Western Graduate&PostdoctoralStudies

Western University
Scholarship@Western

Electronic Thesis and Dissertation Repository

6-15-2011 12:00 AM

Numerical simulations of composition changes in the high latitude thermosphere during disturbed conditions

Albert T. Russell, *The University of Western Ontario*

Supervisor: Dr. J.-P. St.Maurice, *The University of Western Ontario*
Joint Supervisor: Dr. R. J. Sica, *The University of Western Ontario*
A thesis submitted in partial fulfillment of the requirements for the Doctor of Philosophy degree in Physics
© Albert T. Russell 2011

Follow this and additional works at: https://ir.lib.uwo.ca/etd

Part of the Atmospheric Sciences Commons

Recommended Citation

Russell, Albert T., "Numerical simulations of composition changes in the high latitude thermosphere during disturbed conditions" (2011). *Electronic Thesis and Dissertation Repository*. 164.
https://ir.lib.uwo.ca/etd/164

This Dissertation/Thesis is brought to you for free and open access by Scholarship@Western. It has been accepted for inclusion in Electronic Thesis and Dissertation Repository by an authorized administrator of Scholarship@Western. For more information, please contact wlswadmin@uwo.ca.

Numerical simulations of composition changes in the high latitude thermosphere during disturbed conditions

(Spine title: Thermospheric composition changes during disturbed conditions)

(Thesis format: Monograph)

by

Albert T. Russell

Graduate Program in Physics

A thesis submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy

The School of Graduate and Postdoctoral Studies
The University of Western Ontario
London, Ontario, Canada

© Albert T. Russell 2011

THE UNIVERSITY OF WESTERN ONTARIO
SCHOOL OF GRADUATE AND POSTDOCTORAL STUDIES

CERTIFICATE OF EXAMINATION

Supervisors

Dr. Robert J. Sica

Dr. Jean-Pierre St.-Maurice

Examining Board

Dr. Wayne K. Hocking

Dr. Donald R. Moorcroft

Dr. John W. MacDougall

Dr. William E. Ward

The thesis by

Albert T. Russell

entitled

Numerical simulations of composition changes
in the high latitude thermosphere during disturbed conditions

is accepted in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

Date

Chair of Examining Board

ABSTRACT

The closure of magnetospheric currents in the high latitude ionosphere makes the high latitude thermosphere a very dynamic environment. The composition and dynamics of this region become even more complex during geomagnetic disturbances as the electric fields from the magnetosphere now have the ability to substantially alter the winds and composition of this region. This complexity is especially apparent in mass spectrometer observations of composition changes, with heavier gases ($N_2, O_2, \text{and} Ar$) showing substantial enhancements while lighter gases (He and O) normally exhibit moderate to severe depletions. Quantifying the changes in atomic and molecular oxygen can be particularly difficult as most mass spectrometers are not able to discern between ambient molecular oxygen and the molecular oxygen created by atomic oxygen-satellite surface interactions - one usually measures the total oxygen content $O + 2\,O_2$ and assumes that any molecular oxygen above 250 km is due to the recombination of atomic oxygen on a satellite surface. High resolution simulations presented in this thesis suggest that large amounts of molecular oxygen can be transported upwards by vertical winds during geomagnetic storms and that the neglect of this transport effect will lead to substantial overestimations of atomic oxygen number densities by mass spectrometers at higher altitudes. These overestimations can be quite significant; a simulated total oxygen depletion of one-half at 280 km could mean that the atomic oxygen number densities at 280 km are actually one-seventh of their normal values while the simulated molecular oxygen concentrations are 25 times larger than normal.

keywords: thermosphere, composition, mass spectrometer

TABLE OF CONTENTS

CERTIFICATE OF EXAMINATION ii

ABSTRACT iii

TABLE OF CONTENTS iv

LIST OF TABLES vii

LIST OF FIGURES viii

Chapter 1 Introduction 1

1.1 The Neutral Atmosphere 3

1.1.1 Temperatures 3

1.1.2 Density Structure 5

1.2 Absorption of Solar Ultraviolet Radiation 9

1.3 The Ionosphere 14

1.3.1 The D-region 16

1.3.2 The E-region 20

1.3.3 The F-region 21

1.4 High Latitude Electrodynamics 22

1.5 Observations of the disturbed high latitude thermosphere 26

1.5.1 Neutral Density Changes 26

1.5.2 Composition changes 26

1.6 Previous Modelling Efforts 49

Chapter 2 Model Description I : Physics **58**

2.1 Governing Equations 58

2.2 UV/EUV Heating rates 60

2.2.1 Schumann-Runge Heating Rates 60

2.2.2 Ozone Heating Rates 62

2.2.3 Exothermic Chemical Heating Rates 63

2.2.4 Ion/Neutral and Electron/Neutral Heating Rates 66

2.3 Cooling Rates 71

2.4 Composition 76

2.5 Other Thermospheric Parameters 80

2.5.1 Transport Coefficients 80

2.5.2 Body Forces 81

2.6 Ionospheric Parameterizations 82

2.6.1 Ion Density Profiles 82

2.6.2 Geomagnetic field 82

2.6.3 Currents 84

Chapter 3 Model Description II : Numerics **89**

3.1 Introduction 89

3.2 The mathematical behaviour of the Navier-Stokes equations 89

3.3 Grid Discretization 94

3.4 Non-dimensionalization of the equations 99

3.5 Numerical Technique 100

3.6 Smoothing/Filtering/Artificial Dissipation 102

3.7 Boundary Conditions 105

3.8 Determination of the Initial Conditions 106

Chapter 4 Results and Discussion **112**

4.1 Zonal winds 113

4.1.1 In the auroral zone 113

4.1.2 In the sub-auroral zone 119

4.2 Meridional Winds 128

4.3 Vertical winds 145

4.3.1 In the auroral zone 145

4.3.2 In the sub-auroral zone 151

4.4 Density and Composition Changes 152

Chapter 5 Summary, Conclusions and Future Work **168**

5.1 Summary and Conclusions 168

5.2 Future Work 169

REFERENCES **170**

VITA **186**

LIST OF TABLES

2.1 Oxygen/hydrogen chemistry reaction rates 65

2.2 Nitrogen chemistry reaction rates 65

2.3 O(^{1}D) quenching rates 66

2.4 Recombination rates 66

2.5 Ion chemistry reaction rates 67

2.6 Neutral gas polarizabilities 70

2.7 Nonresonant ion-neutral collision frequency coefficients 70

2.8 Numerical coefficients used in the viscosity and thermal conductivity approximations 81

3.1 Common grid stretching functions 95

LIST OF FIGURES

1.1 Temperature structure of the neutral atmosphere 4
1.2 Constituents of the neutral atmosphere 8
1.3 Density structure of the neutral atmosphere 9
1.4 Chapman production functions for different solar zenith angles 13
1.5 Density structure of the ionosphere 15
1.6 Constituents of the ionosphere 16
1.7 Positive ion chemistry of the D-region 19
1.8 Schematic diagram of currents and electric fields in the high latitude ionosphere 24
1.9 OGO-6 electric field measurements in the high latitude thermosphere during disturbed conditions 25
1.10 Atmospheric drag on seven satellites during the November 1960 storm period 27
1.11 Satellite drag as a function of angular distance from perigee 28
1.12 Neutral density measurements from the LOGACS accelerometer during the May 24-25, 1967 time period 29
1.13 SETA neutral mass densities at 200 km during the July 20-26, 1983 time period. 30
1.14 CHAMP neutral mass densities at 410 km during the April 16-22, 2002 time period 31

1.15 Schematic diagram of the ESRO-4 mass spectrometer 34
1.16 Neutral number densities in the ion source of the OGO-6 mass spectrometer 35
1.17 Schematic diagram of the AE-C Open Source Spectrometer 37
1.18 Composition changes at 280 km during the geomagnetic storm of October 29, 1973 39
1.19 Composition changes at 160 km during the November 8-15, 1974 storm period 41
1.20 High latitude O_2 number densities at 200 km during the geomagnetic storm of January 11, 1974 42
1.21 DE-1 images of the FUV dayglow during geomagnetically quiet and disturbed conditions in late 1981 45
1.22 O/N_2 column density ratios derived from a DE-1 image of the FUV dayglow on October 22, 1981 46
1.23 O/N_2 column density ratios derived from GUVI images of the airglow during the April 27-29, 2002 time period 48
1.24 Inferred atomic oxygen column densities over Poker Flat, Alaska on February 13, 1994 50
1.25 Daily average number density ratios during the February 9-15, 1974 time period 55

2.1 CO_2 15 μm cooling rate parameterizations 73
2.2 Composition fractions employed in the *Chang and St.-Maurice* (1991) simulations 77
2.3 Nitric oxide density profiles 78
2.4 Ion density profiles 83

3.1 Variation of flow properties across a normal shock wave 91

3.2 Grid cell sizes for selected grids 97
3.3 Grid cell size ratios for selected grids 98
3.4 Initial composition fractions of molecular nitrogen in selected versions of the model 107
3.5 Initial composition fractions of molecular oxygen in selected versions of the model 108
3.6 Initial composition fractions of atomic oxygen in selected versions of the model 109
3.7 Initial density profiles in selected versions of the model 110
3.8 Initial temperature profiles in selected versions of the model 111

4.1 Zonal winds in the auroral zone for the 'reference run' simulation . . 114
4.2 Zonal momentum balances in the auroral zone at the 15 minute mark of the 'reference run' simulation 115
4.3 Zonal momentum balances in the auroral zone at the 30 minute mark of the 'reference run' simulation 116
4.4 Zonal momentum balances in the auroral zone at the 60 minute mark of the 'reference run' simulation 118
4.5 Zonal wind times series at selected altitudes in the centre of the auroral zone as a function of electric field 'ramp-up' time 120
4.6 Zonal acceleration times series at selected altitudes in the centre of the auroral zone as a function of electric field 'ramp-up' time 121
4.7 Zonal acceleration and neutral density times series in the centre of the auroral region for a simulation with an electric field 'ramp-up' time of 10 minutes 122
4.8 Zonal winds outside of the auroral zone for the 'reference run' simulation 124

4.9 Zonal momentum balances outside of the auroral zone at the 60 minute mark of the 'reference run' simulation 125
4.10 Zonal momentum balances outside of the auroral zone at the 120 minute mark of the 'reference run' simulation 126
4.11 Zonal momentum balances outside of the auroral zone at the 180 minute mark of the 'reference run' simulation 127
4.12 'Reference run' meridional winds for the first hour of the simulation . 128
4.13 'Reference run' meridional winds for the final two hours of the simulation 129
4.14 Meridional momentum balances at the 15 minute mark of the 'reference run' simulation 131
4.15 Meridional pressure gradients at the 15 minute mark of the 'reference run' simulation 132
4.16 Meridional momentum balances at the 30 minute mark of the 'reference run' simulation 134
4.17 Meridional momentum balances at the 45 minute mark of the 'reference run' simulation 135
4.18 Meridional wind speeds at the 60 minute mark of two simulations with different meridional momentum source terms 136
4.19 Meridional wind speeds at the 60 minute mark of a simulation with a 'modified' meridional momentum source term 137
4.20 Pressure perturbations at the 60 minute mark of the 'reference run' simulation 138
4.21 Meridional winds at the 90 and 120 minute marks of the simulation with the modified meridional momentum source term 140
4.22 Meridional winds in the auroral zone at selected times of the 'reference' run simulation 142

4.23 Meridional wind times series at selected altitudes in the auroral zone as a function of electric field 'ramp-up' time 144
4.24 'Reference run' vertical winds for the first hour of the simulation . . . 146
4.25 Vertical wind time series at 70°N for selected altitudes 147
4.26 Vertical momentum balances in the auroral zone at the 15 minute mark of the 'reference run' simulation 148
4.27 Vertical momentum balances in the auroral zone at the 30 minute mark of the 'reference run' simulation 149
4.28 Vertical momentum balances in the auroral zone at the 60 minute mark of the 'reference run' simulation 150
4.29 Vertical wind times series at selected altitudes in the centre of the auroral zone as a function of electric field 'ramp-up' time 151
4.30 Vertical winds outside of the auroral zone for the 'reference run' simulation 152
4.31 Density perturbations for the first hour of the 'reference run' simulation 153
4.32 Density perturbations for the final two hours of the 'reference run' simulation 154
4.33 Atomic oxygen composition fractions for the first hour of the 'reference run' simulation 156
4.34 Atomic oxygen composition fractions for the final two hours of the 'reference run' simulation 157
4.35 Molecular oxygen composition fractions for the first hour of the 'reference run' simulation 158
4.36 Molecular oxygen composition fractions for the final two hours of the 'reference run' simulation 159
4.37 Molecular nitrogen composition fractions for the first hour of the 'reference run' simulation 160

4.38 Molecular nitrogen composition fractions for the final two hours of the 'reference run' simulation 161
4.39 N_2/O and $N_2/(O + 2\,O_2)$ ratios during the first ninety minutes of the 'reference run' simulation 162
4.40 N_2/O and $N_2/(O + 2\,O_2)$ ratios during the third hour of the 'reference run' simulation 163
4.41 N_2/O and $N_2/(O+2\,O_2)$ ratios at 280 km in the 'reference run' simulation 164
4.42 Composition ratios at 280 km for the 'reference run' simulation . . . 165

Chapter 1

Introduction

On March 6, 1989, a very large and complex sunspot group, Active Region 5395, rotated around the eastern limb of the Sun and into full view of the Earth. This region erupted almost immediately, producing one of the most intense solar flares ever observed, an X15/3B class flare that would last over ten hours [*Joshi*, 1993]. Several more flares would be observed over the next three days, with at least one X-class flare being observed daily. The production of high energy protons had also increased dramatically over this time period, with the solar proton flux exceeding its quiet-time value by three orders of magnitude. Then, at 18:37 UT on March 10, another flare erupted that would last for 141 minutes. This flare would eventually be classified as an X4.5/3B flare, but the accompanying coronal mass ejection (CME) was headed towards the Earth at a speed of 770 km/sec [*Feynman and Hundhausen*, 1994] and would change our perceptions about the near-Earth space environment.

The leading edge of this CME reached the Earth on the evening of March 12 and caused the high-energy proton flux to increase to 100 times its normal value. By the morning of March 13, geosynchronous satellites like GOES-6 and GOES-7 were no longer flying in the protective cocoon of the Earth's magnetosphere as the magnetopause had been compressed to 4.7 R_E, half of its typical 10 R_E distance [*Allen et al.*, 1989]. Data corruption, "single-event" upsets (SEUs) and other minor satellite operation anomalies were widespread during this time period, affecting satellites at

all altitudes. Japan's geostationary CS-3B communications satellite was particularly hard-hit by the increased particle flux, losing half of its dual redundant command circuitry [*Allen et al.*, 1989]. Seven other geostationary communications satellites had problems maintaining operational attitude, requiring 177 thruster adjustments that shortened their operational lifetimes by a year [*Allen et al.*, 1989].

The increased particle flux also had a pronounced effect on the thermosphere and ionosphere. Bright red aurora were observed over most of North America and Europe on the nights of March 12-13 and 13-14. The Aurora Australis was also visible at unusually low latitudes, with sky watchers in New Zealand, western Australia, and South Africa being treated to the same vivid red displays seen over most of North America and Europe. Ionospheric densities in the D- and E-regions were substantially enhanced, causing a HF (High-frequency) communications blackout on the sunlit side of the Earth. LORAN (LOng RAnge Navigation), GPS (Global Positioning System), and other navigational systems were also being disrupted as scintillation effects and TEC (Total Electron Content) variations were causing substantial signal degradation. VHF (Very-high-frequency) radio communications, on the other hand, were extremely robust, propagating far beyond their normal line-of-sight ranges (a more complete discussion of the ionospheric response to the March 13-14 storm can be found in *Yeh et al.*, 1992).

The auroral electrojets had also intensified and expanded equatorwards, causing ground induced currents (GICs) to flow in high voltage transformers, power transmission lines, telecommunication cables, and pipelines. High voltage transformers in Europe and North America were particularly susceptible to the effects of these GICs, with the voltage surges causing transformer cores to saturate. This resulted in voltage fluctuations that tripped protective relays in many areas of the United States and Sweden [*Allen et al.*, 1989]. It also caused some transformers to overheat, shortening their operational lifetimes (a ten million dollar transformer at a Salem, New Jersey power plant was so badly damaged that it eventually had to be replaced). The most spectacular effect of these GICs occurred in Quebec - the entire Hydro-Quebec power system collapsed in just 90 seconds - leaving 6 million people without electricity in

sub-zero temperatures.

The enhanced electrojet currents also caused Joule heating rates in the upper atmosphere to increase, producing substantial density enhancements throughout most of the thermosphere. Low Earth orbit (LEO) satellites were now traveling through regions that were five to nine times denser than normal [*Allen et al.*, 1989], causing the decay rates of these objects to accelerate. These accelerated decay rates also tended to be quite erratic, with the U. S. Air Force Space Command temporarily losing track of 1300 satellites (it would eventually take the U. S. Air Force Space Command a week to identify all of the 'missing' satellites).

Since the time of that storm, the use of satellites and other technologies that are susceptible to changes in the near-Earth space environment has mushroomed, leading to a concerted effort by many governments to understand how solar activity affects the near-Earth space environment, e.g. the Canadian Geospace Monitoring (CGSM) program and the U. S. National Space Weather Program (NSWP). Understanding how the upper atmosphere distributes and dissipates energy from the magnetosphere is one of the more challenging problems that these programs must solve and is the primary motivation for this thesis. With this in mind, the remainder of this chapter will provide overviews on the upper atmosphere, relevant observational results, and previous thermospheric modelling efforts.

1.1 The Neutral Atmosphere

1.1.1 Temperatures

The neutral component of the Earth's atmosphere is most commonly described by its thermal characteristics. This results in four distinct regions, with each region being known as a *sphere* and the boundary between adjacent regions being referred to as a *pause* (Figure 1.1). The lowest region in this paradigm is the *troposphere*, which has a typical lapse rate of 6.5 K/km (this "typical" lapse rate often exhibits significant variability as the heating of the Earth's surface by solar radiation frequently results in substantial convective activity and strong vertical motions). These decreasing

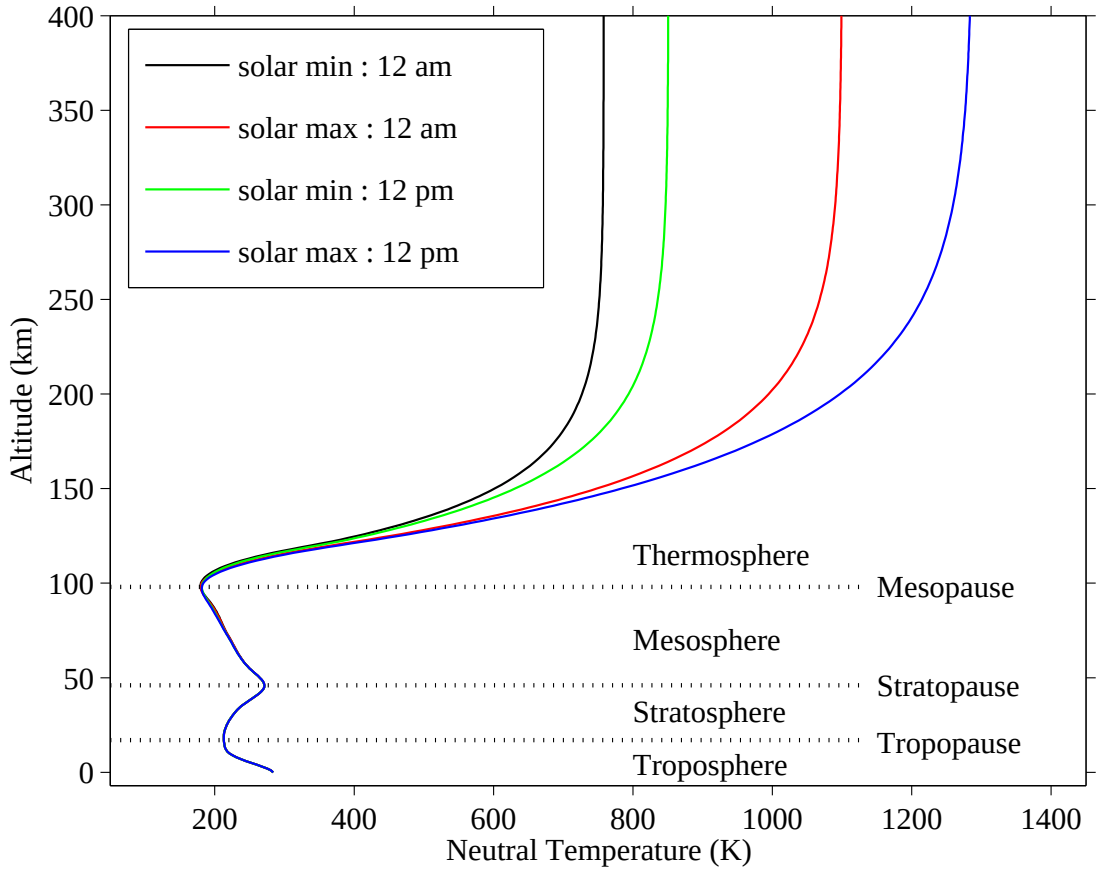

Figure 1.1: Temperature structure of the neutral atmosphere. Neutral temperatures are from the Mass Spectrometer and Incoherent Scatter model [*Hedin*, 1991] at 45°N, 0°E for March 21. The variability of neutral temperatures in the thermosphere is quite evident (F10.7 values for solar minimum and maximum were 80 and 210 respectively).

temperatures persist until one reaches the *tropopause*, which occurs around 18 km in the tropics. The height of the tropopause is significantly lower in the polar regions, with tropopause altitudes typically being in the range of 8-10 km.

The concentrations of some trace atmospheric constituents also change quite abruptly at the tropopause; water vapour number densities decrease sharply while the number density of ozone often increases by an order of magnitude over the next few kilometers. These rapidly increasing concentrations of ozone absorb significant amounts of solar UV radiation, reversing the temperature gradient in the *stratosphere*. This results in a strongly stratified region that inhibits vertical motions -

thunderstorms can only penetrate a few km into the lower stratosphere before being dissipated.

The temperatures in the stratosphere eventually reach a maximum at the *stratopause* and then begin to decrease again in the *mesosphere*. The decreasing temperatures in the mesosphere means that vertical motions in this region are not strongly damped (like the troposphere), allowing dynamical motions and radiation processes to play significant roles in the evolution of this region. The interplay between these two processes allows the second temperature minimum at the *mesopause* to be the coldest place on Earth, with temperatures routinely dropping to 130 K in the summer months [*Theon et al.*, 1967; *Lübken and von Zahn*, 1991; *Lübken*, 1999].

The temperatures in the lower portion of the *thermosphere* increase quite dramatically, with lapse rates of -15 K/km being quite common to altitudes of 150 km. The presence of such large lapse rates in this region of the atmosphere is due to two factors: the absorption of solar ultraviolet (UV) radiation at wavelengths below 175 nm and a paucity of effective cooling mechanisms at these altitudes. The presence of such large temperature gradients cannot be maintained indefinitely and the temperatures in the thermosphere eventually approach an asymptotic value (known as the exospheric temperature) that persists for several hundred kilometers. The dependence of thermospheric temperatures on the absorption of solar radiation also means that the temperatures in the thermosphere are highly variable in space and time. Solar cycle and local time variations are usually the most obvious (recall Figure 1.1), but many other effects can also be seen in thermospheric temperatures, e.g. geomagnetic activity, seasonal variations, etc.

1.1.2 Density Structure

The density and composition of the atmosphere is also quite variable, with variations in latitude, longitude, and time of day being easily discernible in the climatologies of most neutral species. While a complete understanding of a species distribution requires a detailed knowledge of its horizontal variations, considerable insight into a species distribution can be gained if one only considers vertical variations. If one does

this, the *continuity equation* for species s can be written as

$$\frac{\partial n_s}{\partial t} + \nabla \cdot (n_s \mathbf{u}_s) = P_s - L_s \tag{1.1}$$

where n_s is the number density of species s, $\mathbf{u}_s$ is the drift velocity of species s, P_s is the production rate of species s, and L_s is the loss rate of species s. The product $n_s\mathbf{u}_s$ is also known as the *flux* of a species and requires a knowledge of the species drift velocity $\mathbf{u}_s$. *Colegrove et al.* [1965, 1966] have shown that the vertical component of the species flux can, in the absence of strong vertical motions, be written as

$$\begin{aligned}(n_s\mathbf{u}_s)_z &= -D_s \left[\frac{\partial n_s}{\partial z} + \frac{n_s(1+\alpha_s)}{T_s}\frac{\partial T_s}{\partial z} + \frac{n_s m_s g}{kT_s} - n_s \sum_{t \neq s} \frac{(n_t \mathbf{u}_t)_z}{nD_{st}} \right] \\ &\quad -K_z \left[\frac{\partial n_s}{\partial z} + \frac{n_s}{T_s}\frac{\partial T_s}{\partial z} + \frac{n_s \bar{m} g}{kT_s} \right]\end{aligned} \tag{1.2}$$

where α_s is the molecular thermal diffusion coefficient of species s, T_s is the temperature of species s, m_s is the mass of an individual particle in species s, K_z is the eddy diffusion coefficient, g is the gravitational field strength at height z, and k is Boltzmann's constant. The mean mass $\bar{m}$ and molecular diffusion coefficient D_s are determined by the expressions

$$\bar{m} = \sum_s n_s m_s / \sum_s n_s \tag{1.3}$$

$$D_s = \left(\sum_{t \neq s} \frac{n_t}{nD_{st}} \right)^{-1} \tag{1.4}$$

with D_{st} being the binary diffusion coefficient between species s and species t.

While equations (1.1) and (1.2) allow one to determine the vertical distribution of most species, considerable simplifications can still be made if one considers the chemical and dynamical timescales of the species in question. If the chemical timescale of the species is much longer than the dynamical timescale of the species, one can ignore the effects of chemistry, i.e. P_s and L_s are negligible. Most long chemical timescale neutral species are also close to diffusive equilibrium, allowing one to set the vertical flux of these species to zero (atomic hydrogen is the notable exception). Defining $\Lambda = K_z/D_s$ and integrating equation (1.2) from a reference altitude z_o to z

gives the steady-state distribution

$$n_s(z) = n_s(z_o)\left[\frac{T_s(z_o)}{T_s(z)}\right]\exp\left[-\int_{z_o}^{z}\left(\frac{1}{H_s}+\frac{\Lambda}{H_m}\right)(1+\Lambda)^{-1}dz'\right] \times \exp\left[-\alpha_s\int_{T_o}^{T}(1+\Lambda)^{-1}\,d\left(\ln T'\right)\right] \tag{1.5}$$

where the scale heights

$$H_s = \frac{kT_s}{m_s g}\,, \qquad H_m = \frac{kT_s}{\bar{m}g} \tag{1.6}$$

have been introduced. Since mixing processes are much faster than diffusive processes below 80 km, i.e. $\Lambda \gg 1$, one can simplify equation (1.5) to

$$n_s(z) = n_s(z_o)\left[\frac{T_s(z_o)}{T_s(z)}\right]\exp\left[-\int_{z_o}^{z}\frac{dz'}{H_m}\right] \tag{1.7}$$

in this region. This shows that the number density of any species with a long chemical timescale in these regions of the atmosphere will decrease exponentially with the *same* scale height H_m, i.e. any long chemical timescale species will have the same mixing ratio at any altitude where mixing/turbulence dominates. This can be clearly seen in Figure 1.2, where the profiles of N_2, O_2, and Ar parallel each other below 100 km. Since these species are also the dominant constituents in these regions, the mean molecular weight of air does not change appreciably in the mixing dominated region of the atmosphere. Because of this, the mixing dominated region of the atmosphere is known as the *homosphere*.

The effectiveness of mixing processes decreases rapidly above 100 km, allowing molecular diffusion to become the dominant process above 120 km, i.e. $\Lambda \ll 1$. Under these conditions, equation (1.5) simplifies to

$$n_s(z) = n_s(z_o)\left[\frac{T_s(z_o)}{T_s(z)}\right]^{(1+\alpha_s)}\exp\left[-\int_{z_o}^{z}\frac{dz'}{H_s}\right]. \tag{1.8}$$

This shows that the number density of any long chemical timescale species in the diffusion dominated regime of the atmosphere will decrease exponentially with its own scale height H_s. Since the scale heights of the individual species are inversely

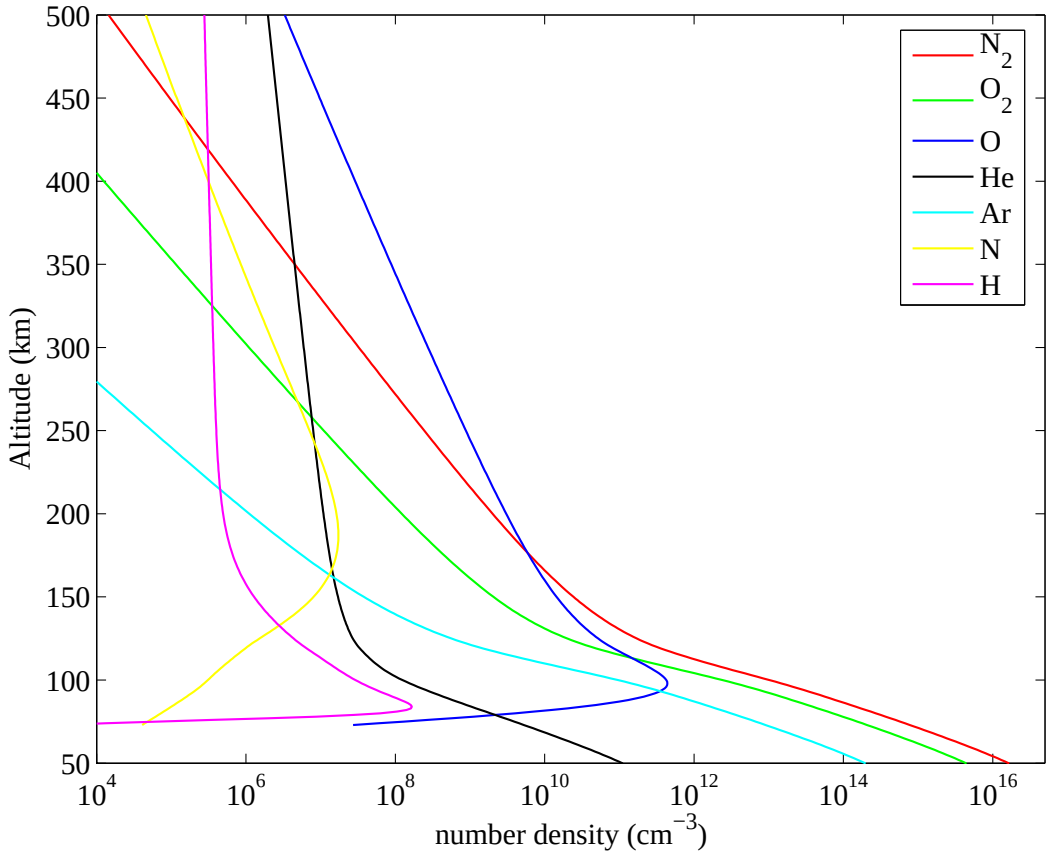

Figure 1.2: Constituents of the neutral atmosphere. Profiles of the individual species are from the Mass Spectrometer and Incoherent Scatter model [*Hedin*, 1991] at 45°N, 0°E for March 21 (the F10.7 value was 80).

proportional to the species mass, lighter species will have larger scale heights, and as a consequence of this, are more abundant than the heavier species at higher altitudes. This is quite evident in Figure 1.2, where the number densities of N_2 and O_2 decrease much more rapidly than the number densities of lighter species like H, He, and O. The effects of vertical transport are also discernible in Figure 1.2, with the number densities of H and He decreasing at a rate that is much slower than that predicted by scale height considerations. This tendency for species in the molecular diffusion dominated regime of the atmosphere to 'separate out' has led to the designation of this region of the atmosphere as the *heterosphere*. The dependence of neutral species scale heights on temperature also means that solar cycle and local time variations

will be present in heterospheric number densities, with order of magnitude variations being quite common (Figure 1.3).

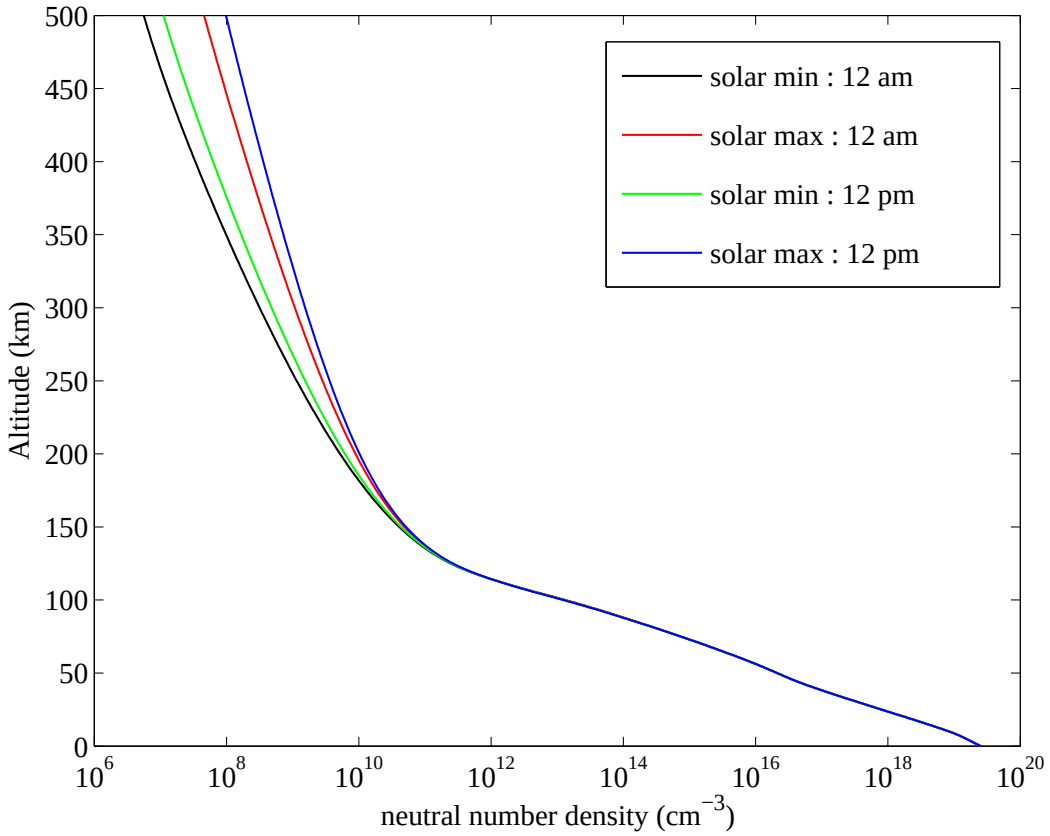

Figure 1.3: Density structure of the neutral atmosphere. Neutral densities are from the Mass Spectrometer and Incoherent Scatter model [*Hedin*, 1991] at 45°N, 0°E for March 21. The variability of the neutral number density in the upper thermosphere is obvious (F10.7 values for solar minimum and maximum were 80 and 210 respectively).

1.2 Absorption of Solar Ultraviolet Radiation

As solar ultraviolet photons penetrate into the atmosphere, they can interact with the ambient species through three basic mechanisms: absorption, emission, and scattering. Scattering is normally negligible at UV wavelengths as the Rayleigh scattering cross section of most thermospheric species is usually several orders of magnitude smaller than the species' absorption cross section [*Huffman*, 1992]. The plethora of

UV emission lines and bands in the thermosphere provides a wealth of information on the structure, dynamics and energetics of this region, but the retrieval of this information is not trivial and will only be discussed as needed in subsequent sections of this thesis. Absorption, which is crucial to understanding the ionosphere and thermosphere, will be presented in the following paragraphs.

The absorption of UV radiation in the thermosphere generally proceeds through three processes: *photoionization*, *photodissociation*, and the combination of these two processes, *dissociative photoionization*. Photodissociation occurs when the energy of the photon is greater than the dissociation threshold of the species in question, i.e.

$$O_2 + h\nu(< 2242\,\text{Å}) \rightarrow O + O \tag{1.9}$$

$$N_2 + h\nu(< 1026\,\text{Å}) \rightarrow N + N \tag{1.10}$$

[*Rees*, 1989]. The wavelengths given in equations (1.9) and (1.10) correspond to the production of two atoms in their electronic ground states; photons with energies greater than the dissociation threshold of that species can also result in electronically excited states, e.g.

$$O_2 + h\nu(< 1749\,\text{Å}) \rightarrow O + O(^1D). \tag{1.11}$$

If the energy of the incident photon is greater than the ionization threshold of the species that the photon is interacting with, photoionization can occur, i.e.

$$N_2 + h\nu(< 796\,\text{Å}) \rightarrow N_2^+ + e^- \tag{1.12}$$

$$O_2 + h\nu(< 1026\,\text{Å}) \rightarrow O_2^+ + e^- \tag{1.13}$$

$$O + h\nu(< 911\,\text{Å}) \rightarrow O^+ + e^- \tag{1.14}$$

[*Rees*, 1989]. The wavelengths given in equations (1.12)-(1.14) correspond to the production of the ion in its electronic ground state; photons with energies greater than the ionization threshold can generate an ion in an electronically excited state and/or an energetic photoelectron. Molecular nitrogen and oxygen can also be photoionized dissociatively if the photons are sufficiently energetic, i.e.

$$N_2 + h\nu(< 510\,\text{Å}) \rightarrow N^+ + N + e^- \tag{1.15}$$

$$O_2 + h\nu(< 662\,\text{Å}) \rightarrow O^+ + O + e^- \tag{1.16}$$

[*Rees*, 1989]. These processes are responsible for atomic oxygen becoming the dominant neutral species above 200 km and for the significant populations of electronically excited species in the thermosphere and ionosphere.

A more quantitative description of solar UV radiation absorption can be obtained through the application of the Beer-Lambert absorption law

$$dI(s,\lambda) = -\sum_t n_t(s)\,\sigma_t^a(s,\lambda)\,I(s,\lambda)\,ds_\lambda \tag{1.17}$$

where

$$\begin{aligned} I(s,\lambda) &= \text{intensity of the photon flux along pathlength } s \text{ at wavelength } \lambda \\ n_t(s) &= \text{number density of species } t \\ \sigma_t^a(s,\lambda) &= \text{absorption cross section of of species } t \text{ at wavelength } \lambda \\ ds_\lambda &= \text{incremental path length traveled by the photons.} \end{aligned}$$

This equation shows that the change in photon flux is proportional to the intensity of the photon flux, the absorption cross section, the number density of the absorbing species and the incremental path length traveled by the photons. Integrating equation (1.17) along the path length of the photons gives

$$I(s,\lambda) = I_\infty(\lambda)\,\exp\left[-\sum_t \int_\infty^s n_t(s')\,\sigma_t^a(s',\lambda)\,ds'_\lambda\right] \tag{1.18}$$

where $I_\infty(\lambda)$ is the unattenuated photon flux at the top of the atmosphere at wavelength λ. The argument of the exponential in equation (1.18) is also known as the *optical depth*

$$\tau(s,\lambda) = \sum_t \int_\infty^s n_t(s')\,\sigma_t^a(s',\lambda)\,ds'_\lambda \tag{1.19}$$

and is a measure of the atmosphere's ability to absorb radiation, i.e. a large optical depth will attenuate the photon flux quickly while a small optical depth will allow most of the photons to pass through.

While equation (1.19) is correct, its evaluation can become quite tedious as a detailed knowledge of the species distributions and absorption cross sections is necessary.

Considerable insight into the absorption of solar UV radiation can be gained if a few simplifying assumptions are made. If the thermosphere is modeled as a horizontally stratified, single species constant scale height gas, i.e.

$$n(z) = n(z_o) \exp\left[-\frac{(z - z_o)}{H}\right]$$

and the photon flux is treated as a monochromatic stream of photons propagating through the atmosphere at a solar zenith angle χ, i.e. s and z are related through the expression $ds = -\, dz\ \sec\chi$, then equation (1.18) becomes

$$\begin{aligned} I(z,\chi) &= I_\infty \exp\left[-\int_z^\infty n(z')\,\sigma\,\sec\chi\,dz'\right] \\ &= I_\infty \exp\left[\,-\sigma\,\sec\chi\,H\,n(z)\,\right]. \end{aligned} \tag{1.20}$$

Since the rates of photodissociation and photoionization are proportional to the photon energy deposition rate, all of these processes can be discussed in terms of the photon energy deposition rate, which, in terms of the preceding approximations, is

$$\begin{aligned} P_c(z,\chi) &= \sigma\, n(z)\, I(z,\chi) \\ &= \sigma\, n(z_o)\, I_\infty \exp\left\{\left[-\frac{(z - z_o)}{H}\right] - n(z_o)\sigma H \sec\chi \exp\left[-\frac{(z - z_o)}{H}\right]\right\}. \end{aligned} \tag{1.21}$$

Equation (1.21), which is also known as the *Chapman production function*, obtains its maximum value at

$$z_m = z_o + H \ln\left[\,n(z_o)\,H\,\sigma\,\sec\chi\right] \tag{1.22}$$

(this expression can be obtained by setting the derivative of equation (1.21) equal to zero). Substituting equation (1.22) into equation (1.21) yields

$$P_c(z_m,\chi) = \frac{I_\infty}{\mathrm{e}\,H\sec\chi} = \frac{I_\infty\,\cos\chi}{\mathrm{e}\,H} \tag{1.23}$$

which, in conjunction with equation (1.22), allows one to rewrite the Chapman production function as

$$P_c(z,\chi) = P_c(z_m,\chi=0) \exp\left\{1 - \left[\frac{(z - z_{mo})}{H}\right] - \sec\chi \exp\left[-\frac{(z - z_{mo})}{H}\right]\right\} \tag{1.24}$$

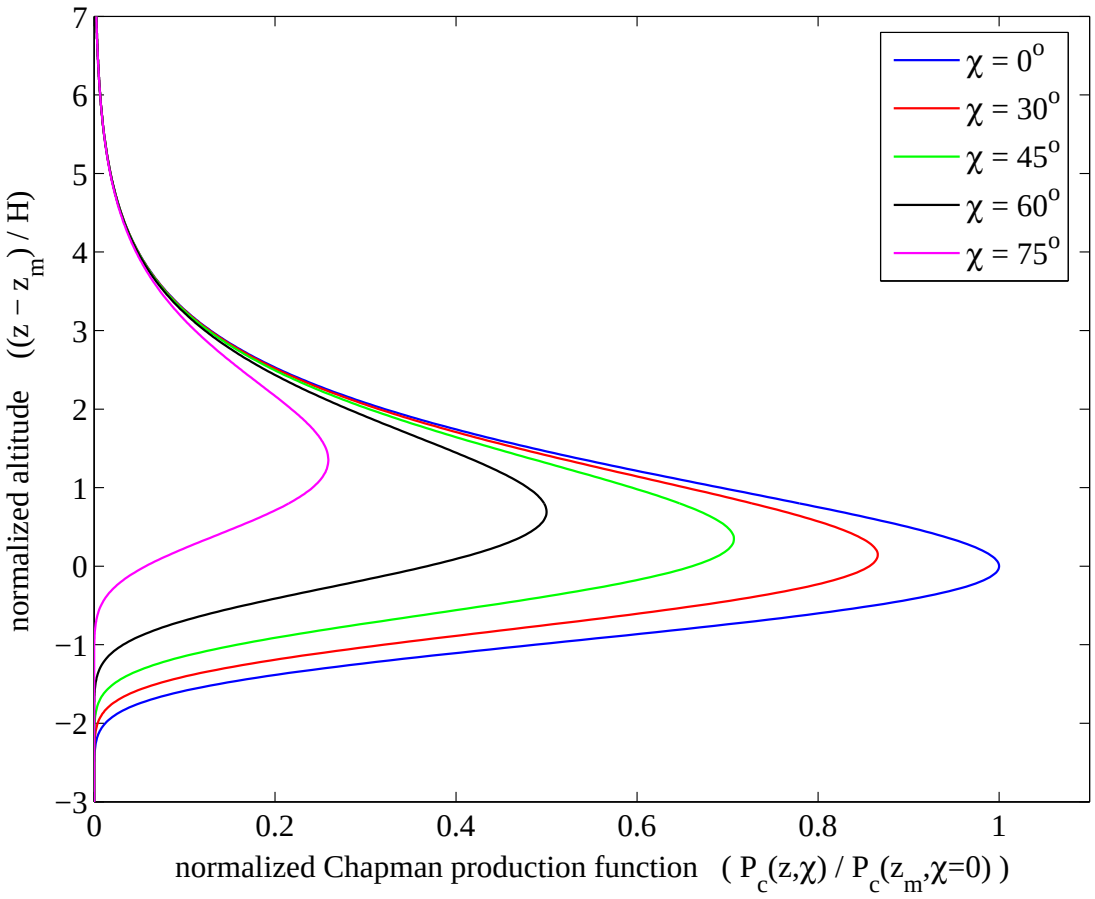

Figure 1.4: Chapman production function for different solar zenith angles. The maximum energy deposition rate decreases and is shifted to higher altitudes as the path length that the radiation must traverse increases.

where z_{mo} is the altitude where the zero zenith angle Chapman production function obtains its maximum value. Plots of equation (1.24) for various values of χ are presented in Figure 1.4. The shift of the maximum energy deposition rate to higher altitudes as the zenith angle increases is quite evident. The merging of the profiles above the peak production rate is also obvious; this is due to the low gas densities at higher altitudes limiting the production rate, i.e. for $z \gg z_{mo}$

$$P_c(z,\chi) \;\rightarrow\; P_c(z_m, \chi = 0)\; \exp\left[-\frac{(z - z_{mo})}{H}\right]. \tag{1.25}$$

The rapid decrease below the production peak is due to the photon flux having been

totally attenuated, i.e. for $z \ll z_{mo}$

$$P_c(z,\chi) \rightarrow P_c(z_m, \chi = 0) \exp\left[- \sec\chi \exp\left[- \frac{(z - z_{mo})}{H} \right] \right]. \quad (1.26)$$

1.3 The Ionosphere

The ionosphere is usually defined as "that region of the atmosphere where significant numbers of free thermal (< 1 eV) electrons and ions exist" [*Schunk and Nagy*, 2000]. The presence of these free electrons and ions in the atmosphere is primarily due to two processes: the ionization of neutral species by solar radiation, which was described in the previous section, and by energetic particle impact, which is usually most pronounced at higher latitudes (describing the passage of electrons through the atmosphere is beyond the scope of this thesis as they are not attenuated like photons - they collide with the ambient neutral species many times before becoming part of the thermal electron population). Once these ions and electrons are produced, they can be influenced by many processes: diffusion, recombination, plasma instabilities, electric fields, magnetic fields, neutral winds and wave disturbances can all play a significant role in the evolution of this region at times.

Despite this complexity, the vertical density structure of the ionosphere exhibits a 'layered' structure at all latitudes. Typical profiles of the mid latitude ionosphere are shown in Figure 1.5 and clearly illustrate this layering phenomenon. The most prominent layer in these profiles is the *F-region*, which, depending on the time of day, has a maximum number density of 10^5-10^6 particles/cm^3 in the 250-350 km region. The lower boundary of the F-region also depends on the time of day, with altitudes of 120-150 km being typical. The second electron number density maximum in the 105-110 km region is associated with the *E-region*, which has a lower boundary near 85-95 km. The 'ledge' in electron number densities below the E-region boundary is, not surprisingly, known as the *D-region*. The large altitude ranges associated with these boundaries is not simply due to solar flux or diurnal variations; the radio wave propagation and ionospheric composition communities employ different criteria in determining the altitudes of these boundaries.

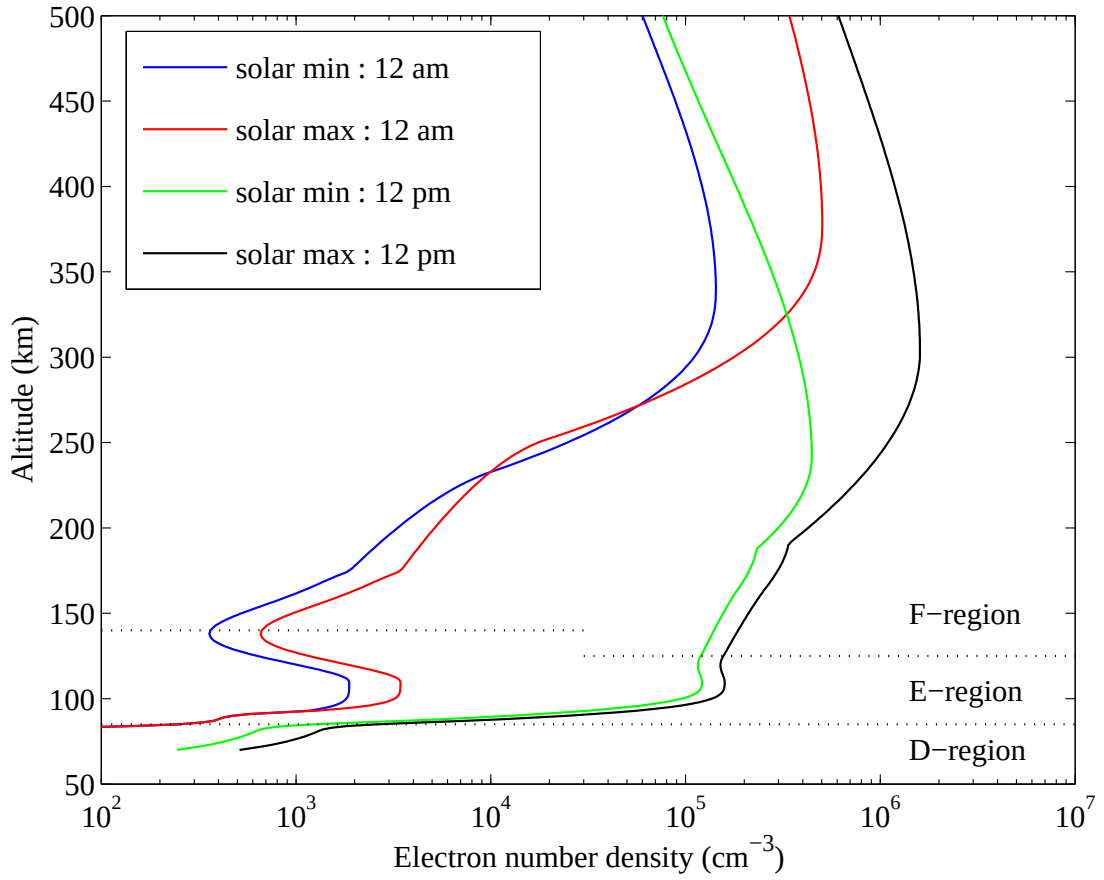

Figure 1.5: Density structure of the ionosphere. Electron densities are from the International Reference Ionosphere [*Blitza*, 2001] at 45°N, 0°E for March 21. The variability of the electron number densities is quite evident (annual average sunspot numbers for solar minimum and maximum were 20 and 120 respectively).

The quasi-neutrality of the terrestrial ionosphere dictates that the number of electrons and negative ions at any altitude must be equal to the number of positive ions at that altitude (negative ions are extremely rare above 85 km, so the number density of positive ions is usually equal to the number density of the electrons above altitudes of 85 km). A 'typical' profile of the positive ion distributions in the terrestrial ionosphere is presented in Figure 1.6. The large concentration of O^+ ions in the F-region is not unexpected as atomic oxygen is the primary neutral constituent at these altitudes. The presence of O_2^+ ions in the D- and E-regions can also be predicted from neutral constituent considerations, but the paucity of N_2^+ ions and the presence of

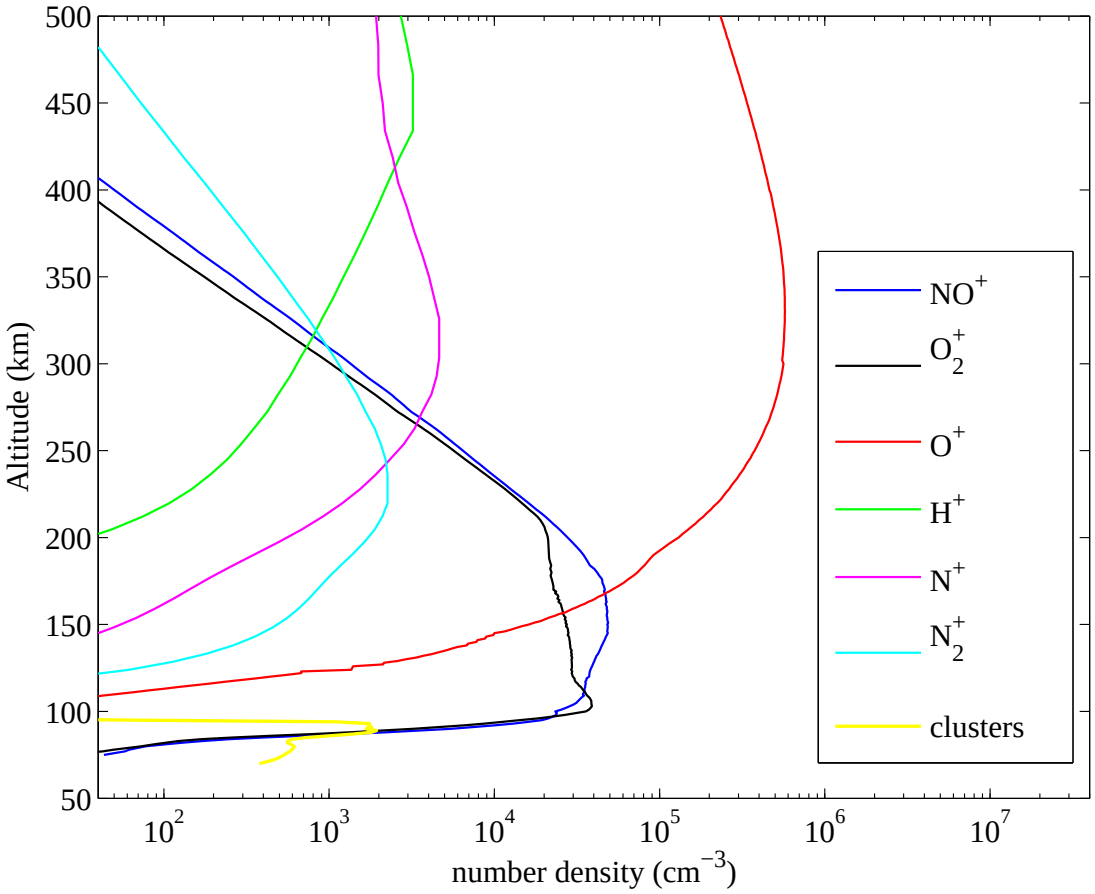

Figure 1.6: Constituents of the ionosphere. Profiles of the ionospheric species are from the International Reference Ionosphere [*Blitza*, 2001] and TRANSCAR [*Blelly et al.*, 1996], a first principles ionospheric model.

significant NO^+ concentrations in the D- and E-regions is somewhat surprising. The presence of $H^+(H_2O)_n$ and other water vapour cluster ions (WVCIs) is even more surprising; the remainder of this section will explain how the ion distributions in the various regions of the terrestrial ionosphere are maintained.

1.3.1 The D-region

The production of positive ions and free electrons in the D-region has been studied extensively since the 1960s and is now fairly well understood. The three most important sources of ionization in this region are the solar Lyman-α line at 1215.7 Å, which

ionizes nitric oxide (NO), solar UV photons in the 1027-1118 Å interval, which ionizes the metastable $O_2(^1\Delta_g)$ state, and solar X-rays in the 2-8 Å interval, which ionizes N_2 and O_2 (solar X-rays actually ionize all of the neutral constituents in the D-region, but N_2 and O_2 are the principal constituents). Cosmic rays and precipitating magnetospheric particles also contribute to the D-region ionization rates, with cosmic rays becoming the primary ionization source below 65 km. Most of these sources exhibit strong solar activity and diurnal variations, but the ionization rates do not drop to zero at night as the galactic cosmic ray flux is essentially constant and the Lyman-α flux always contains a scattered geocorona component, i.e. Lyman-α photons that have been resonantly scattered from atomic hydrogen in the exosphere.

These ionization processes produce N_2^+, O_2^+, NO^+, O^+, and N^+ ions at D-region heights, but the N_2^+, N^+, and O^+ ions are rapidly converted to O_2^+ and NO^+ through the reactions

$$N_2^+ + O_2 \rightarrow O_2^+ + N_2 \tag{1.27}$$

$$N_2^+ + O \rightarrow NO^+ + N \tag{1.28}$$

$$O^+ + O_2 \rightarrow O_2^+ + N \tag{1.29}$$

$$O^+ + N_2 \rightarrow NO^+ + N \tag{1.30}$$

$$N^+ + O_2 \rightarrow O_2^+ + N \tag{1.31}$$

$$N^+ + O_2 \rightarrow NO^+ + O \tag{1.32}$$

leaving NO^+ and O_2^+ as the dominant positive ions in the upper D-region [*Turunen et al.*, 1996]. The O_2^+ ions can also form proton hydrates (the $H^+(H_2O)_n$ group of WVCIs) through the reaction chain

$$O_2^+ + O_2 + M \rightarrow O_4^+ + M \tag{1.33}$$

$$O_4^+ + H_2O \rightarrow O_2^+(H_2O) + O_2 \tag{1.34}$$

$$O_2^+(H_2O) + H_2O \rightarrow H_3O^+(OH) + O_2 \tag{1.35}$$

$$H_3O^+(OH) + H_2O \rightarrow H^+(H_2O)_2 + OH \tag{1.36}$$

$$H^+(H_2O)_2 + H_2O + M \rightarrow H^+(H_2O)_3 + M \tag{1.37}$$

[*Fehsenfeld and Ferguson*, 1969; *Good et al.*, 1970], with the final hydration orders being determined by a balance between the clustering reactions

$$H^+(H_2O)_n + H_2O + M \rightleftharpoons H^+(H_2O)_{n+1} + M \tag{1.38}$$

and dissociative recombination

$$H^+(H_2O)_n + e \rightarrow H + (n-m)(H_2O) + (H_2O)_m. \tag{1.39}$$

The reaction rates associated with clustering reactions (1.38) are very sensitive to the temperature and the water vapour content of the atmosphere, allowing hydration orders of 2-4 to dominate under normal conditions (hydration orders as high as 20 have been observed near the summer mesopause at high latitudes [*Bjorn and Arnold*, 1981]). Atomic oxygen also plays an important role in this reaction chain as it can prevent the switching of O_2 and H_2O in the second step of the reaction sequence by reacting with the O_4^+ ion

$$O_4^+ + O \rightarrow O_2^+ + O_3. \tag{1.40}$$

This reaction increases in importance as one moves higher into the D-region and is believed to be the primary reason why the concentrations of $H^+(H_2O)_n$ decrease so rapidly in the upper D-region.

There is also an appreciable number of NO hydrates, i.e. $NO^+(H_2O)_n$, in the middle of the D-region. The $NO^+(H_2O)$ ion is primarily formed by the reaction sequence

$$NO^+ + N_2 + M \rightarrow NO^+(N_2) + M \tag{1.41}$$

$$NO^+(N_2) + CO_2 \rightarrow NO^+(CO_2) + N_2 \tag{1.42}$$

$$NO^+(CO_2) + H_2O \rightarrow NO^+(H_2O) + CO_2 \tag{1.43}$$

[*Fehsenfeld and Ferguson*, 1969], with the second and third order NO hydrates forming in a similar manner (Figure 1.7). The third order NO hydrate then reacts with water vapour to form the third order proton hydrate,

$$NO^+(H_2O)_3 + H_2O \rightarrow H^+(H_2O)_3 + HNO_2, \tag{1.44}$$

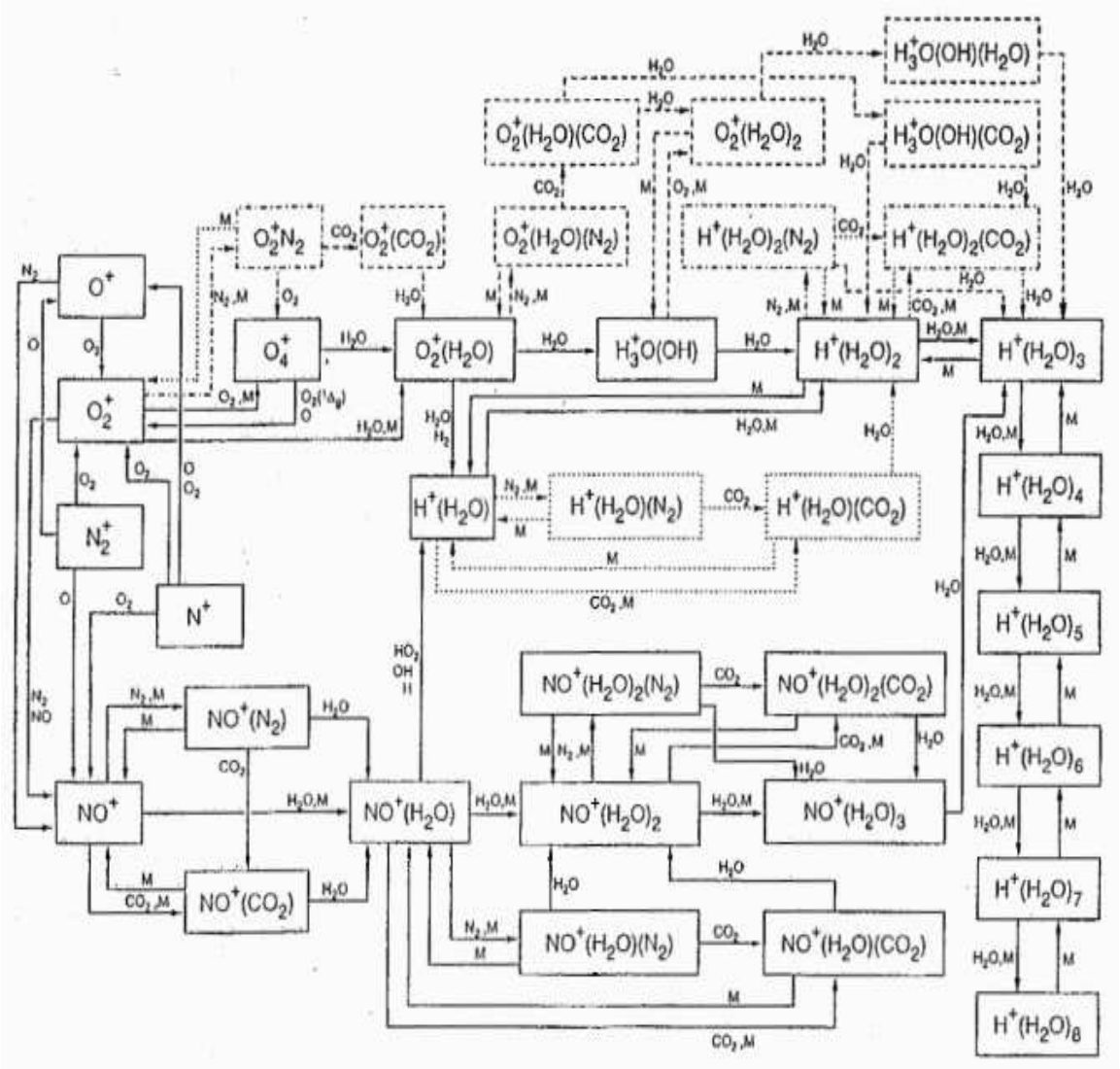

Figure 1.7: Positive ion chemistry of the D-region. There are four reaction channels: the first involves atomic and molecular ions (O^+, O_2^+, N_2^+, NO^+, and O_4^+), the second is associated with proton hydration processes ($H^+(H_2O)_n$, n = 1 to 8), the third is associated with the NO hydration process and the fourth involves clustering between ions and other minor constituents (CO_2, OH, and HO_2). (From *Turunen et al.*, 1996.)

allowing more higher order proton hydrates to form through equation (1.38).

The presence of negative ions in the lower D-region is also due to three-body reactions, with O^- and O_2^- forming through the electron attachment reactions

$$e^- + O_2 + M \rightarrow O_2^- + M \tag{1.45}$$

$$e^- + O_3 \rightarrow O^- + O_2. \tag{1.46}$$

These ions then initiate reaction chains that eventually culminate in the production of NO_3^- ions, with the two most probable chains being

$$O_2^- \xrightarrow{O_2} O_4^- \xrightarrow{CO_2} CO_4^- \xrightarrow{NO} NO_3^- \tag{1.47}$$

$$O^- \xrightarrow{O_2} O_3^- \xrightarrow{CO_2} CO_3^- \xrightarrow{NO} NO_2^- \xrightarrow{O_3} NO_3^- \tag{1.48}$$

[*Wayne*, 2000]. Many other negative ions exist in the lower D-region, e.g. O_4^-, OH^-, and HCO_3^-, but our understanding of negative ion chemistry in the D-region is still quite rudimentary and will not be discussed any further in this thesis.

1.3.2 The E-region

The E-region, like the upper D-region, is composed primarily of NO^+, O_2^+ and electrons (Figure 1.6). The primary sources of ionization in this region are the photoionization of O_2 in the 800-1027 Å interval and N_2, O_2, and O in the 10-100 Å interval, which generates substantial numbers of N_2^+, O_2^+ and O^+ ions. The O^+ and N_2^+ ions are rapidly converted to O_2^+ and NO^+ through reactions (1.27)-(1.30), with the effects of reaction (1.28) becoming more important as one moves higher into the E-region. The O_2^+ ions can also react with N_2 and NO, i.e.

$$O_2^+ + N_2 \rightarrow NO^+ + NO \tag{1.49}$$

$$O_2^+ + NO \rightarrow NO^+ + O_2 \tag{1.50}$$

which increases the NO^+ concentrations at the expense of O_2^+. The electron densities at E-region altitudes are also high enough that one must account for the effects of dissociative recombination

$$O_2^+ + e^- \rightarrow O + O \tag{1.51}$$

$$NO^+ + e^- \rightarrow N + O \tag{1.52}$$

which further exacerbates the NO^+ concentration excess as most O_2^+ ions will collide with NO and form NO^+ through reaction (1.50) before they encounter an electron.

The strong diurnal variation in the E-region number densities is primarily due to the large drop in the UV photon flux at night. Resonantly scattered Lyman-α photons and starlight from O and B type stars in the 911-1026 Å interval become the primary sources of ionization and allow ion densities of 10^2-10^3 cm^{-3} to be maintained throughout the night. Resonantly scattered photons at 304 Å, 584 Å, and Lyman-β also make a small contribution to the nighttime ionization rates at their respective wavelengths (a more detailed discussion on the ionization rates of the nighttime E-region can be found in *Titheridge* [2000]).

1.3.3 The F-region

The F-region is normally divided into three subregions: the F1-region (where photochemistry dominates), the F2-region (a transition region where chemistry and diffusion both play an important role), and the topside ionosphere (where diffusion dominates). The primary source of ionization in the F1-region is the photoionization of atomic oxygen by UV photons in the 200-911 Å interval, which produces large amounts of O^+. The O^+ ions are then converted to NO^+ and O_2^+ through reactions (1.29)-(1.30), which can then recombine through reactions (1.51) and (1.52). The concentrations of molecular oxygen and nitrogen play a pivotal role in this region as they limit the ability of reactions (1.29)-(1.30) to create NO^+ and O_2^+ ions, causing the composition of the F1-region to change from molecular ions in the lower F1-region to O^+ ions in the upper F1-region (the O^+ number density increases exponentially with height in the F1-region as the photoionization rate, which is proportional to the atomic oxygen number density, decreases at a rate that is much slower than the loss rate imposed by reactions (1.29)-(1.30)). The dependence of the ion number densities on the photoionization rates also means that the ion densities in the F1-region exhibit a pronounced diurnal variation, like the D- and E-regions below it.

The inability of an atomic oxygen ion to recombine with an electron means that the photochemical lifetime of an O^+ ion will increase with altitude (atomic recom-

bination reactions cannot conserve energy and momentum simultaneously unless the excited 'intermediate' state decays through a radiative transition). This increasing photochemical lifetime means that the O^+ ions will be influenced more and more by transport processes as one moves higher into the F-region, with the F2-region density peak occuring where transport and photochemical production rates are equal. The subsequent decrease in ion number densities is primarily controlled by a balance between the ionospheric pressure gradient and gravity, which allows H^+ to become the dominant ion by 1000 km.

1.4 High Latitude Electrodynamics

The high latitude ionosphere can exhibit substantial deviations from the behaviour described in the previous section as it is coupled to the magnetosphere through electric fields, particle precipitation, and field-aligned currents (Figure 1.8). There are usually two types of field-aligned currents; Region-1 field-aligned currents and Region-2 field-aligned currents. Region-1 field-aligned currents occur on the polewards edge of the auroral zone (they flow into the ionosphere in the morning sector and away from the ionosphere in the evening sector) and are believed to close near the magnetopause or in the distant magnetotail [*Toffoletto and Hill*, 2000]. Region-2 field-aligned currents occur on the equatorwards edge of the auroral zone (they flow down into the ionosphere on the duskside and up on the dawnside) and close in the inner plasma sheet and ring current regions [*Schield et al.*, 1969; *Wolf*, 1983]. During geomagnetically disturbed periods, the Region-1 and Region-2 current patterns expand to lower latitudes, with the average latitudinal width of the Region-1 and Region-2 current patterns increasing by 20% [*Iijima and Potemra*, 1978].

The upwards flowing components of the Region-1 and Region-2 current systems are primarily composed of precipitating magnetospheric electrons, causing these regions to charge up negatively. Field-aligned currents flowing down into the ionosphere are carried by upwards flowing ionospheric electrons, causing the ionospheric portion of these regions to charge up positively. The resultant charge distribution produces

an electric field pattern in the auroral zone that is directed towards the upwards flowing current regions in the evening and morning sectors. The electric fields in the polar cap region, on the other hand, are quite variable in space and time as they are a consequence of magnetic reconnection near the dayside magnetopause; the upper example in Figure 1.9 exhibits an essentially constant strength across the polar cap region while the lower example exhibits relatively large values in the morning flank of the polar cap region.

At ionospheric heights, the current density $\mathbf{J}$ and the "effective" electric field $\mathbf{E} + \frac{1}{c}(\mathbf{u}_n \times \mathbf{B})$ are normally related through the Ohm's Law expression

$$\mathbf{J} = \sigma \left(\mathbf{E} + \frac{1}{c} (\mathbf{u}_n \times \mathbf{B}) \right) \tag{1.53}$$

where σ is the conductivity tensor, $\mathbf{u}_n$ is the neutral velocity vector, $\mathbf{B}$ is the geomagnetic field vector, and c is the speed of light (a more thorough discussion on the validity of this expression can be found in *Schunk and Nagy* [2000]). If one decomposes the 'effective' electric field into components that are perpendicular to and parallel with the geomagnetic field, then equation (1.53) can also be written as

$$\mathbf{J} = \sigma_{\parallel}\mathbf{E}_{\parallel} + \sigma_P \left(\mathbf{E}_{\perp} + \frac{1}{c}(\mathbf{u}_n \times \mathbf{B}) \right) + \sigma_H \, \mathbf{b} \times \left(\mathbf{E}_{\perp} + \frac{1}{c}(\mathbf{u}_n \times \mathbf{B}) \right) \tag{1.54}$$

where $\mathbf{b}$ is the unit vector parallel to $\mathbf{B}$ and $\sigma_{\parallel}$, σ_P, and σ_H are the parallel, Pedersen and Hall conductivities respectively. Expressions for the parallel, Pedersen, and Hall conductivities are given by

$$\sigma_{\parallel} = e^2 \left(\frac{n_e}{m_e \nu_e} + \sum_i \frac{n_i}{m_i \nu_i} \right) \tag{1.55}$$

$$\sigma_P = e^2 \left(\frac{n_e}{m_e \nu_e} \frac{\nu_e^2}{\nu_e^2 + \omega_e^2} + \sum_i \frac{n_i}{m_i \nu_i} \frac{\nu_i^2}{\nu_i^2 + \omega_i^2} \right) \tag{1.56}$$

$$\sigma_H = e^2 \left(\frac{n_e}{m_e \nu_e} \frac{\nu_e \omega_e}{\nu_e^2 + \omega_e^2} - \sum_i \frac{n_i}{m_i \nu_i} \frac{\nu_i \omega_i}{\nu_i^2 + \omega_i^2} \right) \tag{1.57}$$

where $\omega_i = \frac{eB}{m_i}$ is the gyrofrequency of ion species i, $\omega_e = \frac{eB}{m_e}$ is the electron gyrofrequency, ν_i is the ion-neutral collision frequency for ion species i, ν_e is the electron-neutral collision frequency, n_i is the number density of ion species i, and n_e is the

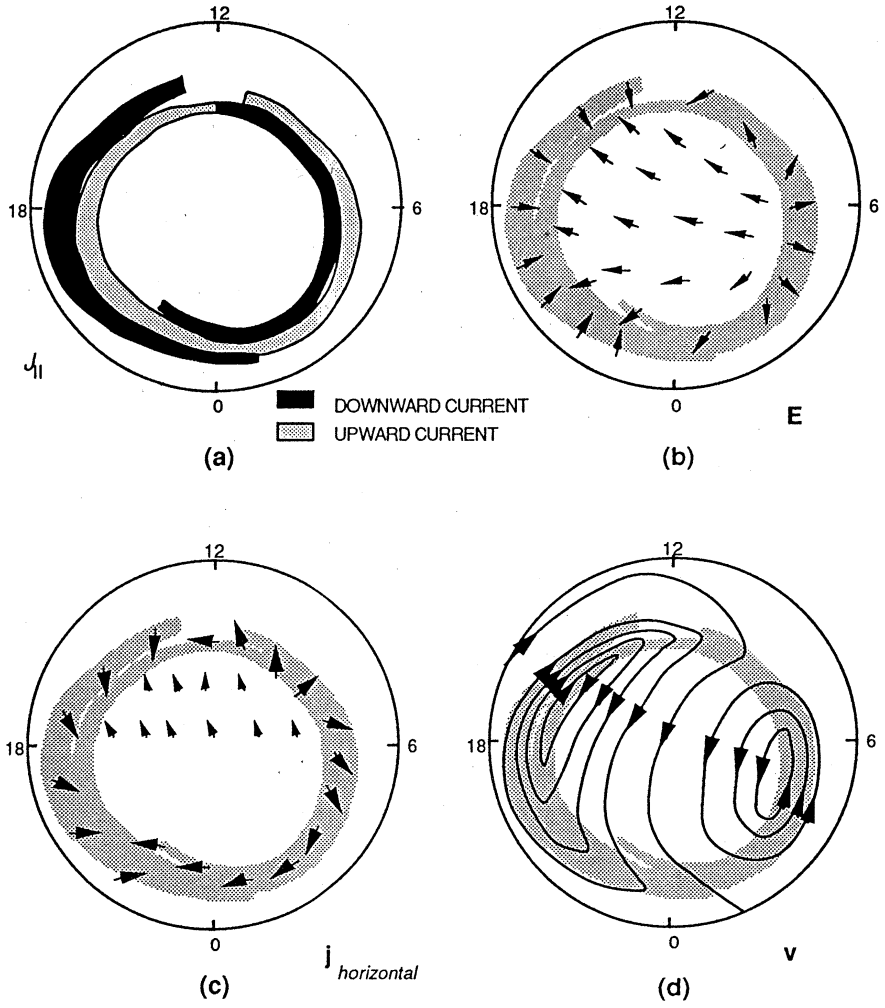

Figure 1.8: Schematic diagram of currents and electric fields in the high latitude ionosphere. The field-aligned current pattern (panel (a)) has been adapted from *Iijima and Potemra* (1978). Regions where these currents flow down into the ionosphere tend to charge up positively and regions where these currents flow upwards tend to charge up negatively, producing an electric field pattern in the auroral zone that is directed towards the upwards flowing current regions in the evening and morning sectors (panel (b)). The electric field pattern that is directed from dawn-to-dusk in the polar cap region is a consequence of magnetic reconnection near the dayside magnetopause. The horizontal currents (panel (c)) connect the upwards and downwards flowing field-aligned current regions. The resulting $\mathbf{E} \times \mathbf{B}$ drift velocities (panel(d)) are anti sunwards in the polar cap region and sunwards in the morningside and afternoonside auroral zones. (From *Wolf*, 1995.)

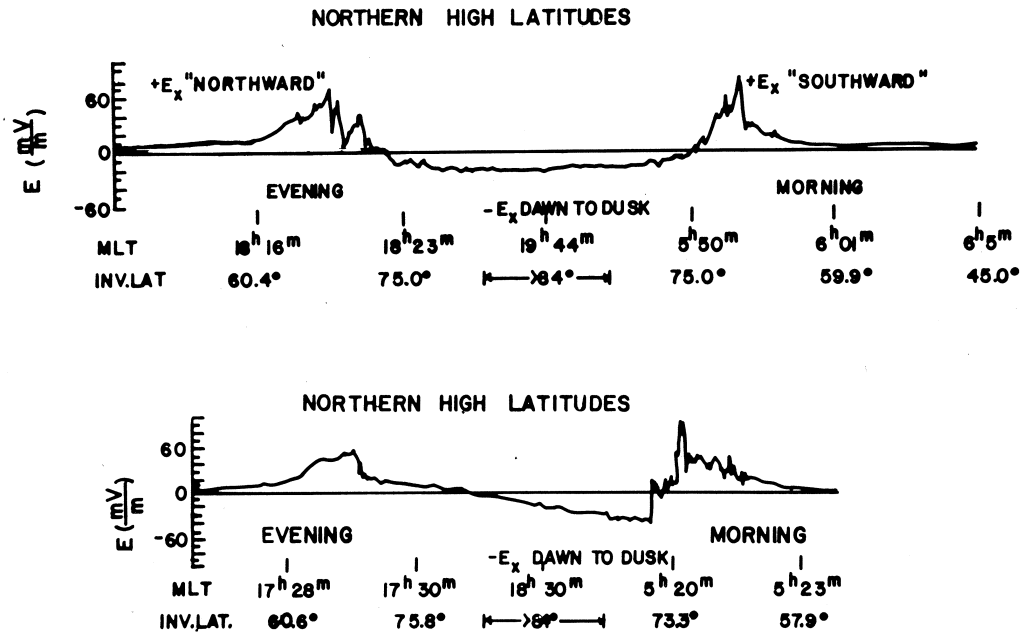

Figure 1.9: OGO-6 electric field measurements in the high latitude thermosphere during disturbed conditions. The electric field in the evening sector is directed northwards (polewards) in both examples and reaches values of 60-80 mV/m in the eveningside auroral zone. The electric field in the morning sector is directed southwards (equatorwards) in both examples and reaches very large values in the auroral zone. The characteristics of the electric field in the polar cap region are very different: the upper example has an essentially uniform strength across the polar cap region (Type A electric field pattern) while the lower example has relatively large values in the morning flank of the polar cap region (Type B electric field pattern). (Reprinted from *Planet. Space Sci.*, Vol. 20, J. W. Heppner, Electric field variations during substorms: OGO-6 measurements, pp. 1475-1498, Copyright 1972, with permission from Elsevier.)

electron number density. It should be noted that there are two components to the perpendicular current: the Pedersen current, which is in the direction of $\mathbf{E}_\perp$, and the Hall current, which is perpendicular to both $\mathbf{E}_\perp$ and $\mathbf{B}$.

1.5 Observations of the disturbed high latitude thermosphere

1.5.1 Neutral Density Changes

Neutral density enhancements during geomagnetically disturbed periods have been inferred from the orbital decay of satellites since the dawning of the space age [*Jacchia*, 1959, 1961; *Groves*, 1961; *Jacchia and Slowey*, 1963, 1964; *Moe*, 1966; *Roemer*, 1966, 1971; *Jacchia, Slowey and Verniani*, 1967]. The large altitude and latitude ranges of these studies revealed several characteristics about these enhancements

i) the relative amplitudes of these enhancements increased with height (Figure 1.10)

ii) the higher latitude density enhancements tended to be larger than those occurring at middle and lower latitudes and

iii) the higher latitude density enhancements normally preceded the density enhancements at middle and lower latitudes

Unfortunately, satellite-decay derived density values also had very poor spatial resolution as the measured values had to be averaged over arcs of 30^o or more (Figure 1.11). The deployment of accelerometers on-board LEO satellites in the late 1960s provided much better spatial resolution (Figures 1.12-1.14) and revealed that density depletions could sometimes occur in the lower regions of the auroral thermosphere. (Figure 1.12).

1.5.2 Composition changes

Most quantitative observations of thermospheric composition changes in the 1970s and early 1980s were obtained from mass spectrometers on-board low Earth orbit (LEO) satellites. Inferring absolute number densities from these measurements was not a trivial task as the calibration factor that related the number densities in the ion source to the ambient number densities had to account for the high orbital speed

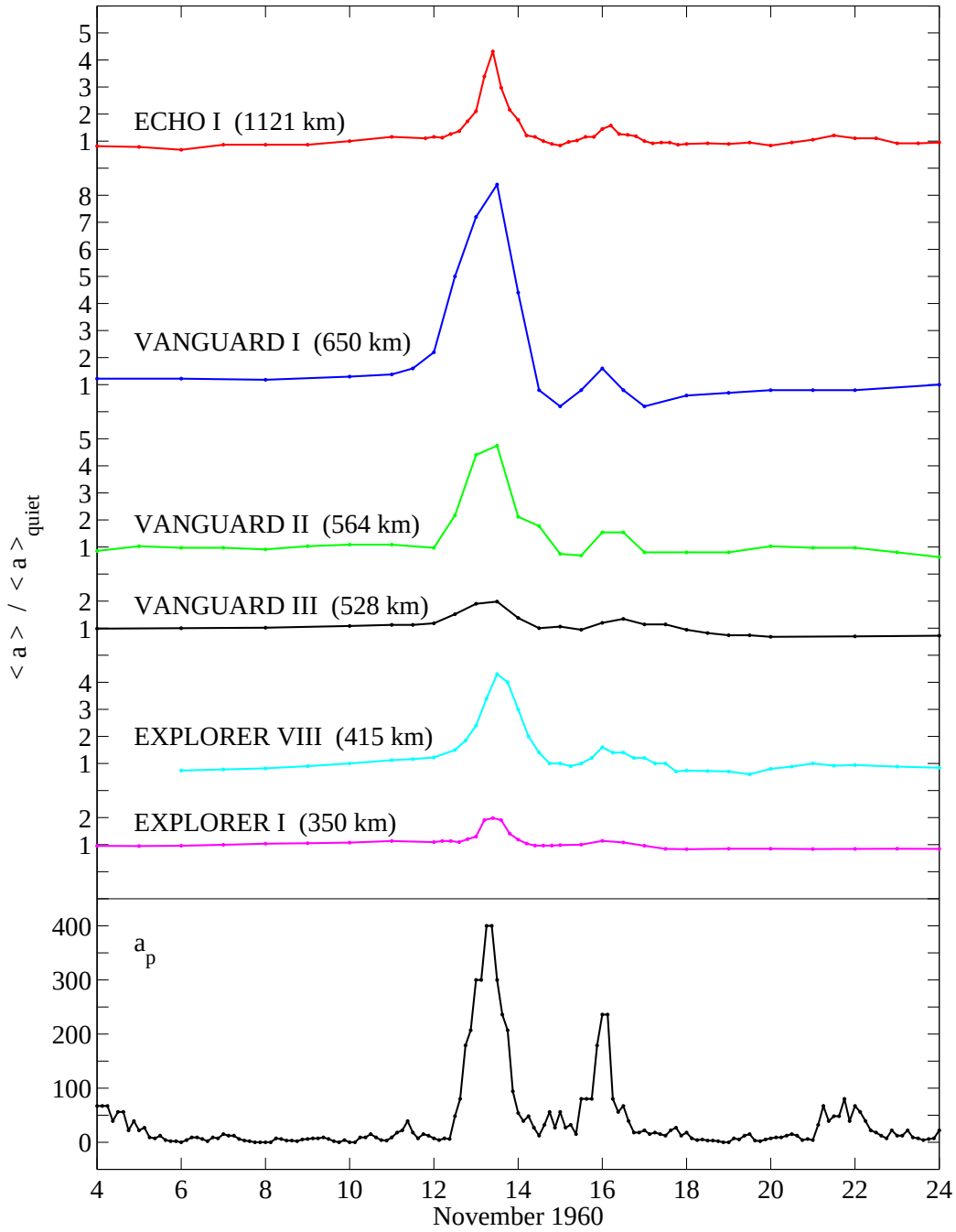

Figure 1.10: Atmospheric drag on seven satellites during the November 1960 storm period. The decelerations experienced by the seven satellites during the storms of November 12-14 and November 15-17 are presented in the upper panel of this figure and have been 'normalized' to the mean deceleration experienced by these satellites during the geomagnetically quiet periods that preceded and followed the November 12-17 time period. The perturbations in the satellite decelerations tend to increase with height and follow the a_p index (lower panel) quite closely. (From *Jacchia*, 1961.)

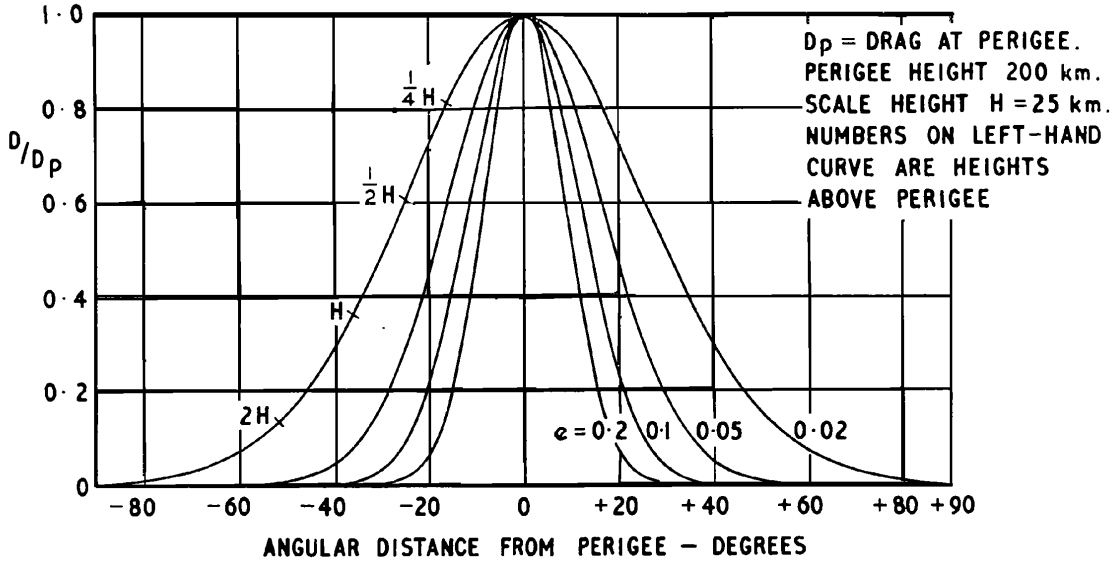

Figure 1.11: Satellite drag as a function of angular distance from perigee. The ratio of the drag D to the drag at perigee D_P is plotted for several orbital eccentricities and highlights the spatial averaging that occurs in any orbital drag determination of the neutral density. (From *King-Hele*, 1966.)

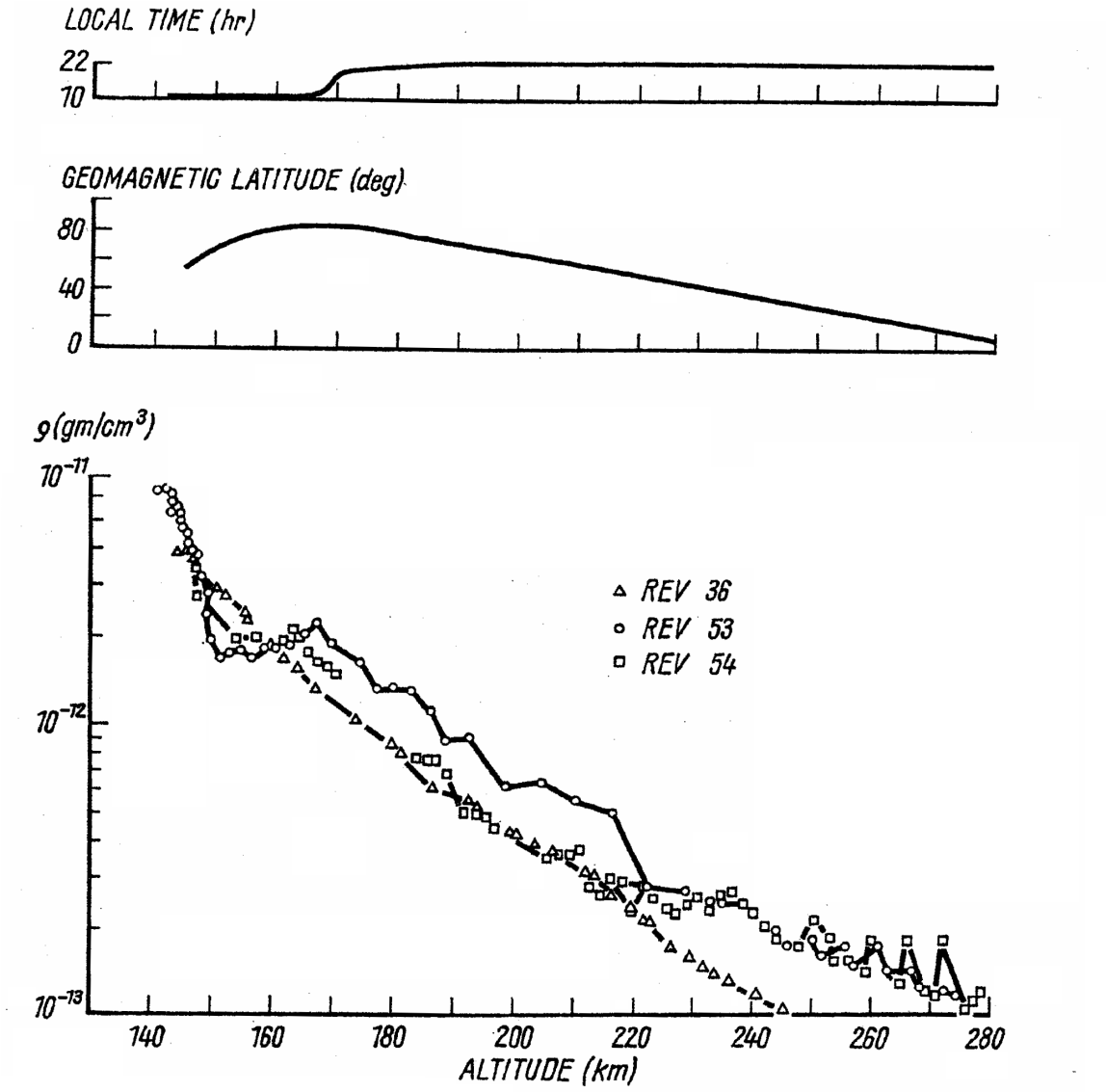

Figure 1.12: Neutral density measurements from the LOGACS accelerometer during the May 24-25, 1967 time period. Orbit 36 occurred during geomagnetically quiet conditions while orbits 53 and 54 were made during disturbed conditions. There is a moderate density depletion in the 140-160 km region and significant density enhancements above 170 km on orbit 53. Both of these features are less prominent on orbit 54. (Reprinted from *Space Research*, Vol. 12, L. L. DeVries, Analysis and interpretation of density data from the Low-G accelerometer calibration system (LOGACS), pp. 777-789, Copyright 1972, with permission from Elsevier.)

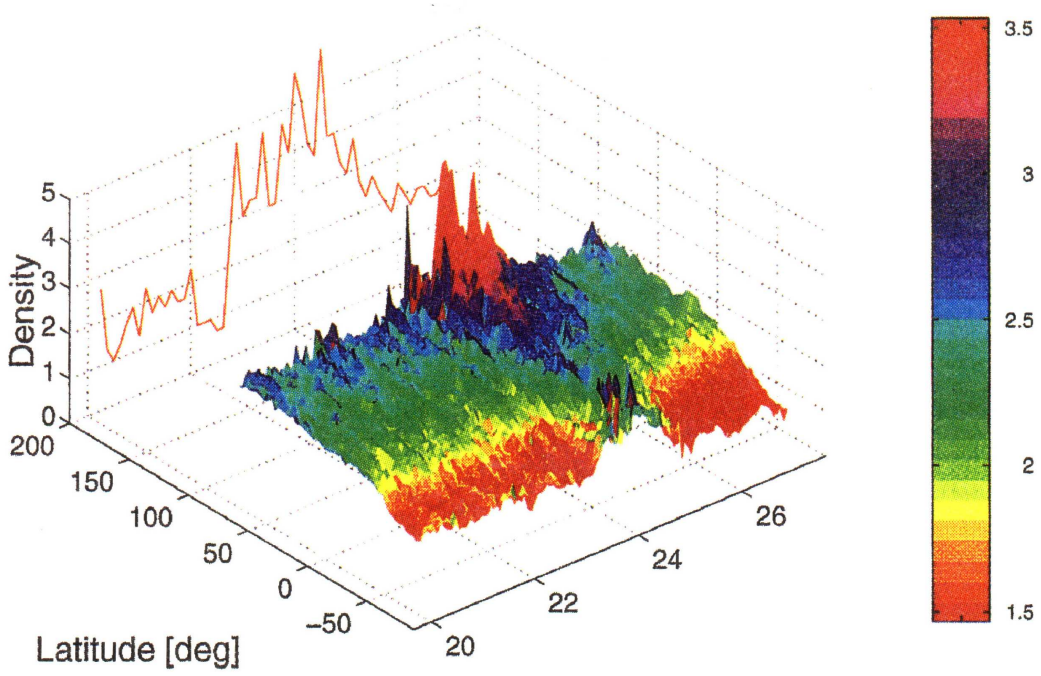

Figure 1.13: SETA neutral mass densities at 200 km during the July 20-26, 1983 time period. The neutral density enhancements during the July 23-25 storm period are on the order of 50% and coincide with the increase in the K_p index. The density scale is in units of $10^{-13}\,\mathrm{g\,cm^{-3}}$ and also serves as the scale for the K_p values in the vertical plane. The neutral density measurements extend from -80^o to 80^o latitude; the latitude scale has been extended past 80^o so the K_p values can be displayed clearly. (From *Forbes et al.*, 1996.)

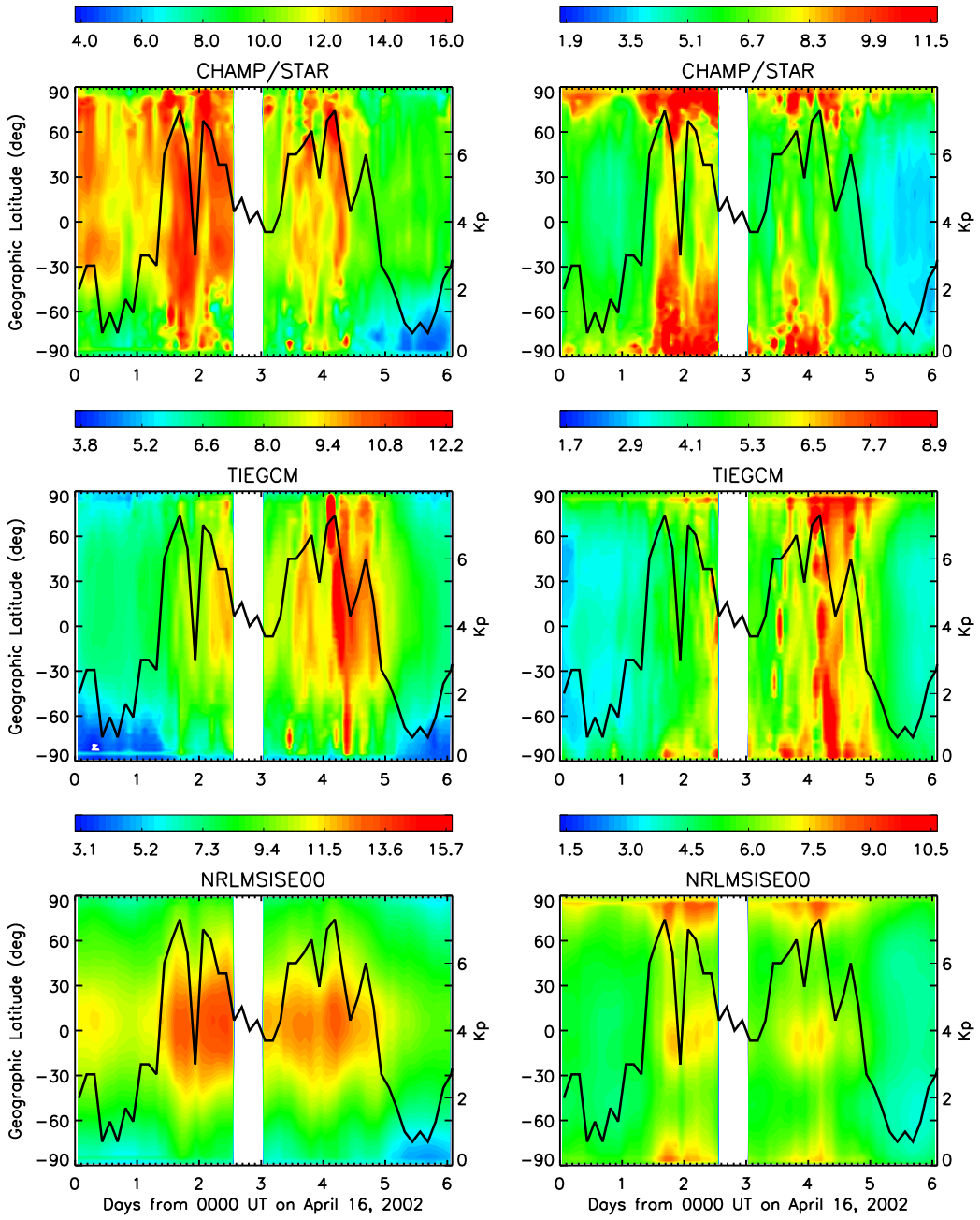

Figure 1.14: CHAMP neutral mass densities at 410 km during the April 16-22 2002 time period. Neutral density enhancements derived from the CHAMP accelerometer (top) near 1530 LT (left) and 0430 LT (right) are on the order of 100% at 410 km and coincide with the increase in the K_p index (black solid line, scaled on the right hand side of each panel). Densities from an NCAR TIEGCM simulation of this period (middle) and the empirical NRLMSISE00 model (bottom) are also shown for comparison purposes. The neutral density colour scales are in units of $10^{-15}\,\mathrm{g\,cm^{-3}}$ and have been individually scaled to illustrate the dynamic range of each data set. (From *Forbes et al.*, 2005.)

($\approx$ 8 km/sec) of the satellite and the large temperature difference that frequently existed between the ambient atmosphere and the ion source region (temperatures in the ion source region were normally 300-350 K [*von Zahn*, 1974] while temperatures in the thermosphere could range from 500-1500 K). Atomic oxygen and other reactive species also had a tendency to react with any satellite surface they collided with, making the determination of these number densities particularly difficult as high speed gas-surface interactions were not very well understood during this time period.

Closed ion source mass spectrometers, e.g. the gas analyzers on OV3-6 [*Philbrick*, 1974], OGO-6 [*Carignan and Pinkus*, 1968], ESRO-4 [*Trinks and von Zahn*, 1975], S3-1 [*Philbrick*, 1976], Dynamics Explorer 2 [*Carignan et al.*, 1981] and the Neutral Atmosphere Composition Experiment on the Atmosphere Explorer series of satellites [*Pelz et al.*, 1973], addressed the calibration issue by allowing the ambient species to be thermally accommodated in an antechamber before entering the ionization region (Figure 1.15). This ensured that any *non-reactive* ambient species entering the ionization region would have been at the same temperature as the accommodation chamber surfaces, eliminating any uncertainties about incomplete accommodation in these measurements. It also allowed one to relate the number densities in the ion source to the ambient number densities through the expression

$$n_s = n_a \left[\left(\frac{T_a}{T_s} \right)^{1/2} F(S) \cos^2 \left(\frac{\alpha}{2} \right) + 1 \right] \tag{1.58}$$

where

n_a = the ambient number density
n_s = the number density inside the mass spectrometer ion source
T_a = the ambient temperature
T_s = the temperature inside the mass spectrometer ion source
α = the angle of attack (the angle between the satellite velocity vector and the normal of the ion source opening)
$F(S)$ = the velocity correction function

$$F(S) = \exp(-S^2) + \sqrt{\pi} \cdot S[1 + \mathrm{erf}(S)]$$

with

$$\begin{aligned} S &= \text{the speed ratio } (\, V \cos\alpha \,/\, V_{th} \,) \\ V &= \text{the velocity of the satellite with respect to the atmosphere} \\ V_{th} &= \text{the thermal (most probable) speed of the ambient species} \end{aligned}$$

[*Hedin et al.*, 1964]. If the angle of attack was small, the large orbital speed of a LEO satellite allowed the number densities in the ionization region to become much larger than the ambient number densities, reducing the statistical uncertainties in the measurements and extending the measurement range of the spectrometer.

Unfortunately, the closed ion source solution also exacerbated the reactive species issue as almost all of the ambient atomic oxygen entering the accommodation chamber would be adsorbed or recombine with an atomic oxygen atom that was already adsorbed on one of the accommodation chamber surfaces. The molecular oxygen created during this recombination process was indistinguishable from the ambient molecular oxygen being thermalized in the accommodation chamber, introducing substantial uncertainties into the interpretation of the atomic oxygen and molecular oxygen measurements (Figure 1.16). However, *Hedin et al.* (1973) were able to demonstrate that ambient atomic oxygen densities could be inferred from the atomic and molecular oxygen densities in the ion source through the expression

$$n_a^{\mathrm{o}} \;=\; \left[\, n_s^{\mathrm{o}} \,+\, \frac{1}{\sqrt{2}}\, 2\, n_s^{\mathrm{o}_2} \,\right] \left(\frac{T_s}{T_a}\right)^{1/2} \,+\, \left(\frac{C}{T_a}\right)^{1/2} \left(\frac{d\, n_w^{\mathrm{o}}}{dt}\right) \tag{1.59}$$

if the ambient molecular oxygen number densities were negligible (n_w^{o} is the surface density of adsorbed O in the accommodation chamber and C is a constant that accounts for geometry of the accommodation chamber). If the ambient molecular oxygen densities were not negligible, the total ambient oxygen content, $\mathrm{O} + 2\mathrm{O}_2$, was determined.

This inability of closed ion source mass spectrometers to differentiate between atomic and molecular oxygen led to the development of quasi-open ion source mass spectrometers, e.g. the Open Source Spectrometer (OSS) on the Atmosphere Explorer series of satellites [*Nier et al.*, 1973], the Neutral and Ion Mass Spectrometer (NIMS) on the Aeros series of satellites [*Krankowsky et al.*, 1974], and the Upper

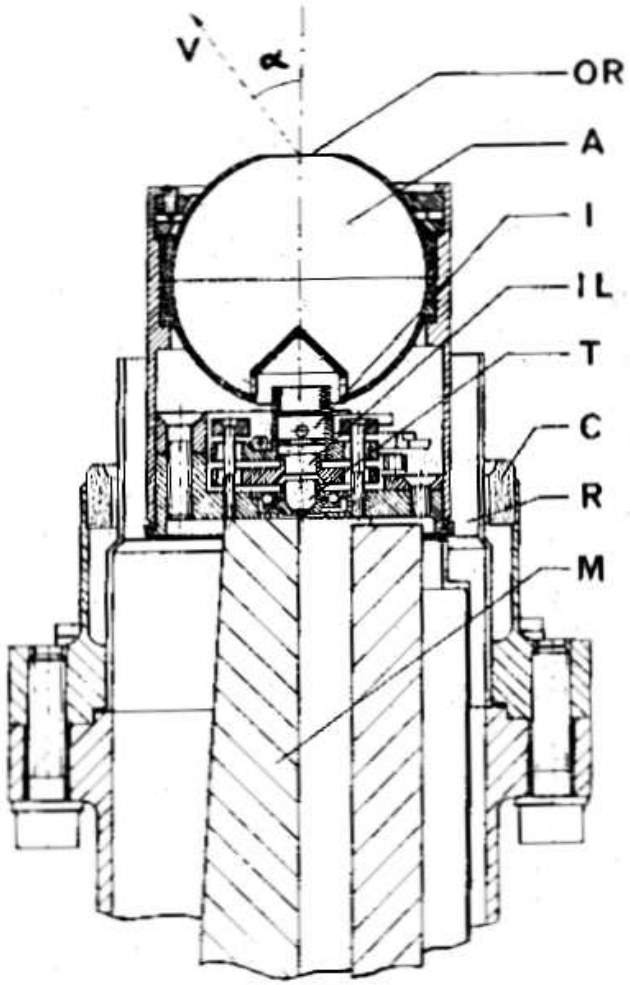

Figure 1.15: Schematic diagram of the ESRO-4 mass spectrometer. Ambient atoms and molecules entered the spherical antechamber (**A**) through a knife-edged orifice (**OR**). The antechamber had an inner diameter of 40 mm, which ensured that any particles entering the ionization region (**I**) would have been thermally accommodated to the temperature of the antechamber walls, i.e. they would have undergone several collisions with the antechamber walls before entering the ionization region. A 75 eV beam of electrons then ionized the thermally accommodated particles, which were focused by the ion lenses (**IL** and **T**) into the monopole spectrometer (**M**). (Reused with permission from H. Trinks and U. von Zahn, *Review of Scientific Instruments*, 46, 213 (1975). Copyright 1975, American Institute of Physics.)

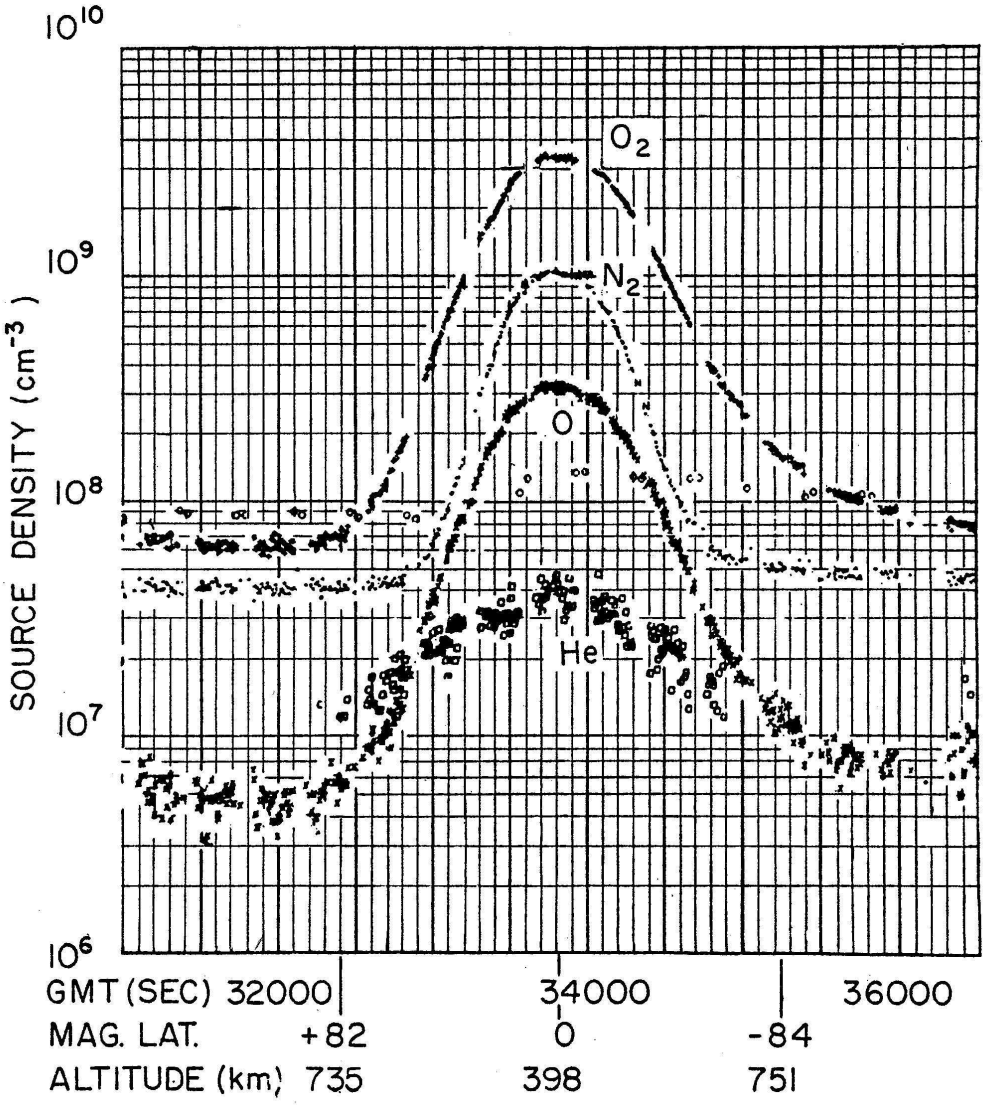

Figure 1.16: Neutral number densities in the ion source of the OGO-6 mass spectrometer. These number densities were obtained during geomagnetically quiet conditions on September 27, 1969 at 1000 GMT and show that most of the atomic oxygen has recombined into molecular oxygen. (From *Taeusch et al.*, 1971.)

Atmosphere Composition Spectrometer (UACS) on the S85-1 satellite [*Kayser et al.*, 1986]. These mass spectrometers allowed the ambient gas into the ionization region after passing through a series of high transmission grids (Figure 1.17), which ensured that a small percentage of the ambient particles would not have collided with *any* of the ion source surfaces before being ionized. Since the ionization process did not change the momentum of the particles appreciably, these ionized particles were still quite energetic, with energies of 0.37 eV/amu being typical [*Nier*, 1985]. Any ambient particles that did collide with an ion source surface usually underwent several more collisions in the ion source region before being ionized, creating a second, much larger population of low energy ions in the ion source.

In the "normal" mode of operation, the low and high energy ion populations were drawn out of the ionization region by the potential difference across the ion source (this potential difference was created in the OSS ion source by keeping the potential of the third grid above the potential of the ion source housing and the potential of the focusing plates below the potential of the ion source housing). The ions were then directed into the analyzer section of the spectrometer where the appropriate *low energy* ion populations were selected and counted. Laboratory studies with high speed molecular beams had demonstrated that most of the low energy ion species produced in the OSS ion source had been largely accommodated [*Hayden et al.*, 1974], which allowed the ambient densities to be determined from the ion source number densities through equation (1.58) if an additional factor that accounted for the effects of incomplete accommodation was incorporated into the data processing.

Quasi-open ion source spectrometers operating in this "normal" mode were not able to differentiate between atomic and molecular oxygen as the low energy ions being measured in this mode had been created from particles that were essentially accommodated, introducing the same ambiguities that plagued closed ion source source measurements. However, if the potential difference across the ion source region was set to zero, the low energy ion species were not able to leave the ion source as they did not have enough energy to overcome the negative space charge of the electron beam. The energetic ion populations, on the other hand, were able to overcome this retarding

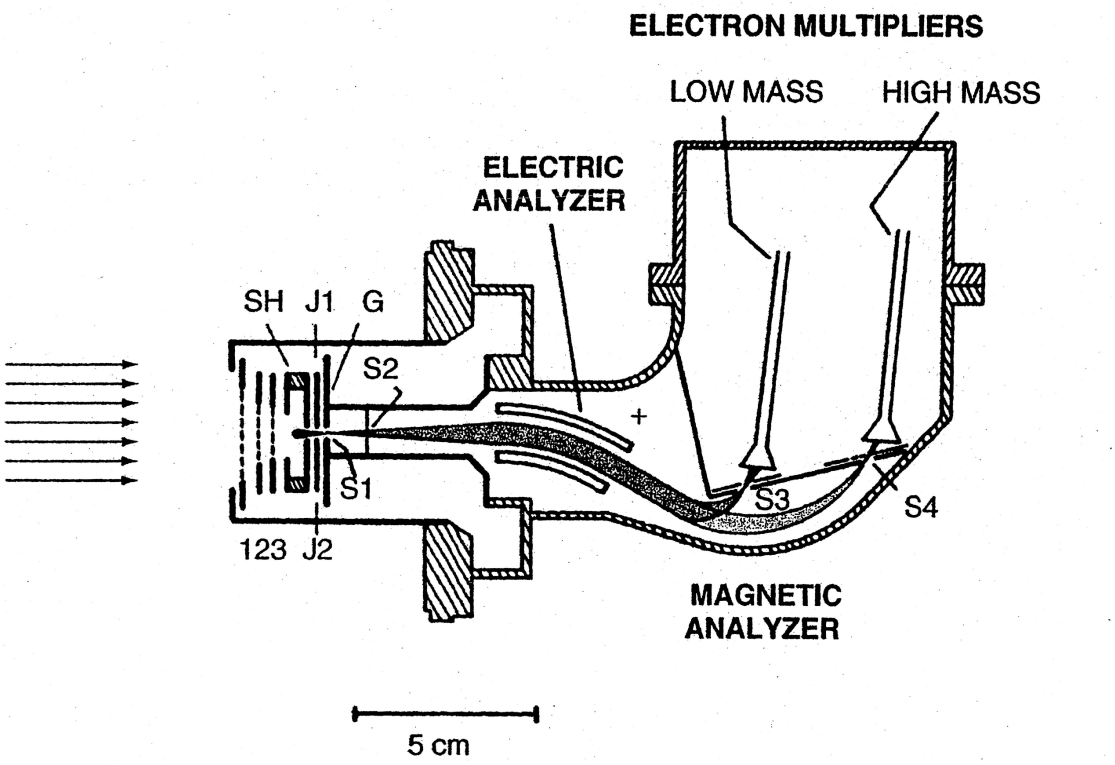

Figure 1.17: Schematic diagram of the AE-C Open Source Spectrometer. The ambient gas particles entered the the spectrometer from the left and passed through three high transmission grids that minimized the effects of any stray electric fields. These gas particles were then ionized by a 75 eV beam of electrons and collimated by the J1-J2 focusing plates (the electron beam is perpendicular to the plane of the figure and is represented by the dot in the middle of the ion source assembly SH). These collimated ions were then directed into the electric and magnetic field analyzers, where the desired ions were selected and counted. (From *Nier et al.*, 1973.)

potential quite easily as the ambient particles they had been created from were totally unaccommodated, i.e. the energetic ion populations had been created from ambient particles that had not collided with any of the ion source surfaces. This exclusion of the low energy ion populations from subsequent analysis in the spectrometer ensured that there were no recombination ambiguities in these "fly-through" mode measurements of atomic and molecular oxygen. More details about the "fly-through" mode of AE-C and the procedures used in converting the ion source densities to ambient number densities can be found in *Nier et al.* (1974).

Now that the potential issues in the interpretation of closed and quasi-open mass spectrometers have been identified, some typical examples of mass spectrometer inferred composition changes in the thermosphere can be presented. The first example (Figure 1.18) was obtained by the ESRO-4 spectrometer over the North American continent during the geomagnetic storm of October 29, 1973. The changes in the number densities of argon (AR), molecular nitrogen (N_2), atomic oxygen (O), and helium (HE) have been plotted in the middle panel and show that there is a well defined "disturbance zone" extending from high to middle latitudes. The number densities of argon were substantially enhanced in this "disturbance zone" while the number densities of molecular nitrogen exhibited a smaller but still significant enhancement. The number densities of helium, on the other hand, were substantially reduced in this zone and the number densities of atomic oxygen appeared to exhibit a minor decrease in this region. However, it must be emphasized that ESRO-4 was a closed source mass spectrometer that could not measure atomic oxygen directly; the atomic oxygen number densities being presented in the middle panel were derived from the total oxygen content densities under the assumption that the ambient O_2 number densities were negligible. Outside of this "disturbance zone", all of the constituents exhibited a minor increase. A more complete discussion on the morphology and evolution of these "disturbance zones" can be found in *Prölss* (1980).

Observations of these "disturbance zone" perturbations in thermospheric composition were usually limited to the upper thermosphere as the low eccentricity orbits required to make these measurements were not very conducive to long satellite life-

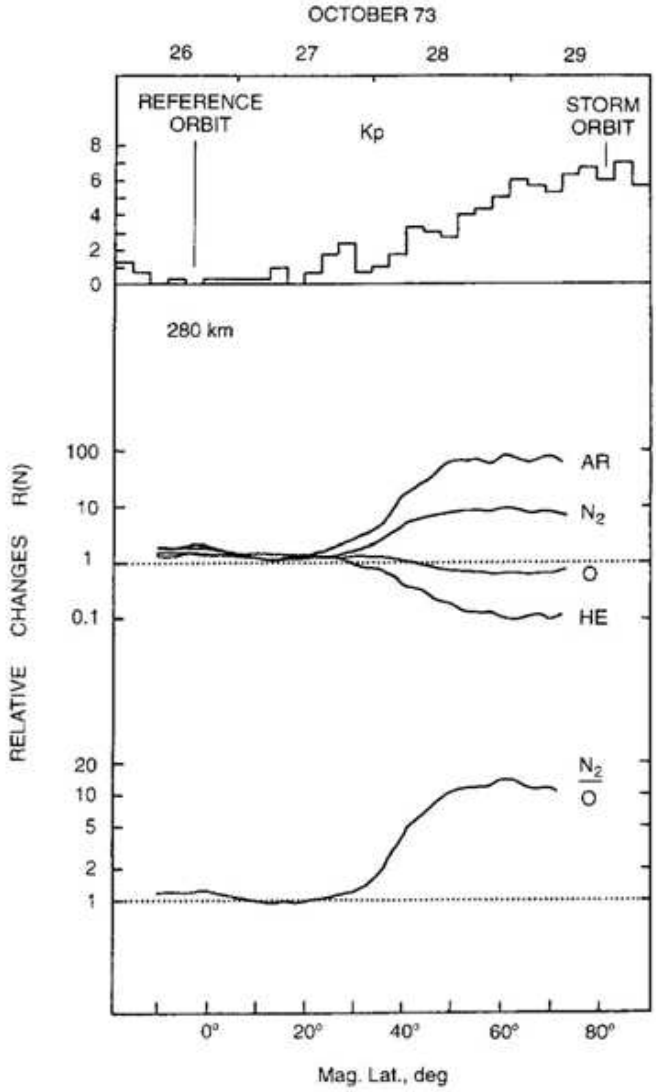

Figure 1.18: Composition changes at 280 km during the geomagnetic storm of October 29, 1973. The development of the geomagnetic storm can be clearly seen in the Kp index (uppermost panel). The associated changes in the number densities of argon (AR), molecular nitrogen (N_2), atomic oxygen (O), and helium (HE) are presented in the middle panel while changes in the N_2/O ratio are presented in the lower panel. In this style of presentation, a relative change of three meant that the disturbed number densities were three times greater than the quiet-time number densities while a relative change of one meant that there has been no change in the number densities. The quiet-time number densities used in the determination of these ratios were obtained from the reference orbit denoted in the upper panel. (Reprinted from *J. Atmos. Terr. Phys.*, Vol. 59, M. Zuzic, L. Scherliess, and G. W. Prölss, Latitudinal structure of thermospheric composition perturbations, pp. 711-724, Copyright 1997, with permission from Elsevier.)

times (a satellite in a low eccentricity orbit at 200 km would have a lifetime of approximately 3 months if its orbit could not be raised). Because of this constraint, measurements of thermospheric composition change below 250 km were usually confined to localized regions around a particular latitude. A typical closed ion source measurement from the mass spectrometer on-board the S3-1 satellite at an altitude of 160 km is presented in Figure 1.19 and shows that the molecular nitrogen densities around 60° N have doubled while the argon number densities have increased by an order of magnitude (bottom panel). The atomic oxygen number densities have been reduced to one-half of their pre-storm values (lower middle panel) and were *inferred* from the total oxygen content densities by removing the molecular oxygen component, i.e.

$$n_{\mathrm{O}} \;=\; n_{(\mathrm{O}+2\mathrm{O}_2)} - 2\left(\frac{\mathrm{O}_2}{\mathrm{N}_2}\right)_{\mathrm{model}} \times n_{\mathrm{N}_2}$$

where $(\mathrm{O}_2/\mathrm{N}_2)_{\mathrm{model}}$ is the $\mathrm{O}_2/\mathrm{N}_2$ ratio from *Jacchia* (1971). The errors introduced by this approximation were believed to be quite small as the effects of vertical transport were assumed to be negligible.

Another example of composition change in the lower thermosphere is presented in Figure 1.20 and was obtained from the OSS on AE-D when it was operating in "fly-through" mode at high latitudes. The *measured* molecular oxygen number densities (solid line, lower panel) were 4 times larger than the the quiet-time values that preceded the storm. The measured molecular oxygen densities were also two times larger than the diffusive equilibrium calculations of O_2 using the measured temperatures and molecular nitrogen number densities (dotted line, lower panel), which suggested that transport by vertical winds may be playing an important role in the evolution of O_2 during this storm. Unfortunately, most OSS measurements of atomic and molecular oxygen were made in the "normal" mode of operation, with the unambiguous "fly-through" mode of operation being restricted to the 200 km region [*Nier et al.*, 1976].

Most of the LEO satellite missions investigating the upper atmosphere in the late 1960s and early 1970s also carried instruments that were capable of measuring at-

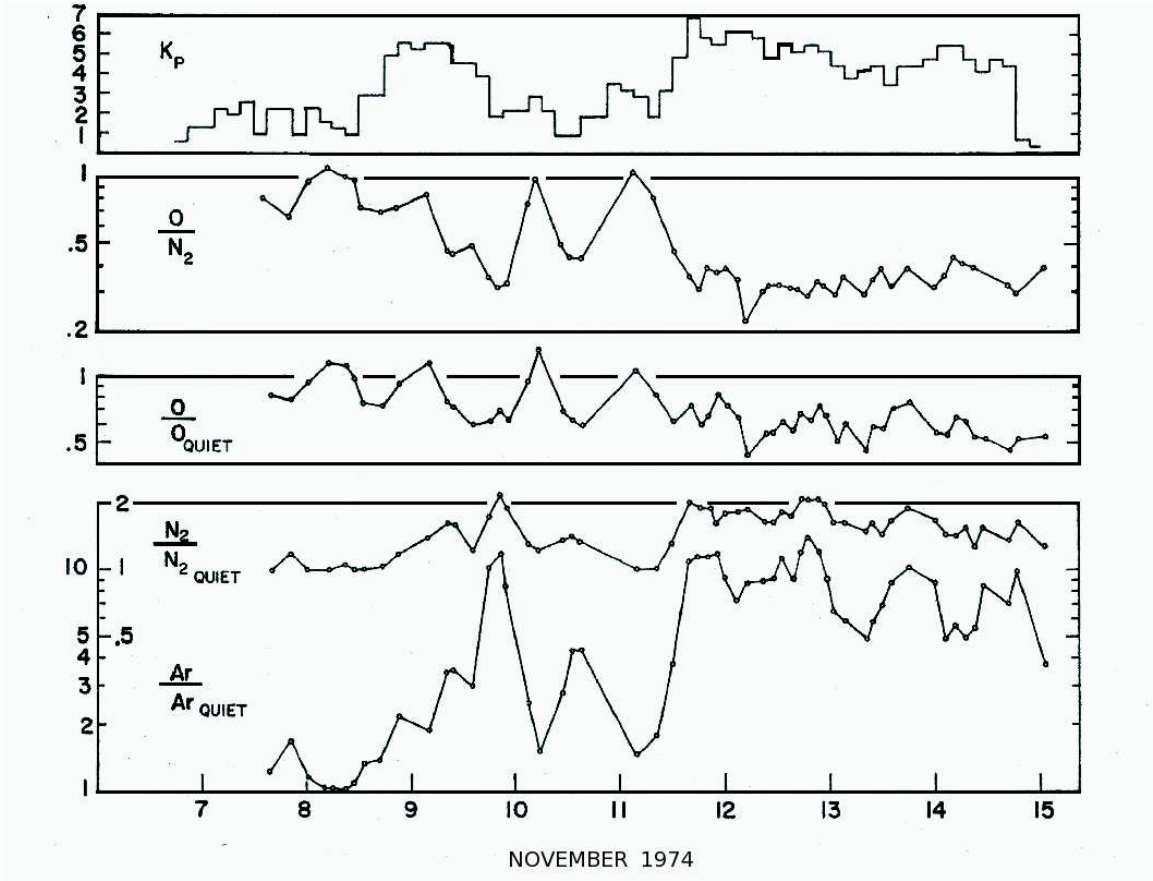

Figure 1.19: Composition changes at 160 km during the November 8-15, 1974 storm period. The evolution of two geomagnetic storms during the November 8-10 and November 11-15 1974 time periods is quite evident in the Kp index (top panel). The associated changes in the number densities of argon ($Ar/Ar_{\,quiet}$) and molecular nitrogen ($N_2/N_{2\,quiet}$) are presented in the bottom panel while the changes in atomic oxygen ($O/O_{\,quiet}$) and the O/N_2 ratio are presented in the middle two panels. The quiet-time number densities used in the determination of these ratios were obtained from five quiet-time orbits that preceded the storms. More details about the determination of the atomic oxygen number densities can be found in the text. (Reprinted from *Space Research*, Vol. 17, C. R. Philbrick, J. P. McIsaac, and G. A. Faucher, Variations in atmospheric composition and density during a geomagnetic storm, pp. 349-353, Copyright 1977, with permission from Elsevier.)

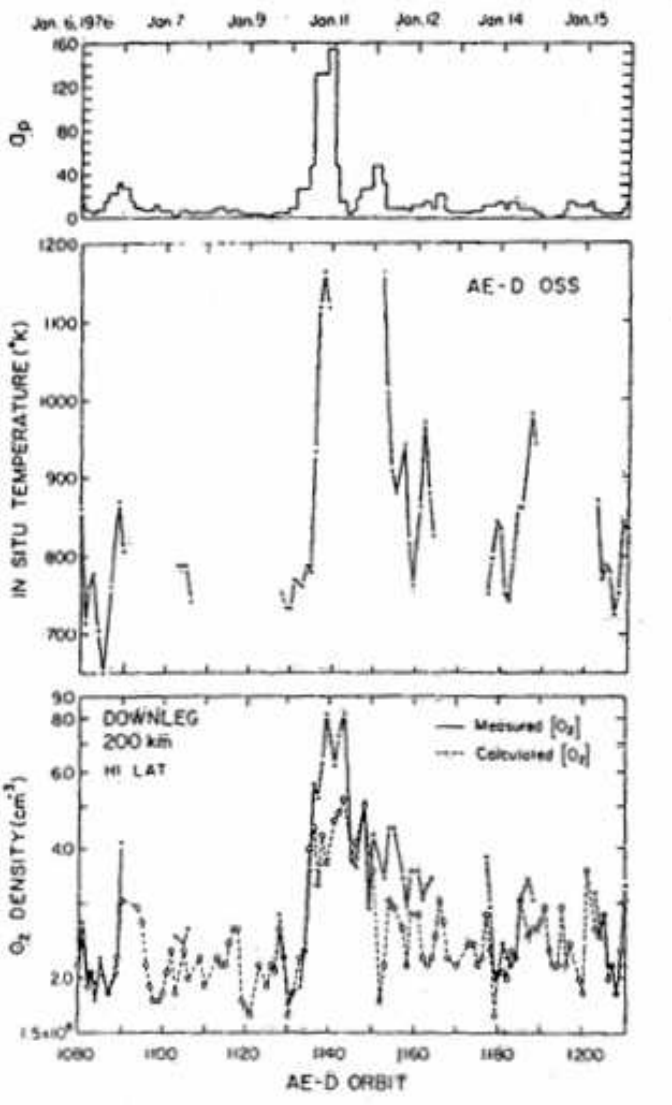

Figure 1.20: High latitude O_2 number densities at 200 km during the geomagnetic storm of January 11, 1974. The evolution of the storm is quite discernible in the a_p index (top panel). The neutral temperatures increased from 780 K to 1160 K during this storm (middle panel) while the "fly-through" mode number densities (solid line, lower panel) increased by a factor of 4. These values are two times larger than the diffusive equilibrium calculations (dotted line, lower panel), which suggests that transport by vertical winds may be playing an important role in the distribution of O_2 during this storm. (Reprinted from *Space Research*, Vol. 19, W. E. Potter, D. G. Kayser, and A. O. Nier, Thermospheric variations as an indicator of magnetic storm heating and circulation, pp. 259-262, Copyright 1979, with permission from Elsevier.)

mospheric emissions in the far ultraviolet (FUV) region. Nadir observations of the 130.4ñm triplet by photometers onboard the OGO-4 satellite [*Meier*, 1970; *Meier and Prinz*, 1971] showed that several depressions in this dayglow were associated with atomic oxygen depletions during geomagnetically active conditions. Subsequent observations of the 130.4ñm triplet by OGO-6 during geomagnetically disturbed conditions [*Strickland and Thomas*, 1976] confirmed the existence of these atomic oxygen depletions in the auroral zone and also showed that atomic oxygen enhancements could exist at lower latitudes. However, the very large line-centre opacity of the atmosphere at the centre of these resonance lines made any quantitative interpretation of the atomic oxygen concentrations in these depletions very difficult.

This issue was eventually circumvented by making observations of the FUV dayglow at several wavelengths; observations made by the S3-4 satellite at 130.4, 135.6, and 164.1 nm during the spring of 1978 [*Conway et al.*, 1988] showed that the *inferred* atomic oxygen column densities in these depletions were approximately 70% of their quiet-time values (the ratios of these emission lines allowed atomic oxygen column densities in the depletions to be deduced as the optical depths of these emission lines were quite different - see *Conway et al.*, 1988 for more details). Despite this success, the low orbital altitudes of these satellite missions meant that most observations of the FUV emission lines were confined to the orbital plane of the satellite, i.e. only a narrow range of latitudes or longitudes was scanned each orbit.

The deployment of spin-scan imaging photometers on the Dynamics Explorer 1 (DE-1) satellite mission produced the first *global scale* images of FUV emissions, allowing composition changes in the thermosphere to be observed simultaneously over a wide range of latitudes and longitudes [*Frank et al.*, 1981; *Frank and Craven*, 1988]. However, the relatively wide bandwidth of the filter used in these airglow observations meant that there were several emissions were present in the FUV images: the OI triplet at 130.4 nm, the OI doublet at 135.6 nm, and emissions from the N_2 Lyman-Birge-Hopfield (LBH) molecular bands. This complication caused the subsequent analysis of these images to be split into two categories: empirical [*Craven et al.*, 1994; *Nicholas et al.*, 1997; *Immel et al.*, 1997, 2000] and first principles modelling

[*Gladstone*, 1994; *Meier et al.*, 1995; *Strickland et al.*, 1999; *Drob et al.*, 1999; *Immel et al.*, 2001; *Strickland*, 2001; *Strickland et al.*, 2001].

The empirical studies were largely focused on quantifying the temporal and spatial distribution of the airglow depletions during geomagnetically disturbed conditions by comparing the FUV images obtained during geomagnetic disturbances to quiet time FUV images/climatologies. The most significant decreases in the FUV brightness ($30-40\%$) were normally observed equatorwards of the auroral oval after sustained periods ($\sim$ 6 hours) of intense geomagnetic activity (average AE greater than 700 nT). These decreases usually extended to the middle latitudes, with measurable decreases exisiting as far south as 30^o N. More moderate decreases ($15-30\%$) were also observed southwards of the auroral oval after prolonged periods of moderate geomagnetic activity (average AE $\sim$ 300-400 nT), but these decreases usually did not extend as far equatorwards. The spatial extent of these depletions also tended to be greater when the B_y component of the IMF was positive; the largest FUV depletion recorded by DE-1 was obtained during the geomagnetic disturbance of October 22, 1981 (Figure 1.21) with the depletion in emissions extending from the terminator in the morning sector to local noon and persisting continuously for several orbits.

Relating the intensity of the airglow image to the composition of the thermosphere was (and still remains) a formidable task as first principles calculations of the photoelectron and photon excited dayglow require a substantial amount of forward modelling. Early studies by *Gladstone* [1994] and *Meier et al.* [1995] were able to reproduce the salient features in the quiet time dayglow images acquired by DE-1, validating the photoelectron and multiple scattering codes used in these airglow models. Subsequent simulations with the airglow model used in the *Meier et al.* [1995] study suggested that the OI 135.6 nm/ N_2 LBH intensity ratio could be used to infer thermospheric composition changes through the O/N_2 column density ratio if the lower boundary in the column density evaluations was below all of the emission rate maxima [*Strickland et al.*, 1995]. The relatively wide bandwidth of the filter used in the DE-1 FUV observations precluded the use of this technique on DE-1 data, but *Strickland et al.* [1999] were able to demonstrate that O/N_2 column density ratios

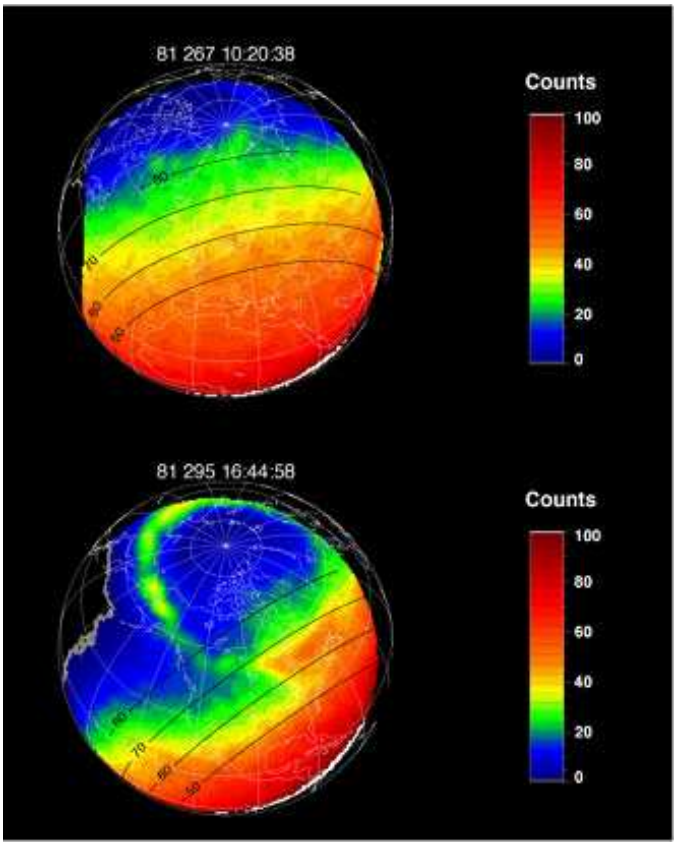

Figure 1.21: DE-1 images of the FUV dayglow during geomagnetically quiet and disturbed conditions in late 1981. The image of the FUV airglow in the upper panel was obtained on September 24, 1981 (Day 267) and shows that the brightness of the FUV dayglow during geomagnetically quiet conditions is primarily controlled by the solar zenith angle. Significant deviations from this pattern can occur during geomagnetically disturbed conditions; the geomagnetic disturbance on October 22, 1981 (Day 295) had a significantly brighter auroral oval and a large depletion in the sub-auroral zone. Contours of constant zenith angle at 50°, 60°, 70°, and 80° have been included on both panels. (From *Strickland et al.*, 1999.)

could still be inferred from DE-1 airglow images (Figure 1.22) if larger uncertainties in the column density ratios were tolerated. The availability of coincident DE-1 and DE-2 data from the solar proton event of July 13, 1982 provided an exemplary validation opportunity for this new technique; the overall consistency between the in-situ DE-2 column density ratios and the column density ratios deduced from the DE-1 airglow images showed that this technique was quite capable of inferring composition

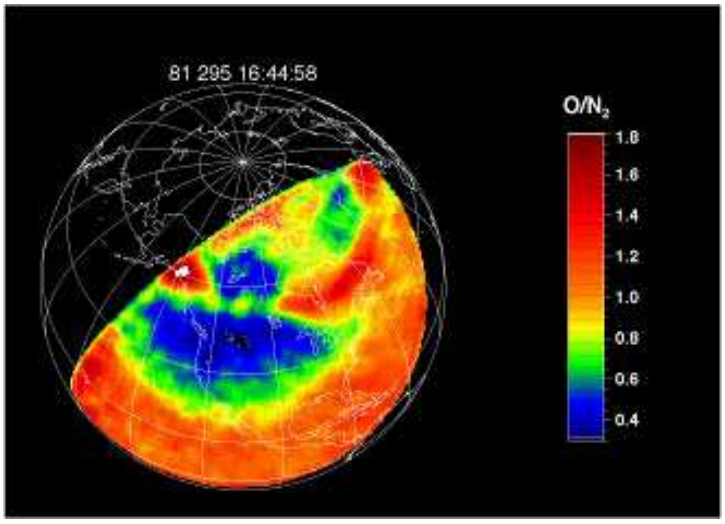

Figure 1.22: O/N_2 column density ratios derived from a DE-1 image of the FUV dayglow on October 22, 1981. The O/N_2 column ratios have been derived from the geomagnetically disturbed airglow image in the previous figure and shows that there is a substantial depletion in the atomic oxygen column density over North America, with the O/N_2 column density ratios being approximately 40% of their quiet time values. (From *Strickland et al.*, 1999.)

changes in the thermosphere. This success also allowed DE-1 derived composition changes to be incorporated into more detailed studies of selected geomagnetic storms from the 1981-1983 time period [*Immel et al.*, 2001; *Strickland*, 2001; *Strickland et al.*, 2001; *Immel et al.*, 2006].

The development of sensitive line scanning imaging spectrographs for satellite-borne FUV imagers, i.e. the Global Ultraviolet Imager (GUVI) on the Thermosphere Ionosphere Mesosphere Energetics and Dynamics (TIMED) satellite [*Christensen et al.*, 1994; *Paxton et al.*, 2004] and the Special Sensor Ultraviolet Spectrographic Imager (SSUSI) on the Defense Meteorological Support Program (DMSP) series of weather satellites [*Paxton et al.*, 1992], finally allowed simultaneous images of the FUV airglow to be acquired at high spectral resolution. However, the large amounts of data being generated by such high resolution observations would quickly exceed the telemetry bandwidth available to the satellite if all of it was transmitted to a ground

station; most GUVI studies focused on a few wavelength regions that provided all of the information necessary to infer thermospheric temperatures and composition: the Lyman-α doublet at 121.6 nm, the OI triplet at 130.4 nm, the OI doublet at 135.6 nm, and two broadband regions of the N_2 LBH spectrum (141.0-152.8 nm and 167.2-181.2 nm). The inclusion of the final three wavelength regions, i.e. the OI doublet at 135.6 nm and the two broadband regions of the N_2 LBH spectrum, allowed the *Strickland et al.* [1995] methodology to be used in the processing of GUVI composition data, producing accurate, high resolution maps of the O/N_2 column density ratio under a wide variety of geophysical conditions [*Christensen et al.*, 2003; *Strickland et al.*, 2004; *Zhang et al.*, 2004; *Meier et al.*, 2005; *Crowley et al.*, 2006; *Goncharenko et al.*, 2006; *Tsugawa et al.*, 2007; *Hecht et al.* 2008; *Stephan et al.*, 2008].

Early GUVI studies on thermospheric composition, e.g. *Christensen et al.* [2003] and *Strickland et al.* [2004], demonstrated that the inferred O/N_2 column density ratios were in very good agreement with empirical climatologies: the O/N_2 column density ratios exhibited a pronounced latitudinal gradient during solstice conditions (Figure 1.23) that reflected the latitudinal gradient present in atomic oxygen concentrations (thermospheric atomic oxygen number densities tend to be greater in the winter hemisphere and lower in the summer hemisphere during solstice condtions). Localized decreases in the O/N_2 column density ratios were also observed around the magnetic poles, with the decreases in the O/N_2 column density ratios reflecting the level of geomagnetic activity. Subsequent composition studies, e.g. *Zhang et al.* [2004], *Meier et al.* [2005], *Goncharenko et al.* [2006], and *Hecht et al.* [2008], were more focused on understanding the changes in the O/N_2 column density ratios during geomagnetic disturbances; large depletions in the O/N_2 column density ratios could extend over a wide range of longitudes, penetrate into the opposite hemisphere, and persist for extended periods of time (Figure 1.23).

Composition changes in the lower thermosphere have also been inferred from ground-based observations of bright auroral emission lines at high latitude observatories since the late 1980s [*Christensen et al.* 1997; *Hecht et al.* 1989, 1991, 1995,

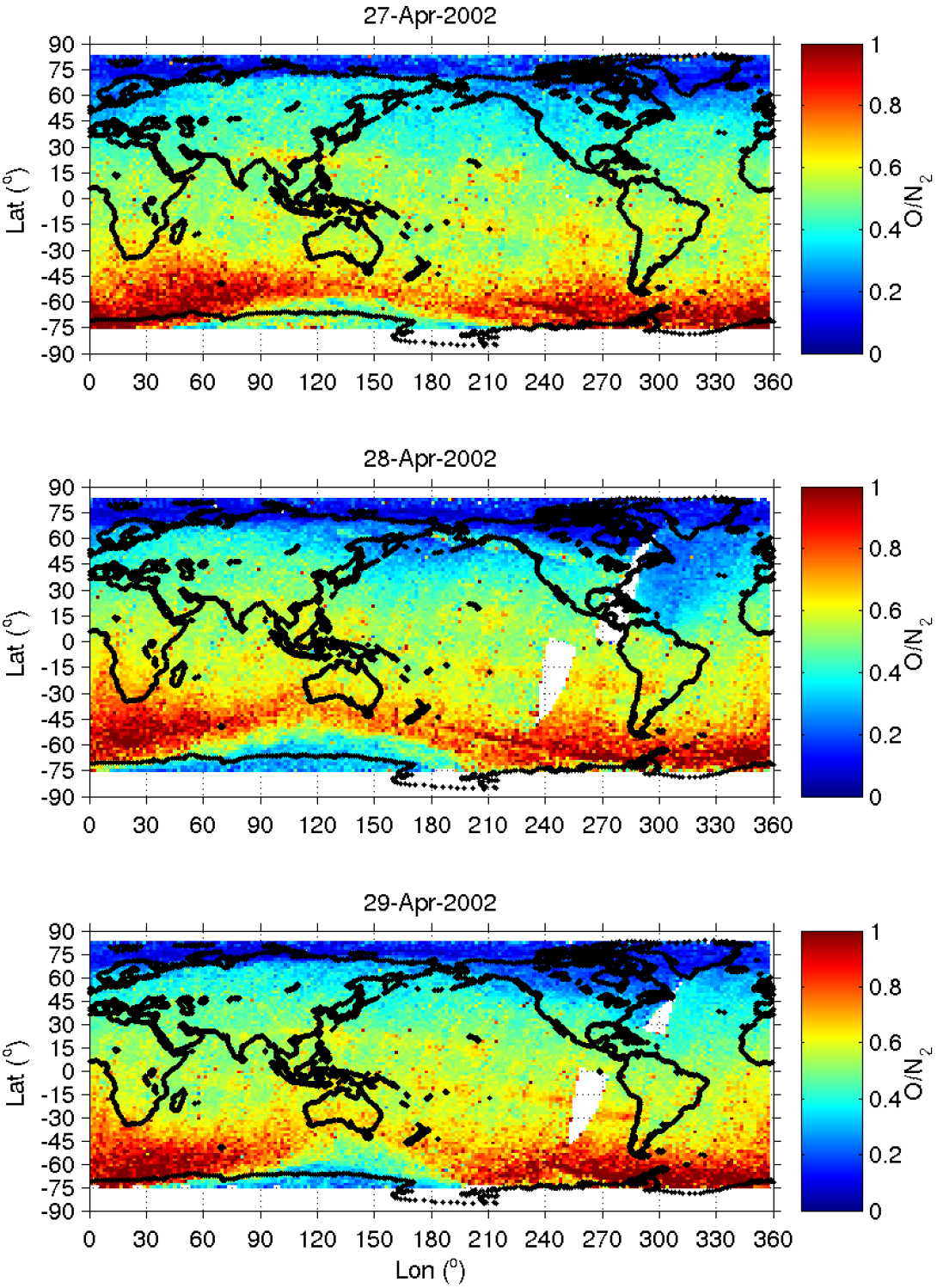

Figure 1.23: O/N_2 column density ratios derived from GUVI images of the airglow during the April 27-29, 2002 time period. There is a pronounced latitudinal gradient in all of the O/N_2 column density maps that reflects the gradients in atomic oxygen during solstice conditions. There is also a large decrease in the O/N_2 column density ratios during the geomagnetic disturbance of April 28 (middle panel) that extended all the way to the equator at 300°E. (From *Goncharenko et al.*, 2006.)

1999, 2000, 2006, 2008]. These changes in thermospheric composition were normally quantified by the forward modelling of the N_2^+ 427.8 nm (blue), OI 630.0 nm (red), and OI 844.6 nm brightness ratios: a model atmosphere and a Maxwellian electron energy distribution with energy flux Q and characteristic energy E_o were adjusted until the brightness ratios generated by the electron transport code of *Strickland et al.* [1989] were in agreement with the observed brightness ratios. This adjustment procedure was usually accomplished by varying three parameters: Q, E_o, and f_o, a dimensionless factor that was used to scale the atomic oxygen number density profile of the model atmosphere (this scaling of the atomic oxygen number density profile by a constant really should have been interpreted as a change in the atomic oxygen column density, but the usage of this terminology is now widespread). A typical plot of f_o as a function of time during moderate auroral activity shows that the inferred atomic oxygen column densities can decrease by 50% in a timespan of 30 minutes (Figure 1.24).

1.6 Previous Modelling Efforts

When the neutral density enhancements produced by geomagnetic disturbances were first observed, two very different mechanisms were proposed to explain how the energy from the solar wind was being transferred to the upper atmosphere:

i) hydromagnetic waves being generated by the interaction of the solar wind with the magnetosphere were propagating into the upper F-region, where the exponentially increasing neutral densities were causing the waves to dissipate [*Dessler*, 1959]

ii) large scale electric fields and high energy particle precipitation were heating the auroral thermosphere, which then transported the energy to lower latitudes through conduction and convection [*Cole*, 1962; 1966]

These widely disparate ideas led *Thomas and Ching* [1969] to develop a Green's function solution to the one-dimensional thermospheric heat flow equation, which allowed

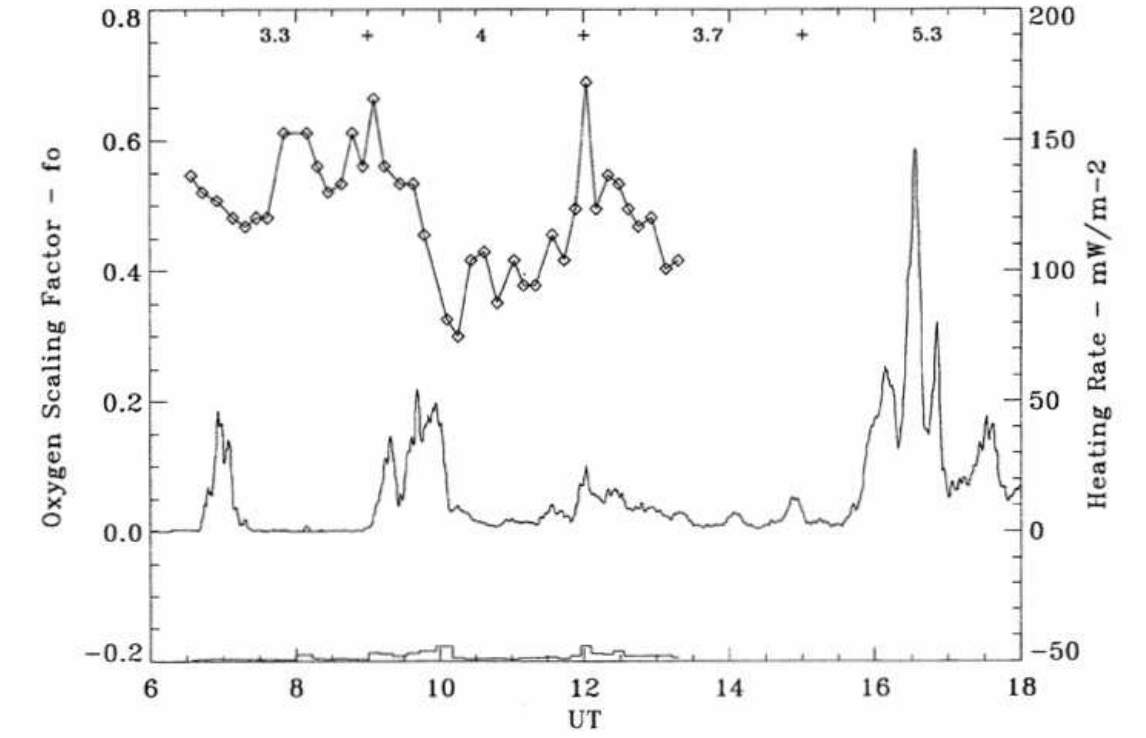

Figure 1.24: Inferred atomic oxygen column densities over Poker Flat, Alaska on February 13, 1994. The atomic oxygen scaling factor f_o (upper trace) is the constant that the atomic oxygen number density profiles have to be multiplied by to reproduce the observed brightness ratios. There was a significant decrease in the atomic oxygen column densities around 10 UT, right after a localized Joule heating event had occurred. (From *Christensen et al.*, 1997.)

them to study the dynamical response of the upper atmosphere to a wide variety of heat sources. By comparing the theoretically predicted density perturbations with a typical satellite drag inferred density perturbation profile, they were able to demonstrate that most of the energy must be deposited in the lower thermosphere, with the maximum of the energy deposition occuring in the 140-160 km region. The six hour time lag between the maximum of the geomagnetic disturbance and the maximum in the satellite drag derived density enhancements could also be reproduced quite readily if the heat source had a duration of 4-5 hours (the six hour time lag between the maximum of the geomagnetic disturbance and the maximum in the inferred den-

sity perturbations was the average time lag obtained from the satellite drag studies of *Roemer* [1966] and *Jacchia, Slowey and Verniani* [1967]). A subsequent study on the thermospheric response to impulsive heating events by *Volland* [1969] confirmed that "the heat input responsible for the geomagnetic activity effect occurs primarily in the lower thermosphere between 100 and 200 km altitude" and that the time delay between the geomagnetic disturbance maximum and the maximum in the inferred density enhancements was "a natural response of the thermosphere to a pulse-type disturbance."

The study by *Thomas and Ching* [1969] was also able to demonstrate that hydrostatic equilibrium was "remarkably valid for atmospheric changes occurring over time scales of the order of tens of minutes or greater", with the largest departures from hydrostatic equilibrium occurring above 300 km shortly after the heat source had reached its maximum value. This partitioning of the thermospheric response into two distinct phases did not gain widespread acceptance in the literature; the one-dimensional study of *Hays et al.* [1973] on thermospheric composition changes during disturbed conditions assumed that hydrostatic equilibrium was always valid and did not allow the altitudes of the pressure levels to change over the course of the simulation, i.e. the atmospheric pressure was kept constant, forcing any changes in the neutral temperature to be balanced by a change in the thermospheric density. These two assumptions substantially reduced the simulated response, limiting the vertical wind speeds in the model to a few metres per second and the changes in the N_2/O_2 and N_2/O ratios at 500 km to factors of 2 and 10 respectively. However, the simulated N_2/O ratios were in excellent agreement with OGO-6 measurements, allowing these simulation results to gain widespread acceptance in the literature (this agreement was quite fortuitous as OGO-6 measurements of atomic oxygen were always derived from the total oxygen content, which assumed that any measured molecular oxygen was due to the recombination of ambient atomic oxygen on ion source surfaces).

A subsequent study by *Bates* [1974] argued that the thermospheric response at the onset of widespread Joule heating would occur primarily at constant pressure, i.e. the thermosphere was free to expand upwards, but the pressure forces acting on

individual fluid elements was not allowed to change. Such an expansion would alter the composition and dynamics of the upper thermosphere in a matter of minutes: neutral densities would double, N_2/O ratios would triple, vertical wind speeds in excess of 100 m/s would exist for a few minutes, and neutral temperatures would increase by several hundred degrees. Bates also argued that the solution presented in the *Hays et al.* [1973] study was

> ... a steady state solution for which the maximum amount of input energy goes into horizontal expansion. Under these conditions, vertical expansion is obviously minimal, and that conclusion is verified by the computations of [*Hays et al.* [1973]]

while

> ... the solution presented in this paper is a dynamical solution for which the maximum possible amount of energy goes into vertical expansion ... after the onset of widespread Joule heating, the solution presented here will be much closer to reality than the steady state solution of [*Hays et al.* [1973]]. At later times, however, the converse will be true ...

Unfortuneately, these results also failed to gain widespread acceptance in the literature as most subsequent studies on the response of the thermosphere to large energy inputs have assumed that the thermospheric response is hydrostatic; any studies that have allowed for non-hydrostatic effects did not examine the inital response of the thermosphere in any detail.

The first two-dimensional study on the thermospheric response to enhanced heating in the auroral zones was undertaken by *Volland and Mayr* [1971], who solved the linearized Navier-Stokes equations using a transfer function approach, i.e. the 'system transfer function' coefficients obtained from the solution to the linearized Navier-Stokes equations were combined with the spherical harmonic coefficients of the enhanced auroral zone heat source to yield spherical harmonic coefficients for the perturbations produced by the enhanced heating rates. They found that the smaller

wavenumber (longer wavelength) components of the density response were able to reproduce most of the salient features in the available satellite drag data: the density perturbations in the auroral zone of the model were two times larger than the density perturbations in the equatorial region of the model and the maximum in the auroral zone density perturbations preceded the maximum in the equatorial region density perturbations by 3.5 hours. The rapid attenuation of the larger wavenumber (shorter wavelength) phenomena in the density response, on the other hand, was not in accord with observations: the simplified treatment of the ion-drag and viscous terms in the linearized Navier-Stokes equations did not provide enough dissipation at higher frequencies, forcing the authors to suppress these frequencies in the system transfer function coefficients. This additional damping ensured that the higher frequency, shorter wavelength components of the density response were being dissipated, but it also precluded any quantitative analysis of these features in the density response as well.

The inability of the *Volland and Mayr* [1971] model to simulate shorter wavelength phenomena at the onset of a geomagnetic disturbance was not considered to be a major issue in its subsequent development as the elimination of the larger wavenumber components in the system transfer function allowed the authors to incorporate a more realistic composition scheme into the model, i.e. the inclusion of the horizontal and vertical transport terms in all of the species' continuity equations allowed the thermospheric number densities to deviate from diffusive equilibrium. Simulations employing this new composition scheme [*Mayr and Volland*, 1972, 1973, 1976; *Mayr and Hedin*, 1977; *Mayr and Trinks*, 1977; *Hedin et al.*, 1977] predicted that the number densities of argon and molecular nitrogen would be enhanced at all altitudes in the auroral thermosphere while the number densities of helium would be depleted throughout the auroral thermosphere. Atomic oxygen concentrations, on the other hand, were expected to exhibit an altitude-dependent behaviour; the atomic oxygen number densities in the lower auroral thermosphere were predicted to decrease (since vertical transport was the dominant process in this region) and increase slightly at higher altitudes (where thermal expansion was the more important process). Outside

of the auroral zone, all of the thermospheric species were expected to exhibit minor increases as the meridional wind system being generated by the enhanced heating rates would transport all of these species equatorwards.

Most of these predictions were in very good agreement with mass spectrometer observations; a simulation of the moderate geomagnetic storm on February 12, 1974 (max $a_p \approx 50$) was able to reproduce the observed changes in the molecular nitrogen (N_2) and helium (He) number densities at different altitudes and latitudes (Figure 1.25). The simulated changes in the atomic oxygen (O) number densities were also in fairly good agreement with measurements at lower latitudes, but the predicted increases in the upper auroral thermosphere were not observed; the atomic oxygen number densities being measured by the Neutral Atmosphere Composition Experiment (a closed ion source mass spectrometer) were actually much lower than the values preceding the storm. This discrepancy between the simulated and measured atomic oxygen density ratios in the upper auroral thermosphere was not a very important issue in this study as most mass spectrometer studies during this time period were still trying to quantify the effects of geomagnetic storms on thermospheric composition. Additional studies with this model in the 1980s [*Mayr et al.*, 1984a, b, 1987, 1990] employed a system transfer function with a much larger range of wavenumbers, which allowed the authors to study gravity wave generation, propagation, and dissipation in the auroral regions.

The ability of the system transfer function approach to reproduce most of the observed composition changes during moderate geomagnetic disturbances could not be extended to more disturbed conditions as the linearization procedures employed in this approach were only valid when the wind speeds in the simulation were much less than the local sound speed. The computational resources needed to solve the two-dimensional Navier-Stokes equations during very disturbed conditions would not be available to upper atmospheric modellers for several more years, forcing most of these modellers to assume that any vertical motions in the regions of enhanced heating were still hydrostatically balanced, i.e. the pressure gradient and the force of gravity were the only forces acting in the vertical direction. This assumption ensured that

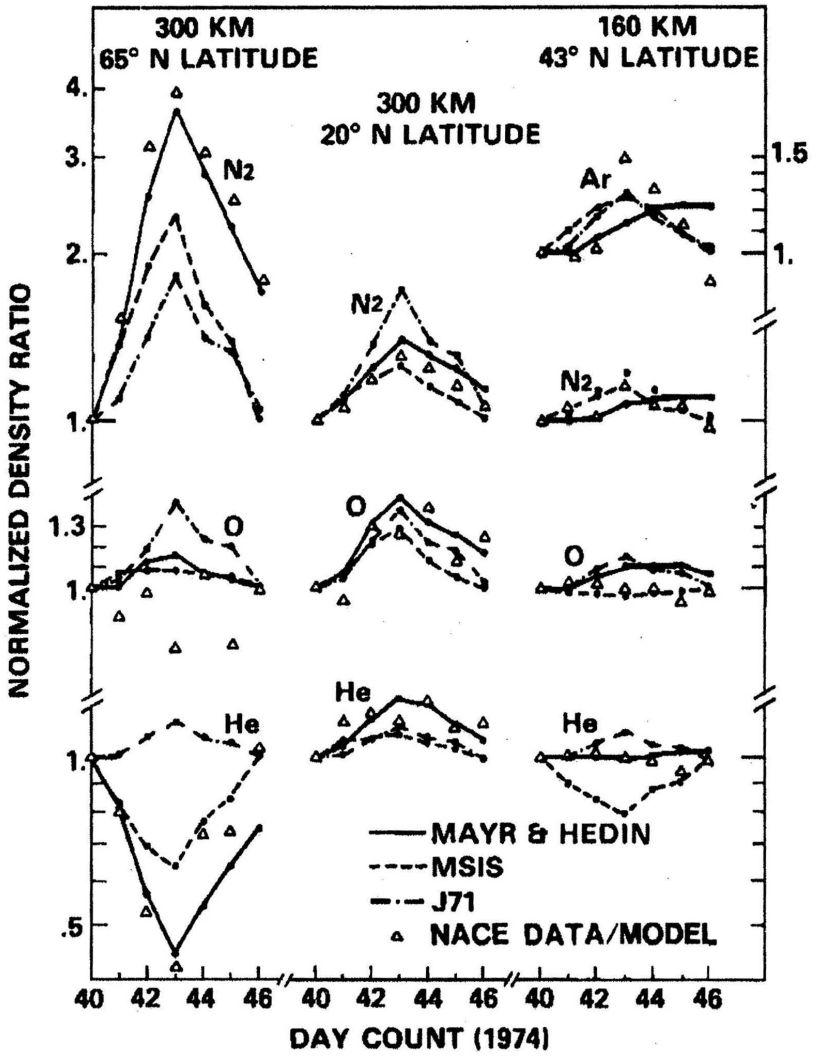

Figure 1.25: Daily average number density ratios during the February 9-15, 1974 time period. The simulated number density ratios (solid lines) are in fairly good agreement with the measured number density ratios (triangles) when one considers non-reactive species at higher altitudes, i.e. molecular nitrogen (N_2) and helium (He) at 65°N, 300 km and 20°N, 300 km. The simulated atomic oxygen ratios are also in fairly good agreement with the measured ratios outside of the auroral zone, but the moderate enhancements seen in the auroral zone of are model are not consistent with the large depletions being measured by the Neutral Atmosphere Composition Experiment on the AE-C satellite. Number density ratios from the MSIS (dashed line) and Jacchia 71 (dot-dash line) empirical models have also been included for comparison purposes. All of the number density ratios were normalized to their pre-storm values on February 9 (day 40). (From *Hedin et al.*, 1977.)

the pressure field in these simulations would always be a monotonically decreasing function of height, allowing a pressure coordinate system to be employed. Besides simplifying the evaluation of the vertical velocity and pressure fields, the assumption of hydrostatic equilibrum also eliminated vertically propagating sound waves from the thermospheric response, permitting explicit time stepping methods to use a much larger time step.

These simplifications, when combined with the rapidly expanding database of ion density and electric field measurements in the early 1970s, e.g. *Heppner* [1972], *Banks et al.* [1973], and *Bates and Hunsucker* [1974], allowed *Richmond and Matsushita* [1975] to simulate the thermospheric response to a large, isolated substorm. They found that the vertical and meridional winds in the auroral zone of the model were an order of magnitude larger than the transfer function results: the vertical winds were now able to reach speeds of 40 m/sec at the peak of the Joule heating rates and the meridional winds were flowing away from the region of enhanced heating at speeds of 100 m/sec. The zonal winds in the auroral oval were also much larger, approaching speeds of 400 m/sec at 400 km. An analysis of these wind fields revealed that the nonlinear advection terms were playing an important role in the evolution of the zonal winds; the vertical and meridional wind fields were transporting significant amounts of momentum upwards and equatorwards, reducing the ion drag forcing in the auroral oval through most of the simulation. Such large wind speeds would have also generated substantial composition changes in the auroral regions but the mixing ratios of the individual species were not allowed to vary during the simulation.

Self-consistent thermosphere-ionosphere general circulation models (TIGCMs) are also able to reproduce these inferred composition changes and have become the defacto standard in modelling composition changes.

This thesis is composed of five chapters. Chapter 2 provides the prerequisite background and parameterizations for the physical processes implemented in the model. A concise overview of the model numerics is then presented in Chapter 3, along with a brief discussion on suggested numerical upgrades. The results obtained from the various versions of the model will be presented in Chapter 4 and the physics of these

processes discussed. Chapter 5 will provide brief summary of the main results of this thesis and suggestions for future work.

Chapter 2

Model Description I : Physics

2.1 Governing Equations

If one assumes that the atmosphere can be represented by an ideal, mean mass gas, the time-dependent Navier-Stokes equations for a flow that is viscous, compressible, and thermally conducting can be written as

$$\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \mathbf{u_n}) = 0 \tag{2.1}$$

$$\frac{\partial \rho\, \mathbf{u_n}}{\partial t} + \nabla \cdot [\rho\, \mathbf{u_n}\, \mathbf{u_n} + p\mathbf{I} - \tau] = \rho \mathbf{f} \tag{2.2}$$

$$\frac{\partial e}{\partial t} + \nabla \cdot [(e+p)\mathbf{u_n} - \mathbf{u_n} \cdot \tau + \mathbf{q}] = \rho(\mathbf{f} \cdot \mathbf{u_n}) + Q - C \tag{2.3}$$

[*Oran and Boris*, 2001] where $\mathbf{I}$ is the unit tensor, τ is the viscous stress tensor, $\mathbf{f}$ is the force per unit volume, $\mathbf{q}$ is the heat flow vector, Q is the heating rate, and C is the cooling rate due to radiative emissions. The viscous stress tensor is given by the expression

$$\tau = 2\mu\mathbf{D} - \frac{2}{3}\mu(\nabla \cdot \mathbf{u_n})\mathbf{I} \tag{2.4}$$

$$= \mu\,[\,\nabla \mathbf{u_n} + (\nabla \mathbf{u_n})^T\,] - \frac{2}{3}\mu(\nabla \cdot \mathbf{u_n})\mathbf{I} \tag{2.5}$$

where $\mathbf{D}$ is the deformation tensor

$$\mathbf{D} = \frac{1}{2}\,[\,\nabla \mathbf{u_n} + (\nabla \mathbf{u_n})^T\,] \tag{2.6}$$

It should be noted that this definition of the deformation tensor is not unique and one must be aware of which form of the deformation tensor is being employed. The heat flow is given by

$$\mathbf{q} = -\lambda \nabla T. \tag{2.7}$$

where λ is the thermal conductivity. These are the equations that were originally used in the model.

If one now assumes that the atmosphere is a 'true' multiple component gas that is viscous, compressible, and thermally conducting, the time-dependent Navier-Stokes equations become

$$\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \mathbf{u_n}) = 0 \tag{2.8}$$

$$\frac{\partial n_s}{\partial t} + \nabla \cdot (n_s \mathbf{u_n}) + \nabla \cdot (n_s \mathbf{u}_{ds}) = P_s - L_s \tag{2.9}$$

$$\frac{\partial \rho \mathbf{u_n}}{\partial t} + \nabla \cdot [\rho \mathbf{u_n} \mathbf{u_n} + p\mathbf{I} - \tau] = \rho \mathbf{f} \tag{2.10}$$

$$\frac{\partial e}{\partial t} + \nabla \cdot [(e+p)\mathbf{u_n} - \mathbf{u_n} \cdot \tau + \mathbf{q}] = \rho(\mathbf{f} \cdot \mathbf{u_n}) + Q - C \tag{2.11}$$

[*Oran and Boris*, 2001] where n_s is the number density of species s, $P_s - L_s$ is the net production rate of species s, and $\mathbf{u}_{ds}$ is the diffusion velocity of species s. The heat flow is now written as

$$\mathbf{q} = -\lambda \nabla T + \sum_s \rho_s h_s \mathbf{u}_{ds} + p \sum_s \alpha_s \mathbf{u}_{ds} \tag{2.12}$$

$$= -\lambda \nabla T + T \sum_s \rho_s c_{ps} \mathbf{u}_{ds} + p \sum_s \alpha_s \mathbf{u}_{ds} \tag{2.13}$$

where h_s is the enthalpy of species s and α_s is the thermal diffusion coefficient of species s. However, the thermal diffusion contribution to the heat flow is usually ignored in most atmospheric simulations, allowing one to simplify the heat flow expression to

$$\mathbf{q} = -\lambda \nabla T + T \sum_s \rho_s c_{ps} \mathbf{u}_{ds}. \tag{2.14}$$

The diffusion velocities $\mathbf{u}_{ds}$ are found by inverting the matrix equation

$$\sum_k \frac{n_s n_k}{n^2 D_{sk}} (\mathbf{u}_{dk} - \mathbf{u}_{ds}) = \mathbf{G}_s \tag{2.15}$$

with the source terms $\mathbf{G}_s$ defined as

$$\mathbf{G}_s = \nabla \left(\frac{n_s}{n}\right) - \left(\left(\frac{\rho_s}{\rho}\right) - \left(\frac{n_s}{n}\right)\right)\left(\frac{\nabla p}{p}\right) + \alpha_s \left(\frac{\nabla T}{T}\right). \tag{2.16}$$

These diffusion velocities $\mathbf{u_{ds}}$ are subject to the constraint that

$$\sum_s \rho_s \mathbf{u}_{ds} = 0 \tag{2.17}$$

which ensures that no net mass flux can be generated by interspecies diffusion. In thermospheric simulations, the horizontal components of the diffusion velocities are frequently ignored and hydrostatic equilibrium is often assumed, which allows one rewrite (2.15) and (2.16) as

$$\sum_k \frac{n_s n_k}{n^2 D_{sk}} (\mathbf{u}_{dk} - \mathbf{u}_{ds}) = \left[\nabla_v n_s + (1 + \alpha_s) \left(\frac{n_s}{T} \right) \nabla_v T + \frac{n_s m_s g}{kT} \right] \tag{2.18}$$

where ∇_v is the vertical component of the gradient operator in the coordinate employed by the model.

2.2 UV/EUV Heating rates

The original version of the model utilized a UV/EUV heating rate profile that was representative of average winter daytime conditions [*Chang and St.-Maurice*, 1991]. This profile was always kept constant throughout the simulation as the original runs of the model were not concerned with compositional feedbacks; they were only interested in having the UV/EUV heating rate provide a realistic background temperature profile for the thermosphere. The current/new emphasis on trying to incorporate compositional feedbacks into the model has dictated that a more realistic UV/EUV heating rate profile would be required. The processes and rate coefficients used in the construction of this new neutral gas heating rate are based on the heating rates and rate coefficients used in the 1995 version of the globally averaged NCAR TIME-GCM [*Roble*, 1995].

2.2.1 Schumann-Runge Heating Rates

The heating rates in the Schumann-Runge continuum (SRC) of molecular oxygen were calculated using scheme of *Zhu* [1994], which is essentially the parameterization of *Strobel* [1978] with updated cross-sections and fluxes (the updated cross-sections and

fluxes are the values recommended by the World Meteorological Organization [1985] and Nicolet [1989] respectively). In these parameterizations, the SRC is split into two regions: SRC1 (125 - 152 nm) and SRC2 (152 - 175 nm). The absorption cross-section of O_2 in the SRC1 region is essentially constant, allowing one to parameterize the heating rate in this region as

$$Q_{\mathrm{SRC1}} \;=\; \epsilon_{\mathrm{SRC}}\, F_{\mathrm{SRC1}}\, \sigma_{\mathrm{SRC1}}\, \exp\left(-\sigma_{\mathrm{SRC1}}\, \mathrm{N}_{\mathrm{O}_2}\right) \times n_{\mathrm{O}_2} \tag{2.19}$$

where σ_{SRC1} is the average absorption cross-section of O_2 in this region ($1.1{\times}10^{-17}\mathrm{cm}^2$), N_{O_2} is the total column abundance of O_2 along the solar radiation path, and $\epsilon_{\mathrm{SRC}}\, F_{\mathrm{SRC1}}$ is the product of the SRC efficiency factor and the integrated flux in the SRC1 region (0.65 ergs cm^{-2} sec^{-1}). In the SRC2 region, the absorption cross-section of O_2 tends to decrease exponentially as the wavelength increases, i.e., $\sigma_{\mathrm{SRC2}} \propto \exp(-M\lambda)$, allowing one to decompose the heating rate in this region into 2 subintervals, i.e.

$$\begin{aligned} Q_{\mathrm{SRC2}} \;=\; & \epsilon_{\mathrm{SRC}} \left[\left(\frac{I_{\mathrm{SRC2},l}}{M} \right) \exp(-\sigma_{\mathrm{SRC2},l}\, N_{\mathrm{O}_2}) \right. \\ & + \left(\frac{I_{\mathrm{SRC2},s} - I_{\mathrm{SRC2},l}}{M} \right) \exp(-\sigma_{\mathrm{SRC2},m}\, N_{\mathrm{O}_2}) \\ & \left. - \left(\frac{I_{\mathrm{SRC2,s}}}{M} \right) \exp(-\sigma_{\mathrm{SRC2,s}}\, N_{\mathrm{O}_2}) \right] \times \frac{n_{\mathrm{O}_2}}{N_{\mathrm{O}_2}} \end{aligned} \tag{2.20}$$

where l refers to the long-wavelength end of the SRC2 interval, s refers to the short-wavelength end of the SRC2 interval, and m refers to the wavelength where the 2 subintervals meet (166 nm). The absorption cross-sections of O_2 at 175, 166, and 152 nm are $3.0 \times 10^{-19}\mathrm{cm}^2$, $2.0 \times 10^{-18}\mathrm{cm}^2$, and $1.5 \times 10^{-17}\mathrm{cm}^2$ respectively while the values of $\epsilon_{\mathrm{SRC}}\, I_{\mathrm{SRC2},l}/M$ and $\epsilon_{\mathrm{SRC}}\, I_{\mathrm{SRC2},s}/M$ are 0.12 ergs cm^{-2} sec^{-1} and 0.04 ergs cm^{-2} sec^{-1} respectively. The heating rates in the SRC1 and SRC2 regions are then added together to obtain the net heating rate for the SRC region.

The heating rate in the Schumann-Runge bands (SRB) is much more difficult to parameterize as the absorption cross-section of O_2 exhibits significant variability in this interval. Most methods that do attempt to parameterize the heating rates in the SRB region divide the SRB into smaller intervals, allowing the transmittances in

these intervals to be more easily determined. The heating rates in the SRB are then derived from these individual transmittances.

In Strobel's [1978] parameterization, the SRB heating rate is calculated by summing the heating rate in each 500 cm^{-1} interval and then fitting this sum to an appropriately chosen expression, i.e.,

$$\begin{aligned} Q_{\mathrm{SRB}} &= \sum_{i=1}^{19} F_{\mathrm{SRB},i}\ \sigma_{\mathrm{SRB},i}\, Tr_i\ n_{\mathrm{O}_2} \\ &= \sum_{i=1}^{19} F_{\mathrm{SRB},i} \left(\frac{1}{\alpha_i + \beta_i N_{\mathrm{O}_2}^{1/2}} \right) Tr_i\, n_{\mathrm{O}_2} \\ &= \frac{n_{\mathrm{O}_2}}{(aN_{\mathrm{O}_2} + bN_{\mathrm{O}_2}^{1/2})} \end{aligned} \tag{2.21}$$

where $F_{\mathrm{SRB},i}$ is the integrated solar flux in interval i, $\sigma_{\mathrm{SRB},i}$ is the absorption cross-section of O_2 in interval i, Tr_i is the transmittance in interval i, and α_i and β_i are the coefficients that determine the absorption cross-section of O_2 in interval i. The values of a and b are then computed from the values of $Q_{\mathrm{SRB}}/n_{\mathrm{O}_2}$ and N_{O_2}. Zhu [1994] has recently developed a more sophisticated SRB parameterization that fits a Malkmus model [*Malkmus*, 1967] to the atmospheric transmittance, i.e.

$$\tau = \exp\left\{ -\frac{\pi y}{2} \left[\left(1 + \frac{4\overline{k}}{\pi y} N_{\mathrm{O}_2} \right)^{1/2} - 1 \right] \right\} \tag{2.22}$$

and then determines the SRB heating rate as

$$Q_{\mathrm{SRB}} = \frac{\epsilon\, F_{\mathrm{SRB}}\, \sigma_{\mathrm{SRB}}\, n_{\mathrm{O}_2}}{\left(1 + \frac{4\sigma_{\mathrm{SRB}}}{\pi y_{\mathrm{SRB}}} N_{\mathrm{O}_2} \right)^{1/2}} \exp\left\{ -\frac{\pi y_{\mathrm{SRB}}}{2} \left[\left(1 + \frac{4\sigma_{\mathrm{SRB}}}{\pi y_{\mathrm{SRB}}} N_{\mathrm{O}_2} \right)^{1/2} - 1 \quad \right] \right\}. \tag{2.23}$$

In the Malkus model of atmospheric transmittance, $\overline{k}$ is the mean band strength and y is a constant that relates the mean line half-width to the mean line spacing. In the SRB heating rate expression, $\epsilon\, F_{\mathrm{SRB}}$ has a value of 1.28 ergs cm^{-2} sec^{-1}, σ_{SRB} has a value of 2.07×10^{-20} cm^2, and y_{SRB} has a value of 0.0152. The SRB parameterization of Zhu [1994] is currently being used in the model.

2.2.2 Ozone Heating Rates

The heating rates due to the absorption of radiation in the Chappius, Hartley, and Huggins bands of ozone are also calculated using Zhu's [1994] parameterization. The

absorption cross-sections of O_3 in the Chappius and Hartley bands are essentially constant, allowing one to approximate these heating rates as

$$Q_{\mathrm{Ch}} \;=\; F_{\mathrm{Ch}}\;\sigma_{\mathrm{Ch}}\;\exp\left(\,-\sigma_{\mathrm{Ch}}\,N_{\mathrm{O_3}}\right)\;\times\; n_{\mathrm{O_3}} \tag{2.24}$$

$$Q_{\mathrm{Ha}} \;=\; \epsilon_{\mathrm{Ha}}\,F_{\mathrm{Ha}}\,\sigma_{\mathrm{Ha}}\;\exp\left(\,-\sigma_{\mathrm{Ha}}\,N_{\mathrm{O_3}}\right)\;\times\; n_{\mathrm{O_3}} \tag{2.25}$$

where F_{Ch} and F_{Ha} are the integrated fluxes in the Chappius and Hartley bands ($3.7\ \times 10^5$ ergs cm^{-2} sec^{-1} and 5130 ergs cm^{-2} sec^{-1} respectively), σ_{Ch} and σ_{Ha} are the average absorption cross-sections of O_3 in these regions (2.85×10^{-21} cm^2 and 8.7×10^{-18} cm^2 respectively), $N_{\mathrm{O_3}}$ is the column density of O_3 along the solar radiation path, and ϵ_{Ha} is the heating efficency in the Hartley band. This heating efficiency is determined by the polynomial

$$\epsilon_{\mathrm{Ha}} = \mathrm{c_0} + \mathrm{c_1 x} + \mathrm{c_2 x^2} + \mathrm{c_3 x^3} \tag{2.26}$$

[Mlynczak and Solomon, 1993], where x is

$$\mathrm{x} = \log 10\,(\,p\,) + 3, \qquad 10^{-4}\,\mathrm{mb} \;<\; p \;<\; 10^{-2}\,\mathrm{mb} \tag{2.27}$$

and the coefficients c_0, c_1, c_2, and c_3 are 0.669650, -0.009682, 0.033093, and -0.017938 respectively. The heating rate in the Huggins band of O_3 can be parameterized as

$$\begin{aligned} Q_{\mathrm{Hu}} \;\;=\;\; & \frac{1}{M N_{\mathrm{O_3}}}[I_1 + (I_2 - I_1)\exp(-\sigma_{\mathrm{Hu}} N_{\mathrm{O_3}} e^{-M\lambda_{long}}) \\ & -I_2 \exp(-\sigma_{\mathrm{Hu}} N_{\mathrm{O_3}} e^{-M\lambda_{short}})] \times n_{\mathrm{O_3}} \end{aligned} \tag{2.28}$$

where I_1 is 70 ergs cm^{-2} sec^{-1} Å^{-1}, I_2 is 50 ergs cm^{-2} sec^{-1}Å^{-1}, M = 0.0127 Å, λ_{short} is 2805 Å, λ_{long} is 3015 Å, and σ_{Hu} is 0.0115 cm^2. It should be noted that all of Zhu's [1994] ozone heating parameterizations are based on Strobel's [1978] parameterizations with updated cross-sections and fluxes, just as in case of the SRC parameterization.

2.2.3 Exothermic Chemical Heating Rates

The heating by exothermic chemical reactions is usually broken down into four categories: heating by neutral oxygen/hydrogen chemistry, heating by neutral nitrogen chemistry, heating by the quenching of O(^{1}D) and heating by ion chemistry.

The heating rate due to exothermic chemical reactions involving hydrogen and oxygen is only important in the mesosphere and lower thermosphere of the model. A summary of the reactions used in the determination of this heating rate can be found in Table 2.1. Since all of the products in these reactions are neutral molecules, the amount of energy being released in a particular reaction can be determined by multiplying the reactant number densities, the reaction rate, and the amount of energy released together. As an example, consider the reaction

$$H + O_3 \xrightarrow{\beta_{14}} OH + O_2 + 3.34\,eV$$

The total heating rate (in ergs cm^{-3} sec^{-1}) due to this chemical reaction will be

$$Q_{H,O_3} = \epsilon_{14} \times \beta_{14} \times n_H \times n_{O_3} \times (3.34 \,/\, 6.2415 \times 10^{11})$$

where ϵ_{14} is the heating efficiency of the reaction and 6.2415×10^{11} is the conversion factor from eV to ergs. The neutral gas heating rate due to hydrogen/oxygen chemistry is then calculated by summing the heating rates of the individual reactions in this category. It should be noted that the heating efficiency of the preceding reaction is 0.6; all of the other reactions in this group have an efficiency of one.

The neutral gas heating rates due to the quenching of $O(^1D)$ and exothermic nitrogen chemistry are important in the lower and middle regions of the thermosphere. A summary of the reactions used in the determination of these heating rates can be found in Tables 2.2 and 2.3. Since all of the products in these reactions are neutral molecules, these heating rates can also be determined by multiplying the reactant number densities, the reaction rates, and the amount of energy released together. This procedure is very similar to the procedure used in the determination of the oxygen/hydrogen chemistry heating rates - the only difference is that all of the reactions in these two categories have a unit heating efficiency.

The heating caused by exothermic ion chemistry is also important in the lower and middle regions of the thermosphere. A summary of the reactions used in the determination of this heating rate can be found in Tables 2.4 and 2.5. The heating

Reaction			Reaction Rate
$O + O + M$	$\xrightarrow{\beta_1}$	$O_2 + M + 5.12$ eV	$\beta_1 = 9.59 \times 10^{-34} \exp(\frac{480}{T})$
$O + O_2 + M$	$\xrightarrow{\beta_2}$	$O_3 + M + 1.10$ eV	$\beta_2 = 6 \times 10^{-34} (\frac{300}{T})^{2.8}$
$O + O_3$	$\xrightarrow{\beta_3}$	$O_2 + O_2 + 4.06$ eV	$\beta_3 = 8 \times 10^{-12} \exp(-\frac{2060}{T})$
$O + OH$	$\xrightarrow{\beta_4}$	$O_2 + H + 0.72$ eV	$\beta_4 = 2.2 \times 10^{-11} \exp(\frac{117}{T})$
$O + HO_2$	$\xrightarrow{\beta_5}$	$OH + O_2 + 2.33$ eV	$\beta_5 = 3.0 \times 10^{-11} \exp(\frac{200}{T})$
$O + H_2$	$\xrightarrow{\beta_6}$	$OH + H + 0.08$ eV	$\beta_6 = 1.6 \times 10^{-11} \exp(-\frac{4570}{T})$
$OH + O_3$	$\xrightarrow{\beta_7}$	$HO_2 + O_2 + 1.73$ eV	$\beta_7 = 1.6 \times 10^{-12} \exp(-\frac{940}{T})$
$OH + OH$	$\xrightarrow{\beta_8}$	$H_2O + O + 0.73$ eV	$\beta_8 = 4.2 \times 10^{-12} \exp(-\frac{242}{T})$
$OH + HO_2$	$\xrightarrow{\beta_9}$	$H_2O + O_2 + 3.06$ eV	$\beta_9 = 4.8 \times 10^{-11} \exp(\frac{215}{T})$
$OH + H_2$	$\xrightarrow{\beta_{10}}$	$H_2O + H + 0.65$ eV	$\beta_{10} = 7.7 \times 10^{-12} \exp(-\frac{2100}{T})$
$HO_2 + O_3$	$\xrightarrow{\beta_{11}}$	$OH + O_2 + O_2 + 1.23$ eV	$\beta_{11} = 1.4 \times 10^{-14} \exp(-\frac{580}{T})$
$HO_2 + HO_2$	$\xrightarrow{\beta_{12}}$	$H_2O_2 + O_2 + 1.71$ eV	$\beta_{12} = 2.3 \times 10^{-13} \exp(\frac{590}{T})$
$H + O_2 + M$	$\xrightarrow{\beta_{13}}$	$HO_2 + M + 2.11$ eV	$\beta_{13} = 5.5 \times 10^{-32} (\frac{300}{T})^{1.6}$
$H + O_3$	$\xrightarrow{\beta_{14}}$	$OH + O_2 + 3.34$ eV	$\beta_{14} = 1.4 \times 10^{-10} \exp(-\frac{470}{T})$
$H + HO_2$	$\xrightarrow{\beta_{15}}$	$H_2 + O_2 + 2.41$ eV	$\beta_{15} = 4.2 \times 10^{-11} \exp(-\frac{350}{T})$
$H + HO_2$	$\xrightarrow{\beta_{16}}$	$OH + OH + 1.61$ eV	$\beta_{16} = 4.2 \times 10^{-10} \exp(-\frac{950}{T})$
$H + HO_2$	$\xrightarrow{\beta_{17}}$	$H_2O + O + 2.34$ eV	$\beta_{17} = 8.3 \times 10^{-11} \exp(-\frac{500}{T})$
$H + H + M$	$\xrightarrow{\beta_{18}}$	$H_2 + M + 4.52$ eV	$\beta_{18} = 5.7 \times 10^{-32} (\frac{300}{T})^{1.6}$

Table 2.1: Oxygen/hydrogen chemistry reaction rates (From *Roble*, 1995)

Reaction			Reaction Rate
$N(^4S) + O_2$	$\xrightarrow{\beta_{19}}$	$NO + O + 1.40$ eV	$\beta_{19} = 1.5 \times 10^{-11} \exp(-\frac{3600}{T})$
$N(^2D) + O_2$	$\xrightarrow{\beta_{20}}$	$NO + O(^1D) + 1.84$ eV	$\beta_{20} = 5 \times 10^{-12}$
$N(^4S) + NO$	$\xrightarrow{\beta_{21}}$	$N_2 + O + 2.68$ eV	$\beta_{21} = 1.6 \times 10^{-10} \exp(-\frac{460}{T})$
$N(^2D) + O$	$\xrightarrow{\beta_{22}}$	$N(^4S) + O + 2.38$ eV	$\beta_{22} = 4.5 \times 10^{-13}$
$N(^2D) + e$	$\xrightarrow{\beta_{23}}$	$N(^4S) + e + 2.38$ eV	$\beta_{23} = 3.6 \times 10^{-10} (\frac{T_e}{300})^{1/2}$
$N(^2D) + NO$	$\xrightarrow{\beta_{24}}$	$N_2 + O + 5.63$ eV	$\beta_{24} = 7 \times 10^{-11}$
$N(^4S) + OH$	$\xrightarrow{\beta_{25}}$	$NO + H + 2.10$ eV	$\beta_{25} = 5 \times 10^{-11}$
$NO + O_3$	$\xrightarrow{\beta_{26}}$	$NO_2 + O + 2.08$ eV	$\beta_{26} = 1.8 \times 10^{-12} \exp(-\frac{1370}{T})$
$NO + HO_2$	$\xrightarrow{\beta_{27}}$	$NO_2 + OH + 0.35$ eV	$\beta_{27} = 3.5 \times 10^{-12} \exp(\frac{250}{T})$

Table 2.2: Nitrogen chemistry reaction rates (From *Roble*, 1995)

Reaction			Reaction Rate
$O(^1D) + O_2$	$\xrightarrow{\beta_{28}}$	$O(^3P) + O_2 + 1.96$ eV	$\beta_{28} = 3.2 \times 10^{-11}\ \exp(\frac{67}{T})$
$O(^1D) + N_2$	$\xrightarrow{\beta_{29}}$	$O(^3P) + N_2 + 1.96$ eV	$\beta_{29} = 1.8 \times 10^{-11}\ \exp(\frac{107}{T})$

Table 2.3: $O(^1D)$ quenching rates (From *Roble*, 1995)

Reaction			Reaction Rate
$NO^+ + e$	$\xrightarrow{\beta_{36}}$	$N(^4S) + O + 2.75$ eV (20%)	$\beta_{36} = 4.2 \times 10^{-7}\ (\frac{300}{T_e})^{0.85}$
$NO^+ + e$	$\xrightarrow{\beta_{36}}$	$N(^2D) + O + 0.38$ eV (80%)	$\beta_{36} = 4.2 \times 10^{-7}\ (\frac{300}{T_e})^{0.85}$
$O_2^+ + e$	$\xrightarrow{\beta_{37}}$	$O(^3P) + O(^3P) + 6.95$ eV (15%)	$\beta_{37} = 1.6 \times 10^{-7}\ (\frac{300}{T_e})^{0.55}$
$O_2^+ + e$	$\xrightarrow{\beta_{37}}$	$O(^1D) + O(^3P) + 4.98$ eV (85%)	$\beta_{37} = 1.6 \times 10^{-7}\ (\frac{300}{T_e})^{0.55}$
$N_2^+ + e$	$\xrightarrow{\beta_{38}}$	$N(^4S) + N(^4S) + 5.82$ eV (10%)	$\beta_{38} = 1.8 \times 10^{-7}\ (\frac{300}{T_e})^{0.39}$
$N_2^+ + e$	$\xrightarrow{\beta_{38}}$	$N(^2D) + N(^4S) + 3.44$ eV (90%)	$\beta_{38} = 1.8 \times 10^{-7}\ (\frac{300}{T_e})^{0.39}$

Table 2.4: Recombination rates (From *Roble*, 1995)

rates due to the reactions in Table 2.4 can be obtained by multiplying the reactant number densities, the reaction rates, and the amount of energy released together as both of the products in these reactions are neutrals. The heating rates due to the reactions in Table 2.5 are a bit more difficult to determine as one must now consider how much of the energy being released in these reactions is being acquired by the neutral species (this was not a concern in the other categories as *both* of the products were neutrals). Following Stolarski [1975,1976], the energy gained by each product was set equal to the inverse ratio of its mass, i.e. $Q_{ion} = E_{reaction} \times m_{neutral}/(m_{neutral} + m_{ion})$ and $Q_{neutral} = E_{reaction} \times m_{ion}/(m_{neutral} + m_{ion})$.

2.2.4 Ion/Neutral and Electron/Neutral Heating Rates

The neutral gas heating rate due to electron and ion collisions with the neutrals was calculated by assuming that the energy being lost by ions and electrons in collisions with the neutrals was equal to the heating rate of the neutrals by these collisions.

The energy lost by electrons in elastic electron-neutral collisions is usually formu-

Reaction			Reaction Rate
$O_2^+ + N(^4S)$	$\xrightarrow{\beta_{30}}$	$NO^+ + O + 4.21$ eV	$\beta_{30} = 1 \times 10^{-10}$
$O_2^+ + NO$	$\xrightarrow{\beta_{31}}$	$NO^+ + O_2 + 2.813$ eV	$\beta_{31} = 4.4 \times 10^{-10}$
$N^+ + O_2$	$\xrightarrow{\beta_{32}}$	$O_2^+ + N(^4S) + 2.486$ eV	$\beta_{32} = 4 \times 10^{-10}$
$N^+ + O_2$	$\xrightarrow{\beta_{33}}$	$NO^+ + O + 6.699$ eV	$\beta_{33} = 2 \times 10^{-10}$
$N^+ + O$	$\xrightarrow{\beta_{34}}$	$O^+ + N + 0.98$ eV	$\beta_{34} = 1 \times 10^{-12}$
$N_2^+ + O_2$	$\xrightarrow{\beta_{35}}$	$O_2^+ + N_2 + 3.52$ eV	$\beta_{35} = 6 \times 10^{-11}$

Table 2.5: Ion chemistry reaction rates (From *Roble*, 1995)

lated in terms of the linear 13-moment energy exchange term

$$L_{e,n} = - \frac{n_e m_e \nu_{en}}{(m_e + m_n)} \, 3k \, (T_e - T_n) \tag{2.29}$$

where ν_{en} is momentum transfer collision frequency between the electrons and the neutral species that the electrons are colliding with [*Schunk and Nagy*, 2000]. Since it has been assumed that all of the energy being lost by the electrons is being gained by the neutral species, one can write the heating rate of the neutral species due to elastic electron collisions as

$$Q_{n,e} = \frac{n_e m_e \nu_{en}}{(m_e + m_n)} \, 3k \, (T_e - T_n). \tag{2.30}$$

Substituting the appropriate expressions for the momentum transfer collision frequencies into the preceding equation yields

$$Q_{N_2,e} = 1.77 \times 10^{-19} n_e n_{N_2} [1 - (1.21 \times 10^{-4}) T_e] T_e (T_e - T_n) \tag{2.31}$$

$$Q_{O_2,e} = 1.21 \times 10^{-18} n_e n_{O_2} [1 + (3.6 \times 10^{-2}) \sqrt{T_e}] \sqrt{T_e} (T_e - T_n) \tag{2.32}$$

$$Q_{He,e} = 2.46 \times 10^{-17} n_e n_{He} \sqrt{T_e} (T_e - T_n) \tag{2.33}$$

$$Q_{O,e} = 7.9 \times 10^{-19} n_e n_O [1 + (5.7 \times 10^{-4}) T_e] \sqrt{T_e} (T_e - T_n) \tag{2.34}$$

$$Q_{H,e} = 9.63 \times 10^{-16} n_e n_H [1 - (1.35 \times 10^{-4}) T_e] \sqrt{T_e} (T_e - T_n) \tag{2.35}$$

[*Schunk and Nagy*, 2000], with the heating rates in units of eV cm^{-3} sec^{-1}. These rates do not have a significant impact on the neutral gas heating rate as the small mass ratio, i.e. $m_e/(m_e + m_n)$, usually inhibits the transfer of energy from electrons to the neutrals.

The energy lost by electrons in inelastic electron-neutral collisions is much more complicated to parameterize as one must now account for the excitation of vibrational and rotational modes in the neutral species. This requires a knowledge of the excitation cross-sections, which are usually not well-known [*Schunk and Nagy*, 2000]. To get around this limitation, theoretically derived cross-sections are frequently used and analytical functions fitted to the calculated cooling rates. These cooling rates are then used to determine the neutral heating rates due to inelastic electron collisions by (again) assuming that all of the energy being lost by the electrons is being gained by the neutral species. The following excitations were considered :

$$\begin{aligned}
\mathrm{N_2\ rotation} : \mathrm{Q_{N_2,rot}} &= 2.9\times10^{-14}\, n_e\, n_{\mathrm{N_2}}\, (T_e - T_n)/\sqrt{T_e} && (2.36)\\
\mathrm{O_2\ rotation} : \mathrm{Q_{O_2,rot}} &= 6.9\times10^{-14}\, n_e\, n_{\mathrm{O_2}}\, (T_e - T_n)/\sqrt{T_e} && (2.37)\\
\mathrm{N_2\ vibration} : \mathrm{Q_{N_2,vib}} &= 2.99\times10^{-12}\, n_e\, n_{N_2}\, \exp\left[\, f(T_e-2000)/2000T_e \,\right] \\
&\quad \times \left[\exp\left(-g\, \frac{T_e - T_n}{T_e T_n} \right) - 1 \right] && (2.38)\\
\mathrm{O_2\ vibration} : \mathrm{Q_{O_2,vib}} &= 5.2\times10^{-13}\, n_e\, n_{\mathrm{O_2}}\, \exp\left[\, h(T_e-700)/700T_e \,\right] \\
&\quad \times \left[\exp\left(-2270\, \frac{T_e - T_n}{T_e T_n} \right) - 1 \right] && (2.39)\\
\mathrm{O(^1D)\ excitation} : \mathrm{Q_{O(^1D)}} &= 1.57\times10^{-12}\, n_e\, n_{\mathrm{O}}\, \exp\left[\, d(T_e-3000)/3000T_e \,\right] \\
&\quad \times \left[\exp\left(-22713\, \frac{T_e - T_n}{T_e T_n} \right) - 1 \right] && (2.40)\\
\mathrm{O\ fine\ structure} : \mathrm{Q_{O,fine}} &= n_e\, n_{\mathrm{O}}\, \left[\, A_1\, (A_2 + A_3 + A_4)\, A_5 \,\right] && (2.41)
\end{aligned}$$

where

$$\begin{aligned}
d &= 2.4\times10^4 + 0.3(T_e - 1500) - 1.947\times10^{-5}(T_e-1500)(T_e-4000)\\
f &= 1.06\times10^4 + 7.51\times10^3\, \tanh\left[1.10\times10^{-3}(T_e-1800)\right]\\
g &= 3300 + 1.233(T_e-1000) - 2.056\times10^{-4}(T_e-1000)(T_e-4000)\\
h &= 3300 - 839\, \sin\left[1.91\times10^{-4}(T_e-2700)\right]\\
A_1 &= 3.0\times10^{10} / \left\{ T_e^{1/2} \left[5 + 3\exp\left(-\frac{0.02}{A_6}\right) + \exp\left(-\frac{0.028}{A_6}\right) \right] \right\}\\
A_2 &= 0.02\left[\, (4.75\times10^{-21})\, T_e + 6.82\times10^{-17} \,\right] \left[\exp\left(-\frac{0.02}{A_5}\right) - \exp\left(-\frac{0.02}{A_6}\right) \right]\\
A_3 &= 0.028\left[(2.3\times10^{-21}) T_e + 1.16\times10^{-17} \right] \left[\exp\left(-\frac{0.028}{A_5}\right) - \exp\left(-\frac{0.028}{A_6}\right) \right]
\end{aligned}$$

$$
\begin{aligned}
A_4 &= 0.008\,[(3.0\times10^{-21})T_e + 2.20\times10^{-17}\,]\left[\exp\left(-\frac{0.008}{A_5}\right) - \exp\left(-\frac{0.008}{A_6}\right)\right] \\
A_5 &= 8.63\times10^{-5}\,T_e \\
A_6 &= 8.63\times10^{-5}\,T
\end{aligned}
$$

The expression for O(^{1}D) is from *Rees and Roble* [1975]; references for all of the other expressions can be found in *Schunk and Nagy* [2000].

There are two types of collisions that contribute to the exchange of energy between ions and neutrals. The first type of collision is a non-resonant collision, which involves an unlike ion and neutral, e.g. O^+ and N_2. This type of collision is usually modeled as an elastic collision that is being caused by a long-range polarization attraction (due to the induced dipole in the neutral) and a short-range repulsion. Since this kind of collision is elastic, one can follow the same arguments presented in the elastic electron neutral collision section to obtain

$$Q_{n,i} = \frac{n_i\, m_i\, \nu_{in}}{(m_i + m_n)}\, 3k\,(T_i - T_n). \tag{2.42}$$

The ion-neutral momentum transfer collision frequency for a non-resonant ion-neutral collision is given by the expression

$$\nu_{in} = 2.21\pi\, \frac{n_n m_n}{m_i + m_n}\, \left(\frac{\gamma_n e^2}{\mu_{in}}\right)^{\frac{1}{2}} \tag{2.43}$$

[*Dalgarno et al.*, 1958], where γ_n is the polarizability of the neutral species and μ_{in} is the reduced mass of the ion and neutral. This now allows one to rewrite equation (2.42) as

$$
\begin{aligned}
Q_{n,i} &= 3\, n_i\, n_n\, \frac{2.21\pi}{(m_i + m_n)}\, (\mu_{in}\gamma_n e^2)^{\frac{1}{2}}\; k(T_i - T_n) && (2.44) \\
&= 3\, \left(\frac{m_i}{m_i + m_n}\right)\, C_{in}\, n_i\, n_n\, k\,(T_i - T_n). && (2.45)
\end{aligned}
$$

The neutral species polarizabilities and C_{in} values can be found in Tables 2.6 and 2.7 respectively. Since O^+ and O_2^+ are the only ions of consequence at altitudes where

Species	$\gamma_n(\mathrm{cm}^3)$
N_2	1.76×10^{-24}
O_2	1.60×10^{-24}
O	0.77×10^{-24}
He	0.21×10^{-24}

Table 2.6: Neutral gas polarizabilities. (From *Schunk and Nagy*, 2000.)

	O_2^+	O^+
N_2	4.13×10^{-10}	6.82×10^{-10}
O_2	-	6.64×10^{-10}
O	2.31×10^{-10}	-
He	0.70×10^{-10}	1.32×10^{-10}

Table 2.7: Nonresonant ion-neutral collision frequency coefficients. (From *Schunk and Nagy*, 2000.)

ion-neutral heating is important, the following heating rates were included :

$$Q_{N_2,O^+} = \left(\frac{12}{11}\right) \times 6.82 \times 10^{-10}\, n_{O^+}\, n_{N_2}\, k(T_i - T_n) \quad (2.46)$$

$$Q_{O_2,O^+} = (1.0) \times 6.64 \times 10^{-10}\, n_{O^+}\, n_{O_2}\, k(T_i - T_n) \quad (2.47)$$

$$Q_{He,O^+} = \left(\frac{24}{9}\right) \times 1.32 \times 10^{-10}\, n_{O^+}\, n_{He}\, k(T_i - T_n) \quad (2.48)$$

$$Q_{N_2,O_2^+} = \left(\frac{24}{15}\right) \times 4.13 \times 10^{-10}\, n_{O_2^+}\, n_{N_2}\, k(T_i - T_n) \quad (2.49)$$

$$Q_{O,O_2^+} = (2.0) \times 2.31 \times 10^{-10}\, n_{O_2^+}\, n_{O}\, k(T_i - T_n) \quad (2.50)$$

The second type of collision that contributes to the exchange of energy between ions and neutrals is a resonant collision, in which an electron from a neutral species is transferred to an ion species, i.e., the ion becomes a neutral and the neutral becomes an ion. This charge exchange usually results in a large transfer of energy and momentum to the neutral species as very little energy is lost in the charge exchange process. Only two resonant charge exchanges were considered, $O^+ - O$ and $O^+ - H$, which contribute to the neutral gas heating rate as

$$Q_{O,O^+} = \left(\frac{3}{2}\right) \times 3.67 \times 10^{-11}\, n_{O^+}\, n_{O}\, k(T_i - T_n)\, T_R^{1/2}\, [\,1 - 0.064 \log_{10}(T_R)\,]^2 \quad (2.51)$$

$$Q_{H,O^+} = \left(\frac{48}{17}\right) \times 6.61 \times 10^{-11}\, n_{O^+}\, n_{H}\, k(T_i - T_n)\, T_i^{1/2}\, [\,1 - 0.047 \log_{10}(T_i)\,]^2 \quad (2.52)$$

[*Schunk and Nagy*, 2000], where T_R is the reduced temperature of the ions and neutrals, i.e. $T_R = (T_i + T_n)/2$. The $O^+ - O$ reaction is the dominant process in determining the neutral gas heating rates at higher altitudes and is quite controversial as

the $O^+ - O$ collision frequency is not well known (a more complete discussion about the controversy surrounding the $O^+ - O$ collision frequency can be found in *Buotsano* [1995]). Once the heating rates for all of the electron-neutral and ion-neutral collisions have been determined, they are added together and converted to ergs cm^{-3} sec^{-1} by multiplication of the appropriate conversion factor.

2.3 Cooling Rates

Despite the veritable plethora of emission lines in the IR region, only a few of these emissions are strong enough to affect the energetics of the thermosphere: NO at 5.3μm, CO_2 at 15μm, and atomic oxygen at 63 μm and 147 μm. Incorporating these emissions into an atmospheric model is usually accomplished through a parameterization scheme as the numerical solution to the radiative transfer equation in the mesosphere and thermosphere is complicated by several factors: the upper mesosphere and thermosphere are not in a state of local thermodynamic equilibrium (LTE), most of the important deactivation rates in the emission rate calculations are not that well known, and most of the deactivation rates also depend upon the atomic oxygen number densities, which are not that well known either.

The importance of the atomic oxygen emissions at 63μm and 147μm was first demonstrated by *Bates* [1951] in his study on the rate of energy loss in an optically thin atmosphere. However, treating the thermosphere as an optically thin medium will lead to a substantial overestimate of the atomic oxygen emissions in the 100-150 km region [*Stewart*, 1968; *Craig and Gille*, 1969]. To compensate for this, *Kockarts and Peetermans* [1970] developed an emission rate parameterization that contained Bates' "optically thin" parameterization and a "masking factor" $(1 - \chi)$ which accounted for the effects of absorption and induced emission, i.e.

$$C_{63\mu m} = (1-\chi) \times \frac{1.69 \times 10^{-18}\, n_O\, \exp(-\frac{228}{T})}{1 + 0.6\exp(-\frac{228}{T}) + 0.2\exp(-\frac{326}{T})} \tag{2.53}$$

$$C_{147\mu m} = (1-\chi) \times \frac{4.59 \times 10^{-20}\, n_O\, \exp(-\frac{326}{T})}{1 + 0.6\exp(-\frac{228}{T}) + 0.2\exp(-\frac{326}{T})}. \tag{2.54}$$

More recent airglow observations by *Grossmann and Offermann* [1978] have suggested

that the observed 63μm emission rates above 120 km are much smaller than those predicted by equations (2.53) and (2.54), with non-LTE effects being the suspected cause. These uncertainties have led many modelers, e.g. *Roble* [1987], to reduce the atomic oxygen 63 μm and 147 μm emissions by factor of 2. The current version of the model does not reduce these emission rates by a factor of 2 and uses a simplified time-independent 'masking factor' that resembles the masking factors used in the CTIP/CMAT models developed by *Fuller-Rowell et al.* (more details on the numerical implementation of these masking factors can be found in the Ph.D. thesis of Matthew Harris]2000]). This simplification will have no effects on the results of this thesis as the masking factor is only important when one is trying to reproduce the temperature climatology of the lower thermosphere.

The CO_2 emissions in the 15 μm band were originally parameterized in the model as

$$\begin{aligned} C_{15\mu m} &= 1.3 \times 10^{-13}\, g_{10}\, \exp\left(-\frac{960}{T}\right)\, n_{CO_2} \left(\sum_M k_M[M]\right) \\ &\quad \times \left[1 - (1+\xi)e^{-\xi}\right] \end{aligned} \tag{2.55}$$

[*Chamberlain and McElroy*, 1966], where

$$\begin{aligned} \sum_M k_M[M] &= (\,4.4 n_{O_2} + 1.467 n_{N_2}\,) \times 10^{-15}\, T\, \exp\left(-\frac{41}{T^{1/3}}\right) \\ &\quad + 1.5 \times 10^{-11}\, \exp\left(-\frac{800}{T}\right)\, n_O \end{aligned} \tag{2.56}$$

is the collisional deactivation rate for CO_2 and

$$\xi = \frac{10^9}{n_{CO_2}} \left(1 + \frac{A}{\sum_M k_M[M]}\right) \tag{2.57}$$

is a factor that accounts for absorption of radiation in the 15 μm band. This expression is valid if the number densities of CO_2 are less than $10^9\,cm^{-3}$, i.e. above 100 km in the Earth's atmosphere. If this formula is used at lower altitudes where the CO_2 number densities are much larger, significant errors can occur (Figure 2.1).

Although these errors were tolerated in the original version of the code, the current emphasis on modeling composition/energetics feedbacks more accurately has demanded that a more realistic parameterization be incorporated into the model. The

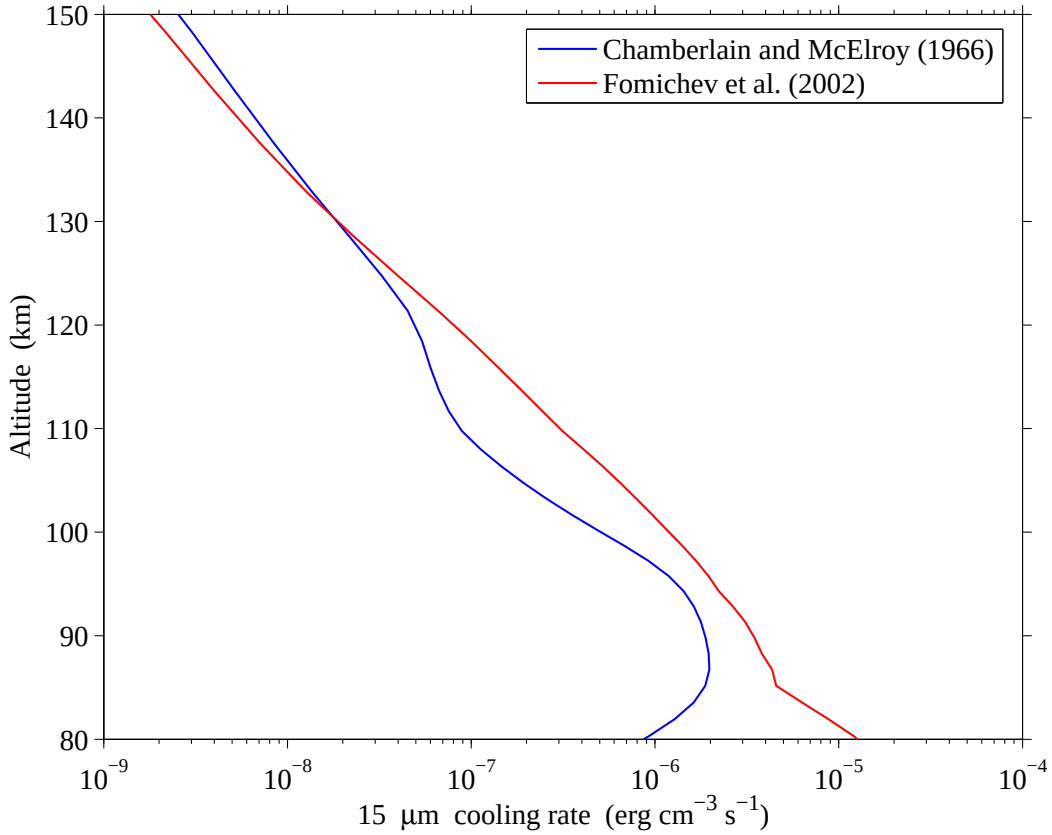

Figure 2.1: CO_2 15 μm cooling rate parameterizations. The parameterization of *Chamberlain and McElroy* [1966] substantially underestimates the cooling rates in the lower thermosphere, with order of magnitude errors occurring near 80 km.

Fomichev et al. [2002] scheme was chosen for this task. This scheme is an improved version of the scheme developed by *Fomichev et al.* [1998], with a more accurate treatment of the non-LTE region, i.e. the errors in the non-LTE regions of the more "extreme" temperature profiles are now less than 5% above 100 km.

The *Fomichev et al.* [2002] scheme determines the 15 μm emissions on the dimensionless log-pressure grid

$$x = \ln\left(\frac{1000}{p(\text{mb})}\right) \tag{2.58}$$

which means that the variables used in the determination of the cooling rates must be interpolated onto the dimensionless log-pressure grid. Once this interpolation has

been performed, the cooling rates from $x = 2.0 - 12.5$ (z < 85 km) are calculated as

$$\epsilon(x_0) = \sum_{j=-5}^{3} [\, a_j(x_0) + \, b_j(x_0)\, \psi_0 \,]\, \psi_j \tag{2.59}$$

where $\epsilon(x_0)$ is the cooling rate (in erg $\mathrm{g}^{-1}\,\mathrm{s}^{-1}$) at grid point x_0 and ψ_j is the exponential part of the source function at grid point j, i.e.

$$\psi_j = \exp(-h\nu/kT_j) = \exp(-960.217/T_j). \tag{2.60}$$

Nine different layers/grid points ($j = -5, -4, ..., 3$) are used in the cooling rate calculation to ensure that radiative exchanges between the different atmospheric layers are being properly accounted for. The distances of these layers/levels from the level where the cooling rate is being determined can be found in Table 1 of *Fomichev et al.* [1998].

In the region $x = 12.75 - 16.50$ (85 km $<$ z $<$ 115 km), the cooling rates are determined by the expression

$$\epsilon(x_j) = 2.55521 \times 10^{11} \times C_{\mathrm{CO_2}}(x_j)\, (1 - \lambda(x_j))\, \overline{\epsilon}(x_j)\, /\, \mu(x_j) \tag{2.61}$$

where $C_{\mathrm{CO_2}}(x_j)$ is the mixing ratio of $\mathrm{CO_2}$ at level j, $\lambda(x_j)$ is the quantum survival probability at level j, $\mu(x_j)$ is the mean molecular mass (in g/mole) at level j, and $\overline{\epsilon}(x_j)$ is the cooling rate derived from the *Kutepov and Fomichev* [1993] recurrence formula (more details about the determination of $\overline{\epsilon}(x_j)$ can be found in *Fomichev et al.*, [1998]). The quantum survival probability at level j, $\lambda(\mathrm{x_j})$, is determined by the expression

$$\lambda(x_j) = 1.5988/(1.5988 + Z_{10}) \tag{2.62}$$

with the collisional deactivation rate, Z_{10}, given by

$$\begin{aligned} Z_{10} &= \rho(x_j)\, [\, C_{\mathrm{N_2}}(x_j)\, k_{\mathrm{N_2}} + C_{\mathrm{O_2}}(x_j)\, k_{\mathrm{O_2}} + C_{\mathrm{O}}(x_j)\, k_{\mathrm{O}} \,] \\ &= \rho(x_j)\, \{ C_{\mathrm{N_2}}(x_j) [\, 5.5 \times 10^{-17}\, \sqrt{T} + 6.7 \times 10^{-10}\, \exp(\, -83.3\, T^{-1/3}) \,] \\ &\quad + C_{\mathrm{O_2}}(x_j) [\, 1.0 \times 10^{-15}\, \exp(\, 23.37 - 230.9\, T^{-1/3} + \, 564\, T^{-2/3}\,) \,] \\ &\quad + C_{\mathrm{O_2}}(x_j)\, [\, 3 \times 10^{-12}] \,\} \ . \end{aligned} \tag{2.63}$$

$C_{N_2}(x_j)$, $C_{O_2}(x_j)$, and $C_O(x_j)$ are the mixing ratios of N_2, O_2, and O, respectively, at level j. The quenching rates of CO_2 by N_2 and O_2, i.e. k_{N_2} and k_{O_2}, are not particularly contentious. The quenching rate of CO_2 by O, on the other hand, is. Laboratory measurements by *Shved et al.* [1991] and *Pollock et al.* [1993] have suggested that the value of k_O is $(1.5 \pm 0.5) \times 10^{-12}$cm^3/s and $(1.2 \pm 0.2) \times 10^{-12}$cm^3/s respectively. Analysis of satellite observations suggest that k_O is much higher, with values of $(3-6) \times 10^{-12}$cm^3/s being common, e.g. *Sharma and Wintersteiner* [1990], *Rogers et al.* [1992], and *Lopez-Puertas et al.* [1992]. A value of 3×10^{-12}cm^3/s was selected for these simulations because this value usually works well in most simulations of planetary thermospheres, e.g. *Bougher et al.* [1994] and *Roble* [2000].

Above $x = 16.5$ (z $>$ 115km), the 15 μm cooling rates are determined by the expression

$$\epsilon(x_j) = 2.55521 \times 10^{11} \times C_{CO_2}(x_j)\ (1 - \lambda(x_j))\ [\Phi(16.5) - \psi_j]\ /\ \mu(x_j) \qquad (2.64)$$

where $\Phi(16.5)$ is the radiative flux at x = 16.5. Once the 15 μm cooling rates have been computed at all of the points on the dimensionless log-pressure grid, they are interpolated back onto the model grid.

The NO emission at 5.3μm also plays an important role in the cooling of the thermosphere [*Gordiets*, 1978]. The non-LTE parameterization developed by *Kockarts* [1980] is currently being used in the model, i.e.

$$C_{5.3\mu m} = 4.985 \times 10^{-12}\ n_{NO}\ \left(\frac{6.5 \times 10^{-11} n_O}{6.5 \times 10^{-11} n_O + 13.3}\right)\ \exp\left(-\frac{2714.7}{T}\right). \qquad (2.65)$$

It assumes that NO can be modeled as a two level molecule ($v = 0, 1$) with the NO($v = 1$) deactivation rates being controlled by atomic oxygen, i.e. $k_{10}(N_2) < k_{10}(O_2) << k_{10}(O)$. While there is little debate about the dominance of the $k_{10}(O)$ deactivation rate, the actual value of the rate is still in question. Recent calculations [*Duff and Sharma*, 1997] and measurements [*Dodd et al.*, 1999] suggest that the deactivation rate is approximately 2.5×10^{-11} cm^3/s at room temperature, which is nearly one-third of the previously accepted value of 6.5×10^{-11} cm^3/s [*Fernando and Smith*, 1979]. *Sharma and Roble* [2001] have implemented this newer value in the globally averaged

NCAR TIME-GCM and found that there are significant changes in the thermal and density structure of the thermosphere unless the $CO_2 - O$ deactivation rate is doubled to $6 \times 10^{-12} cm^3/s$, which is the currently the largest value supported by observation. The 'older' *Fernando and Smith* [1979] value is still being used in the model as there is no general consensus within the thermospheric community on which rate value is correct.

2.4 Composition

To model the physics of the thermosphere and upper mesosphere with any degree of realism, one must have a knowledge on how N_2, O_2, O, NO, and CO_2 are distributed in these regions. The composition fractions of N_2, O_2, and O originally employed in the model are shown in Figure 2.2. The O and O_2 composition fractions were not particularly realistic at the lowest altitudes, but these values were tolerated at the time as this region of the model was not deemed to be that important to the simulation. Profiles of NO and CO_2 were obtained from the measurements of *Trinks et al.* [1978] and *Allen et al.* [1981] respectively.

The current emphasis on the more realistic treatment of compositional effects has led to several improvements in how the various constituents are treated in the model. The *Trinks et al.* [1978] NO profile has been updated with a NO profile obtained from the SME/HALOE climatology of NO [*Siskind et al.*, 1998]. This profile is shown in Figure 2.3 and is representative of disturbed conditions at 60°N in January (disturbed conditions in the SME/HALOE climatology are defined as $Ap > 15$ and $F10.7 > 120$). The CO_2 profile of *Trinks et al.* [1978] has also been updated with the 'average' profile used in the study of *Fomichev et al.* [1998]. These profiles remain constant through a simulation.

The treatment of N_2, O_2, and O have also been improved in the model. The 'background' values of these species are now much more realistic, with the composition fractions below 130 km now being equal to the composition fractions of MSIS-90e [*Hedin*, 1990]. At higher altitudes, the 'background' profiles of N_2, O_2, and O are

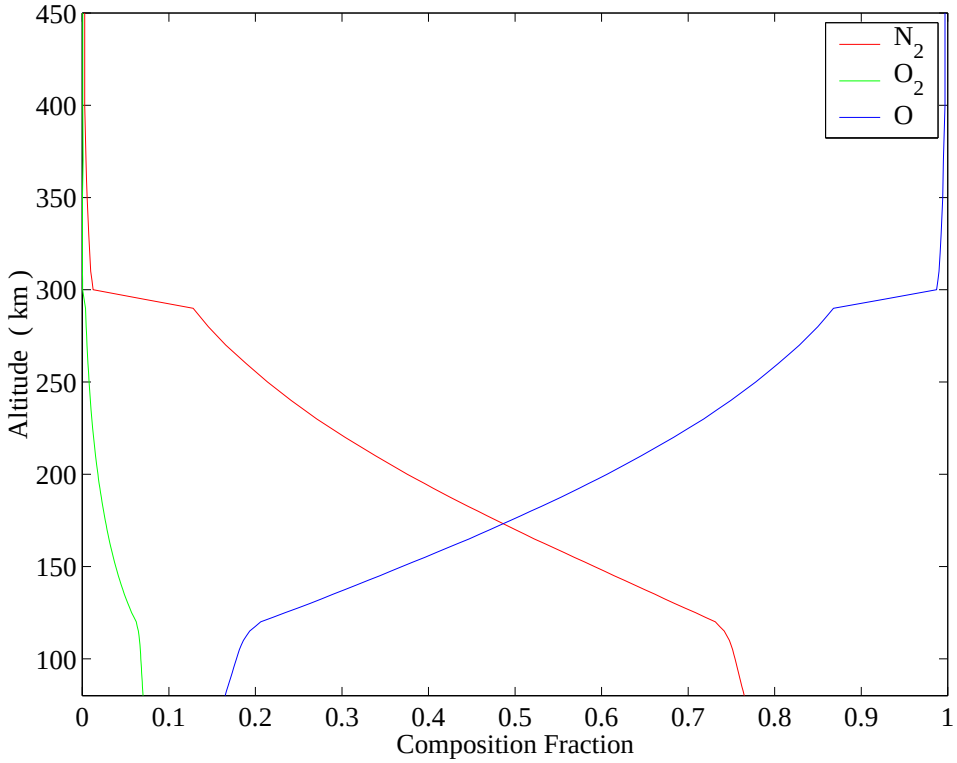

Figure 2.2: Composition fractions employed in the *Chang and St.-Maurice* [1991] simulations. The O and O_2 composition fractions were not particularly realistic at the lowest altitudes. The O and N_2 composition fractions also exhibited unrealistically large "jumps" at 300 km. (Adapted from *Chang and St.-Maurice*, 1991)

assumed to be in diffusive equilibrium. Once the simulation starts, these composition fractions are allowed to evolve as the effects of transport and molecular diffusion are now being accounted for.

The more self-consistent treatment of the neutral gas heating rates in the model also dictates that the distributions of many minor neutral species in the thermosphere and upper mesosphere are now required. The profile of $N(^4S)$ was obtained from the MSIS-90e model [*Hedin*, 1990]. The atomic hydrogen profile is a blend of *Thomas*' [1996] climatology and MSIS; the average December profile for 60° N was used below 94 km and a suitably scaled MSIS profile was used above 94 km. The profiles of O_3, HO_2, and OH were calculated by the methodology of *Brasseur and Solomon* [1986]. In this methodology, the profiles of O, O_2 and H are already known and a very simple

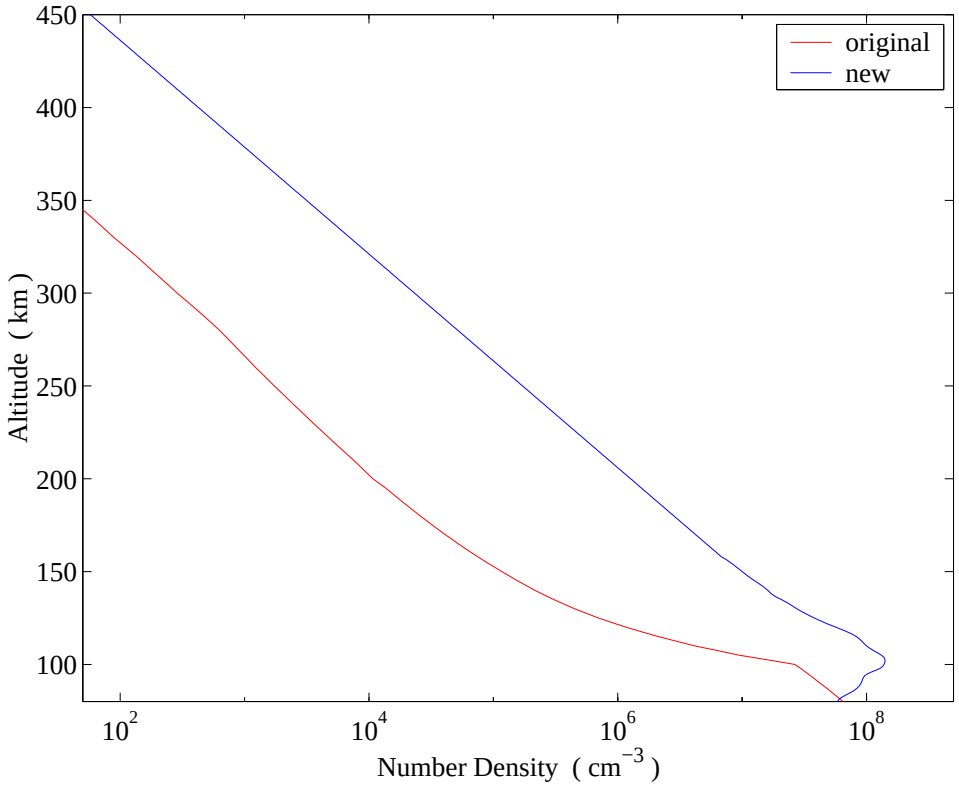

Figure 2.3: Nitric oxide density profiles

hydrogen-oxygen chemistry, i.e.

$$\begin{aligned}
\mathrm{O} + \mathrm{O_2} + \mathrm{M} &\xrightarrow{\beta_2} \mathrm{O_3} + \mathrm{M} \\
\mathrm{O} + \mathrm{O_3} &\xrightarrow{\beta_3} \mathrm{O_2} + \mathrm{O_2} \\
\mathrm{O} + \mathrm{OH} &\xrightarrow{\beta_4} \mathrm{O_2} + \mathrm{H} \\
\mathrm{O} + \mathrm{HO_2} &\xrightarrow{\beta_5} \mathrm{OH} + \mathrm{O_2} \\
\mathrm{O_3} + \mathrm{h}\nu &\xrightarrow{\mathbf{J_{O_3}}} \mathrm{O_2} + \mathrm{O_2} \\
\mathrm{H} + \mathrm{O_2} + \mathrm{M} &\xrightarrow{\beta_{13}} \mathrm{HO_2} + \mathrm{M} \\
\mathrm{H} + \mathrm{O_3} &\xrightarrow{\beta_{14}} \mathrm{OH} + \mathrm{O_2}
\end{aligned}$$

is solved under the assumption of photochemical equilibrium, yielding

$$n_{\mathrm{O_3}} = \frac{\beta_2 n_{\mathrm{O}} n_{\mathrm{O_2}} n_{\mathrm{M}}}{J_{\mathrm{O_3}} + \beta_3 n_{\mathrm{O}} + \beta_{14} n_{\mathrm{H}}} \tag{2.66}$$

$$n_{\mathrm{HO_2}} = \frac{\beta_{13} n_{\mathrm{H}} n_{\mathrm{O_2}} n_{\mathrm{M}}}{\beta_5 n_{\mathrm{O}}} \tag{2.67}$$

$$\begin{aligned}
n_{\mathrm{OH}} &= \frac{\beta_5 n_{\mathrm{HO_2}} n_{\mathrm{O}} + \beta_{14} n_{\mathrm{H}} n_{\mathrm{O_3}}}{\beta_4 n_{\mathrm{O}}} \\
&= \frac{\beta_5 n_{\mathrm{HO_2}} n_{\mathrm{O}} + \beta_{13} n_{\mathrm{H}} n_{\mathrm{O_2}}}{\beta_4 n_{\mathrm{O}}}
\end{aligned} \tag{2.68}$$

The profile of $\mathrm{N}(^2\mathrm{D})$ is also determined under the assumption of photochemical equilibrium, i.e.

$$\begin{aligned}
\mathrm{N_2} + h\nu &\xrightarrow{\mathrm{J_{N_2}}} \mathrm{N}(^2\mathrm{D}) + \mathrm{N}(^4\mathrm{S}) \\
\mathrm{NO^+} + e &\xrightarrow{\beta_{36}} \mathrm{N}(^2\mathrm{D}) + \mathrm{O} \\
\mathrm{N_2^+} + e &\xrightarrow{\beta_{38}} \mathrm{N}(^2\mathrm{D}) + \mathrm{N}(^4\mathrm{S}) \\
\mathrm{N_2^+} + \mathrm{O} &\xrightarrow{\beta_{41}} \mathrm{N}(^2\mathrm{D}) + \mathrm{NO^+} \\
\mathrm{N^+} + \mathrm{O_2} &\xrightarrow{\beta_{32a}} \mathrm{N}(^2\mathrm{D}) + \mathrm{O_2^+} \\
\mathrm{N}(^2\mathrm{D}) + \mathrm{O_2} &\xrightarrow{\beta_{20}} \mathrm{NO} + \mathrm{O}(^1\mathrm{D}) \\
\mathrm{N}(^2\mathrm{D}) + \mathrm{O} &\xrightarrow{\beta_{22}} \mathrm{N}(^4\mathrm{S}) + \mathrm{O}
\end{aligned}$$

$$
\begin{aligned}
\mathrm{N}(^2\mathrm{D}) + \mathrm{NO} &\xrightarrow{\beta_{24}} \mathrm{N_2} + \mathrm{O} \\
\mathrm{N}(^2\mathrm{D}) + \mathrm{e} &\xrightarrow{\beta_{23}} \mathrm{N}(^4\mathrm{S}) + e \\
\mathrm{N}(^2\mathrm{D}) &\xrightarrow{\beta_{42}} \mathrm{N}(^4\mathrm{S}) + h\nu \\
\mathrm{N}(^2\mathrm{D}) + \mathrm{O_2^+} &\xrightarrow{\beta_{43}} \mathrm{N^+} + \mathrm{O_2} \\
\mathrm{N}(^2\mathrm{D}) + \mathrm{O^+} &\xrightarrow{\beta_{44}} \mathrm{N^+} + \mathrm{O}
\end{aligned}
$$

which yields

$$
n_{\mathrm{N}(^2\mathrm{D})} = \frac{0.5 J_{N_2} n_{N_2} + 0.8\beta_{36} n_{NO^+} n_e + 0.9\beta_{38} n_{N_2^+} n_e + \beta_{41} n_{N_2^+} n_O + \beta_{32a} n_{N^+} n_{O_2}}{\beta_{20} n_{O_2} + \beta_{22} n_O + \beta_{23} n_e + \beta_{24} n_{NO} + \beta_{42} + \beta_{43} n_{O_2^+} + \beta_{44} n_{O^+}}
$$

The $\mathrm{O}(^1\mathrm{D})$ profile was modeled as a Chapman profile above 160 km, i.e.

$$
n_{\mathrm{O}(^1\mathrm{D})} = 2.8 \times 10^4 \ \exp\left\{1 - \left(\frac{z-200}{26}\right) - \exp\left[-\left(\frac{z-200}{26}\right)\right]\right\}.
$$

Below 160 km, the $\mathrm{O}(^1\mathrm{D})$ profile was determined by interpolation of the following data

Altitude (km)	Number Density (cm^{-3})
80	1.00×10^2
115	3.00×10^2
140	9.50×10^2
155	2.25×10^3
160	3.36×10^3

It should be noted that no attempt was to made to calculate the $\mathrm{O}(^1\mathrm{D})$ profile using photochemical equilibrium considerations as this would have required a knowledge of the $\mathrm{O_2}$ photolysis rates.

2.5 Other Thermospheric Parameters

2.5.1 Transport Coefficients

Theoretical expressions for the transport coefficients of a multicomponent gas depend on how the interactions between the various species are modeled. For thermospheric

species j	A_j	C_j
N_2	3.43×10^{-6}	56
O_2	4.03×10^{-6}	56
O	3.90×10^{-6}	75.9
He	3.84×10^{-6}	299
H	1.22×10^{-6}	379

Table 2.8: Numerical coefficients used in the viscosity and thermal conductivity approximations. (From *Banks and Kockarts*, 1973.)

simulations, the viscosity and thermal conductivity of the individual species are frequently parameterized as

$$\mu_i = A_i \, T^{0.69} \tag{2.69}$$

$$\lambda_i = C_i \, T^{0.69} \tag{2.70}$$

where A_i and C_i are experimentally fitted coefficients. The values of these coefficients can be found in Table 2.8. These individual expressions are then weighted by their respective composition percentages to obtain the final viscosity and thermal conductivity expressions for the multicomponent mixture, i.e.

$$\mu = \sum_i A_i \, (\frac{n_i}{n}) \, T^{0.69} \tag{2.71}$$

$$\lambda = \sum_i C_i \, (\frac{n_i}{n}) \, T^{0.69}. \tag{2.72}$$

2.5.2 Body Forces

Three forces are being accounted for : gravity, coriolis, and ion drag. The Coriolis force per unit mass, f_c, is determined from the expression $-2\mathbf{\Omega} \times \mathrm{u_n}$, where Ω is the angular rotation rate of the Earth (equal to $7.29 \times 10^{-5} \mathrm{s}^{-1}$) and $\mathrm{u_n}$ is the neutral gas velocity vector. In the coordinate system of the model, this yields

$$\mathbf{f_c} = 2\Omega w \sin\theta \, \mathbf{e_r} \; + \; 2\Omega w \cos\theta \, \mathbf{e}_\theta \; - 2\Omega (u \sin\theta + v \cos\theta) \, \mathbf{e}_\phi \tag{2.73}$$

The ion drag force per unit mass is evaluated as $\frac{1}{\rho c}(\mathbf{J} \times \mathbf{B})$, where $\mathbf{J}$ is the current density in the ionosphere and $\mathbf{B}$ is the magnetic field of the Earth. In the coordinate system of the model, this yields

$$\mathbf{f_{ion}} = \frac{1}{\rho c}[\,(-J_\phi B_\theta)\,\mathbf{e_r} + (J_\phi B_r)\,\mathbf{e}_\theta + (J_r B_\theta - J_\theta B_r)\,\mathbf{e}_\phi\,] \tag{2.74}$$

The gravitational force per unit mass is evaluated using the gravitational field strength g, which is evaluated as

$$\mathrm{g} = \frac{980.6}{(1 + (\frac{z}{R_E})^2)}\ \mathrm{cm/sec^2}. \tag{2.75}$$

2.6 Ionospheric Parameterizations

2.6.1 Ion Density Profiles

The ion density profiles used in these simulations are shown in Figure 2.4. The 'weak case' ion density profile is taken from *St-Maurice and Schunk* [1981] and has an F-region peak ion density of 10^5 cm^{-3} and an E-region peak density of 3.5×10^5 cm^{-3} (the enhanced E-region peak at 110 km is due to the effects of particle precipitation). The 'strong case' ion density profile has an essentially constant ion density of 10^6 cm^{-3} above 200 km and has been chosen to study the response of the thermosphere under extremely disturbed conditions. It should be noted that both of these ion profiles do not carry any information about the composition of the ionosphere; the original version of the model assumed that the ionosphere was composed entirely of NO^+. A more realistic ion composition scheme has been introduced into the model, with the NO^+, O_2^+, and O^+ composition percentages corresponding to a TRANSCAR simulation.

2.6.2 Geomagnetic field

In most geophysical simulations, the geomagnetic field of the Earth is derived from a magnetic potential, i.e.,

$$\mathbf{B} = -\nabla \Phi_M. \tag{2.76}$$

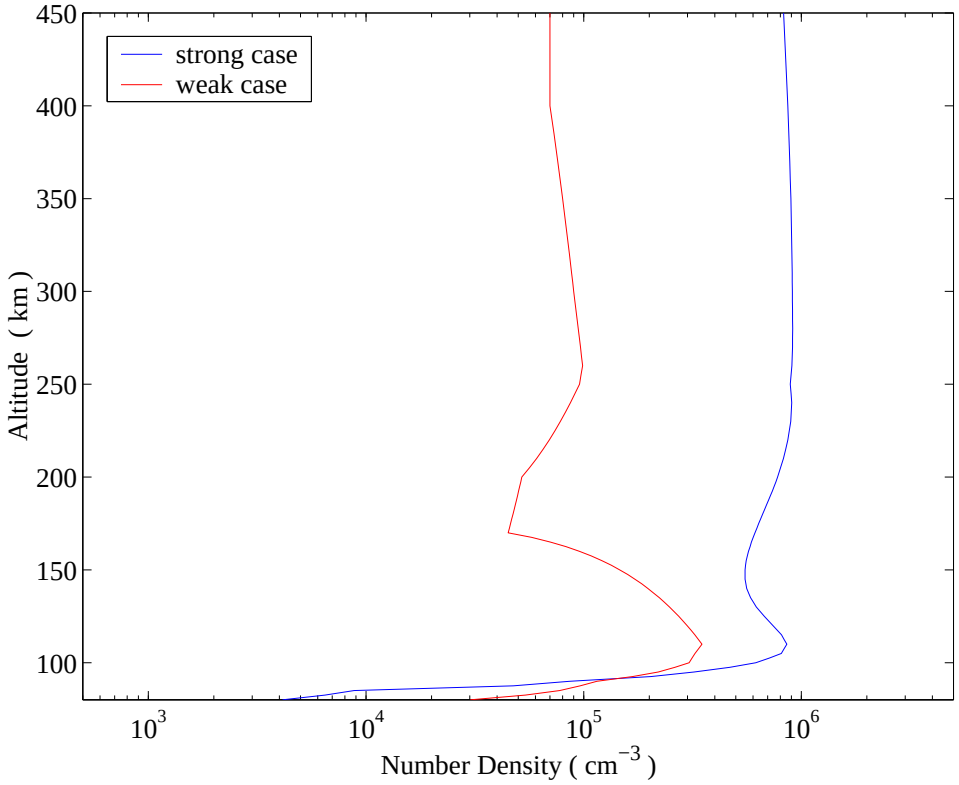

Figure 2.4: Ion density profiles. (Adapted from *Chang and St.-Maurice*, 1991.)

This magnetic potential is compiled from a global distribution of ground-based and satellite magnetometer measurements that are fit to a spherical harmonic series of the form

$$\Phi_M(r,\theta,\phi) = R_E \sum_{n=1}^{\infty}\sum_{m=0}^{n}(\frac{R_E}{r})^{n+1}(g_n^m \cos m\phi + h_n^m \sin m\phi)P_n^m(\cos\theta), \qquad (2.77)$$

(r is the distance from the centre of the Earth, θ is the colatitude with respect to the North pole, and ϕ is east longitude). If one only considers the first term in the expansion of the magnetic potential, i.e.,

$$\Phi_M(r,\theta,\phi) = R_E(\frac{R_E}{r})^2 g_0^1(\cos\theta) = \frac{m\cos\theta}{r^2} \qquad (2.78)$$

then the axial-centred dipole approximation of the geomagnetic field is obtained, ie.

$$\mathbf{B} = \frac{2m\cos\theta}{r^3}\mathbf{e_r} + \frac{m\sin\theta}{r^3}\mathbf{e}_\theta \qquad (2.79)$$

where m is the dipole moment at the Earth's centre. In this approximation, the geographical and geomagnetic poles are coincident and the magnetic field is only a function of latitude and altitude (only $m = 0$ terms can contribute to the magnetic potential series expansion in these simulations as the model assumes that there are no variations in longitude). If variations in altitude are also ignored, the magnitude of the geomagnetic field (in gauss) can be written as

$$B(\theta) = 0.586 \times (0.25 + 0.75 \cos^2 \theta)^{0.5}. \tag{2.80}$$

2.6.3 Currents

The decomposition of ionospheric currents into components that are parallel and perpendicular to the magnetic field can provide great insights into the physics of these currents once the underlying structure and shape of the magnetic field is known. The axially-centred dipole approximation to the magnetic field currently being used in the model allows a particularly simple geomagnetic coordinate system to be adopted: **b**, which is a unit vector in the direction of **B**; ϕ, which is a unit vector parallel to the geographical longitudinal vector; and **a**, which is the unit vector that makes the resulting coordinate system orthogonal. Rewriting equation (1.54) in terms of this coordinate system yields

$$\begin{aligned} \mathbf{J} &= \sigma_{\parallel} E_b \, \mathbf{b} + \left[\sigma_P \left(E_a + \frac{v_\phi}{c} B \right) - \sigma_H \left(E_\phi - \frac{v_a}{c} B \right) \right] \mathbf{a} \\ &\quad + \left[\sigma_P \left(E_\phi - \frac{v_a}{c} B \right) + \sigma_H \left(E_a + \frac{v_\phi}{c} B \right) \right] \phi \end{aligned} \tag{2.81}$$

which can also be written in matrix form as

$$\begin{pmatrix} J_b \\ J_a \\ J_\phi \end{pmatrix} = \begin{pmatrix} \sigma_{\parallel} & 0 & 0 \\ 0 & \sigma_P & -\sigma_H \\ 0 & \sigma_H & \sigma_P \end{pmatrix} \begin{pmatrix} E_b \\ E_a + \frac{v_\phi}{c} B \\ E_\phi - \frac{v_a}{c} B \end{pmatrix}$$

Multiplying both sides of the preceding equation by the transformation matrix from geomagnetic to geographic coordinates gives

$$\begin{pmatrix} J_r \\ J_\theta \\ J_\phi \end{pmatrix} = \begin{pmatrix} \sigma_{rr} & \sigma_{r\theta} & \sigma_{r\phi} \\ \sigma_{\theta r} & \sigma_{\theta\theta} & \sigma_{\theta\phi} \\ \sigma_{\phi r} & \sigma_{\phi\theta} & \sigma_{\phi\phi} \end{pmatrix} \begin{pmatrix} E_r - \frac{w}{c} B_\theta \\ E_\theta + \frac{w}{c} B_r \\ E_\phi + \frac{1}{c}(u B_\theta - v B_r) \end{pmatrix}$$

with

$$\sigma_{rr} = \sigma_{\parallel} \sin^2 I + \sigma_P \cos^2 I \quad (2.82)$$

$$\sigma_{r\theta} = (\sigma_{\parallel} - \sigma_P) \sin I \cos I \quad (2.83)$$

$$\sigma_{r\phi} = - \sigma_H \cos I \quad (2.84)$$

$$\sigma_{\theta r} = (\sigma_{\parallel} - \sigma_P) \sin I \cos I \quad (2.85)$$

$$\sigma_{\theta\theta} = \sigma_P sin^2 I + \sigma_{\parallel} \cos^2 I \quad (2.86)$$

$$\sigma_{\theta\phi} = \sigma_H \sin I \quad (2.87)$$

$$\sigma_{\phi r} = \sigma_H \cos I \quad (2.88)$$

$$\sigma_{\phi\theta} = - \sigma_H \sin I \quad (2.89)$$

$$\sigma_{\phi\phi} = \sigma_P \quad (2.90)$$

The dip angle I is the angle between the geomagnetic field line and the local "horizontal" $\theta - \phi$ surface and is normally evaluated as $\tan(I) = 2 \cot \theta$.

Now that the currents in the ionosphere have been transformed into expressions that are compatible with the grid of the thermospheric model, one can start to consider other issues in the implementation of the model electrodynamics, i.e. how is the electric field going to be parameterized and are there going to be any restrictions placed on how the currents are modeled in the ionosphere.

If one follows the methodology of *Richmond and Matsushita* [1975], one can assume that the atmosphere is "shallow" in the sense of *Phillips* [1966], which implies that the vertical current densities in the ionosphere will be negligible when compared to the horizontal current densities in the ionosphere. This assumption is usually implemented in most numerical models by setting the radial component of the current equal to zero, which implies that

$$E_r - \frac{w}{c} B_\theta = (-1)[\sigma_{r\theta}(E_\theta + \frac{w}{c} B_r) + \sigma_{r\phi}(E_\phi + \frac{1}{c}(uB_\theta - vB_r))]/\sigma_{rr} \quad (2.91)$$

The current density expression can then be written as

$$\begin{pmatrix} J_\theta \\ J_\phi \end{pmatrix} = \begin{pmatrix} \sigma'_{\theta\theta} & \sigma'_{\theta\phi} \\ \sigma'_{\phi\theta} & \sigma'_{\phi\phi} \end{pmatrix} \begin{pmatrix} E_\theta + \frac{w}{c} B_r \\ E_\phi + \frac{1}{c}(uB_\theta - vB_r) \end{pmatrix}$$

with

$$\sigma'_{\theta\theta} = \frac{\sigma_\parallel \, \sigma_P}{\sigma_\parallel \, \sin^2 I + \sigma_P \, \cos^2 I} \tag{2.92}$$

$$\sigma'_{\theta\phi} = \frac{\sigma_\parallel \, \sigma_H \, \sin I}{\sigma_\parallel \, \sin^2 I + \sigma_P \, \cos^2 I} \tag{2.93}$$

$$\sigma'_{\phi\theta} = \frac{-\sigma_\parallel \, \sigma_H \, \sin I}{\sigma_\parallel \, \sin^2 I + \sigma_P \, \cos^2 I} \tag{2.94}$$

$$\sigma'_{\phi\phi} = \frac{\sigma_H \, \cos^2 I}{\sigma_\parallel \, \sin^2 I + \sigma_P \, \cos^2 I} + \sigma_P \tag{2.95}$$

Since $\sigma_\parallel \gg \sigma_P$ and $\sigma_\parallel \gg \sigma_H$, one can approximate the preceding conductivities as

$$\sigma'_{\theta\theta} \approx \frac{\sigma_P}{\sin^2 I} \tag{2.96}$$

$$\sigma'_{\theta\phi} \approx \frac{\sigma_H}{\sin I} \tag{2.97}$$

$$\sigma'_{\phi\theta} \approx -\frac{\sigma_H}{\sin I} \tag{2.98}$$

$$\sigma'_{\phi\phi} \approx \sigma_P \tag{2.99}$$

which now allows one to approximate the current densities as

$$J_\theta = \frac{\sigma_P}{\sin^2 I}\left(E_\theta + \frac{w}{c}B_r\right) + \frac{\sigma_H}{\sin I}\left(E_\phi + \frac{u}{c}B_\theta - \frac{v}{c}B_r\right) \tag{2.100}$$

$$J_\phi = -\frac{\sigma_H}{\sin I}\left(E_\theta + \frac{w}{c}B_r\right) + \sigma_P\left(E_\phi + \frac{u}{c}B_\theta - \frac{v}{c}B_r\right) \tag{2.101}$$

If the zonal component of the electric field is set to zero, i.e. $E_\phi = 0$, the current densities simplify to

$$J_\theta = \frac{\sigma_P}{\sin^2 I}\left(E_\theta + \frac{w}{c}B_r\right) + \frac{\sigma_H}{\sin I}\left(\frac{u}{c}B_\theta - \frac{v}{c}B_r\right) \tag{2.102}$$

$$J_\phi = -\frac{\sigma_H}{\sin I}\left(E_\theta + \frac{w}{c}B_r\right) + \sigma_P\left(\frac{u}{c}B_\theta - \frac{v}{c}B_r\right) \tag{2.103}$$

If one also assumes that the atmosphere is in hydrostatic equilibrium, then the vertical velocity u will be zero and the current density expressions can be further simplified to

$$J_\theta = \frac{\sigma_P}{\sin^2 I}\left(E_\theta + \frac{w}{c}B_r\right) - \frac{\sigma_H}{\sin I}\left(\frac{v}{c}B_r\right) \tag{2.104}$$

$$J_\phi = -\frac{\sigma_H}{\sin I}\left(E_\theta + \frac{w}{c}B_r\right) - \sigma_P\left(\frac{v}{c}B_r\right) \tag{2.105}$$

which are the expressions originally used in *Richmond and Matsushita* [1975] and *Chang and St.-Maurice* [1991]. It should be noted that the assumption of hydrostatic equilibrium in the current expressions is not consistent with the original intentions of the model, but the effects of this assumption are not anticipated to be of much consequence as the terms involving the vertical velocity u also contained B_θ, which was very small in any of the regions where the vertical velocity was large.

If the "shallow" / "thin shell" approximation is not made, then the parallel electric fields in the ionosphere must be set zero as solving equation 2.81 in a self-consistent manner is a daunting task. The easiest way to avoid this complication is to rewrite the geographic electric field components in terms of E_a, E_b, and E_ϕ, i.e.

$$\begin{pmatrix} J_r \\ J_\theta \\ J_\phi \end{pmatrix} = \begin{pmatrix} -\sigma_\parallel \sin I & \sigma_P \cos I & -\sigma_H \cos I \\ -\sigma_\parallel \cos I & -\sigma_P \sin I & \sigma_H \sin I \\ 0 & \sigma_H & \sigma_P \end{pmatrix} \begin{pmatrix} E_b \\ E_a - \frac{1}{c} wB \\ E_\phi + \frac{1}{c}(uB_\theta - vB_r) \end{pmatrix}$$

and then set E_b equal to zero, which yields

$$\begin{pmatrix} J_r \\ J_\theta \\ J_\phi \end{pmatrix} = \begin{pmatrix} \sigma_P \cos I & -\sigma_H \cos I \\ -\sigma_P \sin I & \sigma_H \sin I \\ \sigma_H & \sigma_P \end{pmatrix} \begin{pmatrix} E_a - \frac{1}{c} wB \\ E_\phi + \frac{1}{c}(uB_\theta - vB_r) \end{pmatrix}$$

If the zonal component of the electric field is again set to zero, i.e., $E_\phi = 0$, the current densities become

$$\begin{pmatrix} J_r \\ J_\theta \\ J_\phi \end{pmatrix} = \begin{pmatrix} \sigma_P \cos I & -\sigma_H \cos I \\ -\sigma_P \sin I & \sigma_H \sin I \\ \sigma_H & \sigma_P \end{pmatrix} \begin{pmatrix} E_a - \frac{1}{c} wB \\ \frac{1}{c}(uB_\theta - vB_r) \end{pmatrix}$$

This expression for the current densities is now being employed in the model. When the preceding equation is compared to the *Chang and St.-Maurice* [1991] current density expressions, there are two obvious differences: there is now a radial component to the current densities and the effects of vertical winds are being accounted for. This newer expression for the current densities should have some effect on the thermospheric response as the radial component of the current density will provide an extra acceleration term via ion-drag. The parameterization of the electric field will

be discussed in Chapter 4 as some of its characteristics will vary from simulation to simulation.

Chapter 3

Model Description II : Numerics

3.1 Introduction

The selection of an appropriate numerical scheme is usually the most involved part of the simulation process and normally requires one to consider the following issues in computational fluid dynamics (CFD):

- the mathematical behaviour of the equations
- the spatial and temporal discretizations used in the numerical scheme
- the implementation of appropriate boundary conditions
- and the generation of the numerical grid

The substantial progress made in most of these fields since 1991 suggests that most of the numerics in the model should be substantially rewritten, but time constraints have not allowed for this possibility. Despite this, several easy to implement upgrades have been incorporated into the model, enhancing its numerical stability. An overview of the numerics currently being implemented in the model, as well as possible upgrades, will be presented in the following sections of this chapter.

3.2 The mathematical behaviour of the Navier-Stokes equations

The Navier-Stokes equations are a mixed parabolic-hyperbolic set of equations, i.e. the eigenvalues of the Navier-Stokes equations are a mixture of real and complex eigenvalues. This mixed behaviour of the equations makes the choice of a numerical scheme particularly daunting as the numerical scheme must be able to cope with

the parabolic behaviour in certain regions of the model domain and the hyperbolic behaviour in the remaining regions of the model domain. The hyperbolic behaviour of the solution can be especially difficult to simulate as the solution in a hyperbolic region can develop strong shocks/steep gradients, which may introduce numerical oscillations into the solution. These numerical oscillations are normally reduced by

- using the *conservative* form of the Navier-Stokes equations, and
- introducing small amounts of artificial viscosity into the numerical scheme.

It should be emphasized that the conservative form of the Navier-Stokes equations does not add any new physics to the simulation; it is just a form of the Navier-Stokes equations that is more stable numerically as the solution variables are *physically conserved* quantities, i.e. mass, momentum, and energy.

This stability can be most easily illustrated by considering the propagation of a normal 1-D shock, as presented in Figure 3.1. The conservative and non-conservative forms of the N-S equations for such a situation are

$$\frac{\partial \rho}{\partial t} + \frac{\partial (\rho u)}{\partial r} = 0 \tag{3.1}$$

$$\frac{\partial (\rho u)}{\partial t} + \frac{\partial (\rho u^2 + p)}{\partial r} = 0 \tag{3.2}$$

and

$$\frac{\partial \rho}{\partial t} + \rho \frac{\partial u}{\partial r} + u \frac{\partial \rho}{\partial r} = 0 \tag{3.3}$$

$$\frac{\partial u}{\partial t} + u \frac{\partial u}{\partial r} = -\frac{\partial p}{\partial r} \tag{3.4}$$

respectively. If one attempts to determine the velocity u using the non-conservative form

$$\frac{\partial u}{\partial t} + u \frac{\partial u}{\partial r} = -\frac{\partial p}{\partial r}$$

the discontinuity in u will be amplified by the numerical evaluation of the $u \frac{\partial u}{\partial r}$ and $\frac{\partial p}{\partial r}$ terms. The conservative form

$$\frac{\partial (\rho u)}{\partial t} + \frac{\partial (\rho u^2 + p)}{\partial r} = 0$$

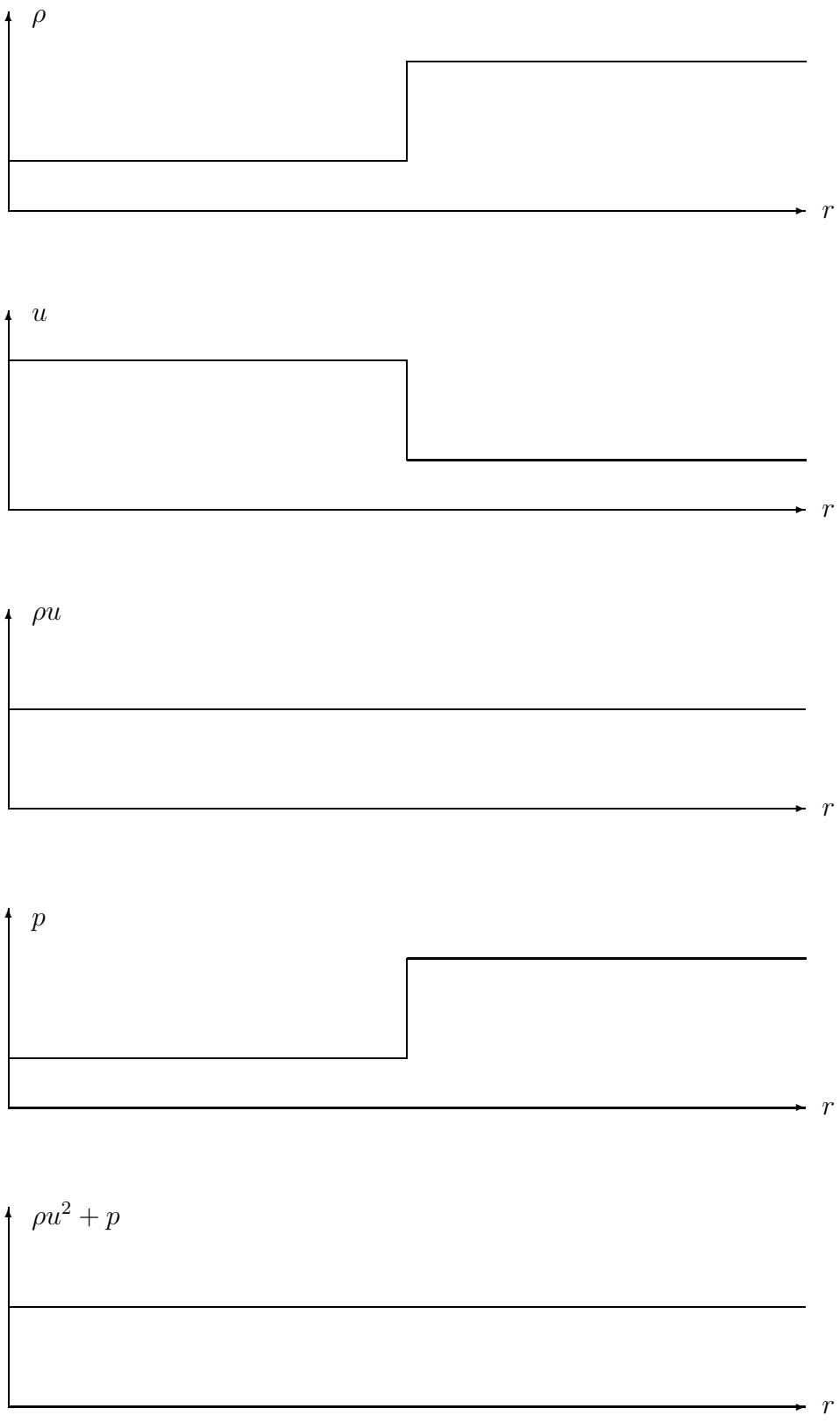

Figure 3.1: Variation of flow properties across a normal shock wave. (From *Anderson*, 1995.)

on the other hand, will not be influenced by any of the discontinuities in ρ or u as the quantities ρu and $\rho u^2 + p$ are continuous across the shock. A more thorough discussion of this topic can be found in the text *Computational Fluid Dynamics* by J. D. Anderson. Another important, but frequently overlooked property of the conservative form is that the locations and speeds of the shocks are accurately captured; non-conservative forms tend to propagate shocks at slower speeds [*Roache*, 1975].

If one now substitutes the spherical form of the divergence and gradient operators into equations (2.1)-(2.3), the Navier-Stokes equations become

$$\frac{\partial \mathbf{U}}{\partial t} + \frac{\partial \mathbf{F}}{\partial r} + \frac{2\mathbf{F}}{r} + \frac{1}{r}\frac{\partial \mathbf{G}}{\partial \theta} + \frac{\mathbf{G}}{r}cot\theta + \frac{1}{r\sin\theta}\frac{\partial \mathbf{H}}{\partial \phi} = \mathbf{S} \tag{3.5}$$

where $\mathbf{U}$ is a vector containing the solution variables, $\mathbf{F}$, $\mathbf{G}$, and $\mathbf{H}$ are the components of the solution vector flux and $\mathbf{S}$ is a vector that contains the sources and sinks associated with the solution vector $\mathbf{U}$, i.e.

$$\mathbf{U} = \begin{pmatrix} \rho \\ \rho u \\ \rho v \\ \rho w \\ e \end{pmatrix} \tag{3.6}$$

$$\mathbf{F} = \begin{pmatrix} \rho u \\ \rho u^2 + p - \tau_{rr} \\ \rho uv - \tau_{r\theta} \\ \rho uw - \tau_{r\phi} \\ (e+p)u - u\tau_{rr} - v\tau_{r\theta} - w\tau_{r\phi} - \lambda\frac{\partial T}{\partial r} \end{pmatrix} \tag{3.7}$$

$$\mathbf{G} = \begin{pmatrix} \rho v \\ \rho uv - \tau_{\theta r} \\ \rho v^2 + p - \tau_{\theta\theta} \\ \rho vw - \tau_{\theta\phi} \\ (e+p)v - u\tau_{\theta r} - v\tau_{\theta\theta} - w\tau_{\theta\phi} - \frac{\lambda}{r}\frac{\partial T}{\partial \theta} \end{pmatrix} \tag{3.8}$$

$$\mathbf{H} = \begin{pmatrix} \rho w \\ \rho uw - \tau_{r\phi} \\ \rho vw - \tau_{\theta\phi} \\ \rho w^2 + p - \tau_{\phi\phi} \\ (e+p)w - u\tau_{r\phi} - v\tau_{\theta\phi} - w\tau_{\phi\phi} - \frac{\lambda}{rsin\theta}\frac{\partial T}{\partial \phi} \end{pmatrix} \tag{3.9}$$

$$\mathbf{S} = \begin{pmatrix} 0 \\ \rho f_r + 2\frac{p}{r} + \rho\frac{v^2}{r} + \rho\frac{w^2}{r} - \frac{\tau_{\theta\theta}}{r} - \frac{\tau_{\phi\phi}}{r} \\ \rho f_\theta + \frac{p}{r}\cot\theta - \rho\frac{uv}{r} + \rho\frac{w^2}{r}\cot\theta + \frac{\tau_{r\theta}}{r} - \frac{\tau_{\phi\phi}}{r}\cot\theta \\ \rho f_\phi - \rho\frac{uw}{r} - \rho\frac{vw}{r}\cot\theta + \frac{\tau_{\phi r}}{r} + \frac{\tau_{\theta\phi}}{r}\cot\theta \\ \rho u f_r + \rho v f_\theta + \rho w f_\phi + Q - C \end{pmatrix} \tag{3.10}$$

Although the preceding form of the Navier-Stokes equations is physically conservative, the numerical efficiency of this particular form is very poor (5 operations have to be performed on the flux vector to advance the solution). Multiplying equation (3.5) by $r^2 \sin\theta$ and combining some like terms gives

$$\frac{\partial}{\partial t}\left[\,(r^2\sin\theta)\,\mathbf{U}\,\right] + \frac{\partial}{\partial r}\left[\,(r^2\sin\theta)\,\mathbf{F}\,\right] + \frac{\partial}{\partial\theta}\left[\,(r\sin\theta)\,\mathbf{G}\,\right] + \frac{\partial}{\partial\phi}\left[\,(r^2\sin\theta)\,\mathbf{H}\,\right] = \left[\,(r^2\sin\theta)\,\mathbf{S}\,\right] \tag{3.11}$$

This form is much more efficient numerically (only 3 operations have to be performed on the flux vector to advance the solution now). It should be emphasized that this form of the Navier-Stokes equations is still physically conservative; the $r\sin\theta$ and $r^2\sin\theta$ terms are only geometry factors that have no affect on the "physics" of the simulations.

If development time and CPU resources were not an issue, equation (3.11) would be the starting point of this thesis. However, such 3-D simulations are still not practical on current workstations and zonal symmetry must be invoked, i.e. any derivatives with respect to ϕ are set to zero. Equation (3.11) becomes

$$\frac{\partial}{\partial t}\left[\,(r^2\sin\theta)\,\mathbf{U}\,\right] + \frac{\partial}{\partial r}\left[\,(r^2\sin\theta)\,\mathbf{F}\,\right] + \frac{\partial}{\partial\theta}\left[\,(r\sin\theta)\,\mathbf{G}\,\right] = \left[\,(r^2\sin\theta)\,\mathbf{S}\,\right]$$

and is the form of the Navier-Stokes equations currently being used in this thesis.

3.3 Grid Discretization

Grid discretization (which is also known as grid generation) is the process of decomposing the physical space that one is modeling into smaller regions/computational cells. Determining the shape and size of these computational cells is an important issue in the design of a numerical model as the stability and accuracy of the numerical scheme being implemented in the model is usually determined by how accurately the variables are represented in the computational domain. However, one must also balance this desire for accuracy with the computational time and/or power available to the modeler (doubling the number of points used in a finite difference scheme typically quadruples the runtime of the program).

Since the spherical coordinate system is an orthogonal coordinate system, one obvious choice in the grid discretization procedure would be to create altitude and latitude grids with uniform spacings, e.g. $\Delta z = 1\text{km}$, $\Delta\theta = 0.1^{\text{o}}$. Unfortunately, such a choice would not be very efficient numerically as very different resolutions are required in the upper and lower regions of the model (a fairly small Δz is needed at the lower altitudes to correctly resolve the changes that are occurring in most of the thermospheric variables at these altitudes while a much coarser Δz is required at the higher altitudes to maintain numerical stability). The original grid in the model tried to accommodate these conflicting requirements by using an unequally spaced altitude grid with a minimum cell size of 2.5 km at the lower boundary and a maximum cell size of 25 km at the upper boundary. An unequally spaced grid was also employed in latitude to resolve features in the electrojet and polar regions of the model more clearly (more details on the grid used in the original version of the model can be found in Table 2 of *Chang and St.-Maurice* [1991]).

Although unequally spaced grids can increase the resolution in a particular region of a model, they can also degrade the stability and accuracy of the the numerical model if the changes in the cell sizes/grid spacing becomes too large. To overcome this, the governing equations can be transformed from 'physical space' to a 'computational space' where the grid spacing is uniform. This transformation from 'physical

Function	Grid : x = x(ξ)
Exponential	$\frac{\exp(a\xi)-1}{\exp(a)-1}$
Hyperbolic Tangent	$1-\frac{\tanh(\alpha(1-\xi)))}{\tanh(\alpha)}$
Hyperbolic Sine	$\frac{\sinh(\alpha\xi)}{\sinh(\alpha)}$
Error Function	$1-\frac{\mathrm{erf}(\alpha(1-\xi))}{\mathrm{erf}(\alpha)}$
Tangent ($0 \leq \alpha \leq \frac{\pi}{2}$)	$\frac{\tan(\alpha\xi)}{\tan(\alpha)}$
Arctangent	$1-\frac{\arctan(\alpha(1-\xi))}{\arctan(\alpha)}$
Sine ($0 \leq \alpha \leq \frac{\pi}{2}$)	$1-\frac{\sin(\alpha(1-\xi))}{\sin(\alpha)}$
Logarithm	$1-\frac{\ln(1+\alpha(1-\xi))}{\ln(1+\alpha)}$
Inverse Hyperbolic Tangent ($0 \leq \alpha \leq \frac{\pi}{2}$)	$\frac{\mathrm{arctanh}(\alpha(1-\xi))}{\mathrm{arctanh}(\alpha)}$
Quadratic ($0 \leq \alpha \leq 1$)	$(1-\alpha)\xi+\alpha\xi^2$

Table 3.1: Common grid stretching functions.

space' to 'computational space' can be accomplished by a variety of analytical functions or by solving anther system of partial differential equations (a more complete discussion of grid generation techniques can be found in the texts *Numerical Grid Generation: Foundations and Applications* and *Grid Generation Methods*). This requirement for extra resolution in the boundary region of a model is fairly common in the simulation of viscous flow, which has led to the proposal of many different stretching functions. Some of the more commonly used functions are presented in Table 3.1. *Thompson and Mastin* [1984] have analyzed the truncation errors and grid point distributions associated with most of these functions and concluded that the hyperbolic sine, hyperbolic tangent, and exponential functions are usually the best-suited functions for resolving boundary-layer type phenomena.

Since computing power has increased dramatically since 1991, it was decided that the vertical discretization of the model should be updated with a newer 121 point grid based on one of the stretching functions recommended by *Thompson and Mastin* [1984]. While the hyperbolic tangent function is usually considered to be the best

overall function, it has one drawback: the cell sizes at the upper and lower boundaries must satisfy the criterion

$$\Delta z_{\text{top}} \Delta z_{\text{bottom}} < (\ H/(N-1))^2$$

where H is the altitude range covered by the discretization and N is the number of points used in the discretization. This means that as one increases the number of points used in the discretization, one must make the product of $\Delta z_{\text{top}}\, \Delta z_{\text{bottom}}$ smaller, limiting the ability of this function to generate stretched grids with large differences between Δz_{top} and Δz_{bottom} ($\Delta z_{\text{top}}\, \Delta z_{\text{bottom}} < 9.507\, km$ for $H = 370$ km and $N = 121$ pts). The hyperbolic sine function, on the other hand, requires the cell sizes at the upper and lower boundaries to satisfy the condition

$$\Delta z_{\text{top}}\, \Delta z_{\text{bottom}} > (H/(N-1))^2$$

which provides much more flexibility in designing stretched grids with a large difference between Δz_{top} and Δz_{bottom}.

After extensive experimentation with the grid parameters, 4 of the most suitable grids were chosen for further analysis. Plots of the cell sizes and of the cell size changes for these grids can be found in Figures 3.2 and 3.3 respectively. The high resolution grid based on the hyperbolic tangent function generates an almost ideal cell size distribution, with minimal changes in the cell sizes and resolution where it is needed. However, it would also be the most expensive to implement numerically as its upper boundary cell size of 9.5 km would reduce the time step of the model substantially. The lower resolution hyperbolic tangent function based grid is also adept at minimizing the changes in the cell sizes, but clusters too many grid points at the lower altitudes (recall that $\Delta z_{\text{top}}\, \Delta z_{\text{bottom}} < 9.507\,\text{km}$). The grid based on the hyperbolic sine function also prevents the cell sizes from changing too quickly, but clusters too many grid points at the lower altitudes like the lower resolution hyperbolic tangent based grid. The exponential based grid, on the other hand, provides resolution that is comparable to the high resolution hyperbolic tangent based grid to heights of 300 km, the principal region of interest in this model. However, the

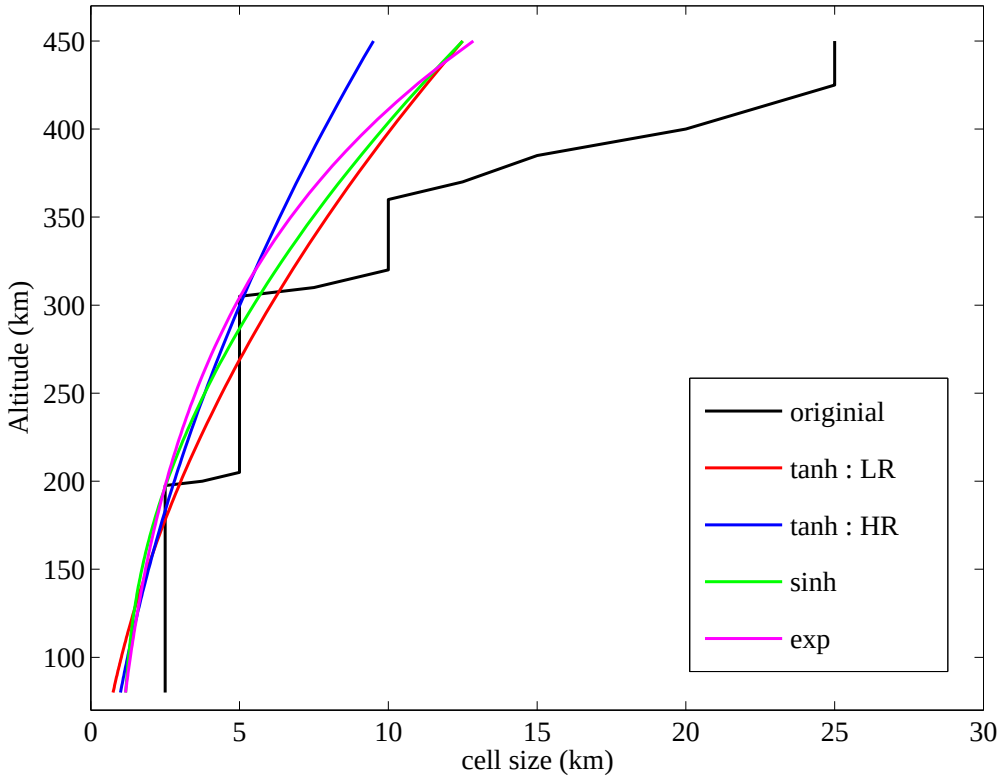

Figure 3.2: Grid cell sizes for selected grids. The high resolution hyperbolic tangent grid (blue curve) offers the best spatial resolution but would require the most CPU time. The exponential grid (magenta curve) offers better spatial resolution than the low resolution hyperbolic tangent (red curve) and hyperbolic sine (green curve) grids below 300 km and would require the same amount of CPU time as these lower resolution grids. The large changes in the cell sizes of the original grid (black curve) are quite evident.

exponential based grid allows the cell sizes to change most quickly, but these rapid changes in cell size are confined to the upper altitudes of the model, where the variables are not expected to change that quickly. Despite this apparent weakness, the time-step associated with the exponential grid is comparable to the hyperbolic sine and low resolution hyperbolic tangent grids as the upper boundary cell sizes of these three grids are all quite comparable. Because of this desired resolution and larger time-step (when compared to the high resolution hyperbolic tangent grid), the grid based on the exponential function was chosen as the new vertical grid for the model.

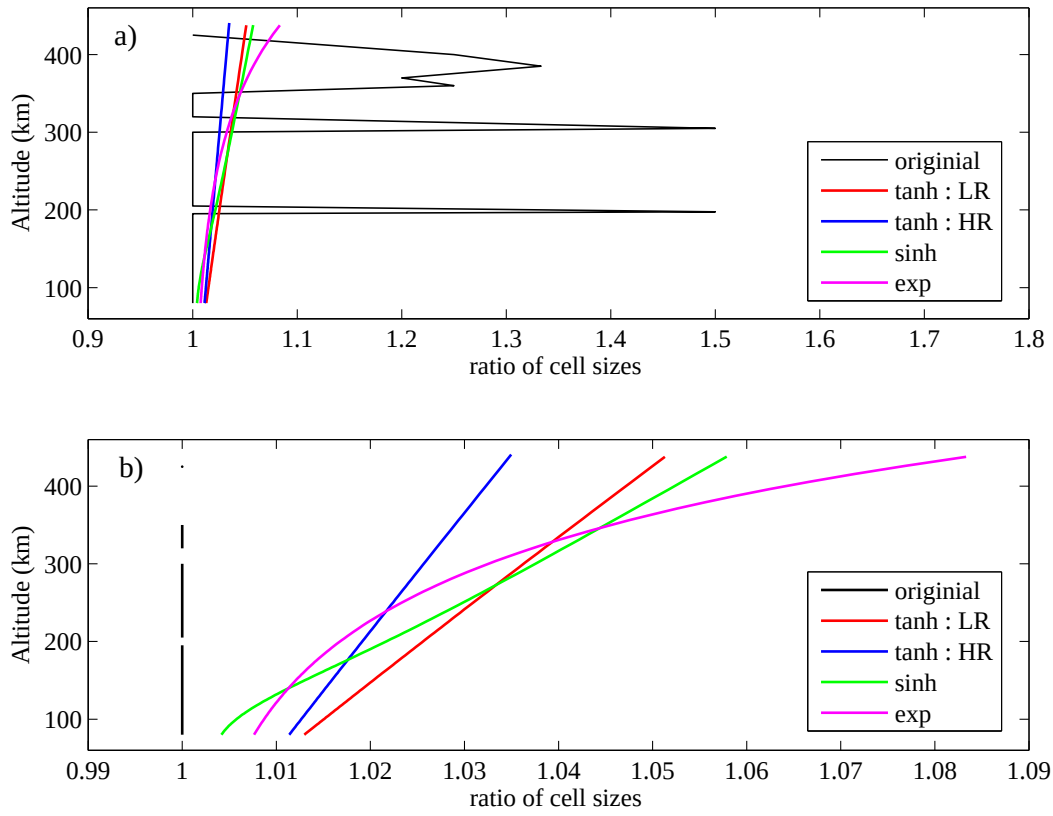

Figure 3.3: Grid cell size ratios for selected grids. The changes in the grid cell size of the high resolution hyperbolic tangent grid (blue curve) are never larger than 3.5% while changes in the grid cell size of the exponential grid (magenta curve) can exceed 8% near the upper boundary. The changes in the grid cell sizes of the low resolution hyperbolic tangent (red curve) and hyperbolic sine (green curve) grids fall in between these two extremes.

Now that the stretching function for the vertical grid has been determined, the mechanics of transforming the equations from physical space to the computational space will be discussed. The chain rule of calculus allows one to write the derivative with respect to altitude as

$$\frac{\partial}{\partial r} = \frac{\partial \xi}{\partial r}\frac{\partial}{\partial \xi} \tag{3.12}$$

which now lets one rewrite the Navier-Stokes equations as

$$\frac{\partial}{\partial t}\left[\,(r^2 \sin\theta)\,\mathbf{U}\,\right] + \frac{\partial \xi}{\partial r}\frac{\partial}{\partial \xi}\left[\,(r^2 \sin\theta)\,\mathbf{F}\,\right] + \frac{\partial}{\partial \theta}\left[\,(r \sin\theta)\,\mathbf{G}\,\right] = (r^2 \sin\theta)\,\mathbf{S}. \tag{3.13}$$

Stretching functions for the latitudinal grid have also been investigated, with the hyperbolic sine formulation showing some promise. However, time constraints did not allow these changes to be incorporated into the model.

3.4 Non-dimensionalization of the equations

Now that the appropriate form of the Navier-Stokes equations have been decided upon, one must then decide if these equations should be put into a non-dimensionalized form, i.e. if the variables should be divided by an appropriate/characteristic value pertinent to that variable. The most obvious benefit of this non-dimensionalization procedure is that all of the numerical values of the variables are now much closer to one, which should reduce the possibility of a numerical overflow/underflow occurring. However, this non-dimensionalization also makes the debugging of the program much more difficult as the numbers are no longer meaningful 'physically'. Despite this possible drawback, the disparate magnitudes of some variables in the model dictated that non-dimensionalization was a necessity (typical values of r are $\sim 10^8$ cm and the mass densities near the upper boundary can be as small as $\sim 10^{-20}$ g/cm^3).

The selection of the characteristic values is usually at the discretion of the modeler. In the original version of the model, the characteristic values were defined in terms of three fundamental quantities :

$$\begin{aligned} \text{length}\,[\,\mathrm{L}\,] &= 10^7 \text{ cm}, \\ \text{mass}\,[\,\mathrm{M}\,] &= 10^{10} \text{ gram}, \\ \text{time}\,[\,\mathrm{t}\,] &= 10^3 \text{ seconds}. \end{aligned}$$

Any other characteristic value could then be derived from these three quantities by dimensional analysis, e.g.

$$\begin{aligned} \text{velocity}\,[\,\mathrm{LT}^{-1}\,] &= 1.0 \times 10^4 \,\text{cm sec}^{-1} \\ \text{density}\,[\,\mathrm{ML}^{-3}\,] &= 1.0 \times 10^{-11} \,\text{g cm}^{-3} \end{aligned}$$

This non-dimensionalization is still being used in the model.

3.5 Numerical Technique

The original version of the code solved the Navier-Stokes equations with an explicit MacCormack scheme [*MacCormack*, 1969]. This scheme is a predictor-corrector method with second-order accuracy in time and space. In the predictor step, a temporary solution at the next time step is obtained by replacing the temporal and spatial derivatives with first order forward differences, i.e.

$$\begin{aligned}\overline{\mathbf{U}}[n+1,i,j] &= \mathbf{U}[n,i,j] - \left(\frac{\Delta t}{\Delta \xi}\right)\cdot(\mathbf{F}[n,i+1,j]-\mathbf{F}[n,i,j]) \\ &\quad - \left(\frac{\Delta t}{\Delta \theta}\right)\cdot(\mathbf{G}[n,i,j+1]-\mathbf{G}[n,i,j]) + \Delta t\cdot\mathbf{S}[n,i,j]\end{aligned} \tag{3.14}$$

where $\mathbf{U}$, $\mathbf{F}$, $\mathbf{G}$ and $\mathbf{S}$ are now defined as

$$\mathbf{U}[n,i,j] = (r^2\sin\theta)\,\mathbf{U}[n,i,j] \tag{3.15}$$

$$\mathbf{F}[n,i,j] = (r^2\sin\theta)\left(\frac{\partial \xi}{\partial r}\right)\mathbf{F}[n,i,j] \tag{3.16}$$

$$\mathbf{G}[n,i,j] = (r\sin\theta)\,\mathbf{G}[n,i,j] \tag{3.17}$$

$$\mathbf{S}[n,i,j] = (r^2\sin\theta)\,\mathbf{S}[n,i,j] \tag{3.18}$$

This estimate of the solution at the next time level is then averaged with the solution at the current time step to form a solution at the intermediate time step $n+\frac{1}{2}$, i.e.

$$\mathbf{U}[n+\frac{1}{2},i,j] = \frac{1}{2}(\,\overline{\mathbf{U}}[n+1,i,j] + \mathbf{U}[n,i,j]\,).$$

The solution at the intermediate time step $n+\frac{1}{2}$ is then advanced to the next time level by replacing the temporal and spatial derivatives with first order backward differences, i.e.

$$\begin{aligned}\mathbf{U}[n+1,i,j] &= \mathbf{U}[n+\frac{1}{2},i,j] - \frac{1}{2}\left(\frac{\Delta t}{\Delta \xi}\right)(\overline{\mathbf{F}}[n+1,i,j]-\overline{\mathbf{F}}[n+1,i-1,j]) \\ &\quad - \frac{1}{2}\left(\frac{\Delta t}{\Delta \theta}\right)(\,\overline{\mathbf{G}}\,[n+1,i,j] - \overline{\mathbf{G}}\,[n+1,i,j-1]\,) \\ &\quad + \left(\frac{\Delta t}{2}\right)\cdot\overline{\mathbf{S}}\,[n+1,i,j])\end{aligned} \tag{3.19}$$

While this presentation of the MacCormack scheme provides some insight into how the scheme was derived, it is not very efficient numerically. It has been implemented in the model in the following manner

$$\begin{aligned}
\overline{\mathbf{U}}\,[n+1,i,j] &= \mathbf{U}[n,i,j] - \left(\frac{\Delta t}{\Delta \xi}\right)\cdot(\mathbf{F}[n,i+1,j] - \mathbf{F}[n,i,j]) \\
&\quad - \left(\frac{\Delta t}{\Delta \theta}\right)\cdot(\,\mathbf{G}\,[n,i,j+1] \,-\, \mathbf{G}\,[n,i,j]\,) \,+\, \Delta t\,\cdot\mathbf{S}[n,i,j] \\
\overline{\overline{\mathbf{U}}}\,[n+1,i,j] &= \overline{\mathbf{U}}[n+1,i,j] - \left(\frac{\Delta t}{\Delta \xi}\right)\cdot(\overline{\mathbf{F}}[n+1,i,j] - \overline{\mathbf{F}}[n+1,i-1,j]) \\
&\quad - \left(\frac{\Delta t}{\Delta \theta}\right)\cdot(\,\overline{\mathbf{G}}\,[n+1,i,j] \,-\, \overline{\mathbf{G}}\,[n+1,i,j-1]\,) \\
&\quad + \left(\frac{\Delta t}{2}\right)\cdot\overline{\mathbf{S}}\,[n+1,i,j]) \\
\mathbf{U}[n+1,i,j] &= \frac{1}{2}\left(\mathbf{U}[n,i,j] + \overline{\overline{\mathbf{U}}}\,[n+1,i,j]\right)
\end{aligned} \tag{3.20}$$

It should be noted that the forward-backward differencing of the spatial derivatives is not sacrosanct; second order accuracy can also be obtained if backwards differences are used in the spatial derivatives of the predictor step and forwards differences are used in the spatial derivatives of the corrector step. To avoid any type of biasing due to the one-sided differencing of the spatial derivatives, the forwards and backwards differencing of the spatial derivatives was alternated between the predictor and the corrector steps and between the r and θ derivatives. A summary of the differentiation sequence can be found in Table 1 of *Chang and St.-Maurice* [1991].

When viscous terms are present in $\mathbf{F}$ and $\mathbf{G}$, the spatial derivatives of the viscous terms must be differenced correctly to maintain the second order accuracy of the MacCormack scheme [*Tannehill et al.*, 1997]. This was accomplished in the following manner. The r-derivative terms appearing in $\mathbf{F}$ are differenced in the opposite direction to that used for $\frac{\partial \mathbf{F}}{\partial r}$, while the θ derivatives are approximated with central differences. Similarly, the θ derivative terms appearing in $\mathbf{G}$ are differenced in the opposite direction to that used for $\frac{\partial \mathbf{G}}{\partial \theta}$, while the r derivatives are approximated with central differences. This sequence can also be found in Table 1 of *Chang and St.-Maurice* [1991].

The robustness, programming simplicity, and second order accuracy of MacCor-

mack's original 1969 scheme led to its widespread adoption in the CFD community [*Anderson*, 1995]. While the second order accuracy of the original 1969 scheme was adequate for most simulations, the simulation of boundary layers, shear layers, and hypersonic flows required much more accurate algorithms to reproduce the observed flow patterns. *Eli and Turkel* [1974] were the first to increase the accuracy of MacCormack's original scheme by increasing the spatial accuracy of the inviscid terms to 4th order. *Carpenter* [1984] then extended MacCormack's original scheme to a true 2-4 (2nd order time, 4th order space) scheme by replacing the inviscid derivatives with 4th order compact differences and the viscous terms with third order upwind finite differences. There have been many other 'extended' MacCormack schemes proposed since 1984, with almost all of these schemes tailoring the treatment of the spatial derivatives to the task at hand: stability, resolution of smaller scale phenomena, low dispersion errors/long time integration, etc. The possibility of incorporating one of these extended MacCormack schemes into the model should be seriously considered in any future upgrades.

3.6 Smoothing/Filtering/Artificial Dissipation

In most simulations of nonlinear phenomena, there is a tendency for the smaller scale features on the computational mesh to grow with time, especially in regions of steep gradients. This spurious growth of the shorter wavelength features can allow positive definite quantities like the mixing ratios of individual species to become negative, which is clearly unacceptable in any physically realistic simulation. Even more disconcerting is the fact that these numerical oscillations sometimes grow to the point where they dominate the entire simulation, creating regions of negative temperatures and pressures in the solution (the solution has now "blown up"in CFD terms).

The most common way of incorporating artificial viscosity into Mac-Cormack-

based schemes is to add the artificial viscosity at each step of the simulation, i.e.,

$$\begin{aligned}
\overline{\mathbf{U}}\,[n+1,i,j] &= \mathbf{U}[n,i,j] - \left(\frac{\Delta t}{\Delta \xi}\right)\cdot(\mathbf{F}[n,i+1,j]-\mathbf{F}[n,i,j]) \\
&\quad - \left(\frac{\Delta t}{\Delta \theta}\right)\cdot(\,\mathbf{G}\,[n,i,j+1]\;-\;\mathbf{G}\,[n,i,j]\,) \\
&\quad + \Delta t\;\cdot\mathbf{S}[n,i,j]\;+\;\mathbf{D}\,[n,i,j] \qquad (3.21)\\
\overline{\overline{\mathbf{U}}}\,[n+1,i,j] &= \overline{\mathbf{U}}[n+1,i,j] - \left(\frac{\Delta t}{\Delta \xi}\right)\cdot(\overline{\mathbf{F}}[n+1,i,j]-\overline{\mathbf{F}}[n+1,i-1,j]) \\
&\quad - (\frac{\Delta t}{\Delta \theta})\cdot(\,\overline{\mathbf{G}}\,[n+1,i,j]\;-\;\overline{\mathbf{G}}\,[n+1,i,j-1]\,) \\
&\quad + \left(\frac{\Delta t}{2}\right)\cdot\overline{\mathbf{S}}\,[n+1,i,j])\;+\;\overline{\mathbf{D}}\,[n+1,i,j]) \qquad (3.22)
\end{aligned}$$

where

$$\begin{aligned}
\mathbf{D}\,[n,i,j] &= C_r\,\frac{|\,p[n,i+1,j]-2p[n,i,j]+p[n,i-1,j]\,|}{p[n,i+1,j]+2p[n,i,j]+p[n,i-1,j]} \\
&\quad\times(\mathbf{U}[n,i+1,j]-2\mathbf{U}[n,i,j]+\mathbf{U}[n,i-1,j]) \\
&\quad + C_\theta\,\frac{|\,p[n,i,j+1]-2p[n,i,j]+p[n,i,j-1]\,|}{p[n,i,j+1]+2p[n,i,j]+p[n,i,j-1]} \\
&\quad\times(\mathbf{U}[n,i,j+1]-2\mathbf{U}[n,i,j]+\mathbf{U}[n,i,j-1])
\end{aligned}$$

$$\begin{aligned}
\overline{\mathbf{D}}\,[n,i,j] &= C_r\,\frac{|\,\overline{p}[n,i+1,j]-2\overline{p}[n,i,j]+\overline{p}[n,i-1,j]\,|}{\overline{p}[n,i+1,j]+2\overline{p}[n,i,j]+\overline{p}[n,i-1,j]} \\
&\quad\times(\overline{\mathbf{U}}[n,i+1,j]-2\overline{\mathbf{U}}[n,i,j]+\overline{\mathbf{U}}[n,i-1,j]) \\
&\quad + C_\theta\,\frac{|\,\overline{p}[n,i,j+1]-2\overline{p}[n,i,j]+\overline{p}[n,i,j-1]\,|}{\overline{p}[n,i,j+1]+2\overline{p}[n,i,j]+\overline{p}[n,i,j-1]} \\
&\quad\times(\overline{\mathbf{U}}[n,i,j+1]-2\overline{\mathbf{U}}[n,i,j]+\overline{\mathbf{U}}[n,i,j-1])
\end{aligned}$$

[*Anderson*, 1995]. This method of adding artificial viscosity also has to be tuned for each type of simulation (via the C_r and C_θ parameters) and is activated when pressure oscillations are detected.

The artificial viscosity methods considered so far are 'naive methods' that smear any discontinuities/shocks over several computational cells and ignore any physics that may already be known about the governing equations. Overcoming these limitations has been an active area of research in the CFD community for the last thirty-five years and has led to the development of algorithms that add artificial viscosity so intelligently that shocks and discontinuities can now be resolved within one or two

computational cells without oscillations. These newer, more intelligent schemes can usually be incorporated into existing MacCormack schemes by adding the correct amount of artificial viscosity after the corrector step, i.e.

$$\mathbf{U}[n+1,i,j] = \frac{1}{2}(\mathbf{U}[n,i,j] + \overline{\mathbf{U}}[n+1,i,j]) + (\mathbf{D}_{i+\frac{1}{2},j+\frac{1}{2}} - \mathbf{D}_{i-\frac{1}{2},j-\frac{1}{2}}). \qquad (3.23)$$

Most of these schemes determine the "correct" amount of artificial viscosity, $\mathbf{D}_{i+\frac{1}{2},j+\frac{1}{2}} - \mathbf{D}_{i-\frac{1}{2},j-\frac{1}{2}}$ by imposing monotonicity and positivity constraints on the numerical fluxes in the computational cells (i.e., the *convective* portion of the algorithm cannot allow maxima to increase, minima to decrease, or new extremum to be created). Some of these schemes also account for the "flow" of information in the solution by constructing solutions to the Euler system of equations, which ignore the effects of viscosity and thermal conduction. Despite this fact, these schemes often generate solutions that are often superior to those obtained with Navier-Stokes schemes utilizing naive finite difference schemes [*Toro*, 1999].

The second way of dealing with the spurious growth of shorter wavelength features on the computational grid is to filter these features out of the solution with an appropriately designed filter/smoother. One of the most commonly used filters/smoothers is the one dimensional 3 point operator

$$\mathbf{U}_f[n,i,j] = (1-\hat{A})\mathbf{U}[n,i,j] + \frac{\hat{A}}{2}(\mathbf{U}[n,i+1,j] \; - \; \mathbf{U}[n,i-1,j]\,)$$

where $\hat{A}$ is a constant that determines the strength of the damping rate (a more thorough discussion of this filter can be found in *Haltiner and Williams* [1980]). Increasing the number of points employed in the filter/smoother allows one to control the range of wavenumbers dissipated by the filter much more effectively, with filter lengths of 10 - 20 points being typical. There are also more sophisticated filters that identify where the shorter scale oscillations are occurring and then smooth these regions only, e.g. *Forester* [1977]. However, these more sophisticated filters are rarely used in practice as the computational overhead associated with them can easily double the runtime of a simulation.

When the original version of the code was written, it was decided that artificial viscosity would not be incorporated into the predictor and corrector steps; all shorter

wavelength phenomena would be removed by applying the one-dimensional filter

$$f_f[n,i] \;=\; 0.5\,(f[n,i+1] + f[n,i-1]),\;\; i = 2,4,6,... \tag{3.24}$$

as most of the aforementioned intelligent artificial viscosity schemes were still considered to be 'cutting-edge' numerics (*St.-Maurice*, private communication). Since the temperature, density, and pressure exhibit significant variations with height, these variables can only be smoothed latitudinally; all of the other variables are smoothed in both directions. The incorporation of a more up-to-date scheme should also be considered in any future upgrades, but time constraints have not allowed this endeavour to be considered.

3.7 Boundary Conditions

The 2 dimensionality of the model dictates that boundary conditions have to be applied at the upper and lower boundaries of the model, as well as at the equatorward and poleward edges of the grid. The boundary conditions at the lower boundary are the easiest to implement: all of the variables at the lower boundary have been assumed to be constant throughout the simulation, i.e.,

$$\begin{aligned} u &= v = w = 0 \\ T &= 221.3 \text{ K} \\ \rho &= 1.306 \times 10^{-8} \text{ g/cm}^3 \\ p &= 8.304 \text{ dynes/cm}^2. \end{aligned}$$

This choice is consistent with having a large pool of matter at the lowest altitudes that will absorb any disturbance coming from above [*Chang and St.-Maurice*, 1991]. The upper boundary conditions are much more problematic as these boundary conditions must provide a physically realistic background and cope with the waves being generated in the model. To accomplish this, the radial derivatives of the neutral winds and temperature were set to zero through a first order extrapolation, i.e.,

$$u(n, I-1, j) = u(n, I, j)$$

$$v(n, I-1, j) = v(n, I, j)$$

$$w(n, I-1, j) = w(n, I, j)$$

$$T(n, I-1, j) = T(n, I, j)$$

where I is the index of the upper boundary cell in the r direction and $I-1$ is the index of the cell immediately beneath the upper boundary cell in the r direction. These conditions are consistent with the assumption that there is no heat flow or diffusion of momentum at the upper boundary. The neutral pressures and densities at the upper boundary were assumed to satisfy hydrostatic equilibrium. The original version of the model implemented this by setting the ratio

$$\rho\,(n, I-1, j)/\rho\,(n, I, j)$$

equal to the *same* constant for the entire simulation. This has now been changed to allow neutral temperature changes at the upper boundary to be incorporated into the assumption of hydrostatic equilibrium.

3.8 Determination of the Initial Conditions

The initial conditions of the model were obtained by setting all of the meridional derivatives equal to zero and then allowing the model to run until the temperature changes between two consecutive time-steps at *every* altitude was $< 10^{-6}$ K. The resulting temperature and number density profiles were then substituted back into the 2-D model and allowed to run with a zero electric field for 3 hours. No appreciable differences were noted between the initial and final profiles, verifying the validity of the initial condition profiles. It should be emphasized that the initial conditions are intended to provide a reasonably realistic background; no attempts have been made to 'tune' the model so that it will reproduce climatological mean values.

Profiles of the new mixing ratios are shown in Figures 3.4-3.6 and are much more realistic as the atomic oxygen mixing ratios are now much more reasonable in the lower thermosphere and the 'discontinuities' in the atomic oxygen and molecular nitrogen mixing ratios at 300 km have been eliminated. The initial neutral density

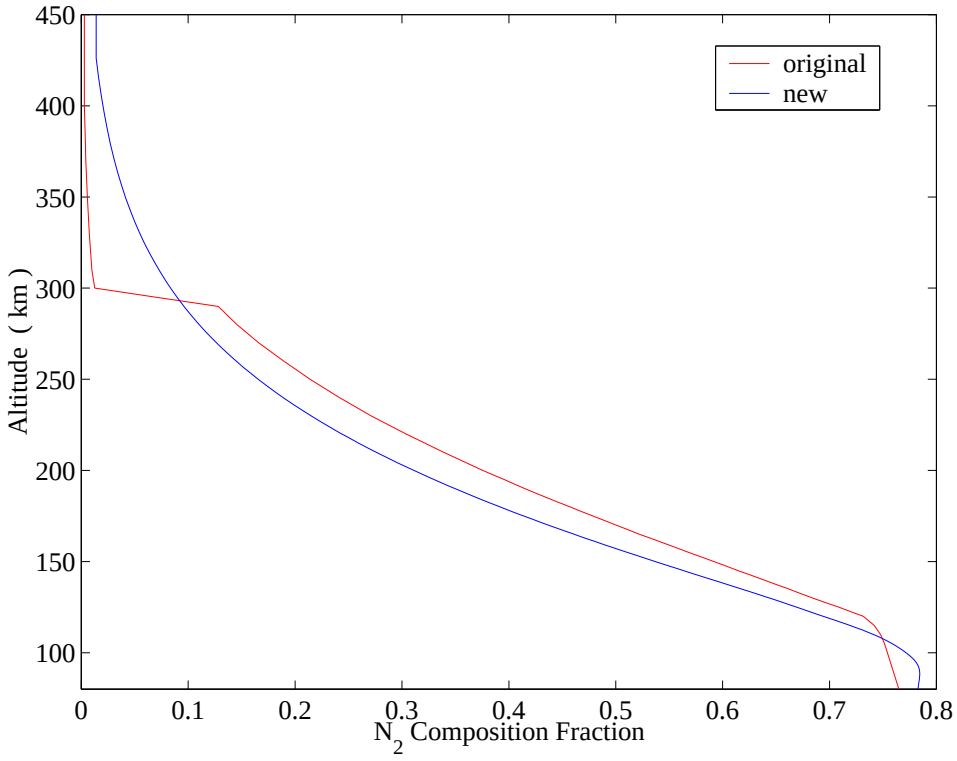

Figure 3.4: Initial composition fractions of molecular nitrogen in selected versions of the model. The N_2 composition fractions in the current version of the model (blue curve) are more realistic in the lower thermosphere and do not exhibit a large discontinuity at 300 km.

profile is also much more reasonable as it now exhibits an exponential increase as one moves into the mesosphere, unlike the original profile which exhibited a much gentler increase in mass density (Figure 3.7). The new temperature profile is a bit cooler at most altitudes, but is still quite reasonable (Figure 3.8).

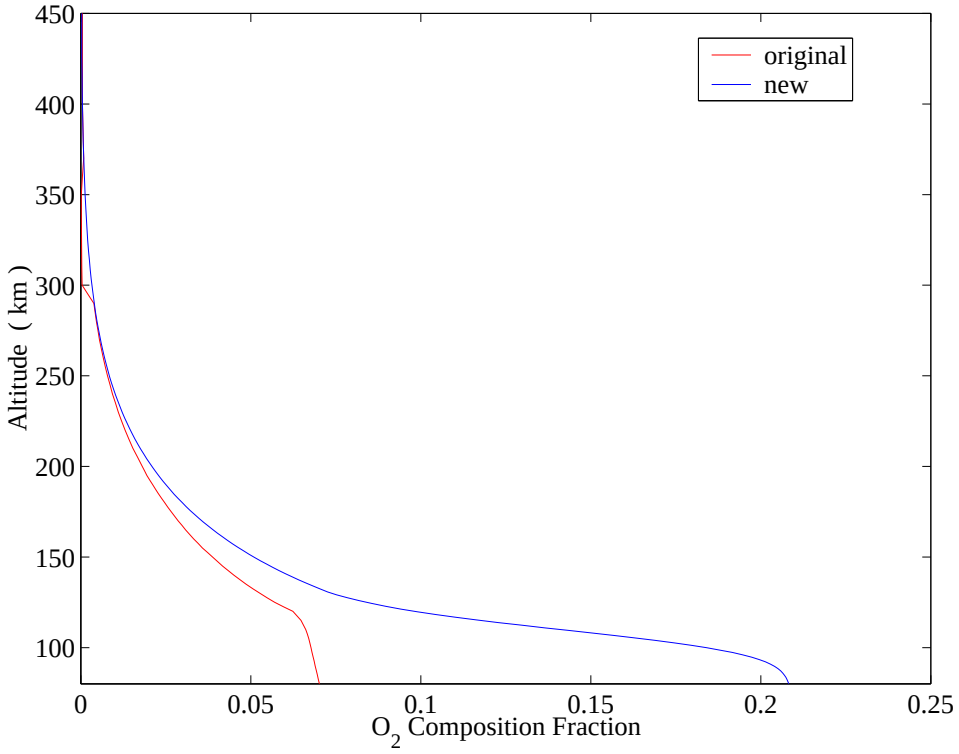

Figure 3.5: Initial composition fractions of molecular oxygen in selected versions of the model. The O_2 composition fractions in the current version of the model (blue curve) are more realistic in the lower thermosphere and do not exhibit a small discontinuity at 300 km.

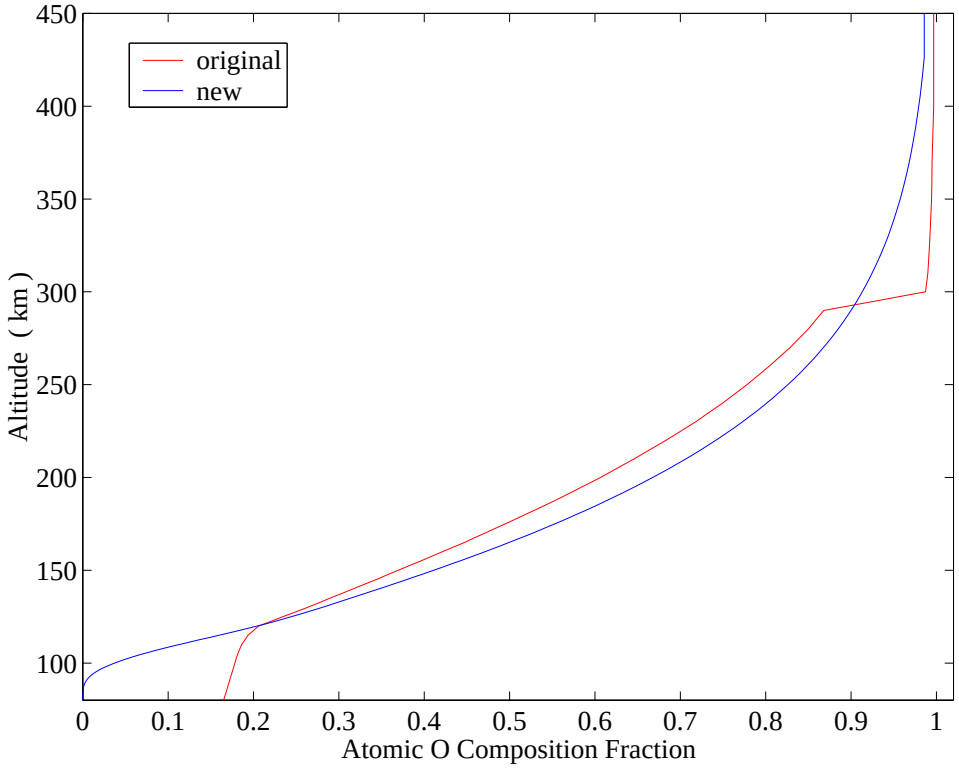

Figure 3.6: Initial composition fractions of atomic oxygen in selected versions of the model. The atomic oxygen composition fractions in the current version of the model (blue curve) are more realistic in the lower thermosphere and do not exhibit a large discontinuity at 300 km.

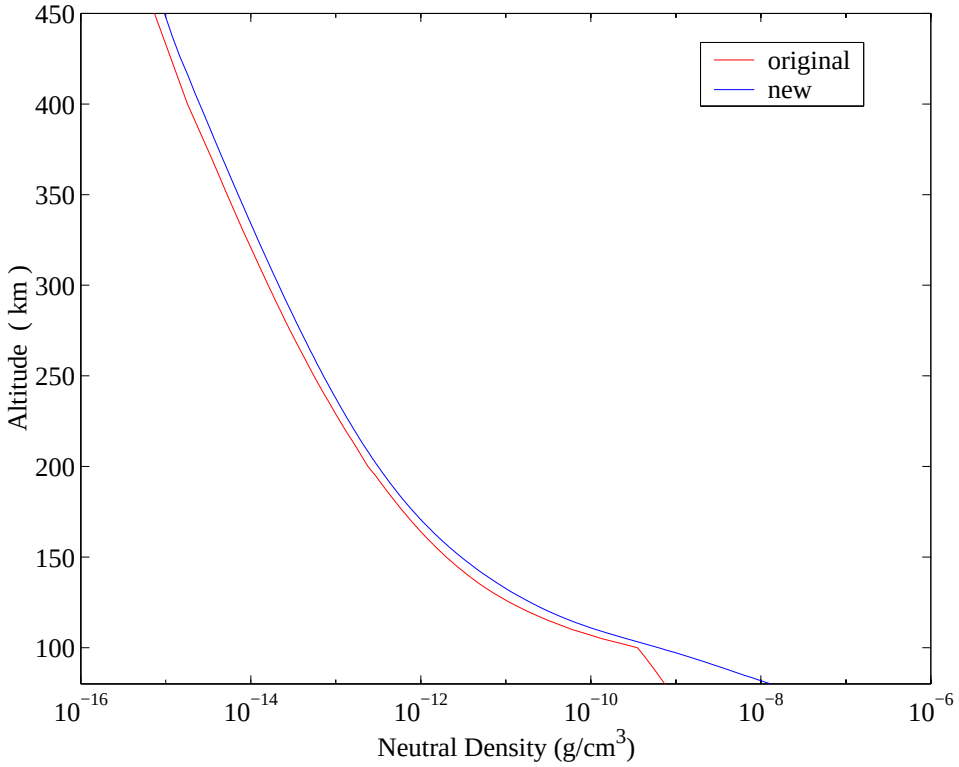

Figure 3.7: Initial density profiles in selected versions of the model. The initial density profile in the current version of the model (blue curve) is much more realistic near the lower boundary.

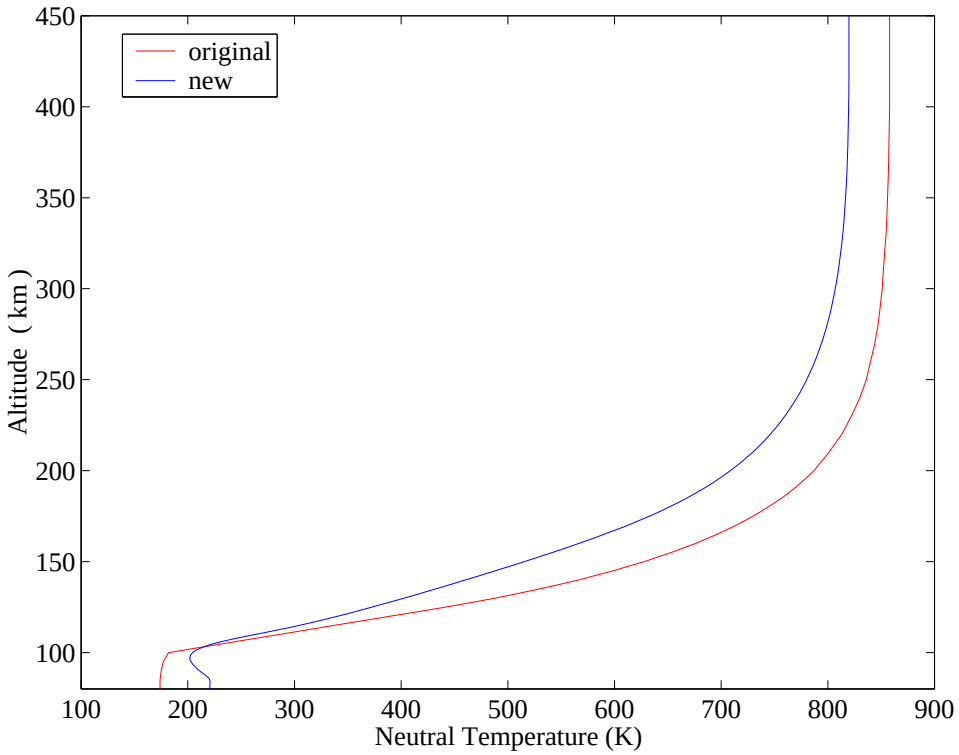

Figure 3.8: Initial temperature profiles in selected versions of the model. The initial temperature profile in the current version of the model (blue curve) is slightly cooler at most altitudes.

Chapter 4

Results and Discussion

In this chapter, several model runs will be presented: a 'reference run' that will be examined in some detail so the physics of the processes that control the response of the thermosphere can be elucidated and several other runs that will highlight how the temporal characteristics of the electric field can influence the thermospheric response. The location and width of the imposed electric field will be the same for all of these simulations: a Gaussian function of 0.5° half-width, centred on the geomagnetic field line that emerges from the Earth's surface at 70°N. The electric field will always attain a maximum value of 100 mV/m, but the rate at which it increases, i.e. the 'ramp-up' time, will vary. The 'reference run' will have a 'ramp-up' time of 1000 seconds so that the numerical results from this version of the model can be compared with the results obtained by *Chang and St.-Maurice* [1991]. Additional runs with quicker 'ramp-up' times of 60 seconds (1 minute), 300 seconds (5 minutes) and 600 seconds (10 minutes) will also be presented and compared to the 'reference run' results. Once the electric field achieves its maximum value, it will be left on at this value for the simulations that are investigating the effects of variable 'ramp-up' times. It should be emphasized that large changes in the electric field strength over time spans of 5-10 minutes are not uncommon during disturbed conditions; the 60 second 'ramp-up' time, on the other hand, is a bit more contentious, but will allow one to see how quickly the thermosphere can respond to rapid energy inputs.

4.1 Zonal winds

4.1.1 In the auroral zone

Contour plots of the 'reference run' zonal winds at 15 minutes, 30 minutes, 45 minutes, 60 minutes, 90 minutes, and 120 minutes are shown in Figure 4.1 (a positive wind corresponds to an eastward wind while a negative wind corresponds to a westward wind). The most obvious feature of these contour plots is that the zonal wind speeds increase quite quickly during the first 45 minutes of the simulation. This rapid acceleration is initially due to the ions being driven in the E x b direction by the electric field. The zonal momentum balance plots at the 15 minute mark of the simulation clearly illustrate this, with the ion drag terms (the green curves in Figure 4.2) being significantly larger than any of the other terms present. It should be noted that the advection terms are also playing a role; the meridional advection terms (the red curves in Figure 4.2) act as a source of momentum near the edges of the electric field at most heights while the vertical advection terms (the blue curves in Figure 4.2) act as a sink of momentum in the electric field region at all altitudes. Viscosity is only playing a minor role at this point in the simulation; it is a small sink of momentum at the higher heights that is comparable to the vertical advection contribution at these heights.

As time progresses, the contributions of the advection terms to the overall zonal momentum balance increase, with the meridional advection terms becoming comparable to the ion drag terms at the higher altitudes by the 30 minute mark of the simulation (Figure 4.3). Although the advection terms are becoming more important, it should be noted that the ion drag term is not as large as it was at the 15 minute mark of the simulation as the differences between the ion and neutral velocities are not as pronounced now. The Coriolis contribution to the momentum balances is also starting to become more important at the lower heights, essentially balancing the ion drag outside of the auroral zone (this will be discussed more thoroughly in the next section).

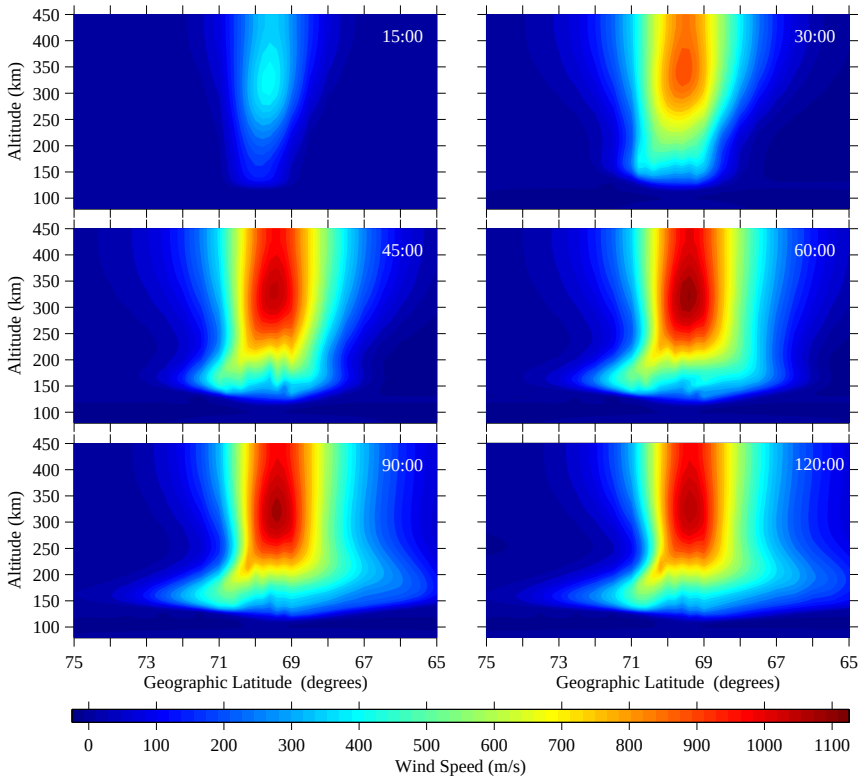

Figure 4.1: Zonal winds in the auroral zone for the 'reference run' simulation. The wind speeds have been contoured in 25 m/s intervals and reach their maximum values around the 60 minute mark of the simulation.

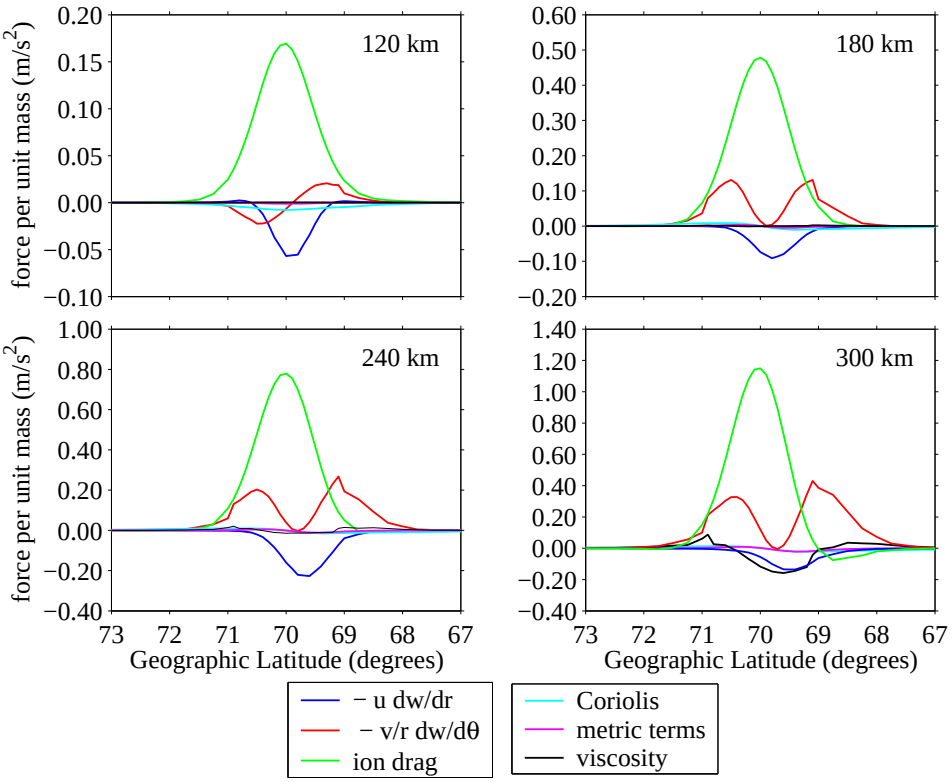

Figure 4.2: Zonal momentum balances in the auroral zone at the 15 minute mark of the 'reference run' simulation. The ion drag terms (the green curves) are the dominant acceleration mechanism at the 15 minute mark of the simulation, with the meridional and vertical advection terms (the red and blue curves respectively) making making minor contributions.

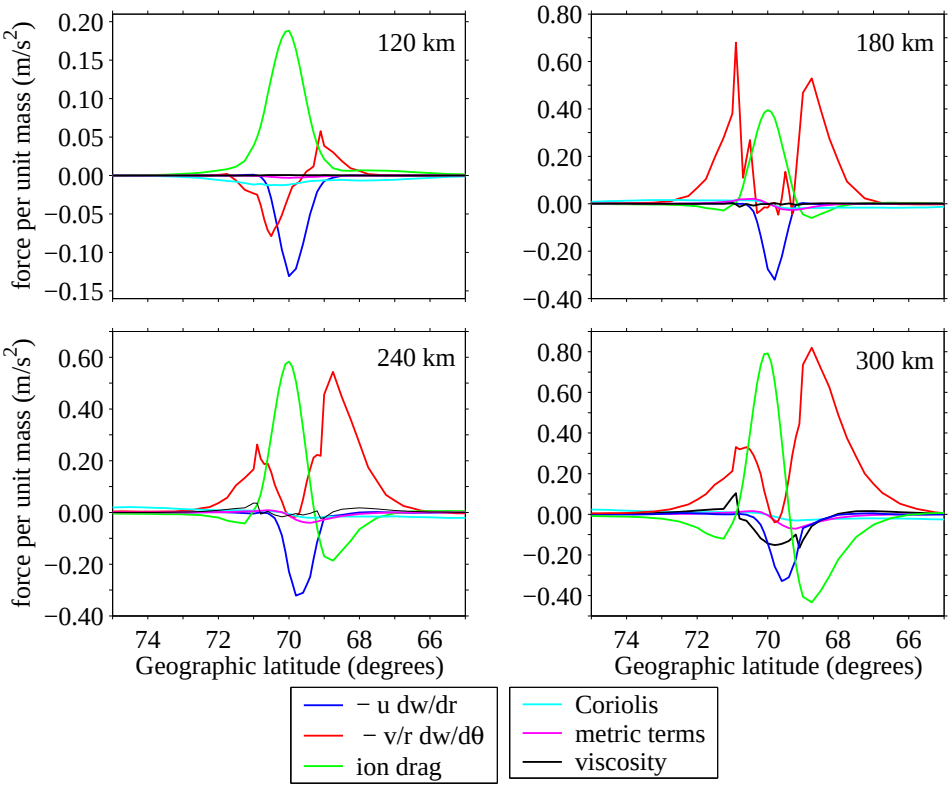

Figure 4.3: Zonal momentum balances in the auroral zone at the 30 minute mark of the 'reference run' simulation. The ion drag and the meridional advection terms (the green and red curves respectively) are providing most of the zonal acceleration at higher altitudes while the vertical advection terms (the blue curves) are acting as sink of momentum in the electric field region at all altitudes.

By the 60 minute mark of the simulation, the zonal winds in the auroral zone have achieved an essentially steady state solution, with an approximate balance between the ion drag, advection, and viscous terms occurring at all heights (Figure 4.4). However, the relative contributions of the meridional advection, vertical advection, and viscous terms to the zonal momentum balance are a function of height; the vertical and meridional advection terms balance ion drag at the lower altitudes, vertical advection balances ion drag at 180 km, while viscosity and meridional advection are balancing ion drag at the highest heights. It should be emphasized that these results are not new; *Richmond and Matsushita* [1975], *Mikkelsen et al.* [1981], *Fuller-Rowell* [1984,1985], and *Chang and St.-Maurice* [1991] have all noted the prominent role played by advection terms in the development of zonal winds during disturbed conditions.

It should also be noted that the assumption of zonal symmetry in these simulations will cause the zonal wind speeds in the auroral zone to be larger than those obtained by most other three dimensional models. The most obvious reason for this is the lack of dissipation in the zonal direction: any dynamical responses produced by the model can only be dissipated as they propagate meridionally. Another much more subtle reason is that the total energy input in a two dimensional model will almost always be larger than the total energy input in an equivalent three dimensional model (the energy sources in a three dimensional model are not zonally symmetric under most circumstances as this defeats the purpose of making the model three dimensional). *Sun et al.* [1995] have carried out a two dimensional/three dimensional model comparison to determine how large the differences in the modeled responses could be and found that the zonal winds could be overestimated by a factor of two. However, their results were obtained in Cartesian geometry on a much coarser grid, which makes a comparison of their results to the results of this thesis problematic. Despite all of these issues in the simulation of zonal winds during disturbed conditions, observations of high localized *transonic* zonal winds in the dawn sector of the polar cap region have been obtained by the Wind And Temperature Spectrometer (WATS) on the DE-2 satellite [*Balthazor and Bailey*, 2006], with most of these ob-

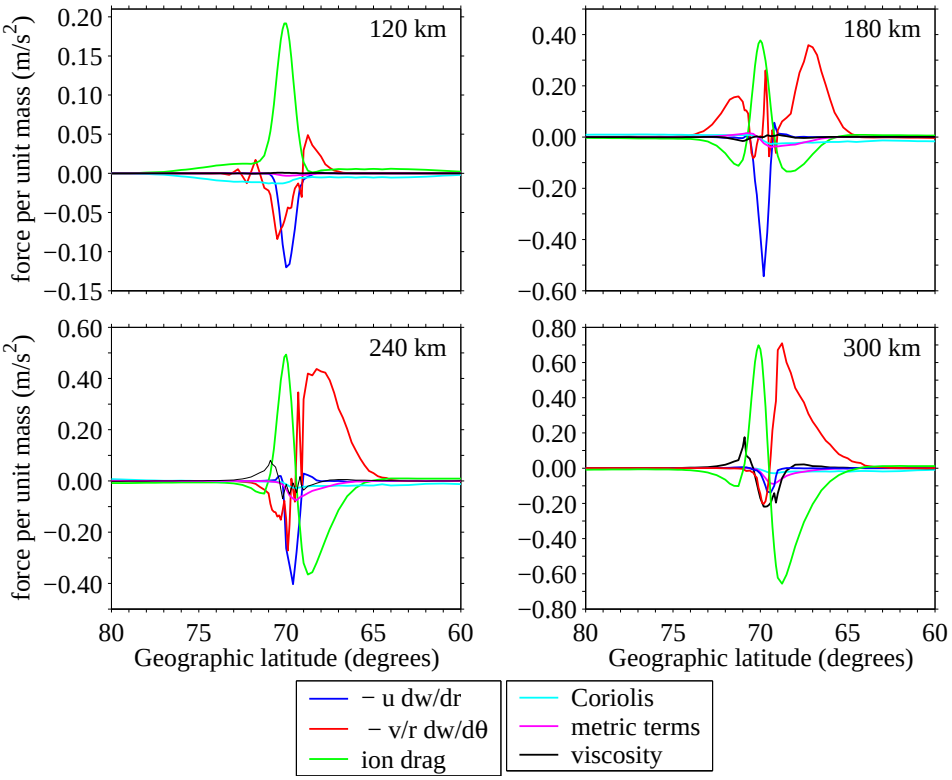

Figure 4.4: Zonal momentum balances in the auroral zone at the 60 minute mark of the 'reference run' simulation. The ion drag, meridional advection, vertical advection, and viscosity terms (the green, red, blue, and black curves respectively) are all playing a role in the determination of the final wind speeds.

servations occuring in the 300-600 km region. The simulation results presented in Figure 4.1 are in fairly good agreement with these observations.

The 'ramp-up' time of the electric field has very little effect on the evolution of the zonal winds once the wind speeds exceed 200 m/s (Figure 4.5). The zonal acceleration time series derived from these wind speed time series (through 2nd order finite differencing) are presented in Figure 4.6 and show that the maximum zonal accelerations occur when the electric field has obtained its maximum value. They also show that there are three stages in the evolution of the zonal acceleration time series: an initial stage where the zonal acceleration is linear, an intermediate stage where the zonal acceleration continues to increase at a slower, non-linear rate, and a final stage, where the zonal acceleration has already reached its maximum value and is decreasing in an approximately exponential manner. The first two stages in the zonal acceleration time series are caused by ion-drag and occur before the electric field has obtained its maximum value; the differences between these two stages are largely due to the increases in neutral density (Figure 4.7), i.e. the ions are not able to transfer as much momentum to the neutrals during the second stage as the enhanced neutral densities do not allow the ions to be accelerated to the velocities that they were obtaining in the first stage. This means that, for the conditions of these simulations, the effects of Joule heating (through the increases in density) play an important role in the evolution of the zonal winds. A more thorough analysis of this effect will be pursued at a later date.

4.1.2 In the sub-auroral zone

Contour plots of the 'reference run' zonal winds at 30 minutes, 60 minutes, 90 minutes, 120 minutes, 150 minutes, and 180 minutes are shown in Figure 4.8. The maximum and minimum values of these contour plots have been set to 40 m/s and -40 m/s respectively so the zonal winds in the auroral zone do not 'wash out' the zonal wind system that exists outside of the auroral zone.

There are three features to this wind system : winds that radiate outwards from the auroral region during the first hour of the simulation, a second, slowly propa-

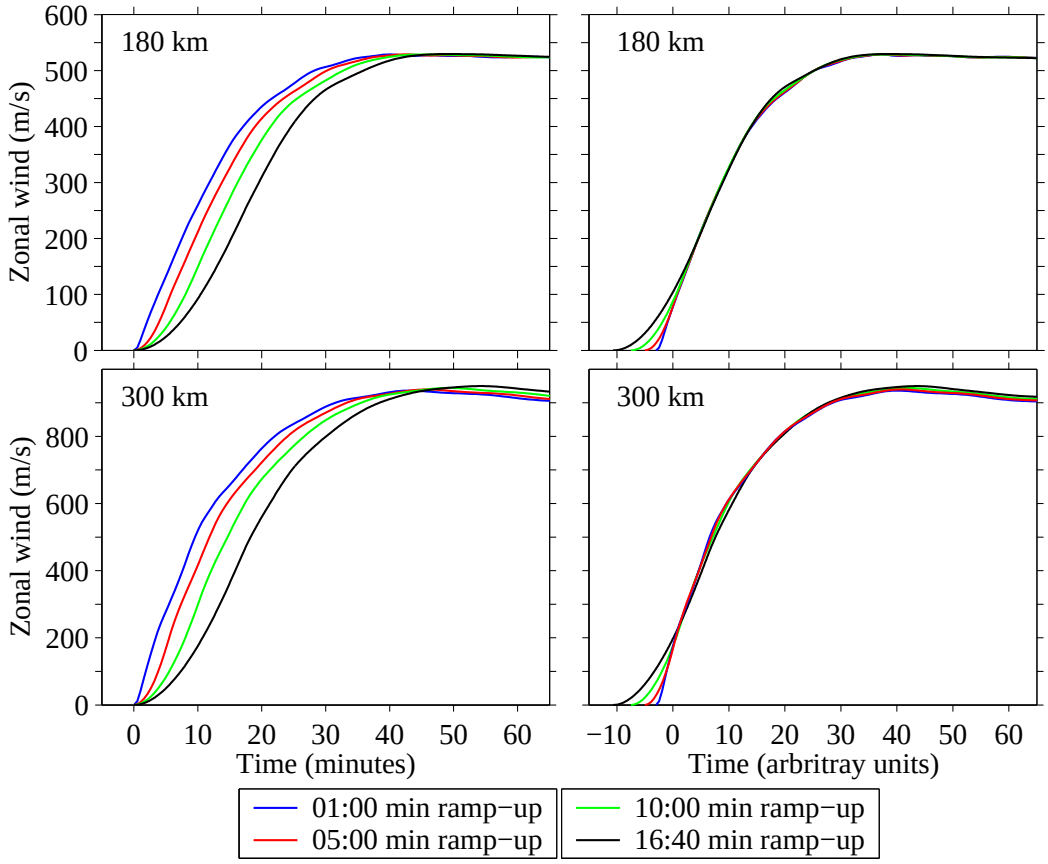

Figure 4.5: Zonal wind time series at selected altitudes in the centre of the auroral zone as a function of electric field 'ramp-up' time. The zonal winds at 180 km (upper left panel) and 300 km (lower left panel) are not very sensitive to the 'ramp-up' time of the electric field as all of the wind speed time series exhibit the same behaviour once the wind speed exceeds 200 m/s. This can be more clearly seen in the right panel zonal wind time series, which have all been 'time-lagged' to coincide with each other.

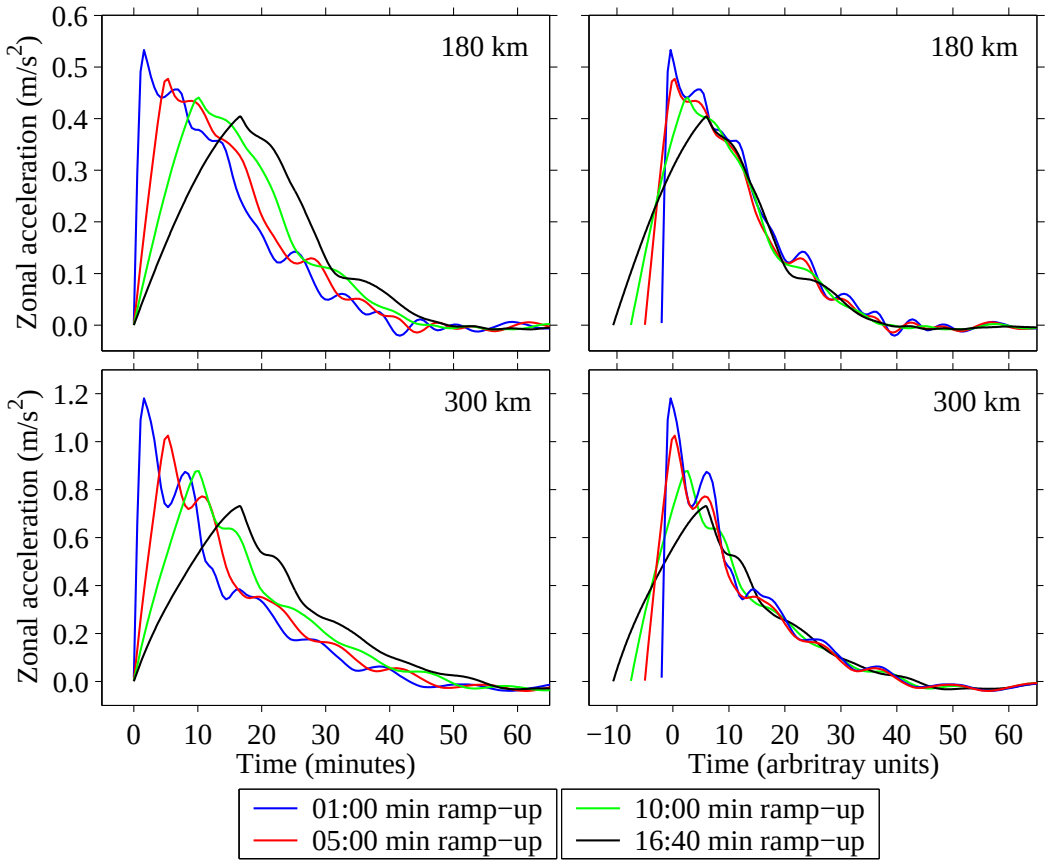

Figure 4.6: Zonal acceleration time series at selected altitudes in the centre of the auroral zone as a function of electric field 'ramp-up' time. The zonal accelerations at 180 km (upper left panel) and 300 km (lower left panel) are not very sensitive to the 'ramp-up' time of the electric field as all of the acceleration time series exhibit the same behaviour. This can be more clearly seen in the right panel zonal acceleration time series, which have all been 'time-lagged' to coincide with each other.

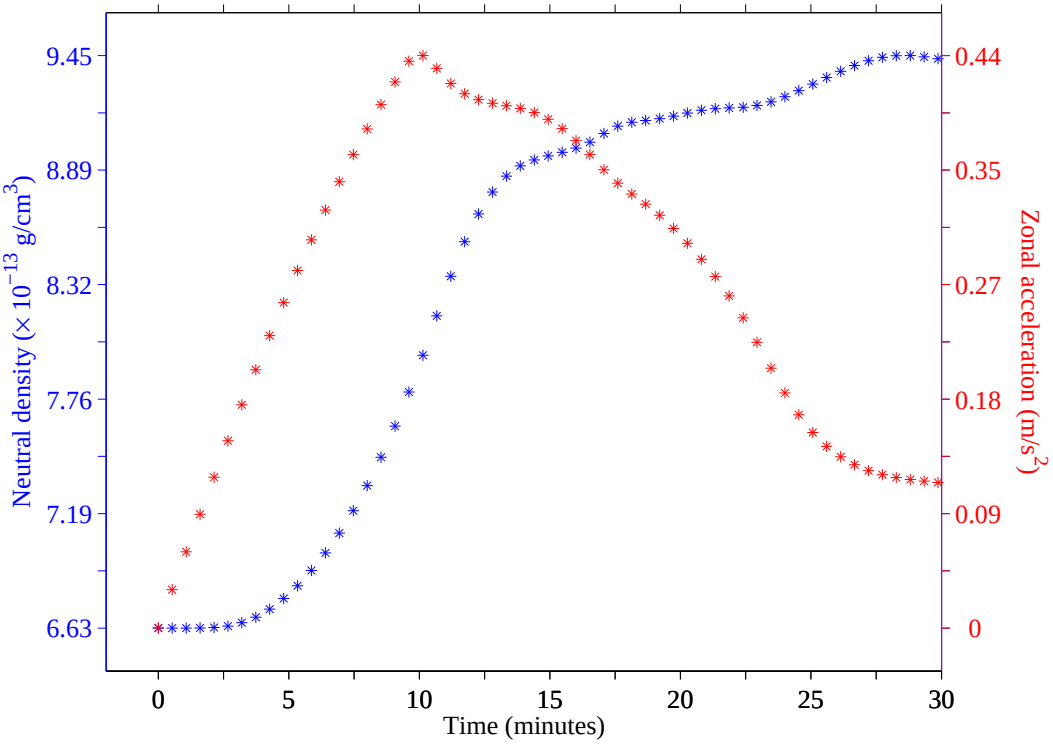

Figure 4.7: Zonal acceleration and neutral density time series in the centre of the auroral region for a simulation with an electric field 'ramp-up' time of 10 minutes. The zonal acceleration time series deviates from its linear behaviour when the neutral densities begin to increase and reaches its maximum value when the electric field finishes 'ramping-up' to its maximum value.

gating, westwards wind that propagates equatorwards in the final two hours of the simulation, and a weak "jet" that appears on the equatorial side of auroral zone in the final hour of the simulation. A quick inspection of the zonal momentum balance plots indicates that the ion drag and Coriolis terms are the dominant terms outside of the auroral zone in the latter part of the simulation (Figures 4.9-4.11), which suggests that the slowly propagating wind is a geostrophically-balanced flow. This conclusion is not particularly surprising as the assumption of zonal symmetry does not allow any zonal pressure gradients to exist in the model. If one were to allow zonal pressure gradients to exist in the model, one would observe the fore mentioned slowly propagating wind and superposed wave-like perturbations outside the auroral zone, substantially altering the geostrophic balance of this flow. The wind system that radiates outwards during the first hour of the simulation and the weak 'jet' on the equatorial side of the auroral zone in the final hour of the simulation are due to the Coriolis deflection of meridional wind system features that will be discussed more thoroughly in the next section of this chapter.

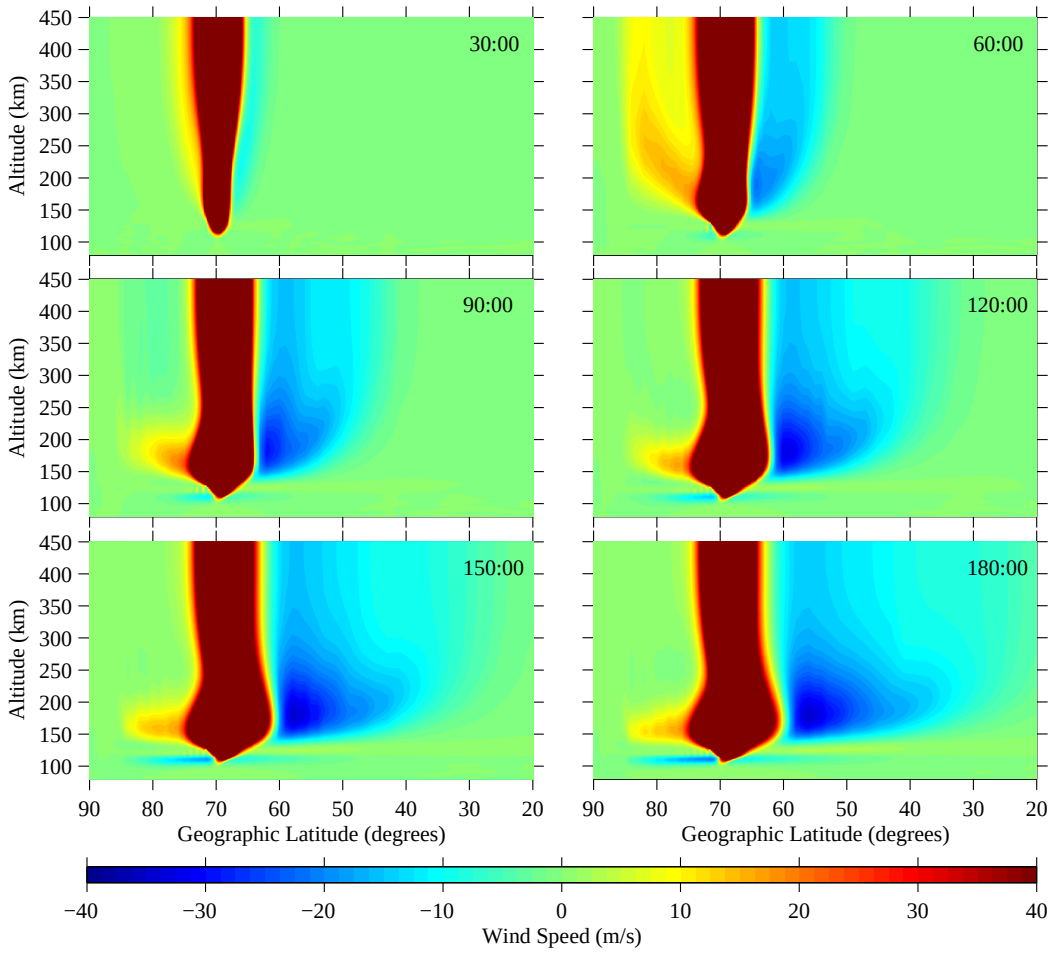

Figure 4.8: Zonal winds outside of the auroral zone for the 'reference run' simulation. The wind speeds have been contoured in 1 m/s intervals and are predominantly westwards outside of the auroral region.

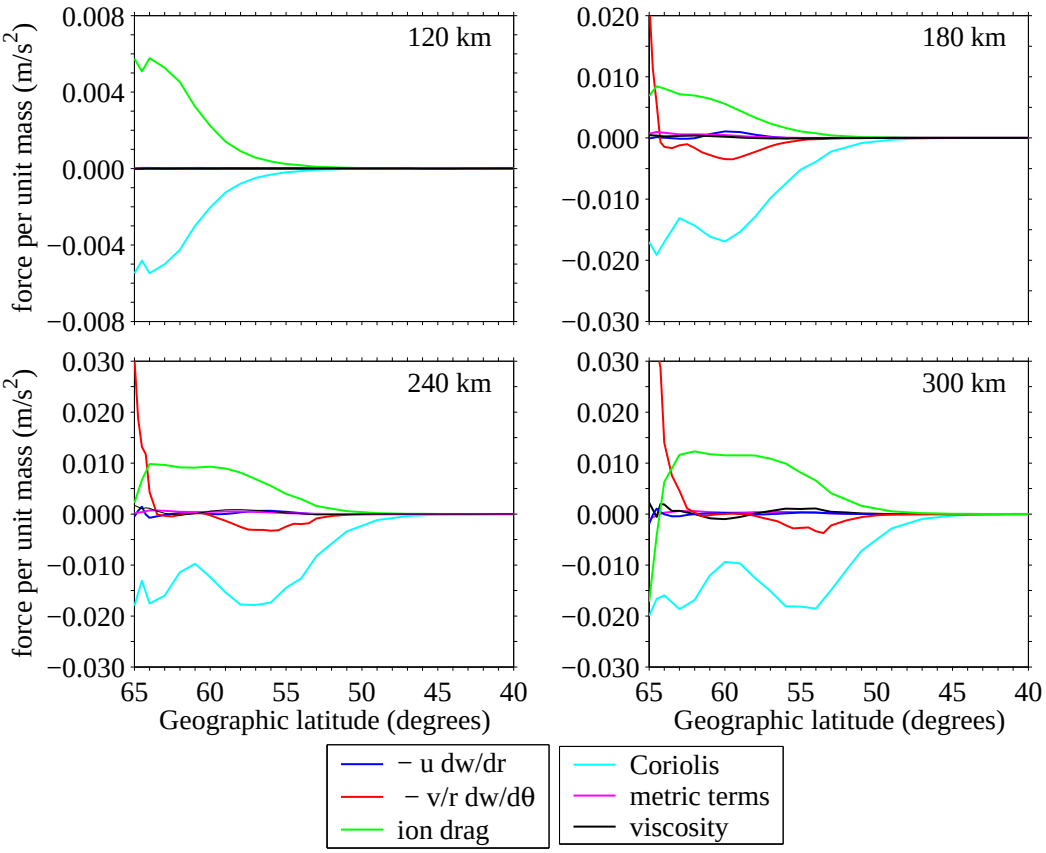

Figure 4.9: Zonal momentum balances outside of the auroral zone at the 60 minute mark of the 'reference run' simulation. The Coriolis terms (the cyan curves) are slightly larger than the ion drag terms (the green curves) at most altitudes, producing a small, westwards zonal acceleration that drives the geostrophically balanced flow of Figure 4.8.

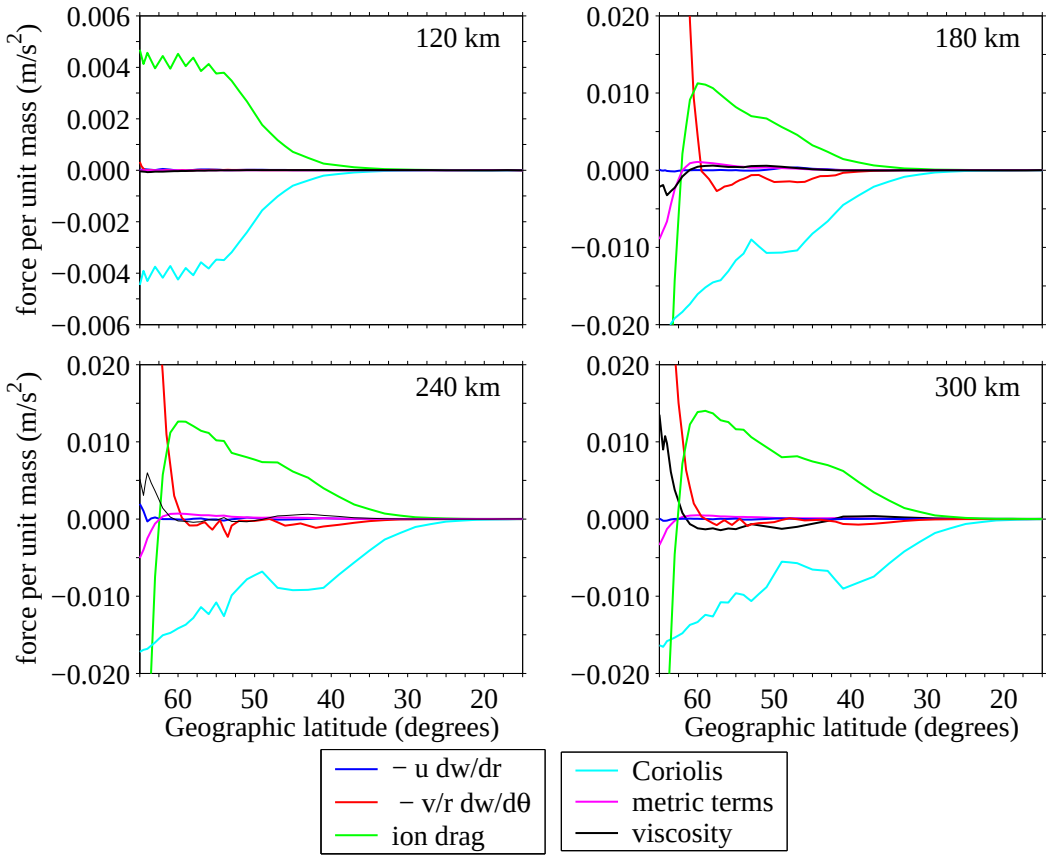

Figure 4.10: Zonal momentum balances outside of the auroral zone at the 120 minute mark of the 'reference run' simulation. The Coriolis terms (the cyan curves) are slightly larger than the ion drag terms (the green curves) at most altitudes, producing a small, westwards zonal acceleration that drives the geostrophically balanced flow of Figure 4.8.

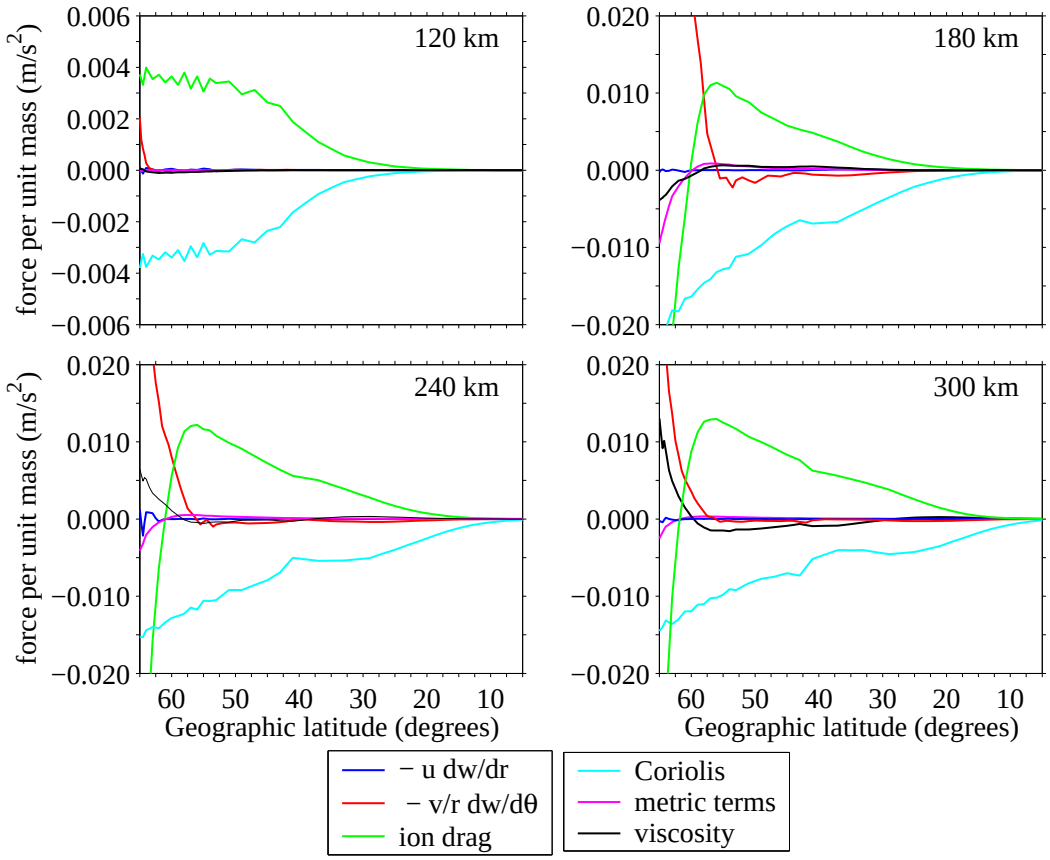

Figure 4.11: Zonal momentum balances outside of the auroral zone at the 180 minute mark of the 'reference run' simulation. The Coriolis terms (the cyan curves) are slightly larger than the ion drag terms (the green curves) at most altitudes, producing a small, westwards zonal acceleration that drives the geostrophically balanced flow of Figure 4.8.

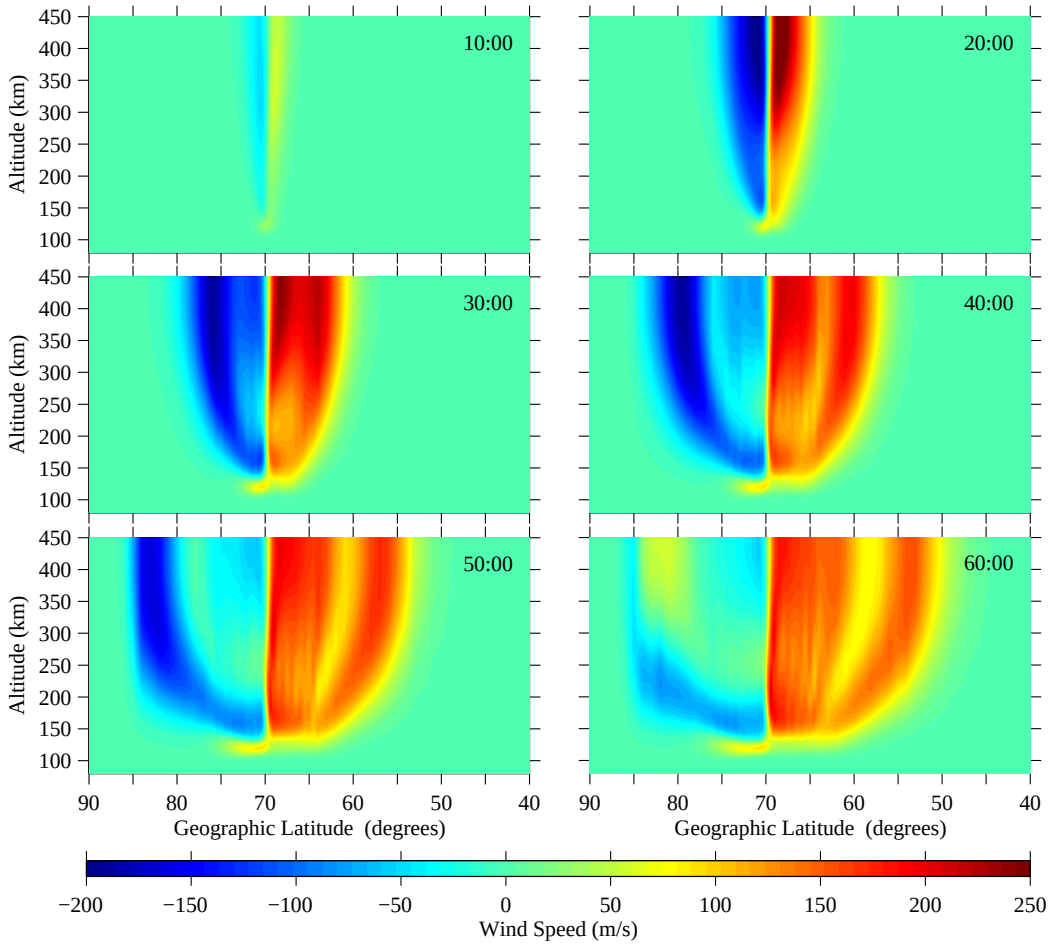

Figure 4.12: 'Reference run' meridional winds for the first hour of the simulation. The wind speeds have been contoured in 5 m/s intervals and reach their maximum values in the upper thermosphere near the 20 minute mark of the simulation.

4.2 Meridional Winds

Contour plots of the 'reference' run meridional winds at 10 minutes, 20 minutes, 30 minutes, 40 minutes, 50 minutes, 60 minutes, 75 minutes, 90 minutes, 120 minutes, and 180 minutes are shown in Figures 4.12-4.13. It should be noted that a positive (negative) meridional wind corresponds to a southward (northward) wind and that the meridional winds at 60 minutes appear in both figures.

During the first 20 minutes of the simulation, the meridional winds are being driven by Joule heating above altitudes of approximately 150 km. This is quite evident

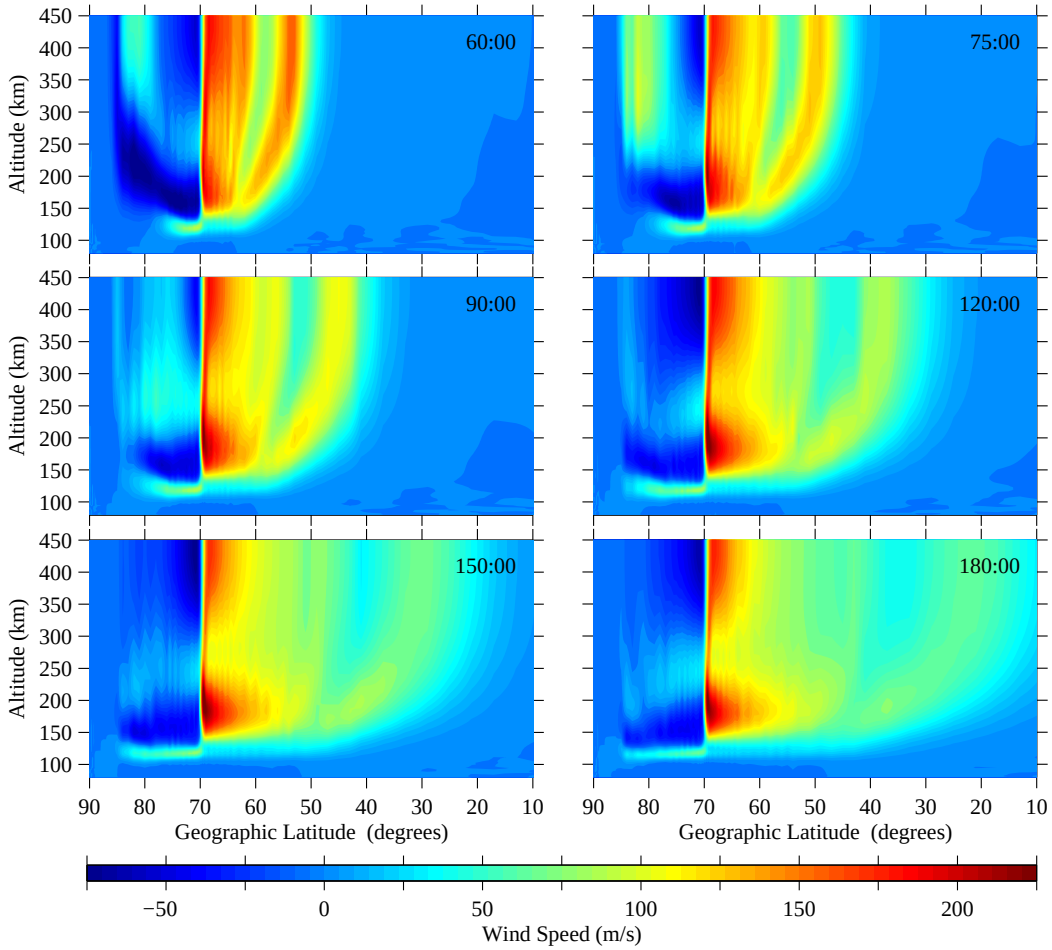

Figure 4.13: 'Reference run' meridional winds for the final two hours of the simulation. The wind speeds have been contoured in 5 m/s intervals and are predominantly equatorwards during this part of the simulation.

in the meridional momentum balance plots of Figure 4.14, where the large pressure gradient terms (the blue curves in Figure 4.14) are driving the meridional winds away from the centre of the auroral zone, i.e. there is a large positive pressure gradient on the equatorial side of the auroral zone that is driving the winds equatorwards and a large negative pressure gradient on the polar side of the auroral zone that is driving the winds polewards. It should be emphasized that this pressure gradient is due to changes in thermospheric density; the changes in neutral temperature are only making a significant contribution to the pressure gradient in the centre of the auroral zone (Figure 4.15). The ion drag terms (the green curves in Figure 4.14) and the advection terms (the red curves in Figure 4.14) are opposing this pressure gradient driven motion, i.e. the ion drag and the advection terms are essentially negative when the pressure gradient term is positive and vice-versa, while the viscosity term, once again, is acting as a minor sink of momentum at the higher altitudes (the black curves in Figure 4.14). The zonal winds also play a role in the auroral region through the Coriolis and metric terms (the cyan and magenta curves in Figure 4.14 respectively); these effects are fairly small at the fifteen mark of the simulation but will become much more important as the zonal wind speeds increase.

At altitudes between 110 km and 130 km, the meridional winds are moving equatorwards on both sides of the auroral zone, in contrast to the Joule heating driven winds at the higher altitudes. This equatorwards motion is due to a large ion drag acceleration in this region (the green curve in Figure 4.14a), which is a consequence of the ion-neutral collision frequency allowing the ions to be accelerated in the electric field direction (the ion-neutral collision frequency keeps the ions 'bound' to the neutrals at lower altitudes and allows them to $\mathbf{E} \times \mathbf{B}$ drift at higher altitudes). The pressure gradient also has a role to play at these altitudes; it acts as a source of momentum outside the convection channel and as a sink of momentum inside the convection channel.

By the 30 minute mark of the simulation, the meridional winds on the equatorial side of the auroral zone have reached speeds of 150 m/s while the winds on the polar side of the auroral zone have only been accelerated to speeds of approximately 75

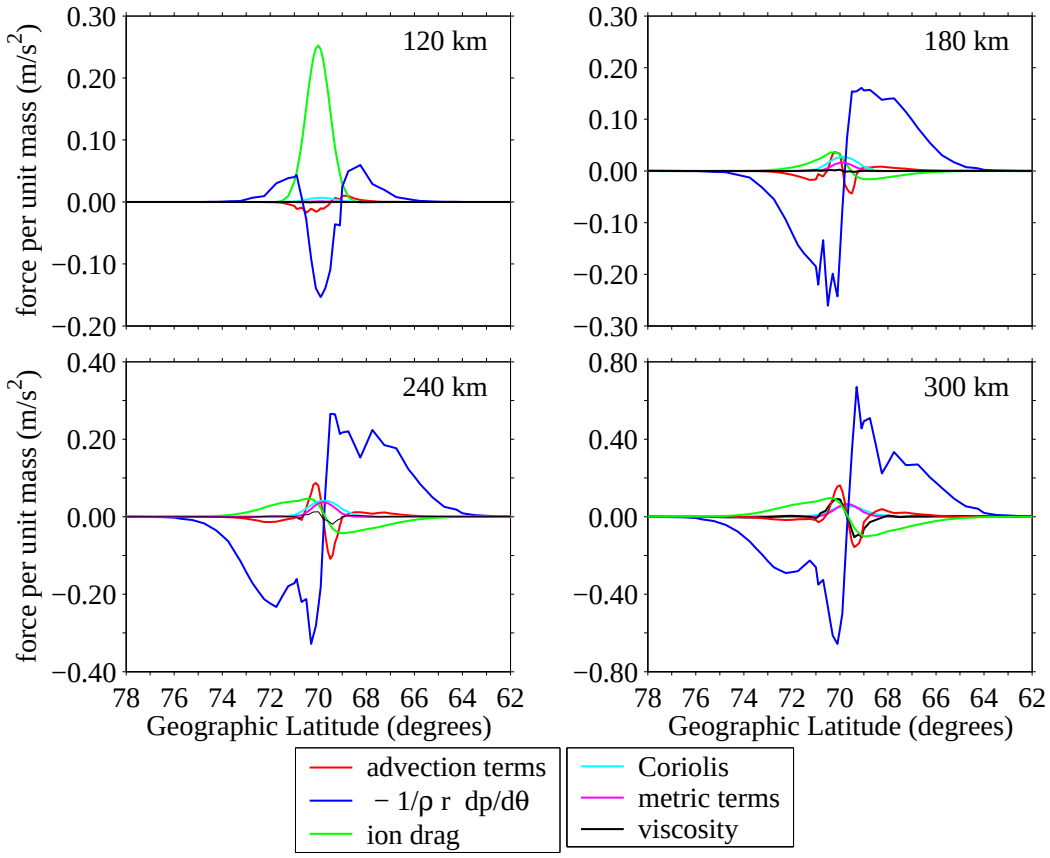

Figure 4.14: Meridional momentum balances at the 15 minute mark of the 'reference run' simulation. The meridional pressure gradients (the blue curves) are the dominant momentum term above 150 km, causing the meridional winds to radiate away from the auroral zone. The ion drag term at 120 km (the green curve in the uppermost left panel) is larger than the pressure gradient at this altitude, creating an equatorwards wind that flows across the entire auroral zone at this height.

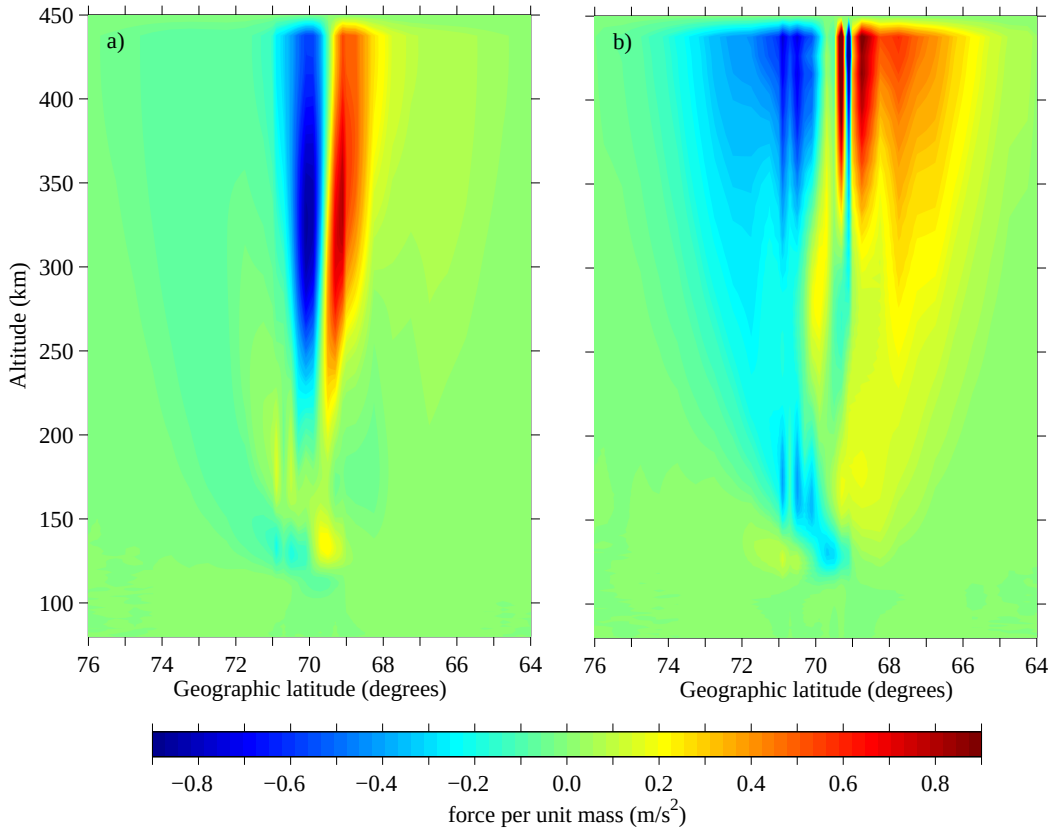

Figure 4.15: Meridional pressure gradients at the 15 minute mark of the 'reference run' simulation. The temperature gradient component of the meridional pressure gradient (panel a) is normally much smaller than the density gradient component (panel b); the upper auroral zone is the only exception.

m/s (Figure 4.12). This asymmetry in the meridional wind speeds of the auroral zone becomes even more pronounced at the 40 minute mark of the simulation, with the wind speeds on the polar side of the auroral zone only reaching values of 20 m/s (wave-like perturbations on the order of 200 m/s can still be found *outside* of the auroral zone). Most of this asymmetry can be attributed to the rapidly increasing zonal wind speeds in the auroral zone; the Coriolis and centrifugal ($w^2 \cot\theta/r$) terms are both generating substantial equatorwards accelerations that, when added together, can almost overcome the pressure gradient term on the polar side of the auroral zone (Figures 4.16 -4.17). These two terms actually do overcome the Joule heating driven pressure gradient in the 200-300 km region by the 50 minute mark of the simulation, with the meridional winds on the polar side of the upper auroral zone now moving equatorwards (Figure 4.12). The reflection of gravity waves off of the polar boundary can also be observed at the 60 minute of the simulation, which suggests that the boundary conditions at the polar boundary can still be improved upon.

The importance of the zonal winds in the evolution of the meridional wind system can also be demonstrated by setting the Coriolis and centrifugal ($w^2 \cot\theta/r$) terms in the meridional momentum equation to zero. A contour plot of the meridional wind field at the 60 minute mark of such a simulation is presented on the right hand side of Figure 4.18. The meridional winds on the polar side of auroral zone are now much stronger as the Coriolis and centrifugal terms are not decelerating the gas now (an equatorwards acceleration will decelerate the gas if it is moving polewards). The meridional winds on the equatorial side of the auroral zone, on the other hand, are now weaker as the Coriolis and centrifugal terms are not being allowed to accelerate the gas equatorwards.

If the contour values of Figure 4.18 are adjusted to better represent the maximum and minimum of the wind field, one can see that the magnitudes of the meridional winds are much more symmetrical below altitudes of 220 km (Figure 4.19), which suggests that these winds are primarily being driven by Joule heating. The asymmetry in the auroral zone above 250 km is still quite pronounced and is due to a large pressure build-up in the polar region of the model (Figure 4.20). This large pressure build-up

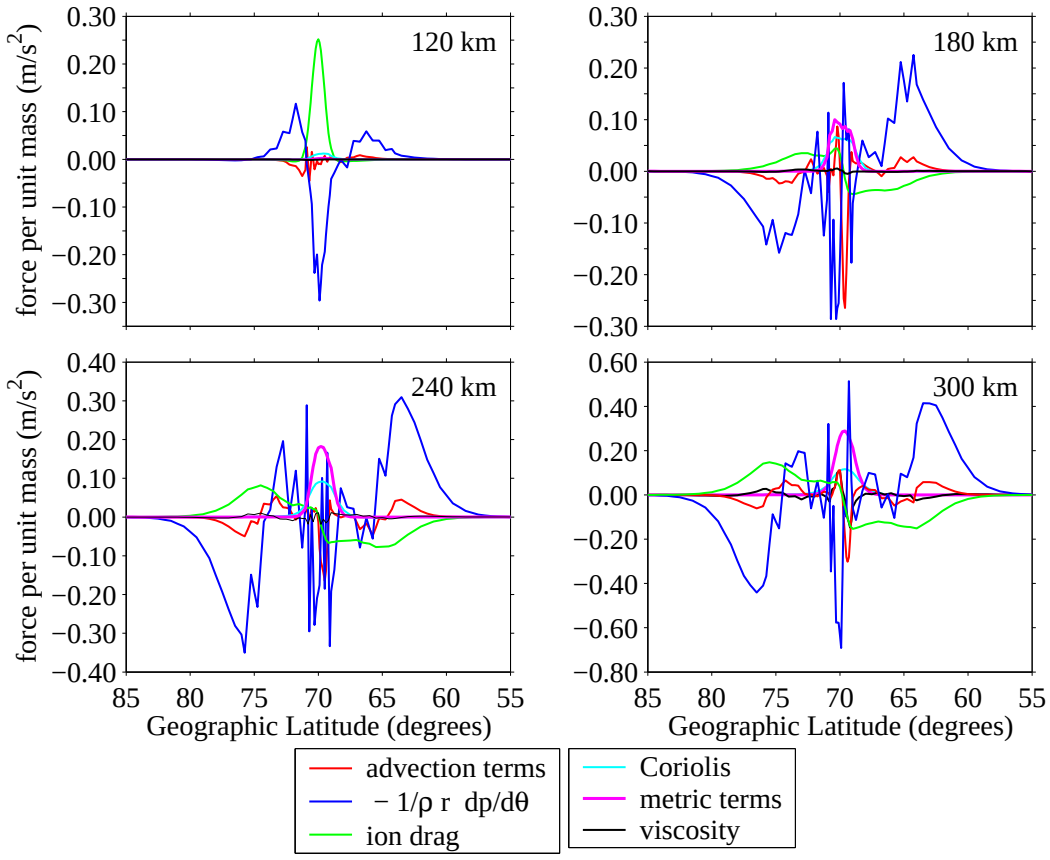

Figure 4.16: Meridional momentum balances at the 30 minute mark of the 'reference run' simulation. The meridional pressure gradients (the blue curves) are the dominant momentum term outside of the auroral zone. They also play an important role in the auroral zone, but they are not as effective in this region as the Coriolis (cyan) and centrifugal (magenta) terms are opposing them.

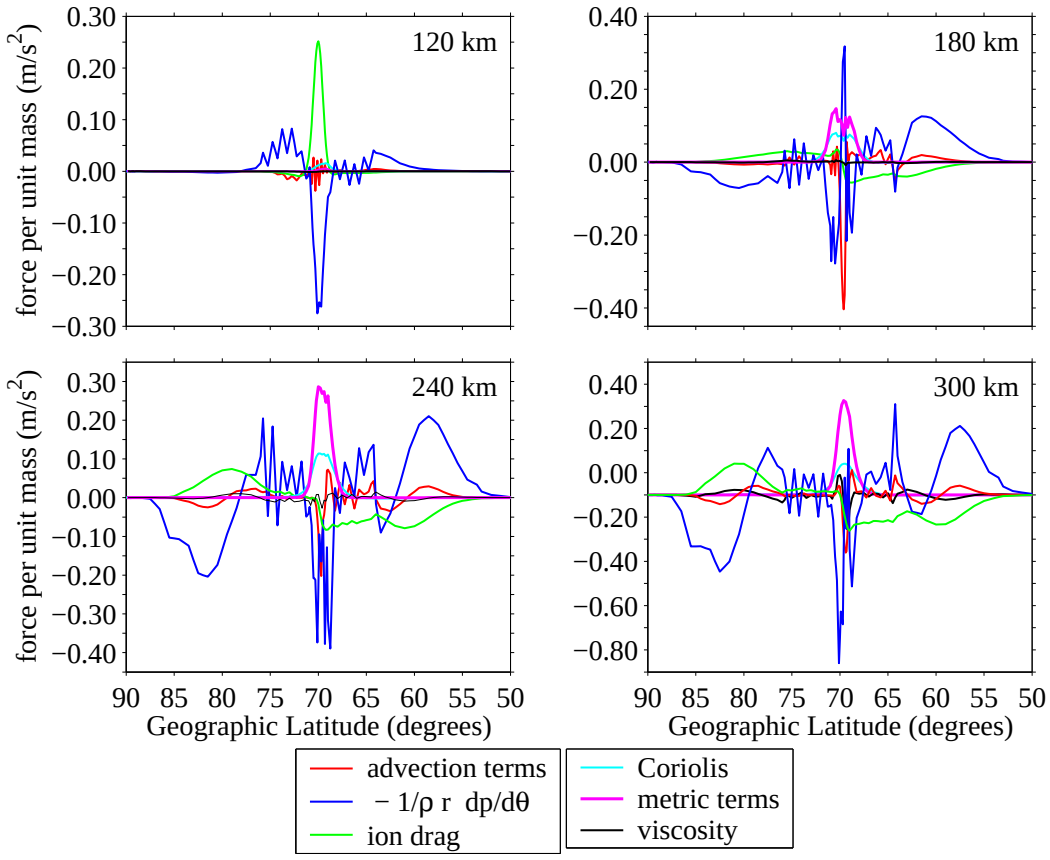

Figure 4.17: Meridional momentum balances at the 45 minute mark of the 'reference run' simulation. The meridional pressure gradients (the blue curves) are the dominant momentum term outside of the auroral zone. They also play an important role in the auroral zone, but they are not as effective in this region as the Coriolis (cyan) and centrifugal (magenta) terms are opposing them.

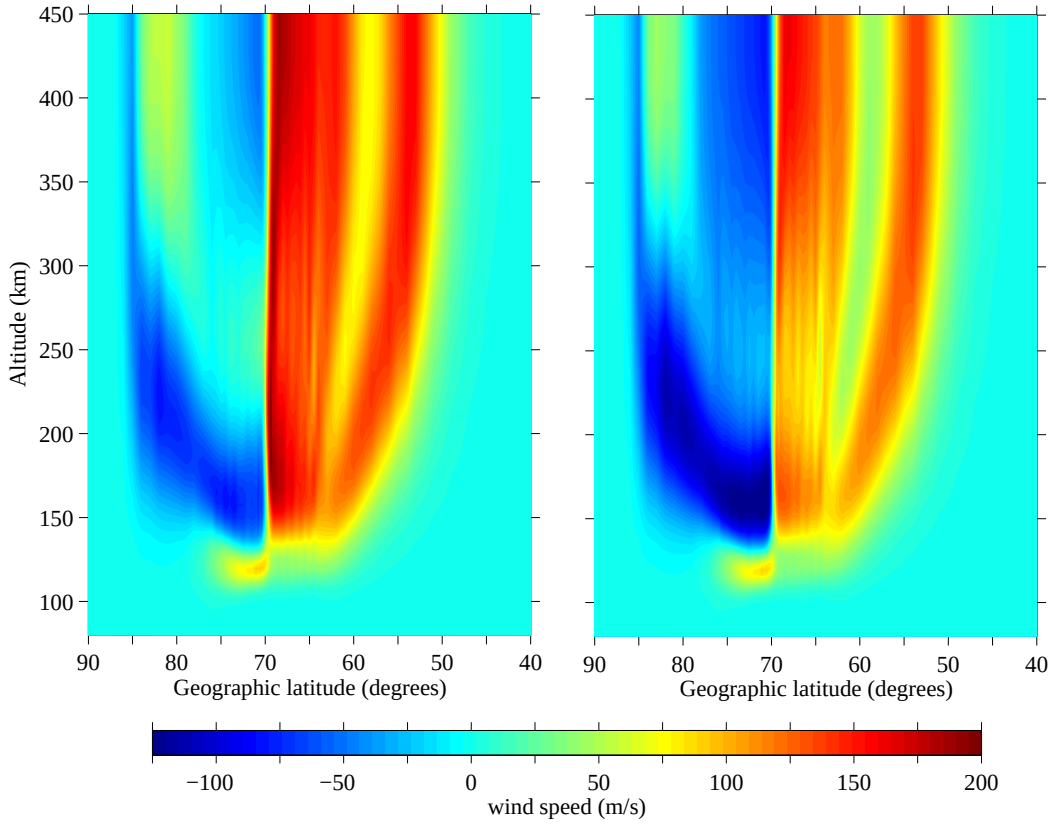

Figure 4.18: Meridional wind speeds at the 60 minute mark of two simulations with different meridional momentum source terms. The meridional winds produced by the simulation with the 'modified' meridional momentum source term are presented in the right panel and are weaker (stronger) on the equatorial (polar) side of the auroral zone (see the text for more details about the 'modified' meridional momentum source term).

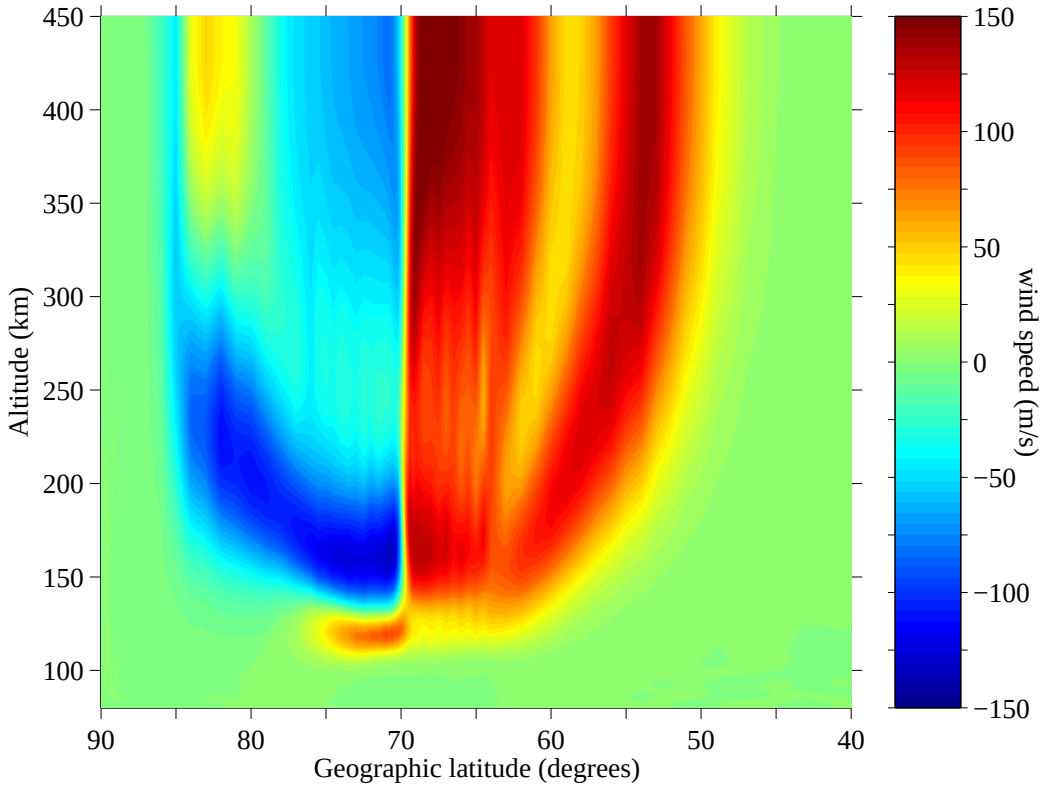

Figure 4.19: Meridional wind speeds at the 60 minute mark of a simulation with a 'modified' meridional momentum source term. The meridional wind speeds in this simulation are much more symmetric about the auroral zone.

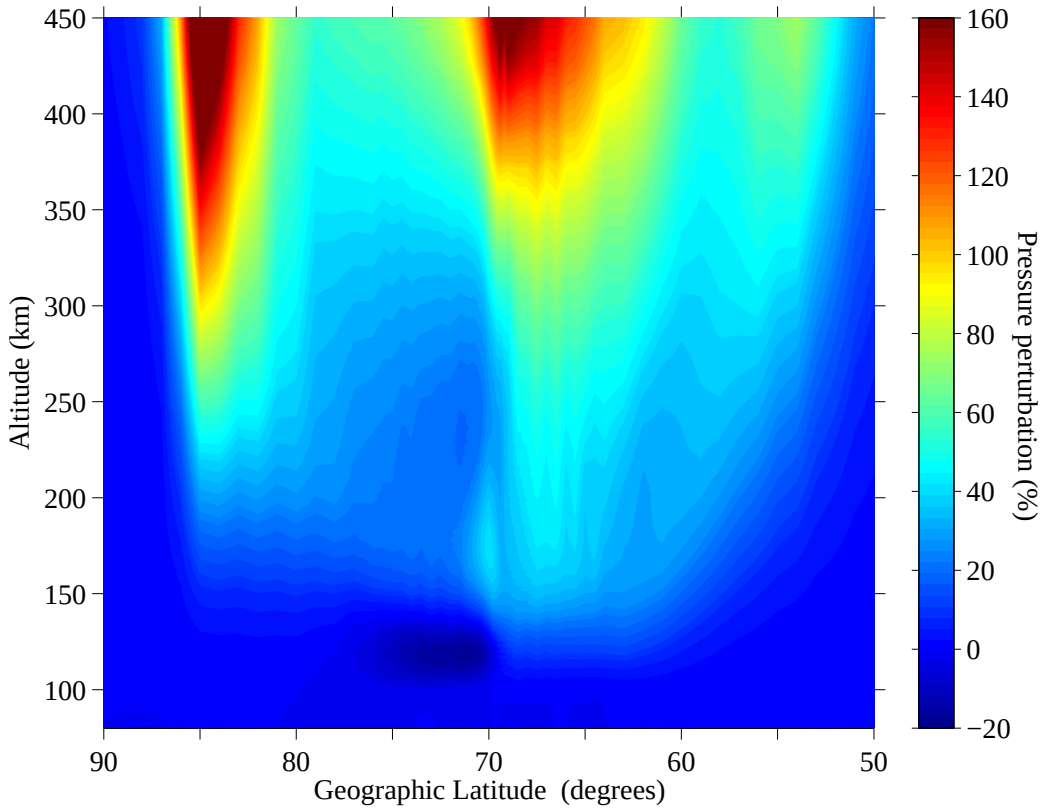

Figure 4.20: Pressure perturbations at the 60 minute mark of the 'reference run' simulation.

is caused by the polewards transport of material, which, in the zonal symmetry of the model, must 'pile up' in the polar region. The only way to avoid this pressure build-up is to replace the coordinate system of the model with one that avoids pole singularities and/or relaxes the assumption of zonal symmetry, i.e. makes the model three dimensional so material can be advected over the polar region. Small scale oscillations in the pressure gradient term are also quite apparent in some of the meridional momentum balance plots, but these oscillations are most likely numerical artifacts as the finite difference methods currently being employed in the model cannot simulate these smaller scale features with any degree of accuracy.

The most identifiable feature in the meridional wind field in the latter parts of the simulation is the jet that develops on the equatorward side of the auroral zone (Figure 4.13). This jet dominates the meridional wind system in the 150 km - 250 km altitude interval and, by the end of the simulation, has reached the 30 degree meridian. The appearance of this jet after the development of large zonal winds in the convection channel is no coincidence: most of the equatorwards acceleration driving this jet is being provided by the Coriolis and centrifugal terms. The importance of the zonal winds in the evolution of this jet can, once again, be demonstrated most easily by setting the Coriolis and centrifugal terms to zero in the meridional momentum equation. Contour plots of the meridional wind field at the 90 and 120 minute marks of such a simulation are presented in Figure 4.21.

The effects of Joule heating at the lower altitudes is quite obvious and no jet has formed in the lower auroral zone, as expected (see previous paragraph). This explanation of the meridional jet contradicts the explanation of *Fuller-Rowell* [1985], who proposed that

> ... the lower thermosphere meridional jet on the equatorward side peaking at 135 km appears to be 'fed' by a weaker jet at 115 km on the poleward side. The altitude, and hence mass density difference, of the layers explains the velocity differences on the poleward and equatorward sides; i.e., mass continuity can be satisfied by a smaller wind velocity on the

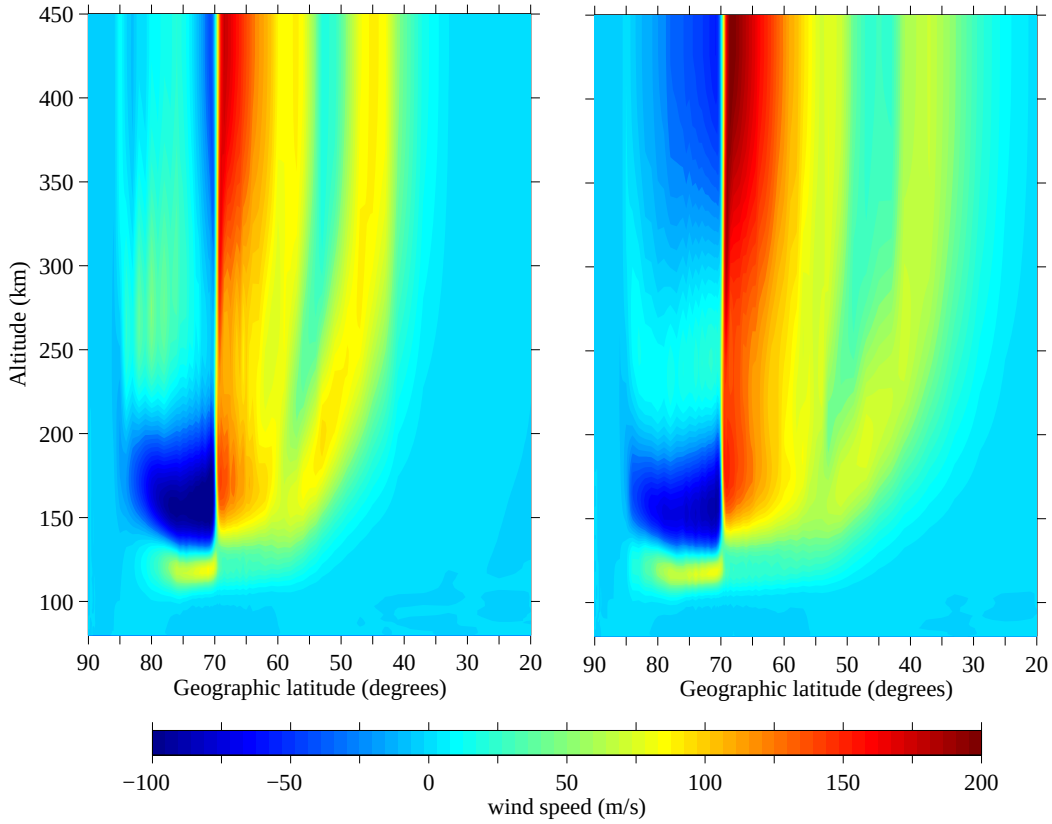

Figure 4.21: Meridional winds at the 90 and 120 minute marks of the simulation with the modified meridional momentum source term.

> poleward side due to the higher density at the lower altitude.

and *Chang and St.-Maurice* [1991], who state that the 'jet' is

> ... a remnant of the circulation pattern induced by Joule heating during the simulation.

The asymmetry about the convection channel at higher altitudes also becomes more pronounced when the Coriolis and centrifugal terms are set to zero in the meridional momentum equation; the lack of an equatorwards acceleration allows a larger pressure build-up to occur in the polar regions, which in turn, generates stronger equatorwards winds. It should now be quite obvious that the Coriolis and centrifugal terms are playing a pivotal role in the dynamics of the meridional wind field when large zonal winds are present and should not be neglected in *any* high resolution simulations that involve large zonal wind speeds, i.e., any high resolution simulations that employ Cartesian or cylindrical coordinates should incorporate a centrifugal effect into the source term of the meridional momentum equation. Even more insight into the evolution of the meridional and zonal wind fields could be obtained by decomposing the wind fields into rotational and divergent components, but such an undertaking is beyond the scope of this thesis.

The meridional jet and the equatorwards winds at 110 km also make the meridional winds in the lower thermosphere quite variable; vertical profiles of the meridional winds at 69.7° and 70.3° latitude are plotted in Figure 4.22 and exhibit significant differences above 130 km. The wind shears on the poleward side of the auroral zone are especially large, but such wind shears are not that uncommon in the lower thermosphere (*Larsen* [2002] and references therein). While it is tempting to compare the wind shears obtained in these runs with actual observations, the omission of the background winds in this simulation severely limits its ability to reproduce such observations; *Fuller-Rowell* [1985], *Brinkmann et al.* [1992], and *Parish et al.* [2003] have all demonstrated the importance of the background winds in the evolution of the lower thermosphere wind system during disturbed conditions.

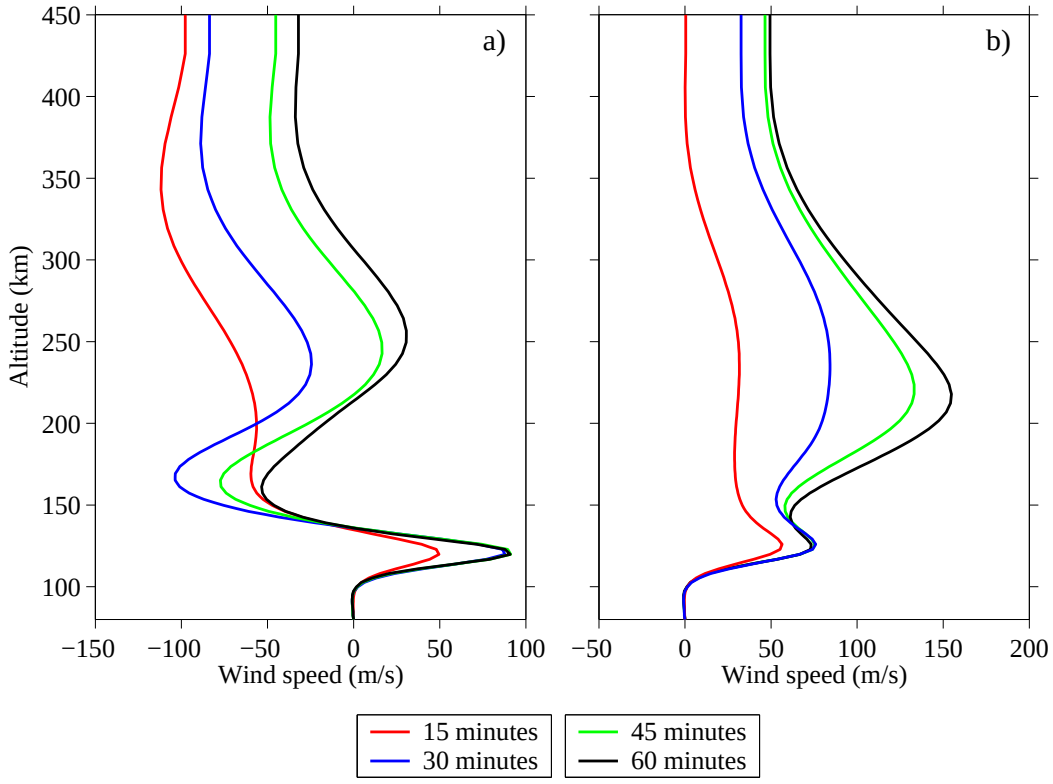

Figure 4.22: Meridional winds in the auroral zone at selected times of the 'reference run' simulation. The wind shears on the poleward side of the auroral zone at 70.3^{o}N (panel a) are much stronger than the wind shears on the equatorial side of the auroral zone at 69.7^{o}N (panel b).

The 'ramp-up' time of the electric field has a very significant impact on the meridional wind speeds (Figure 4.23), with the quicker 'ramp-up' times generating substantially larger wind speeds in the early parts of the simulation. The different 'ramp-up' times also excited waves of different frequencies in the lower thermosphere as the meridional acceleration time series derived from the meridional wind time series at 180 km could not be 'time-lagged' to coincide with each other. The waves being generated at 300 km were essentially monochromatic, allowing the meridional acceleration time series to be aligned (not shown). When the meridional accelerations at 300 km were plotted with the neutral densities at 300 km, they exhibited the same 3 stage behaviour exhibited by the zonal acceleration time series: an initial stage where the meridional acceleration increased *quadratically*, an intermediate stage where the meridional acceleration continued to increase at a slower, non-linear rate, and a final stage, where the meridional acceleration has already reached its maximum value and is decreasing (not shown). The transition between the first and second stages of the meridional acceleration coincided with the neutral density increase, so the same arguments made in the zonal wind section can be applied to this situation.

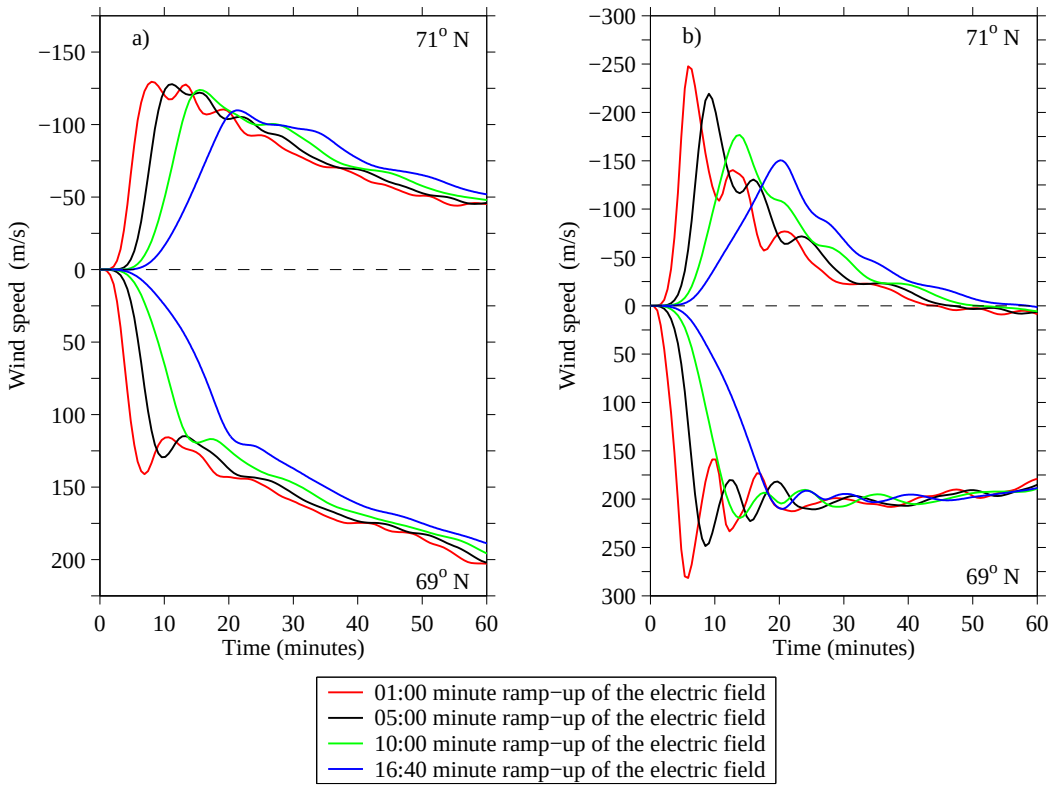

Figure 4.23: Meridional wind time series at selected altitudes in the auroral zone as a function of electric field 'ramp-up' time. The meridional winds at 180 km (right panel) and 300 km (left panel) are quite sensitive to the 'ramp-up' time of the electric field; the maximum wind speeds at 180 and 300 km can vary by 25 and 100 m/s during the initial acceleration phase.

4.3 Vertical winds

4.3.1 In the auroral zone

Contour plots of the 'reference' run vertical winds at 10 minutes, 20 minutes, 30 minutes, 40 minutes, 50 minutes, and 60 minutes are presented in Figure 4.24. The most obvious feature of these contour plots is that the vertical winds reach their maximum values very early in the simulation. To highlight this rapid acceleration of the vertical winds, time series of the vertical wind speeds at altitudes of 120 km, 180 km, 240 km, and 300 km at 70°N have been plotted in Figure 4.25. There are essentially two components to these time series: a 'DC' component that follows the evolution of the electric field in this simulation which eventually 'levels off' at the 1000 second mark of the simulation, and an oscillating component in the latter part of the time series that can be attributed to waves in the simulation.

This rapid acceleration of the vertical winds in the first 20 minutes of the simulation is, not surprisingly, due to Joule heating. This can be clearly seen in the vertical momentum balance plots of Figure 4.26, where the vertical pressure gradient term exceeds the gravitational field strength (the black curves) at all of the heights presented. It should also be noted that there are other sources of vertical momentum in the auroral zone: the meridional advection term, $-\frac{v}{r}\left(\frac{\partial u}{\partial \theta}\right)$, is a source of momentum at the higher altitudes (the red curves) while the Coriolis force (the cyan curves) and the w^2/r 'metric' term (the magenta curves) are making minor contributions at all of the heights being presented. Ion drag (the green curves), viscosity (the yellow curves), and vertical advection (the blue curves) are all sinks of vertical momentum, with ion drag being the most important at lower altitudes and viscosity becoming important at higher altitudes.

The dominance of the Joule heating driven pressure gradients in the first 20 minutes of the simulation persists throughout the simulation in the lower regions of the auroral zone (Figures 4.27-4.28). The ion drag (green) term also remains the principal sink of momentum at these altitudes, with the meridional advection (red), vertical advection (blue) and Coriolis (cyan) terms making minor contributions to the overall

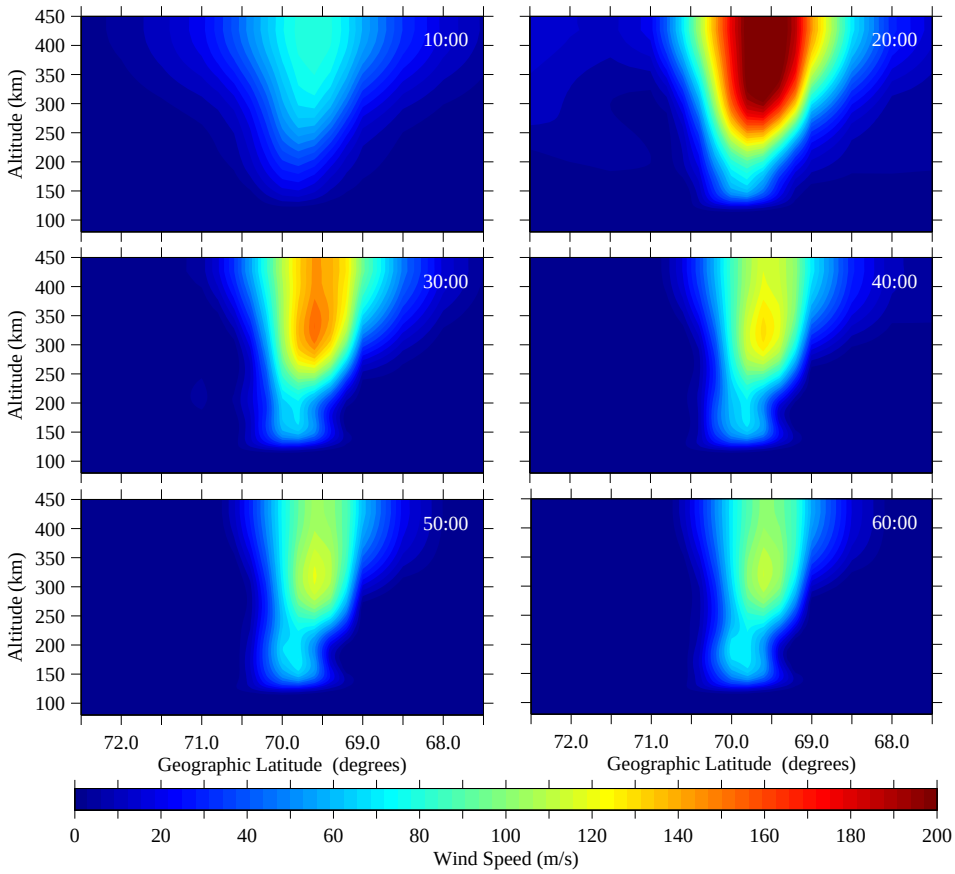

Figure 4.24: 'Reference run' vertical winds for the first hour of the simulation. The wind speeds have been contoured in 5 m/s intervals and reach their maximum values around the 20 minute mark of the simulation. The downwards (negative) component of the vertical wind speed field has been set to zero to enhance the legibility of the plot.

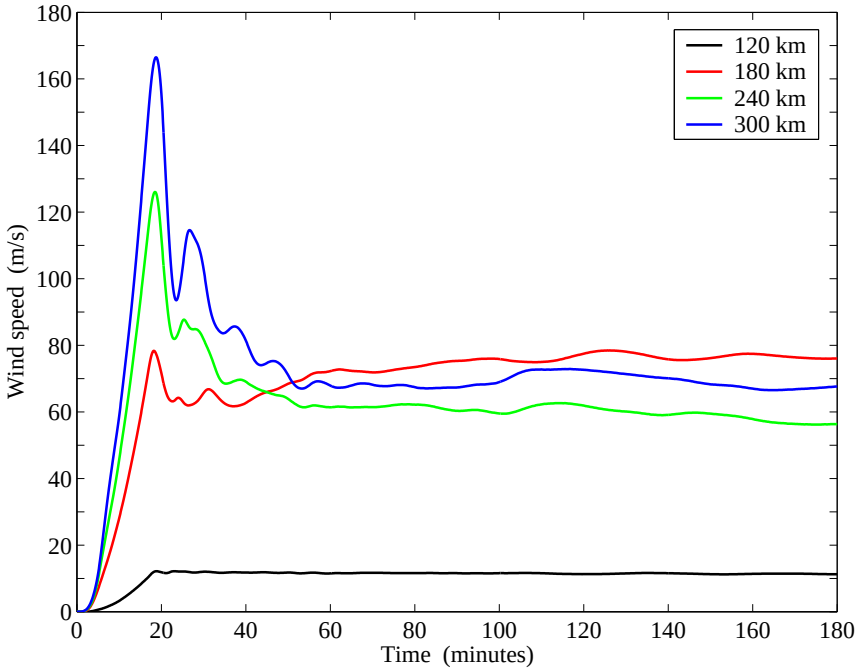

Figure 4.25: Vertical wind time series at 70°N for selected altitudes.

balance. It should be emphasized that an essentially steady state balance is achieved in this region by the 30 minute mark of the simulation, resulting in a persistent positive (radially upwards) acceleration of 0.015 m/s^2. The evolution of the vertical wind field at higher altitudes is not as easy to interpret as the pressure gradient term cannot be used as a proxy for Joule heating as the pressure gradient is not greater than the gravitational field strength at these heights. The Coriolis, meridional advection and w^2/r terms are opposing the pressure gradient-gravitational field differences at the higher altitudes (Figures 4.27-4.28).

The vertical wind speeds are also very sensitive to the effects of the electric field 'turn-on' time (Figure 4.29). Vertical wind speeds in excess of 100 m/s have been observed by ground-based Fabry-Perot interferometers (FPIs) at high latitudes since the early 1980s [*Rees et al.*, 1984; *Sica et al.*, 1986; *Conde and Smith*, 1995; *Smith and Hernandez*, 1995; *Price et al.*, 1995; *Innis et al.*, 1996, 1997; *Ishii et al.*, 2001], so the results being presented here are not unreasonable. It should be re-emphasized that the 1 minute 'ramp-up' time results may not be the most physically realistic -

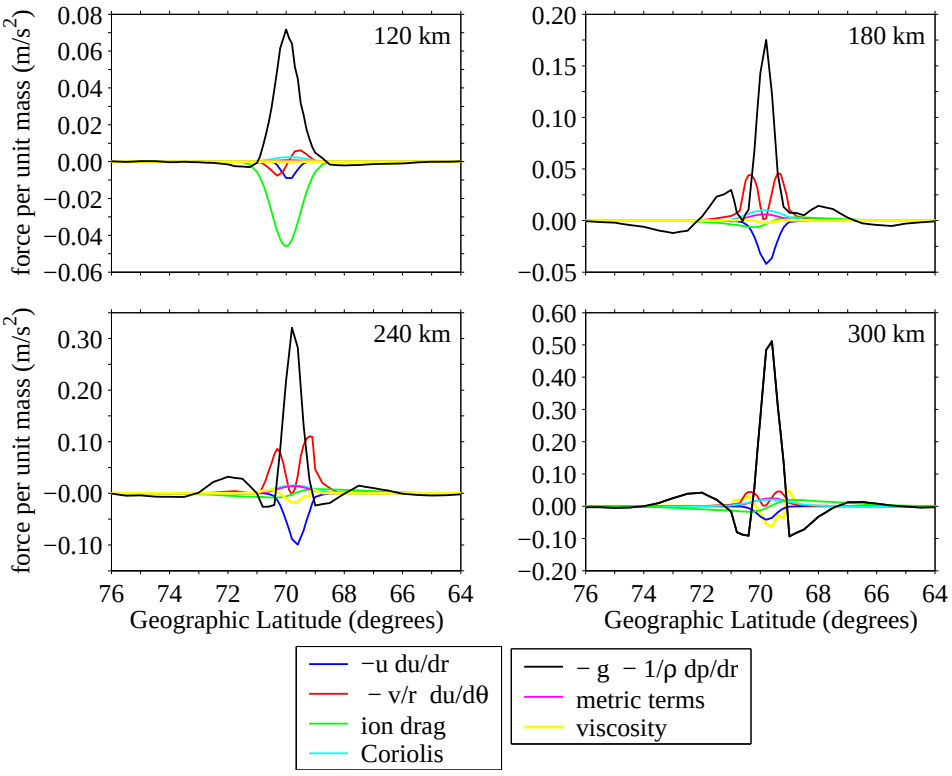

Figure 4.26: Vertical momentum balances in the auroral zone at the 15 minute mark of the 'reference run' simulation. The vertical pressure gradient terms (the black curves) are the dominant acceleration mechanism at the 15 minute mark of the simulation, with the meridional advection terms (the red curves) making making minor contributions at most heights.

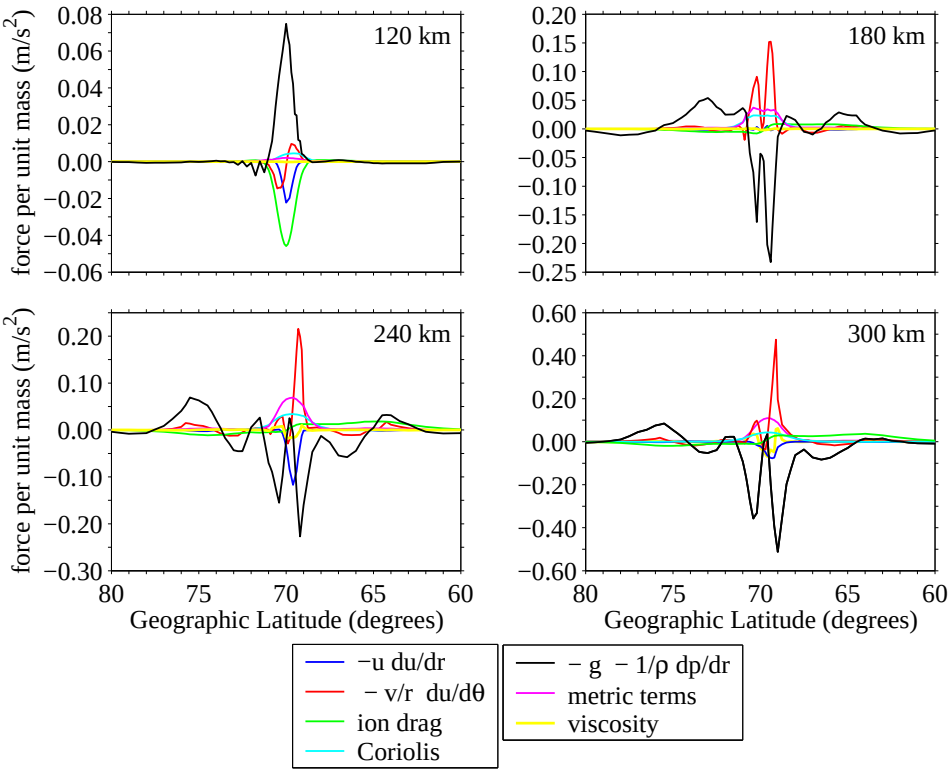

Figure 4.27: Vertical momentum balances in the auroral zone at the 30 minute mark of the 'reference run' simulation. The vertical pressure gradient terms (the black curves) are the dominant acceleration mechanism at the minute mark of the simulation, with the meridional advection terms (the red curves) making making minor contributions at most heights.

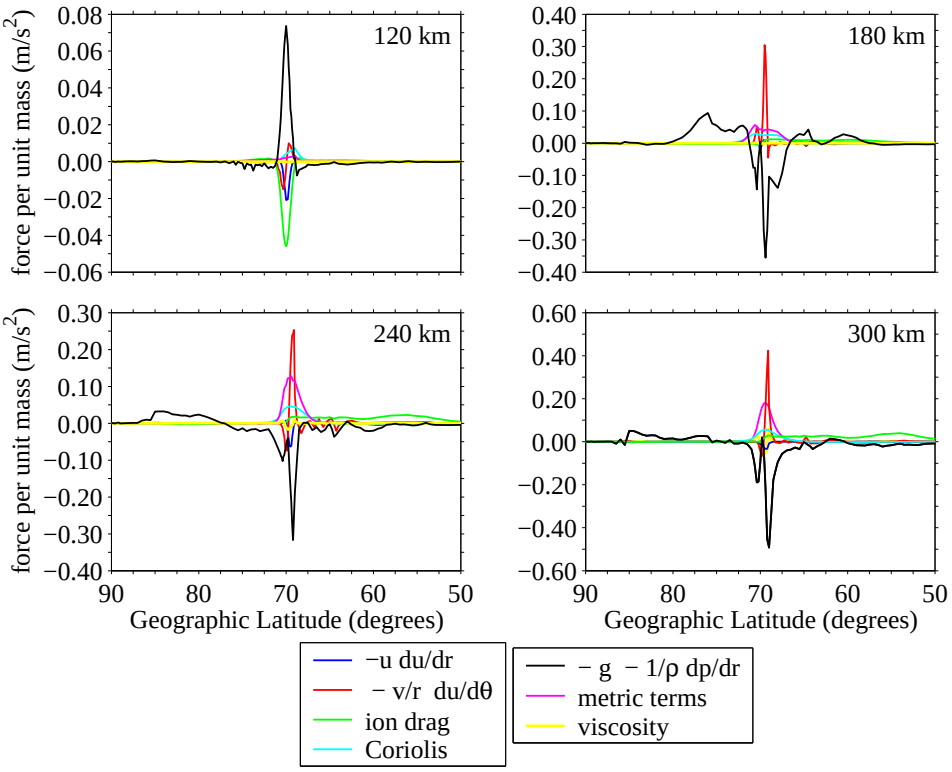

Figure 4.28: Vertical momentum balances in the auroral zone at the 60 minute mark of the 'reference run' simulation. The vertical pressure gradient terms (the black curves) are the dominant acceleration mechanism at the minute mark of the simulation, with the meridional advection terms (the red curves) making making minor contributions at most heights.

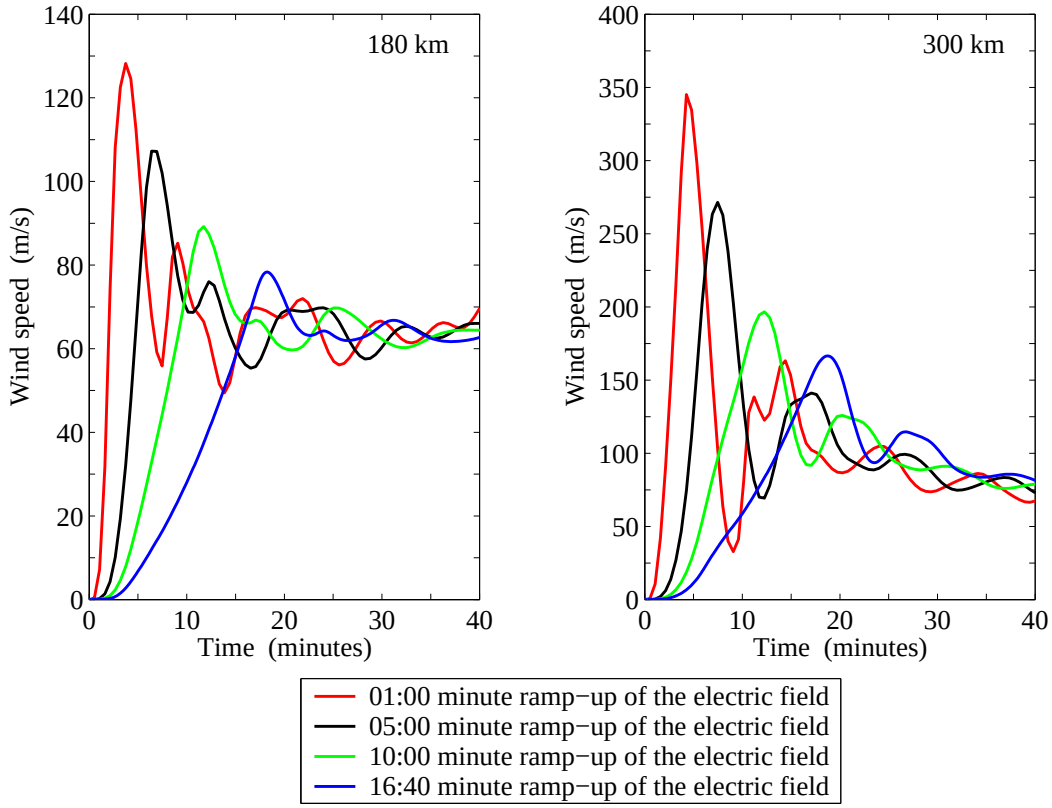

Figure 4.29: Vertical wind time series at selected altitudes in the auroral zone as a function of electric field 'ramp-up' time. The vertical winds at 180 km (right panel) and 300 km (left panel) are quite sensitive to the 'ramp-up' time of the electric field; the maximum wind speeds at 180 and 300 km can vary by 50 and 200 m/s during the initial acceleration phase.

they are being presented to see how quickly the thermosphere can respond to large energy inputs.

4.3.2 In the sub-auroral zone

Contour plots of the 'reference run' vertical winds at 30 minutes, 60 minutes, 90 minutes, 120 minutes, 150 minutes, and 180 minutes are presented in Figure 4.30 and clearly show the propagation of waves away from the auroral region. The generation and propagation of gravity waves in the auroral regions of the thermosphere is an active area of research, but time limitations have not allowed this topic to be explored

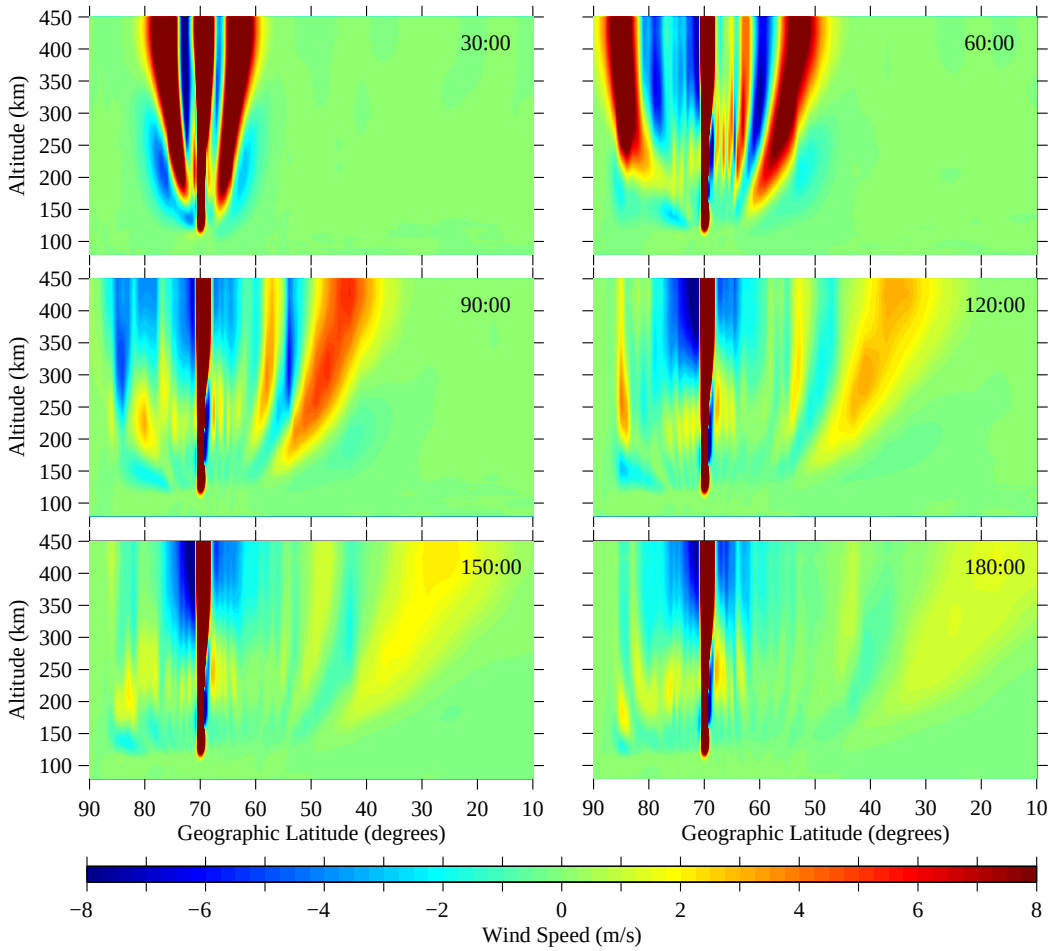

Figure 4.30: Vertical winds outside of the auroral zone for the 'reference run' simulation. The wind speeds have been contoured in 0.2 m/s intervals and clearly illustrate that waves are propagating away from the auroral region.

in any detail.

4.4 Density and Composition Changes

Contour plots of the 'reference' run density perturbations at 10 minutes, 20 minutes, 30 minutes, 40 minutes, 50 minutes, 60 minutes, 75 minutes, 90 minutes, 120 minutes, 150 minutes, and 180 minutes are presented in Figures 4.31-4.32 (the density perturbation is defined as the percentage change in neutral density). Associated changes in the composition fractions of atomic oxygen, molecular oxygen and molecular nitrogen

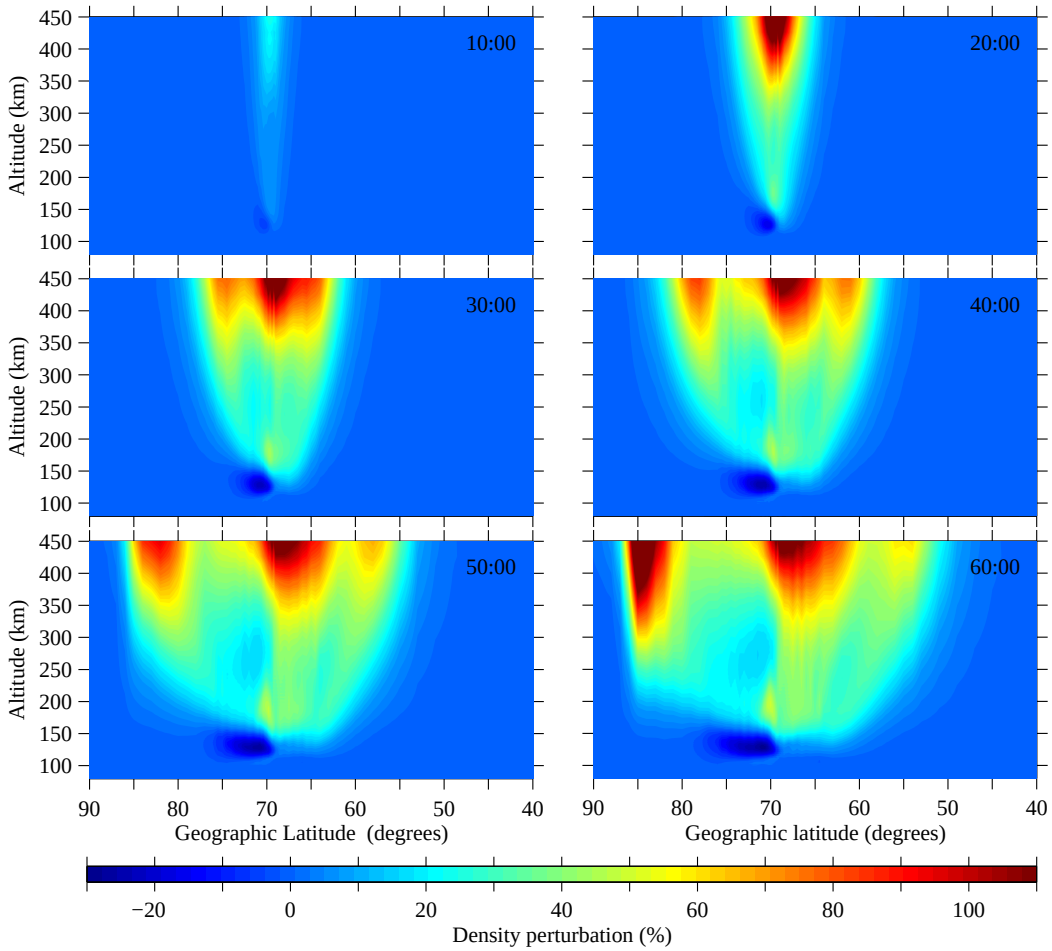

Figure 4.31: Density perturbations for the first hour of the 'reference run' simulation. The density perturbations have been contoured in intervals of 1% and reach their maximum values at the upper boundary of the auroral zone.

are plotted in Figures 4.33-4.34, 4.35-4.36, and 4.37-4.38 respectively.

During the first 20 minutes of the 'reference run' simulation there is a significant increase in neutral densities above 150 km in the auroral zone (Figure 4.31). This rapid increase in the neutral densities 'mirrors' the increase in the strength of the electric field, which strongly suggests that these neutral density increases are due to the effects of Joule heating. An examination of the continuity balances at 180 km and 300 km confirms this, with the velocity divergence terms almost balancing the vertical advection term (the velocity divergence terms represent the expansion of the neutral gas due to Joule heating - see Figure 24 in *Chang and St.-Maurice* [1991]).

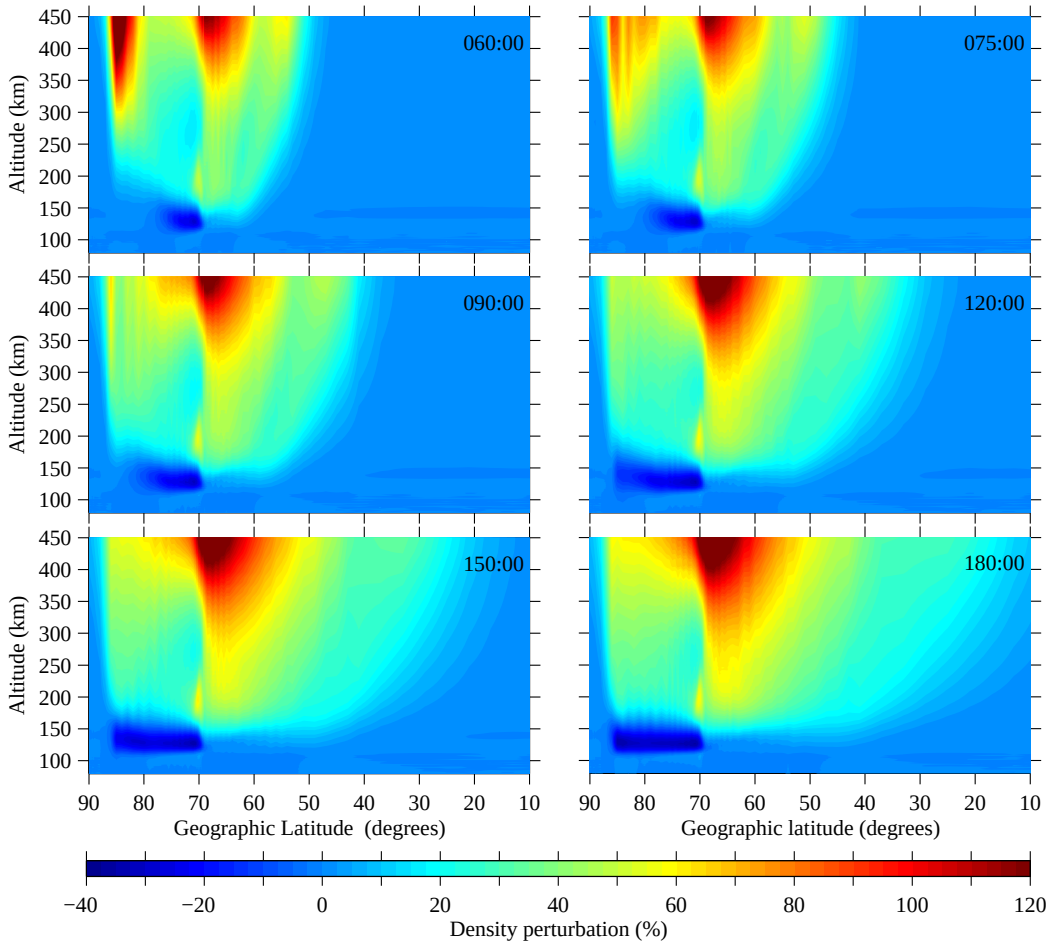

Figure 4.32: Density perturbations for the final two hours of the 'reference run' simulation. The density perturbations have been contoured in intervals of 1%.

This upwelling in the auroral region can also be seen in the composition fractions of the various species, with the atomic oxygen composition fractions decreasing in the auroral region and the molecular oxygen and nitrogen composition fractions increasing in the auroral region (Figures 4.33,4.35, and 4.37 respectively). At altitudes between 110 km and 130 km, one observes a density depletion on the poleward side of the convection channel. This depletion was also obtained in the model runs of *Chang and St.-Maurice* [1991] and is due to the meridional wind 'jet' at these altitudes (see section 4.3) transporting the neutral gas into the auroral region where it is then transported upwards. The LOGACS accelerometer observed a similar density depletion during a strong geomagnetic storm in May 1967 (Figure 1.12), providing some evidence that this phenomenon is real.

As the simulation progresses, the density depletion on the polar side of the auroral zone continues to grow. The neutral densities outside of the auroral region also continue to increase, with wave-like perturbations 'superimposed' on the slowly propagating enhancement. The large density enhancement near the polar boundary should be ignored as this feature is due to the assumption of zonal symmetry in the model (see the meridional wind section). The composition fractions in the auroral zone continue to exhibit signs of upwelling during this period, with the atomic oxygen and molecular nitrogen composition fractions becoming approximately equal at 400 km 60 minutes into the simulation. Wave-like perturbations can also be observed in the composition fractions outside of the auroral zone, with the 'lead' wave packet generating substantial mixing and overturning as it propagates. A composition fraction 'tongue' follows this lead wave packet, with the composition fractions in the 'tongue' being approximately equal to the composition fractions in the auroral zone at those altitudes. A second much broader and slower moving enhancement/depletion can also be seen in the composition fractions around the auroral zone; this enhancement/depletion in the composition fractions is being driven by the large meridional winds that now exist on the equatorward side of the auroral region. It should also be noted that the changing composition fractions do not affect the neutral densities at the present time as the MacCormack scheme currently being used is not positive

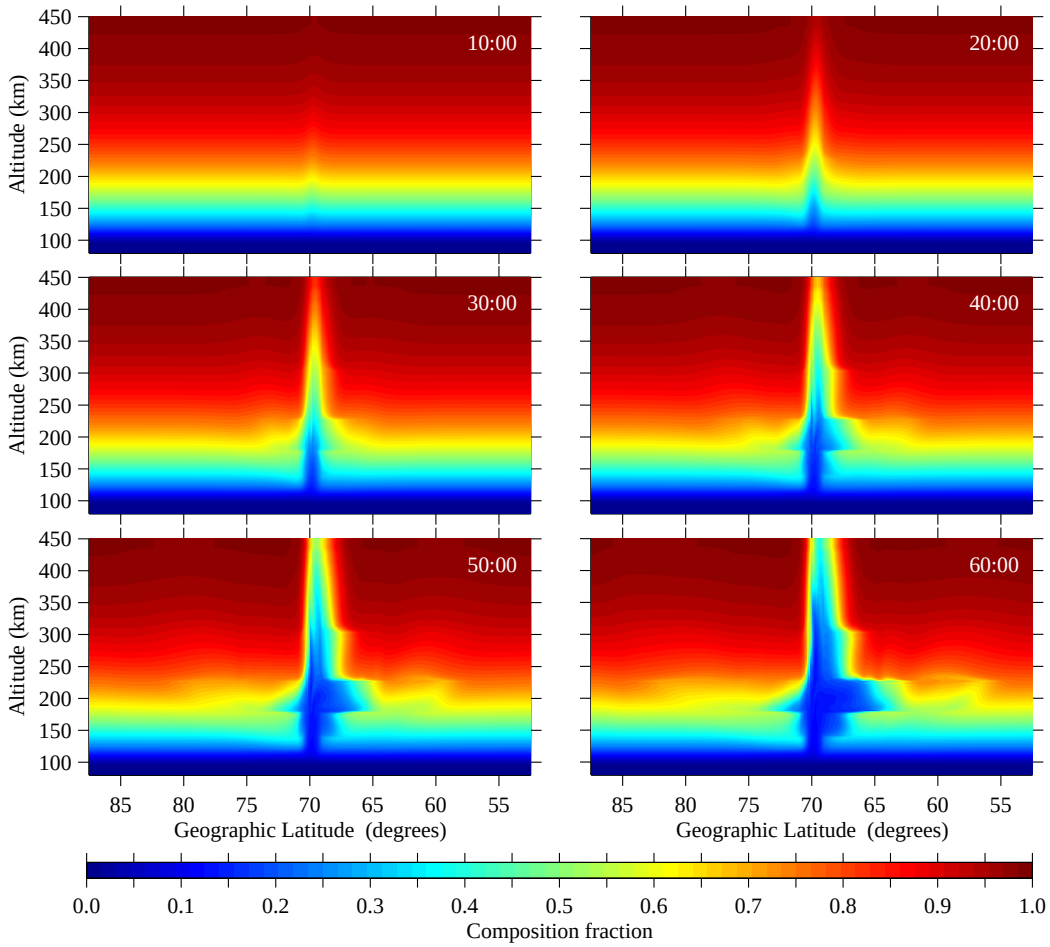

Figure 4.33: Atomic oxygen composition fractions for the first hour of the 'reference run' simulation.

definite for the entire simulation, i.e. negative composition fractions occur in localized regions near the end of the simulation.

The development of the meridional wind jet on the equatorial side of the auroral zone in the latter part of the 'reference run' simulation accelerates the changes that are already occurring in the density field and composition fractions. The composition fraction 'tongue' is affected most, with the 'tongue' eventually being accelerated to the front of the 'lead' wave packet.

The number density ratios of N_2 (blue), O_2 (red), O (green) and $O + 2\,O_2$ (black) have been plotted as a function of latitude at the 90 minute mark of the 'reference run' simulation in Figure 4.42. The number densities of N_2 are 6.5 times greater in

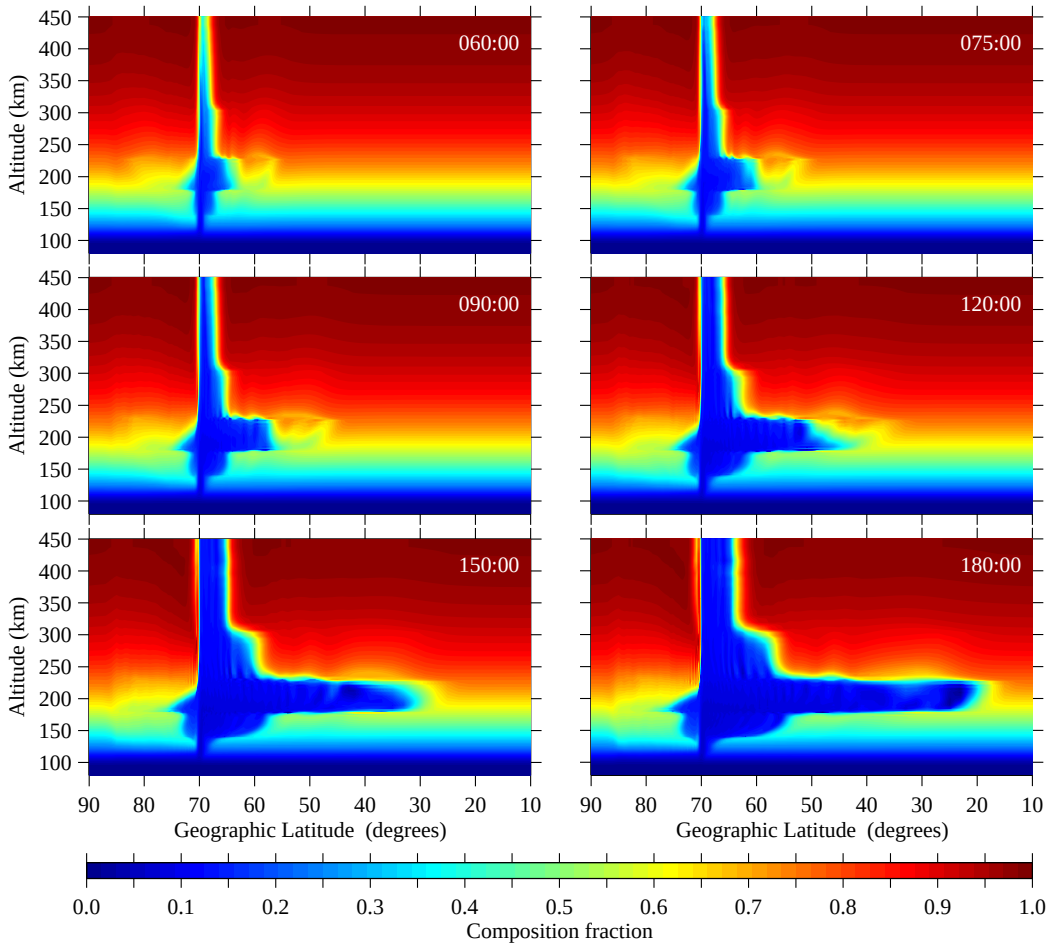

Figure 4.34: Atomic oxygen composition fractions for the final two hours of the 'reference run' simulation.

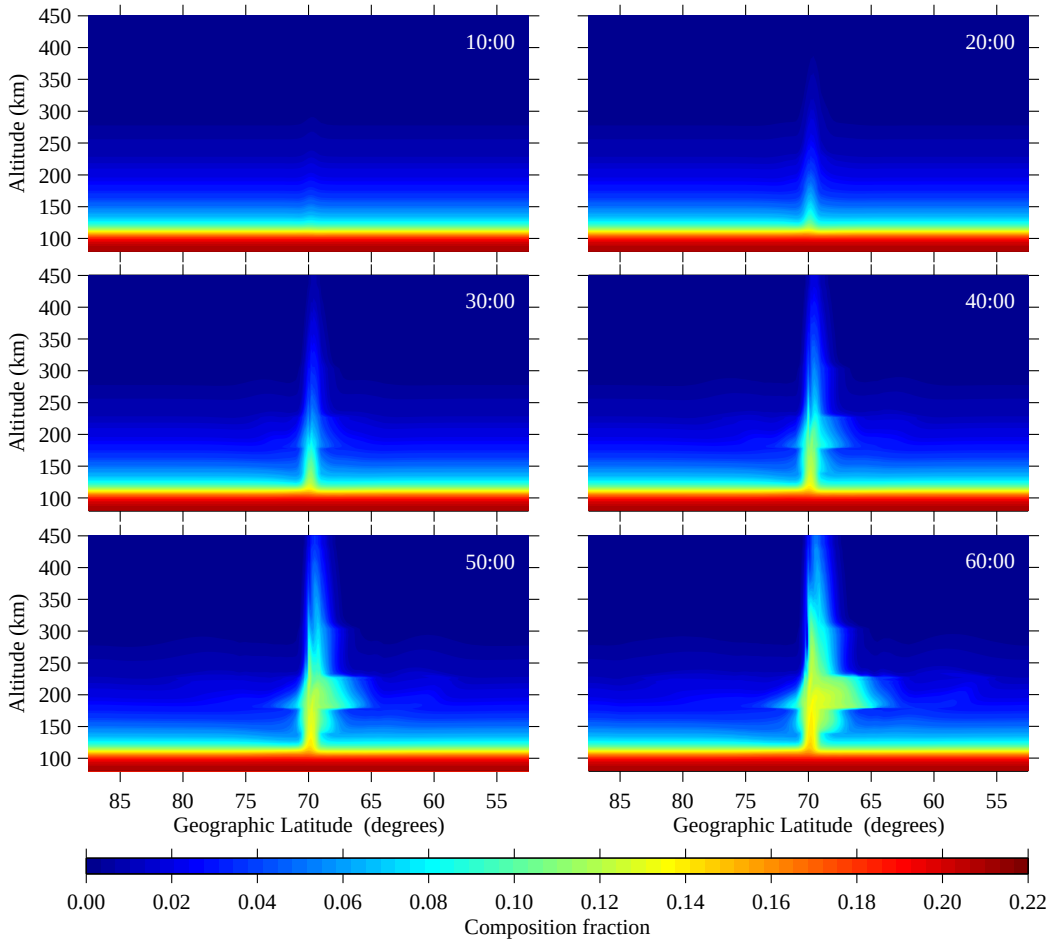

Figure 4.35: Molecular oxygen composition fractions for the first hour of the 'reference run' simulation.

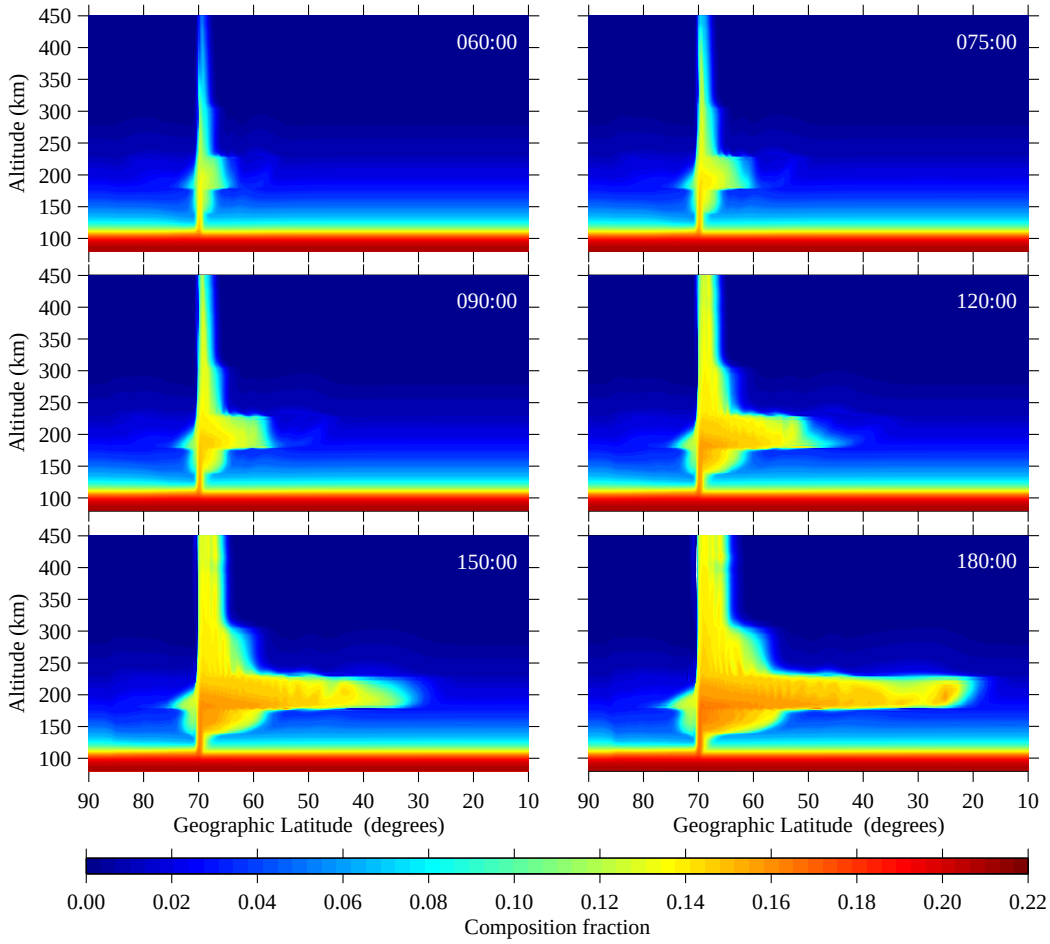

Figure 4.36: Molecular oxygen composition fractions for the final two hours of the 'reference run' simulation.

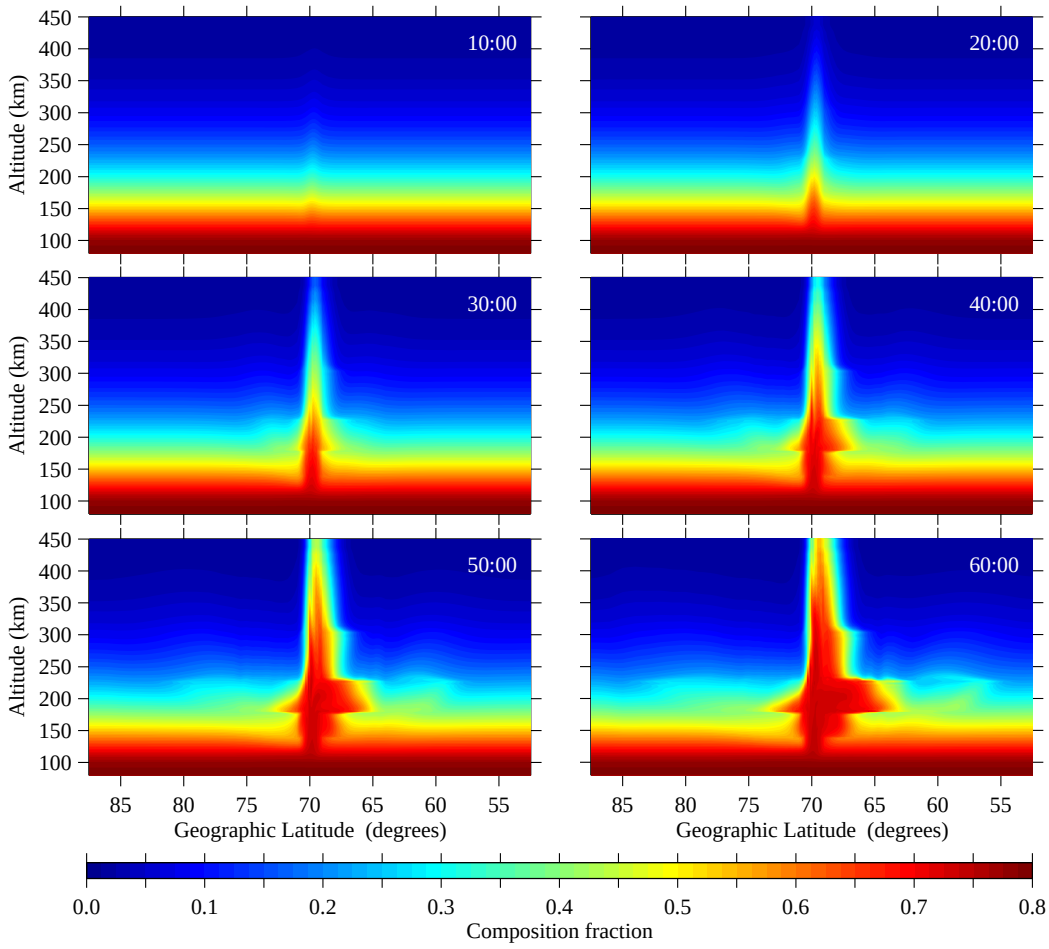

Figure 4.37: Molecular nitrogen composition fractions for the first hour of the 'reference run' simulation.

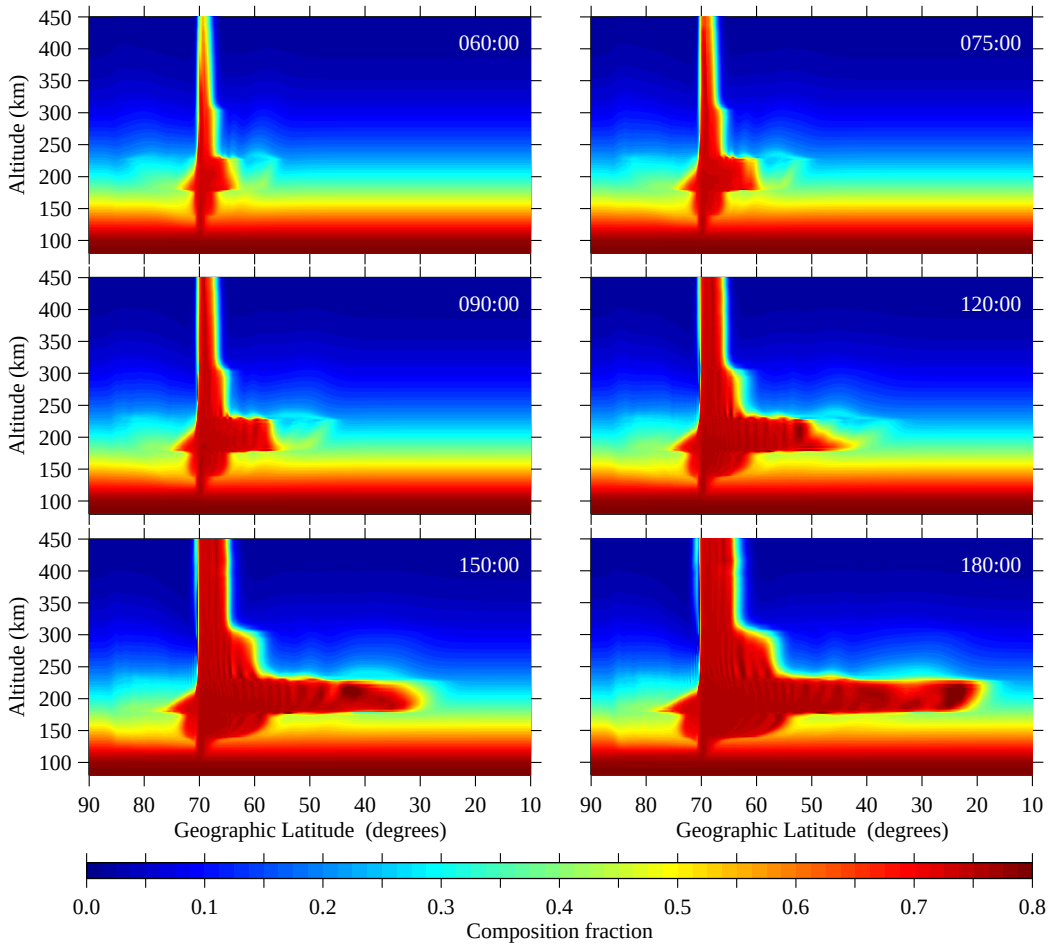

Figure 4.38: Molecular nitrogen composition fractions for the final two hours of the 'reference run' simulation.

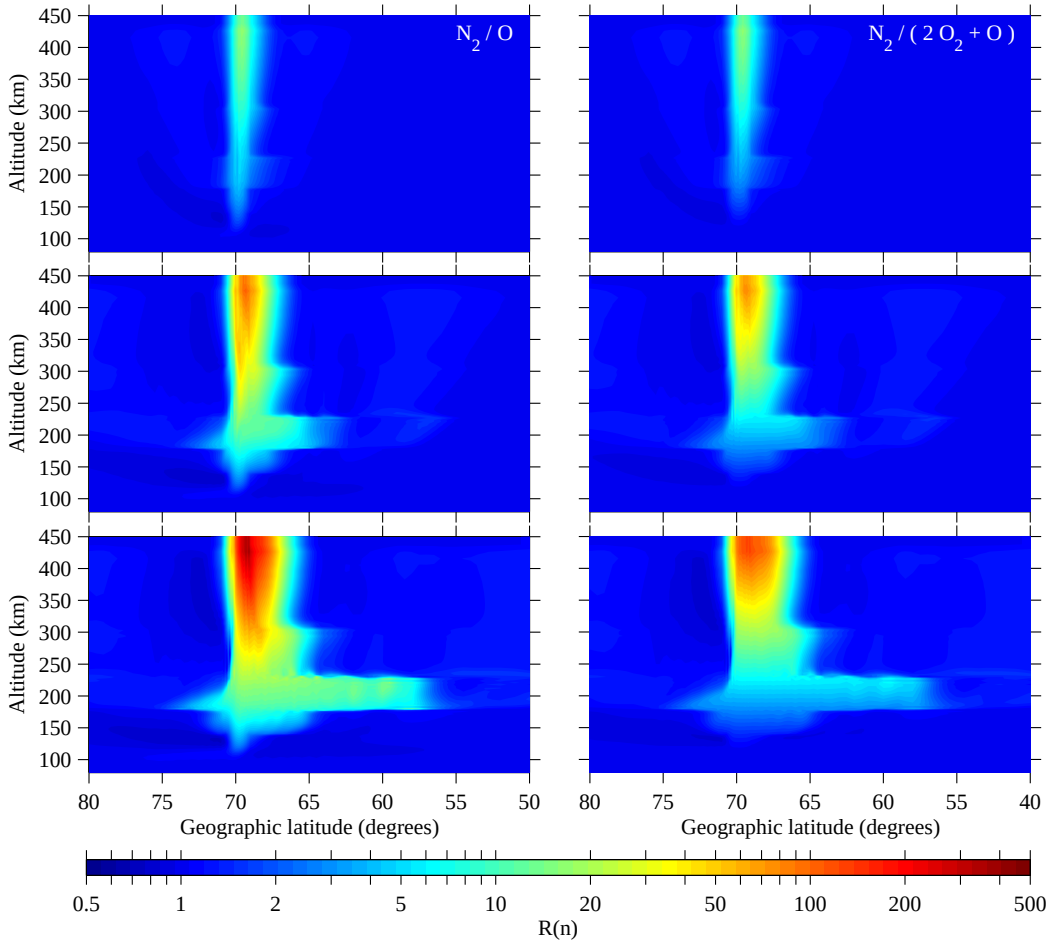

Figure 4.39: N_2/O and $N_2/(O + 2\,O_2)$ ratios during the first ninety minutes of the 'reference run' simulation. The ratios of N_2/O (left panels) and $N_2/(O + 2\,O_2)$ (right panels) with respect to their quiet-time values have been plotted at the 30 minute (top panels), 60 minute (middle panels), and 90 minute (bottom panels) marks of the simulation. There are no appreciable differences between these two ratios at the 30 minute mark of the simulation. By the 90 minute mark of the simulation, the N_2/O ratios are significantly larger than the $N_2/(O + 2\,O_2)$ ratios in the auroral zone and in the composition 'tongue'.

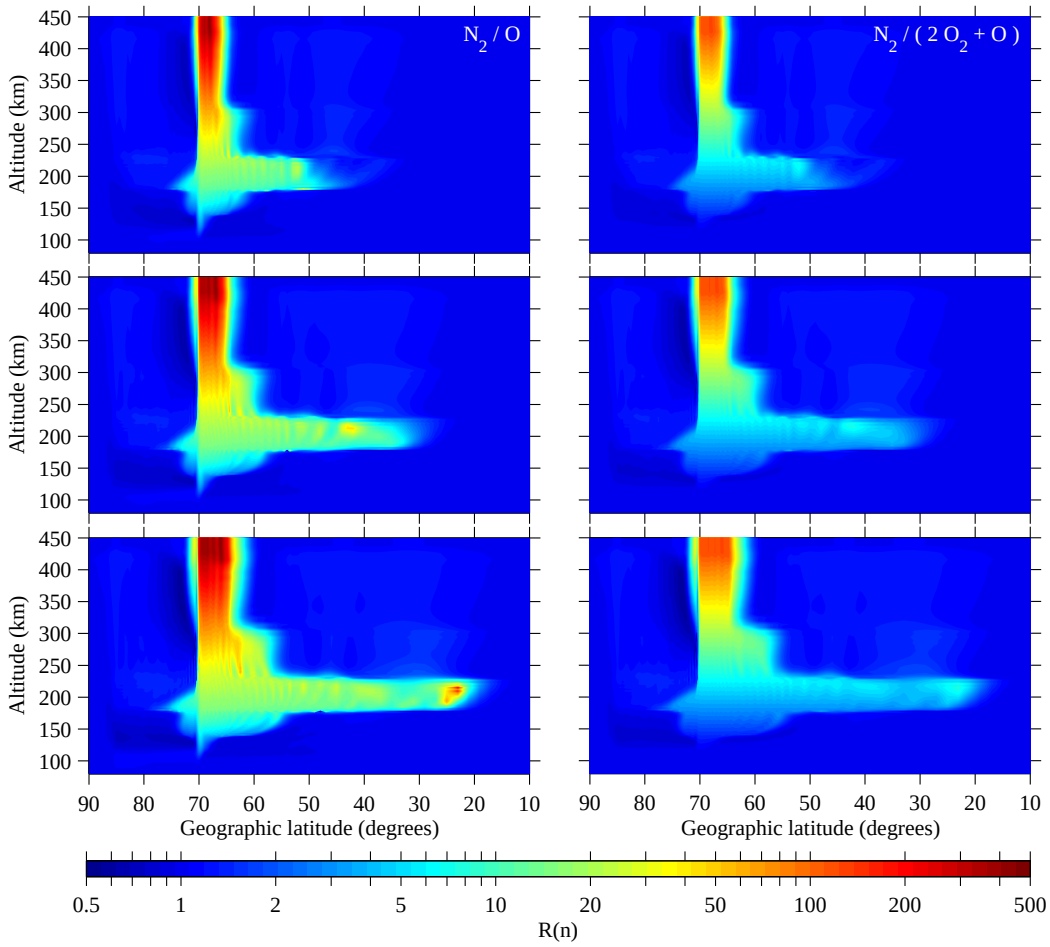

Figure 4.40: N_2/O and $N_2/(O + 2\,O_2)$ ratios during the third hour of the 'reference run' simulation. The ratios of N_2/O (left panels) and $N_2/(O + 2\,O_2)$ (right panels) with respect to their quiet-time values have been plotted at the 120 minute (top panels), 150 minute (middle panels), and 180 minute (bottom panels) marks of the simulation. The N_2/O ratios are much larger than the $N_2/(O + 2\,O_2)$ ratios in the auroral zone and in the composition 'tongue'.

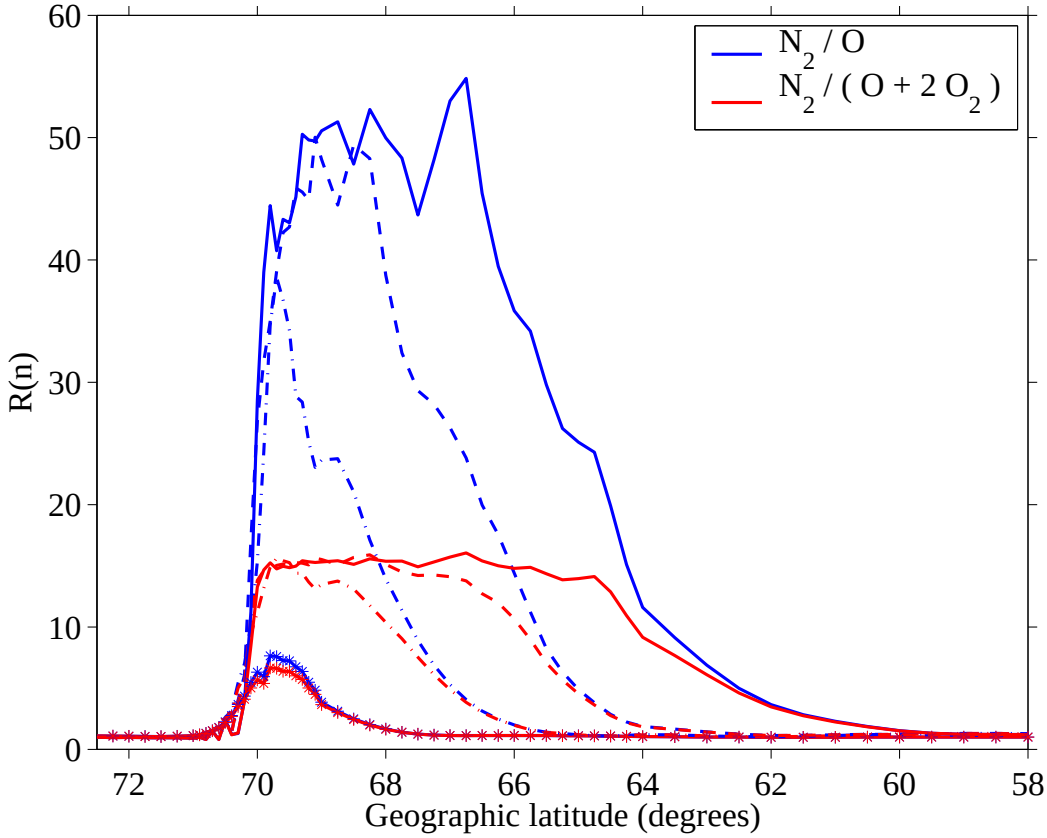

Figure 4.41: N_2/O and $N_2/(O + 2\,O_2)$ ratios at 280 km in the 'reference run' simulation. The ratios of N_2/O (blue) and $N_2/(O + 2\,O_2)$ (red) with respect to their quiet-time values are plotted at the 30 minute (star), 60 minute (dash-dot), 90 minute (dashed), and 120 minute (solid) marks of the simulation. The N_2/O ratios are much larger than the $N_2/(O + 2\,O_2)$ ratios in the second hour of the simulation.

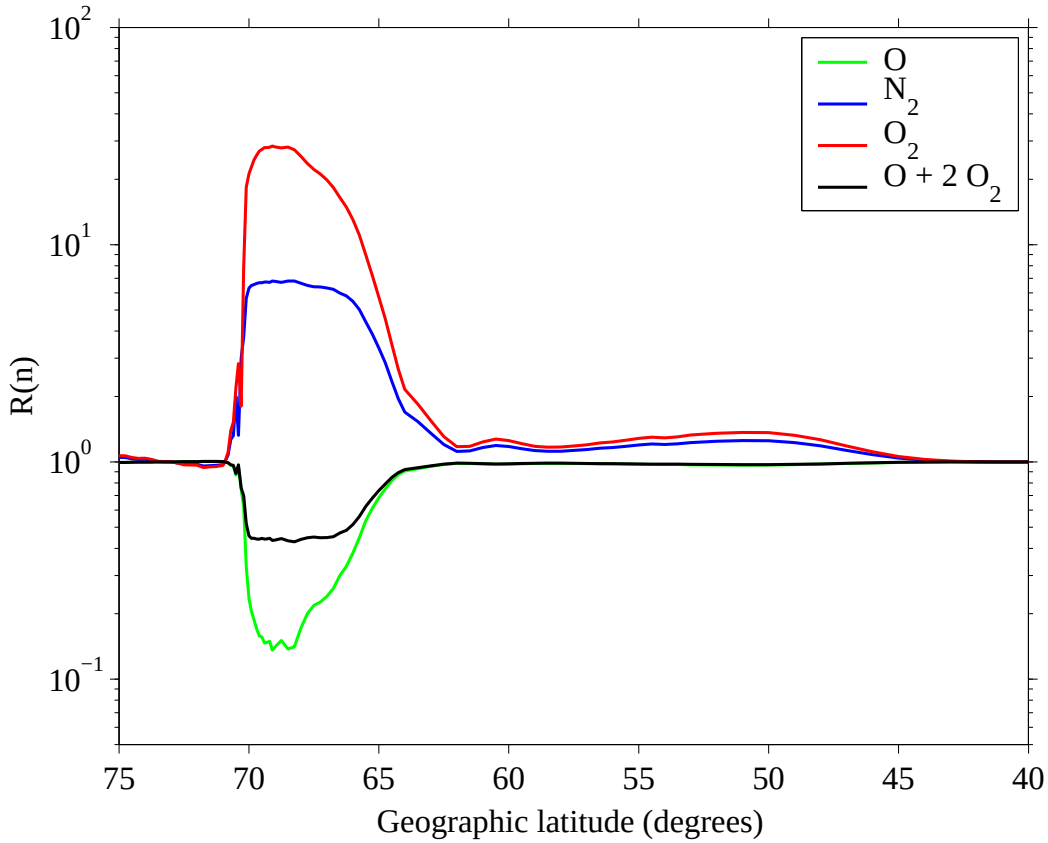

Figure 4.42: Composition ratios at 280 km for the 'reference run' simulation. The ratios of N_2 (blue), O_2 (red), O (green) and $O + 2\,O_2$ (black) have been plotted as a function of latitude at the 90 minute mark of the simulation. The $O + 2\,O_2$ ratios are much larger than the atomic oxygen ratios in the region of heating as significant amounts of molecular oxygen have been transported upwards.

the region of heating and exhibit a steep latitudinal gradient, in good agreement with ESRO-4 observations of the geomagnetic storm on October 29, 1973 (Figure 1.18). The simulated $O + 2\,O_2$ ratios are also consistent with the observed ESRO-4 atomic oxygen ratios for October 29, 1973, which is not surprising if one remembers that ESRO-4 was a closed source mass spectrometer. The ability of these simulations to separate the $O + 2O_2$ ratio into atomic and molecular oxygen shows that the concentrations of molecular oxygen are not negligible at high latitudes during strongly disturbed conditions; the molecular oxygen number densities are 25 times greater in the region of heating while the atomic oxygen concentrations are actually one-seventh of their quiet-time values. GUVI measurements have also shown that significant atomic oxygen depletions/molecular nitrogen enhancements can propagate to lower latitudes (Figure 1.22), providing additional support that the composition change mechanism being presented can occur.

This apparent overestimation of atomic oxygen number densities by ESRO-4 means that the N_2/O ratios inferred from ESRO-4 could also be in error. Plots of the $N_2/(O + 2\,O_2)$ ratio at 280 km at the 30, 60, 90, and 120 minute marks of the simulation are presented in Figure 4.43 and approach a value of 15, in good agreement with ESRO-4 observations on October 29, 1973 [*Prölss*, 1980]. The actual N_2/O ratios may be much higher, with values of 45-55 being more realistic if the majority of the molecular oxygen being detected by ESRO-4 at these altitudes is due to vertical transport.

While the amounts of molecular nitrogen and oxygen transported upwards are sensitive to the heating rates, the fact that significant amounts of molecular oxygen can be transported upwards with molecular nitrogen suggests that high latitude mass spectrometer measurements may need to be reexamined in light of these new results. Localized enhancements of molecular nitrogen are routinely observed in the polar cap region at all levels of geomagnetic activity [*Hedin and Reber*, 1972; *Reber and Hedin*, 1974; *Taeusch and Hinton*, 1975; *Laux and von Zahn*, 1979] and, if they are accompanied by enhancements in molecular oxygen, then mass spectrometer observations of atomic oxygen in this region may be in significant error. The retrieval of atomic

hydrogen densities from mass spectrometer measurements at high latitudes may also be in error as the number densities of atomic hydrogen are normally derived from atomic oxygen measurements (see *Sanatani et al.* [1995] and references therein for more details).

Chapter 5

Summary, Conclusions and Future Work

5.1 Summary and Conclusions

Several upgrades were incorporated into the model of *Chang and St.-Maurice* [1991]. The 'stretched grid' and improved upper boundary conditions enhanced the numerical stability of the model. The new 'UV/EUV heating' profile and CO_2 cooling parameterization allowed the initial conditions to be much more representative of a geomagnetically quiet thermosphere. The improved ion composition and altitude dependent **B** also improved the 'reality' of the simulations at higher altitudes.

While all of these upgrades improved the reality and robustness of the simulations, they did little to advance our state of knowledge about the disturbed thermosphere. The relaxation of the mean molecular mass gas assumption in the model, in conjunction with its nonhydrostatic formulation, showed that significant amounts of molecular nitrogen and oxygen could be transported upwards during geomagnetic disturbances. This transport of molecular oxygen to higher altitudes during geomagnetic disturbances had already been noted by several authors, but this thesis showed that the amounts of molecular oxygen being transported upwards could have a significant impact on the *interpretation* of mass spectrometer measurements during these conditions, i.e. a simulated total oxygen depletion of 50% at 280 km could mean that the atomic oxygen number densities at 280 km are actually one-seventh of their normal values while the simulated molecular oxygen concentrations are 25 times larger than normal. It was also shown that the neutral atmosphere can respond very quickly to large energy inputs; the simulated vertical and meridional wind speeds were very sensitive to how quickly the energy was deposited into the thermosphere.

5.2 Future Work

The composition changes generated in the model runs presented are very dependent upon the vertical winds in the auroral region. More model runs with variations in the strength, location, width and duration of **E** will provide some insight into the generation of the vertical winds, but a more realistic coupling of the thermosphere-ionosphere system will be necessary to capture the physics of this process correctly, i.e. heating due to particle precipitation and a self-consistent odd nitrogen chemistry scheme. The generation of localized negative composition fractions in the model should also be eliminated; this will most likely require the implementation of a more sophisticated transport scheme and/or stretched latitudinal grid.

The presence of wave-like perturbations in most of the dynamical fields of the model has been noted but not explored in any detail in this thesis. A closer examination of these perturbations should be undertaken and the results compared to available linear/nonlinear theories of acoustic wave / gravity wave generation and propagation. The decomposition of the dynamical fields into a rotational and a divergent should also provide some insight into the response of the thermosphere-ionosphere system.

REFERENCES

[1] Allen, J., L. Frank, H. Sauer, and P. Reiff, Effects of the March 1989 solar activity, *EOS Transactions*, *70*, 1486-1488, 1989.

[2] Allen, M., Y. L. Yung, and J. W. Waters, Vertical transport and photochemistry in the terrestrial mesosphere and lower thermosphere (50 - 120 km), *J. Geophys. Res.*, *86*, 3617-3627, 1981.

[3] Anderson, J. D., *Computational Fluid Dynamics: The Basics with Applications*, McGraw-Hill, New York, 1995.

[4] Banks, P. M., and G. Kockarts, *Aeronomy*, Academic Press, New York, 1973.

[5] Bates, D. R., The temperature of the upper atmosphere, *Proc. Phys. Soc. B*, *64*, 805-821, 1951.

[6] Bates, H. F., Thermospheric changes shortly after the onset of daytime joule heating, *Planet. Space Sci.*, *22*, 1625-1636, 1974.

[7] Bjorn, L. G., and F. Arnold, Mass spectrometric detection of precondensation nuclei at the arctic summer mesopause, *Geophys. Res. Lett.*, *8*, 1167-1170, 1981.

[8] Blelly, P.-L., A. Robineau, J. Lilensten, and D. Lummerzheim, 8-moment fluid models of the terrestrial high latitude ionosphere between 100 and 3000 km, in *Solar-Terrestrial Energy Program : Handbook of Ionospheric Models*, R. W. Schunk, ed., pp. 53-72, SCOSTEP Secretariat, Boulder, Colorado, 1996.

[9] Blitza, D., IRI 2000, *Radio Sci.*, *36*, 261-276, 2001.

[10] Blum, P., C. Wulf-Mathies, and H. Trinks, Interpretation of local thermospheric disturbances of composition observed by ESRO-4 in the polar region, *Space Res.*, *15*, 209-214, 1975.

[11] Bougher, S. W., D. M. Hunten, and R. G. Roble, CO_2 cooling in the terrestial planet thermospheres, *J. Geophys. Res.*, *99*, 14,609-14,622, 1994.

[12] Brinkman, D. G., R. L. Walterscheid, A. D. Richmond, and S. V. Venkateswaran, Wave-mean flow Interaction in the storm-time thermosphere: a two-dimen-sional model simulation, *J. Atmos. Sci.*, *49*, 660-680, 1992.

[13] Brinkman, D. G., R. L. Walterscheid, L. R. Lyons, D. C. Kayser, A. B. Christensen, J. R. Sharber, R. A. Frahm, and M. F. Larsen, E region neutral winds in the postmidnight diffuse aurora during the Atmospheric Response in Aurora 1 rocket campaign, *J. Geophys. Res.*, *100*, 17,309-17,320, 1995.

[14] Brasseur, G., and S. Solomon, *Aeronomy of the Middle Atmosphere*, D. Reidel Publishing Company, Dordrecht, 1986.

[15] Buonsanto, M. J., D. P. Sipler, G. B. Davenport, and J. M. Holt, Estimation of the $O - O^+$ collision frequency from coincident radar and Fabry-Perot observations at Millstone Hill, *J. Geophys. Res.*, *102*, 17,267-17,274, 1997.

[16] Carignan, G. R., and W. H. Pinkus, OGO-F-04 experiment description, *Tech. Note 08041-3-T*, Space Phys. Res. Lab., Univ. of Mich., Ann Arbor, 1968.

[17] Carignan, G. R., B. P. Block, J. C. Maurer, A. E. Hedin, C. A. Reber, and N. W. Spencer, The neutral mass spectrometer on Dynamics Explorer B, *Space Sci. Instrum.*, *5*, 429-441, 1981.

[18] Carpenter, M. H., A high-order compact numerical algorithm for supersonic flow, in *Twelfth International Conference on Numerical Methods in Fluid Dynamics*, Lecture Notes in Physics, Vol. 371, K. W. Morton, ed., 254-258, 1984.

[19] Chamberlain, J. W., and M. B. McElroy, Martian Atmosphere: The Mariner Occultation Experiment, *Science*, *152*, 21-25, 1966.

[20] Chang, C. A., and J.-P. St.-Maurice, Two-dimensional high-latitude thermospheric modeling: A comparison between moderately and extremely disturbed conditions, *Can. J. Phys.*, *69*, 1007-1031, 1991.

[21] Christensen, A. B., J. H. Hecht, R. L. Walterscheid, M. F. Larsen, and W. E. Sharp, Depletion of oxygen in aurora: Evidence for a local mechanism, *J. Geophys. Res.*, *102*, 22,273-22,278, 1997.

[22] Colegrove, F. D., W. B. Hanson, and F. S. Johnson, Eddy diffusion and oxygen transport in the lower thermosphere, *J. Geophys. Res.*, *70*, 4931-4941, 1965.

[23] Colegrove, F. D., F. S. Johnson, and W. B. Hanson, Atmospheric composition in the lower thermosphere, *J. Geophys. Res.*, *71*, 2227-2236,1966.

[24] Conway, R. R., R. R. Meier, and R. E. Huffman, Satellite observations of the OI 1304, 1356 and 1641 Å dayglow and the abundance of atomic oxygen in the thermosphere, *Planet. Space Sci.*, *36*, 963-973, 1988.

[25] Craig, R. A., and J. C. Gille, Cooling of the thermosphere by atomic oxygen, *J. Atmos. Sci.*, *26*, 205-209, 1969.

[26] Craven, J. D., A. C. Nicholas, L. A. Frank, D. J. Strickland, and T. J. Immel, Variations in the FUV dayglow after intense auroral activity, *Geophys. Res. Lett.*, *21*, 2793-2796, 1994.

[27] Dalgarno, A. M., R. C. McDowell, and A. Williams, The mobilities of ions in unlike gases, *Phil. Trans. Roy. Soc. London A*, *250*, 411-425, 1958.

[28] DeVries, L. L., Analysis and interpretation of density data from the Low-G acceleromter calibration system (LOGACS), *Space Res.*, *12*, 777-789, 1972.

[29] Dodd, J. A., R. B. Lockwood, E. S. Hwang, S. M. Miller, and S. J. Lipson, Vibrational relaxation of NO($\nu = 1$) by oxygen atoms, *J. Chem. Phys.*, *111*, 3498-3507, 1999.

[30] Drob, D. P., R. R. Meier, J. M. Picone, D. J. Strickland, R. J. Cox, and A. C. Nicholas, Atomic oxygen in the thermosphere during the July 13, 1982 solar proton event deduced from far ultraviolet images, *J. Geophys. Res.*, *104*, 4267-4278, 1999.

[31] Duff, J. W., and R. D. Sharma, Quasi-classical trajectory study of NO vibrational relaxation by collisions with atomic oxygen, *J. Chem. Soc., Faraday Trans.*, *93*, 2645-2649, 1997.

[32] Fehsenfeld, F. C., and E. E. Ferguson, Origin of water cluster ions in the D-region, *J. Geophys. Res.*, *74*, 2217-2222, 1969.

[33] Fernando, R. P. and I. M. W. Smith, Vibrational relaxation of NO by atomic oxygen, *Chem. Phys. Lett*, *66*, 218-222, 1979.

[34] Feynman, J., and A. J. Hundhausen, Coronal mass ejections and major solar flares: The great active center of March 1989, *J. Geophys. Res.*, *99*, 8451-8464, 1994.

[35] Fomichev, V. I., J.-P. Blanchet, and D. S. Turner, Matrix parameterization of the 15μm CO_2 band cooling in the middle and upper atmosphere for variable CO_2 concentration, *J. Geophys. Res.*, *103*, 11,505-11,528, 1998.

[36] Fomichev, V. I., W. E. Ward, S. R. Beagley, C. McLandress, J. C. McConnell, N. A. McFarlane and T. G. Shepherd, Extended Canadian Middle Atmosphere Model: Zonal-mean climatology and physical parameterizations, *J. Geophys. Res.*, *107*, 4087, doi: 10.1029/2001JD000479, 2002.

[37] Forbes, J. M., R. Gonzalez, F. A. Marcos, D. Revelle, and H. Parish, Magnetic storm response of lower thermosphere density, *J. Geophys. Res.*, *101*, 2313-2320, 1996.

[38] Forbes, J. M., G. Lu, S. Bruinsma, S. Nerem, and X. Zhang, Thermosphere density variations due to the 15-24 April 2002 solar events from CHAMP/STAR accelerometer measurements, *J. Geophys. Res.*, *110*, A12S27, doi: 10.1029/2004JA010856, 2005.

[39] Forester, C. K., Higher order monotonic convective difference schemes, *J. Comput. Phys.*, *23*, 1-22, 1977.

[40] Frank, L. A., and J. D. Craven, Imaging results from Dynamics Explorer 1, *Rev. Geophys.*, *26*, 249-283, 1988.

[41] Frank, L. A., J. D. Craven, K. L. Ackerson, M. R. English, R. H. Eather, and R. L. Carovillano, Global auroral imaging instrumentation for the Dynamics Explorer mission, *Space Sci. Instrum.*, *5*, 369-393, 1981.

[42] Fuller-Rowell, T. J., A two-dimensional, high-resolution, nested-grid model of the thermosphere. I. Neutral response to an electric field "spike", *J. Geophys. Res.*, *89*, 2971-2990, 1984.

[43] Fuller-Rowell, T. J., A two-dimensional, high-resolution, nested-grid model of the thermosphere. II. Response of the thermosphere to narrow and broad electrodynamic features, *J. Geophys. Res.*, *90*, 6567-6586, 1985.

[44] Gladstone, G. R., Simulations of DE-1 UV airglow images, *J. Geophys. Res.*, *99*, 11,441-11,448, 1994.

[45] Goncharenko, L. P., J. E. Salah, J. C. Foster, and C. Huang, Variations in lower thermosphere dynamics at midlatitudes during intense geomagnetic storms, *J. Geophys. Res.*, *109*, A04304, doi: 10.1029/2003JA010244, 2004.

[46] Good, A., D. A. Durden, and P. Kebarle, Mechanism and rate constants of ion-molecule reactions leading to the formation of $H^+(H_2O)_n$ in moist oxygen and air, *J. Chem. Phys.*, *52*, 222-229, 1970.

[47] Gordiets, B. F., M. N. Markov, and L. A. Shelepin, IR radiation of the upper atmosphere, *Planet. Space Sci.*, *26*, 933-948, 1978.

[48] Gordiets, B. F., Yu. N. Kulikov, M. N. Markov, and M. Ya. Marov, Numerical modeling of the thermospheric heat budget, *J. Geophys. Res.*, *87*, 4504-4514, 1982.

[49] Gottlieb, D., and E. Turkel, Dissipative two-four methods for time dependent problems, *Math. Comput. 30*, 703-723, 1976.

[50] Grossmann, K. U., and D. Offermann, Atomic oxygen emission at 63 μm as a cooling mechanism in the thermosphere and ionosphere, *Nature*, *276*, 594-595, 1978.

[51] Groves, G. V., Correlation of upper atmosphere density with geomagnetic activity, in *Space Research II, Proceedings of the Second International Space Science Symposium*, H. C. van de Hulst, C. de Jager, and A. F. Moore, eds., pp. 751-753, North-Holland Pub. Co., Amsterdam, 1961.

[52] Hays, P. B., R. A. Jones, and M. H. Rees, Auroral heating and the composition of the neutral atmosphere, *Planet. Space Sci.*, *21*, 559-573, 1973.

[53] Haltiner, G. J., and R. T. Williams, *Numerical Prediction and Dynamic Meteorology*, 2nd edition, John Wiley & Sons, New York, 1980.

[54] Hayden, J. L., A. O. Nier, J. B. French, N. M. Reid, and R. J. Duckett, The characteristics of an open source mass spectrometer under conditions simulating upper atmosphere flight, *Int. J. Mass Spectrom. Ion Phys.*, *15*, 37-47, 1974.

[55] Hedin, A. E., Extension of the MSIS thermospheric model into the middle and lower atmosphere, *J. Geophys. Res.*, *96*, 1159-1172, 1991.

[56] Hedin, A. E.; C. P. Avery, and C. D. Tschetter, An analysis of spin modulation effects on data obtained with a rocket-borne mass spectrometer, *J. Geophys. Res.*, *69*, 4637-4648, 1964.

[57] Hedin, A. E., B. B. Hinton, and G. A. Schmitt, Role of gas-surface interactions in the reduction of OGO-6 neutral particle mass spectrometer data, *J. Geophys. Res.*, *78*, 4651-4668, 1973.

[58] Heppner, J. P., Electric field variations during substorms: OGO-6 measurements, *Planet. Space Sci.*, *20*, 1475-1498, 1972.

[59] Huffman, R. E., *Atmospheric ultraviolet remote sensing*, Academic Press, Orlando, FL, 1992.

[60] Iijima, T., and T. A. Potemra, Large-scale characteristics of field-aligned currents associated with substorms, *J. Geophys. Res.*, *83*, 599-615, 1978.

[61] Immel, T. J., J. D. Craven, and L. A. Frank, Influence of IMF B_y on large-scale decreases of O column density at middle latitudes, *J. Atmos. Terr. Phys.*, *59*, 725-737, 1997.

[62] Immel, T. J.; J. D Craven, and A. C. Nicholas, An empirical model of the OI FUV dayglow from DE-1 images, *J. Atmos. Solar-Terr. Phys.*, *62*, 47-64, 2000.

[63] Jacchia, L. G., Corpuscular radiation and the acceleration of artificial satellites, *Nature*, *183*, 1662-1663, 1959.

[64] Jacchia, L. G., Satellite drag during the events of November 1960, in *Space Research II, Proceedings of the Second International Space Science Symposium*, H. C. van de Hulst, C. de Jager, and A. F. Moore, eds., pp. 747-750, North-Holland Pub. Co., Amsterdam, 1961.

[65] Jacchia, L. G, Revised static models of the thermosphere and exosphere with empirical temperature profiles, *Smithson. Astrophys. Obs. Spec. Rep. 332*, 104 pp., 1971.

[66] Jacchia, L. G., and J. Slowey, An analysis of the atmospheric drag of the Explorer IX satellite from precisely reduced photographic observations, *Smithson. Astrophys. Obs. Spec. Rep. 125*, 57 pp., 1963.

[67] Jacchia, L. G., and J. Slowey, Atmospheric heating in the auroral zones: A preliminary analysis of the atmospheric drag on the Injun 3 satellite, *J. Geophys. Res.*, *69*, 905-910, 1964.

[68] Jacchia, L. G., J. Slowey, and F. Verniani, Geomagnetic perturbations and upper atmosphere heating, *J. Geophys. Res.*, *72*, 1423-1434, 1967.

[69] Joshi, A., Superactive Region AR 5395 of solar cycle 22, *Solar Phys.*, *147*, 269-285, 1993.

[70] Kayser, D. C., W. T. Chater, C. K. Howey, and J. B. Pranke, The Upper Atmosphere Composition Spectrometer, *J. Spacecraft Rockets*, *23*, 336-341, 1986.

[71] King-Hele, D. G., Methods of determining air density from satellite orbits, *Ann. Geophys.*, *22*, 40-52, 1966.

[72] Kockarts, G., Nitric oxide cooling in the terrestrial thermosphere, *Geophys. Res. Lett.*, *7*, 137-140, 1980.

[73] Kockarts, G., and W. Peetermans, Atomic oxygen infrared emission in the Earth's upper atmosphere, *Planet. Space Sci.*, *18*, 271-285, 1970.

[74] Krankowsky, D.; P. Laemmerzahl, F. Bonner, and H. Wieder, The AEROS neutral and ion mass spectrometer, *J. Geophys.*, *40*, 601-611, 1974.

[75] Kutepov, A. A., and V. I. Fomichev, Application of the second-order escape probability approximation to the solution of the NLTE vibration-rotational band radiative transfer problem, *J. Atmos. Terr. Phys.*, *55*, 1-6, 1993.

[76] Larsen, M. F., Winds and shears in the mesosphere and lower thermosphere: Results from four decades of chemical release wind measurements, *J. Geophys. Res.*, *107*, 1215, doi: 10.1029/2001JA000218, 2002.

[77] Liseikin, V. D., *Grid Generation Methods*, Springer-Verlag, Berlin, 1999.

[78] Lopez-Puertas, M., M. A. Lopez-Valverde, C. P. Rinsland, and M. R. Gunson, Analysis of the upper atmosphere $CO_2(\nu_2)$ vibrational temperatures retrieved from ATMOS/Spacelab 3 observations, *J. Geophys. Res.*, *97*, 20,469-20,478, 1992.

[79] Lübken, F.-J., and U. von Zahn, Thermal structure of the mesopause region at polar latitudes, *J. Geophys. Res.*, *96*, 20,841-20,857, 1991.

[80] Lübken, F.-J., M. J. Jarvis, and G. O. L. Jones, First in situ temperature measurements at the Antarctic summer mesopause, *Geophys. Res. Lett.*, *26*, 3581-3584, 1999.

[81] MacCormack, R. W., The effect of viscosity in hypervelocity impact cratering, *AIAA Paper No. 69-354*, 1969.

[82] Malkmus, W., Random Lorentz band model with exponential-tailed S^{-1} line-intensity distribution function, *J. Opt. Soc. Amer.*, *57*, 323-329, 1967.

[83] Mayr, H. G., and H. Volland, Magnetic storm effects in the neutral composition, *Planet. Space Sci.*, *20*, 379-393, 1972.

[84] Mayr, H. G., and A. E. Hedin, Significance of large-scale circulation in magnetic storm characteristics with application to AE-C neutral composition data, *J. Geophys. Res.*, *82*, 1227-1234, 1977.

[85] Meier, R. R., Depressions in the far-ultraviolet airglow over the poles, *J. Geophys. Res.*, *75*, 6218-6232, 1970.

[86] Meier, R. R., and D. K. Prinz, Observations of the OI 1304 Å airglow from OGO-4, *J. Geophys. Res.*, *76*, 4608-4620, 1971.

[87] Meier, R. R., R. J. Cox, D. J. Strickland, J. D. Craven, and L. A. Frank, Interpretation of Dynamics Explorer far UV images of the quiet time thermosphere, *J. Geophys. Res.*, *100*, 5777-5794, 1995.

[88] Mikkelsen, I. S., T. S. Jorgensen, M. C. Kelley, M. F. Larsen, and E. Pereira, Neutral winds and electric fields in the dusk auroral oval II : Theory and model, *J. Geophys. Res.*, *86*, 1525-1536, 1981.

[89] Mlynczak, M., and S. Solomon, A detailed evaluation of the heating efficiency in the middle atmosphere, *J. Geophys. Res.*, *98*, 10,517-10,541, 1993.

[90] Nicholas, A. C., J. D. Craven, and L. A. Frank, A survey of large-scale variations in thermospheric oxygen column density with magnetic activity as inferred from observations of the FUV dayglow, *J. Geophys. Res.*, *102*, 4493-4510, 1997.

[91] Nier, A. O., Mass spectrometry in planetary research, *Int. J. Mass Spectrom. Ion Processes*, *66*, 55-73, 1985.

[92] Nier, A. O., W. E. Potter, D. R. Hickman, and K. Mauersberger, The open-source neutral-mass spectrometer on Atmosphere Explorer-C -D, and -E, *Radio Sci.*, *8*, 271-276, 1973.

[93] Nier, A. O., W. E. Potter, D. C. Kayser, and R. G. Finstad, The measurement of chemically reactive atmospheric constituents by mass spectrometers carried on high-speed spacecraft, *Geophys. Res. Lett.*, *1*, 197-200, 1974.

[94] Nier, A. O., W. E. Potter, and D. C. Kayser, Atomic and molecular oxygen densities in the lower thermosphere, *J. Geophys. Res.*, *81*, 17-24, 1976.

[95] Oran, E. S., and J. P. Boris, *Numerical Simulation of Reactive Flow*, 2nd edition, Cambridge University Press, New York, 2001.

[96] Pelz, D. T., C. A. Reber, A. E. Heden, and G. R. Carignan, A neutral atmosphere composition experiment for the Atmosphere Explorer-C -D, and -E, *Radio Sci.*, *8*, 277-285, 1973.

[97] Philbrick, C. R., Satellite measurements of neutral atmospheric composition in the altitude range 150-450 km, *Space Res.*, *14*, 151-156, 1974.

[98] Philbrick, C. R., Recent satellite measurements of upper atmospheric composition, *Space Res.*, *16*, 289-295, 1976.

[99] Philbrick, C. R., J. P. McIsaac, and G. A. Faucher, Variations in atmospheric composition and density during a geomagnetic storm, *Space Res.*, *17*, 349-353, 1977.

[100] Phillips, N. A., The equations of motion for a shallow rotating atmosphere and the "traditional approximation", *J. Atmos. Sci.*, *23*, 626-627, 1966.

[101] Pollock, D. S., G. B. I. Scott, and L. F. Phillips, Rate constant for the quenching of $CO_2(01^10)$ by atomic oxygen, *Geophys. Res. Lett.*, *20*, 727-729, 1993.

[102] Potter, W. E., D. G. Kayser, and A. O. Nier, Thermospheric variations as an indicator of magnetic storm heating and circulation, *Space Res.*, *19*, 259-262, 1979.

[103] Prölss, G. W., Magnetic storm associated perturbations of the upper atmosphere - Recent results obtained by satellite-borne gas analyzers, *Rev. Geophys. Space Phys.*, *18*, 183-202, 1980.

[104] Rees, M. H., *Physics and Chemistry of the Upper Atmosphere*, Cambridge University Press, Cambridge, U.K., 1989.

[105] Rees, M. H., and R. G. Roble, Observations and theory of the formation of stable auroral red arcs, *Rev. Geophys. Space Phys. 13*, 201-242, 1975.

[106] Richmond, A. D. and S. Matsushita, Thermospheric response to a magnetic substorm, *J. Geophys. Res.*, *80*, 2839-2850, 1975.

[107] Rishbeth, H., R. J. Moffet, and G. J. Bailey, Vertical diffusion of O and N_2 in the thermosphere, *Planet. Space Sci.*, *22*, 189-192, 1974.

[108] Roache, P. J., *Computational Fluid Dynamics*, Hermosa, Albuquerque, New Mexico, 1972.

[109] Roble, R. G., Energetics of the mesosphere and thermosphere, in *The Upper Mesosphere and Lower Thermosphere: A Review of Experiment and Theory*, Geophysical Monograph Series, Vol. 87, R. M. Johnson and T. L. Killeen, eds., pp. 1-21, AGU, Washington, D. C., 1995.

[110] Roble, R. G., On the feasibility of developing a global atmospheric model extending from the surface to the exosphere, in *Atmospheric Science Across the Stratopause*, Geophysical Monograph Series, Vol. 123, D. E. Siskind, S. D. Eckermann, and M. E. Summmers, eds., pp. 53-67, AGU, Washington, D. C., 2000.

[111] Roble, R. G., E. C. Ridley, and R. E. Dickinson, On the global mean structure of the thermosphere, *J. Geophys. Res.*, *92*, 8745-8758, 1987.

[112] Rodgers, C. D., F. W. Taylor, A. H. Muggeridge, M. Lopez-Puertas, and M. A. Lopez-Valverde, Local thermodynamic equilibrium of carbon dioxide in the upper atmosphere, *Geophys. Res. Lett.*, *19*, 589-592, 1992.

[113] Roemer, M., Atmospheric densities and temperatures from precisely reduced observations of the Explorer IX satellite, *Smithson. Astrophys. Obs. Spec. Rep. 199*, 81 pp., 1967.

[114] Roemer, M., Geomagnetic activity effect on atmospheric density in the 250 to 800 km altitude region, in *Space Research XI*, H. C. van de Hulst, ed., pp. 965-974, North-Holland Pub. Co., Amsterdam, 1971.

[115] St.-Maurice, J.-P., and R. W. Schunk, Ion-neutral momentum coupling near discrete high-latitude ionospheric features, *J. Geophys. Res.*, *86*, 11299-11321, 1981.

[116] Schield, M. A., J. W. Freeman, and A. J. Dessler, A source for field-aligned currents at auroral latitudes, *J. Geophys. Res.*, *74*, 247-256, 1969.

[117] Schunk, R. W., and A. F. Nagy, *Ionospheres: Physics, plasma physics, and chemistry*, Cambridge University Press, New York, 2000.

[118] Sharma, R. D., and R. G. Roble, Impact of the new rate coefficients for the O atom vibrational deactivation and photodissociation of NO on the temperature and density structure of the terrestial atmosphere, *J. Geophys. Res.*, *106*, 21,343-21,350, 2001.

[119] Sharma, R. D., and P. P. Wintersteiner, Role of carbon dioxide in cooling planetary atmospheres, *Geophys. Res. Lett.*, *17*, 2201-2204, 1990.

[120] Shved, G. M., L. E. Khvorostovskaya, I. Yu. Potekhin, A. I. Demynikov, A. A. Kutepov, and V. I. Fomichev, Measurement of the quenching rate for collisions $CO_2(01^10) - O$: The importance of the rate constant magnitude for the thermal regime and radiation of the lower thermosphere, *Atmos. and Oceanic Phys.*, *27*, 431-437, 1991.

[121] Sinha, A. K., and S. Chandra, Seasonal and magnetic storm related changes in the thermosphere induced by eddy mixing, *J. Atmos. Terr. Phys.*, *36*, 2055-2066, 1974.

[122] Siskind, D. E., C. A. Barth, and J. M. Russell III, A climatology of nitric oxide in the mesosphere and thermosphere, *Adv. Space Res.*, *21*, 1353-1362, 1998.

[123] Stewart, R. W., Radiative terms in the thermal conduction equation for planetary atmospheres, *J. Atmos. Sci.*, *25*, 744-749, 1968.

[124] Stolarski, R. S., Energetics of the mid-latitude thermosphere, *J. Atmos. Terr. Phys.*, *38*, 863-868, 1976.

[125] Stolarski, R. S., P. B. Hays, and R. G. Roble, Atmospheric heating by solar EUV radiation, *J. Geophys. Res.*, *80*, 2266-2276, 1975.

[126] Strickland, D. J., and G. E. Thomas, Global atomic oxygen density derived from OGO-6 1304 Å airglow measurements, *Planet. Space Sci.*, *24*, 313-326, 1976.

[127] Strickland, D. J., R. J. Cox, R. R. Meier, and D. P. Drob, Global O/N_2 derived

from DE-1 FUV imaging dayglow data: Technique and examples from two storm periods, *J. Geophys. Res.*, *104*, 4251-4266, 1999.

[128] Strobel, D. F., Parameterization of the atmospheric heating rate from 15 to 120 km due to O_2 and O_3 absorption of solar radiation, *J. Geophys. Res.*, *83*, 6225-6230, 1978.

[129] Sun, Z.-P., R. P. Turco, R. L. Walterscheid, S. V. Venkateswaran and P. W. Jones, Thermospheric response to morningside diffuse aurora: high-resolution three-dimensional simulations, *J. Geophys. Res.*, *100*, 23,779-23,794, 1995.

[130] Taeusch, D. R., G. R. Carignan and C. A. Reber, Neutral composition variations above 400 kilometers during a magnetic storm, *J. Geophys. Res.*, *76*, 8318-8325, 1971.

[131] Tannehill, J. C., D. A. Anderson, and R. H. Pletcher, *Computational Fluid Mechanics and Heat Transfer*, 2nd edition, Taylor and Francis, Washington, D. C., 1997.

[132] Theon, J. S., W. Nordberg, L. B. Katchen, and J. J. Horvath, Some Observations on the Thermal Behavior of the Mesosphere, *J. Atmos. Sci.*, *24*, 428-438, 1967.

[133] Thomas, R. J., Atomic hydrogen and atomic oxygen density in the mesopause region: Global and seasonal variations deduced from Solar Mesosphere Explorer near-infrared observations, *J. Geophys. Res.*, *95*, 16,457-16,476, 1990.

[134] Thompson, J. F., and C. W. Mastin, Order of difference expressions in curvilinear coordinate systems, *ASME J. of Fluids Engineering*, *107*, 241-250, 1985.

[135] Thompson, J. F., Z. U. A. Warsi, and C. W. Mastin, *Numerical Grid Generation - Foundations and Applications*, Elsevier, New York, 1985.

[136] Toffoletto, F. R., and T. W. Hill, Field line mapping and Birkeland Currents, in *Magnetospheric Current Systems*, Geophysical Monograph Series, Vol. 118, Shin-

ichi Ohtani, Ryoichi Fujii, Michael Hesse and Robert L. Lysak, eds., pp. 71-79, AGU, Washington, D. C., 2000.

[137] Toro, E. F., *Riemann Solvers and Numerical Methods for Fluid Dynamics: A Practical Introduction*, Springer, Berlin, 1999.

[138] Trinks, H., and U. von Zahn, The ESRO 4 gas analyzer, *Rev. Sci. Instru.*, *46*, 213-217, 1975.

[139] Trinks, H., U. von Zahn, C. A. Barth, and K. K. Kelly, A joint nitric oxide measurement by rocket-borne ultraviolet photometer and mass spectrometer in the lower thermosphere, *J. Geophys. Res.*, *83*, 203-206, 1978.

[140] Turunen, E., H. Matveinen, J. Tolvanen, and H. Ranta, D-region ion chemistry model, in *Solar-Terrestrial Energy Program : Handbook of Ionospheric Models*, R. W. Schunk, ed., pp. 1-25, SCOSTEP Secretariat, Boulder, Colorado, 1996.

[141] Vinokur, M., Conservation equations of gas dynamics in curvilinear coordinate systems, *J. Comput. Phys.*, *14*, 105-125, 1974.

[142] von Neumann, J., and R. D. Richtmyer, A model for the numerical calculation of hydrodynamic shocks, *J. Appl. Phys.*, *21*, 232-237, 1950.

[143] von Zahn, U., Composition studies in the thermosphere by means of mass spectrometers, in *Structure and Dynamics of the Upper Atmosphere*, Developments in Atmospheric Science, Vol. 1, F. Verniani, ed., pp. 389-434, Elsevier Scientific, New York, 1974.

[144] Walterscheid, R. L., and D. G. Brinkman, Spin-up circulation of high-latitude ion drag-driven gyres, *J. Geophys. Res.*, *108*, 1304, doi: 10.1029/2002JA009642, 2003.

[145] Walterscheid, R. L., L. R. Lyons, and K. E. Taylor, The perturbed neutral circulation in the vicinity of a symmetric stable auroral arc, *J. Geophys. Res.*, *90*, 12,235-12,248, 1985.

[146] Wolf, R. A., The quasi-static (slow-flow) region of the magnetosphere, in *Solar-terrestrial physics: Principles and theoretical foundations*, R. L. Carovillano and J. M. Forbes, eds., pp. 303-368, D. Reidel, Dordrecht, 1983.

[147] Wolf, R. A., Magnetospheric configuration, in *Introduction to Space Physics*, M. G. Kivelson and C. T. Russell, eds., pp. 288-329, Cambridge University Press, New York, 1995.

[148] World Meteorological Organization, Atmospheric Ozone 1985 : Assessment of our Understanding of the Processes Controlling its Present Distribution and Change, Global Ozone Research and Monitoring Project - Report No. 16, Geneva, Switzerland, 1985.

[149] Yeh, K. C., K. H. Lin, and R. O. Conkright, The global behavior of the March 1989 ionospheric storm, *Can. J. Phys.*, *70*, 532-543, 1992.

[150] Zhu, X., An accurate and efficient radiation algorithm for middle atmosphere models, *J. Atmos. Sci.*, *51*, 3593-3614, 1994.

[151] Zuzic, M., L. Scherliess, and G. W. Prölss, Latitudinal structure of thermospheric composition perturbations, *J. Atmos. Solar-Terr. Phys.*, *59*, 711-724, 1997.

VITA

NAME:	Albert T. Russell
PLACE OF BIRTH:	Mississauga, Ontario
YEAR OF BIRTH:	1967
POST-SECONDARY EDUCATION AND DEGREES:	The University of Western Ontario London, Ontario 1998 - 2011 (Ph. D. Candidate)
	The University of Western Ontario London, Ontario 1998 M.Sc. (Physics)
	University of Waterloo Waterloo, Ontario 1992 B.Sc. (Honours Physics)
HONOURS AND AWARDS:	Special University Scholarship The University of Western Ontario 1995-2002
	Award of Excellence in Teaching Department of Physics and Astronomy The University of Western Ontario 1996
PUBLICATIONS:	Russell, A. T., J.-P. St.-Maurice, R. J. Sica, and J.-M. Noël (2007). "Composition changes during disturbed conditions: Are mass spectrometers overestimating the concentrations of atomic oxygen?", *Geophys. Res. Lett.*, **34**, L21106, doi:10.1029/2007GL030607.

Sica, R. J., T. Thayaparan, P. S. Argall, A. T. Russell, and W. K. Hocking (2002). "Modulation of upper mesospheric temperature inversions due to tidal-gravity wave interactions", *J. Atmos. Sol. Terr. Phys.,* **64**, pp. 915 - 922.

Russell, A. T. and R. J. Sica (2001). "Observations of high-frequency temporal gravity wave spectra in the middle upper stratosphere", *J. Geohys. Res.,* **106**, D11, pp. 11,849 - 11,858.

Sica, R. J., P. S. Argall, A. T. Russell, C. R. Bryant, and M. M. Mwangi (2000). "Dynamics and composition measurements in the lower and middle atmosphere with the Purple Crow Lidar" in *Recent Research Developments in Geophysical Research*, S.G. Pandalia (ed.).

Sica, R. J. and A. T. Russell (1999). "How many waves are in the gravity wave spectrum?", *Geohys. Res. Lett.,* **26**, pp. 3617 - 3620.

Sica, R. J. and A. T. Russell (1999). "Measurements of the effects of gravity waves in the middle atmosphere using parametric models of density fluctuations I. Vertical wavenumber and temporal spectra", *J. Atmos. Sci.,* **56**, pp. 1308 - 1329.