each mission.

Varying Temperature

While the vertical distributions, associated with constant temperature and changing mixing ratio, have a sporadic structure characteristic of isolated sampling events, the broad distribution of clear air RH_i data in Figure 6.6a bears

the signature of being driven primarily by changes in temperature. The lines of constant mixing ratio, at 3 and 6 ppmv respectively, reveal the dramatic impact that relatively small changes in temperature have on the ambient relative humidity. In particular, the tropical clear air data at temperatures <200 K, that extend to, and in some instances, push beyond the homogeneous threshold toward the limit of liquid water saturation, look like airmasses with relatively constant mixing ratios that are cooling as they ascend through the TTL.

Figure 6.6b shows the same clear air RH_i data plotted as a function of temperature and frequency (color bar). The data in Figure 6.6b have been sorted into bins 1 K wide by 5% tall. The number of observations in each bin is normalized by the total number of observations in each 1 K wide column. Thus, the frequency of RH_i observations in each bin is expressed as a fraction of the data within each temperature slice. The data show that the distribution of in-cloud RH_i shifts in response to temperature, with the median value of the distribution increasing as temperature decreases.

Case Study Analysis

We further explore the role of temperature in determining ambient RH_i by contrasting two flights from the CRYSTAL-FACE mission. A general cooling of the UT was observed during the middle of the CRYSTAL-FACE mission. This cooling, in combination with a warming of the tropopause region at the end of the mission, provided an excellent demonstration of the dominance of temperature in determining the dynamic range of clear air RH_i.

The high cold tropopause, observed on the flight of July 13, 2002, generated some of the coldest UT temperatures and highest clear air supersaturations of the mission, despite the fact that this flight recorded average to below average water vapor mixing ratios around 360 K. Furthermore, the cooling on this day led to the formation of a thin tropopause cirrus layer [*Jensen et al.*, 2005; *Webster and Heymsfield*, 2003], one of the few in situ cirrus sampled during the mission. In contrast, the flight of July 28, 2002, recorded some of the warmest tropopause temperatures, and combined with a lower than average water profile, it yielded some of the least saturated UT air observed during the CRYSTAL-FACE mission.

Profiles of measured temperature, water vapor and corresponding RH_i from CRYSTAL-FACE are plotted in Figure 6.7a-c. The mission means and standard deviations of the measurements (in 3 K potential temperature bins) are represented by the large solid black circles, and the data from the two flights under comparison, July 13, 2002, and July 28, 2002, are plotted in blue and red dots, respectively. The most striking feature of Figure 6.7b is the large temperature difference, >10 K, between the flight profiles in the near tropopause region (~365 K ±5 K), with convergence toward the average temperature profile above and below. This temperature difference, likely caused by synoptic scale disturbances, is the primary factor driving the difference in RH_i between these two flight profiles.

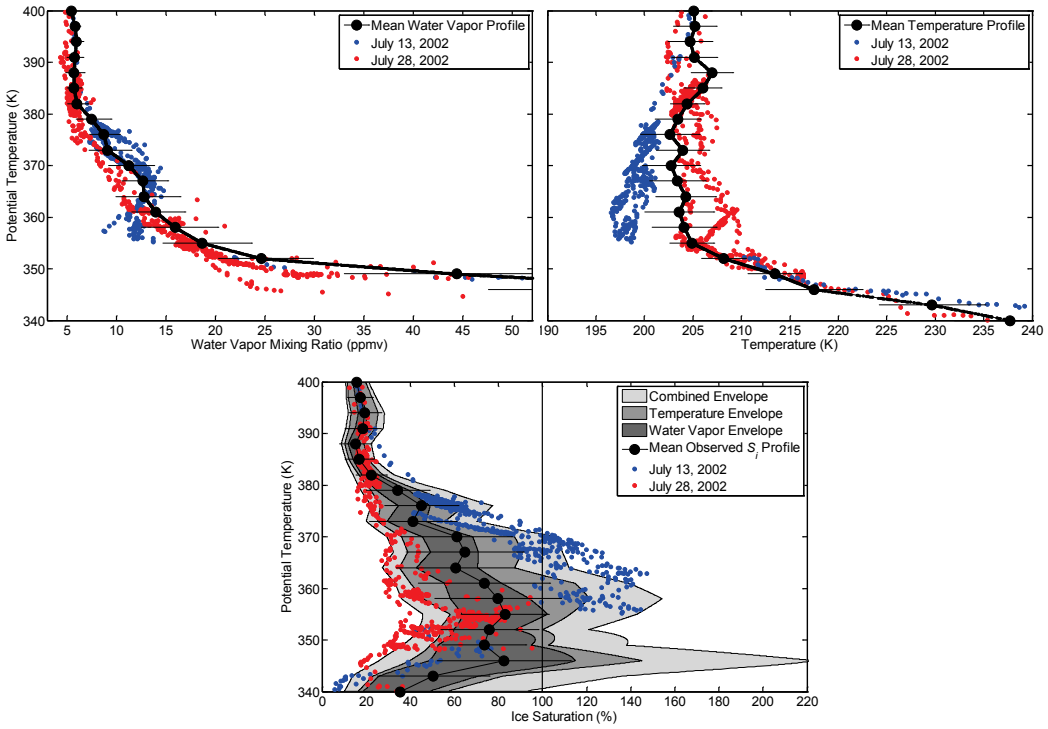


Figure 6.7. Examination of the ambient temperature and water vapor control of clear air RH_i using data from the CRYSTAL-FACE flights of July 13, and July 28, 2002. Panel (a) shows water vapor mixing ratios for each flight date plotted versus potential temperature as the vertical coordinate, as well as the bin-averaged vertical profile for sub-tropical clear air data for the entire mission. The horizontal bars correspond to one standard deviation about the mean, and are plotted to indicate the range of measured values. Panel (b) shows the comparable vertical profiles of temperature; and panel (c) shows the resultant RH_i profiles. The shading in panel (c) captures the range of RH_i values resulting from variability in water vapor alone, given the mean temperature profile (dark gray). The envelopes accounting for temperature variability alone, using the mean water vapor profile (medium gray), and accounting for variability in both temperature and water vapor (light gray) are also shown.

We applied a simple analysis to the CRYSTAL-FACE flight data to explore the dependence of the clear air RH_i distribution on water vapor and temperature. The average profiles plotted in Figure 6.7a-b were used to generate simulated average relative humidity profiles. The envelope of RH_i resulting solely from variations in water vapor (dark gray shading in Figure 6.7c) was calculated

using the mean temperature profile and the mean water vapor profile ± its standard deviation. The range of the simulated RH_i resulting from variability in water alone is too narrowly distributed to match the extremes of the observed distribution represented by the data from the two flights examined above. The corresponding temperature envelope (medium gray), generated using the mean water vapor profile and the range of temperatures bounded by the mean temperature ± its standard deviation in each bin, is roughly twice as wide. Therefore, variability in ambient temperatures, ~±3 K, is the dominant driver of the RH_i distribution in clear air. As expected, the range of simulated RH_i that arises due to variations in both ambient temperature and water vapor (light gray) captures the full range of the observations.

6.8. In-Cloud RH_i

6.8.1 In-Cloud RH_i: Persistent Ice Supersaturation

The empirical distribution function of in-cloud, upper tropospheric RH_i data from CRYSTAL-FACE, is plotted in blue in Figure 6.2. The corresponding empirical distribution functions for all of the missions are plotted in Figure 6.8. There were no in-cloud data for the AVE-0506 mission. Both Figures 6.2 and 6.8 show that supersaturation is commonly observed within clouds. Figure 6.2 shows that the CRYSTAL-FACE in-cloud RH_i data are symmetrically distributed around a peak value of ~110%, with broad tails extending into undersaturated air with RH_i ~70% and supersaturated air with values as high as ~180%. The fact that the center of the distribution is ~110%, rather than the equilibrium value of 100% is

not statistically significant given the combined uncertainties in the water vapor and temperature measurements.

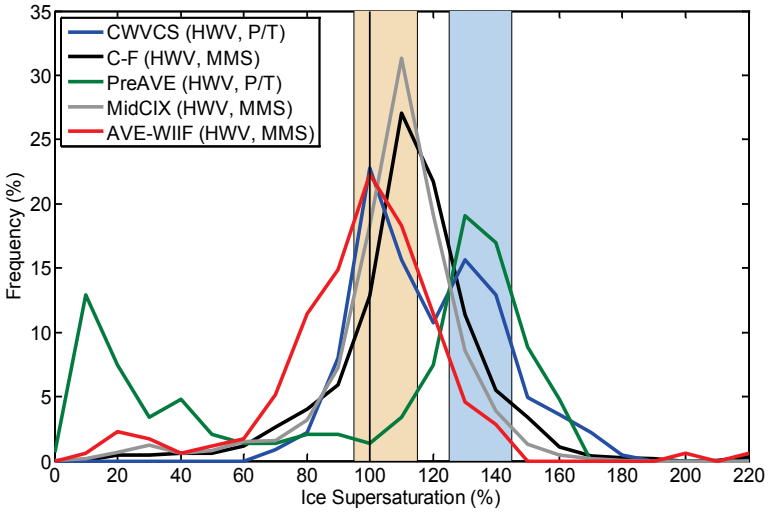


Figure 6.8. Empirical distribution functions of in-cloud RH_i data for each mission. Data were selected to be within the UT/LS.

Figure 6.8 shows the CRYSTAL-FACE data in the context of the complete data set. This figure reveals an underlying bi-modal character to the distribution of in-cloud RH_i, with one mode centered near 100%, and the second centered near 130%. The first group is dominated by midlatitude spring and summer data from the CRYSTAL-FACE (black), MidCIX (gray) and AVE-WIIF (red) missions, and the second is dominated by wintertime tropical data from PreAVE (green). The summertime tropical data from the CWVCS (blue) mission explicitly shows this bi-modal character with peaks in both regions. The two regions are shaded light pink and blue respectively, corresponding to clouds sampled in warmer and colder regions.

Similar to Figure 6.6, Figure 6.9a-b provides a better context for viewing the in-cloud RH_i results. In this figure in-cloud RH_i is plotted as a function of temperature. In Figure 6.9a the data are color-coded by mission, and in Figure 6.9b the data are color-coded by frequency. The data in Figure 6.9b have been sorted into bins 1 K wide by 5% tall. The number of observations in each bin is normalized by the total number of observations in each 1 K wide column. Thus the frequency of RH_i observations in each bin is expressed as a fraction of the data within each temperature slice.

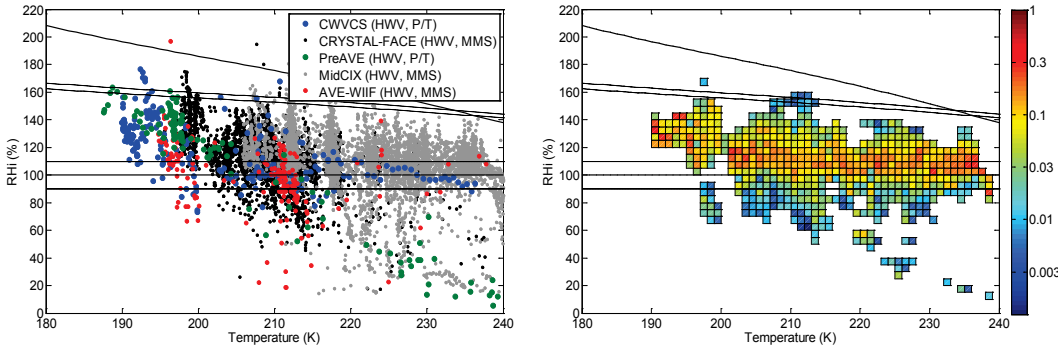


Figure 6.9. In-cloud relative humidity observations for each mission plotted as a function of temperature. Panel (a) is the in-cloud companion to Figure 6.6, each mission is color-coded according to the legend. In panel (b) the color coding represents the RH_i frequency distribution in each 1 K temperature slice.

Both panels of Figure 6.9 show that the distribution of in-cloud RH_i shifts in response to temperature, with the median value of the distribution increasing as temperature decreases. In clouds with ambient temperatures below ~200 K, the median RH_i value is ~130%. The in-cloud data from each mission show evidence of this shift, with a convergence toward 100% in lower, warmer cirrus (>210 K), and an increased frequency of supersaturated values in high, cold (<200 K), near tropopause cirrus. Furthermore, the data indicate that supersaturation in clouds is

a common phenomenon, with the distribution in cold clouds often entirely in supersaturated air.

Figure 6.10 shows that these supersaturations persist in cold clouds despite high in-cloud surface area densities. We might expect high relative humidities to correlate with low surface area density because of limitations on vapor diffusion to individual cloud particles, however, the distribution is relatively insensitive to cloud density. Figure 6.10a shows in-cloud relative humidity plotted as a function of temperature and binned by cloud ice water content, as measured by the HWV/HTW instrument pair. Substantial supersaturation is present at cold temperatures despite of a two order of magnitude range in cloud ice water content. A similar analysis (not shown) reveals that this persistent in-cloud supersaturation observed at cold temperatures is effectively insensitive to a two order of magnitude change in surface area density.

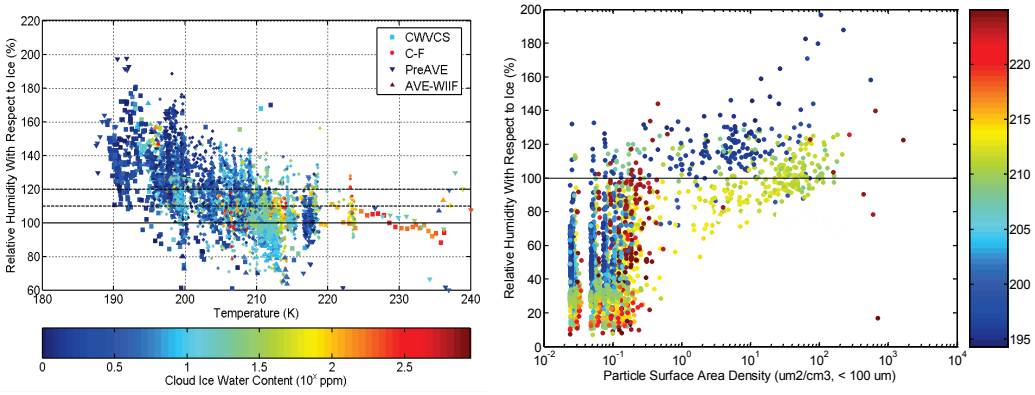


Figure 6.10. In-cloud relative humidity data. Panel (a) shows the RH_i data for each mission plotted as function of temperature, and cloud ice water content determined from the pair of Harvard hygrometers. The solid black line denotes the value for water vapor in equilibrium with ice, RH_i = 100%, and the color-coding is representative of the cloud ice water content ($\log_{10}$ scale). Panel (b) shows in-cloud relative humidity observations from the AVE-WIIF flight of July 7, 2005, plotted as a function of particle surface area density and color-coded by ambient temperature. The solid black line again denotes RH_i = 100%.

The AVE-WIIF dataset afforded another illustrative case-study. Figure 6.10b shows in-cloud RH_i from AVE-WIIF plotted as a function of particle surface area density measured by CAPS. The data are color-coded by ambient temperature. The figure features two distinct groups of in-cloud RH_i data, where the difference is clearly dictated by ambient temperature. Both sets span comparable particle surface area densities, and both sets appear to span a similar (±20%) range in RH_i. Typical of the warm clouds in the combined dataset, the in-cloud RH_i converges toward 100%.

We conclude, therefore, that there is a strong temperature dependence to in-cloud RH_i, and that this temperature dependence appears to be unexpectedly insensitive to cloud thickness. Significant in-cloud supersaturation is present and persistent even when cloud particle number densities and surface areas are high. This result has significant implications for the dehydration efficiency of near tropopause cirrus. Sustained supersaturations, on the order of 30%, in high cold laminar cirrus have the potential to raise estimates of $[H_2O]_e$ by the same fraction, again leading to a moister stratosphere than that predicted by efficient thermal control.

6.8.2. In-Cloud RH_i: Dependence on Temperature & H_2O

Though the mean in-cloud RH_i appears to be driven by ambient temperature, much of the variability about the mean is driven by variability in water vapor mixing ratios within the cloud. The vertical spread of in-cloud RH_i data at a given temperature, evident in Figures 6.9 and 6.10, is not due to a lack of

precision in the water vapor measurements. Instead, it is caused by measurable inhomogeneities in the water vapor mixing ratio within single cloud encounters.

Figure 6.11 shows data from a representative encounter from AVE-WIIF. RH_i, water vapor, temperature, cloud ice water content and particle number density are plotted as a function of time for a single cloud encounter. The close correspondence between the variability of the measured water vapor mixing ratios and the resulting RH_i observations is evident. The variability in the calculated ice saturation mixing ratio, SMR_i, which represents the effect of temperature changes on RH_i, is negligible by comparison.

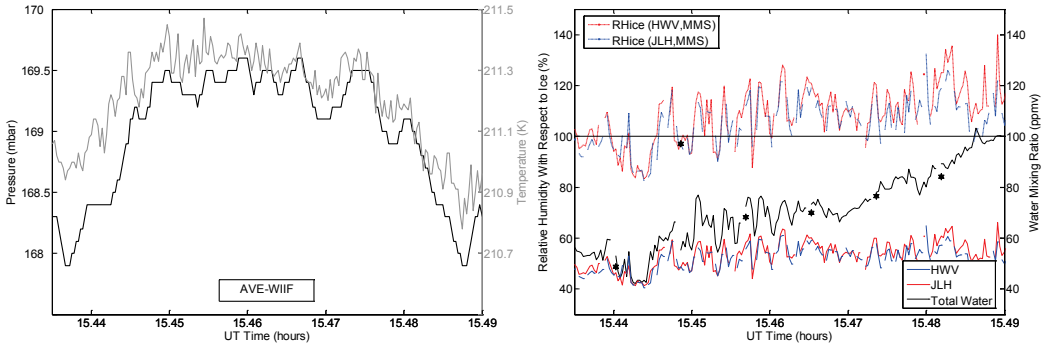


Figure 6.11. Representative cloud encounter from the July 5, 2005, flight during AVE-WIIF. Panel (a) shows the measurements of ambient pressure and temperature as a function of time during the respective flight segments. Panel (b) shows the corresponding measurements of water vapor, HWV (red) & JLH (blue), and the resulting relative humidity values. In panel (b) agreement between HWV and HTW (black) denotes clear air, and the difference between the two measurements yields the cloud ice water content (IWC). The black stars are scaled cloud particle density data from CAPS.

To establish that this variability is not due to poor instrument precision in HWV, we plot data acquired by both HWV and JLH. These two instruments are entirely independent, and thus the close correlation between the measurements,

noticeable even at a 1 Hz data rate, is representative of true fine scale heterogeneity within the cloud. The calculated correlation coefficient between the JLH and HWV water vapor measurements for the selected cloud encounter is 0.94. Indeed, Figure 6.11b reveals the excellent precision of both instruments at 1 Hz. Within the cloud, the short term variations in water vapor are on the order of ±5 (4) ppmv as measured by HWV (JLH), which corresponds to a ~±10% (~7%) range in in-cloud RH_i.

The concurrent temperature structure during the cloud encounter, shown in Figure 6.11a, is vastly different. The fine scale variability is on the order of ±0.04 K, which corresponds to a variability of less than ±1% in SMR_i. The broad scale structure in temperature is clearly caused by pressure, i.e., altitude, changes along the flight track, and does not contribute to RH_i variability within the cloud transect. Furthermore, a simple examination of the correlation coefficient between in-cloud relative humidity and water vapor mixing ratios unequivocally demonstrates the role of water vapor heterogeneity in determining the range of RH_i observed within the cloud. The calculated correlation coefficient between the water measurements and the derived RH_i is ~0.78. In contrast, an analogous correlation coefficient between in-cloud RH_i and SMR_i, which quantitatively represents the effect of ambient temperature changes, yields a value of ~0.10.

6.9. Discussion & Implications

In this section, we evaluate our RH_i data within the framework of the conventional understanding of ice nucleation and cirrus formation, examining

when and where our results agree with or challenge the standard model. Furthermore, we examine the consequences of our results within the context of the processes controlling lower stratospheric humidity, cirrus frequency and coverage, and the net chemical and radiative impacts of water vapor and clouds in the UT/LS.

6.9.1 Clear Air RH_i

We discuss first the clear air RH_i results. There are three factors that dictate the distribution of RH_i in clear air: ambient water vapor concentrations, ambient temperature, and the threshold for ice nucleation, RH_{nuc}, which, as we have discussed, functions as a nominal upper bound to RH_i in clear air. Our simple analysis of CRYSTAL-FACE data showed that the distribution of RH_i in clear air is largely driven by variations in ambient temperature. In comparison, the sensitivity of the RH_i distribution to variations in the ambient water vapor mixing ratio was significantly smaller. This result is consistent with prior studies, e.g., those of [*Gierens et al.*, 1999; *Haag et al.*, 2003; *Karcher and Haag*, 2004].

The high tail of the clear air RH_i distribution is truncated by cirrus formation, i.e., when RH_i reaches RH_{nuc}. As many earlier studies have demonstrated, atmospheric measurements of this "cutoff" clear air RH_i in cirrus forming regions can provide a means of inferring the threshold saturation required for ice nucleation, and thus provide information regarding the dominant nucleation mode. In the present analysis, we discussed the limitations of inferring RH_{nuc} from datasets not specifically designed to sample in regions of cirrus

formation. However, focusing upon specific examples within the composite dataset we showed, 1) data consistent with in situ cirrus formation resulting from heterogeneous nucleation on particles seeded by aircraft contrails, 2) data from midlatitudes to the tropics that are in agreement with the expected threshold for homogeneous nucleation, and 3) data with saturations significantly exceeding the homogeneous threshold especially at the cold temperatures encountered in the upper tropical troposphere. These latter observations challenge the accepted model of ice nucleation and cirrus formation and are discussed in more detail below.

There is ample evidence both from laboratory and in situ campaigns, including the data examined here, that shows that ice nucleation in liquid aerosol often requires significant supersaturation. Indeed, it is fully expected that an airmass will become supersaturated with respect to ice as it ascends and cools. The work of [*Koop et al.*, 2000], which provides the foundation for the prevailing model of ice nucleation, showed that for a wide range of liquid aerosol compositions the temperature at which ice nucleates homogeneously is a function only of water activity, which is essentially equivalent to the relative humidity with respect to liquid water at equilibrium. Furthermore, [*Koop et al.*, 2000] showed that significant supercooling is required, and that the threshold humidity with respect to ice increases as temperature decreases. At the temperatures of the UT/LS the equivalent ice supersaturation threshold for nucleation is ~65%. Thus, nucleation is expected to commence at values greater than the equilibrium value, RH_i = 100%, but RH_i is not expected to exceed ~165% under the conditions of the

UT/LS, i.e., this value should serve as a constraint on the maximum RH_i observed in clear air [*Peter et al.*, 2006].

Our clear air data acquired at temperatures >200 K agree well with this model. However, at temperatures <200 K our data exceed the expected threshold, approaching the equilibrium for water vapor over supercooled liquid. These data, acquired almost exclusively in the cold upper tropical troposphere, suggest that there is a mechanism inhibiting homogeneous ice nucleation in the "real" atmosphere, and that kinetic (non-equilibrium) processes may be important.

[*Koop et al.*, 2000] examined the freezing behavior of sulfuric acid, and other salt solutions, but considered only a small selection of organic species. More recent laboratory studies and a growing body of in situ evidence suggest that chemical composition, particularly the presence of organic components, can impact both heterogeneous and homogeneous nucleation mechanisms. Organics, either as coatings of otherwise efficient heterogeneous ice nuclei, or as components of liquid sulfate-organic aerosols may inhibit nucleation, and thus raise ice nucleation thresholds under conditions representative of the UT/LS.

Laboratory studies conducted at the AIDA facility at temperatures ~205 K found that the presence of organics significantly reduced the nucleation efficiency of dust particles [*Mohler et al.*, 2008]. As expected, the uncoated dust particles were highly efficient deposition mode heterogeneous ice nuclei, with nearly 100% nucleating ice at RH_i ~110%. With the introduction of organic coatings designed to mimic the effect of secondary organic aerosol (SOA) in the atmosphere, the efficiency of heterogeneous nucleation was substantially reduced. The data

suggest that the effect becomes more significant as the SOA surface coverage increases, with pure SOA particles requiring RH_i near or even in excess of liquid water saturation for strong ice-activation. These results unequivocally demonstrate that organic material can reduce the effectiveness of heterogeneous nuclei. Furthermore, they provide some evidence that pure SOA may require a RH_i significantly higher than the expected homogeneous nucleation threshold.

Similarly, recent laboratory studies examining the homogeneous freezing behavior of liquid aerosol have shown that under specific thermal and chemical conditions, the aerosol may become ultra-viscous and undergo a phase transition to a glassy, i.e., amorphous, solid state rather nucleate ice [*Bogdan*, 2006; *Murray*, 2008; *Zobrist et al.*, 2008]. In particular, the work of [*Murray*, 2008], and [*Zobrist et al.*, 2008], demonstrate this behavior for certain organic-rich particles. The glass transition occurs during cooling when molecular diffusion effectively stops and the aerosol becomes locked in a "liquid-like" amorphous state without crystallization. For solutions with viscosities orders of magnitude greater than that of water at room temperature, diffusion of liquid water molecules is expected to be very slow. Thus, even if an ice embryo forms in a viscous liquid aerosol its growth may be severely delayed if not altogether prevented. Laboratory evidence also suggests that as the ice germ grows in the bulk of aerosol the solute (a salt or organic species) is expelled, and the remaining liquid solution can then form a concentrated brine and/or ultra-viscous liquid coating [*Bogdan et al.*, 2006; *Murray*, 2008; *Zobrist et al.*, 2008]. Furthermore, both highly viscous liquid

aerosol and glassy solids are expected to have greatly reduced efficiency for water vapor uptake [*Zobrist et al.*, 2008].

The results of in situ studies also indicate that organics can affect ice nucleation. [*DeMott et al.*, 2003a] explored the freezing behavior of free tropospheric aerosol under well-defined temperature and humidity conditions, and subsequently analyzed the chemical composition of the frozen ice crystal nuclei. The chemical composition of individual sub-micrometer sized particles was determined by examining the ionic spectra of ablated aerosol using the NOAA Aeronomy Laboratory Particle Analysis by Laser Mass Spectrometery (PALMS) instrument [*Cziczo et al.*, 2003]. [*DeMott et al.*, 2003a], found that the homogeneous freezing of aerosols with an unusually high organic fraction required measurably higher saturations than purer sulfate particles. Similarly, PALMS data acquired in and around UT cirrus during the CRYSTAL-FACE mission showed that the organic signal in the interstitial aerosol, small unfrozen particles within ice clouds, was consistently higher than that in the ice crystal residue [*Cziczo et al.*, 2004]. This unequal partitioning of organics between the aqueous and ice phases, i.e., the apparent preference for aerosols with organics to remain liquid, is consistent with the emerging theory that internally mixed organic-sulfate aerosol do not nucleate as readily as pristine sulfates.

[*Froyd et al.*, 2009], examined the chemical composition of aerosol in the TTL (~13 – 16.5 km) using the PALMS instrument aboard the NASA WB-57 aircraft during the PreAVE campaign, and the Costa Rica Aura Validation Experiment (CRAVE) campaign during January and February of 2006. At the

time of publication, their results constituted the most complete set of aerosol composition measurements in the TTL and tropical lower stratosphere. Consistent with a change in convective influence, the data near the base of the TTL showed a sharp change in aerosol chemical composition. Organic-rich signals, associated with the direct injection of particles and chemical precursors from deep continental convection, were evident above ~12 km. Over 99% of particles sampled in the TTL during these campaigns had a measurable organic signal, and 80% of sulfate-organic mixtures had organic mass fractions >10% [*Froyd et al.*, 2009].

Thus, the ubiquitous presence of organics in the cold upper tropical troposphere, in combination with the results of laboratory nucleation studies, offers a coherent microphysical explanation for the inhibition of ice nucleation and the unexpectedly high clear air RH_i observations at temperatures ~<200 K. The combination of results from laboratory nucleation studies on organic-rich aerosol, and the observation of a high organic signal in tropical UT suggest that if an air parcel contains a range of sulfate-rich to organic-rich particles, homogeneous ice nucleation will occur preferentially in the purer sulfate aerosols, which have lower viscosities and can adapt more readily to changes in ambient relative humidity [*Zobrist et al.*, 2008]. These newly crystallized particles will then rapidly deplete the excess vapor leaving the organic-rich aerosol particles unfrozen, in a process virtually identical to that observed by [*DeMott et al.*, 2003a] and [*Cziczo et al.*, 2004] in the UT. This selective ice nucleation will then influence the composition of the aged upper tropospheric aerosol, leading to an

enrichment in organic-rich, glassy aerosols that are less likely to form cirrus at the homogeneous ice nucleation threshold, and consequently cause clear air relative humidities to increase consistent with our clear air RH_i observations.

The scenario outlined here alters the expected balance between water in the vapor and condensed phases in the tropical upper troposphere, by shifting the balance toward the vapor phase. If representative, these results could have a significant impact on model estimates of cirrus cloud frequency and coverage, as well as on cirrus particle sizes and number densities, with further consequences for our understanding of the radiative and chemical properties of the region

Furthermore, our clear air RH_i results call in to question the "freeze-drying" efficiency of the tropical tropopause. If representative of conditions in the upper tropical troposphere, they could lead to higher estimates of $[H_2O]_e$. The nucleation of ice particles and their subsequent growth and sedimentation near the cold point tropopause leads to the dehydration of airmasses prior to entry into the stratosphere. Our observations of prevalent clear air supersaturation suggest that these processes are inhibited, compromising the efficient scavenging of water vapor in excess of the ambient SMR_i. If irreversible transport into the stratosphere occurs in regions of significant ice supersaturation, the entry level value will be $[H_2O]_e = SMR_i \cdot (1 + f_{SS})$, where f_{SS} is equal to the degree of supersaturation. A simple analysis of parcel temperatures versus time and height along simulated back-trajectories in the tropics showed that most (~75%) of the parcels that underwent irreversible transport into the lower tropical stratosphere got colder than 190 K [Eric Jensen, *personal communication*, 2010]. Thus, our clear air RH_i

data, which show a simultaneous increase in both the frequency and degree of supersaturation at the coldest temperatures, are particularly relevant for understanding and forecasting $[H_2O]_e$.

6.9.2 In-Cloud RH_i

We turn now to a discussion of the in-cloud RH_i results. The accepted model of RH_i within clouds assumes that, given sufficient surface area, i.e., where vapor diffusion is not limiting, the RH_i is expected to be near equilibrium, i.e., ~100%. The theory predicts that once the threshold relative humidity has been reached, ice nucleation promptly commences (within ~<1 minute), the particles grow rapidly via condensation, and deplete the excess vapor until equilibrium is established [*Peter et al.*, 2006] *and references therein*.

Our in-cloud RH_i data acquired at temperatures warmer than ~200 K exhibit the expected convergence toward 100%. Figure 6.9b shows that the composite distribution of RH_i as a function of temperature (for temperatures >200 K) is nominally centered at ~110%, in agreement with the equilibrium condition within the combined experimental uncertainty of the water vapor and temperature measurements.

The in-cloud data acquired at ambient temperatures lower than 200 K, however, challenge the conventional understanding of the evolution of RH_i within cirrus. Figure 6.9b shows that at low temperatures the distribution of RH_i appears to be centered around ~130% ± 10%. Furthermore, Figures 6.12a-b, demonstrate that these high humidity results are consistently insensitive to cirrus ice density.

Though the rate of relaxation toward equilibrium is expected to vary depending upon vertical ascent rates, as well as on vapor diffusion rates to, and accommodation rates on, ice crystals, our measurements of ubiquitous supersaturation in the presence of cold (<200 K) cirrus cannot be explained with the prevailing model.

[*Gao et al.*, 2006; *Gao et al.*, 2004], examined HWV measurements of sustained supersaturation within cold near tropopause cirrus and within an aged contrail from the WB-57 flight of July 13, 2002, during the CRYSTAL-FACE mission. The relaxation time in the contrail, with particle number densities between ~100 and 200 cm^{-3}, is expected to be ~<1 minute. Thus, the contrail data, acquired 5 to 40 minutes after its generation, are assumed to be representative of an equilibrium state. Our measurements of RH_i within the contrail, as well as within the surrounding cirrus, which had surface area densities in excess of 100 $\mu m^2/cm^{-3}$, were ~130% suggesting that there is some mechanism either inhibiting relaxation toward equilibrium or altering the equilibrium condition at temperatures less than ~200 K.

[*Gao et al.*, 2004], proposed that HNO_3 adsorbed on the ice particles can enhance ambient RH_i. HNO_3 is readily adsorbed onto low-temperature surfaces of liquid water and ice, and simultaneous measurements of both gas and condensed phase HNO_3 showed that it was present on all the ice particles sampled [*Gao et al.*, 2004], *and references therein*. Furthermore, [*Gao, et al.*, 2004], suggest that because HNO_3 has low bulk solubility, once it is adsorbed it remains at the surface and combines with H_2O to form NAT or NAT-like molecules, which they

designate Δ-ice. They theorize that this change in particle surface characteristics affects the rates of water vapor accommodation.

The laboratory study of [*Delval and Rossi*, 2005], examined the evaporation rate of bulk ice coated with monolayer amounts of HNO_3 at temperatures between 179 K and 208 K. They found the evaporation rate gradually decreased with the evaporation of H_2O and the concomitant increase in the HNO_3 weight percent. The final rates of evaporation of coated ice were substantially slower than that of pure ice. In order to maintain thermodynamic consistency, the rates of evaporation and condensation over the sample must scale such that their ratio is constant. A slower evaporation rate implies a slower rate of condensation, leading [*Delval and Rossi*, 2005], to conclude that H_2O/HNO_3 complexed surface structures can substantially slow the rate of relaxation toward the equilibrium ice condition. The relaxation times they calculate for the [*Gao et al.*, 2004], dataset range from ~1 minute to 1 day. If, as they suggest, three to five relaxation periods are needed to effectively establish vapor equilibrium, HNO_3 adsorbed on the surface of cirrus may provide an explanation of the sustained in-cloud supersaturation via changes in kinetic parameters.

An alternative explanation for the high in-cloud RH_i is supported by the laboratory studies of [*Murray et al.*, 2005; *Shilling et al.*, 2006]. [*Murray et al.*, 2005], show that when micrometer-sized droplets freeze homogeneously, the meta-stable form of ice, called cubic ice (Ic), can form at temperatures below ~210 K. Furthermore, their results indicate that Ic is the preferred form at colder temperatures (~<190 K) and higher solute concentrations, forming with nearly

100% probability. [*Shilling et al.*, 2006], subsequently demonstrated that the equilibrium vapor pressure over cubic ice is ~10.5% higher than that of hexagonal ice (Ih). Together, these results support the existence of Ic under atmospherically relevant conditions, and offer a thermodynamic argument for the elevated vapor pressure in cold clouds. [*Shilling et al.*, 2006], reevaluated the cold in-cloud data from CRYSTAL-FACE and found it to be consistent, within experimental error, with their laboratory vapor pressure results.

Our composite dataset, however, shows that below 195 K the RH_i distribution in each 1 K temperature bin has a significant, if not dominant fraction of data exceeding ~125%, i.e., falling outside the bounds of the combined measurement uncertainties in ambient RH_i and the laboratory vapor pressure results. It is possible, therefore, that the high frequency of supersaturation in cold cirrus is due to a combination of kinetic and thermodynamic effects.

6.10. Conclusions

We have presented in situ observations of upper tropospheric and lower stratospheric relative humidity with respect to ice, RH_i, from six recent measurement campaigns aboard NASA's WB-57 high-altitude research aircraft. These data, determined from accurate measurements of water vapor, temperature and pressure, extend the latitudinal extent of aircraft-borne RH_i observations. The data cover a range from the tropics to midlatitudes, and were acquired in different seasons. Furthermore, the full complement of data sets spans a fifty-degree

temperature range, and three orders of magnitude in both water vapor and cloud ice water content.

The data show that supersaturation in clear air is common with values in the UT approaching the level required for homogeneous nucleation. Furthermore, our cold temperature (<200 K) tropical UT data show supersaturation in excess of this threshold, approaching the limit of water vapor in equilibrium with supercooled liquid. These high clear air supersaturations have important implications for our understanding of, and ability to model, cloud nucleation, high cloud cover, and net cloud radiative forcing. The data plainly show that cloud formation in the UT does not simply occur when ambient RH_i = 100%, and models operating under this assumption may be prone to erroneously high estimates of cloudiness. Furthermore, these results challenge the strict thermal control of $[H_2O]_e$ by the cold-point tropopause. If regions of clear air ice-supersaturation are large and frequent, the assumption that $[H_2O]_e = SMR_i(T_{CP})$ may under-predict the amount of water vapor entering the stratosphere.

Our in-cloud results show that supersaturation is also present within clouds. Sustained supersaturations appear to be increasingly common as temperatures decrease below 200 K independent of cloud particle densities. Similar to the results in clear air, our in-cloud challenge the dehydration efficiency of cold near tropopause cirrus. Enduring supersaturations in cirrus of ~30% yield stratospheric entry levels of water vapor 30% higher than those predicted by the saturation mixing ratio at the cold-point.

Consistent with the extreme sensitivity of RH_i to small changes in temperature, we showed that the distribution of clear air RH_i data is largely driven temperature variations caused both by large synoptic-scale disturbances and smaller variability due to wave activity.

Analogous to our clear air results, the broad structure of in-cloud RH_i is also determined by temperature. However, much of the variability of RH_i in individual cloud samples results from measurable inhomogeneities in the water vapor field. The degree of observed variability, often ~±10%, is well outside the limits of our instrument precision. This inhomogeneity should not be surprising given the turbulent mixing processes operating within clouds and the long equilibration times for water vapor in cold cirrus.

References

Abbatt, J. P. D., S. Benz, D. J. Cziczo, Z. Kanji, U. Lohmann, and O. Mohler (2006), Solid ammonium sulfate aerosols as ice nuclei: A pathway for cirrus cloud formation, *Science*, *313*(5794), 1770-1773.

Baumgardner, D., H. Jonsson, W. Dawson, D. O'Connor, and R. Newton (2001), The cloud, aerosol and precipitation spectrometer: a new instrument for cloud investigations, *Atmos Res*, *59*, 251-264.

Bogdan, A. (2006), Reversible formation of glassy water in slowly cooling diluted drops, *J Phys Chem B*, *110*(25), 12205-12206.

Bogdan, A., M. J. Molina, K. Sassen, and M. Kulmala (2006), Formation of low-temperature cirrus from H2SO4/H2O aerosol droplets, *J Phys Chem A*, *110*(46), 12541-12542.

Carslaw, K. S., B. P. Luo, and T. Peter (1995), An Analytic-Expression for the Composition of Aqueous Hno3-H2so4 Stratospheric Aerosols Including Gas-Phase Removal of Hno3, *Geophys Res Lett*, *22*(14), 1877-1880.

Comstock, J. M., T. P. Ackerman, and D. D. Turner (2004), Evidence of high ice supersaturation in cirrus clouds using ARM Raman lidar measurements, *Geophys Res Lett*, *31*(11).

Cziczo, D. J., D. M. Murphy, P. K. Hudson, and D. S. Thomson (2004), Single particle measurements of the chemical composition of cirrus ice residue during CRYSTAL-FACE, *J Geophys Res-Atmos*, *109*(D4).

Cziczo, D. J., P. J. DeMott, C. Brock, P. K. Hudson, B. Jesse, S. M. Kreidenweis, A. J. Prenni, J. Schreiner, D. S. Thomson, and D. M. Murphy (2003), A method for single particle mass spectrometry of ice nuclei, *Aerosol Sci Tech*, *37*(5), 460-470.

Delval, C., and M. J. Rossi (2005), Influence of monolayer amounts of HNO3 on the evaporation rate of H2O over ice in the range 179 to 208 K: A quartz crystal microbalance study, *J Phys Chem A*, *109*(32), 7151-7165.

DeMott, P. J., D. J. Cziczo, A. J. Prenni, D. M. Murphy, S. M. Kreidenweis, D. S. Thomson, R. Borys, and D. C. Rogers (2003a), Measurements of the concentration and composition of nuclei for cirrus formation, *P Natl Acad Sci USA*, *100*(25), 14655-14660.

DeMott, P. J., K. Sassen, M. R. Poellot, D. Baumgardner, D. C. Rogers, S. D. Brooks, A. J. Prenni, and S. M. Kreidenweis (2003b), African dust aerosols as atmospheric ice nuclei, *Geophys Res Lett*, *30*(14).

Froyd, K. D., D. M. Murphy, T. J. Sanford, D. S. Thomson, J. C. Wilson, L. Pfister, and L. Lait (2009), Aerosol composition of the tropical upper troposphere, *Atmos Chem Phys*, *9*(13), 4363-4385.

Fueglistaler, S., A. E. Dessler, T. J. Dunkerton, I. Folkins, Q. Fu, and P. W. Mote (2009), Tropical Tropopause Layer, *Rev Geophys*, *47*.

Gaines, S. E., S. W. Bowen, R. S. Hipskind, T. P. Bui, and K. R. Chan (1992), Comparisons of the Nasa Er-2 Meteorological Measurement System with Radar Tracking and Radiosonde Data, *J Atmos Ocean Tech*, *9*(3), 210-225.

Gao, R. S., et al. (2006), Measurements of relative humidity in a persistent contrail, *Atmos Environ*, *40*(9), 1590-1600.

Gao, R. S., et al. (2004), Evidence that nitric acid increases relative humidity in low-temperature cirrus clouds, *Science*, *303*(5657), 516-520.

Gierens, K., and P. Spichtinger (2000), On the size distribution of ice-supersaturated regions in the upper troposphere and lowermost stratosphere, *Ann Geophys-Atm Hydr*, *18*(4), 499-504.

Gierens, K., U. Schumann, M. Helten, H. Smit, and A. Marenco (1999), A distribution law for relative humidity in the upper troposphere and lower stratosphere derived from three years of MOZAIC measurements, *Ann Geophys-Atm Hydr*, *17*(9), 1218-1226.

Haag, W., B. Karcher, J. Strom, A. Minikin, U. Lohmann, J. Ovarlez, and A. Stohl (2003), Freezing thresholds and cirrus cloud formation mechanisms inferred from in situ measurements of relative humidity, *Atmos Chem Phys*, *3*, 1791-1806.

Hanisco, T. F., et al. (2007), Observations of deep convective influence on stratospheric water vapor and its isotopic composition, *Geophys Res Lett*, *34*(4).

Heymsfield, A. J., and L. M. Miloshevich (1995), Relative humidity and temperature influences on cirrus formation and evolution: Observations from wave clouds and FIRE II, *J Atmos Sci*, *52*(23), 4302-4326.

Heymsfield, A. J., L. M. Miloshevich, C. Twohy, G. Sachse, and S. Oltmans (1998), Upper-tropospheric relative humidity observations and implications for cirrus ice nucleation, *Geophys Res Lett*, *25*(9), 1343-1346.

Immler, F., K. Kruger, S. Tegtmeier, M. Fujiwara, P. Fortuin, G. Verver, and O. Schrems (2007), Cirrus clouds, humidity, and dehydration in the tropical tropopause layer observed at Paramaribo, Suriname (5.8 degrees N, 55.2 degrees W), *J Geophys Res-Atmos*, *112*(D3).

Jensen, E. J., and L. Pfister (2005), Implications of persistent ice supersaturation in cold cirrus for stratospheric water vapor, *Geophys Res Lett*, *32*(1).

Jensen, E. J., L. Pfister, A. S. Ackerman, A. Tabazadeh, and O. B. Toon (2001a), A conceptual model of the dehydration of air due to freeze-drying by optically thin, laminar cirrus rising slowly across the tropical tropopause, *J Geophys Res-Atmos*, *106*(D15), 17237-17252.

Jensen, E. J., O. B. Toon, S. Kinne, G. W. Sachse, B. E. Anderson, K. R. Chan, C. H. Twohy, B. Gandrud, A. Heymsfield, and R. C. Miake-Lye (1998), Environmental conditions required for contrail formation and persistence, *J Geophys Res-Atmos*, *103*(D4), 3929-3936.

Jensen, E. J., O. B. Toon, S. A. Vay, J. Ovarlez, R. May, T. P. Bui, C. H. Twohy, B. W. Gandrud, R. F. Pueschel, and U. Schumann (2001b), Prevalence of ice-supersaturated regions in the upper troposphere: Implications for optically thin ice cloud formation, *J Geophys Res-Atmos*, *106*(D15), 17253-17266.

Jensen, E. J., L. Pfister, T. Bui, A. Weinheimer, E. Weinstock, J. Smith, J. Pittman, D. Baumgardner, P. Lawson, and M. J. McGill (2005), Formation of a tropopause cirrus layer observed over Florida during CRYSTAL-FACE, *J Geophys Res-Atmos*, *110*(D3).

Karcher, B., and W. Haag (2004), Factors controlling upper tropospheric relative humidity, *Ann Geophys-Germany*, *22*(3), 705-715.

Koop, T., B. P. Luo, A. Tsias, and T. Peter (2000), Water activity as the determinant for homogeneous ice nucleation in aqueous solutions, *Nature*, *406*(6796), 611-614.

Kramer, M., et al. (2009), Ice supersaturations and cirrus cloud crystal numbers, *Atmos Chem Phys*, *9*(11), 3505-3522.

Mohler, O., S. Benz, H. Saathoff, M. Schnaiter, R. Wagner, J. Schneider, S. Walter, V. Ebert, and S. Wagner (2008), The effect of organic coating on the heterogeneous ice nucleation efficiency of mineral dust aerosols, *Environ Res Lett*, *3*(2).

Mohler, O., et al. (2003), Experimental investigation of homogeneous freezing of sulphuric acid particles in the aerosol chamber AIDA, *Atmos Chem Phys*, *3*, 211-223.

Murphy, D. M. (2005), Commentary on "Measurements of ice supersaturations exceeding 100% at the cold tropical tropopause" by E. Jensen et al, *Atmos. Chem. Phys. Discuss.*, *5*(2), 2463-2468.

Murphy, D. M., and T. Koop (2005), Review of the vapour pressures of ice and supercooled water for atmospheric applications, *Q J Roy Meteor Soc*, *131*(608), 1539-1565.

Murray, B. J. (2008), Inhibition of ice crystallisation in highly viscous aqueous organic acid droplets, *Atmos Chem Phys*, *8*(17), 5423-5433.

Murray, B. J., D. A. Knopf, and A. K. Bertram (2005), The formation of cubic ice under conditions relevant to Earth's atmosphere, *Nature*, *434*(7030), 202-205.

Ovarlez, J., J. F. Gayet, K. Gierens, J. Strom, H. Ovarlez, F. Auriol, R. Busen, and U. Schumann (2002), Water vapour measurements inside cirrus clouds in Northern and Southern hemispheres during INCA, *Geophys Res Lett*, *29*(16).

Peter, T., C. Marcolli, P. Spichtinger, T. Corti, M. B. Baker, and T. Koop (2006), When dry air is too humid, *Science*, *314*(5804), 1399-+.

Scott, S. G., T. P. Bui, K. R. Chan, and S. W. Bowen (1990), The Meteorological Measurement System on the Nasa Er-2 Aircraft, *J Atmos Ocean Tech*, *7*(4), 525-540.

Shilling, J. E., M. A. Tolbert, O. B. Toon, E. J. Jensen, B. J. Murray, and A. K. Bertram (2006), Measurements of the vapor pressure of cubic ice and their implications for atmospheric ice clouds, *Geophys Res Lett*, *33*(17).

Smith, J. B., E. J. Hintsa, N. T. Allen, R. M. Stimpfle, and J. G. Anderson (2001), Mechanisms for midlatitude ozone loss: Heterogeneous chemistry in the lowermost stratosphere?, *J Geophys Res-Atmos*, *106*(D1), 1297-1309.

Webster, C. R., and A. J. Heymsfield (2003), Water isotope ratios D/H, O-18/O-16, O-17/O-16 in and out of clouds map dehydration pathways, *Science*, *302*(5651), 1742-1745.

Weinstock, E. M., J. B. Smith, D. Sayres, J. V. Pittman, N. Allen, and J. G. Anderson (2006a), Measurements of the total water content of cirrus clouds. Part II: Instrument performance and validation, *J Atmos Ocean Tech*, *23*(11), 1410-1421.

Weinstock, E. M., et al. (2006b), Measurements of the total water content of cirrus clouds. Part I: Instrument details and calibration, *J Atmos Ocean Tech*, *23*(11), 1397-1409.

Weinstock, E. M., et al. (2009), Validation of the Harvard Lyman-alpha in situ water vapor instrument: Implications for the mechanisms that control stratospheric water vapor, *J Geophys Res-Atmos*, *114*.

Zobrist, B., C. Marcolli, D. A. Pedernera, and T. Koop (2008), Do atmospheric aerosols form glasses?, *Atmos Chem Phys*, *8*(17), 5221-5244.

Chapter 7

Mechanisms of Midlatitude Ozone Loss: Heterogeneous Chemistry in the Lowermost Stratosphere?

7.1 Introduction

We use simultaneous, high-resolution observations of ClO, H_2O, tropopause height, particle reactive surface area, and ice saturation, obtained from the NASA ER-2 aircraft in order to address the question of midlatitude ozone erosion by chlorine free radical catalysis. The objective is to test the hypothesis that the existence of cirrus clouds or cold aerosols in the first few kilometers above the tropopause at midlatitudes is responsible for increasing the ratio of chlorine free radicals to total inorganic chlorine, thus amplifying the rate of catalytic ozone destruction. The observations reveal a sharp decrease in ice saturation frequency at the tropopause, a marked degree of under-saturation just above the tropopause, a corresponding sharp gradient in the product of cold aerosol reactive surface area and reaction probability, γS_a, and, finally, no significant enhancement of ClO concentrations in the region immediately above

the tropopause. As a whole, these results suggest that the observed midlatitude ozone losses are not controlled in situ by the mechanism of cirrus cloud and/or cold aerosol enhancement of chlorine radicals in the vicinity of the tropopause.

Losses of ozone over midlatitudes of the northern hemisphere are well documented in the scientific literature [*Logan et al.*, 1999; *Randel et al.*, 1999]. Accurately defining secular trends in the vertical distribution of ozone, and establishing the mechanism responsible for the observed losses are two dominant and enduring issues. Considerable progress has recently been made with respect to determining ozone trends. In an analysis of global ozonesonde stations, [*Logan et al.*, 1999] report "significant decreases in stratospheric ozone at all stations in middle and high latitudes of the northern hemisphere from 1970 to 1996, with the largest decreases located between 12 and 21 km, and trends of −3 to −10%/decade near 17 km." They report the largest ozone losses in the late winter and spring and in the lowermost stratosphere. These results are summarized in Figure 7.1, which displays the seasonal mean profiles for sonde stations between 36° – 53°N calculated from data collected from 1970 to 1996. [*Bojkov and Fioletov*, 1997] reduce the uncertainty of calculated trends in the lowermost stratosphere by analyzing ozone changes as a function of altitude from the locally determined tropopause. Their analysis shows that negative stratospheric trends become significant 1 − 2 km above the tropopause. In the present analysis, it is the proposed mechanisms responsible for the observed ozone erosion that concern us.

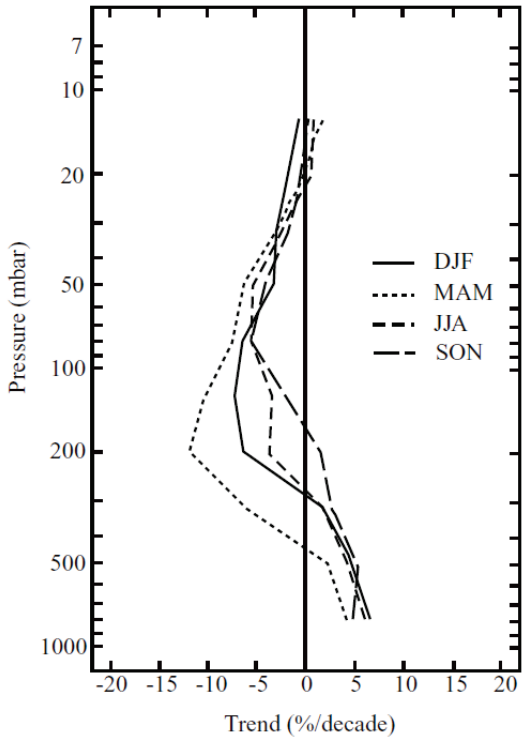


Figure 7.1. Seasonal mean ozone trend profiles for ozonesonde stations located between 36° − 53°N. Adapted from *Logan et al.* [1999], Figure 13.

The changes to midlatitude ozone are largely centered in the "lowermost stratosphere" [*Holton et al.*, 1995], the region bounded at the top by the potential temperature surface corresponding to the tropical tropopause (~380 K) and at the lower extreme by the local tropopause. Figure 7.2a shows a superposition of the mean annual midlatitude ozone trend, calculated by [*Logan et al.*, 1999], with a meridional cross section of the structure and dynamics of the troposphere and stratosphere in a framework of pressure, altitude, and potential temperature.

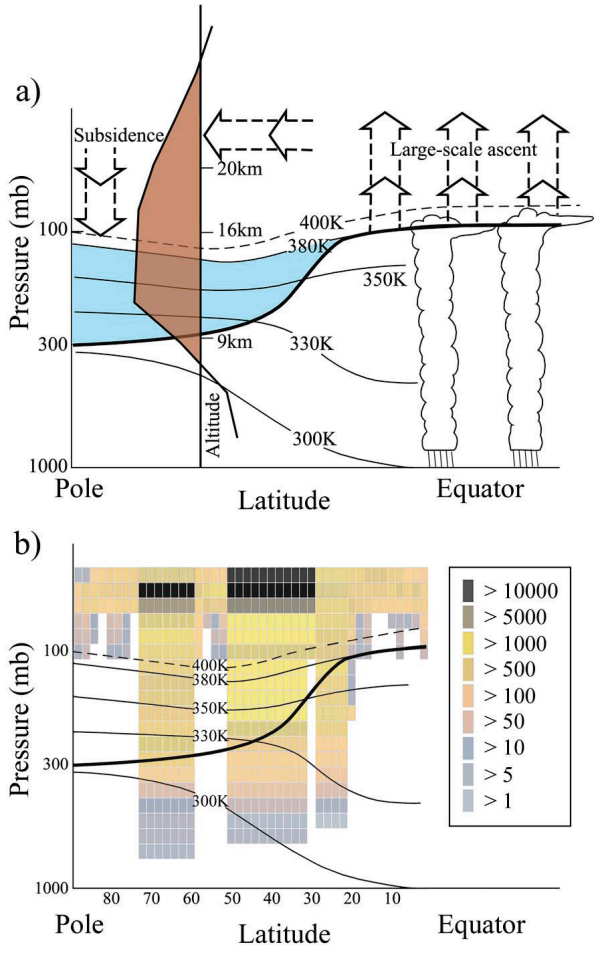


Figure 7.2. Panel (a) shows a superposition of the annual mean midlatitude ozone trend (shaded red) on a meridional cross section of the structure and dynamics of the stratosphere/ troposphere. The ozone loss profile is located using the relationship between pressure, altitude, and potential temperature at midlatitudes. The "lowermost stratosphere" (shaded blue) is the region between the tropopause (thick black line) and the 380 K surface. The "overworld" is the region above 380 K. Panel (b) illustrates the effective sampling area of the ER-2 during the STRAT and POLARIS missions. The structural schematic of the stratosphere and troposphere is included for reference. The large variability in the position of the subtropical jet and the "tropopause break" increases our effective atmospheric sampling range from ~7°N latitude, centered around Moffett Field, CA (37°N), to ~20°N. The color bar denotes the number of observations in each altitude/ latitude bin.

The predominance of ozone loss in the lowermost stratosphere presents a problem with both chemical and dynamical components because of the long

chemical lifetime of ozone in this region. Given that residence times in this region are ~6 months or less [*Hintsa et al.*, 1998; *Rosenlof and Holton*, 1993], that the region is dynamically coupled to high latitudes, and that ozone is rapidly destroyed at high latitudes in the winter and early spring, these coupled chemical and dynamical considerations are almost certainly important. However, the emphasis in this analysis is on evaluating the chemical mechanisms occurring in situ at midlatitudes that could be responsible for the observed changes utilizing the high spatial and temporal resolution observations obtained primarily from the NASA ER-2. Specifically, we consider the proposal by *Solomon et al*. [1997] and [1998] that cirrus clouds and/or cold aerosols act to increase the ratio of halogen radicals to total inorganic halogens enough to make a significant contribution to the catalytic destruction of ozone at midlatitudes. Thus, we 1) analyze the water-temperature relationships in this region that define the existence of cirrus clouds and the prevalence of cold liquid aerosols, 2) review the heterogeneous processes that alter the radical-molecule partitioning in this region, and 3) examine observations of the free radicals that rate limit the catalytic ozone loss processes.

7.2 Chemical Processes of Rate-Limiting Radicals

The catalytic loss of ozone in the lower stratosphere at midlatitudes has been shown to be controlled by a combination of cycles involving HO_2, ClO, BrO, and NO_2 [*Stimpfle et al.*, 1999; *Wennberg et al.*, 1994]. In particular, the catalytic destruction of ozone is rate-limited by

$HO_2 + O_3 \rightarrow OH + O_2 + O_2$ in the HO_x cycle (7.1)

$NO_2 + O \rightarrow NO + O_2$ in the NO_x cycle (7.2)

and

$ClO + O \rightarrow O_2$

$ClO + HO_2 \rightarrow HOCl + O_2$

$ClO + NO_2 \rightarrow ClONO_2$

$ClO + BrO \rightarrow Cl + Br + O_2$

$BrO + HO_2 \rightarrow HOBr + O_2$

$BrO + O \rightarrow Br + O_2$

$BrO + NO_2 \rightarrow BrONO_2$ in the ClO_x/BrO_x cycles. (7.3)

In situ observations of the rate-limiting free radicals in the middle/lower stratosphere by the ER-2 have led to the following conclusions:

1. The dominant loss mechanism for midlatitude ozone at altitudes ≤21 km, the region of interest here, is rate-limited by hydrogen radicals, as reported by [*Wennberg et al.*, 1994]. There is, however, a large seasonal variation in the fractional contribution of the HO_x ($OH + HO_2$) catalyzed loss process. In particular, because of the demonstrated NO_x ($NO_x \equiv NO + NO_2$) control over the fraction of HO_x tied up in the rate limiting radical, HO_2, [*Hanisco et al.*, 2001; *Lanzendorf et al.*, 2001; *Wennberg et al.*, 1994], and because of the strong heterogeneous control over the NO_x to NO_y ratio [*Fahey and Kawa*, 1993], the system is both extremely nonlinear and highly sensitive to the balance between insolation and the rate of N_2O_5 hydrolysis, the heterogeneous reaction that dominates the conversion of NO_x to NO_y ($NO_y \equiv HNO_3 + HNO_4 + 2N_2O_5 + NO_x + NO_3$

+ $ClONO_2$ + $BrONO_2$ +…) .

2. Halogen radicals play a significant role in this region and they too are highly sensitive to latitude, altitude, and season. This is, of course, most dramatically evident at high latitudes in the winter and early spring [*Brune et al.*, 1991; *Toohey et al.*, 1993], when halogen-catalyzed ozone loss not only exceeds other catalytic cycles, but is amplified such that ozone lifetimes drop from months to weeks [*Anderson et al.*, 1991]. In sharp contrast to the late winter/early spring conditions, high-latitude summer conditions reflect the dominance of gas phase reactions coupled with 24-hour illumination that serves to amplify the concentrations of NO and NO_2 [*Del Negro et al.*, 1999; *Perkins et al.*, 2001], thereby suppressing the concentrations of the halogen rate-limiting radicals, particularly ClO [*Stimpfle et al.*, 1999].
3. The role of nitrogen radicals shifts from the primary “buffering agent” at subtropical and mid/high latitudes, where the primary rate-limiting radical concentrations, [HO_2] and [ClO], are the dominant catalytic agents, to the primary "catalytic agent" at high latitudes in summer where intense and nearly continuous insolation overwhelms the heterogeneous loss of NO_x to HNO_3 and leaves NO_2 the primary rate-limiting radical in the photochemical destruction of ozone.

Measured concentrations of the radicals involved in each of the rate-limiting steps defined in equations 7.1 − 7.3 above, permit calculation of the

lifetime of ozone against catalytic destruction by the dominant cycles. This is done to first order by dividing the ozone concentration in the midlatitude lower stratosphere by the sum of the rate-limiting steps defining the loss rate of odd oxygen (O_3 and O):

$$-d[O_3]/dt = 2k_{NO2+O}[NO_2][O] + 2k_{HO2+O3}[HO_2][O_3] \} \text{ NO}_x \text{ \& HO}_x \text{ catalysis}$$
$$+ 2k_{ClO+O}[ClO][O] + 2k_{ClO+HO2}[ClO][HO_2]$$
$$+ 2k_{BrO+O}[BrO][O] + 2k_{BrO+HO2}[BrO][HO_2] \} \text{ ClO}_x \text{ \& BrO}_x \text{ catalysis}$$
$$+ 2k_{ClO+BrO}[ClO][BrO]$$
$$+ \text{minor terms involving } ClONO_2 \text{ and } BrONO_2 \text{ photolysis.} \quad (7.4)$$

Outside of high-latitude winter/spring, this yields a lifetime for ozone against photochemical loss in the lowermost stratosphere in the range of 1 year to many years, depending on latitude and season. This is considerably longer than estimates of the timescale for transport. Because we are focused here on decade-scale changes in the ozone concentration in the lower stratosphere, there are three possibilities to describe the observed decreases in ozone, and that may impact future concentrations:

1. Changes to the chemical composition of the lowermost stratosphere have occurred such that the rate of ozone loss by the catalytic cycles in Equation (7.4) has increased, driving down ozone concentrations in this region.
2. The dynamical structure of the atmosphere has changed such that ozone is removed by transport processes more rapidly from the midlatitude lowermost stratosphere or is supplied more slowly relative to conditions in

previous decades.

3. Ozone is destroyed elsewhere in the system, and those volume elements depleted in ozone are transported into the midlatitude lowermost stratosphere.

Although each of these possibilities requires careful consideration, we restrict our discussion here to the first of the three.

7. 3 Susceptibility of Midlatitude Lower Stratospheric Free Radical Concentrations to Cirrus Clouds

[*Borrmann et al.*, 1996; *Borrmann et al.*, 1997] propose that cirrus cloud surfaces in the vicinity of the tropopause, composed of water ice or liquid sulfuric acid/water supercooled droplets, might activate chlorine via heterogeneous processing, and subsequently lead to ozone depletion. [*Solomon et al.*, 1997] offer a more detailed examination of the chemistry of the tropopause region. Using Stratospheric Aerosol and Gas Experiment II (SAGE) satellite observations and a modeling analysis, they advance the hypothesis that heterogeneous reactions on cirrus clouds and/or liquid aerosols in the vicinity of the tropopause could enhance local ClO mixing ratios by up to a factor of 30 near the tropopause at midlatitudes. They state that if cirrus clouds occur with sufficient frequency and spatial extent, they could considerably influence not only the chemical composition, but also the loss rate of ozone in the region near the tropopause. [*Solomon et al.*, 1997] infer stratospheric cirrus cloud frequencies and surface

areas from satellite observations of cloud optical depths and a climatological, or seasonally averaged, tropopause altitude, and then calculate the episodic enhancements of ClO that result from these surface areas using a two-dimensional numerical model of the chemistry and dynamics of the atmosphere. The key heterogeneous reactions that occur on the inferred cirrus are,

$$HCl + ClONO_2 \xrightarrow{cirrus} Cl_2 + HNO_3$$

$$HCl + HOCl \xrightarrow{cirrus} Cl_2 + H_2O$$

$$H_2O + ClONO_2 \xrightarrow{cirrus} HOCl + HNO_3$$

$$HCl + HOBr \xrightarrow{cirrus} H_2O + BrCl.$$

Given that ClO is the rate-limiting radical in halogen-catalyzed ozone loss, the issue centers on whether these heterogeneous reactions amplify ClO concentrations in the vicinity of the tropopause enough to significantly affect ozone loss in this region. Specifically, the integrated enhancement of ClO concentrations must be sufficient to substantially increase the photochemical loss rate of ozone, and make the resulting photochemical lifetime of ozone competitive with transport in and out of the lowermost stratosphere.

Figure 7.3a is a reproduction of [*Solomon et al.*, 1997], Figure 7, and summarizes their principal results. This figures captures the modeled, oscillatory behavior of ClO just above the tropopause at midlatitudes that results from estimates of cirrus cloud/reactive surface area occurrence based on satellite observations. These results show both the magnitude and frequency of ClO enhancements required to account for the observed ozone trends. The altitude dependence of the cirrus-enhanced ClO from [*Solomon et al.*, 1997] is shown in

Figure 7.3b for summer midlatitudes (49°N). The figure indicates that diurnally averaged ClO mixing ratios reach 25 parts per trillion by volume (pptv), which translates into noontime enhancements of ClO of approximately 50 pptv.

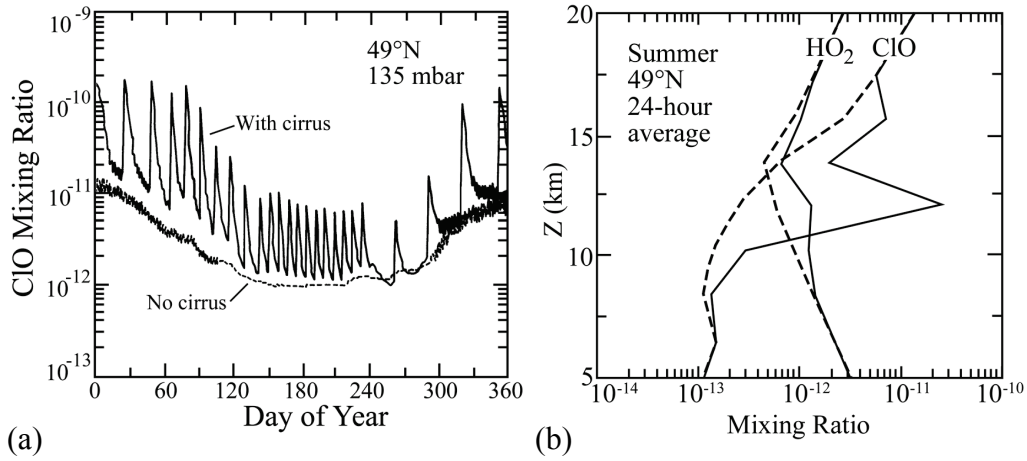


Figure 7.3. Panel (a) shows ClO mixing ratio in the tropopause region calculated by the [*Solomon et al.*, 1997] model. The solid line shows the modeled response of ClO to the presence of inferred cirrus clouds. Adapted from [*Solomon et al.*, 1997], Figure 7. Panel (b) shows diurnally averaged ClO and HO_2 profiles from 5 − 20 km at 49°N in summer calculated by the model. Solid and dashed lines show calcuations with and without cirrus clouds, respectively. Adapted from [*Solomon et al.*, 1997], Figure 14.

We turn now to the question of what the high-resolution observations obtained from the NASA ER-2 contribute to this issue: Is there evidence of cirrus and/or significant ClO enhancements in the midlatitude lowermost stratosphere in flight data acquired during a series of campaigns extending from 1993 – 1997?

7.4 Evidence From In Situ Aircraft Observations

The ER-2 aircraft provides a platform for simultaneous high-resolution, in situ measurement of a host of chemical species and families. The following

analysis uses data from 63 flights obtained during the Stratospheric Tracers of Atmospheric Transport (STRAT) and Photochemistry of Ozone Loss in the Arctic Region in Summer (POLARIS) missions. These missions covered the period from May 1995 through September 1997, and were flown out of NASA Ames Research Center, Moffett Field, California (37°N); Fort Wainwright Army Base, Fairbanks, Alaska (65°N); and Barbers Point Naval Air Station, Barbers Point, Hawaii (24°N). Figure 7.2b shows the sampling frequency in the upper troposphere and lower stratosphere during these missions, with a structural schematic of the stratosphere and troposphere superimposed for reference. Although all midlatitude ascents, descents, and stairstep flights from Moffett Field are geographically centered about 37°N, with a range of ~7°, our effective atmospheric sampling is significantly wider, with an estimated range of ~20°, because of the large day-to-day variability in the latitudinal position of the subtropical jet and the tropopause.

7.4.1 Cloud Occurrence Frequency

The in situ data allow for precise determination of the occurrence frequency of ice saturation, measurement of its vertical extent, and a broader characterization of the airmass during ice-supersaturated events. The flight data were binned by latitude, with 0° − 30°N representative of equatorial/ subtropical data; 30° − 50°N midlatitude data; and 50° − 90°N polar data. Figure 7.4 shows the percentage of observations of ice saturation, evaluated for midlatitude data, as a function of altitude referenced to the local tropopause.

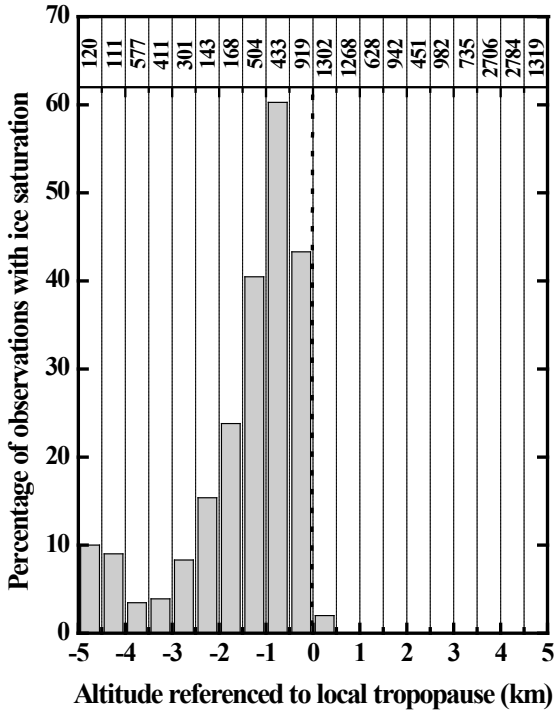


Figure 7.4. Histogram displaying the percentage of observations of ice saturation, evaluated for midlatitude data only (30° − 50°N), as a function of altitude referenced to the measured local tropopause. The total number of observations for each 0.5 km altitude bin is shown in the corresponding box at the top.

Because the height of the tropopause can vary dramatically over short timescales, we reference the altitude associated with each data point to the local tropopause. Tropopause heights are reported for every flight (B. Gary, personal communication, 1999), and are determined using measurements of temperature profiles and altitude from a combination of data from the Meteorological Measurement System (MMS) [*Scott et al.*, 1990], Microwave Temperature Profiler (MTP) [*Denning*, 1989], and radiosonde data. The tropopause is calculated using the standard World Meteorological Organization (WMO) definition: The tropopause is the lowest altitude at which the thermal lapse rate ($-dT/dz$) is 2 K/km or less for all layers bounded by the candidate tropopause

altitude and all altitudes within the next +2 km. The uncertainty associated with the tropopause calculations is typically ±0.1 − 0.3 km. The number of tropopause values reported per flight depends on local meteorology and the flight profile, but is typically two, one each for ascent and descent. We divide the data set for each flight accordingly and redefine altitude to kilometers above (plus) or below (minus) the observed tropopause position, and then sort the data into 0.5 km altitude bins from −5 km to +5 km about the measured tropopause position.

In order to ascertaion cloud occurrence, we examine the frequency of water vapor/ice saturation. Thus, in this analysis, the ratio of data points showing saturation to the total number of data points in each 0.5 km altitude bin defines the frequency of cloud occurrence. This is an extremely generous criterion, as significant supersaturation may be required to actually form clouds and maintain cloud presence for extended periods. Furthermore, a data point is taken to indicate saturation if the measured temperature is up to 1 K warmer than the water vapor/ice saturation temperature calculated from the measured water vapor mixing ratio and measured pressure. The equation used to determine the saturation tempersture is derived from the water vapor saturation pressure over ice data in the *CRC Handbook of Chemistry and Physics* [*Dessler et al.*, 1995; *Weast et al.*, 1970]. Again, this is a generous criterion for saturation. It fully envelops the uncertainty associated with the in situ temperature and water vapor measurements. At saturation, 1 K at stratospheric temperatures and pressures corresponds to a change of approximately 15% in water vapor. The Harvard water vapor instrument (HWV-E) reports data with an estimated ±5% accuracy [*Hintsa*

et al., 1999] (see also Chapters 2 and 3), while the accuracy of temperature and pressure are ~1 K and ±0.3 hPa, respectively.

The midlatitude histogram in Figure 7.4 shows proxy cloud occurrence frequencies building with altitude in the upper troposphere, reaching 60% just below the tropopause, but dropping abruptly to 2% within the first 0.5 km above the tropopause. In over 1200 midlatitude measurements in the next highest 0.5 km altitude bin there is no evidence of saturation. The same is true for all subsequent altitude bins. Results from the ER-2 do not support the existence of stratospheric cirrus.

Cirrus clouds have occasionally been observed in the stratosphere, but these observations have been sporadic, just above the tropopause, and largely in the tropics [*Beyerle et al.*, 1998; *Jensen et al.*, 1996; *Rinsland et al.*, 1998; *Winker and Trepte*, 1998]. [*Poulida et al.*, 1996] observed overshoot from an intense thunderstorm leading to clouds in the midlatitude stratosphere. However, these clouds were accompanied by tropospheric air with low ozone, and their lifetime was likely short. Thus, historic observations are consistent with our finding that ice saturation is extremely rare in the midlatitude stratosphere at altitudes >500 m above the tropopause.

To further explore the possibility of ice saturation in the lowermost stratosphere, we used the time and position of selected data points/air parcels and the National Centers for Environmental Prediction (NCEP) analyses of balanced winds to calculate air parcel back-trajectories. The trajectory output includes the position and temperature of the specified air parcel for the 10 days prior to the

aircraft encounter. Whether or not the air parcel saw saturation conditions in the past can be determined assuming the water vapor mixing ratio of the air parcel remains constant over the 10-day period.

We ran back trajectories on data points selected for their saturation potential, that is, data points in the lowermost stratosphere, between 1 − 3 km above the local tropopause, with a separation of 10 K or less between the measured temperature and the calculated temperature of saturation. About 0.5% of the data points in the selected altitude range met these criteria, and the back trajectories showed no evidence of prior saturation. In every case the degree of separation between the back trajectory temperature and the temperature of saturation was significantly larger than the temperature fluctuations along the trajectory.

7.4.2 Water Vapor and Temperature in the Lower Stratosphere

The absence of ice saturation in the lowermost stratosphere may be understood by examining the behavior of water vapor and temperature in the upper troposphere/lower stratosphere. Relative humidity falls off by an order of magnitude immediately above the tropopause, from values close to 100% in the upper troposphere to values below 10% in the first kilometer of the lower stratosphere, except where the upper troposphere is already dry. Figure 7.5 shows vertical profiles of both the measured water vapor mixing ratio and the calculated saturation mixing ratio. These data, acquired on descent into Moffett Field, California, on February 1, 1996, are representative of midlatitude conditions. The

dryness of air within 1 km above the tropopause explains the scant supply of points fulfilling our back-trajectory criteria. The plot captures the following points, 1) as altitude increases above the tropopause (within the first 0.5 km), there is a transition to much lower water vapor mixing ratios, and 2) the measured temperature at the tropopause usually remains constant or increases, driving the lowermost stratosphere further from saturation.

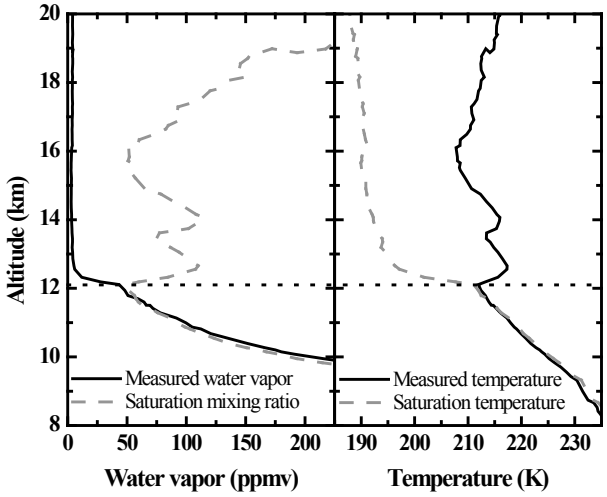


Figure 7.5. Conditions near the tropopause for descent into Moffett Field, California, on February 1, 1996. Measured temperature and water vapor (HWV-E) are solid curves. Water vapor saturation mixing ratio and temperature of saturation are dashed curves. The horizontal dotted line indicates the position of the tropopause.

Figure 7.6 consists of a series of histograms showing normalized distributions of the degree of undersaturation, $T_{\text{saturation}} - T_{\text{measured}}$, for altitude bins from 3 km below to 3 km above the tropopause. The number of observations in each degree bin is divided by the total number of observations, and multiplied by 100 to yield occurrence frequency in percent. In the transition region, ±1 km about the tropopause, we increase our altitude resolution to 0.5 km from 1 km.

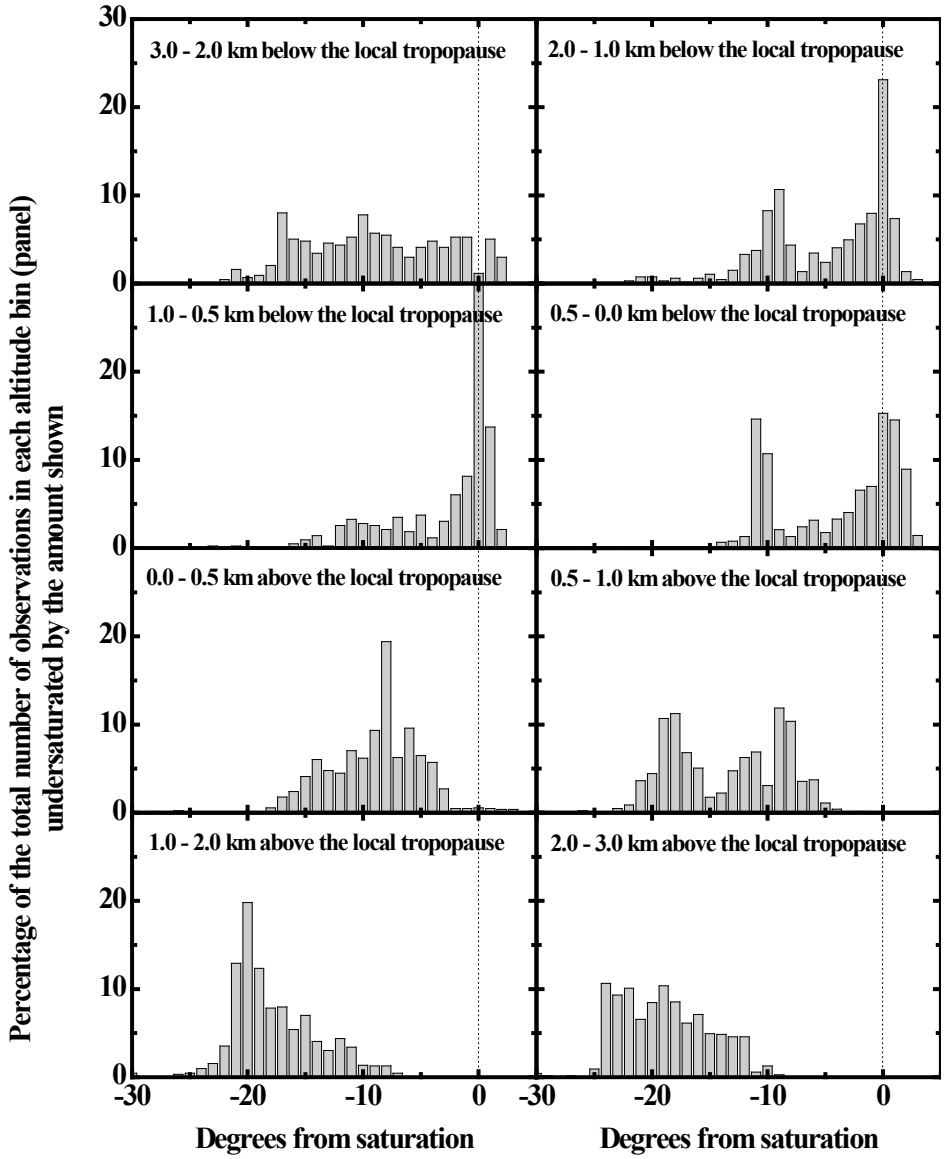


Figure 7.6. Histograms showing normalized distributions of measured degrees from saturation, $T_{\text{saturation}} - T_{\text{measured}}$, for altitude bins ±3 km around the local tropopause. The resolution in the range ±1 km about the tropopause increases from 1 km altitude bins to 0.5 km bins. Only midlatitude data are shown.

Below the tropopause the histograms peak at zero degrees separation, i.e., the most probable temperature is the saturation temperature. However, in the 0.5 km altitude bin just above the tropopause, the peak of the histogram shifts abruptly away from saturation, indicating the sharp onset of undersaturation in the

lowermost stratosphere. The data show that the stratosphere is undersaturated by a minimum of 10 K only 2 km above the tropopause. Figure 7.7 shows the same trend. In this figure all the data in each altitude bin are plotted on a phase diagram of measured water vapor partial pressure versus measured temperature. The water vapor/ice transition is the solid line in each plot. The majority of the data lie along the saturation line just below the tropopause, but fall away from the line just above the tropopause. These plots, in combination with the backtrajectory analysis, underscore the fundamental characteristics of the thin layer immediately above the tropopause, 1) there is a sharp divergence between the observed temperature and the calculated temperature of saturation, and 2) air parcels within 10 K of saturation located in the first 3 km above the tropopause do not approach the saturation condition along backward (or forward) trajectories. In summary, the in situ observations reveal a sharply diminishing probability of ice saturation and clouds in the lower stratosphere, driven by a precipitous decrease in relative humidity.

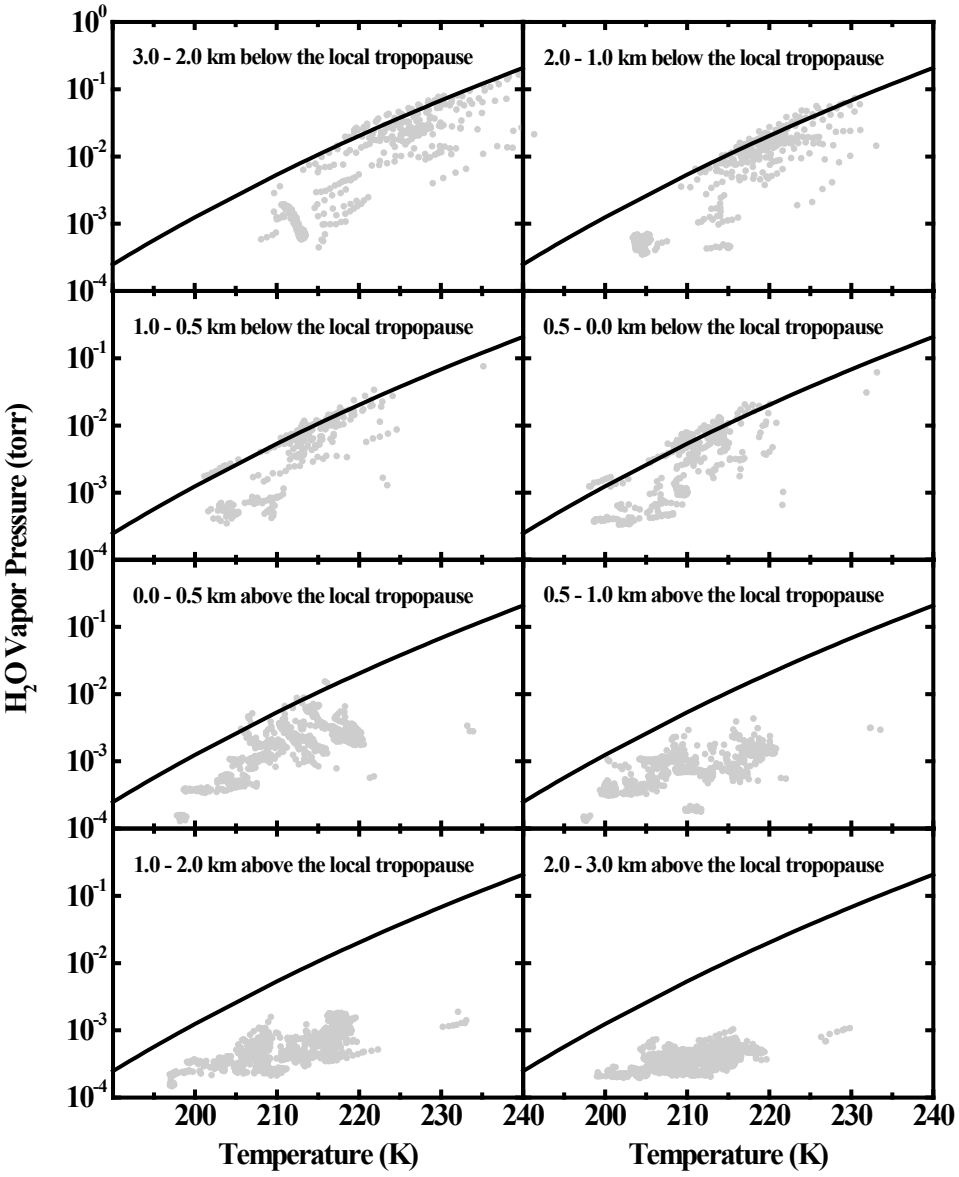


Figure 7.7. Phase diagram displaying the relationship between measured water vapor pressure and temperature referenced to the saturation line (solid line in plots) for altitude bins ranging ±3 km around the local tropopause. The resolution in the range ±1 km about the tropopause increases from 1 km altitude bins to 0.5 km bins. Only midlatitude data are shown.

The lowermost stratosphere, as defined in the introduction, is a region distinct from the rest of the stratosphere, or overworld [*Dessler et al.*, 1995; *Holton et al.*, 1995]. The water vapor mixing ratios observed in the lowermost

stratosphere are often notably higher than those of the overworld [*Pan et al.*, 1997], but this analysis of in situ data shows they are not sufficient for ice saturation, or sustaining cirrus. Adiabatic transport along isentropic surfaces and/or vertical diabatic transport across isentropes to and from the troposphere influences the chemical properties of this region. In particular, water vapor concentrations in this region are consistent with a mixture of dry air from the overworld, determined primarily by the low temperatures encountered at the tropical tropopause, and moister air that has been transported isentropically with either tropical/subtropical or midlatitude origin [*Dessler et al.*, 1995; *Hintsa et al.*, 1998]. The mixing of dehydrated air from the overworld coupled with the temperature trend at the tropopause explains much of the transition in relative humidity at the tropopause, and the observed scarcity of saturation events in the lower stratosphere. It has also been suggested that, regardless of their exact origin, air parcels with high water vapor mixing ratios have often reached saturation temperatures near the tropopause and have undergone dehydration prior to their entry into the lowermost stratosphere [*Hintsa et al.*, 1998].

7.4.3. Discussion of the *Solomon et al.* [1997] Analysis

The success of the [*Solomon et al.*, 1997] model of heterogeneous chemistry on stratospheric cirrus is contingent upon an accurate representation of the frequency of cloud events. They used a seasonal climatology of cloud observations from SAGE II satellite data [*Wang et al.*, 1996] coupled with a climatological tropopause position to determine the frequency of cirrus cloud

events in the midlatitude lowermost stratosphere. Their results yield summer cloud occurrence frequencies on the order of 1 − 10% at altitudes as high as 13 − 15 km, and above the average summer tropopause evaluated at 45°N [*Solomon et al.*, 1997]. The effect of these clouds on modeled ClO was summarized in Figures 7.3a&b.

The high-resolution ER-2 observations suggest a very different picture. First, when keyed to simultaneous observations of tropopause position, the probability for cirrus cloud occurrence in the midlatitude stratosphere drops to a few percent in the first 0.5 km above the tropopause, and to immeasurably low levels above that. Second, the altitude of the tropopause varies markedly on short timescales, with the tropopause altitude shifting by 2 – 3 km about its mean.

Thus we consider the following issue: Suppose the observed high-resolution distribution of cirrus cloud frequency is convolved with a function which represents both the intrinsic vertical resolution of the SAGE II instrument (~1 km full width at half its maximum observational sensitivity) and the natural "climatological smear" resulting from the observed short-term vertical translation of the actual tropopause about its climatological mean. What then, is the inferred cloud occurrence frequency?

To examine this, we derived a spatial weighting function comprised of a Gaussian function, measuring 1 km full width at half maximum, convolved with a normalized histogram of the frequency of tropopause occurrence above and below its average value. The histogram was evaluated using temperature and pressure altitude measurements from a compilation of radiosonde data from the summer

season, and the WMO tropopause definition (P. Newman, private communication, 1999). The data were selected to be coincident with the ER-2 missions, that is, from 1995 to 1997, and were chosen from a narrow band centered at ~49° ±1°N at all longitudes to best compare with the cloud occurrence frequencies inferred by [*Solomon et al.*, 1997] (see their Figures 3 and 7). The subsequent convolution of the spatial weighting function with the observed high-resolution cloud occurrence data from the ER-2 yields an enhanced distribution of cloud frequency in the lower stratosphere.

This convolution sequence is depicted in Figure 7.8. The process conserves the area under the original curve, i.e., the integrated cloud occurrence frequency is a constant. The largest effect results, of course, from the "climatological smear," since that assumption dominates the spatial weighting function. The combined result yields occurrence frequencies above the tropopause markedly greater that those actually observed in the high-resolution data. The low vertical resolution resulting from the combination of the remote sensing technique, and a tropopause height locked to the climatological mean, effectively "creates" reactive surface area in the lower stratosphere.

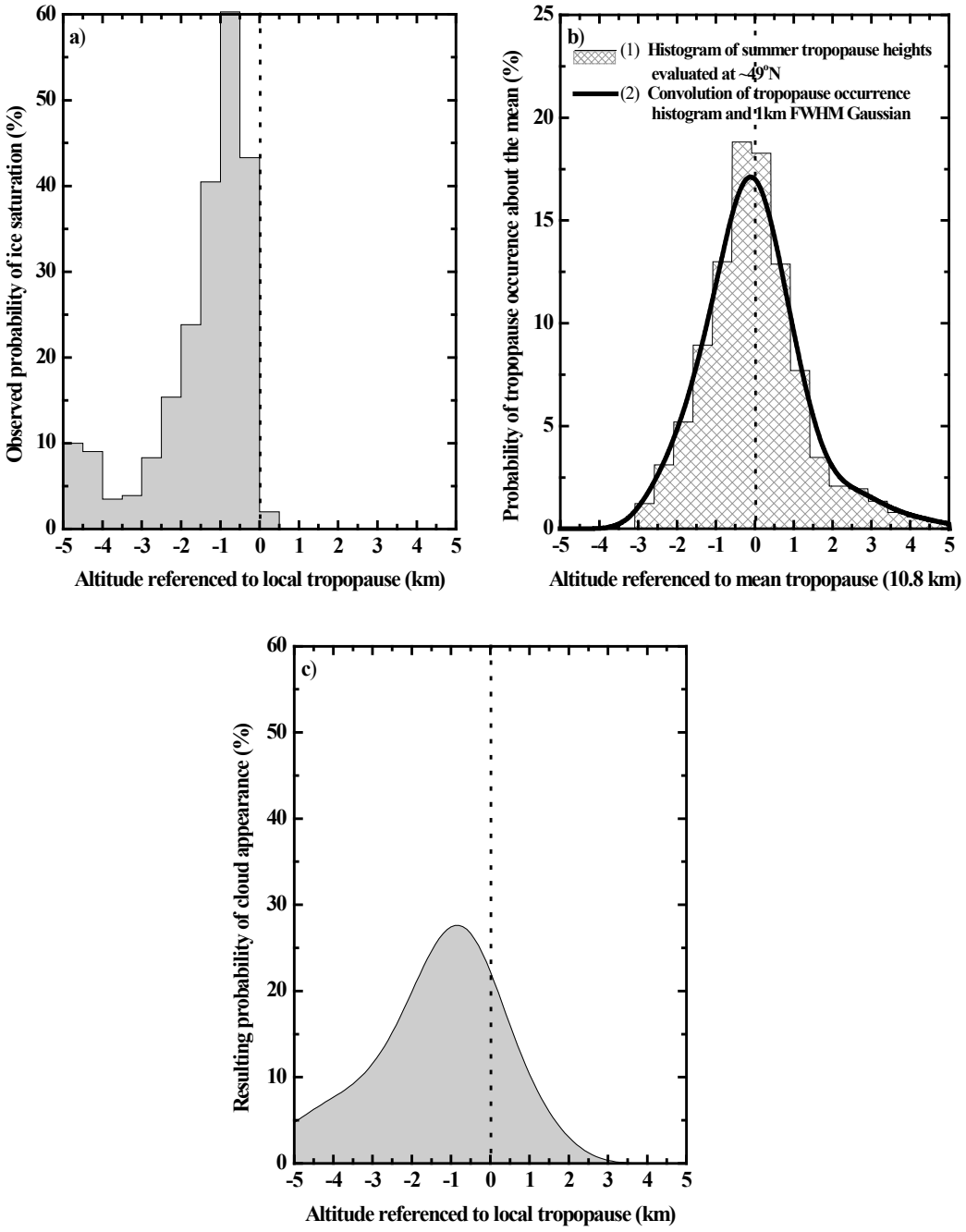


Figure 7.8. The sequence showing (a) the high-resolution ER-2 observations of ice saturation frequency (a proxy for cloud occurrence frequency); (b) the spatial weighting function, evaluated at ~49°N, that results from the convolution of the "climatological smear" resulting from assuming a climatological tropopause, and a Gaussian representing the intrinsic vertical resolution of the SAGE II satellite; and (c) the resulting distribution of cloud occurrence frequency in the lower stratosphere, obtained by the convolution of Figures 7.8a and 7.8b.

In contrast to the radiosonde data centered at 49°N, midlatitude tropopause data from the ER-2 reveal a strong bimodal distribution. This is a result of the fact that nearly all the ER-2 tropopause data are centered at 38°N (above NASA Ames), a latitude range where the tropopause altitude has a strong seasonal dependence, that is, a higher "summer" (subtropical) tropopause, from northward excursions of the subtropical jet, and a lower "winter" (midlatitude) tropopause. The highest tropopause values are found in summer, with some observations showing a tropopause nearly as high as values typically found in the tropics/subtropics (360 − 380 K potential temperature, ~16 km). The bimodal character is reproduced in histograms of NCEP radiosonde data from sites selected for their proximity to NASA Ames. Use of summer ER-2, or radiosonde data centered at 38°N in a calculation similar to that shown in Figure 7.8, further increases the relative cloud occurrence above the tropopause due to the broader distribution of tropopause heights at this latitude.

The seasonal variation of cirrus cloud frequency reported by [*Solomon et al.*, 1997], in particular the increased frequency of cloud observations in summer, is most likely another artifact of using a climatological tropopause. As mentioned above, data from the midlatitude ER-2 flights show that the position of the summer tropopause has a wide dynamic range. Using a fixed climatological tropopause will not reflect the frequently lifted local tropopause values in summer and will consequently overestimate stratospheric cloud events in that season. In contrast, our analysis of the midlatitude ER-2 data showed that the 26 data points that showed evidence of saturation in the first 0.5 km bin above the tropopause

(of a total 1268 observations in that bin) occurred between December and mid-May, and 21 of the 26 occurred between December and the beginning of February. This observation of saturation events in the winter season is consistent with the results of [*Murphy et al.*, 1990]. Their analysis of ER-2 data taken over Moffett Field, California (37°N) and Wallops Island, Virginia (31°N) yields evidence of saturation just above the tropopause in the winter season, but none in summer.

7.4.4. Particle Reactive Surface Area

Particle reactive surface area is critical because the rate constant defining the heterogeneous conversion of inorganic chlorine to free radical form is given by $k_{het} = (v/4)\gamma S_a$, where v is the average molecular speed, γ is the reaction probability per collision, and S_a is the available aerosol/cirrus cloud reactive surface area. Particle surface area was measured by an array of instruments aboard the ER-2 [*Jonsson et al.*, 1995]. The data were again dissected by latitude and altitude in 0.5 km bins keyed to the local tropopause. We averaged all the data in each altitude bin, and also recorded the highest observed value of S_a in each bin. These results are displayed in Figure 7.9. The mean measured particle surface area peaks just below the tropopause, and decreases steadily above it. Similarly, the maximum observed surface area drops dramatically, and the distribution of particle areas narrows sharply at the tropopause.

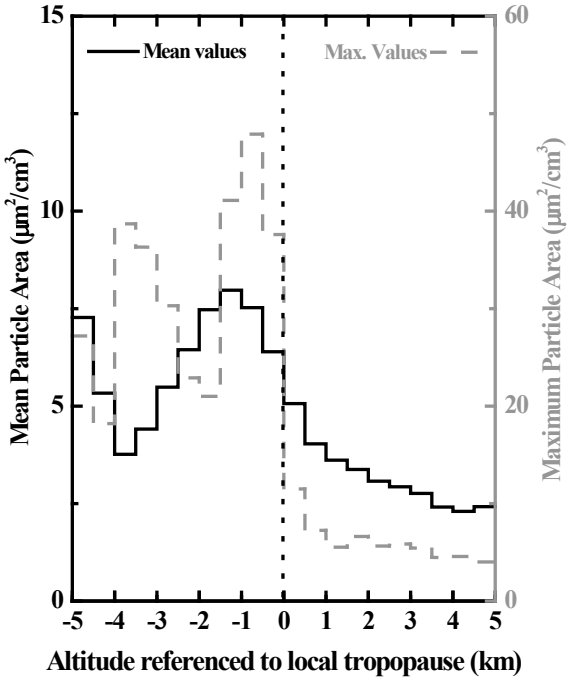


Figure 7.9. Observed particle surface area as a function of altitude referenced to the local tropopause. The mean values for each bin are represented by the solid curve, and the dashed curve is a plot of the maximum observed value in each bin. Note the mean and maximum plots use the left and right y axes, respectively. Also, mean values in this figure and all subsequent figures are the bin-wise averages of individually calculated values.

7.4.5. Reaction Probability on Cold Aerosols

The possibility that heterogeneous reactions occur on aerosol particles, not strictly cirrus clouds, in the lowermost stratosphere is suggested as an alternative mechanism for the ozone loss observed at midlatitudes [*Solomon et al.*, 1996; *Solomon et al.*, 1998]. Aerosol surface area, S_a, is a function of aerosol number density, temperature, pressure, and water vapor mixing ratio. Furthermore, laboratory results defining the sharp increase of γ with decreasing temperature, in combination with model results, suggest that cloud presence is not required for heterogeneous chemistry to be significant [*Michelsen et al.*, 1999].

The ER-2 in situ data allow for direct investigation of this question. We used measurements of temperature, pressure, and water vapor mixing ratio to calculate H_2SO_4 weight percents (adapted from [*Carslaw et al.*, 1995]). These H_2SO_4 weight percents in combination with simultaneous in situ measurements of HCl [*Webster et al.*, 1994], water vapor mixing ratios, particle surface areas, and the parameterization of [Hanson, 1998], allowed us to calculate the corresponding reactive uptake coefficients for the following heterogeneous reactions on sulfuric acid aerosol:

$H_2O + ClONO_2 \xrightarrow{S_a} HOCl + HNO_3$, and

$HCl + ClONO_2 \xrightarrow{S_a} Cl_2 + HNO_3$.

We used measured HCl whenever it was reported. Where measurements were scant, at and below the tropopause, we used a default value of 100 pptv. The resulting data were then binned by latitude and altitude.

Figure 7.10 shows the results of our calculations of sulfuric acid weight percent and γ for the midlatitude data. All the calculated values are plotted as a function of altitude about the local tropopause, with the median values for each bin and the relevant extrema highlighted (minimum values for sulfuric acid weight percent, and maximum values for γ). The sulfuric acid weight percent decreases up to the tropopause and increases rapidly within the first kilometer above it. This is consistent with the increasing relative humidity and low temperatures just below the tropopause and the sudden transition to drier air above. The reaction probabilities, γ, which are extremely sensitive to both temperature and water vapor mixing ratio, also change dramatically as a function

of distance above the tropopause. They are affected by sulfuric acid weight percent because the presence of sulfuric acid "poisons" the key heterogeneous reactions, e.g., [*Kolb et al.*, 1996; *Ravishankara et al.*, 1998] *and references therein*. Our results show that the chlorine activation potential on aerosols falls off by an order of magnitude or more just as the chlorine begins to emerge from the organic source molecules and where calculated ozone trends begin to be statistically significant at 1 − 2 km above the tropopause.

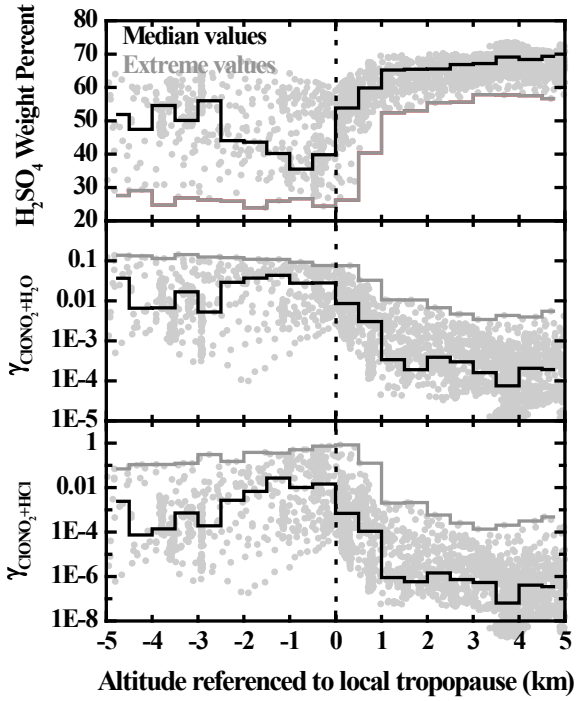


Figure 7.10. Calculated H_2SO_4 weight percents, and reaction probabilities, γ, for the heterogeneous activation of chlorine by the reaction of $ClONO_2$ with H_2O and with HCl as a function of height about the tropopause. These values were calculated using simultaneous observations of temperature, pressure, H_2O, HCl, and calculated H_2SO_4 weight percent. Where measurements of HCl were not available, below and just above the tropopause, we used a constant value of 100 pptv. The bin-wise median and extreme values are the solid curves.

We calculated the time constants for both heterogeneous reactions using the calculated reactive uptake coefficients, γ, measured particle surface areas, S_a, and mean molecular speeds. The first-order time constants for both reactions, shown in Figure 7.11, exhibit an order of magnitude increase in the region just above the tropopause. The minimum calculated time constants for both reactions are significantly longer than the time constant for $ClONO_2$ photolysis >0.5 km above the tropopause. In sharp contrast, [*Solomon et al.*, 1997] estimate the time constants to be on the same order as photolysis, a few hours or less, for the lowermost stratosphere.

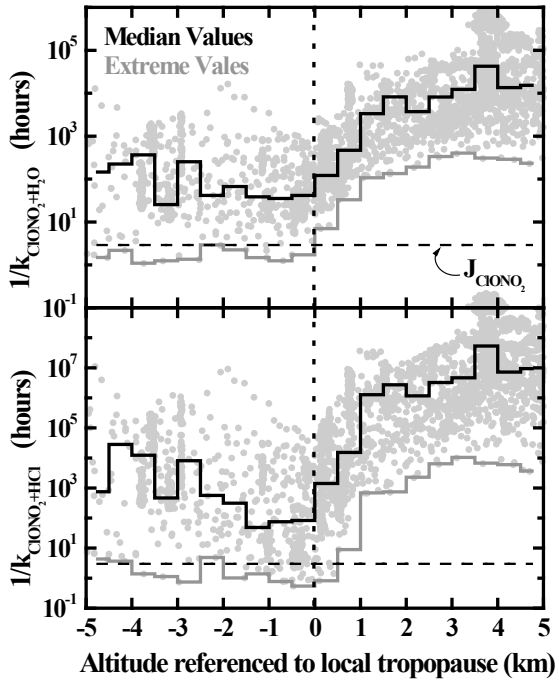


Figure 7.11. Calculated first-order time constants for the heterogeneous activation of chlorine by reaction of $ClONO_2$ with H_2O and HCl as a function of distance from the local tropopause, based on simultaneous observations of H_2O, HCl, and S_a, and calculated γ values.

7.4.6. Observation of ClO in the Vicinity of the Tropopause

A critical question remains: What evidence exists for the presence of enhanced concentrations of the rate limiting chlorine radical, ClO, in the lowermost stratosphere at midlatitudes? To this end, we analyzed ClO data from the most recent ER-2 aircraft missions that carried the Harvard ClO instrument, the Stratospheric Photochemistry, Aerosols and Dynamics Expedition (SPADE), the Airborne Southern Hemisphere Ozone Experiment/Measurements for Assessing the Effects of Stratospheric Aircraft (ASHOE/MAESA), and POLARIS. The ClO instrument has a detection threshold of ~2 pptv in a 30-s sample at cruise altitude, and a detection threshold of ~5 pptv in stair-step flights or ascent/descent flights in the vicinity of the tropopause. In most cases, longer integration times are possible, improving the detection limit. We examined all midlatitude (30° − 50°N/S) ClO data that were within the first +4 km of the measured tropopause, excluding only volume elements that showed the distinct signature of vortex processing, that is, where the simultaneously measured N_2O was less than 250 ppbv. In all, there were 1210 midlatitude data points from 56 flights.

Figure 7.12 shows the results of our analysis. Figure 7.12a shows all the midlatitude data in black, except that of the April 30, 1993, SPADE flight which is discussed below. The data for all latitudes, in total 2661 data points, are shown in the background for comparison. The mean for each 0.5 km bin is plotted in green and the standard error associated with each mean is covered by the size of the dot.

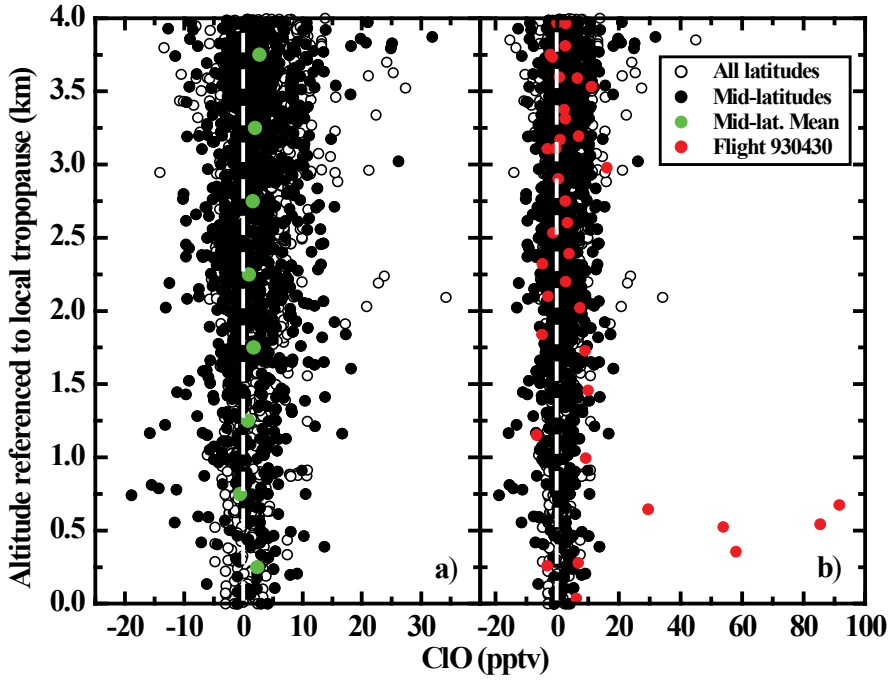


Figure 7.12. Panel (a) shows a plot of ClO, measured aboard the ER-2 during SPADE (except the flight of April 30, 1993), ASHOE/MAESA, and POLARIS, as a function of altitude above the tropopause. The midlatitude (30° − 50°N/S) data are plotted in black; data from all other latitudes are open circles. The midlatitude means, evaluated for each 0.5 km bin, are plotted in green. Panel (b) shows ClO data with the inclusion of the April 30, 1993, flight in red.

The first result that emerges from the data set is that the ClO mole fraction in the vicinity of the tropopause appears evenly distributed about zero. We are unable to distinguish between 0.1 and 1 or 2 pptv because of the detection threshold of the measurement, and the possibility remains that ClO levels are slightly above zero in the lowermost stratosphere. However, we can distinguish between a few pptv and the levels of ClO required to substantially increase the photochemical loss rate of ozone in the critical region from the tropopause to >3 km above it. Note that in the [*Solomon et al.*, 1997], Figure 14, the model with cirrus clouds shows a diurnal average of ~10 pptv ClO at 12 km. The observational dataset does not support ClO enhancements of this order.

The ClO data from the SPADE flights of 1993, during which there was a large enhancement in background aerosol loading from the eruption of Mount Pinatubo, show a single outstanding example of an enhanced ClO layer above the tropopause. [*Keim et al.*, 1996] reported ClO mixing ratios of 90 pptv ~0.5 km above the midlatitude tropopause, and attributed them to reactions on cold aerosols in conjunction with moderate levels of ambient water vapor. Figure 7.12b shows the data from this flight in red. This figure demonstrates 1) that the ClO instrument can detect significant enhancements of ClO near the tropopause, and 2) that this phenomenon was not observed on any other SPADE flight, or in >40 midlatitude flights following the decay of the Pinatubo aerosol. Furthermore, D.W. Toohey (personal communication, 1999) analyzed ClO data from 1991 − 1992 and found no evidence of enhanced ClO near the midlatitude tropopause. If judged by the frequency of in situ observation, the event highlighted in Figure 7.12b must be considered exceedingly rare.

The limited degree of variability that ClO exhibits in the first few kilometers above the tropopause reveals two key aspects about the control of active chlorine in this region. First, the fraction of organic chlorine released into inorganic form is both small and tightly coupled to the long-lived tracer field in the midlatitude tropopause region [*Bonne et al.*, 2000; *Woodbridge et al.*, 1995]. Second, the product of the gradient in maximum reactive surface area, S_a, above the tropopause (Figure 7.9) and the "poisoning" of the key heterogeneous reactions resulting from the increase in H_2SO_4 weight percent in the cold aerosols of this region, strongly constrains the release of ClO from total available chlorine.

7.5. Summary and Conclusions

ER-2 flights extending over a period of 4 years provide simultaneous, high-resolution observations of (1) tropopause height; (2) ice saturation frequency (a proxy for cloud occurrence frequency) in 0.5 km bins from 5 km below the tropopause to 5 km above the tropopause; (3) the degree of undersaturation in the same array of altitude bins; (4) particle reactive surface area density S_a; (5) the reaction probability on cold aerosols, γ; and finally, (6) ClO concentrations. The observations reveal a sharply diminishing probability of observed ice saturation in the lower stratosphere, severely limiting the possibility of sustained cirrus cloud occurrence and/or enhanced aerosol surface area. The observations also show a rapid decrease in γ for the two representative chlorine activation reactions. The decline in heterogeneous reactivity is driven by the sharp decrease in relative humidity and the resultant increase in the sulfuric acid weight percent of the reactive aerosol. In summary, the atmospheric conditions that would enable the heterogeneous activation of chlorine do not exist in the lowermost stratosphere. Furthermore, high-resolution observations of ClO show the absence of substantially enhanced chlorine radical concentrations in the first few kilometers above the local tropopause. The in situ data lead us to conclude that the observed loss of ozone at midlatitudes does not occur locally in the lowermost stratosphere via the heterogeneous conversion of inorganic chlorine into free radical form as a result of cirrus or cold aerosols.

References

Anderson, J. G., D. W. Toohey, and W. H. Brune (1991), Free-Radicals within the Antarctic Vortex - the Role of CFCs in Antarctic Ozone Loss, *Science*, *251*(4989), 39-46.

Beyerle, G., H. J. Schafer, R. Neuber, O. Schrems, and I. S. McDermid (1998), Dual wavelength lidar observation of tropical high-altitude cirrus clouds during the ALBATROSS 1996 campaign, *Geophys Res Lett*, *25*(6), 919-922.

Bojkov, R. D., and V. E. Fioletov (1997), Changes of the lower stratospheric ozone over Europe and Canada, *J Geophys Res-Atmos*, *102*(D1), 1337-1347.

Bonne, G. P., et al. (2000), An examination of the inorganic chlorine budget in the lower stratosphere, *J Geophys Res-Atmos*, *105*(D2), 1957-1971.

Borrmann, S., S. Solomon, J. E. Dye, and B. P. Luo (1996), The potential of cirrus clouds for heterogeneous chlorine activation, *Geophys Res Lett*, *23*(16), 2133-2136.

Borrmann, S., S. Solomon, L. Avallone, D. Toohey, and D. Baumgardner (1997), On the occurrence of ClO in cirrus clouds and volcanic aerosol in the tropopause region, *Geophys Res Lett*, *24*(16), 2011-2014.

Brune, W. H., J. G. Anderson, D. W. Toohey, D. W. Fahey, S. R. Kawa, R. L. Jones, D. S. Mckenna, and L. R. Poole (1991), The Potential for Ozone Depletion in the Arctic Polar Stratosphere, *Science*, *252*(5010), 1260-1266.

Carslaw, K. S., B. P. Luo, and T. Peter (1995), An Analytic-Expression for the Composition of Aqueous HNO3-H2SO4 Stratospheric Aerosols Including Gas-Phase Removal of Hno3, *Geophys Res Lett*, *22*(14), 1877-1880.

Del Negro, L. A., et al. (1999), Comparison of modeled and observed values of NO2 and JNO2 during the Photochemistry of Ozone Loss in the Arctic Region in Summer (POLARIS) mission, *J. Geophys. Res.*, *104*(D21), 26687-26703.

Denning, R., S. Guidero, G. Parks, and B. Gary (1989), Instrument Description of the Airborne Microwave Temperature Profiler, *J. Geophys. Res.*, *94*(D14), 16757-16765.

Dessler, A. E., E. J. Hintsa, E. M. Weinstock, J. G. Anderson, and K. R. Chan (1995), Mechanisms Controlling Water-Vapor in the Lower Stratosphere - a Tale of 2 Stratospheres, *J Geophys Res-Atmos*, *100*(D11), 23167-23172.

Fahey, D. W., and S. R. Kawa (1993), In situ measurements constraining the role of sulphate aerosols in mid-latitude ozone depletion, *Nature*, *363*(6429), 509.

Hanisco, T. F., et al. (2001), Sources, sinks, and the distribution of OH in the lower stratosphere, *J Phys Chem A*, *105*(9), 1543-1553.

Hanson, D. R. (1998), Reaction of ClONO2 with H2O and HCl in sulfuric acid and HNO3/H2SO4/H2O mixtures, *J Phys Chem A*, *102*(25), 4794-4807.

Hintsa, E. J., E. M. Weinstock, J. G. Anderson, R. D. May, and D. F. Hurst (1999), On the accuracy of in situ water vapor measurements in the troposphere and lower stratosphere with the Harvard Lyman-alpha hygrometer, *J Geophys Res-Atmos*, *104*(D7), 8183-8189.

Hintsa, E. J., et al. (1998), Troposphere-to-stratosphere transport in the lowermost stratosphere from measurements of H2O, CO2, N2O and O-3, *Geophys Res Lett*, *25*(14), 2655-2658.

Holton, J. R., P. H. Haynes, M. E. Mcintyre, A. R. Douglass, R. B. Rood, and L. Pfister (1995), Stratosphere-Troposphere Exchange, *Rev Geophys*, *33*(4), 403-439.

Jensen, E. J., O. B. Toon, L. Pfister, and H. B. Selkirk (1996), Dehydration of the upper troposphere and lower stratosphere by subvisible cirrus clouds near the tropical tropopause, *Geophys Res Lett*, *23*(8), 825-828.

Jonsson, H. H., et al. (1995), Performance of a Focused Cavity Aerosol Spectrometer for Measurements in the Stratosphere of Particle Size in the 0.06–2.0-μm-Diameter Range, *J Atmos Ocean Tech*, *12*(1), 115-129.

Keim, E. R., et al. (1996), Observations of large reductions in the NO/NOy ratio near the mid-latitude tropopause and the role of heterogeneous chemistry, *Geophys Res Lett*, *23*(22), 3223-3226.

Kolb, C. E., D. R. Worsnop, M. S. Zahniser, and P. Davidovits (1996), A spectroscopic tour through the liquid aerosol interface: Implications for atmospheric chemistry, *J Geophys Res-Atmos*, *101*(D17), 23039-23043.

Lanzendorf, E. J., T. F. Hanisco, P. O. Wennberg, R. C. Cohen, R. M. Stimpfle, J. G. Anderson, R. S. Gao, J. J. Margitan, and T. P. Bui (2001), Establishing the dependence of [HO2]/[OH]] on temperature, halogen loading, O3, and NOx based on in situ measurements from the NASA ER-2, *J Phys Chem A*, *105*(9), 1535-1542.

Logan, J. A., et al. (1999), Trends in the vertical distribution of ozone: A comparison of two analyses of ozonesonde data, *J Geophys Res-Atmos*, *104*(D21), 26373-26399.

Michelsen, H. A., C. M. Spivakovsky, and S. C. Wofsy (1999), Aerosol-mediated partitioning of stratospheric Cly and NOy at temperatures above 200 K, *Geophys Res Lett*, *26*(3), 299-302.

Murphy, D. M., K. K. Kelly, A. F. Tuck, M. H. Proffitt, and S. Kinne (1990), Ice Saturation at the Tropopause Observed from the ER-2 Aircraft, *Geophys Res Lett*, *17*(4), 353-356.

Pan, L., S. Solomon, W. Randel, J.-F. Lamarque, P. Hess, J. Gille, E.-W. Chiou, and M. P. McCormick (1997), Hemispheric asymmetries and seasonal variations of the lowermost stratospheric water vapor and ozone derived from SAGE II data, *J. Geophys. Res.*, *102*(D23), 28177-28184.

Perkins, K. K., et al. (2001), The NOx-HNO3 system in the lower stratosphere: Insights from in situ measurements and implications of the J(HNO3)-[OH] relationship, *J Phys Chem A*, *105*(9), 1521-1534.

Poulida, O., R. R. Dickerson, and A. Heymsfield (1996), Stratosphere-troposphere exchange in a midlatitude mesoscale convective complex, *J Geophys Res-Atmos*, *101*(D3), 6823-6836.

Ravishankara, A. R., G. Hancock, M. Kawasaki, and Y. Matsumi (1998), Photochemistry of ozone: Surprises and recent lessons, *Science*, *280*(5360), 60-61.

Rinsland, C. P., et al. (1998), ATMOS/ATLAS 3 infrared profile measurements of clouds in the tropical and subtropical upper troposphere, *J Quant Spectrosc Ra*, *60*(5), 903-919.

Rosenlof, K. H., and J. R. Holton (1993), Estimates of the Stratospheric Residual Circulation Using the Downward Control Principle, *J Geophys Res-Atmos*, *98*(D6), 10465-10479.

Scott, S. G., T. P. Bui, K. R. Chan, and S. W. Bowen (1990), The Meteorological Measurement System on the Nasa Er-2 Aircraft, *J Atmos Ocean Tech*, *7*(4), 525-540.

Solomon, S., R. W. Portmann, R. R. Garcia, L. W. Thomason, L. R. Poole, and M. P. McCormick (1996), The role of aerosol variations in anthropogenic ozone depletion at northern midlatitudes, *J Geophys Res-Atmos*, *101*(D3), 6713-6727.

Solomon, S., S. Borrmann, R. R. Garcia, R. Portmann, L. Thomason, L. R. Poole, D. Winker, and M. P. McCormick (1997), Heterogeneous chlorine chemistry in the tropopause region, *J Geophys Res-Atmos*, *102*(D17), 21411-21429.

Solomon, S., R. W. Portmann, R. R. Garcia, W. Randel, F. Wu, R. Nagatani, J. Gleason, L. Thomason, L. R. Poole, and M. P. McCormick (1998), Ozone depletion at mid-latitudes: Coupling of volcanic aerosols and temperature variability to anthropogenic chlorine, *Geophys Res Lett*, *25*(11), 1871-1874.

Stimpfle, R. M., et al. (1999), The coupling of ClONO2, ClO, and NO2 in the lower stratosphere from in situ observations using the NASA ER-2 aircraft, *J Geophys Res-Atmos*, *104*(D21), 26705-26714.

Toohey, D. W., L. M. Avallone, L. R. Lait, P. A. Newman, M. R. Schoeberl, D. W. Fahey, E. L. Woodbridge, and J. G. Anderson (1993), The Seasonal Evolution of Reactive Chlorine in the Northern-Hemisphere Stratosphere, *Science*, *261*(5125), 1134-1136.

Wang, P.-H., P. Minnis, M. P. McCormick, G. S. Kent, and K. M. Skeens (1996), A 6-year climatology of cloud occurrence frequency from Stratospheric Aerosol and Gas Experiment II observations, *J. Geophys. Res.*, *101*(D23), 29407-29429.

Weast, R. C., S. M. Selby, and C. Chemical Rubber (1970), *CRC handbook of tables for mathematics*, 4th ed., xvi, 1120 p. pp., Chemical Rubber Co., Cleveland, Ohio.

Webster, C. R., R. D. May, C. A. Trimble, R. G. Chave, and J. Kendall (1994), Aircraft (Er-2) Laser Infrared-Absorption Spectrometer (Alias) for In-Situ Stratospheric Measurements of HCl, N2O, CH4, NO2, and HNO3, *Appl Optics*, *33*(3), 454-472.

Wennberg, P. O., et al. (1994), Removal of Stratospheric O3 by Radicals - In-Situ Measurements of OH, HO2, NO, NO2, ClO, and BrO, *Science*, *266*(5184), 398-404.

Winker, D. M., and C. R. Trepte (1998), Laminar cirrus observed near the tropical tropopause by LITE, *Geophys Res Lett*, *25*(17), 3351-3354.

Woodbridge, E. L., et al. (1995), Estimates of Total Organic and Inorganic Chlorine in the Lower Stratosphere from In-Situ and Flask Measurements during AASE-II, *J Geophys Res-Atmos*, *100*(D2), 3057-3064.

Chapter 8

In Situ Observations of OH: The Signature of Heterogeneous Processing in the Polar Vortex

8.1 Introduction

Stratospheric ozone protects life on Earth by absorbing harmful ultraviolet radiation from the Sun. One of the crowning scientific and societal achievements of the last century was the discovery of the anthropogenically induced Antarctic "ozone hole", followed by the rapid development of the scientific case unequivocally linking the release of chlorofluorocarbons (CFCs) to chlorine-catalyzed ozone loss, and the subsequent ratification of the Montreal Protocol in 1987 which led to global control of CFCs. New challenges face this century. The rate of climate forcing via unrestrained anthropogenic greenhouse gas (GHG) emissions, combined with the potential for extreme, near term amplification from a daunting array of irreversible feedbacks, is changing the fundamental radiative, dynamical and chemical condition of the stratosphere. Specifically, the long-wave forcing due to CO_2, CH_4, and other well mixed GHGs has the potential 1) to substantially warm the troposphere while cooling the stratosphere [*Ramaswamy et al.*, 2006], 2) to alter the rate at which air is transported from regions of ozone

production to regions of loss [*Austin et al.*, 2007; *Engel et al.*, 2009; *Garcia and Randel*, 2008; *Hegglin and Shepherd*, 2009; *Randel et al.*, 2006; *Stevenson*, 2009], and 3) to modify the fundamental chemical composition of the stratosphere, most notably through predicted increases in stratospheric water vapor concentrations [*Austin and Reichler*, 2008; *Kirk-Davidoff et al.*, 1999]. Individually and collectively, these changes have the potential to significantly alter the forecast recovery of stratospheric ozone. In particular, the combination of colder stratospheric temperatures and increased water vapor concentrations enhances the availability of reactive surface area, e.g., polar stratospheric clouds (PSCs) and/or liquid sulfuric acid aerosol, which is responsible for the heterogeneous conversion of chlorine from its reservoir species, HCl and $ClONO_2$, into its catalytically active forms, Cl and ClO, which promote ozone loss.

Additionally, the threat of abrupt and irreversible climate change has provoked a wide range of proposals aimed at controlling different aspects of the Earth's climate system. One proposal which has garnered significant attention because of its economic and technological feasibility, involves exerting control over the amount of incoming solar radiation by delivering sulfur directly to the lower stratosphere where it undergoes in situ chemical and microphysical conversion into sub-micrometer sulfate aerosol [*Crutzen*, 2006; *Rasch et al.*, 2008]. This intentional introduction of *reflective* surface area comes at the cost of adding *reactive* surface area. [*Tilmes et al.*, 2008] use empirically derived parameterizations to examine the impact of deliberately increasing stratospheric

surface area density on ozone loss through the end of this century. They conclude that such an enterprise will 1) have a strong impact on wintertime Arctic ozone losses, with future losses exceeding those presently observed despite the expected decrease in halogen loading, and 2) considerably delay the predicted recovery of the Antarctic ozone hole.

Forecast changes in stratospheric chemistry and dynamics triggered by long-wave forcing due to increasing GHG emissions, coupled with the intentional elevation of reactive surface area densities, not only increases the probability for chemical activation in the cold, isolated conditions of the polar vortices, but also raises the probability of heterogeneous activation and consequent ozone loss in the midlatitude lower stratosphere. In order to produce robust forecasts of the global distribution of stratospheric ozone as a function of altitude, latitude, and season, for future generations, our models must have the capability of accurately adapting to the rapidly changing boundary conditions of future climate scenarios. As a result, there is a renewed urgency to review and confirm our understanding of the chemical production and loss cycles for stratospheric ozone, with particular emphasis given to understanding the factors that influence heterogeneous chemical activation.

Credible atmospheric chemistry models are validated through 1) a potent combination of laboratory rate constant measurements, and 2) carefully executed in situ and remote studies. The in situ observations of OH obtained on the sunrise flight of January 23, 2000, fulfill the second requirement, providing the most direct observational evidence to date for the occurrence of the heterogeneous

reaction HOCl + HCl → H_2O + Cl_2. This reaction is central to the release of reactive chlorine from the HCl reservoir, and prolonging ozone loss in both the Antarctic and Arctic late winter/early spring vortices.

8.2 The Chemical Context

8.2.1 Heterogeneous Chlorine Activation and O_3 Loss

Ozone loss in the winter/spring Antarctic and Arctic polar vortices is driven by the heterogeneous conversion of chlorine from its reservoir species, HCl, $ClONO_2$ and HOCl, to its catalytically active forms, Cl, Cl_2, ClO, ClOOCl, and OClO (denoted Cl_x). The elevated levels of active chlorine, often measured as a fraction of total inorganic chlorine ($Cl_y = Cl_x + HCl + ClONO_2 + HOCl$), observed in the lower stratosphere during the polar winter, result from activation via one or more of the following reactions on liquid or solid $HNO_3/H_2SO_4/H_2O$ surfaces of varying composition:

$$ClONO_2 + HCl \xrightarrow{PSC/aerosol} Cl_2 + HNO_3 \quad (8.1)$$

$$ClONO_2 + H_2O \xrightarrow{PSC/aerosol} HOCl + HNO_3 \quad (8.2)$$

$$HOCl + HCl \xrightarrow{PSC/aerosol} H_2O + Cl_2 \quad (8.3)$$

In the dark early to mid-winter polar stratosphere where temperatures are low (<190 K), the heterogeneous reaction of $ClONO_2$ with HCl, (Reaction 8.1), is fast and will dominate the release of active chlorine. Assuming, however, that the HCl reservoir is larger than the $ClONO_2$ reservoir going into the winter vortex [*Santee et al.*, 2008; *Wilmouth et al.*, 2006], this reaction will rapidly deplete

$ClONO_2$, and Cl_x production will slow. Under the late winter/early spring depleted $ClONO_2$ conditions, however, excursions into sunlight allow for the rapid production of HOCl. Reaction 8.3 provides a means of continually supplying reactive chlorine from the remaining HCl reservoir, and prolonging the catalytic destruction of ozone within the polar vortex. The model studies of [*Crutzen et al.*, 1992; *Prather*, 1992] independently infer the presence of Reaction 8.3 by examining the evolution of O_3, HCl, $ClONO_2$, and Cl_x in simulated airmasses as they emerge from the Antarctic polar night. In order to reproduce the observed O_3 losses in the early spring Antarctic stratosphere, the models require Reaction 8.3. Furthermore, the model results show that Reaction 8.3 not only impacts Cl_x and O_3, but also perturbs the chemistry of the hydrogen radical family OH, HO_2 and HOCl. Thus, they suggest that in situ OH data can provide a diagnostic fingerprint of heterogeneous activation via Reaction 8.3, and the in situ OH data analyzed here offer the best observational confirmation of this chemical mechanism.

8.2.2 HO_x and Cl_x

Under the elevated active chlorine concentrations that characterize the late winter/early spring polar vortex, fast gas phase exchange occurs between the hydroxyl radicals OH and HO_2 (HO_x) and the chlorine radicals Cl and ClO. The following reaction sequence shows the gas phase equations linking these two radical families through HOCl, and the resulting ozone loss:

$OH + ClO \rightarrow HO_2 + Cl$

$2 \cdot (Cl + O_3 \rightarrow ClO + O_2)$

$HO_2 + ClO \rightarrow \mathbf{HOCl} + O_2$ (8.4)

$\underline{\mathbf{HOCl} + h\nu \rightarrow OH + Cl}$ (8.5)

Net: $2 \cdot O_3 \rightarrow 3 \cdot O_2$ (Null Cl_x, Null HO_x, Loss O_3)

Because, Reactions 8.4 and 8.5 are the dominant production and loss terms for HOCl, additional reactions that produce or consume HOCl, e.g., Reactions 8.2 and 8.3, have the potential to significantly alter HO_x concentrations. Indeed, in the model analyses of both [*Prather*, 1992] and [*Crutzen et al.*, 1992], Reaction 8.3 becomes the dominant sink for HO_x, and is capable of significantly reducing the concentrations of OH and HO_2. Consequently, [*Prather*, 1992] concluded that observations of OH and HO_2 could be important diagnostics of heterogeneous processing, with the suppression of OH indicating the occurrence of Reaction 8.3.

[*Jaegle et al.*, 1997] were the first to use in situ data to examine these processes in detail. They had access to a comprehensive suite of simultaneous in situ observations within the Antarctic vortex for model validation. Comparisons between the measurements and the results of the [*Jaegle et al.*, 1997] photochemical trajectory model provide compelling evidence for all three heterogeneous reactions, with the model requiring Reaction 8.3 to reproduce the observed suppression of HO_x within vortex.

In the present analysis, we examine in detail in situ observations of OH and HO_2 obtained in the Arctic polar vortex aboard the NASA ER-2 aircraft during the NASA SAGE III Ozone Loss and Validation Experiment (SOLVE) mission, staged out of Kiruna, Sweden (68°N, 20°E) during the winter/spring of

1999-2000. Specifically, the concentrations of HO_x measured at sunrise on January 23, 2000, are the lowest encountered in the lower stratosphere at these solar zenith angles, and are only a small fraction of the levels expected from known gas phase production and loss processes. The data from this flight allow for a full exploration of the in situ production and loss chemistry of HO_x, and when combined with a simple photochemical model, offer the most direct evidence to date of the complete removal of HOCl by Reaction 8.3.

8.3. The Signature of Heterogeneous Chemistry

The SOLVE mission provided data for the detailed study of stratospheric ozone loss in the Arctic polar vortex during the winter/spring of 1999 – 2000. In situ measurements of OH and HO_2 were obtained, along with simultaneous measurements of a number of stratospheric gases and meteorological parameters including ClO, $ClONO_2$, HCl, NO, NO_2, HNO_3, N_2O, CH_4, H_2O, O_3, temperature, potential temperature (PT) and particle surface areas and number densities.

On January 23, 2000, the ER-2 flew a trajectory designed to study the chemical evolution of a vortex airmass at sunrise. For the first 2:45 hours at altitude the plane flew along a constant PT surface. Figure 8.1a shows both potential temperature and N_2O as a function of time. The relatively constant measurements of N_2O provide confirmation that the plane was flying through a continuous, uniform airmass. Figure 8.1b shows the monotonic increase of solar zenith angle (SZA) along the flight trajectory, as well as simultaneous measurements of OH. Normally, OH increases proportionally with SZA at SZA's

< 95° [*Hanisco et al.*, 2002b; *Wennberg et al.*, 1998], thus the complete absence of OH, evident in Figure 8.1b, is at once indicative of unusual chemistry.

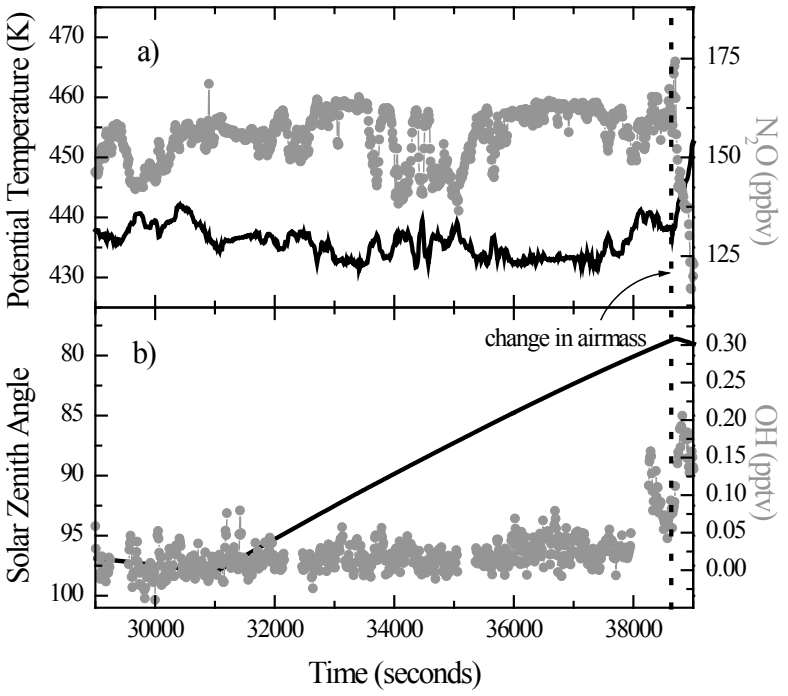


Figure 8.1. Sunrise flight diagnostics. Panel (a) shows measured potential temperature and N_2O as a function of time during the flight of January 23, 2000. The data are indicative of a level (isentropic) flight path through a homogeneous airmass. Panel (b) shows the SZA and measurements of OH as a function of time during the same interval. The ER-2 observed the chemical evolution of this airmass at sunrise.

To put these OH measurements in perspective, we plot in Figure 8.2, OH data from the entire SOLVE mission versus SZA. The solid black line is a fit to the OH data, and is generated solely as a function of SZA. The tight correlation between OH and SZA is apparent. The OH data from the flight of January 23, 2000, are plotted in the solid gray circles. (The data from the flight of March 5, 2000, plotted in the open circles, are the focus of [*Hanisco et al.*, 2002b].) The concentrations of HO_x observed during the sunrise segment of the January 23

flight are the lowest encountered in the lower stratosphere at these solar zenith angles.

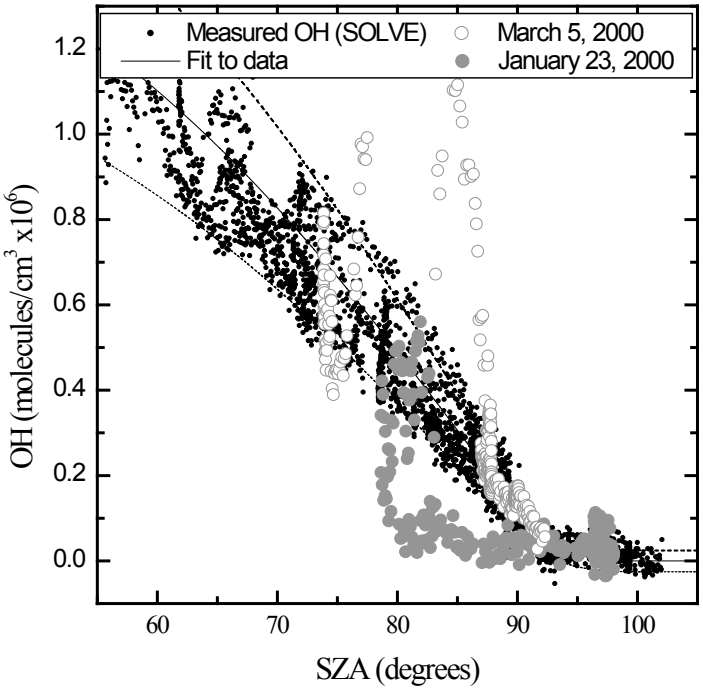


Figure 8.2. Measured OH plotted as a function of SZA. Observations of OH from the entire SOLVE data set are plotted in the black dots. The anomalous flights of January 23, and March 5, 2000, are plotted in the larger closed and open circles respectively.

8.4. The Proposed Chemical Mechanism

The dramatic suppression of OH at sunrise suggests either the loss of a dominant source, or the presence of an additional sink. Under vortex conditions there are four primary gas phase production terms for HO_x (P_{HOx}): $Cl + CH_4$ (~50%), which is dominant because of the high concentrations of Cl_x in the polar vortex, and the extremely fast photolysis of Cl_2; $HNO_3 + h\nu$ (30%); $O^1D + [H_2O, CH_4, H_2]$ (20%); and $OH + CH_4$. In situ measurements of ClO, ClOOCl, and

HNO_3 during the sunrise segment show that all of the largest HO_x sources were present. Furthermore, because there are numerous source terms, the removal of any single source will have at most a ~50% effect. We conclude, therefore, that an additional sink is needed to achieve the near complete removal of HO_x.

Gas phase HO_x loss (L_{HOx}) is dominated by the reaction of $OH + HNO_3 \rightarrow H_2O + NO_3$ (~80%), and $OH + ClO \rightarrow HCl + O_2$ (~20%), which is important for chlorine recovery through the production of HCl. Additional contributions to HO_x loss include the reaction of $HO_2 + ClO \rightarrow HCl + O_3$, another pathway for recovery, the self-reaction, $OH + HO_2 \rightarrow H_2O + O_2$, and $OH + HCl \rightarrow H_2O + Cl$. The first-order lifetime for HO_x with respect to its dominant loss term is typically on the order of 1 minute. In order to achieve the near complete removal of HO_x along the flight track, an additional sink is needed with a rate more than five times that of $OH + HNO_3$[1].

It is useful to treat the reactions that inter-convert HO_x species as distinct from the primary production and loss reactions for HO_x, and a natural analytical separation arises because of the difference in their rates. The reactions involved in the exchange of HO_2 and OH, e.g., Reactions 8.4 and 8.5, participate in null sequences with respect to HO_x, and typically have much faster rates than HO_x production/loss [*Hanisco et al.*, 2002a].

[1] If we assume that HO_x is in photochemical steady state, and define the ratio of the observed concentration, $[HO_x]_{obs}$, to the expected concentration, $[HO_x]_{exp}$, as $F = [HO_x]_{obs} / [HO_x]_{exp}$, where $F < 1$, it can be shown that the additional loss term, L_{add}, must be equal to $(1 - F)/F \cdot L_{exp}$. During the sunrise segment, the observed concentration is ~6 to 8 times less than the concentration expected from standard gas phase production and loss chemistry. Thus, the additional loss term must be at a minimum ~5 times faster than the sum of the normal loss terms.

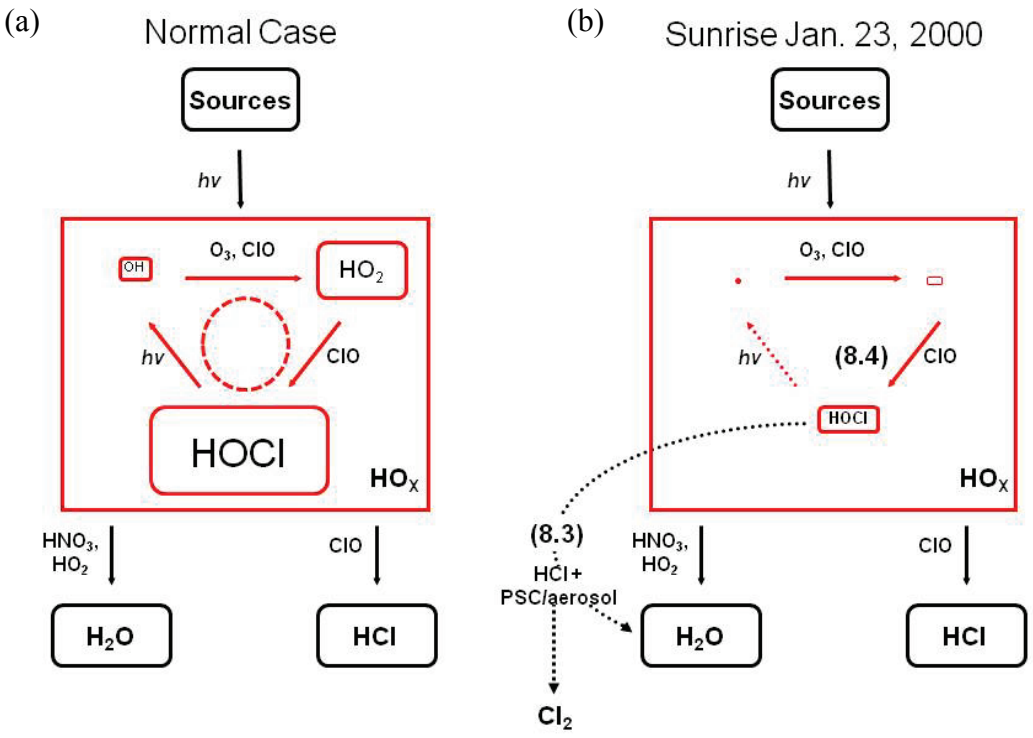


Figure 8.3. Schematic of HO_x chemistry in normal and perturbed airmasses. This figure contrasts the normal scenario, panel (a), in which HOCl acts simply as a reaction intermediate in the fast gas phase exchange of HO_x species, and the special case, panel (b), invoking the heterogeneous loss of HO_x through Reaction 8.3, representing the conditions on the flight of January 23, 2000.

This distinction, emphasized by the red boxes in Figure 8.3, reveals a few important aspects of HO_x chemistry. First, it is reasonable to study HO_x using measurements of OH. This is both because HO_x production/loss occurs primarily through OH, and also because the production/loss of any HO_x species is equivalent since repartitioning is effectively instantaneous. Simply put, changes in P_{HOx} or L_{HOx} are reflected in [OH]. Second, the fast OH ↔ HO_2 exchange reactions, Reactions 8.4 and 8.5, are the dominant gas phase production and loss terms for HOCl. Thus, measurements of OH and HO_2 can be used to obtain steady-state concentrations of HOCl, where $HOCl_{SS} = P_{HOCl}/L_{HOCl} = [HO_2] \cdot k_4/J_5$.

Under vortex conditions diurnal steady-state concentrations of HOCl are typically ~80 pptv ±20 pptv. Finally, because of the tight coupling between OH, HO_2, and HOCl, it is convenient to treat HOCl as an active member of the HO_x family.

Typically, the gas phase chemistry of HOCl is not included as a source or sink of OH because its gas phase production and loss reactions constitute a null cycle with respect to HO_x, and it functions simply as a steady state intermediate. This "normal case" represented by the reaction sequence in Section 8.2.2 is depicted in Figure 8.3a. The addition of Reaction 8.3, however, creates a pathway for HO_x loss through HOCl, which is represented by the following reaction sequence, and illustrated in the "perturbed" scenario in Figure 8.3b.

$OH + ClO \rightarrow HO_2 + Cl$

$2\cdot(Cl + O_3 \rightarrow ClO + O_2)$

$HO_2 + ClO \rightarrow HOCl + O_2$ (8.4)

$HOCl + HCl \xrightarrow{PSC/aerosol} H_2O + Cl_2$ (8.3)

$\underline{Cl_2 + h\nu \rightarrow 2\cdot Cl}$

Net: $2\cdot O_3 + OH + HCl \rightarrow 3\cdot O_2 + H_2O + Cl$ (Increase Cl_x, Loss HO_x, Loss O_3)

In contrast to the "normal case" in which HO_x is recycled through HOCl photolysis, the net result of the "perturbed case" is a loss of HO_x and the oxidation of HCl. If the rate of Reaction 8.3 is sufficiently fast, it may be the primary cause of HO_x suppression along the flight track.

The scenario described above implicitly assumes that the loss via Reaction 8.3 occurs along the flight track. An alternative explanation for the suppression of

OH at sunrise involves the near complete removal of HOCl via Reaction 8.3 during the cold, dark night prior to sunrise.

OH is sensitive to large perturbations in HOCl, i.e., when HOCl concentrations are far from the diurnally averaged steady state value of ~60 pptv [*Hanisco et al.*, 2002a]. This is particularly true at sunrise, when OH levels are low (~0.1 pptv). In this second scenario, the depletion of the nighttime HOCl reservoir effectively short circuits the fast exchange cycle that commences at sunrise, and turns Reaction 8.4 into a sink for HO_x until the concentration of HOCl has had time to recover and resume its role as a reactive intermediate. This is essentially a two-step modification of the "perturbed" case in Figure 8.3.

In the following sections we use the in situ data in conjunction with a simple photochemical model to examine in greater detail the mechanism of HO_x loss in this airmass.

8.5 Heterogeneous Rates in the Late Winter Vortex

The two mechanisms, 1) the suppression of HO_x via heterogeneous loss of HOCl through Reaction 8.3 along the flight track, or 2) the suppression of HO_x due to the complete removal of the nighttime reservoir of HOCl prior to sunrise, depend critically upon the factors controlling heterogeneous processing with the airmass.

Heterogeneous processing is complex and depends on a number of factors, among them, the available reactive surface area, S_a, and the reaction probability per collision, γ. Moreover, γ depends upon temperature and surface type,

specifically aerosol phase and chemical composition. In the present analysis, specifically in the development of our photochemical model, we have made a few simplifying assumptions which require discussion.

8.5.1 Reactive Surface Area

The flight of January 23, 2000, is punctuated by three distinct surface area enhancements, probable PSC encounters, and total surface area is moderately elevated throughout. Particle surface area densities, measured aboard the ER-2 by the Multi-Angle Scattering Probe (MASP) and the Focused Cavity Spectrometer (FCAS) instruments, are plotted in Figure 8.4 [*Baumgardner et al.*, 2001].

Three particle modes are evident in this airmass. The ubiquitous liquid phase sulfate aerosol is the smallest, with particle diameters, D, equal to ~0.1 um. Two additional particle modes are evident in the MASP data set: (1) a population of small particles with diameters sized $0.3 < D \sim 0.5 < 0.6$ um, and (2) a population of larger particles with diameters sized $0.7 < D \sim 1.0 < 1.25$ um. A small population of very large particles, 'rocks', with diameters >10.0 um, is also evident [*Brooks et al.*, 2003; *Fahey et al.*, 2001]. The presence of the very large particles gives rise to spikes in total aerosol volume, however, these particles are so few in number that they rarely contribute significantly to measurements of total surface area.

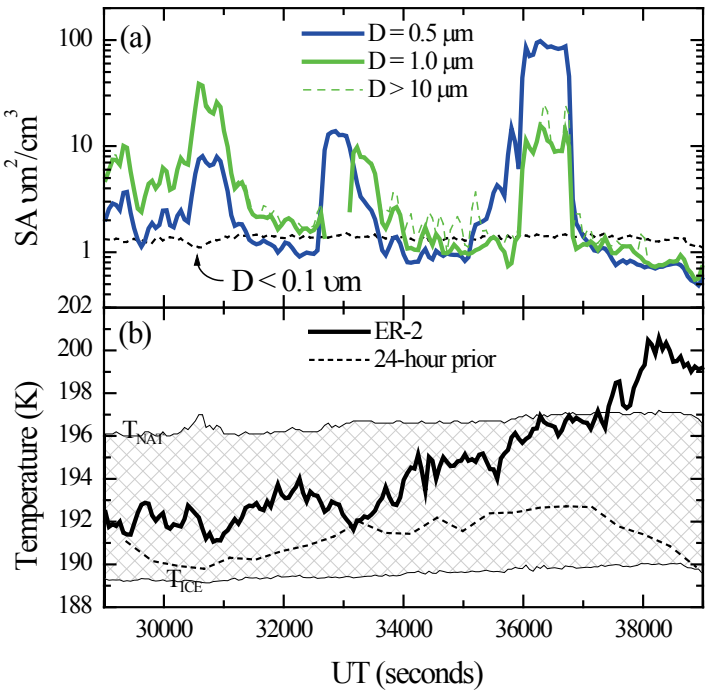


Figure 8.4. Particle surface area and temperature on January 23, 2000. Panel (a) shows measured particle surface area density for the three primary modes, D < 0.1 um (ubiquitous sulfuric acid aerosol), D = 0.5 um (small), D = 1.0 um (big), plus "rocks", D > 10 um. Panel (b) shows the temperatures measured aboard the ER-2 during the sunrise portion of the flight, and 24-hour minimum temperatures from back-trajectory data. The shaded area is bounded by the nitric acid trihydrate (NAT) and water ice condensation temperatures respectively.

Consistent with the convention used in the heterogeneous analysis of [*Hanisco et al.*, 2002b], we assume that 1) the small particle mode, (*D* ~0.5 um), is predominantly supercooled liquid aerosol, and that 2) the big mode, which will include the 'rocks', is predominantly solid nitric acid trihydrate, NAT, (*D* ~ 1.0 um). Though the composition of the liquid aerosol is not measured directly, at ambient temperatures and HNO_3 concentrations (~12 ppbv), it is likely that stratospheric aqueous sulfuric acid aerosols will absorb HNO_3 from the gas phase to form the ternary solution aerosols, $HNO_3/H_2SO_4/H_2O$, hereafter referred to as STS [*Brooks et al.*, 2004; *Carslaw et al.*, 1995; *Peter*, 1997; *Voigt et al.*, 2000].

Temperatures measured along the flight trajectory, plotted in Figure 8.4b, and the extended temperature histories of the airmass are consistent with these assumptions. The minimum temperature encountered by the air parcels over the previous 24-hour period, determined from back-trajectory data, shows that the air has been well below the NAT condensation temperature, T_{NAT}, for an extended period.

8.5.2 Heterogeneous Rates on Solid NAT

[*Carslaw and Peter*, 1997] and [*Carslaw et al.*, 1997] discuss several parameterizations for Reactions 8.1 − 8.3 on NAT, determined by fitting to laboratory data of [*Abbatt and Molina*, 1992a; b]. These laboratory data show that reactive uptake on NAT surfaces is positively correlated with relative humidity as well as the fraction of HCl taken up by the hydrate. Thus, the rates of reaction slow as the ambient relative humidity with respect to ice, RH_i, decreases, and as the partial pressure of HCl, p_{HCl}, decreases. A second laboratory study by [*Hanson and Ravishankara*, 1993] demonstrates a weaker dependence on p_{HCl}, by as much as a factor of 100, for Reactions 8.1 and 8.2.

In this analysis, therefore, we consider two separate schemes for Reactions 8.1 − 8.3 on NAT. In the first scheme, we use the parameterizations of [*Carslaw et al.*, 1997], which are tied to the laboratory data of [*Abbatt and Molina*, 1992a; b], and thus sensitive to p_{HCl}. In the second scheme, we use the identical parameterizations without the modification for low HCl concentrations. In the second scheme, the NAT reactivities are a few orders of magnitude less sensitive

to p_{HCl}, and the reaction probabilities approximately equal to the heterogeneous rate recommendations from the Jet Propulsion Laboratory 2006 evaluation (JPL-06) at temperatures below 192 K [*Sander et al.*, 2006]. Figure 8.5 shows the reactive uptake coefficients, γ, for Reaction 8.3 on solid NAT for both parameterizations. The JPL-06 recommendation is included for reference.

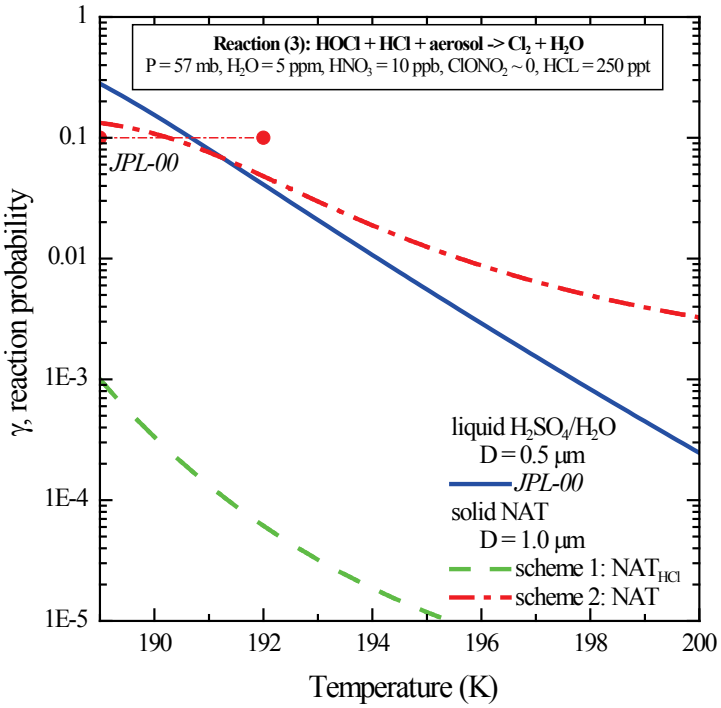


Figure 8.5. Reaction probability for HCl + HOCl $\xrightarrow{PSC/aerosol}$ $H_2O + Cl_2$. The reaction probability, γ, for Reaction 8.3 is shown as a function of temperature for conditions representative of those encountered on January 23, 2000. We assume that particles with diameters ≤ 0.5 μm are ternary liquid aerosols, and particles with diameters ≥ 1.0 μm are solid NAT.

8.5.3. Heterogeneous Rates for Liquid Binary & Ternary Aerosol

Heterogeneous reactivity on liquid binary or ternary aerosols is a function of pressure, temperature, H_2O partial pressure, H_2SO_4 weight percent, HCl, $ClONO_2$, and aerosol radius. The parameterizations for the rates of Reactions 8.1

− 8.3 on binary H_2SO_4/H_2O aerosols are taken directly from the JPL-06 recommendations [*Sander et al.*, 2006], and include terms that account for low concentrations of HCl [*Shi et al.*, 2001]. Both laboratory and in situ data show that the rates of Reactions 8.1 and 8.2 decrease in ternary solutions with significant HNO_3 weight percent [*Hanisco et al.*, 2002b; *Hanson*, 1998], however, Reaction 8.3 is presumed to be insensitive to the presence of HNO_3 and is expected to have the same rate on both binary and ternary aerosols. Due to the very low $ClONO_2$ concentrations along this flight segment[2] we expect the role of Reactions 8.1 and 8.2 to be diminished relative to Reaction 8.3. Thus, we have chosen to apply the JPL-06 parameterizations for Reactions 8.1 − 8.3 on binary aerosols without modification for HNO_3 uptake.

8.6 Models Versus Measurements

8.6.1 Model 1: Heterogeneous Loss Along the Flight Track

The simplest way of modeling OH is to assume it is in instantaneous photo-chemical steady state. Under this assumption, $(d[OH]/dt) = P_{HOx} - [OH]\cdot L_{HOx} \approx 0$, and the steady state concentration is given by $[OH]_{SS} \approx P_{HOx}/L_{HOx}$, where total production (P_{HOx}) and loss (L_{HOx}) represent the sum of the rate limiting steps from each of the dominant HO_x production and loss terms. Recall HO_x is primarily produced and lost through OH. Using this approximation we can

[2] $ClONO_2$ concentrations are below the detection threshold of the instrument, and a steady-state analysis of partitioning among the reactive chlorine species reveals that >90% of Cl_y is either ClO or ClOOCl [*Stimpfle et al.*, 2005; *Wilmouth et al.*, 2006].

model the expected steady state OH concentration as a function of time along the flight track.

We determine the instantaneous production and loss rates from in situ measurements of temperature, pressure, O_3, H_2O, CH_4, H_2, CO, HCl, ClO, BrO, ClOOCl, $ClONO_2$, NO, NO_y, and particle surface areas and densities, interpolated to a common 10 second time grid. Concentrations of $O(^1D)$ and Cl are derived from their steady state relations, and except for the heterogeneous rate parameterizations discussed in the previous section, the chemical rate constants are taken from the JPL-06 recommendations. Photolysis rates for various species along the flight track are calculated using recommended cross-sections and a combination of overhead ozone column data from the TOMS satellite, profiles from ozone sondes, and ER-2 observations [R. J. Salawitch, *personal communication*].

Figure 8.6a shows $[OH]_{SS}$ (open circles) expected from standard gas-phase production and loss chemistry plotted as a function of time along the flight trajectory. OH expected from a SZA fit (shaded region) and observed OH (solid circles) are shown for comparison. The agreement between the instantaneous 'gas-phase only' steady state model and OH determined solely as a function of solar zenith angle is evident, as is the disagreement between the model results and the measurements at sunrise.

At UT > 38000, the plane enters a different airmass. Following the change in aimass the agreement between the model and the measurements is consistent with expectations, i.e., $P_{HOx}/L_{HOx} \approx [OH]_{SS}$ is within ±30% of the measured

concentration. This region is circled in Figure 8.6a. The convergence in agreement between the instantaneous steady-state result, the SZA prediction, and the in situ measurements, demonstrates the quality of our understanding of HO_x gas phase production and loss processes.

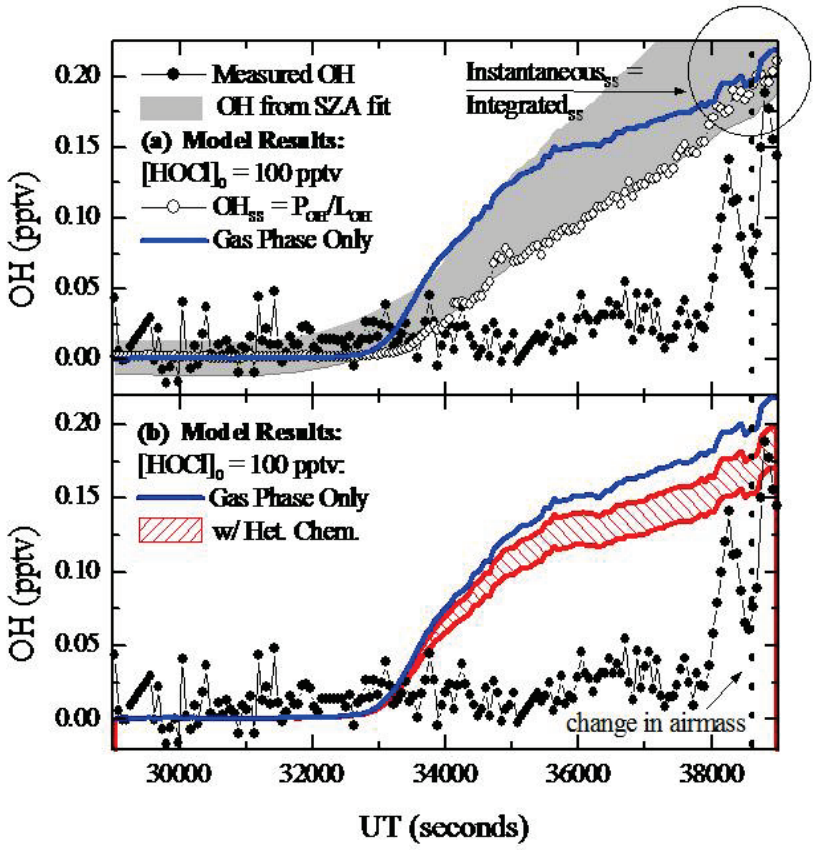


Figure 8.6. Measured and modeled OH as a function of time during sunrise on January 23, 2000. Panel (a) shows the agreement between the instantaneous steady state model (open circles), the integrating model (blue line), and OH determined solely as a function of SZA (shaded area). Also evident is the disagreement between the 'gas-phase only' model results, and the in situ observations (solid circles). Following the change in aimass the agreement between model and measurement is consistent with expectations. Panel (b) shows the results of the integrating model initialized with the expected nighttime reservoir of $HOCl_0$ = 100 pptv run with and without the inclusion of Reaction 8.3.

In order to account for the perturbations to HO_x that arise when intermediary species such as HOCl, or HOBr are not strictly in steady state, we developed a time-dependent, integrating photochemical model. Note that this model is only valid for a homogenous, invariant airmass, e.g., UT = 29000 to ~38000 seconds on January 23rd, 2000.

We initialize the time-dependent model at a point early in the flight where the ER-2 is in total darkness, and set the concentrations of OH and HO_2 equal to zero. We then determine the change in OH at each time step, dOH/dt, from the calculated instantaneous production and loss rates, and integrate over the sunrise segment of the flight. We can adjust the time step of this model to be on the order of seconds, or roughly equal to the rate of the fast exchange reactions 8.4 and 8.5. This is much shorter than the effective time constant of the instantaneous steady state model, which is, by definition, the time required for HO_x to achieve photochemical steady state. At sunrise this is ~1 hour. Deviations between the time-dependent model and instantaneous steady-state model are attributed to the rapid repartitioning of HO_x species at sunrise.

The time-dependent results for gas-phase chemistry only, plotted in Figure 8.6a, show the sensitivity of OH at sunrise to the initial/nighttime value of HOCl, e.g., $HOCl_0$ = 100 pptv, the approximate diurnal steady-state value[3]. At sunrise OH is enhanced relative to its instantaneous steady state value because of the burst of HOCl photolysis. After the first hour of solar exposure, however, the integrating and instantaneous models converge.

[3] Model sensitivity studies (not shown) demonstrate that initial values of HOCl ~100 ± 10 pptv satisfy agreement with the SZA parameterization, and "unperturbed" in situ OH data.

Figure 8.6b shows the model results with heterogeneous processing occurring along the flight track, where the shaded range reflects the results of the two heterogeneous parameterizations. The response in modeled OH is relatively minor. Model runs initialized with 100 pptv of HOCl show that heterogeneous chemistry occurring along the flight track is not fast enough to reproduce the OH observations. Even the heterogeneous parameterization which includes a fast reaction rate on nitric acid trihydrate (NAT) particles, is not sufficient to reduce modeled OH by more than ~30% over gas phase chemistry alone. Furthermore, the percent reduction of OH, due to the occurrence of Reaction 8.3 along the flight track, is greater for larger initial values of HOCl, and diminishes with lower $HOCl_0$ values. The results of this model suggest that heterogeneous processing along the flight track is not fast enough to explain the observed suppression.

An examination of the rate of Reaction 8.3 relative to the standard gas phase loss processes shows that gas phase chemistry is dominant along the flight track. Figure 8.7 shows the lifetime of OH with respect to the dominant gas phase loss term ($OH + HNO_3$), and with respect to heterogeneous removal via Reaction 8.3 plotted versus temperature. The range in the first-order heterogeneous lifetime reflects the two reaction parameterizations. The data show that when temperatures are low, and during periods of significant surface area enhancement, the heterogeneous and gas phase loss rates are comparable. However, for the sunrise dataset as a whole, heterogeneous removal cannot compete with standard gas phase chemistry. Reaction 8.3 is too slow to cause significant additional OH loss during flight. Moreover, observed OH appears to be insensitive to the structure

evident in the aerosol data, see Figure 8.4. If fast heterogeneous chemistry were the cause of OH suppression during flight, we might expect to see corresponding structure in OH[4].

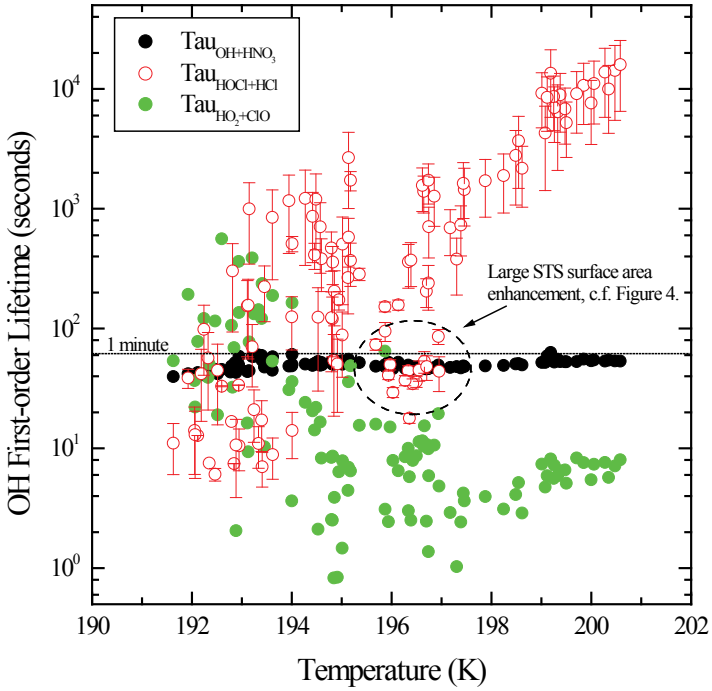


Figure 8.7. The lifetime of OH with respect to the dominant gas phase loss term, $OH + HNO_3 \rightarrow H_2O + NO_3$, and with respect to heterogeneous removal via Reaction 8.3, plotted versus temperature. The top and bottom caps of the data capture the range of the calculated heterogeneous loss rate using schemes 1 & 2 respectively. Also shown is the instantaneous lifetime of OH with respect to the fast exchange Reaction 8.4.

Figure 8.7 also shows the time constant for Reaction 8.4. This gas phase reaction, which is the primary production term for HOCl, is a full order of magnitude faster than $OH + HNO_3$. If the steady-state reservoir of HOCl has been depleted, via Reaction 8.3 in the cold, dark period prior to sunrise, the balance between HOCl photolysis, Reaction 8.5, and Reaction 8.4 will favor the

[4] For example, on the flight of March 3, 2000, episodic OH enhancements are consistent with heterogeneous processing occurring at the time of flight [*Hanisco et al.*, 2002b].

production of HOCl, and effectively lead to the temporary suppression of OH. An airmass depleted of HOCl will be depleted of HO_x, due to the tight photochemical coupling between OH, HO_2, and HOCl.

8.6.2 Model 2: Heterogeneous Loss the Night Prior

We examined the extended history of the airmass in order to determine the likelihood that heterogeneous processing occurred prior to flight. Figure 8.8 shows back-trajectory SZA and temperature data for selected parcels from the airmass with suppressed OH.

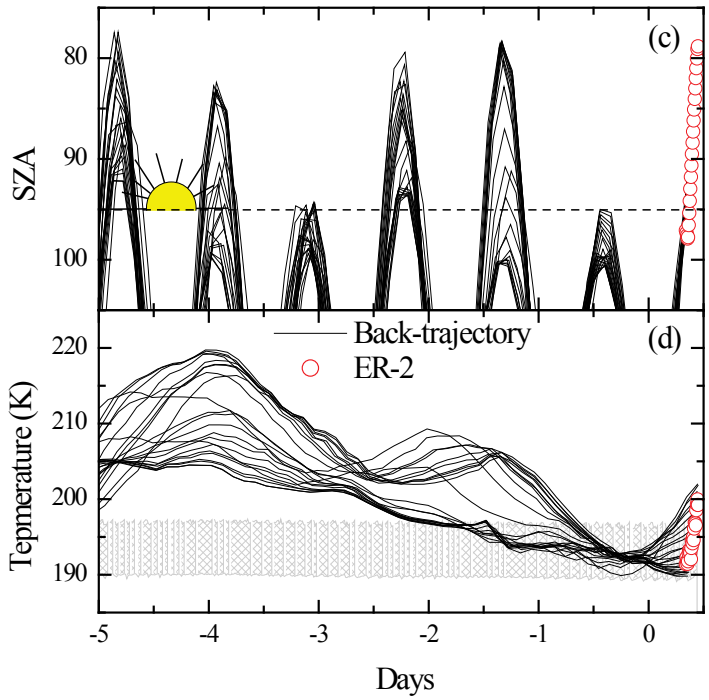


Figure 8.8. Back-trajectory data. Panel (a) shows the solar zenith angles, and panel (b) the temperatures the air parcels encountered for five days prior to flight intercept. The shaded area is bounded by the NAT and water ice condensation temperatures.

Figure 8.8a shows the SZAs these air parcels encountered over the five days prior to the flight encounter. Figure 8.8b shows the corresponding temperature history. Back-trajectory SZAs reveal that all of the air parcels have been exposed to direct sunlight within the last 3 days. However, just prior to flight, the airmass as a whole has been subject to ~36 hours of sustained darkness and simultaneous cooling, with temperatures well below T_{NAT} for >24 hours. Therefore, the heterogeneous loss of HOCl prior to flight intercept is consistent with the solar exposure and temperature histories of the air parcels.

In order to simulate a nighttime loss of HOCl we ran the time-dependent, integrating model with $HOCl_0 = 0$. The results, plotted in Figure 8.9 (red dashed line), show that modeled OH at sunrise is suppressed. Furthermore, there is no significant difference between the gas-phase and heterogeneous result initialized with $HOCl_0 = 0$. Similar to the sunrise burst of OH from HOCl photolysis via Reaction 8.5, when $HOCl_0 = 100$ pptv, the suppression results from the rapid repartitioning of HO_x species. With HOCl depleted, the cycle converting $HO_2 \leftrightarrow$ OH is broken, and the gas phase reaction $HO_2 + ClO \rightarrow HOCl + O_2$, which has a rate 5 to 10 times faster than $OH + HNO_3$, becomes an efficient, but temporary, sink for HO_x. Sensitivity tests (not shown) demonstrated that initial values of HOCl = 10 ± 10 pptv were consistent with the in situ data. The results of this model suggest, therefore, that the OH observations are best explained by a loss of HOCl via heterogeneous processing the night prior to flight intercept. This scenario, the near complete loss of HOCl through heterogeneous reaction with HCl during the cold, dark 24-hour interval prior to sunrise, is consistent with all

aspects of the in situ data set, and is the only way to successfully model the observed OH suppression on the flight of January 23, 2000.

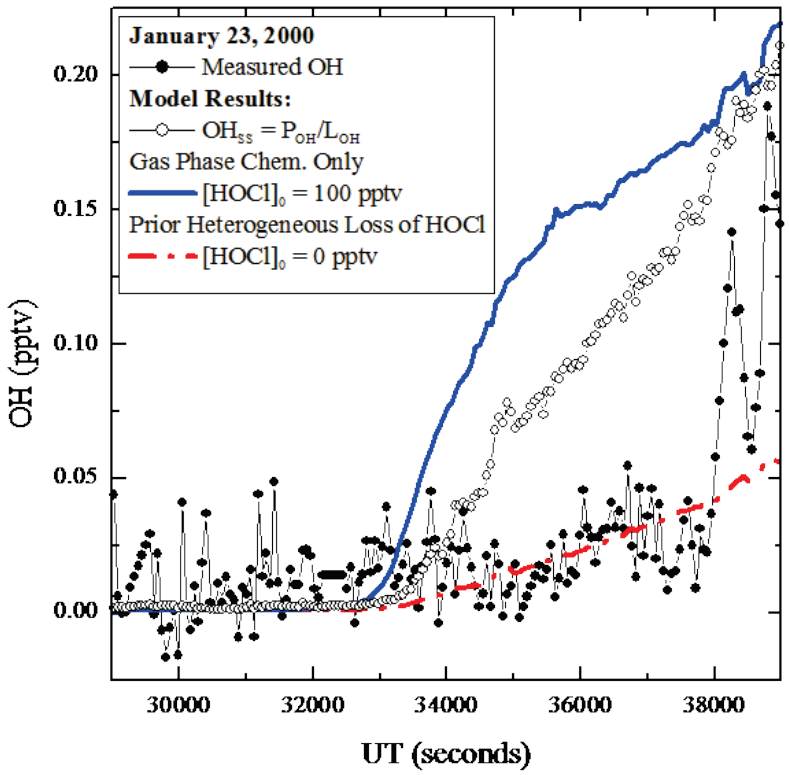


Figure 8.9. Measured and modeled OH as a function of time during sunrise on January 23, 2000. Modeled OH from the integrating model initialized with $HOCl_0$ = 0 pptv (red dashed line) is plotted as a function of time. The initial value of $HOCl_0 = 10 \pm 10$ pptv simulates the complete heterogeneous loss of HOCl via Reaction 8.3 in the cold dark period prior to sunrise. The results of the integrating model initialized with $HOCl_0 = 100$ pptv (blue line), and the instantaneous steady state model (open circles) are shown for comparison.

8.6.3 Additional Considerations

In this section we review our reasoning and consider a few alternative explanations. As discussed in prior sections, the in situ dataset allows us to determine the primary HO_x production and loss rates along the flight trajectory,

and to model OH concentrations. An inspection of the agreement between our model results and the measurements in both the perturbed and normal (at UT >38000 seconds) airmasses is informative. Figure 8.6 shows that in the normal airmass, the OH measurements are in excellent agreement with modeled OH using 1) the gas phase instantaneous steady state relation, i.e., $OH_{SS} = P_{HOx}/L_{HOx}$, 2) the results of our integrating model initialized with $HOCl_0 = 100$ pptv, and 3) the canonical OH:SZA correlation. The agreement between our models' results and the OH measurements in this airmass demonstrates that the gas phase sources and sinks of OH are well characterized. Furthermore, there is no discernable difference in the calculated gas phase production or loss terms between the perturbed and normal airmasses, indicating that there is nothing unique about the gas phase production and loss processes in the perturbed airmass. Finally, as we determined in Section 8.4, in order to reproduce the observations in the perturbed airmass, it is necessary to invoke an additional loss mechanism for HO_x with an effective rate ~5 times faster than the standard gas phase loss reactions.

Under typical vortex conditions, HOCl is continually being supplied by the reaction of HO_2 with ClO (Reaction 8.4), and recycled back to OH via photolysis (Reaction 8.5). Because these two reactions constitute the primary production and loss terms for HOCl, it is convenient to include HOCl as a member of the HO_x family. The differences between the instantaneous steady-state model results and those of the integrating model demonstrate the close coupling of HO_x and HOCl, and the sensitivity of OH to variations in HOCl. Concentrations of HOCl higher than its instantaneous steady-state value can

temporarily enhance OH, because of the rapid repartitioning through Reaction 8.5, while low or zero concentrations of HOCl lead to OH suppression, because of the temporary suspension of Reaction 8.5. Our model results reveal that if OH is suppressed at SZA's < 90°, HOCl must either be suppressed, and/or there must be an additional loss term for HOCl, with a rate competitive with photolysis.

The heterogeneous loss of HOCl through Reaction 8.3 is the only known mechanism for HOCl loss under the low $ClONO_2$ conditions of this airmass. The in situ data show, however, that HOCl loss via Reaction 8.3 along the flight track is too slow to significantly alter OH, assuming a typical nighttime HOCl reservoir of ~100 pptv. No realistic parameterization of heterogeneous rates can reproduce the observed OH concentrations if we rely solely on heterogeneous loss during the time of flight. A longer integration time is required for the complete removal of HOCl. Thus, the only way to successfully model the OH measurements, within the constraints of the complete in situ dataset, is to conclude that HOCl has been eliminated prior to sunrise. Specifically, we find that in order to reproduce the OH measurements in the perturbed airmass, HOCl levels can be no higher than ~20 pptv.

The sharp contrast between the temperature histories of the perturbed and normal airmasses confirms this theory. The airparcels which constitute the perturbed airmass experienced a ~15º cooling and >24 hours of darkness prior to flight intercept, rather than the ~15° warming experienced by the normal airmass. The suppression of OH in the perturbed airmass is consistent, therefore, with heterogeneous processing and the removal of HOCl. Reaction 8.3 can only occur

in the presence of significant reactive surface area, e.g., PSCs, and thus at temperatures below the NAT threshold. In contrast, the OH observations in the normal airmass are consistent with a thaw and subsequent return to standard gas phase conditions. Unfortunately, the long period of darkness with temperatures below T_{NAT}, limits our ability to constrain the rate of Reaction 8.3, because any reasonable combination of particles and parameterizations will result in the near complete removal of HOCl.

Additional evidence for recent heterogeneous processing is provided by the in situ measurements of ClO and ClOOCl [*Stimpfle et al.*, 2004]. The sunrise segment of this flight is remarkable for both its high concentrations of ClO, up to ~1600 pptv, and low concentrations of $ClONO_2$, which are below the detection threshold of the instrument [*Wilmouth et al.*, 2006]. These are the highest ClO concentrations ever recorded by the Harvard ER-2 instrument, including all data from the SOLVE mission, as well as data from the Antarctic polar vortex. Accordingly, the Cl_x/Cl_y ratio for this flight is ~0.90, with ClO and ClOOCl the dominant species, an unequivocal signature of recent heterogeneous activation. Under these late winter depleted $ClONO_2$ conditions, excursions into sunlight allow for the production of HO_x and thus the formation of HOCl. Subsequent cooling, coincident with a return to darkness, leads to the formation of PSCs. This reactive surface area then allows Reaction 8.3 to play a significant role in the release of reactive chlorine from the remaining HCl reservoir.

Alternative explanations for the suppression of OH include 1) the direct reaction of OH or HO_2 on aerosols, and 2) the heterogeneous loss of HOBr.

Though aerosol surface areas are enhanced for periods during the sunrise segment of the flight, the sustained suppression is not likely the result of OH or HO_2 reacting with, or on, aerosols. The lack of structure in the OH concentrations along the flight track is inconsistent with the large deviations in reactive surface area, ~2 orders of magnitude. Furthermore, the OH data of the March 5, 2000, ER-2 flight do not support this theory [*Hanisco et al.*, 2002b]. Unusually high concentrations of OH are evident during this flight, coincident with large surface area enhancements.

The heterogeneous loss of HOBr as the sole cause of, or even a significant contributor to the suppression of HO_x, is also unlikely given the low concentration of HOBr relative to HOCl within the vortex. We estimate the concentration of total inorganic bromine, Br_y ~18 pptv, from its correlation with N_2O and in situ N_2O measurements, and assume that BrO is roughly 50% of Br_y, i.e., ~9 pptv. The consequent ambient steady state concentration of HOBr, determined from measurements of HO_x, where $HOBr = P_{HOBr}/L_{HOBr} = k_{HO2+BrO \rightarrow HOBr+O2}/J_{HOBr}$, is estimated to be ~0.5 pptv, or 0.5% of HOCl for the majority of the vortex data. For the flight of January 23, 2000, HOBr is estimated to be only ~0.3% of HOCl. Under these conditions, modeled OH is practically insensitive to the presence or absence of HOBr. Neither the heterogeneous reaction of HOBr on aerosols during flight, nor a complete loss of HOBr prior to flight is capable of significantly altering OH concentrations along this flight segment.

8.7 Conclusions

The OH concentrations measured in situ aboard NASA's ER-2 aircraft on the flight of January 23, 2000, bear the distinct signature of recent heterogeneous processing. They are the lowest ever encountered at these solar zenith angles and deviate markedly from the values expected from standard gas phase production and loss processes. Model results reproduce measured OH only when initialized with low (~<20 pptv) or zero HOCl, indicating that the nighttime reservoir of HOCl has been decimated via heterogeneous reaction with HCl during the preceding cold, dark >24 hour interval. As such, these measurements provide the most direct evidence to date for Reaction 8.3.

In addition, both the results of this analysis and that of [*Hanisco, et al.*, 2002b], demonstrate that because of 1) the role of HOCl in two of the key heterogeneous reactions leading to the release of catalytically active chlorine, 2) the tight coupling between the HO_x and Cl_x families through HOCl, and 3) the high sensitivity of HO_x to perturbations in HOCl, in situ measurements of OH provide unique insight into the chemistry of chlorine activation and/or recovery within the vortex. Reaction 8.3, the focus of the present analysis, provides a reactive partner for HCl under low $ClONO_2$ conditions and is central to the sustained release of reactive chlorine and prolonging ozone loss through the late winter/early spring vortex. Finally, a thorough understanding of heterogeneous chemistry, founded on measurements and analysis of this sort, is critical to forecasts of ozone recovery, particularly given a future with rapidly changing stratospheric boundary conditions driven by anthropogenic climate forcing.

References

Abbatt, J. P. D., and M. J. Molina (1992a), Heterogeneous Interactions of Clono2 and Hcl on Nitric-Acid Trihydrate at 202-K, *J Phys Chem-Us*, *96*(19), 7674-7679.

Abbatt, J. P. D., and M. J. Molina (1992b), The Heterogeneous Reaction of Hocl+Hcl-]Cl2+H2o on Ice and Nitric-Acid Trihydrate - Reaction Probabilities and Stratospheric Implications, *Geophys Res Lett*, *19*(5), 461-464.

Austin, J., and T. J. Reichler (2008), Long-term evolution of the cold point tropical tropopause: Simulation results and attribution analysis, *J Geophys Res-Atmos*, *113*.

Austin, J., J. Wilson, F. Li, and H. Vomel (2007), Evolution of water vapor concentrations and stratospheric age of air in coupled chemistry-climate model simulations, *J Atmos Sci*, *64*(3), 905-921.

Baumgardner, D., H. Jonsson, W. Dawson, D. O'Connor, and R. Newton (2001), The cloud, aerosol and precipitation spectrometer: a new instrument for cloud investigations, *Atmos Res*, *59*, 251-264.

Brooks, S. D., O. B. Toon, M. A. Tolbert, D. Baumgardner, B. W. Gandrud, E. V. Browell, H. Flentje, and J. C. Wilson (2004), Polar stratospheric clouds during SOLVE/THESEO: Comparison of lidar observations with in situ measurements, *J. Geophys. Res.*, *109*(D2), D02212.

Brooks, S. D., D. Baumgardner, B. Gandrud, J. E. Dye, M. J. Northway, D. W. Fahey, T. P. Bui, O. B. Toon, and M. A. Tolbert (2003), Measurements of large stratospheric particles in the Arctic polar vortex, *J. Geophys. Res.*, *108*(D20), 4652.

Carslaw, K. S., and T. Peter (1997), Uncertainties in reactive uptake coefficients for solid stratospheric particles .1. Surface chemistry, *Geophys Res Lett*, *24*(14), 1743-1746.

Carslaw, K. S., B. P. Luo, and T. Peter (1995), An Analytic-Expression for the Composition of Aqueous Hno3-H2so4 Stratospheric Aerosols Including Gas-Phase Removal of Hno3, *Geophys Res Lett*, *22*(14), 1877-1880.

Carslaw, K. S., T. Peter, and R. Muller (1997), Uncertainties in reactive uptake coefficients for solid stratospheric particles .2. Effect on ozone depletion, *Geophys Res Lett*, *24*(14), 1747-1750.

Crutzen, P. J. (2006), Albedo enhancement by stratospheric sulfur injections: A contribution to resolve a policy dilemma?, *Climatic Change*, *77*(3-4), 211-219.

Crutzen, P. J., R. Muller, C. Bruhl, and T. Peter (1992), On the Potential Importance of the Gas-Phase Reaction Ch3o2 + Clo -] Cloo + Ch3o and the Heterogeneous Reaction Hocl + Hcl -] H2o + Cl2 in Ozone Hole Chemistry, *Geophys Res Lett*, *19*(11), 1113-1116.

Engel, A., et al. (2009), Age of stratospheric air unchanged within uncertainties over the past 30 years, *Nat Geosci*, *2*(1), 28-31.

Fahey, D. W., et al. (2001), The detection of large HNO3-containing particles in the winter arctic stratosphere, *Science*, *291*(5506), 1026-1031.

Garcia, R. R., and W. J. Randel (2008), Acceleration of the Brewer-Dobson circulation due to increases in greenhouse gases, *J Atmos Sci*, *65*(8), 2731-2739.

Hanisco, T. F., J. B. Smith, R. M. Stimpfle, D. M. Wilmouth, J. G. Anderson, E. C. Richard, and T. P. Bui (2002a), In situ observations of HO2 and OH obtained on the NASA ER-2 in the high-ClO conditions of the 1999/2000 Arctic polar vortex, *J Geophys Res-Atmos*, *107*(D20).

Hanisco, T. F., et al. (2002b), Quantifying the rate of heterogeneous processing in the Arctic polar vortex with in situ observations of OH, *J Geophys Res-Atmos*, *107*(D20).

Hanson, D. R. (1998), Reaction of ClONO2 with H2O and HCl in sulfuric acid and HNO3/H2SO4/H2O mixtures, *J Phys Chem A*, *102*(25), 4794-4807.

Hanson, D. R., and A. R. Ravishankara (1993), Uptake of Hcl and Hocl onto Sulfuric-Acid - Solubilities, Diffusivities, and Reaction, *J Phys Chem-Us*, *97*(47), 12309-12319.

Hegglin, M. I., and T. G. Shepherd (2009), Large climate-induced changes in ultraviolet index and stratosphere-to-troposphere ozone flux, *Nat Geosci*, *2*(10), 687-691.

Jaegle, L., et al. (1997), Evolution and stoichiometry of heterogeneous processing in the Antarctic stratosphere, *J Geophys Res-Atmos*, *102*(11D), 13235-13253.

Kirk-Davidoff, D. B., E. J. Hintsa, J. G. Anderson, and D. W. Keith (1999), The effect of climate change on ozone depletion through changes in stratospheric water vapour, *Nature*, *402*(6760), 399-401.

Peter, T. (1997), Microphysics and heterogeneous chemistry of polar stratospheric clouds, *Annu Rev Phys Chem*, *48*, 785-822.

Prather, M. J. (1992), More Rapid Polar Ozone Depletion through the Reaction of Hocl with Hcl on Polar Stratospheric Clouds, *Nature*, *355*(6360), 534-537.

Ramaswamy, V., M. D. Schwarzkopf, W. J. Randel, B. D. Santer, B. J. Soden, and G. L. Stenchikov (2006), Anthropogenic and natural influences in the evolution of lower stratospheric cooling, *Science*, *311*(5764), 1138-1141.

Randel, W. J., F. Wu, H. Vomel, G. E. Nedoluha, and P. Forster (2006), Decreases in stratospheric water vapor after 2001: Links to changes in the tropical tropopause and the Brewer-Dobson circulation, *J Geophys Res-Atmos*, *111*(D12).

Rasch, P. J., P. J. Crutzen, and D. B. Coleman (2008), Exploring the geoengineering of climate using stratospheric sulfate aerosols: The role of particle size, *Geophys Res Lett*, *35*(2).

Sander, S. P., et al. (2006), Chemical Kinetics and Photochemical Data for Use in Atmospheric Studies, Evaluation Number 15*Rep.*, Jet Propulsion Laboratory, Pasadena, CA.

Santee, M. L., I. A. MacKenzie, G. L. Manney, M. P. Chipperfield, P. F. Bernath, K. A. Walker, C. D. Boone, L. Froidevaux, N. J. Livesey, and J. W. Waters (2008), A study of stratospheric chlorine partitioning based on new satellite measurements and modeling, *J Geophys Res-Atmos*, *113*(D12).

Shi, Q., J. T. Jayne, C. E. Kolb, D. R. Worsnop, and P. Davidovits (2001), Kinetic model for reaction of ClONO2 with H2O and HCl and HOCl with HCl in sulfuric acid solutions, *J Geophys Res-Atmos*, *106*(D20), 24259-24274.

Stevenson, D. S. (2009), Putting the wind up ozone, *Nat Geosci*, *2*(10), 677-679.

Stimpfle, R. M., D. M. Wilmouth, R. J. Salawitch, and J. G. Anderson (2004), First measurements of ClOOCl in the stratosphere: The coupling of ClOOCl and ClO in the Arctic polar vortex, *J Geophys Res-Atmos*, *109*(D3).

Tilmes, S., R. Muller, and R. Salawitch (2008), The sensitivity of polar ozone depletion to proposed geoengineering schemes, *Science*, *320*(5880), 1201-1204.

Voigt, C., et al. (2000), Nitric acid trihydrate (NAT) in polar stratospheric clouds, *Science*, *290*(5497), 1756-1758.

Wennberg, P. O., et al. (1998), Hydrogen radicals, nitrogen radicals, and the production of O3 in the upper troposphere, *Science*, *279*(5347), 49-53.

Wilmouth, D. M., R. M. Stimpfle, J. G. Anderson, J. W. Elkins, D. F. Hurst, R. J. Salawitch, and L. R. Lait (2006), Evolution of inorganic chlorine partitioning in the Arctic polar vortex, *J Geophys Res-Atmos*, *111*(D16).

Chapter 9

Revisiting Ozone Chemistry in the Midlatitude Lower Stratosphere

9.1 Introduction

Building on the themes addressed in the prior chapters, we conclude with a recommendation for a targeted airborne mission to probe the chemistry of the stratosphere over the central United States during summer. With the development of airborne instrumentation with 1) the requisite sensitivity to measure BrO (and ClO) in the lower stratosphere, 2) the ability to measure water vapor concentrations with high precision and unassailable accuracy, 3) the capacity of determining the ambient ice water content, in the unlikely event of encountering direct convective remnants, and 4) the ability to detect the unambiguous isotopic signature of convectively sourced evaporated ice, we have the means to obtain high resolution in situ observations of the role of heterogeneous processing, and halogen activated ozone loss in the presence of high water vapor and low temperatures in the midlatitude lower stratosphere in summer. Such a mission provides a critical benchmark for understanding the probability and extent of these mechanisms under current climate conditions, as well as, a model for

examining the sensitivity of the system to future states brought about by 1) climate forcing due to anthropogenic GHG emissions, and 2) counteractive measures, e.g., the intentional introduction of sulfuric acid aerosol into the lower stratosphere to offset the forcing from rising GHG concentrations. The prospect of a wetter and colder stratosphere in conjunction with moderate halogen loading reopens the case for ozone loss, and consequent increases in UV dosage, over the populated regions of the globe. Furthermore, there is mounting pressure to examine the sensitivity of the current system to localized perturbations in humidity and/or aerosol loading in order to better quantify potential tradeoffs between decreasing shortwave forcing and increasing ozone loss.

Recall that the primary elements of anthropogenic ozone loss consist of 1) halogens, specifically the combination of ClO and BrO, 2) high humidity, caused by the combination of cold temperatures and high water vapor mixing ratios, and 3) reactive surface area, e.g., aerosol and/or clouds, which is necessary for the key reactions that make anthropogenic chlorine available for catalytic ozone loss.

Therefore, in the subsequent sections, we revisit the conclusions of Chapters 6, 7, and 8, and examine what happens to the chemistry of the midlatitude lower stratosphere in the presence of both higher concentrations of BrO, consistent with recent estimates, and convectively sourced water vapor, that leads to large episodic enhancements, as well as an integrated summertime enrichment of water vapor in this region. We estimate the likelihood for heterogeneous processing in these airmasses under present day conditions, and

discuss the sensitivity of the chemistry of this region to future perturbations arising from anthropogenic climate forcing, as well as efforts to mitigate it.

9.2 Halogen Photochemistry

Figure 9.1 shows the results of a quick estimation of the photochemical lifetime of odd oxygen ($O_x = O_3 + O$) in the midlatitude lower stratosphere as a function of ClO mixing ratio. The black lines represent the total estimated photochemical loss rate, while the green and red lines capture the independent contributions from the ClO/BrO and HO_2/NO_2 cycles respectively. To generate these values, we used the photochemical cycles enumerated in Equation 7.4, and reasonable estimates for the diurnal concentrations of the important radical species in the lower stratosphere. Because of the strong dependence of the mixing ratios on latitude, altitude and season, this "back-of-the-envelope" calculation is not meant to provide a precise representation of the rates, however, it is adequate for illustrative purposes. For comparison, the shaded pink rectangle represents an estimate of the photochemical loss rate for O_x in the lowermost stratosphere at ~30°N by [*Solomon et al.*, 1997] [1], in the absence of elevated ClO. The shaded blue rectangle represents their estimate for vertical advection, i.e., "dynamics", in the same region.

[1] Figure 10 of [*Solomon et al.*, 1997] shows estimates of the photochemical lifetime of odd oxygen in the midlatitude lower stratosphere, i.e., 135 hPa at 30°N, to be approximately 3.5×10^7 seconds (~14 months) in the summer and 5.0×10^7 seconds (~19 months) in the winter, in the absence of their prescribed cirrus cloud chemistry. In comparison, their calculated rates for vertical advection in the region are 2×10^7 seconds (~8 months) and 1×10^7 seconds (~4 months) in the summer and winter respectively. These values compare well with our estimates.

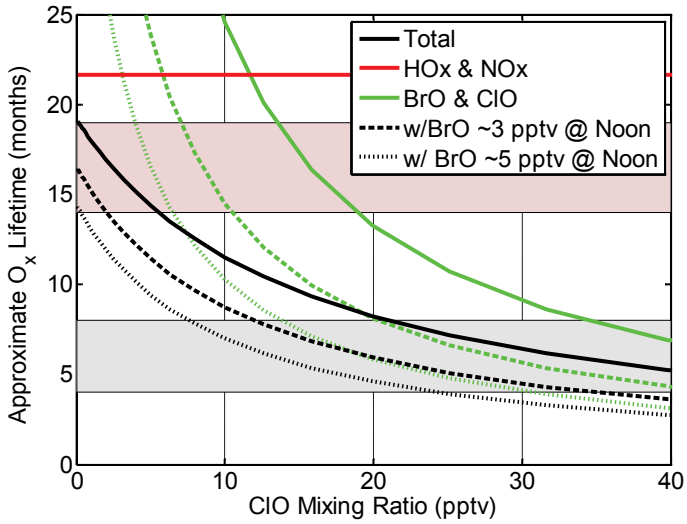


Figure 9.1. Estimated photochemical lifetime of O_x in the midlatitude lower stratosphere. The approximate photochemical lifetime for O_x with respect to the HO_x and NO_x catalytic cycles (solid red line), with respect to the halogen (Cl_x and Br_x) catalytic cycles (green lines), and with respect to the sum (black lines) is plotted as a function of ClO mixing ratio. The solid lines are estimates with diurnally averaged BrO concentrations of ~0.5 pptv, and the dashed and dotted lines show the O_x lifetime response to increases in BrO roughly equivalent to noontime concentrations of ~3 pptv and 5 pptv respectively. The shaded pink region represents the estimated photochemical lifetime of O_x in the midlatitude lower stratosphere by [*Solomon et al.*, 1997] in the absence of ClO enhancements, and the shaded blue region shows their estimate of the dynamical lifetime in the same region.

In Chapter 7, we examined the proposal of [*Solomon et al.*, 1997] and analyzed the humidity structure of the midlatitude lower stratosphere. We found no evidence to support sustained cirrus in this region. In situ temperature and water vapor data, acquired by the MMS and HWV-E instruments flown aboard NASA's ER-2 aircraft, show that the midlatitude lower stratosphere is very dry, well below ice saturation just a few hundred meters above the local tropopause. While approximately 60% of the data just below the tropopause showed evidence of ice saturation, only 2% of the data in the first 0.5 km above the tropopause

were within 1 K of ice saturation (see Figure 7.4). All data at altitudes greater than 0.5 km above the local tropopause were under-saturated. Furthermore, the in situ airborne particle instruments showed no evidence of clouds in this region. Under these conditions, cirrus lifetimes are expected to be exceedingly short, on the order of minutes to hours [*Eric Jensen, personal communication*].

The low humidity in this region also severely limits the potential for heterogeneous activation. The rate constant defining the heterogeneous conversion of inorganic chlorine to free radical form is given by $k_{het} = (v/4)\gamma S_a$, where v is the average molecular speed, γ is the reaction probability per collision, and S_a is the available aerosol/cirrus cloud reactive surface area. Figure 7.9 shows the rapid decline in measured reactive surface area, S_a, above the tropopause at midlatitudes, and Figure 7.10 shows simultaneous estimates of the sulfuric acid weight percent in the ambient aerosol, and γ's for two of the primary heterogeneous reactions using in situ data. The reaction probabilities, γ, are extremely sensitive to sulfuric acid weight, with high values "poisoning" the reactive potential of the aerosol, e.g., [*Kolb et al.*, 1996; *Ravishankara et al.*, 1998] *and references therein*. Consistent with the increasing relative humidity and low temperatures just below the tropopause, the sulfuric acid weight percent decreases up to the tropopause as water is taken up by the aerosol, and the reaction probabilities increase. With the sudden transition to under-saturated air above the local tropopause, the sulfuric acid weight percent increases and the reaction probabilities, γ, decrease dramatically. The data show that the chlorine activation potential on aerosols falls off by an order of magnitude or more,

thereby inhibiting the release of ClO from its reservoir species, in the very region where midlatitude ozone trends become statistically significant, i.e., at 1 − 2 km above the tropopause.

Moreover, ClO measurements acquired in situ by the Harvard ClO instrument in the same region do not show evidence of episodic or sustained enhancements (see Figure 7.12). Though the precision of the instrument limits our ability to distinguish between 0.1 and 1 or 2 pptv, we can distinguish between a few pptv and the levels of ClO required to substantially increase the photochemical loss rate of ozone in the critical region from the tropopause to >~3 km above it. Except for a single encounter with an airmass remarkable for unusually high aerosol loading from Mt. Pinatubo eruption, the observational dataset does not support the requisite ClO concentrations.

The flight of April 30, 1993, during the SPADE mission, is the only example with elevated ClO mixing ratios in the lower stratosphere (see Figure 7.12b). In this airmass, ClO mixing ratios measured 90 pptv at an altitude ~0.5 km above the midlatitude tropopause. [*Keim et al.*, 1996] attribute this event to fast heterogeneous activation on a plume of cold aerosol in conjunction with moderate levels of ambient water vapor. The aerosol plume is thought to be a remnant from the Mt. Pinatubo eruption nearly two years prior. As we stated in Chapter 7, this phenomenon was not observed again in any of the other SPADE flights, or in >40 midlatitude flights following the decay of the Mt. Pinatubo aerosol. If judged by the frequency of in situ observation, the event highlighted in Figure 7.12b must be considered exceedingly rare. However, the [*Keim et al.*, 1996] analysis directly

demonstrates the extreme sensitivity of the chemistry in the midlatitude lower stratosphere to localized perturbations in aerosol surface area density and humidity.

Additional BrO in the lower stratosphere, equivalent to noontime concentrations of~3 pptv and 5 pptv, however, greatly increases the effectiveness of catalytic loss by ClO. This is evident in the dramatic decrease in the estimated O_x lifetime with respect to the halogen cycles for a fixed concentration of ClO, indicated by the dashed and dotted green lines in Figure 9.1, and in the simultaneous decrease in the total photochemical lifetime (dashed and dotted black lines). The enhancements in ClO needed to influence in situ ozone loss are reduced in the presence of BrO. Thus, while the mechanism involving ClO alone does not appear to be capable of explaining the observed ozone loss in the midlatitude lower stratosphere, the link between the BrO and ClO cycles, rate limited by the reaction $ClO + BrO \rightarrow Cl + Br + O_2$ may provide an explanation, especially when considered in conjunction with convectively sourced humidity over the central United States in summer, see Section 9.2.

The Cl_x (Cl and ClO) catalytic cycle requires a reaction partner for ClO following the reaction $Cl + O_3 \rightarrow ClO + O_2$, in order to complete the cycle and return a Cl radical. In the upper stratosphere, this cycle is completed by reaction of $ClO + O \rightarrow Cl + O_2$. In the polar vortices, where ambient temperatures are extremely cold and ClO concentrations are elevated due to heterogeneous processing, the cycle is completed through the self reaction, $ClO + ClO \rightarrow ClOOCl$, followed by photolysis, $ClOOCl + h\nu \rightarrow ClOO + Cl$, and

decomposition, $ClOO + M \rightarrow Cl + O_2$. In the lower stratosphere, i.e., <20 km, the reaction of ClO with BrO provides a means of completing the cycle through the following series of reactions, e.g., [*Yung et al.*, 1980] *and references therein*.

$Cl + O_3 \rightarrow ClO + O_2$

$Br + O3 \rightarrow BrO + O_2$

$\underline{ClO + BrO \rightarrow Cl + Br + O_2}$ (rate-limiting step)

Net: $2O_3 \rightarrow 3O_2$

This reaction mechanism, and the extent of its associated ozone loss, depends critically upon the availability of BrO.

Measurements of BrO column amounts, as well as vertical profiles, from satellites, aircraft and balloons, consistently show values higher than estimates obtained from long-lived stratospheric bromine sources, e.g., CH_3Br and the halons [*Salawitch et al.*, 2005], *and references therein*. One example of vertical profile measurements of BrO in the stratosphere, obtained at 43.8°N in October, 2002, using both balloon-borne and satellite-based instruments shows daytime values of ~4 pptv near the tropopause (~14 km), growing to ~12 pptv at 20 km [*Sioris et al.*, 2004]. These measurements, combined with observations of ~4 pptv of BrO in the tropical transition layer, suggest that some combination of convectively sourced short lived, or very short lived (VSL) organic bromine compounds, and/or their dissociated products, are directly injected near the base of, or into, the TTL or the lowermost stratosphere, e.g., [*Salawitch et al.*, 2005; *Sioris et al.*, 2004], *and references therein*. These data and subsequent analyses suggest, therefore, that inorganic bromine concentrations (Br_y) at and above the

tropopause are significantly, i.e., 4 to 8 pptv, higher than the values traditionally assumed in chemical models used for ozone trend assessments.

[*Salawitch et al.*, 2005] explore the impact of higher concentrations of Br_y on vertical ozone trends from 1980 to the end of 2000 with the Atmospheric and Environmental Research (AER) 2D chemical model. The additional Br_y, ~20 – 50% of which exists as BrO, is imposed as a constant enhancement at all levels over the modeled time period. The model results with enhanced Br_y yield better overall agreement with the measured trend in midlatitude ozone.

The modeled effect is greatest in the lower stratosphere, with a peak between 15 and 16 km. This is evident in Figure 5d from their paper, which is reproduced as Figure 9.2 here. As expected, the maximum increase in modeled ozone loss, plotted in Figure 9.2d, is coincident with an increase in the fractional contribution of the BrO + ClO loss mechanism shown in Figures 9.2a-c. The additional BrO in the model, which is consistent with recent measurements, significantly enhances the effectiveness of Cl_x catalyzed ozone loss in the lower stratosphere, by providing a reaction partner for ClO. This is particularly evident in the modeled airmass of March, 1993, where ClO concentrations were enhanced because of the increase in aerosol loading following the eruption of Mt. Pinatubo. The modeled O_3 loss at ~15 to 16 km is approximately 30% greater for an addition of 4 pptv Br_y, and ~60% greater with an addition of 8 pptv. Given that dynamics within the model is unchanged between runs, this response is entirely due to the chemical perturbation. The response of the total loss rate for O_x plotted in Figure 9.1 is of a similar magnitude. [*Salawitch et al.*, 2005] note that the

model is poor at explaining shorter term year-to-year variability, but the improved agreement with the observed trends lends credence to the importance of the coupled BrO/ClO mechanism for amplifying ozone loss in the lower stratosphere.

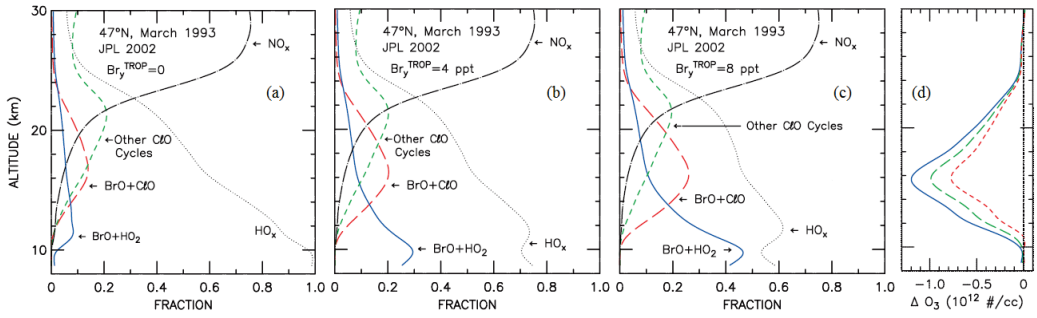


Figure 9.2. Fraction of odd oxygen loss by various catalytic cycles. Panels (a)-(c) show AER model results at 47°N, March 1993, for runs with additional Br_y equal to 0, 4, and 8 pptv respectively. Panel (d) shows the difference between modeled ozone profiles at 47°N, from March 1980 and March 1993, capturing the integrated ozone loss over this period. The red short-dashed line had 0 pptv additional Br_y, the green long-dashed line had 4 pptv additional Br_y, and the solid blue line at had 8 pptv additional Br_y. Figure adapted from Figure 5d in [*Salawitch et al.*, 2005].

The importance of convection as a significant source of stratospheric Br_y introduces an important link between the future of stratospheric photochemistry and climate. Increases in convective frequency and/or in convective strength in response to a warmer and moister troposphere have the potential to amplify this source term. Thus, even with the phaseout of the longer-lived halons, bromine levels may be slow to decline.

9.3 Stratospheric Moisture

9.3.1 Moist Plumes

Convective intrusions primarily increase the humidity of the lower stratosphere. There is ample evidence of convective moistening in the lower stratosphere over the central United States during the summer [*Dessler*, 2009; *Dessler et al.*, 2007; *Hanisco et al.*, 2007], with moist plumes extending into the stratospheric overworld. These localized increases in humidity have the potential to promote halogen activation in this region, and consequently, to enhance in situ ozone loss particularly in the presence of simultaneously elevated levels of BrO.

Figure 9.3 shows water vapor data acquired by HWV at midlatitudes on two summertime missions (AVE and TC4) staged out of Houston, TX. Ambient mixing ratio is plotted as a function of potential temperature, and color-coded by the day of the year. The data show evidence of frequent hydration plumes occurring throughout the stratosphere, with the magnitude of the perturbation, above the nominal ~5 ppmv background, decreasing as a function of altitude. This figure also demonstrates the consistency in background water vapor profiles from year to year. These hydration events, attributed to the evaporation of convectively injected ice, occur up to the highest altitudes that the aircraft flew, i.e., 19 km or ~440 K. In the most extreme cases, they are seen to add ~>15 ppmv of water vapor directly into the stratospheric "overworld", i.e., above 380 K, circumventing the temperature control of vertical transport through the tropical tropopause.

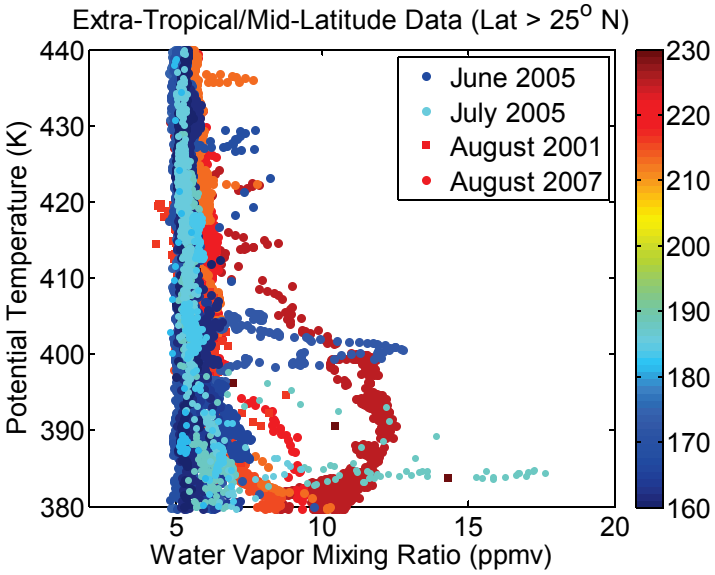


Figure 9.3. Hydration events in the extra-tropical stratosphere binned by day of year. Water vapor data acquired over the central United States during three summertime missions, AVE, AVE-WIIF and TC4, are plotted as a function of potential temperature. The data are color-coded by the day of the year. Hydration plumes of varying depth are evident throughout the stratosphere, with the magnitude of the perturbation, above the nominal ~5 ppmv background, decreasing as a function of altitude.

The proposed mechanism for these moist plumes is that ice is lofted from the troposphere during powerful continental thunderstorms, e.g., [*Wang*, 2003], and rapidly evaporates in the severely undersaturated stratosphere. The expected lifetime for ice in this environment is extremely short, and the layer-like quality of most of these events suggests that evaporation is faster than sedimentation. Furthermore, this mechanism allows for the deposition of large quantities of water vapor as ice, while causing little change in other tracers.

Figure 9.4 shows vertical profiles of extra-tropical H_2O, O_3 and CO_2 (blue dots). The enhancements in water vapor are highlighted in the red circles. The concomitant O_3 and CO_2 data are unaltered, lying directly in line with their respective "background" profiles. Other perturbations are evident in the dataset,

but they are not associated with the hydration events. This observation substantiates the assertion that the water vapor enhancements arise from the injection of air with a disproportionately elevated total water content.

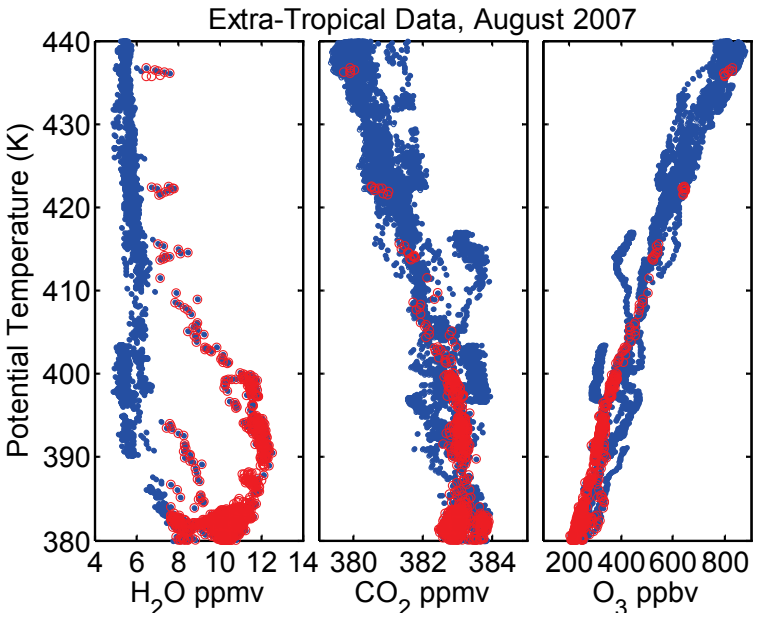


Figure 9.4. Extra-tropical profiles of H_2O, CO_2, and O_3. All of the extra-tropical data are plotted in the blue dots. The moist plume encounters, evident only in the water vapor profile, are highlighted in the red circles.

In situ observations of H_2O and HDO provide unique insight into the thermal histories of sampled airmasses [*Keith*, 2000a; *Moyer et al.*, 1996]. Because HDO is heavier than H_2O, it is preferentially removed during condensation. Sedimentation of the isotopically enriched condensate results in isotopically depleted air, and the ratio of HDO/H_2O decreases as air ascends from the boundary layer through the tropopause. Measurements made during an encounter with a moist plume, on July 7, 2005, during the AVE-WIIF mission, bear the distinct isotopic signature of evaporated ice that originated in the

troposphere. These data provide incontrovertible evidence for the mechanism of convective injection into the midlatitude stratosphere.

Isotopic depletions, δD, are typically reported as the ratio of HDO to H_2O in the airmass, relative to a reference ratio, and scaled by 1000. Equation 9.1 shows the derivation of δD, with SMOW equal to the ratio of HDO to H_2O in Standard Mean Ocean Water, i.e., 3.115×10^{-4}.

$$\delta D = ([HDO]/[H2O] \text{ x } (1/SMOW) - 1) \text{ x } 1000 \text{ [‰, per mille]} \quad (9.1)$$

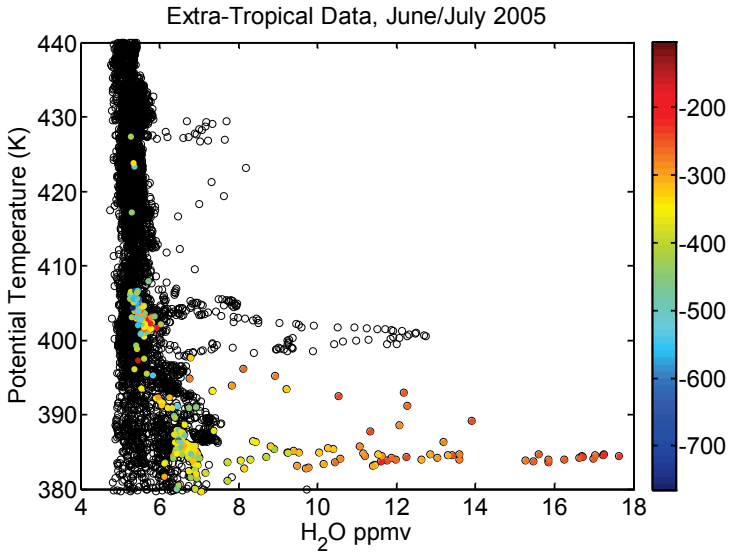


Figure 9.5. Hydration events in the extra-tropical stratosphere binned by δD. Water vapor mixing ratios acquired in June/July of 2005 over the central United States by HWV (black open circles) are plotted as a function of potential temperature. δD values measured by the HOXotope instrument are plotted in the colored circles for the plume encounter on July 7, 2005, with the colorbar providing the scale for δD.

Figure 9.5 shows in situ water vapor data (black open circles) from the AVE and AVE-WIIF missions as a function of potential temperature. The colored circles correspond to mixing ratio data from the flight of July 7, 2005, where the color indicates the level of HDO depletion in the airmass. In the plume, we see

that δD values measure ~−280‰, and are significantly elevated above the nominal background value of ~−450‰. Utilizing average measurements of H_2O in the plume, ~15.6 ppmv, and in the background, ~6.5 ppmv, we calculate that the water vapor added in this event was ~9.1 ppmv. Similarly, normalized HDO* measurements (HDO* = HDO/SMOW) , yield values of ~11.2 ppmv in the plume, and ~3.5 ppmv in the background, indicating that the added HDO* was ~7.7 ppmv, and that the convective component had an average δD = −150‰. The isotopic enrichment in the plume is consistent with the evaporation of tropospheric ice particles generated in, and lofted by powerful convective storms.

Furthermore, the observations show that water vapor in the overworld stratosphere is, in general, isotopically enriched relative to the values expected from the descending arm of the Brewer-Dobson circulation at midlatitudes. This enrichment is consistent with the continual contribution of isotopically heavy water from extra-tropical convection. Following the example of [*Hanisco et al.*, 2007], we can estimate the contribution of convective moisture to the water vapor budget of midlatitude summertime stratosphere.

In the limit that water vapor in this region is simply a combination of two sources, 1) a tropical source (corrected for methane oxidation), and 2) a local convective source, the fractional contribution of the convective source, $\chi_{H2OCb} = [H_2O]_{Cb}/[H_2O]_{ML}$, can be computed utilizing the following simple relation:

$\chi_{H2OCb} \cdot \delta D_{Cb} + (1 - \chi_{H2OCb}) \cdot \delta D_e = \delta D_{ML}$, such that

$$\chi_{H2OCb} = (\delta D_{ML} - \delta D_e) / (\delta D_{Cb} - \delta D_e), \quad (9.2)$$

where $[H_2O]_{Cb}$ is the absolute water vapor contribution from convection, $[H_2O]_{ML}$ is the observed "background" concentration of water vapor at midlatitudes, and δD_{Cb}, δD_e, δD_{ML}, refer to δD's of the convectively lofted ice, air crossing irreversibly into the stratosphere in the tropics, and the summertime observations at midlatitudes, respectively. If we assume $\delta D_{Cb} = -100$ ‰ to −200 ‰, for the ice, $\delta D_e = -560$‰ for the lighter tropical air based on measurements made during CRAVE and TC4 in the tropics, and $\delta D_{ML} = -450$‰, based measurements acquired during AVE-WIIF, the approximate convective contribution to the water vapor budget of the midlatitude, summertime lower stratosphere is ~27% ± ~15%[2].

Back-trajectory analyses also demonstrate convective influence in this region [*Pfister et al.*, 2010; *Sayres et al.*, 2010]. Back-trajectories for the flight of August 13, 2007, were analyzed to determine the fractional convective influence (FCI) for a cluster of 15 trajectories around selected aircraft times. A cluster consists of points at 3 altitudes, i.e., one above, one below, and one at the level of the aircraft, with 5 points at each altitude. The 5 points are aligned perpendicular to the aircraft track. Each trajectory is run back for 20 days. To evaluate the impact of convection, the trajectories are mapped onto a time varying field of global geostationary infrared imagery, i.e., brightness temperatures. Convective interaction is said to occur if the cloud altitude, based upon thermally adjusted brightness temperatures, is equal to or greater than the trajectory altitude. Parcels

[2] This is lower than the original estimate of [*Hansico et al.*, 2007], because the value for the tropics used in this analysis is heavier than their estimate by ~90‰.

must come within 30 km (~0.25°) of a given pixel to be considered impacted by convection.

In Figure 9.6, the fractional convective influence is shown (colored dots) for selected data points during the flight of August 13, 2007. The moist plume event shows a high fraction of convective influence, ~70%, while the background profile shows a correspondingly low fraction. A plot of convective influence as a function of theta for data from the whole TC4 mission (not shown) demonstrates that the fraction decreases rapidly above 380 K, but, consistent with the water vapor and HDO data, is non-zero up to ~440 K in both the tropics and extra-tropics.

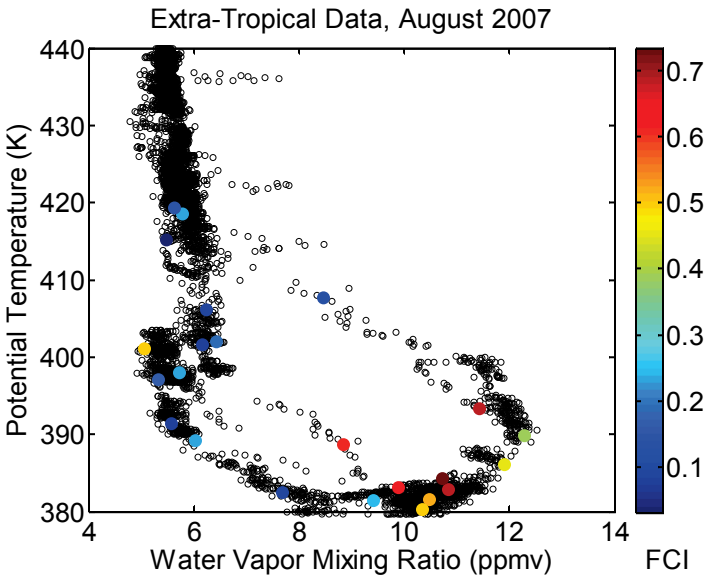


Figure 9.6. Hydration event in extra-tropical stratosphere binned by convective influence. Water vapor mixing ratios acquired in August of 2007 over the central United States by HWV (black open circles) are plotted as a function of potential temperature. Selected water vapor data for the moist plume encounter on August 13, 2007, are plotted in the colored circles with the colorbar providing the scale for the fractional convective influence. The method for determining the fractional convective influence is described in the text.

The importance of convection to the dual delivery of moisture and short-lived chemical compounds to the lower stratosphere renders its evolution, in response to anthropogenic climate forcing, especially pertinent to the future of stratospheric ozone and forecasts of UV dosage. As mentioned above, this mechanism has the potential to become more significant as both surface temperature and tropospheric specific humidity increase in a warming climate. Thus, many questions arise, e.g., What is the frequency of the deep convective events that deposit ice into the UT/LS? How does the frequency of these events vary with respect to season, and geographic location? What is the net impact of these events on the water budget of the lower stratosphere? What other chemical species are transported from the boundary layer to the stratosphere via this mechanism? And, how do we expect the frequency of these events to evolve in response to climate forcing?

9.3.2 Inefficient Dehydration in the TTL

While it is clear that convection provides a means of bypassing the thermal regulation of water vapor entering the stratosphere that occurs at the cold-point tropical tropopause, frequent supersaturation, observed in both clear air and in clouds in the extra-tropics and in the TTL, also limits the efficiency of thermal control. These observations suggest that models, which assume that the water vapor entering the stratosphere is simply equal to the ice saturation mixing ratio set by the temperature of the tropical tropopause, i.e., $[H_2O]_e = SMR_i(T_{trop})$, will

systematically under-predict the moisture content of the stratosphere, and may be missing terms responsible for both short-term variability and long-term trends.

In Chapter 6, we examined the humidity structure of the TTL and UT/LS utilizing HWV measurements acquired in situ during six aircraft campaigns that covered the tropics and the midlatitudes. Observations of water vapor and relative humidity with respect to ice (RH_i) in the tropical and sub-tropical upper troposphere show that supersaturation with respect to ice is common in both clear air and in the presence of cirrus. The observation of frequent clear air supersaturation, particularly prevalent at the cold temperatures in the vicinity of the tropopause, with RH_i measuring as high as ~160%, impacts model estimates of stratospheric entry level water vapor, $[H_2O]_e$, as well as, modeled cloud occurrence frequencies. The existence of clear air ice supersaturation, particularly in regions of ascent into the stratosphere, implies that dehydration via ice nucleation and sedimentation is not occurring, and that the actual $[H_2O]_e$ will be greater than the saturation mixing ratio with respect to ice at the cold-point temperature, SMR_i. Similarly, the observation of sustained supersaturation in the presence of cold, thin near tropopause cirrus, also imply that thermal control is not a wholly efficient dehydration mechanism. It is important, therefore, that models attempting to capture the evolution of stratospheric water carefully consider the impact of supersaturation.

Furthermore, observations in the highest and coldest regions of the troposphere of extreme supersaturion, i.e., relative humidity values exceeding ~>165%, which is the nominal threshold for homogeneous ice nucleation,

challenge the current understanding of ice nucleation and cloud formation in this region. Recent laboratory and in situ data examining the effect of chemical composition on aerosol nucleation properties show that the presence of organic components in aerosol may inhibit nucleation. In some studies, there is evidence that organic-rich liquid aerosol may become ultra-viscous with cooling and undergo a phases transition to a glassy solid state, rather than nucleate ice, e.g., [*Murray*, 2008; *Zobrist et al.*, 2008]. In addition, these glassy particles are expected to have a reduced efficiency for water vapor uptake. The impact of chemical composition on the freezing behavior of aerosol, introduces a mechanism for anthropogenic activities to impact the humidity structure of the UT/LS, as well as the frequency and distribution of cirrus clouds.

As a whole, the results from the humidity analysis of Chapter 6, and the evidence for convective influence discussed in Section 9.2.1, highlight the complexity of modeling stratospheric water concentrations, and underscore the need for accurate in situ measurements for both detailed mechanistic analyses, and for long-term trend detection.

9.3.3 Water Vapor Accuracy

Demonstrably accurate water vapor measurements are crucial to the success of any future missions and modeling efforts. Measurement accuracy is essential for both mechanistic confirmation, and for the unambiguous quantification and attribution of long-terms trends and/or interannual variability in stratospheric concentrations. The measurement discrepancies plaguing the UT/LS

water vapor community currently provide limited constraint on important mechanistic controls of water vapor entering the stratosphere, e.g., 1) the humidity threshold for ice nucleation in the UT, 2) the dehydration efficiency of the cold-point tropical tropopause, and 3) the contribution of continental convection to water vapor budgets at midlatitudes. Furthermore, there is no consensus on the existence of a long-term trend. This enduring uncertainty limits our understanding of the processes currently controlling stratospheric water, with model validation dependent on which dataset is chosen, and undermines our ability to predict the evolution of stratospheric water vapor in response to climate forcing.

In Chapter 3, we provided a strategy for rigorous laboratory calibration and in-flight validation, with Chapters 3 and 4 providing a practical realization of this approach with HWV as the example. The methodology calls attention to the unavoidable accumulation of uncertainty in the chain from laboratory calibration, with an accepted reference standard or standards, to the attainment of the in situ measurement, and emphasizes the importance of identifying and documenting of these sources of uncertainty, as well as the need for transparency at every step. A community-wide adoption of these methods, which are generally applicable to every in situ instrument, is required in order to resolve the remaining in situ measurement discrepancies.

9.4 Heterogeneous Chemistry

9.4.1 Response to High Water Vapor & Cold Temperatures

In order to quantify the combined effect of cold temperatures and high moisture content on heterogeneous chemistry in the lower stratosphere, we analyzed the data from the moist plume encounter at 33°N, 90°W over the south-central U.S., on August 13, 2007. This event, evident in Figures 9.3, 9.4, and 9.6, is notable for both for the magnitude of the enhancement, and for its broad vertical extent. Water vapor mixing ratios in the plume, which extends from 380 K, i.e., just above the position of the tropopause, to ~410 K, are elevated by ~6 ppmv above the nominal extra-tropical background.

Recall from the discussion in Section 9.2, that the rate constant for heterogeneous reaction on sulfuric acid aerosol is given by $k_{het} = (v/4)\gamma S_a$, where v is the average molecular speed, γ is the reaction probability per collision, and S_a is the available aerosol/cirrus cloud reactive surface area density. The units of k_{het} are s^{-1}, thus, $1/k_{het}$ provides yields the first-order time constant of the reaction. Utilizing the in situ temperature, pressure, and water vapor data we calculate the sulfuric acid weight percent of the aerosol, and the water activity, which is effectively equivalent to the relative humidity with respect to liquid water, using the parameterization of [*Carslaw et al.*, 1995]. These values are then fed into a program, adapted from the JPL-06 recommendations, i.e., [*Sander et al.*, 2006], and the work of [*Hanson*, 1998], which determines the reaction probabilities, γ, for the three primary heterogeneous reactions, responsible for the conversion of chlorine from its reservoir species, HCl, $ClONO_2$, and HOCl, to its catalytically

active forms, refer to Equations 8.1 – 8.3 Finally, the rate is determined using the calculated γ, and a reasonable estimate for the ambient surface area density, e.g., ~1 $\mu m^2/cm3$.

The results of our calculations for the extra-tropical stratosphere (>380K, >25°N) are shown in Figure 9.7. The first-order time constant for the reaction $ClONO_2 + HCl + aerosol \rightarrow HNO_3 + Cl_2 + aerosol$ is plotted as a plotted as a function of ambient temperature. The data are color-coded by water vapor mixing ratio. Consistent with the results of similar calculations shown in Figure 7.11[3], the majority of the data yield time constants greater than 100 hours. For the elevated mixing ratios in the plume, however, the time constants are significantly decreased, with the highest mixing ratio data at the coldest temperatures showing time constants of a few hours, comparable to the photolysis lifetime for $ClONO_2$.

Note also, that the rate is directly proportional to surface area density. A conservative estimate of 1 $\mu m^2/cm^3$ was used in this analysis. It is not unusual for values in this region of the lower stratosphere to be a factor of 2 to 4 higher. Furthermore, with an increase of surface area densities to ~10 – 20 $\mu m^2/cm^3$, consistent with the aerosol plume encounter on April 30, 1993, it is evident that heterogeneous processing is not only possible, but highly probable in regions where water vapor mixing ratios have been enhanced.

[3] Note that the geographical coverage of the lower stratospheric humidity analysis in Chapter 7 was largely confined to regions around Moffett Field, CA, at 37°N, 122°W, with few measurements obtained over the convectively active regions in the central United States.

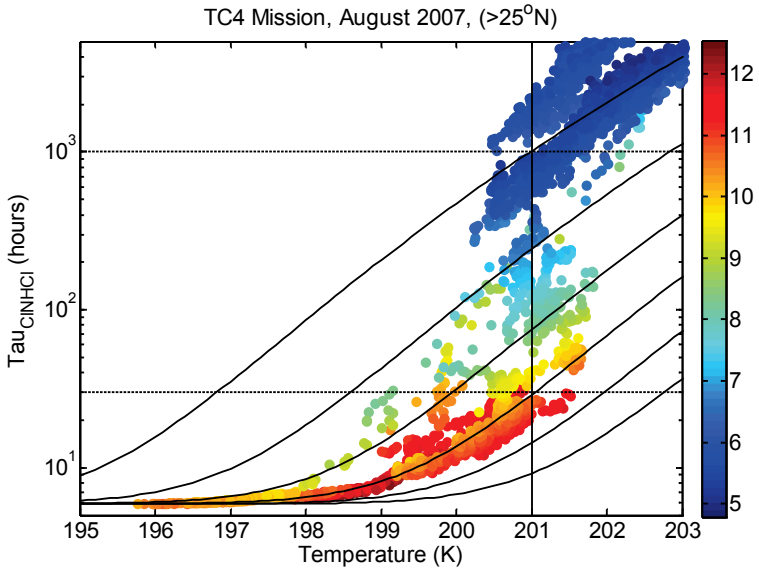


Figure 9.7. First-order lifetime of $ClONO_2$ with respect to heterogeneous loss via $ClONO_2 + HCl + aerosol \rightarrow HNO_3 + Cl_2 + aerosol$. The calculated lifetimes are plotted as a function of ambient temperature, and color-coded by mixing ratio. Only estimates for extra-tropical, stratospheric data are shown, i.e., latitude >25°N, and potential temperature >380 K. The solid black lines correspond to lifetimes calculated at fixed mixing ratios from 5 to 15 ppmv.

9.4.2 Probing Heterogeneous Chemistry with HO_x

The WB-57 aircraft in TC4 was equipped primarily with cloud property instrumentation and a selection of atmospheric tracers, including water vapor, and its isotopologues, however, the radical instruments to measure ClO, BrO, and the HO_x species, OH and HO_2, were not represented. As a result, there was no direct evidence for or against in situ heterogeneous processing in these airmasses.

Analyses of in situ measurements of OH and HO_2 from the SOLVE mission show that HO_x measurements are a highly sensitive probe of recent heterogeneous processing [*Hanisco et al.*, 2002], and discussion in Chapter 8. OH concentrations throughout the lower stratosphere are typically well represented by a parameterization with solar zenith angle (SZA) as the only input. Deviations by

more than 30% from the correlation with SZA are extremely rare and are an immediate indication of unusual sources or sinks for HO_x. Thus, because of the close coupling between the HO_x and Cl_x radical families through the fast photochemical production and loss of HOCl (see Reactions 8.4 and 8.5), additional HOCl production via the heterogeneous reaction $ClONO_2 + H_2O +$ aerosol $\rightarrow HOCl + HNO_3 +$ aerosol (Reaction 8.2), or HOCl loss via $HOCl + HCl +$ aerosol $\rightarrow Cl_2 + H_2O +$ aerosol (Reaction 8.3), can lead to large deviations in OH concentrations.

In Chapter 8, we showed that the dramatic suppression of OH at sunrise on the flight of January 23, 2000, provided unmistakable evidence for the recent occurrence of Reaction 8.3. This reaction is critical under the low $ClONO_2$ conditions characteristic of the late winter/early spring polar vortex, because it provides a reaction partner for the remaining HCl reservoir, thus prolonging chlorine activation and extending ozone loss. The results of a simple photochemical model, constrained by the observations, showed that the only way to generate the very low OH concentrations in this airmass, was through the near complete removal of the nighttime HOCl reservoir in the cold, dark >24 hour interval prior to sunrise. Note that the high concentrations of ClO in this airmass, though indicative of heterogeneous processing, are not sufficient to confirm the existence of Reaction 8.3. The OH data, therefore, provide the most direct evidence to date for this reaction.

Furthermore, the analysis of [*Hanisco et al.*, 2002] uses both high and low perturbations in OH on the flight of March 5, 2000, to constrain the rates of

heterogeneous processing within a PSC. The positive and negative excursions of OH along the flight track from a fit to SZA are consistent with the heterogeneous production of HOCl via Reaction 8.2, and removal via Reaction 8.3. The rates of these reactions, however, are highly sensitive to assumptions regarding the microphysical properties of the particle surface area within the PSC, and, as the steady-state analysis of [*Hanisco et al.*, 2002] demonstrates, so is the response of OH. Comparisons, therefore, between calculated OH concentrations during the processing event and direct observations of OH, provide a means of testing the mechanistic details of these heterogeneous reactions in situ. [*Hanisco et al.*, 2002] conclude that both Reactions 8.2 and 8.3 occur, and that the total heterogeneous reaction rate is dominated by reaction on supercooled ternary solution ($HNO_3/H_2SO_4/H_2O$) liquid aerosols. Furthermore, they are able to recommend modifications to the parameterized rates for both high HNO_3 content in the aerosol, and low HCl coverage on the solid nitric acid trihydrate (NAT) particles.

Together, these analyses of in situ HO_x data demonstrate 1) the remarkable utility of in situ OH measurements for detailed mechanistic studies of heterogeneous processing and, 2) that cold liquid aerosol are extremely effective at promoting heterogeneous processing.

9.5 Reactive Surface Area

One of the most economically and technologically feasible geoengineering proposals, intended to counter the longwave climate forcing due to rising GHG (CO_2/CH_4) concentrations, is to reduce the incoming shortwave radiation through

the deliberate injection of sulfur into the lower stratosphere to form aerosols, i.e., to emulate a volcanic eruption, e.g., [*Crutzen*, 2006; *Keith*, 2000b; *Robock*, 2000]. It is well documented that the large burden of sulfur injected into the stratosphere by the eruption of Mt. Pinatubo in 1991 led to the rapid formation of sulfuric acid aerosols, a substantial reduction in direct solar radiation, and a subsequent cooling at the Earth's surface that persisted for ~2 years, e.g., [*Robock*, 2000].

The creation of reflective surface area, however, is equal to the creation of reactive surface area, which promotes the heterogeneous conversion of inorganic halogen species to their photochemically active forms. Such a scheme, therefore, has the potential to greatly amplify the probability of chemical ozone loss. In fact, post Pinatubo, the destruction of polar ozone was noticeably enhanced for a few years. Furthermore, dramatic ozone column reductions were observed at midlatidues following the eruption, and, as was noted in Section 9.1, the single largest enhancement of ClO in the midlatitude lower stratosphere was observed in an aerosol plume attributed to Mt. Pinatubo nearly two years after the eruption [*Keim et al.*, 1996; *Smith et al.*, 2001], see also Figure 7.12.

The modeling work of [*Tilmes et al.*, 2008], examines the sensitivity of polar ozone depletion to the continuous injection of sulfur into the stratosphere. They find that such a scheme would appreciably increase the extent of Arctic ozone depletion during the present century and considerably delay the expected recovery of the Antarctic ozone hole. While the [*Tilmes et al.*, 2008], study focuses upon the effects of intentional aerosol enhancement at the poles, the [*Keim et al.*, 1996] analysis of in situ data directly demonstrates the extreme

sensitivity of the chemistry in the midlatitude lower stratosphere to localized aerosol density perturbations.

9.6 Summary

Of the many significant human and environmental vulnerabilities to irreversible climate change, variations in UV dosage levels, dependent upon the evolution of stratospheric ozone, are among the most important for life at Earth's surface. Figure 9.7 reveals that modest increases in water vapor concentrations, and/or small decreases in temperature can trigger the extremely rapid conversion of HCl and $ClONO_2$ to ClO, and Figures 9.1 and 9.2 show that, in the presence of BrO, ClO enhancements can become a significant driver of catalytic ozone loss in the midlatitude lower stratosphere. The extreme sensitivity of this mechanism to subtle changes in stratospheric humidity calls attention to the tight coupling between climate, chemistry and UV dosage.

Projected increases in stratospheric water vapor concentrations due to a combination of warmer tropical tropopause temperatures, e.g., [*Austin and Reichler*, 2008; *Kirk-Davidoff et al.*, 1999], increases in the direct convective injection of water as ice into the sub-saturated lower stratosphere, and/or the oxidation of rising stratospheric methane concentrations, combined with the projected cooling of the stratosphere due to rising GHG concentrations (including water vapor), dramatically increase the probability of heterogeneous chlorine activation in the lower stratosphere, and consequently enhance the prospect for future ozone loss even as halogen concentrations are expected to decline.

Furthermore, quantifying the potential trade-off between intentional climate engineering, i.e., the injection of sulfur into the lower stratosphere, and its deleterious chemical impacts, demands that the sensitivity of this mechanism be fully characterized.

We have numerous recent observations of high water and low temperatures at midlatitudes, evident in Figure 9.3. However, we do not have the requisite BrO and ClO observations in the presence of high water vapor and low temperatures, to 1) assess the impact of these events on ozone chemistry under current climate conditions, and 2) constrain projections of their future impact as the fundamental stratospheric boundary conditions change in response to anthropogenic climate forcing. The combination of instrumentation currently available provides 1) direct in situ observations of BrO and ClO concentrations in the lower stratosphere, 2) a measure of the episodic moistening associated with convective incursions – H_2O, Total Water, HDO, 3) the means of determining the integrated effect of convection on the water vapor budget of this region – H_2O, HDO, 4) the ability to detect the impact of this added moisture on the heterogeneous reactions that liberate ClO from its reservoir species – ClO, H_2O, OH, and 5) the capacity to quantify the net effect of these processes on ozone in the midlatitude lower stratosphere.

References

Austin, J., and T. J. Reichler (2008), Long-term evolution of the cold point tropical tropopause: Simulation results and attribution analysis, *J Geophys Res-Atmos*, *113*.

Carslaw, K. S., B. P. Luo, and T. Peter (1995), An Analytic-Expression for the Composition of Aqueous Hno3-H2so4 Stratospheric Aerosols Including Gas-Phase Removal of Hno3, *Geophys Res Lett*, *22*(14), 1877-1880.

Crutzen, P. J. (2006), Albedo enhancement by stratospheric sulfur injections: A contribution to resolve a policy dilemma?, *Climatic Change*, *77*(3-4), 211-219.

Dessler, A. E. (2009), Clouds and water vapor in the Northern Hemisphere summertime stratosphere, *J Geophys Res-Atmos*, *114*.

Dessler, A. E., T. F. Hanisco, and S. Fueglistaler (2007), Effects of convective ice lofting on H2O and HDO in the tropical tropopause layer, *J Geophys Res-Atmos*, *112*(D18).

Hanisco, T. F., et al. (2002), Quantifying the rate of heterogeneous processing in the Arctic polar vortex with in situ observations of OH, *J Geophys Res-Atmos*, *107*(D20).

Hanisco, T. F., et al. (2007), Observations of deep convective influence on stratospheric water vapor and its isotopic composition, *Geophys Res Lett*, *34*(4).

Hanson, D. R. (1998), Reaction of ClONO2 with H2O and HCl in sulfuric acid and HNO3/H2SO4/H2O mixtures, *J Phys Chem A*, *102*(25), 4794-4807.

Keim, E. R., et al. (1996), Observations of large reductions in the NO/NOy ratio near the mid-latitude tropopause and the role of heterogeneous chemistry, *Geophys Res Lett*, *23*(22), 3223-3226.

Keith, D. W. (2000a), Stratosphere-troposphere exchange: Inferences from the isotopic composition of water vapor, *J Geophys Res-Atmos*, *105*(D12), 15167-15173.

Keith, D. W. (2000b), GEOENGINEERING THE CLIMATE: History and Prospect, *Annual Review of Energy & the Environment*, *25*(1), 245.

Kirk-Davidoff, D. B., E. J. Hintsa, J. G. Anderson, and D. W. Keith (1999), The effect of climate change on ozone depletion through changes in stratospheric water vapour, *Nature*, *402*(6760), 399-401.

Kolb, C. E., D. R. Worsnop, M. S. Zahniser, and P. Davidovits (1996), A spectroscopic tour through the liquid aerosol interface: Implications for atmospheric chemistry, *J Geophys Res-Atmos*, *101*(D17), 23039-23043.

Moyer, E. J., F. W. Irion, Y. L. Yung, and M. R. Gunson (1996), ATMOS stratospheric deuterated water and implications for troposphere-stratosphere transport, *Geophys Res Lett*, *23*(17), 2385-2388.

Murray, B. J. (2008), Inhibition of ice crystallisation in highly viscous aqueous organic acid droplets, *Atmos Chem Phys*, *8*(17), 5423-5433.

Pfister, L., H. B. Selkirk, D. O. Starr, K. Rosenlof, and P. A. Newman (2010), A meteorological overview of the TC4 mission, *J Geophys Res-Atmos*, *115*.

Ravishankara, A. R., G. Hancock, M. Kawasaki, and Y. Matsumi (1998), Photochemistry of ozone: Surprises and recent lessons, *Science*, *280*(5360), 60-61.

Robock, A. (2000), Volcanic eruptions and climate, *Rev Geophys*, *38*(2), 191-219.

Salawitch, R. J., D. K. Weisenstein, L. J. Kovalenko, C. E. Sioris, P. O. Wennberg, K. Chance, M. K. W. Ko, and C. A. McLinden (2005), Sensitivity of ozone to bromine in the lower stratosphere, *Geophys Res Lett*, *32*(5).

Sander, S. P., et al. (2006), Chemical Kinetics and Photochemical Data for Use in Atmospheric Studies, Evaluation Number 15*Rep.*, Jet Propulsion Laboratory, Pasadena, CA.

Sayres, D. S., L. Pfister, T. F. Hanisco, E. J. Moyer, J. B. Smith, J. M. St Clair, A. S. O'Brien, M. F. Witinski, M. Legg, and J. G. Anderson (2010), Influence of convection on the water isotopic composition of the tropical tropopause layer and tropical stratosphere, *J Geophys Res-Atmos*, *115*.

Sioris, C. E., T. P. Kurosu, R. V. Martin, and K. Chance (2004), Stratospheric and tropospheric NO2 observed by SCIAMACHY: first results, *Trace Constituents in the Troposphere and Lower Stratosphere*, *34*(4), 780-785.

Smith, J. B., E. J. Hintsa, N. T. Allen, R. M. Stimpfle, and J. G. Anderson (2001), Mechanisms for midlatitude ozone loss: Heterogeneous chemistry in the lowermost stratosphere?, *J Geophys Res-Atmos*, *106*(D1), 1297-1309.

Solomon, S., S. Borrmann, R. R. Garcia, R. Portmann, L. Thomason, L. R. Poole, D. Winker, and M. P. McCormick (1997), Heterogeneous chlorine chemistry in the tropopause region, *J Geophys Res-Atmos*, *102*(D17), 21411-21429.

Tilmes, S., R. Muller, and R. Salawitch (2008), The sensitivity of polar ozone depletion to proposed geoengineering schemes, *Science*, *320*(5880), 1201-1204.

Wang, P. K. (2003), Moisture plumes above thunderstorm anvils and their contributions to cross-tropopause transport of water vapor in midlatitudes, *J Geophys Res-Atmos*, *108*(D6).

Yung, Y. L., J. P. Pinto, R. T. Watson, and S. P. Sander (1980), Atmospheric Bromine and Ozone Perturbations in the Lower Stratosphere, *J Atmos Sci*, *37*(2), 339-353.

Zobrist, B., C. Marcolli, D. A. Pedernera, and T. Koop (2008), Do atmospheric aerosols form glasses?, *Atmos Chem Phys*, *8*(17), 5221-5244.