

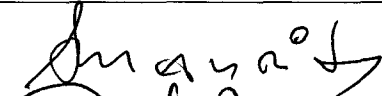
APPLICATIONS OF A TIME-DEPENDENT POLAR IONOSPHERE MODEL FOR
RADIO MODIFICATION EXPERIMENTS

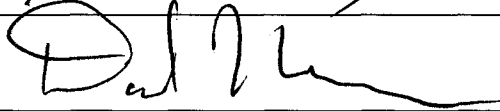
By

Christopher Thomas Fallen

RECOMMENDED:









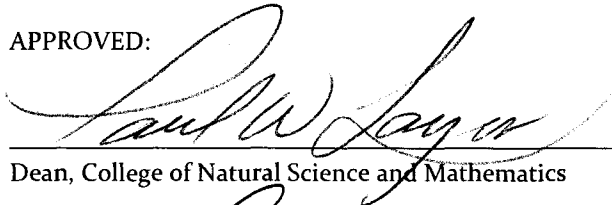


Advisory Committee Chair

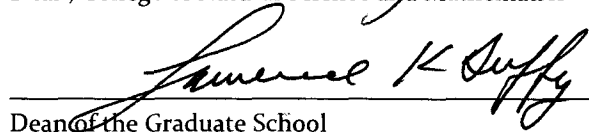


Chair, Department of Physics

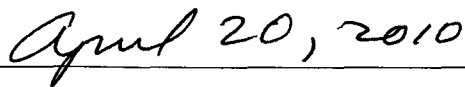
APPROVED:



Dean, College of Natural Science and Mathematics



Dean of the Graduate School



Date

**APPLICATIONS OF A TIME-DEPENDENT POLAR IONOSPHERE MODEL FOR
RADIO MODIFICATION EXPERIMENTS**

A
THESIS

Presented to the Faculty
of the University of Alaska Fairbanks
in Partial Fulfillment of the Requirements
for the Degree of

DOCTOR OF PHILOSOPHY

By
Christopher Thomas Fallen, B.A., B.S., M.A.

Fairbanks, Alaska

May 2010

UMI Number: 3421519

All rights reserved

INFORMATION TO ALL USERS

The quality of this reproduction is dependent upon the quality of the copy submitted.

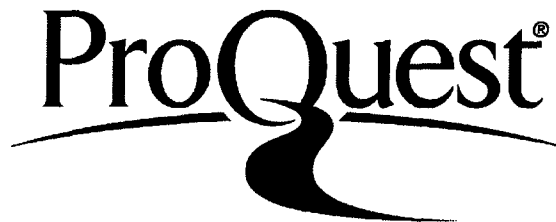
In the unlikely event that the author did not send a complete manuscript and there are missing pages, these will be noted. Also, if material had to be removed, a note will indicate the deletion.



UMI 3421519

Copyright 2010 by ProQuest LLC.

All rights reserved. This edition of the work is protected against unauthorized copying under Title 17, United States Code.



ProQuest LLC
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106-1346

Abstract

A time-dependent self-consistent ionosphere model (SCIM) has been developed to study the response of the polar ionosphere to radio modification experiments, similar to those conducted at the High-Frequency Active Auroral Research Program (HAARP) facility in Gakona, Alaska. SCIM solves the ion continuity and momentum equations, coupled with average electron and ion gas energy equations; it is validated by reproducing the diurnal variation of the daytime ionosphere critical frequency, as measured with an ionosonde. Powerful high-frequency (HF) electromagnetic waves can drive naturally occurring electrostatic plasma waves, enhancing the ionospheric reflectivity to ultra-high frequency (UHF) radar near the HF-interaction region as well as heating the electron gas. Measurements made during active experiments are compared with model calculations to clarify fundamental altitude-dependent physical processes governing the vertical composition and temperature of the polar ionosphere. The modular UHF ionosphere radar (MUIR), co-located with HAARP, measured HF-enhanced ion-line (HFIL) reflection height and observed that it ascended above its original altitude after the ionosphere had been HF-heated for several minutes. The HFIL ascent is found to follow from HF-induced depletion of plasma surrounding the F-region peak density layer, due to temperature-enhanced transport of atomic oxygen ions along the geomagnetic field line. The lower F-region and topside ionosphere also respond to HF heating. Model results show that electron temperature increases will lead to suppression of molecular ion recombination rates in the lower F region and enhancements of ambipolar diffusion in the topside ionosphere, resulting in a net enhancement of slant total electron content (TEC); these results have been confirmed by experiment. Additional evidence for the model-predicted topside ionosphere density enhancements via ambipolar diffusion is provided by in-situ measurements of ion density and vertical velocity over HAARP made by a Defense Meteorological Satellite Program (DMSP) satellite.

Table of Contents

	Page
Signature Page	i
Title Page	ii
Abstract	iii
Table of Contents	iv
List of Figures	vi
List of Tables	xiv
Acknowledgements	xv
Chapter 1. Introduction	1
1.1 Historical review.....	2
1.2 This work.....	6
Chapter 2. Ionosphere model equations	9
2.1 Boltzmann equation.....	9
2.1.1 Continuity equation	11
2.1.2 Momentum equation.....	20
2.1.3 Energy equation	31
2.1.3.1 Electron gas collisional energy transfer rates	37
2.1.3.2 Ion gas collisional energy transfer rates	41
Chapter 3. Solution of model equations	44
3.1 Numerical solution.....	44
3.1.1 Solver equations.....	49
3.1.2 Boundary conditions	54
3.1.3 Simulation structure	57
3.2 Model validation and sensitivity analysis.....	60
3.2.1 Vertical structure of the ionosphere	63
3.2.2 Diurnal variation of daytime critical frequency.....	69
3.2.3 Diurnal variation of daytime TEC.....	75
3.2.4 Model sensitivity to semidiurnal thermal tides.....	79
Chapter 4. Ascent of the ion line	85
4.1 Overview	86
4.2 Method	87
4.2.1 Geophysical conditions	89

4.2.2	HF modulation for ionosphere modification and diagnostics.....	93
4.2.2.1	HF <i>heat-scan</i> modulation sequence	95
4.2.2.2	HF <i>heat-pulse</i> modulation sequence.....	97
4.2.3	Analysis of MUIR coded long-pulse data	97
4.2.4	Ionosphere HF modification simulation	112
4.3	Results	116
4.3.1	HFIL returns observed with UHF radar.....	116
4.3.2	Time-dependent simulation of HFIL reflection height	122
4.3.3	Concurrent radar and optical measurements of HF-induced structures	129
4.4	Discussion	131
Chapter 5.	Total electron content enhancements	138
5.1	Experiment.....	139
5.2	Model.....	142
5.3	Discussion on HF-induced topside ionosphere perturbations.....	148
5.3.1	In-situ measurements of topside density enhancements.....	149
5.3.2	HF frequency dependence of ionosphere modification.....	153
Chapter 6.	Summary and conclusions	160
6.1	Conclusions.....	162
6.2	Future work.....	164
Appendix.	Distributed radar data processing on <i>Midnight</i>	166
A.1	Matlab overview	166
A.2	Review of Parallel Matlab implementations	168
A.2.1	Multithreaded Matlab.....	168
A.2.2	MIT-LL MatlabMPI and pMatlab Parallel Toolbox	170
A.2.3	Shell-script distributed Matlab.....	171
A.2.4	Parallel Computing Toolbox and Matlab Distributed Computing Server.....	172
A.3	Multi-node parallel Matlab on ARSC's <i>Midnight</i> via the PCT and MDCS.....	175
A.3.1	Initialization and configuration on <i>Midnight</i>	175
A.3.2	Case study: Modular UHF ionosphere radar data processing	178
A.3.2.1	Objective.....	178
A.3.2.2	DCT and MDCE configuration.....	178
A.3.2.3	Numerical experiment	179
A.3.2.4	Results and discussion.....	181
References	185

List of Figures

		Page
Figure 2.1	Altitude profiles of neutral species and electron number density above Gakona, Alaska at 01:00 UT (hh:mm) on 24 October 2008.....	13
Figure 2.2	Solar photon flux at approximately 10 Å intervals calculated with the Hinteregger model for two values of the F107 solar radio flux index.....	15
Figure 2.3	Altitude profiles of major and minor ion photoproduction rates modeled above Gakona, Alaska at 0100 UT on 24 October 2008.	17
Figure 2.4	Model altitude profiles of the loss frequency of major ion species due to chemical reactions or electron recombination over Gakona, Alaska on 24 October 2008 at 0100 UT. Model altitude profiles of electron n_e and neutral species n_n number density are plotted for reference.	19
Figure 2.5	Altitude profiles of the total ion-neutral collision frequency ν_{in} calculated for selected ion species over Gakona, Alaska. The associated ion cyclotron frequencies $f_{ic} = \omega_{ic}/2\pi$ and the electron-neutral collision frequency ν_{en} are also plotted for reference.....	26
Figure 2.6	Altitude profiles of the diffusion and ambipolar diffusion coefficients for selected neutral and ionized gas species calculated over Gakona, Alaska. An altitude-independent eddy diffusion coefficient from Kellogg [1964] is plotted for comparison.	29
Figure 2.7	Altitude profiles of selected species scale heights (solid line) and diffusion times (dotted line) calculated over Gakona, Alaska at 01:00 UT (hh:mm) on 24 October 2008.	30
Figure 2.8	Altitude profiles of the thermal conductivity λ (solid lines), downward thermal energy flux (dotted), and upward field-aligned thermal energy flux ϕ (dashed lines) for the ion (blue) and electron (red) gas modeled over Gakona, Alaska at 01:00 UT (hh:mm) on 24 October 2008.....	35
Figure 2.9	Altitude profiles of power transfer rates to and from the ion (blue) and electron (red) gases over Gakona, Alaska at 01:00 UT (hh:mm) on 24 October 2008. Total power transfer into (thick) and from (thin) each gas includes the sums of energy transfer rates via elastic and inelastic collisions between the neutral, ion, and electron gases. An imposed 1 mV/m perpendicular electric field results in partial power density transfer to the ion gas (dotted) via Pedersen-current Joule dissipation.....	37

Figure 3.1	Flow-chart schematic of the 1-D Self-Consistent Ionosphere Model (SCIM).	59
Figure 3.2	Altitude profiles of ion species number density modeled over Gakona, Alaska. The electron number density is calculated by the summing of the ion species number densities. Respective ion species number density profiles calculated with the IRI 2007 empirical model are printed in lighter colors relative to their respective SCIM-calculated profiles.	64
Figure 3.3	Altitude profiles of the electron temperature (dark red) and ion temperature (dark blue) modeled over Gakona, Alaska on 24 October 2008. The neutral gas temperature profile (black) is calculated with the MSIS empirical model. Corresponding temperature profiles (light colors) calculated with the IRI2007 model are provided for comparison.	66
Figure 3.4	Scatter plot of ISR measurements of plasma density and temperature (black points) above Poker Flat, Alaska, on 23 October 2008. Altitude profiles of corresponding quantities at 21:00, 22:00, and 23:00 UT calculated with the IRI model (red lines) and the SCIM model (dark blue lines) are overlaid for comparison. The SCIM model output is from the Gakona diurnal equilibrium run defined by the parameters listed in Table 3.1 except that the electron gas heat flux at the top simulation boundary is set to 0.6 $\mu\text{W}/\text{m}^2$ downward. Altitude profiles at 22:00 UT from a SCIM model run using a 0.0 $\mu\text{W}/\text{m}^2$ electron gas thermal flux at the top simulation boundary (light blue lines) are also shown.	68
Figure 3.5	Altitude profile of plasma frequency (black) measured with the HAARP ionosonde over Gakona, Alaska at 01:00 UT (hh:mm) on 24 October, 2008. The corresponding frequency profiles calculated by the SCIM model (dark blue) and IRI 2007 empirical model (cyan) are overlaid on the ionogram for comparison. The horizontal axis is measured in MHz and the vertical axis is measured in vertical km.	71
Figure 3.6	Time-dependence of the ionosphere critical frequency over Gakona, Alaska on 24 October 2008 calculated with the SCIM model (dotted) and with the IRI 2007 empirical model (solid). The solar zenith angle (dashed) is plotted against time for reference. Each dot in the SCIM model output represents 15 minutes of simulation time, the red dots correspond to the first 24 hours of simulation time following an initial ionization-free atmosphere at 20:00 UT and the overlaid black dots correspond to the to the following 48 hours of simulation time.	73
Figure 3.7	Ionosphere critical frequency vs. universal time over Gakona, Alaska on 24 October 2008 as measured by the HAARP ionosonde (blue), as calculated by the self-consistent ionosphere model (red), and as calculated by the IRI 2007 empirical model (cyan).....	74

Figure 3.8	Time-dependence of the ionosphere total electron content (TEC) between 100 and 1000 km altitude over Gakona, Alaska on 24 October 2008 calculated with the SCIM model (dotted) and with the IRI 2007 empirical model (solid). The solar zenith angle (dashed) is plotted against time for reference. Each dot in the SCIM model output represents 15 minutes of simulation time, the red dots correspond to the first 24 hours of simulation time following an initial ionization-free atmosphere at 20:00 UT and the overlaid black dots correspond to the to the following 48 hours of simulation time	77
Figure 3.9	HAARP GPS measurements of the time-dependent total electron content (TEC) over Gakona, Alaska on 24 October 2008. The TEC below 1000 km as calculated by SCIM (red) and the IRI 2007 model (solid blue) are overlaid for comparison, the TEC below 2000 km as calculated by IRI (dashed blue) is also shown. Short line segments (various colors) represent the GPS data from individual satellite passes.	79
Figure 3.10	Height-integrated electron volume photoproduction rate (red) and height-integrated volume number density of molecular nitrogen and oxygen (blue) as parametric functions of solar zenith modeled over Gakona, Alaska on 24 October 2008. The electron column production rate due to impact with non-thermal photoelectrons (light red) and the column number density of atomic oxygen (light blue) are shown for comparison. The column number density of atomic oxygen has been scaled by a factor of five.	81
Figure 3.11	Parametric dependence on solar zenith angle of height-integrated electron volume heat transfer rate (dark red) due to collisions with energetic photoelectrons and of height-integrated electron volume thermal energy density. The column heat transfer rate of the ion gas (light red) due to joule heat dissipation from a perpendicular magnetospheric electric field of 1.0 mV/m and the column thermal energy density of the ion gas (light blue) are provided for comparison.	83
Figure 4.1	Daily measurements of the Geomagnetic Averaged Planetary Index (A_p) and the 10.7 cm Solar Radio Flux (F107) during October of 2008 and 2009.	91
Figure 4.2	Frequency, time and effective radiated power (ERP) variation of the <i>heat-scan</i> HF on-off modulation sequence initiated at 0020 UT on 24 October 2007 by the HAARP IRI. Each discrete point represents a 10 ms pulse, and the inter-pulse period is 1 s. The horizontal lines represent times of continuous HF transmission. The right-side axis is labeled with the plasma frequency number densities corresponding to the labeled HF frequencies on the left-side axis.....	96

Figure 4.3	Annotated range-time diagram of the peak Fourier power signal-to-noise ratio (SNR) received measured with MUIR. The UHF radar was aligned parallel to the geomagnetic field and the Terra SAR-X satellite entered the radar volume at about 02:49:56 UT (hh:mm:ss).	104
Figure 4.4	Received power signal-to-noise ratio (SNR) vs. time (range) of two MUIR 996 μ s pulses phase-modulated with a 4 μ s baud code. The two pulses were transmitted parallel to the geomagnetic field at approximately 02:49:55 (black) and 02:49:58 (red) UT (hh:mm:ss) on 22 March 2009. The TERRASAR-X satellite was traversing the radar pulse volume during the second UHF pulse, as indicated in Figure 4.3. SNR vs. range is plotted before (a) and after (b) applying the <i>basic range-decoding</i> algorithm. SNR vs. range calculated with the <i>normalized code-muted range-decoding</i> algorithm is plotted before (c) and after (d) muting the first half of the phase code used in the decoding algorithm, reducing the maximum signal range affected by ground clutter.	107
Figure 4.5	1100-point Fast Fourier Transform (FFT) power spectrum of samples taken from the two decoded MUIR pulses illustrated in Figure 4.4 (b). The pulse samples begin at (blue) 0 km range and (red, black) 538 km range; the sample rate is 250,000 kHz. Samples starting at 538 km range are 684 μ s in duration and the sample starting at 0 km range is 996 μ s in duration. A positive frequency offset indicates that the relative distance between the receiver and the reflecting target is decreasing.	111
Figure 4.6	Range-time diagram of HAARP HF-enhanced ion-line (HFIL) power measured with MUIR during a HAARP <i>heat-scan</i> cycle on 24 October 2008. Both HAARP and MUIR are directed toward magnetic zenith. HAARP transmitted three 4.2 MHz <i>scan</i> sequences (1, 3, and 4) consisting of ten 1 ms pulses with 1 s IPP at 25% power starting at 4.1 MHz and ending at 4.298 MHz in 22 KHz steps, three 4.2 MHz <i>scan</i> sequences (5, 7, and 8) at 12.5% power, and two 4.2 MHz <i>heat</i> sequences (2, 6) at full power. The 4.55 MHz <i>scan</i> sequence (9) transmits 4.5 to 4.6 MHz in 11 kHz steps and with the 4.55 MHz <i>heat</i> sequence (10) is part of the next <i>heat-scan</i> cycle. MUIR transmitted a 996 μ s phase-coded pulse with 4 μ s baud length. Range to the HFIL return was calculated using the normalized decoding algorithm after muting the first 109 bits of the code. The HFIL during <i>scan</i> sequences (3-5) following the first <i>heat</i> sequence (2) are several kilometers further from MUIR than the original HFIL <i>scan</i> sequence (1).	118

Figure 4.7	Range-time diagram of HAARP HF-enhanced ion-line (HFIL) power measured with MUIR during a HAARP <i>heat-pulse</i> cycle on 21 October 2009. Both HAARP and MUIR are directed toward magnetic zenith. HAARP transmitted a 6 min <i>pulse</i> sequence consisting of 0.1 s pulses with 1 s IPP at 4.06 MHz from 01:43:00 to 01:49:00 hours UT, a 5 min heat sequence at 4.06 MHz from 01:49:00 to 01:54:00 UT, and a 5.5 min 4.06 MHz pulse sequence from 01:54:00 to 01:59:30 UT. The HFIL return beginning at 02:00:00 UT is from an unrelated subsequent experiment, where HAARP transmitted O-mode at 2.75 MHz along the geomagnetic field. MUIR data processing parameters are the same as those used for Figure 4.6.	120
Figure 4.8	Inverted ionogram showing the O-mode 4.06 MHz reflection height over HAARP before the start of the 21 Oct 2009 experiment. The horizontal axis is expressed in MHz and the vertical axis is expressed in km.	121
Figure 4.9	Altitude profile of the electron plasma frequency in a diurnal-equilibrium model ionosphere on 24 October 2008 over Gakona, Alaska.	123
Figure 4.10	Electron (left) and ion (right) temperature profiles calculated at the times (hhmm UT) indicated by the labels. Simulated HF-electron heating began at 0000 UT and ended at 0005 UT.	125
Figure 4.11	Electron number density profiles calculated at the times (hhmm UT) indicated by the labels. Simulated HF-electron heating began at 0000 UT and ended at 0005 UT. The approximate electron number density with a plasma frequency equal to the chosen HF heating frequency is identified with a dashed line.	127
Figure 4.12	Field-aligned 557.7 nm airglow (white) and HF-enhanced ion-line intensity (copper) over the HAARP facility in Gakona, Alaska. A wide-field imager located about 160 km north of HAARP measured the airglow intensity. MUIR, collocated with HAARP, measured ion-line intensity. The IRI transmitted O-mode polarized waves parallel to the geomagnetic field at either 2.85 or 3.16 MHz in an on-off cycle, as indicated by the vertical lines. White horizontal specks are stars passing through the imager field of view. Copper horizontal lines are artifacts from ground-clutter returns from MUIR side-lobes. The dashed horizontal line indicates the altitude where the second electron gyro-harmonic is 2.85 MHz, as calculated by an empirical model of the geomagnetic field. See <i>Pedersen et al. [2010]</i> for the original airglow image and description, including an overlay of reflection heights measured with the HAARP ionosonde.	130

Figure 4.13	Time-variation of the 446 MHz HFIL matching height (red) perturbed by 5 min of simulated HF pumping at the bottom-side 4.2 MHz layer and (black) without HF pumping. Sensitivity of the matching height perturbation to the HF-electron volume heating rate is indicated by shading. The red bar indicates the time range of simulated HF-heating. A dashed line shows the natural variation of the 4.2 MHz reflection height 133	133
Figure 4.14	Altitude profile of the electron photoproduction rate in a diurnal-equilibrium model ionosphere on 24 October 2008 over Gakona, Alaska..... 137	137
Figure 5.1	Slant TEC data derived from GPS signals from satellite PRN 30 observed from a GPS receiver at the HAARP facility during two heater “on” periods. 140	140
Figure 5.2	Ionogram taken with the HAARP Digisonde during the heater “off” period in the center of Figure 5.1. The solid black lines on this figure are estimates of the O-mode trace derived from the sounding and the true-height profile (with uncertainty estimates) derived from the estimated trace. This analysis was made using the University of Massachusetts at Lowell ARTIST program (Version 5). Plasma frequency profiles for 00:00 (black dashed line) and 00:05 UTC (orange dashed line) calculated in the 94 nW/m ³ simulated heating experiment described in sections 4.2.4 and 5.2 are overlaid for comparison..... 141	141
Figure 5.3	Map of the area near the HAARP facility showing the geometry of the experiment. Blue and red ovals indicate the 3dB (inner) and 6dB (outer) contours for the IRI beams for the first (blue) and second (red) “on” periods. The blue and red sections of the Ionospheric Penetration Point (IPP) track indicate the satellite PRN 30 locations corresponding to the two “on” periods. 142	142
Figure 5.4	Time variation of TEC calculated below 1000km with the self consistent ionosphere model during and immediately following five minutes of simulated HF heating. The simulation time step is 1 min..... 144	144
Figure 5.5	Temporal evolution of TEC calculated below 1000 km altitude during simulated HF heating experiments. Contours indicate sensitivity of the TEC perturbation to variations in the HF-to-electron power transfer density during five minutes of heating starting at 00:00 UTC..... 145	145
Figure 5.6	(a) Cumulative TEC calculated at selected times during a simulated HF heating experiment. (b) Differences between the ambient and modified profiles are emphasized with shading. Ion and electron temperature profiles at selected times. In each figure, the ambient electron volume heating rate was increased by 94 nW/m ³ from 00:00 to 00:05 UTC at the 4.2 MHz layer..... 147	147

Figure 5.7	Time and altitude variation of the electron number density perturbation from ambient values. Purple contours indicate regions of density enhancement with intervals of $10,000 \text{ cm}^{-3}$ and yellow contours indicate regions of density depletion with intervals of $20,000 \text{ cm}^{-3}$	148
Figure 5.8	DMSP SSIES data from F15 (25 February 2008) showing (a) horizontal and (b) vertical cross-track velocities from the DMSP Drift Meter instrument; (c) total ion density data from the DMSP Scintillation Meter instrument. The vertical line indicates the nearest approach to HAARP. The horizontal lines indicate times when the satellite was within the 3dB (heavy lines) and 6dB (light lines) contours of the IRI beam at 840km altitude.....	150
Figure 5.9	DMSP F15 orbit track and IRI heater beam contours (3dB and 6dB). Left panel is at 200km altitude (DMSP locations traced down field lines); right panel is at 840km altitude. Heavy line along DMSP track indicated region of enhanced ion density. Tic marks along the track are at 5 second intervals.....	152
Figure 5.10	Time evolution of ion and electron temperature. Ion and electron temperature variation following HF heating at 200 km (left) and 250 km (right) altitudes. The ion temperature profiles are separated by 90 sec intervals and the electron temperature profiles are at 9 sec intervals; time increases in the direction of darker shading.....	156
Figure 5.11	Time evolution of the major ion ambipolar diffusion velocity relative to the unperturbed velocity. Change in the vertical component of geomagnetic field-aligned O^+ ambipolar diffusion velocity during HF heating at 200 km (a) and 250 km (b) altitude. The simulated HF heater is switched on at time $t = 0$ s and switched off at $t = 900$ s.....	157
Figure 5.12	Ion and electron density and temperature profiles before (black curves) and after (red curves) 900 sec of simulated HF heating at (a, c) 200 km altitude and (b, d) 250 km altitude. The neutral gas temperature profiles in (c) and (d) are provided by the NRLMSIS00 model.....	159
Figure A.1	Speedup of selected multithreaded Matlab functions using two processor cores per thread (light colors) and one core per thread (dark colors) measured on a 4-core (blue squares) and on a 16-core (red circles) <i>Midnight</i> compute node. The functions operated on one, or two if necessary, 500x500 element arrays of real numbers.....	170
Figure A.2	Workflow overview of the Parallel Computing Toolbox and Matlab Distributed Computing Server. As configured on ARSC's <i>Midnight</i> cluster, both the Matlab client with the Parallel Computing Toolbox (PCT) and the Matlab Distributed Computing Server (MDCS) are started and run on a set of compute nodes allocated to a user by the PBS Pro batch queue system. The image is provided by The Mathworks.....	173

- Figure A.3 Probability distribution function of wall time (sec) required to complete one *parfor* loop iteration in the example program code. Colors represent experiments performed on combinations of *Midnight* node class (4-core X2200 nodes vs. 16-core X4600 nodes) and worker density of ρ workers per processor core. Dark (light) colored curves are probability distribution functions of the time to calculate the decoded signal spectrum at a single row—representing a single radar range bin—estimated for random 1024-by-6000 element (8192-by-6000 element) complex matrices. The mean time μ to complete a loop iteration in each experiment is reported in seconds. 182
- Figure A.4 Speedup factor (dark colors) of the total execution wall time to complete all loop iterations concurrently over an experiment run using *parfor* with a single-worker *matlabpool* on the indicated *Midnight* compute node type. Each color represents a sequence of experiment trials where the number of workers in the *matlabpool* was doubled after processing each 1024-by-6000 element random complex matrix while the number of workers per available processor core was maintained below the indicated threshold ρ . The total execution wall time (light colors) required to process each matrix does not include the time to start Matlab or to connect to the *matlabpool*. 183

List of Tables

	Page
Table 3.1	Parameters used in a self-consistent ionosphere model (SCIM) model run configured to simulate the vertical structure and time-evolution of the natural ionosphere over Gakona, Alaska on 24 October 2008.62
Table 4.1	Geophysical conditions observed during ionosphere HF modification experiments at Gakona, Alaska.90
Table 4.2	MUIR field-aligned coded long pulse configuration parameters and values. Symbolic names used in the text to refer to individual parameters are provided for reference. 100

Acknowledgements

The Arctic Region Supercomputing Center (ARSC) provided funding for the graduate student research stipend and for conference travel from 2005 to 2010. ARSC also provided grants of high-performance computing resources including tens of thousands of CPU hours on the Midnight supercomputer for ionosphere simulations and radar data processing. Midnight is a shared resource allocated through the Department of Defense (DoD) High-Performance Computing Modernization Program (MPCMP).

Several tens of hours of experiment time with the High-Frequency Active Auroral Research Program (HAARP) Ionospheric Research Instrument (IRI) and the Modular UHF Ionosphere Radar (MUIR), including related travel expenses, were funded through Defense Advanced Research Projects Agency (DARPA) under Contract No. HR0011-09-C-0099, and Office of Naval Research grant N000140711082. The experiment time occurred during the October 2009 BRIOCHE campaign, the October 2008 Irregularities campaign, and the Polar Aeronomy and Radio Science (PARS) summer schools in 2007 and 2008. The HAARP open house during the summer of 2006 first stimulated my interest in radio modification of the ionosphere.

The University of Alaska Fairbanks Graduate School provided partial international travel support in 2008 through the Graduate Student Travel Grant.

Chapter 5 is a joint work with James Secan of NorthWest Research Associates, Inc. The measurements of total electron content made during HAARP experiments shown in section 5.1 were supported by the DARPA under Contract No. HR0011-09-C-0099. F. J. Rich of the Massachusetts Institute of Technology Lincoln Laboratories and K. R. Martin of the Institute for Scientific Research at Boston College provided the Defense Meteorological Satellite Program measurements shown in section 5.3.1 from the Air Force Research Laboratory database.

I would like to acknowledge the support, both intangible and tangible, of my advisor Brenton Watkins; my advisory committee members Mark Conde, Sergei Maurits, David Newman, Davis Sentman, and Heinz Weichen; the ARSC staff, Greg Newby and Anton Kulchitsky, in particular; the staff of the Geophysical Institute, the Physics Department, and the Graduate School; and finally, my family, including my parents Ray and Michelle, my wife Vanessa, and the various other critters (e.g., Kitty).

"Essentially, all models are wrong, but some are useful."

—George. E. P. Box

Chapter 1. Introduction

The term *ionosphere* refers exclusively to the ionized gas species in the *thermosphere*, the terrestrial upper atmosphere between approximately 100 and 1000 km altitude consisting primarily of molecular and atomic oxygen and nitrogen. Hard ultra-violet and soft x-ray radiation incident from the sun, in addition to energetic particles accelerated from the magnetosphere, ionize the neutral gases in the thermosphere resulting in the partially ionized plasma that makes up the ionosphere. The state of the ionosphere at any given point in time and space is coupled to the bulk motion, chemical composition, and temperature of the surrounding background thermosphere gases.

The polar ionosphere affects radio communication and navigation in high-latitude regions by modifying or absorbing electromagnetic waves reflecting from or transiting through the partially-ionized atmosphere above about 80 km altitude. It is dynamic and is subject to its own space weather effects. Knowledge of the current and future state of the polar ionosphere is particularly important, not only because of the ionosphere physics unique to the polar region, but also since communication and navigation at high latitudes rely heavily on a well-behaved ionosphere. Signals from geostationary communication satellites above the equator as well as Global Positioning System (GPS) satellites in medium earth orbit must travel obliquely through the ionosphere to reach high-latitude observers, thus magnifying the effects from ionospheric irregularities.

The limited availability of suitable satellites means that point-to-point communication beyond the line-of-sight at polar latitudes still often requires reflecting high-frequency (HF) radio waves off the bottomside ionosphere, and the signal may be degraded or interrupted entirely during geomagnetic storms or other ionospheric weather events. Commercial and military transoceanic flights, for example, regularly use the HF band for communication with land-based stations. However, continuous measurements of basic ionosphere properties like temperature or pressure, needed to make accurate and reliable predictions of ionosphere events affecting radio wave propagation, are spatially sparse due to the difficulty and expense of monitoring near-earth space on a global scale. But even with reliable data, the physics governing the state of the ionosphere at any given time and place is complicated. Relevant considerations include both collisional and collision-less plasma physics, thermodynamics and statistical mechanics, atmospheric dynamics, magnetospheric physics, and temperature-dependent chemistry applied to over a dozen different gases and ions.

1.1 Historical review

Aeronomy was of little practical importance before the invention of radio communication. Measurements were limited to optical observations of the aurora and careful monitoring of suspended magnetized needles to determine variations in the earth's magnetic field. With radio transmitting and receiving devices came the discovery in 1901 by Marconi, explained by Kennelly and Heaviside, of a reflective atmospheric layer, extending upwards from approximately 80 km in altitude [Whitten and Poppoff, 1971], the quality and character of reflectivity varying with latitude, local time, radio wave frequency, and wave polarization. Appleton was the first to label the ionosphere E layer after a series of experiments in the 1920s, and concurrent theoretical development provided a quantitative description for the formation of ionosphere layers. Fortuitous experiments with powerful transmissions at wavelength 1190 m by the new Radio Luxembourg, intended for commercial audiences in 1933, found that radio waves not only interact passively with the ionosphere, but that they can actually modify it as well [Tellegen, 1933].

The development of reliable suborbital sounding rockets in the mid 20th century brought the first opportunities to make in-situ measurements of the upper atmosphere composition and temperature, further motivating theoretical development. The V2 rockets launched from Germany starting in 1944 were the first sounding rockets to take basic measurements of atmospheric density, pressure, temperature, and composition, with the goal of improving rocket reliability [Gunther, 2006]. Scientific sounding rocket flights with V2's were suspended during the War, to be resumed from White Sands, New Mexico in 1949. But the number of chemical species involved, and the complexity of equations describing magnetized plasma physics in a stratified collisional and chemically reactive medium, meant that theoretical studies of the upper atmosphere were severely limited to drastically simplified models (e.g., see Lettau [1951]) before automatic computation with commercially produced computers became available in the 1960s.

Bailey was among the first to recognize practical and scientific applications of radio-based ionosphere modification [Bailey and Goldstein, 1958], beginning theoretical work in 1925. Farley [1963] continued the theoretical predictions of F-region heating by electromagnetic waves, initiated by Bailey, deriving and numerically solving steady-state formulas of electron density and temperature for fixed transmitter parameters and initial ionosphere states. He found that F-region electron temperatures could be enhanced 10% by a 1000 kW transmitter near Lima, Peru, leading to a 5% electron density depletion.

Meanwhile, calculations of diffusion coefficients from rocket measurements [Colegrove *et al.*, 1965; Colegrove *et al.*, 1966] and collisional energy transfer rates between electrons, ions, and neutral species [Banks, 1966a; b; c; Dalgarno and Degges, 1968] provided parameterizations necessary for self-consistent calculations of ionosphere composition and temperature, given a suitable model of the neutral atmosphere. Rees and Walker [1968] calculated rates of electron and ion heating by electric fields in the polar ionosphere. Nagy and Banks [1970] derived an equation to describe the principal ionospheric electron heat source: collisions with non-thermal photoelectrons liberated by solar extreme ultraviolet (EUV) radiation. An early example of a self-consistent ionosphere model incorporating much of the physics known at the time is described by Stubbe [1970].

A uniform and system framework describing the physics of multispecies gas or plasma transport was consolidated by Burgers [1969]. Shortly thereafter, these equations were adapted to the upper atmosphere [Schunk and Walker, 1970a; b; c; 1971; 1973; Schunk, 1975; 1977], aiding the development of self-consistent ionosphere models. Rees and Roble [1975] used an ionosphere model to study stable auroral red arcs, and calculated the resulting changes in the electron temperature profile and the necessary energy input from the aurora.

Incoherent scatter radar (ISR) techniques for remote measurement of ionosphere density and temperature [Evans, 1969] based on theoretical work on scattering of radio waves by plasma density fluctuations [Dougherty and Farley, 1960; Fejer, 1960; Hagfors, 1961; Jahn, 1962; Salpeter, 1960; 1963] dramatically increased the accessibility and repeatability of ionosphere measurements. ISR provided new opportunities to validate model calculations without the use of rockets [Schunk and Nagy, 1978]. Model results could be combined with ISR measurements to indirectly estimate neutral gas temperatures [Nisbet, 1967; Swartz and Nisbet, 1971; Watkins and Banks, 1974]. Rocket flights continued to provide useful measurements, however. In-situ rocket measurements were used by Gerard and Rusch [1979] to validate model thermosphere and ionosphere profiles, initialized by electron flux and neutral gas parameters measured by satellite.

Ionosphere modification theory [Meltz and LeLevier, 1970; Meltz *et al.*, 1974] and experiments at mid-latitude locations near Platteville, Colorado and Arecibo, Puerto Rico continued throughout the 1970s [Meltz and Perkins, 1974]. The later station included a large ISR, allowing measurements of additional HF-modification phenomena such as the parametric decay instability [Perkins *et al.*, 1974], providing new diagnostics of the underlying plasma physics. The Arecibo ISR has the capability to measure ionosphere electron density and temperature above the facility with high-

spatiotemporal resolution (e.g., [Duncan and Gordon, 1982]; [Hansen et al., 1992]), providing a means to directly evaluate model calculations, similar to that described by [Mantas et al., 1981; Perkins and Roble, 1978], of the ionospheric response to HF heating.

Computational power, the number of incoherent scatter radars, and the size of HF modification facilities increased throughout the 1980s. Three dimensional ionosphere and thermosphere general circulation models [Dickinson et al., 1981; Roble and Ridley, 1987; Roble et al., 1987; Schunk, 1988] took full advantage of computer power available at the time. The 3D models used grid sizes appropriate for studying the circulation of plasma and its response to changes in driving solar or magnetospheric inputs, but the grid sizes are much too coarse to gain insight to relatively small-scale HF- or aurora-induced perturbations. One-dimensional models continued to be refined and were capable of days-long comparative studies in conjunction with ISR measurements at Chatanika, Alaska [Rasmussen and Schunk, 1988], and Millstone Hill, Massachusetts [Rasmussen et al., 1988].

The newly installed heating facility installed with the European Incoherent Scatter Scientific Association (EISCAT) radar facilities near Tromsø in Norway, provided the first combined ISR measurements of the high-latitude ionosphere response to HF-induced perturbations [Djuth et al., 1987; Frey et al., 1984; Hagfors et al., 1983; Robinson, 1989]. The High Power Auroral Stimulation (HIPAS) observatory in Alaska [Wong et al., 1990] included a heating facility (incorporating equipment originally installed at the Platteville facility) but not an ISR. Coded pulse techniques (e.g., [Sulzer, 1986]) increased the spatial resolution of ISR measurements and leading to high-resolution studies HF-induced plasma turbulence [Cheung et al., 2001; Djuth et al., 1990; Djuth et al., 1994; DuBois et al., 2001].

Ionosphere models flourished in the 1990s with the widespread availability of relatively powerful desktop computer workstations and freely available empirical models of the neutral atmosphere composition and temperature [Hedin, 1979; 1983; 1987; Picone et al., 2002], neutral winds [Hedin et al., 1996], and geomagnetic field [McLean et al., 2004]. Empirical ionosphere models like the International Reference Ionosphere (IRI) provided another means validating physics-based models [Bilitza and Reinisch, 2008; Lai, 1999]. A high-latitude ionosphere model was used to investigate the effects of parallel electric fields and the divergence of the high-latitude geomagnetic field on auroral- and photo-electron energy distributions by incorporating an improved electron transport equation calculation [Min, 1993; Min et al., 1993]. Uncertainties in numerical solutions of the

electron transport equation solution were investigated by *Lummerzheim and Lilensten* [1994]. One-dimensional numerical modeling of the ionosphere over the EISCAT facility [*Diloy et al.*, 1996b; *Robineau et al.*, 1996] took advantage of ISR measurements, and focused on explaining radar observations like intermittent ion outflow events [*Blelly et al.*, 1996]. High-latitude HF-heating experiments in a convecting ionosphere were simulated by *Mingaleva and Mingalev* [1997].

Work on improving parameterizations of thermal electron energy transfer rates, largely abandoned since the 1970s, was updated just before the turn of the century [*Pavlov*, 1998a; b; *Pavlov and Berrington*, 1999] and the resulting difference in physics-based model calculations is described by *Pavlov et al.* [2000].

A new 3D simulation of the polar ionosphere, by *Maurits et al.* [2000], improved coverage of a region not well served by previous 3D models. *Kulchitsky et al.* [2005] further improved the model by incorporating the total variation diminishing advection scheme. The accuracy and efficiency of photoionization calculations, a significant computational burden of general circulation models, was improved by *Solomon and Qian* [2005].

Perrine et al. [2006] simulated the response of convecting ducts of ionospheric plasma connecting geomagnetic conjugate points, disturbed by HF-heating at with power transfer rates available at the High Frequency Active Auroral Research Program (HAARP) facility. The model used SAMI2, a two-dimensional model of the low-latitude ionosphere [*Huba et al.*, 2000]. A high-latitude ionosphere model was used by *Mingaleva and Mingalev* [2008] to study large-scale density perturbations over the Sura facility near Nizhny Novgorod, Russia as a function of heater power.

Recent experiments at the HAARP and EISCAT heating facilities have focused on creating and understanding the mechanism behind HF-induced airglow [*Gustavsson et al.*, 2006; *Kosch et al.*, 2004; *Kosch et al.*, 2007; *Pedersen et al.*, 2010; *Pedersen and Carlson*, 2001; *Pedersen et al.*, 2003]. Model calculations described by *Gustavsson et al.* [2005] fitted model electron energy distribution functions to observed distributions in order to estimate physical conditions that could produce the observed airglow.

1.2 This work

The objective of this thesis is to understand the response of the ionosphere to high-power HF radio waves. To facilitate this work, a numerical model of the ionosphere has been developed for comparative studies with experiment data.

A fluid description of the ionosphere is used in the *self-consistent ionosphere model* (SCIM) to model the state of the ionosphere at each point in time and space by the number density of each ion constituent, the average ion temperature, and the electron temperature. The model is used to simulate the response of the ionosphere to powerful HF electromagnetic waves incident from the ground, similar those broadcast during experiments conducted at the High Frequency Active Auroral Research Program (HAARP) facility in Gakona, Alaska. In particular, ionosphere parameters including the HF reflection height, total electron content, and vertical ion flux are used to relate model calculations to measurements made during HAARP radio modification experiments. Model results are used to interpret field measurements made during controlled experiments so that additional insight to the relevant physical processes underlying the measurements can be made. This information may be used later to plan future experiments or predict effects on communication systems sensitive to perturbations in the ionosphere.

Chapter 2 describes the self-consistent ionosphere model (SCIM) in detail by listing the relevant equations used to model a one-dimensional, slab-symmetric, high-latitude daytime ionosphere over Gakona, Alaska. The model equations are validated by plotting altitude profiles of equation terms like the electron photoproduction rate or ion-neutral collision frequency that significantly influence the model output and determine the vertical structure of the ionosphere. Altitude profiles of these terms can be compared with published profiles that were, for example, measured remotely by incoherent scatter radar (ISR), measured in-situ by sounding rockets, or calculated with a computer model.

Chapter 3 describes the numerical method and computer program used to initialize and solve the self-consistent ionosphere model equations. This is the first reported use of the Skeel-Berzins numerical “method of lines” to solve the coupled parabolic continuity-momentum equations—for the individual ion species—that are fully coupled to the ion and electron gas energy equations. Physics as modeled by selected chemical species and their interactions is comparable to the physics included in other one-dimensional models [Blelly *et al.*, 1996; Diloy *et al.*, 1996a; Diloy *et al.*, 1996b;

Robineau et al., 1996; *Roble*, 1975; *Stubbe*, 1970]. Note in particular that the model by *Robineau et al.* [1996] uses either a method of lines solution or a flux corrected transport method to solve the model equations. The model by *Bailey et al.* [2002], like SCIM, makes extensive use of the GLOW [*Solomon et al.*, 1988; *Solomon and Abreu*, 1989] and MSIS [*Hedin*, 1987; *Picone et al.*, 2002] models for calculations of photo- and impact-excitation and ionization effects. The model solution is validated by comparing the daytime ionosphere critical frequency (f_oF_2) and total electron content (TEC) calculated over a diurnal cycle with SCIM to f_oF_2 measured over Gakona with the HAARP ionosonde and TEC measured with a global positioning system (GPS) receiver at HAARP. The International Reference Ionosphere (IRI), an empirical rather than physics-based model, is also used to calculate f_oF_2 and TEC over a diurnal cycle for comparison with the measured and SCIM-calculated quantities. Finally, the sensitivity of SCIM results to the semidiurnal variation in thermosphere number density is explored.

Chapter 4 describes experiments conducted at HAARP during October 2008 and October 2009, times of very quiet solar and geomagnetic activity that are ideal for simulating with a one-dimensional ionosphere model. A coded long-pulse technique is presented for measuring the HF reflection height with high spatiotemporal resolution using the Modular UHF Ionosphere Radar (MUIR), co-located with the HAARP Ionosphere Research Instrument (IRI). The technique is verified by measuring the range to a satellite that passed over MUIR on 22 March 2009. Validation of the technique is provided by comparing time-dependent observations of HF-enhanced airglow along the HAARP geomagnetic field line with MUIR coded long-pulse measurements of the HF reflection height, near where the HF-driven plasma turbulence leads to electron gas heating and strongly enhanced radar reflectivity. The HF reflection height is observed with MUIR via the HF-enhanced ion line (HFIL) to ascend in the minutes following strong HF transmissions lasting tens to hundreds of seconds. Measurements of the spatiotemporal evolution of enhanced ion-line and plasma-line reflection heights during HF-heating are similar to earlier experiments [*Djuth et al.*, 1990; *Djuth et al.*, 1994; *Rietveld et al.*, 2000]. Radio modification experiments are simulated with SCIM by introducing a new heat source term in the electron energy equation. Others have followed a similar approach to modeling radio modification experiments in the Eulerian frame [*Mantas et al.*, 1981; *Perkins and Roble*, 1978]; in particular, *Perrine et al.* [2006] solves model equations on the entire mid-latitude geomagnetic field line. *Mingaleva and Mingalev* [1997] simulates HF heating effects in a Lagrangian frame by solving the model equations on a column of the convecting ionosphere. SCIM calculations indicate that the ascent of the ion line (and

reflection height) are caused by a HF-induced depletions of plasma in the F-region peak plasma density that are apparent when the HF-induced plasma enhancements in the lower F region dissipate following cessation of continuous HF transmissions. Comparison of model ionosphere response to radio modification experiments to measured variations in HFIL reflection heights extends prior results [Ashrafi *et al.*, 2006; Ashrafi *et al.*, 2007].

Chapter 5 further explores effects from HF heating of the electron gas on the field-aligned electron density profile. In particular, time-dependent HF-induced TEC perturbations measured at HAARP are found with SCIM to result from electron density enhancements in the lower F region and the topside ionosphere that offset the electron density depletion surrounding the critical layer. The topside plasma density enhancement follows first from temperature-enhanced ambipolar diffusion along the geomagnetic field; atomic oxygen ions lifted to the topside ionosphere persist where chemical loss reactions with the molecular neutrals are less efficient. Measurements of topside ion flux measured with a Defense Meteorological Satellite Program (DMSP) satellite during an overpass of HAARP while the IRI was operating are presented and compared with SCIM calculations of the ion flux perturbation. The electron temperature, electron density, and ion flux perturbation profiles are also found with SCIM to be sensitive to the altitude of HF heating, determined by the chosen HF frequency and ambient ionosphere density profile along the HF ray path.

Chapter 6 summarizes the main results of this thesis, and describes ideas for future work that build on the main results.

High Performance Computing (HPC) resources were used extensively to perform SCIM calculations and process MUIR coded long-pulse data. The Appendix discusses the scalability and performance of these tasks on the *Midnight* supercomputer at the Arctic Region Supercomputing Center (ARSC).

Chapter 2. Ionosphere model equations

This chapter describes the fluid transport equations used to model the time-dependent state of the ionosphere by introducing the equations through their relation with the Boltzmann equation. Simplifications and approximations applicable to a slab-symmetric and time-dependant one-dimensional numerical model of the ionosphere are then applied to the transport equations. Input parameters derived from external empirical models are noted and the application of the parameterizations in the model ionosphere is discussed. The model equations are verified throughout the chapter by altitude profile plots of selected model equation quantities. These quantities include inputs from empirical models such as the major neutral gas number density or incident ionizing radiation spectrum; intermediate quantities like the electron photoproduction rate or electron heating rate; and final solution variables like the ion species number densities and electron temperature. The profiles of the various quantities are in agreement of values published in literature or measurements.

The numerical solution of the transport equations with a parabolic and elliptic partial differential equation solver software library is discussed in section 3.1. There, the following main functions of the SCIM software program are described: generation of time-dependent model boundary conditions or parameters from empirical model data, solution of the model transport equations, and archiving the model calculation results to non-volatile storage. The time-dependent numerical solution to the model equations are validated in section 3.2 where the daytime diurnal variation of the modeled ionosphere critical frequency f_oF_2 and total electron content (TEC) is compared with ionosonde and GPS measurements made at the HAARP facility in Gakona, Alaska. Sensitivity of the model to neutral atmosphere semi-diurnal tidal variations is also studied, with model results compared to those expected through physical intuition.

2.1 Boltzmann equation

The continuity, momentum, and energy transport equations from Grad's 13-moment approximation to the Boltzmann equation, as applied to the upper atmosphere [Schunk, 1975], are used here to model the steady-state and time-evolution of a slab-symmetric ionosphere. A gas mixture composed of multiple species of atomic, molecular, and ionized constituents can be described by a set of Boltzmann equations, one for each species of gas. The transport equations solved by the self-consistent ionosphere model are related through *moments* of the Boltzmann equation.

In an inertial reference frame, the Boltzmann equation for a species s in a gas mixture is

$$\frac{\partial f_s}{\partial t} + \mathbf{v}_s \cdot \nabla f_s + \left[\mathbf{G} + \frac{e_s}{m_s} (\mathbf{E} + \mathbf{v}_s \times \mathbf{B}) \right] \cdot \nabla_{\mathbf{v}_s} f_s = \frac{\delta f_s}{\delta t} \quad (2.1)$$

where $f_s(\mathbf{r}, \mathbf{v}_s, t)$ is the distribution function of species s at position \mathbf{r} and time t with velocity \mathbf{v}_s , $\mathbf{G}(\mathbf{r})$ is the position-dependent acceleration due to gravity, e_s is the electric charge of a single s particle, m_s is the mass of a single s particle, $\mathbf{E}(\mathbf{r}, t)$ is the electric field, and $\mathbf{B}(\mathbf{r}, t)$ is the magnetic field. Generally the electric and magnetic fields at any given point in plasma depend on the distribution and velocity of the surrounding plasma, but in this work it is assumed that the fields are imposed externally upon the ionosphere and are independent of the species distribution functions. Interaction effects such as collisions between particles affect the time derivative of the distribution function through the *collision term* $\delta f_s / \delta t$.

The closed 13-moment approximation to the Boltzmann equation is derived by taking velocity moments of equation (2.1) and substituting the velocity distribution of each gas species by a truncated orthogonal series expansion, e.g. see *Schunk and Nagy* [2000]. Of the resulting set of transport equations (continuity, momentum, energy, pressure tensor, and heat flow), the SCIM software solves the *continuity equation*:

$$\frac{\partial n_s}{\partial t} + \nabla \cdot (n_s \mathbf{u}_s) = \frac{\delta n_s}{\delta t} \quad (2.2)$$

The *momentum equation*:

$$n_s m_s \frac{D_s \mathbf{u}_s}{Dt} + \nabla p_s + \nabla \cdot \boldsymbol{\tau}_s - n_s e_s (\mathbf{E} + \mathbf{u}_s \times \mathbf{B}) - n_s m_s \mathbf{G} = \frac{\delta \mathbf{M}_s}{\delta t} \quad (2.3)$$

And the *energy equation*:

$$\frac{D_s}{Dt} \left(\frac{3}{2} p_s \right) + \frac{5}{2} p_s (\nabla \cdot \mathbf{u}_s) + \nabla \cdot \mathbf{q}_s + \boldsymbol{\tau}_s : \nabla \mathbf{u}_s = \frac{\delta E_s}{\delta t} \quad (2.4)$$

The terms on the right hand side of equations (2.2), (2.3), and (2.4) are the collision terms that can be evaluated if the interaction potentials between species particles are known or given, the species volume number density is $n_s(\mathbf{r}, t) = \int d^3 v_s f_s(\mathbf{r}, \mathbf{v}_s, t)$, the species drift velocity is

$\mathbf{u}_s(\mathbf{r}, t) = \int d^3v_s \mathbf{v}_s f_s(\mathbf{r}, \mathbf{v}_s, t) / \int d^3v_s f_s(\mathbf{r}, \mathbf{v}_s, t)$, the operator $\frac{D_s}{Dt} = \frac{\partial}{\partial t} + \mathbf{u}_s \cdot \nabla$ is the convective derivative, $p_s = n_s k_B T_s$ is the partial species pressure, k_B is the Boltzmann constant, and $\boldsymbol{\tau}_s$ is the species stress tensor. The stress, or viscosity, tensor will only have non-negligible contribution to the model equations in the terrestrial domain of interest for non-Maxwellian ion velocity distributions so it will be suppressed [Prölss, 2004]. Even this reduced set of transport equations is too difficult to solve analytically, and worse, the equations depend significantly on quantities that are generally difficult or expensive to measure. Consequently, a number of approximations will be applied to each equation in the following sections.

2.1.1 Continuity equation

The species number density continuity equation relates the time rate of change of the species number density in an arbitrarily small volume, or neighborhood, of space to the balance of the divergence of the mean species number density flux with the local density production and loss rates:

$$\frac{\partial n_s}{\partial t} = -\nabla \cdot (n_s \mathbf{u}_s) + \frac{\delta n_s}{\delta t} \quad (2.5)$$

The continuity equation cannot be solved at this point because the species drift velocity is still an unknown function of species number density, and of other quantities. Instead, the continuity equation will be directly coupled to the species momentum equations, describing the time dependence of the species drift velocities, and it will also be used to simplify the energy equations. Both the momentum and energy equations are described in sections 2.1.2 and 2.1.3, respectively. Much of the physics describing the E - and F_1 -region ionosphere can be described by balancing local production and loss rates due to collisions between species (chemical reactions) or between solar photons and species (photoionization). These sorts of physical processes, including the ionization and dissociation of major neutral atmosphere species by energetic photons and electrons as well as chemical reactions between the resulting ionization and dissociation products, are incorporated in the collision term on the right hand side of equation (2.5). Physics described through the collision term of the continuity equation will be briefly reviewed in the context of the one-dimensional continuity equation describing a vertical profile of the ionosphere.

The model will be used in Chapter 3 and Chapter 4 to simulate the (field-aligned) ionosphere structure and its response to driven perturbations, so the *continuity equation* (2.5) will be approximated with a scalar equation by assuming slab symmetry and neglecting the horizontal components of the species velocity vector

$$\frac{\partial n_s}{\partial t} = -\frac{\partial}{\partial z}(n_s w_s) + \frac{\delta n_s}{\delta t}. \quad (2.6)$$

While this one-dimensional treatment is not a good approximation of the neutral atmosphere, the plasma motion is restricted to the geomagnetic field line (in the absence of perpendicular electric fields), so the approximation is reasonable for a model the ionosphere. The species vertical velocity w_s is determined by the momentum equation, described in section 2.1.2. The local production and loss term $\delta n_s / \delta t$ includes production sources driven by solar and magnetospheric interaction with the neutral atmosphere as well as local sources and sinks due to chemical reactions between the molecular and ionic species.

$$\frac{\delta n_s}{\delta t} = \left. \frac{\delta n_s}{\delta t} \right|_{\text{driven production}} + \left. \frac{\delta n_s}{\delta t} \right|_{\text{chemical production and loss}} \quad (2.7)$$

Major sources of species production driven by photo-ionization, photo-dissociation, and photo-excitation of neutral atmospheric species include ultraviolet and soft X-ray solar radiation incident on the background neutral atmosphere, secondary impacts of neutral molecules with photoelectrons, and collisions of neutral molecules with high-speed auroral electrons:

$$\left. \frac{\delta n_s}{\delta t} \right|_{\text{driven production}} = \left. \frac{\delta n_s}{\delta t} \right|_{\text{UV and EUV photo-production}} + \left. \frac{\delta n_s}{\delta t} \right|_{\text{photoelectron impact-production}} + \left. \frac{\delta n_s}{\delta t} \right|_{\text{auroral electron impact-production}} \quad (2.8)$$

The photo-production term in equation (2.8) depends on the vertical density profile of the major neutral species density profiles and the intensity spectrum of the solar radiation incident from space. In this work, the time-dependent number-density altitude profiles of the major neutral atmosphere species $\{N_2, O_2, O\}$ are provided by the NRLMSISE-00 (MSIS) empirical model [Picone *et al.*, 2002] as functions of geographic latitude (GLAT) and longitude (GLONG), universal time (UT), the geomagnetic index A_p , and the daily solar 10.7 cm wavelength radio flux index ($F_{10.7}$). Typical altitude profiles of neutral atmosphere species number densities over Gakona, Alaska during an early autumn evening of quiet geophysical conditions as calculated by MSIS is illustrated

in Figure 2.1. A SCIM-calculated electron density profile is shown in the figure for reference and comparison.

Time-dependent altitude profiles of the photoionization and photo-excitation rates of the major neutral atmosphere species are calculated with the GLOW model (version 0.97) [Solomon *et al.*, 1988; Solomon and Abreu, 1989]. The solar extreme ultraviolet (EUV and XUV) photon flux spectrum has been measured and recorded since 22 January 2002 with the Solar EUV Experiment (SEE) [Woods *et al.*, 2005], part of the National Aeronautics and Space Administration (NASA) Thermosphere Ionosphere Mesosphere Energetics and Dynamics (TIMED) mission. However, this thesis uses a proxy measurement of the EUV flux based on the $F_{10.7}$ index. The parameterization

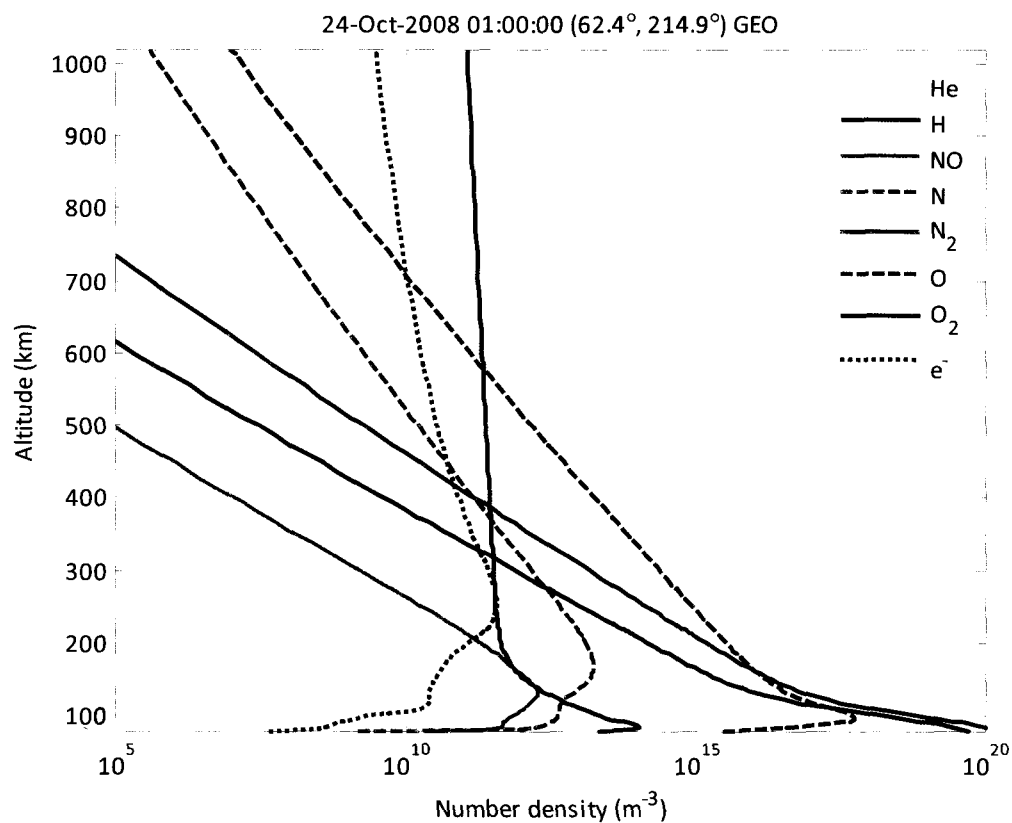


Figure 2.1 Altitude profiles of neutral species and electron number density above Gakona, Alaska at 01:00 UT (hh:mm) on 24 October 2008.

facilitated efficient studies of the ionosphere response to solar irradiance variations, aiding model development. The Hinteregger parameterization [Hinteregger and Fukui, 1981] included with the GLOW model is used here to provide a spectrum of solar ionizing radiation spectrum as a function of the $F_{10.7}$ index measurements. Typical solar photon flux spectra calculated using the Hinteregger parameterization in the GLOW model for low and moderate input values of the $F_{10.7}$ index are illustrated in Figure 2.2. Altitude profiles of the resulting major neutral species photoionization rates are plotted in Figure 2.3. The photoproduction rates in the figure are calculated using the same model neutral atmosphere illustrated in Figure 2.1 and the low-activity model solar photon flux spectrum illustrated in Figure 2.2.

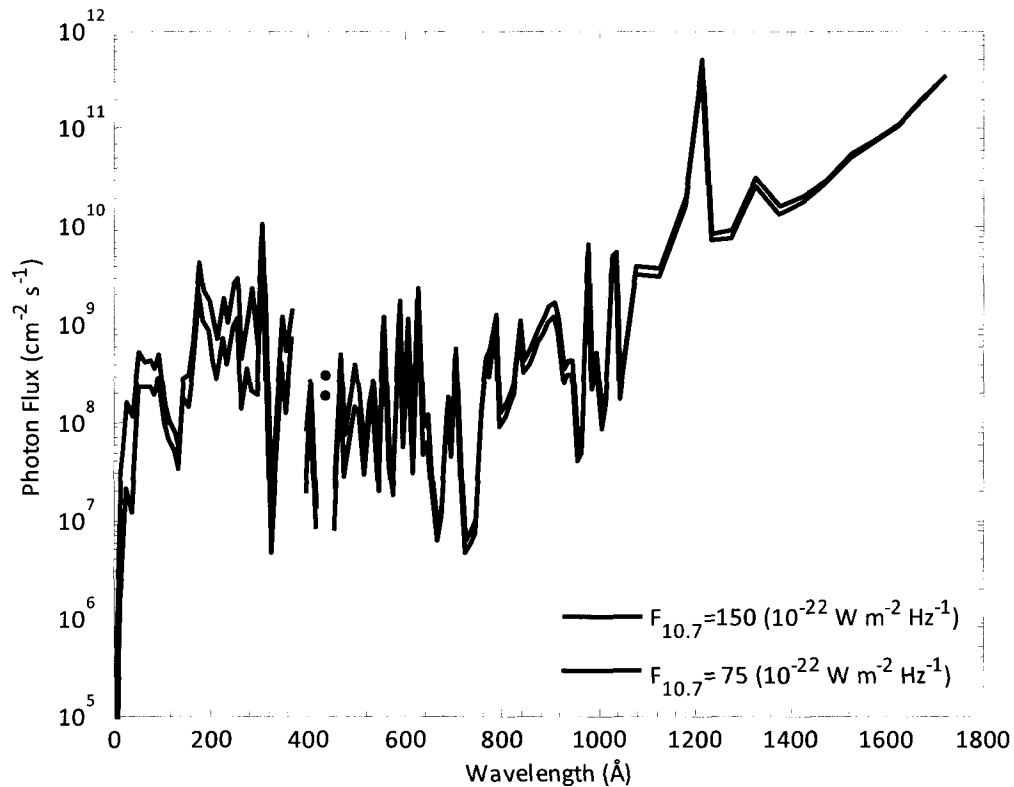


Figure 2.2 Solar photon flux at approximately 10 Å intervals calculated with the Hinteregger model for two values of the F107 solar radio flux index.

Impact ionization caused by the motion of high-speed auroral electrons through the high-latitude neutral atmosphere can contribute significantly to the total neutral atmosphere ionization rate but will not be considered in this work due to the high spatiotemporal variability of such phenomenon. However, the electron production rate including impact ionization and excitation contributions resulting from collisions of neutral gas molecules with energetic photoelectrons is included in the SCIM solution of the transport equations. For example, the altitude profile of the photoelectron impact ionization rate for the conditions illustrated in Figure 2.3 reaches its peak value and is comparable in magnitude to the total photoionization rate in the range of 120 to 220 km altitude. At the top simulation boundary near 1000 km altitude, the secondary electron production rate is about an order of magnitude less than the primary electron production rate.

Ionospheric ions interact with surrounding molecules in the neutral atmosphere through chemical reactions or they can recombine with free electrons. The production and loss of rates of each species can be described with reaction stoichiometry through sets of chemical reaction equations. Each reaction relates the production rates of the reaction products to the number density of the reaction reactants. For example, the two major exothermic chemical reactions that remove atomic oxygen ions and produce molecular oxygen ions and nitric oxide ions in the ionosphere can be described by the following equations from *Rees* [1989]:

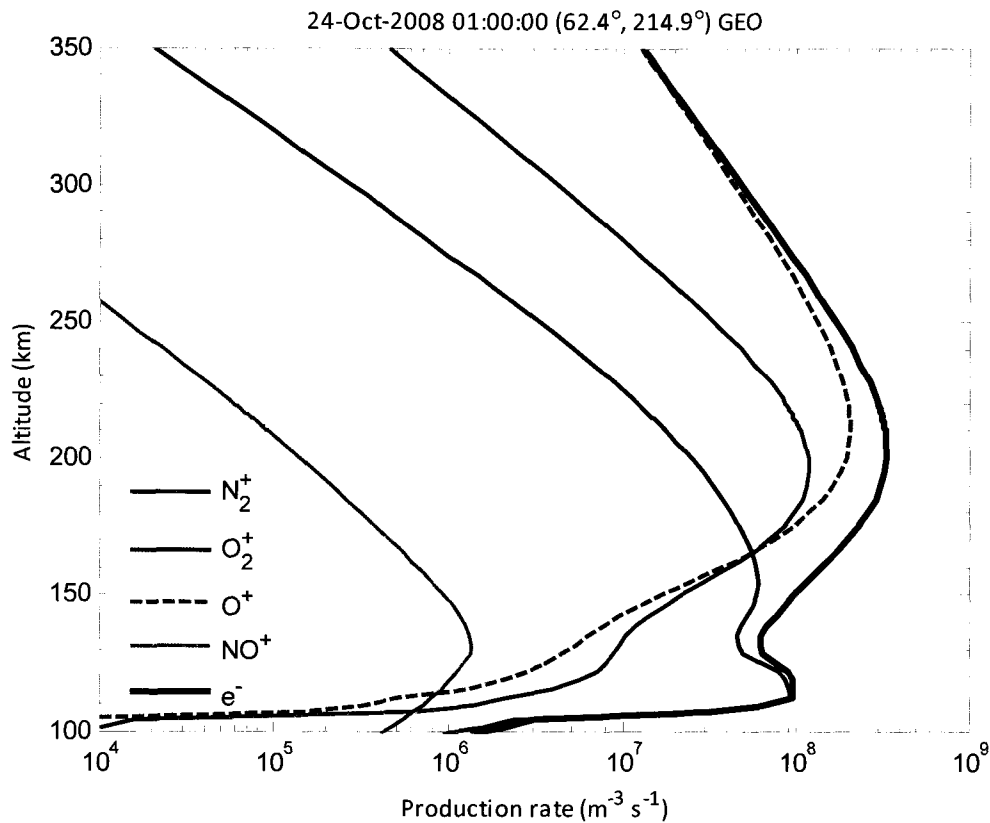


Figure 2.3 Altitude profiles of major and minor ion photoproduction rates modeled above Gakona, Alaska at 0100 UT on 24 October 2008.

$$\begin{aligned}
 & \text{O}^+ (^4\text{S}) + \text{O}_2 \rightarrow \text{O}_2^+ + \text{O} & \gamma_2 = 2 \times 10^{-17} \left(\frac{T_R}{300} \right)^{-0.4} \\
 & \text{O}^+ (^4\text{S}) + \text{N}_2 \rightarrow \text{NO}^+ + \text{N} (^4\text{S}) & \gamma_1 = \begin{cases} 5 \times 10^{-19}, & T_R < 1000 \text{ K} \\ 4.5 \times 10^{-19} \left(\frac{T_R}{300} \right)^2, & T_R \geq 1000 \text{ K} \end{cases}
 \end{aligned} \tag{2.9}$$

The chemical reaction rate coefficients γ_2 and γ_1 in the equations have units of m^3s^{-1} ; the reduced temperature $T_R = (T_i + T_n)/2$ in the equations is the mean of the ion and neutral temperature and is expressed in degrees Kelvin. Since the chemical reaction rate coefficients are temperature dependent, the production and loss terms in the continuity equation are coupled to

the energy equations. Experimental measurement of the chemical reaction rates for species under pressures and temperatures typically found in the natural ionosphere is difficult and subject to significant uncertainty. Consequently, published reaction rate coefficients for a single chemical reaction may vary. The chemical reactions and respective reaction rate coefficients used in the model described here include the equations tabulated by *Rees* [1989], a source that contains a relatively complete set of equations and includes excited state chemical reaction rate coefficients. A smaller set of chemical reaction equations and rate coefficients based on more recent measurements is given by *Schunk and Nagy* [2000] while a very complete -- and actively maintained online -- set of equations is cataloged by *Anicich* [1993].

Regardless of the specific reaction rates used, the time rate of change of volume number density for a particular species s , neglecting transport into or out of the region of interest, can be expressed with the continuity equation

$$\frac{\delta n_s}{\delta t} = P_s - n_s L_s \quad (2.10)$$

where P_s is the production rate including direct ionization from equation (2.7) and chemical production; L_s is the loss frequency due to chemical reactions and spontaneous emission. Both P_s and L_s generally depend on the neutral and ion temperatures as well as on the number densities of the other species. The loss frequency altitude profiles for major ions calculated by SCIM are plotted in Figure 2.4. The model electron density profile is included as well in order to illustrate that electron recombination is the dominant loss process for certain ion species (e.g., NO^+ and O_2^+) while chemical reactions with the neutral atmosphere is the dominant loss process for other ion species (e.g., H^+ , O^+ , and N_2^+).

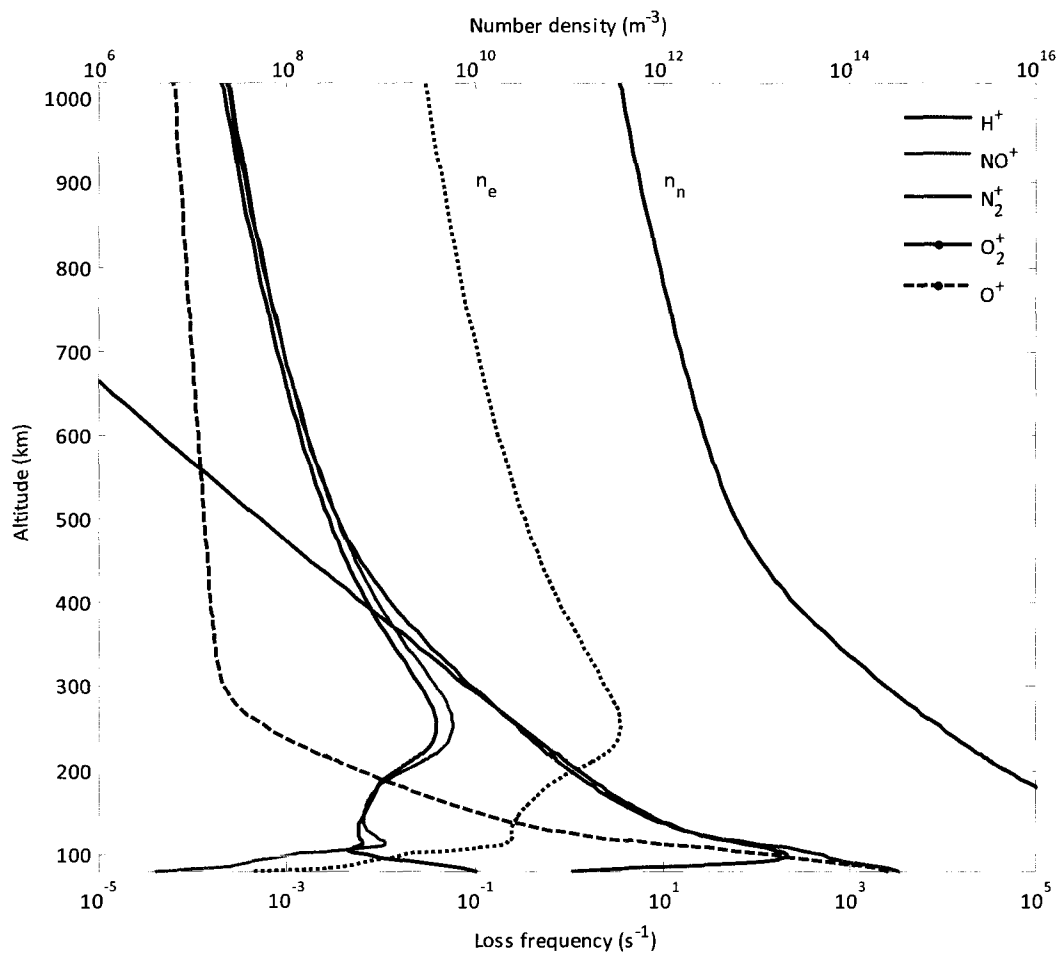


Figure 2.4 Model altitude profiles of the loss frequency of major ion species due to chemical reactions or electron recombination over Gakona, Alaska on 24 October 2008 at 0100 UT. Model altitude profiles of electron n_e and neutral species n_n number density are plotted for reference.

In the *E* region, where the ionosphere physics is dominated by collisional effects, direct ionization and chemical production rates of each ion species are nearly in balance with the chemical and recombination loss rates; the number density of each ion species adjusts within tens of seconds to changes in the production and loss rates or solar zenith angle. Another consequence of frequent ion-neutral collisions is that an ion in the *E* region is destroyed—chemically—near the location it was created so transport into and out of a small neighborhood, relative to the ionosphere scale height, is negligible and can be neglected in the continuity equation. An O^+ ion created in the *F* region, however, may travel a vertical distance comparable to the local ionosphere scale height due

the low collision frequency in the F region and the long characteristic time of O^+ (4S) chemical loss processes. Therefore the transport term in the vertical continuity equation (2.6), the divergence of the mass flux, is comparable to the sum of the local production and loss terms and so the number density of each species in the F region cannot be estimated simply by equating the production and loss terms as described in the section 3.1.2 discussion on the lower model boundary condition. Instead, an expression for the species vertical drift velocity w_s is also needed in order to solve for the density. The altitude and time-dependent vertical drift velocity of a species can be determined by solving the momentum equation for each species.

2.1.2 Momentum equation

The momentum equation is used in the SCIM calculations to calculate the vertical drift velocity of each ionized and minor neutral species in response to external forces, number density gradients, and temperature gradients. The momentum equation of each species is coupled to the continuity, momentum, and energy equations of the other species through collisions and electrostatic forces. Simplifications will be applied to the momentum equation with the end goal of deriving a system of parabolic and elliptic partial differential equations in the general form required by the numerical solver described in section 3.1. In this section, a momentum equation that describes the time-dependent state of the ionosphere along a vertical profile that co-rotates with the Earth over a set of pre-defined geographic coordinates during a specified period of time is derived from equation (2.3). The momentum equation, like the energy equation, will eventually be directly coupled with the continuity equation to reduce the number of equations to solve and put each equation into parabolic form.

Although the momentum equation (2.3) is expressed in an inertial frame and the ionosphere model domain is assumed to co-rotate with the Earth, non-inertial effects such as the centrifugal and Coriolis accelerations will be neglected in the model calculations and are generally insignificant in a high-latitude model of the vertical ionosphere. Recall that the stress term in equation (2.3) is also assumed to be insignificant in the simulation domain so the momentum equation for species s can be written:

$$n_s m_s \frac{D_s \mathbf{u}_s}{Dt} + \nabla p_s - n_s e_s (\mathbf{E} + \mathbf{u}_s \times \mathbf{B}) - n_s m_s \mathbf{G} = \frac{\delta \mathbf{M}_s}{\delta t} \quad (2.11)$$

In the following derivation of the model momentum equations, it is convenient to separate the electron momentum equation from the ion and neutral species momentum equations using the index i to represent a specific ion species, the index n to represent a specific neutral species, the index t will range over the major neutral species $\{N_2, O, O_2\}$, and the index e to represent the electron gas (the index s will also be used later to refer to either an ion or neutral species)

$$\begin{aligned} n_e m_e \frac{D_e \mathbf{u}_e}{Dt} + \nabla p_e - n_e e_e (\mathbf{E} + \mathbf{u}_e \times \mathbf{B}) - n_e m_e \mathbf{G} &= \frac{\delta \mathbf{M}_e}{\delta t} \\ n_i m_i \frac{D_i \mathbf{u}_i}{Dt} + \nabla p_i - n_i e_i (\mathbf{E} + \mathbf{u}_i \times \mathbf{B}) - n_i m_i \mathbf{G} &= \frac{\delta \mathbf{M}_i}{\delta t} \\ n_n m_n \frac{D_n \mathbf{u}_n}{Dt} + \nabla p_n - n_n m_n \mathbf{G} &= \frac{\delta \mathbf{M}_n}{\delta t}. \end{aligned} \quad (2.12)$$

The spatial resolution of the ionosphere model – about 500 meters in the E region -- is significantly larger than the Debye length of the ionosphere plasma so the plasma is assumed to be electrically neutral. Furthermore, it is assumed here that all ions in the simulation domain are singly charged and that there are no negatively charged ions. Then the electron density is equal to the sum of the ion densities

$$n_e = \sum_i n_i \quad (2.13)$$

and the Lorentz force terms cancel in equation (2.12) by adding the ion and electron equations because $e_e = -e_i$. The resulting combined ion-electron momentum equation is

$$n_e m_e \frac{D_e \mathbf{u}_e}{Dt} + \sum_i n_i m_i \frac{D_i \mathbf{u}_i}{Dt} + \nabla \left(p_e + \sum_i p_i \right) - n_e m_e \mathbf{G} - \sum_i n_i m_i \mathbf{G} = \frac{\delta \mathbf{M}_e}{\delta t} + \sum_i \frac{\delta \mathbf{M}_i}{\delta t} \quad (2.14)$$

The collision terms on the right hand side of the momentum equation (2.11) and the subsequent equations (2.12) and (2.14) can be expressed using Burgers linear collision theory [Burgers, 1969]—summarized by Schunk [1975] using notation similar to that used here—since the flow velocities are typically much smaller than the species thermal velocities:

$$\frac{\delta \mathbf{M}_s}{\delta t} = - \sum_t n_s m_s \nu_{st} (\mathbf{u}_s - \mathbf{u}_t) \quad (2.15)$$

In equation (2.15), ν_{st} is the momentum-transfer collision frequency between one molecule of species s with molecules from target species t , and \mathbf{u}_t is the flow velocity of target species t . Note that the contribution of heat flow to the collision term has been neglected. The right hand side of (2.14) then expands to

$$\frac{\delta \mathbf{M}_e}{\delta t} + \sum_i \frac{\delta \mathbf{M}_i}{\delta t} = -\sum_t n_e m_e \nu_{et} (\mathbf{u}_e - \mathbf{u}_t) - \sum_i \sum_t n_i m_i \nu_{it} (\mathbf{u}_i - \mathbf{u}_t) \quad (2.16)$$

and by the quasi-neutrality assumption, the electrons in any sufficiently large volume of space (typically any volume larger than a Debye sphere) can be partitioned such that $n_e = \sum_i n_{e_i}$ where $n_{e_i} = n_i$ is the number density of the electrons associated with ion species i . Similarly, to maintain quasi-neutrality, the electrons associated with ion species i must drift with those ions so $\mathbf{u}_{e_i} = \mathbf{u}_i$. Then by expanding the electron number density into the associated ion densities and collecting terms equation (2.16), the electron mass, which is at least four orders of magnitude less than the mass of each ion species, can be used to eliminate the electron collision terms:

$$\begin{aligned} \frac{\delta \mathbf{M}_e}{\delta t} + \sum_i \frac{\delta \mathbf{M}_i}{\delta t} &\approx -\sum_i \sum_t n_{e_i} m_e \nu_{et} (\mathbf{u}_{e_i} - \mathbf{u}_t) - \sum_i \sum_t n_i m_i \nu_{it} (\mathbf{u}_i - \mathbf{u}_t) \\ &= -\sum_i \sum_t \left[n_{e_i} m_e \nu_{et} (\mathbf{u}_{e_i} - \mathbf{u}_t) + n_i m_i \nu_{it} (\mathbf{u}_i - \mathbf{u}_t) \right] \\ &= -\sum_i (n_i m_i) \sum_t (\mathbf{u}_i - \mathbf{u}_t) \left(\frac{m_e}{m_i} \nu_{et} + \nu_{it} \right) \\ &\approx -\sum_i (n_i m_i) \sum_t (\mathbf{u}_i - \mathbf{u}_t) \nu_{it} \\ &\approx \sum_i \frac{\delta \mathbf{M}_i}{\delta t} \end{aligned} \quad (2.17)$$

In a typical gas or plasma mixture, the momentum transfer collision frequency between a molecule with a velocity near the species thermal velocity and the molecules of a target species is, to a first order approximation, proportional to the number density of the target species. Consequently, the ion-ion momentum-transfer collision frequency between an ion species with any target ion species is assumed to be negligible compared to the major ion-neutral collision frequencies in the model domain. The result in equation (2.17) can be simplified further by assuming that the major neutral species all drift at the same velocity, $\mathbf{u}_t = \mathbf{u}_N$ for $t \in \{\text{O}, \text{O}_2, \text{N}_2\}$, so that the difference in ion and neutral flow velocities can be factored out of the sum over the target species

$$\sum_i \frac{\delta \mathbf{M}_i}{\delta t} = - \sum_i n_i m_i (\mathbf{u}_i - \mathbf{u}_N) \sum_t \nu_{it} \quad (2.18)$$

Momentum transfer collision frequencies between ion species and target major neutral species, expressed in proportion to the number density of each target species, are compiled in the literature [Schunk, 1988; Schunk and Nagy, 2000]. The collision frequencies used in the SCIM calculations include resonant and non-resonant ion-neutral collision frequencies.

Quasi-neutrality and the small electron to ion mass ratio can be used with the electron-density partitioning process described above to simplify the left hand side of equation (2.14). First, the electron gravitational force term $-n_e m_e \mathbf{G}$ is small relative to the respective ion gravitational force terms so it may be neglected. Similarly, the electron force term is small relative to the respective ion force terms so it also will be neglected. And finally, by applying the ideal gas equation of state $p_s = n_s k_B T_s$, the pressure gradients can be written in terms electron and ion number densities and temperatures

$$\begin{aligned} \nabla \left(p_e + \sum_i p_i \right) &= \sum_i \nabla (n_e k_B T_e + n_i k_B T_i) \\ &= \sum_i \left\{ n_i k_B \nabla (T_e + T_i) + [k_B (T_e + T_i) \nabla n_i] \right\} \\ &= 2k_B \sum_i [n_i \nabla T_p + T_p \nabla n_i] \end{aligned} \quad (2.19)$$

The *plasma temperature* $T_p = (T_e + T_i)/2$ in (2.19) is the mean of the electron and ion temperatures. After applying these simplifications, the combined electron and ion species momentum equation (2.14) reduces to:

$$\sum_i n_i m_i \frac{D_i \mathbf{u}_i}{Dt} + 2k_B \sum_i (n_i \nabla T_p + T_p \nabla n_i) - \sum_i n_i m_i \mathbf{G} = - \sum_i n_i m_i (\mathbf{u}_i - \mathbf{u}_N) \sum_t \nu_{it} \quad (2.20)$$

Recall that the target species index now ranges over the simulation major neutral species $t \in \{\text{O}, \text{O}_2, \text{N}_2\}$ unless specified otherwise. Rearranging the finite sum over the ion species index

$$\sum_i \left[n_i m_i \frac{D_i \mathbf{u}_i}{Dt} + 2k_B (n_i \nabla T_p + T_p \nabla n_i) - n_i m_i \mathbf{G} + n_i m_i (\mathbf{u}_i - \mathbf{u}_N) \sum_t \nu_{it} \right] = \mathbf{0} \quad (2.21)$$

and note that each term in the sum must balance individually so the momentum equation for each ion species is

$$n_i m_i \frac{D_i \mathbf{u}_i}{Dt} + z k_B (n_i \nabla T_p + T_p \nabla n_i) - n_i m_i \mathbf{G} = -n_i m_i (\mathbf{u}_i - \mathbf{u}_N) \sum_t \nu_{it} \quad (2.22)$$

Likewise, the momentum equation for the minor neutral constituents included in the model is

$$n_n m_n \frac{D_n \mathbf{u}_n}{Dt} + k_B (n_n \nabla T_n + T_n \nabla n_n) - n_n m_n \mathbf{G} = -n_n m_n (\mathbf{u}_n - \mathbf{u}_N) \sum_t \nu_{nt} \quad (2.23)$$

and is derived in a similar process leading to the ion equation above, however there is no Lorentz-force term to eliminate.

Vertical gradients in species density and temperature are typically much greater than the horizontal gradients during quiet ionosphere conditions so the horizontal components of equations (2.22) and (2.23) will be neglected. The one-dimensional minor neutral species vertical momentum equation is

$$n_n m_n \left(\frac{\partial}{\partial t} + w_n \frac{\partial}{\partial z} \right) w_n + k_B \left(T_n \frac{\partial n_n}{\partial z} + n_n \frac{\partial T_n}{\partial z} \right) - n_n m_n G = -n_n m_n (w_n - w_N) \sum_t \nu_{nt} \quad (2.24)$$

where $G = G(z)$ is the acceleration due to gravity in the vertical direction, assumed here to be equal to the value obtained from Newton's law of gravitation given the altitude of interest and the current estimated the mass off the earth. The hydrostatic approximation will be used to eliminate the convective derivative

$$k_B \left(T_n \frac{\partial n_n}{\partial z} + n_n \frac{\partial T_n}{\partial z} \right) - n_n m_n G = -n_n m_n (w_n - w_N) \sum_t \nu_{nt}. \quad (2.25)$$

For a vertical geomagnetic field, the one-dimensional ion species vertical momentum equation is identical in form to equation (2.25). For a magnetic field that is inclined at the dip angle I to the horizontal, the absence of external forces in the horizontal plane ensures that the Lorentz force will confine the ion motion along the magnetic field direction when the ion cyclotron frequency is significantly greater than the ion-neutral collision frequency, i.e. in the F region. The ion-neutral collision rate greatly exceeds the ion cyclotron frequency in the E region—e.g., see Figure 2.5—so no adjustment to the model equations is made for this because vertical diffusive transport is small

relative to the local production and loss rates. The projection of the vertical force terms in the left hand side of the ion species momentum equation (2.22) onto the field-parallel direction introduces a factor of $\sin I$ to each term; the field-aligned relative drift velocities that determine the resistive collision force to field-parallel ion motion introduces a factor of $1/\sin I$ to the right hand side of the equation. The one-dimensional hydrostatic approximation to the vertical ion momentum equation then has the form

$$\sin^2 I \left[z k_B \left(n_i \frac{\partial T_p}{\partial z} + T_p \frac{\partial n_i}{\partial z} \right) - n_i m_i G \right] = -n_i m_i (w_i - w_N) \sum_t v_{it} \quad (2.26)$$

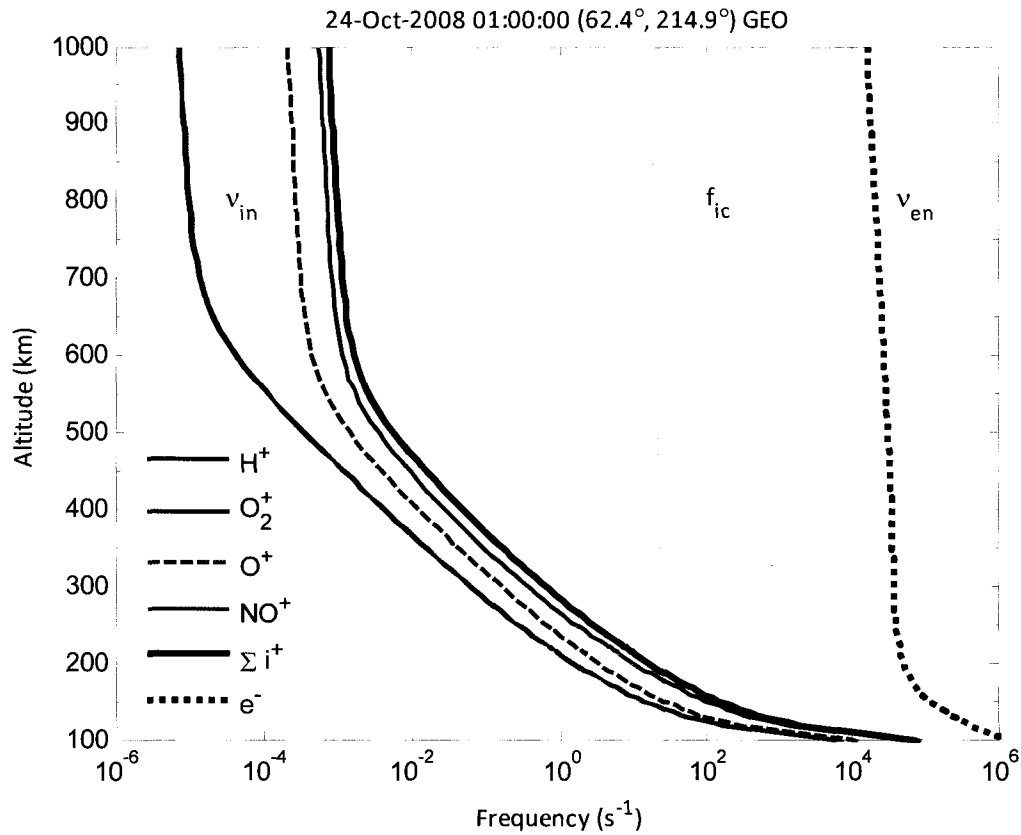


Figure 2.5 Altitude profiles of the total ion-neutral collision frequency ν_{in} calculated for selected ion species over Gakona, Alaska. The associated ion cyclotron frequencies $f_{ic} = \omega_{ic}/2\pi$ and the electron-neutral collision frequency ν_{en} are also plotted for reference.

Equation (2.26) and the corresponding equation for the minor neutrals relates the diffusive flow velocity that maintains the force balance between the pressure gradient force and the resistive collision force to the plasma temperature, the species collision frequencies, and the ion species number density.

In order to reduce the number of equations that must be solved and to also put the equations in a form required by the numerical solver described in section 3.1, the ion and minor-neutral momentum equations can be coupled to the continuity equation (2.6) by solving equations (2.25) and (2.26) for the species velocity or density flux and substituting the resulting expression into the flux term of the continuity equation. The resulting ion flow velocity is

$$w_i = w_N - \sin^2 I \left(m_i \sum_t v_{it} \right)^{-1} \left[2k_B \left(\frac{T_p}{n_i} \frac{\partial n_i}{\partial z} + \frac{\partial T_p}{\partial z} \right) - m_i G \right] \quad (2.27)$$

Substituting the flow velocity expressions into the 1-D continuity equation (2.6) yields the flux form of the species momentum equations for the ion and minor neutral species:

$$\begin{aligned} \frac{\partial n_i}{\partial t} &= -\frac{\partial}{\partial z} \left\{ n_i w_N - \sin^2 I \left(m_i \sum_t v_{it} \right)^{-1} \left[2k_B \left(T_p \frac{\partial n_i}{\partial z} + n_i \frac{\partial T_p}{\partial z} \right) - n_i m_i G \right] \right\} + \frac{\delta n_i}{\delta t} \\ \frac{\partial n_n}{\partial t} &= -\frac{\partial}{\partial z} \left\{ n_n w_N - \left(m_n \sum_t v_{nt} \right)^{-1} \left[k_B \left(T_n \frac{\partial n_n}{\partial z} + n_n \frac{\partial T_n}{\partial z} \right) - n_n m_n G \right] \right\} + \frac{\delta n_n}{\delta t} \end{aligned} \quad (2.28)$$

The transport term of equation (2.28) can be written in terms of the diffusion coefficient and scale height of each species

$$\begin{aligned} \frac{\partial n_i}{\partial t} &= -\frac{\partial}{\partial z} \left[n_i w_N - (\sin^2 I) D_i \left(\frac{\partial n_i}{\partial z} + \frac{n_i}{T_p} \frac{\partial T_p}{\partial z} - \frac{n_i}{H_i} \right) \right] + \frac{\delta n_i}{\delta t} \\ \frac{\partial n_n}{\partial t} &= -\frac{\partial}{\partial z} \left[n_n w_N - D_n \left(\frac{\partial n_n}{\partial z} + \frac{n_n}{T_n} \frac{\partial T_n}{\partial z} - \frac{n_n}{H_n} \right) \right] + \frac{\delta n_n}{\delta t} \end{aligned} \quad (2.29)$$

where $D_i = 2kT_p / (m_i \sum_n v_{in})$ is the ion gas species ambipolar diffusion coefficient,

$D_n = kT_n / (m_n \sum_t v_{nt})$ is the neutral gas diffusion coefficient, $H_i = 2k_B T_p / (m_i G)$ is the ion gas species scale height, and $H_n = 2k_B T_n / (m_n G)$ is the neutral gas scale height.

Ion ambipolar diffusion coefficients used in the model equations were calculated from the ion-neutral collision frequencies referenced above. The diffusion coefficients for minor dynamic neutral species $\{N(^4S), N(^2D), N(^2P), NO\}$ drifting through a background major neutral gas $\{N_2, O, O_2\}$ were approximated with the parameterization used by *Strobel et al.* [1970]:

$$D_n = a_n \frac{\sqrt{T_n}}{\sum_t n_t}, \quad \begin{aligned} a_N &= 4.55 \times 10^{17} \text{ cm}^{-1} \text{ s}^{-1} \text{ K}^{-1} \\ a_{NO} &= 2.98 \times 10^{17} \text{ cm}^{-1} \text{ s}^{-1} \text{ K}^{-1} \end{aligned} \quad (2.30)$$

The index t of the summation term in equation (2.30) ranges over the driven major neutral species $\{N_2, O, O_2\}$, the neutral gas temperature T_n is expressed in units of degrees Kelvin, and the neutral number density is in cm^{-3} .

Above the turbopause, thermospheric turbulence is another mechanism of gas species transport that is included in the ionosphere model. Vertical flux of species number density due to turbulent transport exceeds the density flux due to molecular diffusion in the D and E regions. A technique for incorporating turbulent mixing transport into ionosphere and thermosphere model equations of the form (2.29) through the use of an eddy diffusion coefficient K was suggested [Colegrove *et al.*, 1965; Lettau, 1951]. A recent example of the use of an eddy diffusion coefficient in a self-consistent model is provided by Bailey *et al.* [2002]. An additional number density flux term is added to the flux term in the momentum-continuity equations that relates the species flux to the gradient in composition of the background gas

$$(n_s u_s)_{\text{diffusion}}^{\text{eddy}} = -K \left(\frac{\partial n_s}{\partial z} + \frac{n_s}{\langle H \rangle_n} + \frac{n_s}{T_n} \frac{\partial T_n}{\partial z} \right) \quad (2.31)$$

where $\langle H \rangle_n$ is the scale height averaged over the major neutral gas constituent species. The ionosphere momentum-continuity equations then take the following form

$$\begin{aligned} \frac{\partial n_i}{\partial t} &= -\frac{\partial}{\partial z} \left\{ n_i w_N - (\sin^2 I) \left[D_i \left(\frac{\partial n_i}{\partial z} + \frac{n_i}{T_p} \frac{\partial T_p}{\partial z} - \frac{n_i}{H_i} \right) - K \left(\frac{\partial n_i}{\partial z} + \frac{n_i}{\langle H \rangle_n} + \frac{n_i}{T_n} \frac{\partial T_n}{\partial z} \right) \right] \right\} + \frac{\delta n_i}{\delta t} \\ \frac{\partial n_n}{\partial t} &= -\frac{\partial}{\partial z} \left[n_n w_N - D_n \left(\frac{\partial n_n}{\partial z} + \frac{n_n}{T_n} \frac{\partial T_n}{\partial z} - \frac{n_n}{H_n} \right) - K \left(\frac{\partial n_n}{\partial z} + \frac{n_n}{\langle H \rangle_n} + \frac{n_n}{T_n} \frac{\partial T_n}{\partial z} \right) \right] + \frac{\delta n_n}{\delta t} \end{aligned} \quad (2.32)$$

The value of the eddy diffusion coefficient has been inferred from rocket experiments involving in-situ measurements [Allen *et al.*, 1981] or by observing the evolution of a tracer [Chamberlain and Hunten, 1989; Kellogg, 1964]. Reference atmospheres were used to fit an eddy diffusion coefficient parameter to model calculations [Colegrove *et al.*, 1965; Dickinson, 1984]. An altitude-dependent exponential parameterization of an eddy diffusion coefficient is given by Dickinson [1984]; a constant diffusion coefficient is estimated by Kellogg [1964] to be $K = 4 \times 10^7 \text{ cm}^2 \text{ s}^{-1}$, applicable over the altitude range of interest, while Colegrove *et al.* [1966] estimated the coefficient to be $K = 4.5 \times 10^6 \text{ cm}^2 \text{ s}^{-1}$. The SCIM calculations use the eddy diffusion constant provided by Kellogg

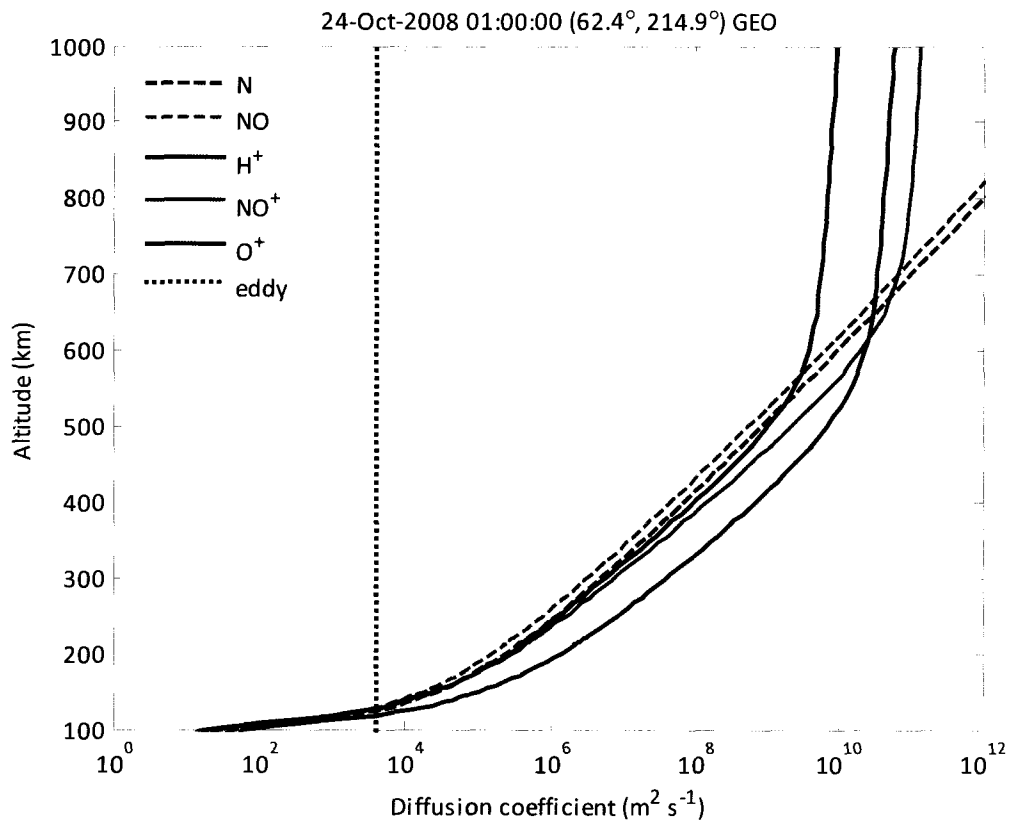


Figure 2.6 Altitude profiles of the diffusion and ambipolar diffusion coefficients for selected neutral and ionized gas species calculated over Gakona, Alaska. An altitude-independent eddy diffusion coefficient from Kellogg [1964] is plotted for comparison.

[1964]. Altitude profiles of selected diffusion coefficients are plotted in Figure 2.6 and note, in particular, that eddy diffusion is not significant relative to molecular diffusion at most heights in the model domain.

A useful measure of the time required for a perturbation in the equilibrium number density to disperse through diffusive processes is the *diffusion time*. The diffusion time is calculated through the Fickian law of diffusion, that relates the mean square distance $\langle \Delta z_s^2 \rangle$, or distance variance, that each particle will travel in time Δt to the diffusion coefficient [Kellogg, 1964]: $\langle \Delta z_s^2 \rangle = 2D_s \Delta t$.

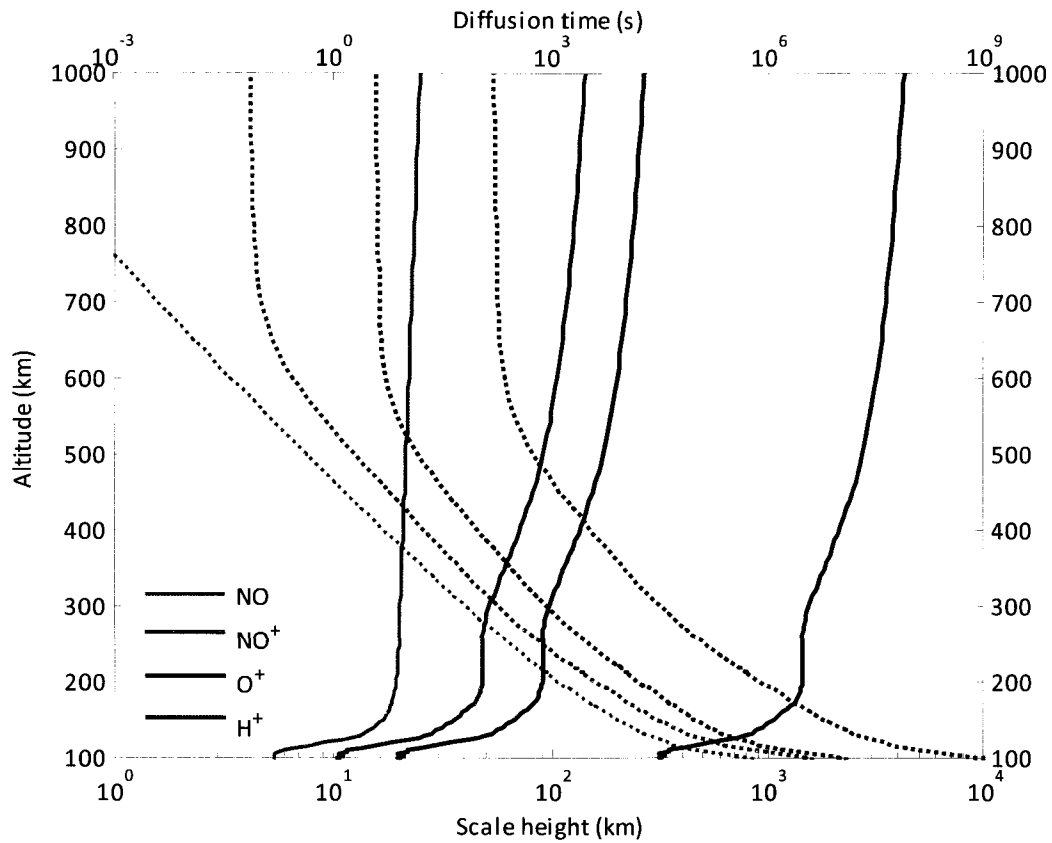


Figure 2.7 Altitude profiles of selected species scale heights (solid line) and diffusion times (dotted line) calculated over Gakona, Alaska at 01:00 UT (hh:mm) on 24 October 2008.

Setting the distance variance to the square of the species scale height $\langle \Delta z_s^2 \rangle = H_s^2$ and solving for Δt yields the species diffusion time

$$\tau_s = \frac{H_s^2}{2D_s} \quad (2.33)$$

Diffusion times and scale heights calculated with the ionosphere model are plotted in Figure 2.7. In particular, note that some diffusion times are many thousands of seconds near the lower simulation boundary at 100 km and yet others are less than a second near the upper simulation boundary at 1000 km altitude. This means the number density of a particular species is primarily determined by the balance between local production and loss rates at low altitudes; but at higher altitudes the effects of transport into or out of a region are more important than local production or loss since

the neutral gas number density decreases exponentially with increasing altitude. Consequently, when boundary conditions are specified for the coupled momentum and continuity equation in section 3.1.2 the number density of each species will be set at the lower boundary and the number density flux will be set at the upper boundary.

2.1.3 Energy equation

In this section, energy equations expressed in terms of ion and electron gas temperatures are derived from the energy equation (2.4) and cast into a parabolic form similar to the flux form of the species momentum equations above. First, the equation of state is used to write the species energy density in terms of number density and temperature thus transforming the energy equation to an equation for temperature. Then the continuity equation is used to eliminate the resulting number-density total derivative; several remaining terms are dropped as justified by physical arguments. The physics of plasma energy transport is again assumed to be aligned with the geomagnetic field so the three dimensional energy equations will be projected onto the vertical direction. Finally, profiles of relevant equation parameters are illustrated for particular ionosphere conditions.

Although an ion species may maintain a temperature independent of the temperatures of the other ion species, at least in principle, but in practice the ion temperatures do not differ much between species in the model domain of interest [Roble and Hastings, 1977]. Therefore the ion thermal energy is assumed to be uniform with respect to ion species and so the index i of the species energy will distinguish the energy of the composite ion gas rather than the energy of a particular ion species. Similarly, the neutral gas energy is identified by the index n . The label e refers to the electron gas as above. For now, the generic index s will be used until approximations specific to the ion or electron gas are applied. The neutral gas temperature is directly driven by the MSIS model so the neutral gas energy equation will not be considered here.

Recall that the energy equation (2.4) for the energy density of a species s , is

$$\frac{D_s}{Dt} \left(\frac{3}{2} p_s \right) + \frac{5}{2} p_s (\nabla \cdot \mathbf{u}_s) + \nabla \cdot \mathbf{q}_s = \frac{\delta E_s}{\delta t} \quad (2.34)$$

where the thermal stress tensor term has been dropped. Substituting the equation of state $p_s = n_s k_B T_s$ for the pressure term, applying the product rule for derivatives, and expanding the number density convective derivative leads to the following equation for temperature.

$$\frac{3}{2}k_B T_s \left(\frac{\partial n_s}{\partial t} + \mathbf{u}_s \cdot \nabla n_s \right) + \frac{3}{2}n_s k_B \frac{D_s T_s}{Dt} + \frac{5}{2}n_s k_B T_s (\nabla \cdot \mathbf{u}_s) + \nabla \cdot \mathbf{q}_s = \frac{\delta E_s}{\delta t} \quad (2.35)$$

Equation (2.35) can be simplified further by using the continuity equation (2.5) to substitute an expression for the partial time derivative of the species number density, thus cancelling the number density convection term and combining the velocity divergence terms:

$$\frac{3}{2}n_s k_B \frac{D_s T_s}{Dt} + n_s k_B T_s (\nabla \cdot \mathbf{u}_s) + \nabla \cdot \mathbf{q}_s = \frac{\delta E_s}{\delta t} - \frac{3}{2}k_B T_s \frac{\delta n_s}{\delta t} \quad (2.36)$$

Of the two collision terms collected on the right hand side of the equation above, the first term is the species s time-rate of volume heating or cooling resulting from collisions with other molecules; the second term accounts for the extensive increase or decrease in energy density due to the respective loss or production of s particles through collisions. Generally the density production and loss rates nearly balance in the E region, and divergence of particle transport drives the local volume density changes in the F region and topside ionosphere; thus the extensive time-rate of energy change due to particle production or loss will be dropped from equation (2.36). The advection term in the temperature derivative will also not be considered here, nor will the energy changes due to adiabatic expansion or compression; so the resulting vector equation for temperature is

$$\frac{3}{2}n_s k_B \frac{\partial T_s}{\partial t} + \nabla \cdot \mathbf{q}_s = \frac{\delta E_s}{\delta t}. \quad (2.37)$$

Temperature changes due adiabatic expansion and compression are important in the neutral atmosphere during vertical transport of an air parcel, but vertical transport in the quiet daytime ionosphere is primarily diffusive, so neglecting these effects is a reasonable approximation in this model.

Fourier's law of heat conduction law, in differential formulation, expresses the species heat flux vector \mathbf{q} as the product of a thermal conduction coefficient and the temperature gradient:

$\mathbf{q}_s = -\lambda_s \nabla T_s$. Therefore the simplified vector equation of temperature can be written without explicit dependence on the heat flux.

$$\frac{3}{2}n_s k_B \frac{\partial T_s}{\partial t} - \nabla \cdot (\lambda_s \nabla T_s) = \frac{\delta E_s}{\delta t} \quad (2.38)$$

The species thermal conduction coefficient λ_s is a tensor quantity; but where the electron or ion cyclotron frequencies are large—relative to the respective collision rates with the major neutral species—energy flux in the direction parallel to the imposed geomagnetic field is much greater for a given temperature gradient than the energy flux perpendicular to the field. In the E region where the collision rate with the major neutrals greatly exceeds the cyclotron frequencies, energy transport is small relative to local collisional heating and cooling. Thus energy transport is negligible in the E region where the field-aligned transport approximation breaks down. The one-dimensional vertical temperature equation is

$$\frac{3}{2} n_s k_B \frac{\partial T_s}{\partial t} - \sin I \frac{\partial}{\partial z} \left(\lambda_s \sin I \frac{\partial T_s}{\partial z} \right) = \frac{\delta E_s}{\delta t} \quad (2.39)$$

Factors of the dip angle sine are necessary because the transport physics is assumed to be parallel to the field and the temperature gradient is parallel to the vertical; i.e. if distance along the field (parallel to the energy flow) is measured by l , then $\partial/\partial l = \sin I (\partial/\partial z)$.

The ion species thermal conductivity coefficient λ_s for heat flux parallel to the geomagnetic field is approximately proportional to $n_s k_B T_s / v_{ii}$ [Schunk, 1975]. For the average ion gas considered in the temperature equation above, the thermal conduction coefficient parameterization used in the model equations is from Banks [1966c] and is a function of major ion densities, ion temperature, and electron density:

$$\lambda_i = 1.2 \times 10^4 \left(n_{O^+} + 2n_{He^+} + 4n_{H^+} \right) \frac{T_i^{5/2}}{n_e} \text{ eV cm}^{-1} \text{ s}^{-1} \text{ K}^{-1} \quad (2.40)$$

The units of the number density quantities in (2.40) are expressed in cm^{-3} . The parameterization of the thermal conduction coefficient of the electron gas used in the model is from Banks [1966b]

$$\lambda_e = \frac{7.7 \times 10^5 T_e^{5/2}}{1 + 3.22 \times 10^4 \frac{T_e^2}{n_e} \sum_n n_n \bar{Q}_n} \text{ eV cm}^{-1} \text{ s}^{-1} \text{ K}^{-1} \quad (2.41)$$

where the temperature in (2.41) is expressed in units of degrees Kelvin, the electron and neutral densities are expressed in units of cm^{-3} , the index n in the summation runs over the neutral

species $n \in \{O, O_2, N_2, H, He\}$, and the quantity \bar{Q}_n is the velocity-average momentum transfer cross section, parameterizations calculated by *Banks* [1966b]:

$$\begin{aligned}
 \bar{Q}_O &= 1.10 \times 10^{-16} + 7.92 \times 10^{-20} T_e \quad \text{cm}^2 \\
 \bar{Q}_{O_2} &= 2.20 \times 10^{-16} + 7.92 \times 10^{-18} T_e^{1/2} \quad \text{cm}^2 \\
 \bar{Q}_{N_2} &= 2.82 \times 10^{-17} T_e^{1/2} - 3.41 \times 10^{-21} T_e^{3/2} \quad \text{cm}^2 \\
 \bar{Q}_H &= 5.47 \times 10^{-15} - 7.45 \times 10^{-19} T_e \quad \text{cm}^2 \\
 \bar{Q}_{He} &= 5.69 \times 10^{-16} \quad \text{cm}^2
 \end{aligned} \tag{2.42}$$

The electron temperatures in (2.42) are expressed in degrees Kelvin. The electron thermal conductivity parameterization (2.41) is an approximation based on a simple mean free path argument; *Schunk and Walker* [1970c] compared this approximation to a more complete treatment by *Shkarofsky* [1961]. Altitude profiles of the thermal conductivity coefficient $\lambda_{\{e,i\}}$ and energy flux parallel to the geomagnetic field through the ion and electron gases

$\phi_{\{e,i\}} = n_{\{e,i\}} w_{\{e,i\}} = (\sin I) (\lambda_{\{e,i\}} \partial T_{\{e,i\}} / \partial z)$ as calculated in a SCIM simulation for quiet daytime conditions over Gakona, Alaska are plotted in Figure 2.8. Note that the energy flux is predominantly downward except in the 50 km layer above the altitude of maximum photoionization: heat from the magnetosphere conducts downward through the thermal electrons, from the top simulation near 1000 km altitude, to the critical layer at about 250 km altitude that is also heated from below by collisions with upward-directed photoelectrons.

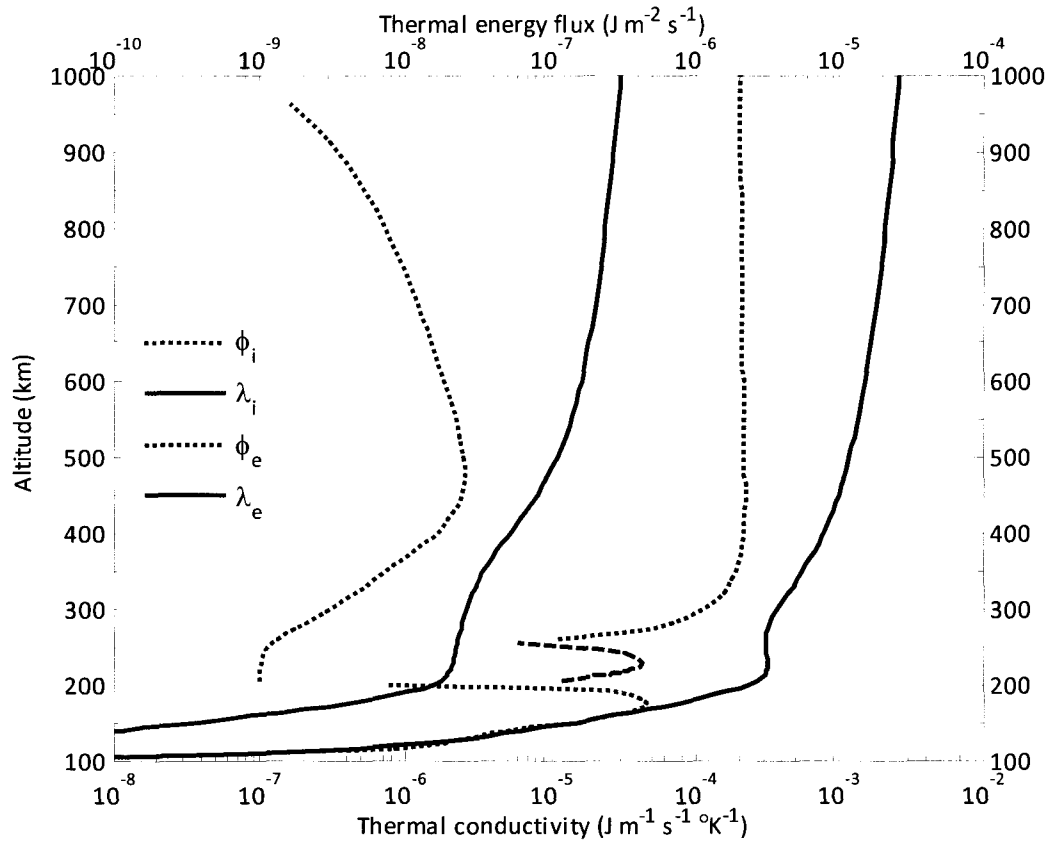


Figure 2.8 Altitude profiles of the thermal conductivity λ (solid lines), downward thermal energy flux (dotted), and upward field-aligned thermal energy flux ϕ (dashed lines) for the ion (blue) and electron (red) gas modeled over Gakona, Alaska at 01:00 UT (hh:mm) on 24 October 2008.

The final term in the temperature equation is the power per volume added or removed from the species due to collisions with other species particles or due to collisions with a small population of highly energetic (non-Maxwellian velocity distribution) electrons. For the ion gas, significant collisional heat sources include elastic collisions with hot (thermal) electrons and Joule heating due to motion relative to the neutral species induced by perpendicular electric fields of magnetospheric origin. Significant ion gas cooling sources include elastic and inelastic collisions with the neutral gas species.

$$\frac{\delta E_i}{\delta t} = \left. \frac{\delta E_i}{\delta t} \right|_{\text{Joule heating}} + \left. \frac{\delta E_i}{\delta t} \right|_{\text{elastic electron collisions}} + \left. \frac{\delta E_i}{\delta t} \right|_{\text{elastic neutral collisions}} \quad (2.43)$$

For the electron gas, significant collisional heat sources considered in the model include collisions with energetic photoelectrons and precipitating auroral electrons. Electron heating due to Joule heating from magnetospheric-imposed perpendicular electric fields is smaller—by a factor of the electron-ion mass ratio—than the corresponding ion Joule heating rate [Schunk and Nagy, 1978] so it is not included in the SCIM calculations. Significant electron cooling sources include elastic collisions with ion and neutral species and inelastic collisions with the neutral species.

$$\frac{\delta E_e}{\delta t} = \frac{\delta E_e}{\delta t} \Big|_{\substack{\text{supra-} \\ \text{thermal} \\ \text{collisions}}} + \frac{\delta E_e}{\delta t} \Big|_{\substack{\text{elastic} \\ \text{ion} \\ \text{collisions}}} + \frac{\delta E_e}{\delta t} \Big|_{\substack{\text{elastic} \\ \text{neutral} \\ \text{collisions}}} + \frac{\delta E_e}{\delta t} \Big|_{\substack{\text{inelastic} \\ \text{neutral} \\ \text{collisions}}} \quad (2.44)$$

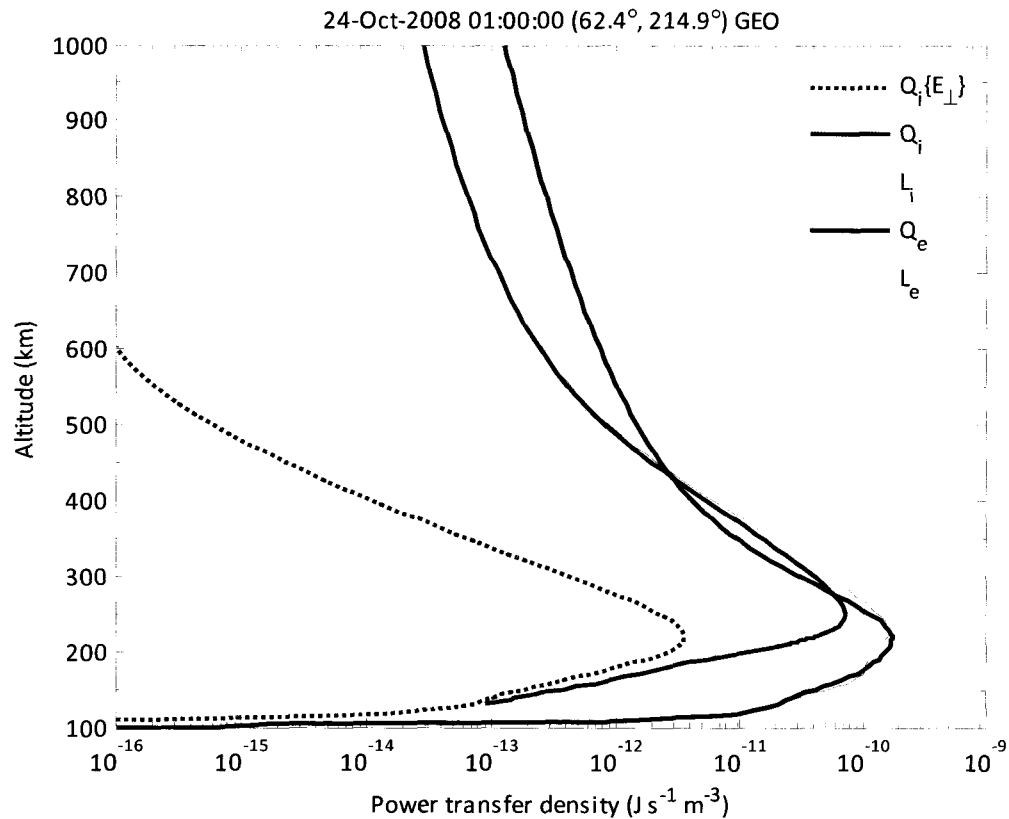


Figure 2.9 Altitude profiles of power transfer rates to and from the ion (blue) and electron (red) gases over Gakona, Alaska at 01:00 UT (hh:mm) on 24 October 2008. Total power transfer into (thick) and from (thin) each gas includes the sums of energy transfer rates via elastic and inelastic collisions between the neutral, ion, and electron gases. An imposed 1 mV/m perpendicular electric field results in partial power density transfer to the ion gas (dotted) via Pedersen-current Joule dissipation.

Figure 2.9 contains altitude profiles of the suprathermal electron collision heating rate and the total electron gas cooling rate in equation (2.44) as calculated by the self-consistent ionosphere model.

2.1.3.1 Electron gas collisional energy transfer rates

Collisional heating and cooling of the ionospheric electron gas is described in the model through the collision term on the right hand side of equation (2.38). Two sources of electron heating are considered: heating through elastic collisions with the relatively small population of energetic photoelectrons or precipitating auroral electrons, and (discussed in Chapter 4 and Chapter 5) heating by interaction with powerful HF electromagnetic waves. A complete derivation of the

thermal electron volume heating rate due to collisions with non-thermal electrons is given by Hoegy [1983] but the expression used here is a simplified heating rate expression [Stamnes and Rees, 1983; Swartz *et al.*, 1971]:

$$\left. \frac{\delta E_e}{\delta t} \right|_{\text{supra-thermal heating}} = n_e(z) \int_{E_c}^{\infty} \Phi(E_e, z) L_e(E_e, z) dE_e \quad (2.45)$$

where Φ_e is the total suprathermal electron intensity in units of $\text{cm}^{-2} \text{s}^{-1} \text{sr}^{-1} \text{eV}^{-1}$, the suprathermal electron stopping power $L_e = -n_e^{-1} dE/dl$ is the energy loss per unit path length in the medium, and the lower limit of integration E_c is the crossover or transition energy where the electron energy distribution departs from a Maxwellian distribution. The electron flux distribution is provided as a function of altitude by the GLOW model via solution of the Banks and Nagy two-stream electron transport code [Banks and Nagy, 1970; Banks and Chappell, 1974; Nagy and Banks, 1970; Solomon *et al.*, 1988]. The parameterization of the suprathermal electron stopping power used in the model equations has been calculated by Swartz *et al.* [1971]

$$L_e = \frac{3.37 \times 10^{-12}}{E^{0.94} n_e^{0.03}} \left(\frac{E - E_c}{E - 0.53 E_c} \right)^{2.36} \text{ eV cm}^2 \quad (2.46)$$

The electron number density in equation (2.46) is expressed in cm^{-3} , the suprathermal energy term E is expressed in units of eV, and $E_c = (3/2)k_B T_e = 8.618 \times 10^{-5} T_e$ is the temperature of the thermal electrons expressed in electron volts.

Sinks of electron thermal energy include elastic collisions with the neutral and ionized species particles. Electron-neutral elastic collision energy exchange rates for each of the major neutral species in the model domain are tabulated by Schunk and Nagy [1978] as parameterizations of the neutral species number density, neutral gas temperature, and electron gas temperature.

$$\begin{aligned} L(e, \text{N}_2) &= 1.77 \times 10^{-19} n_e n_{\text{N}_2} \left(1 - 1.21 \times 10^{-4} T_e \right) T_e (T_e - T_n) \text{ eV cm}^{-3} \text{ s}^{-1} \\ L(e, \text{O}_2) &= 1.21 \times 10^{-16} n_e n_{\text{O}_2} \left(1 + 3.6 \times 10^{-2} T_e^{1/2} \right) T_e^{1/2} (T_e - T_n) \text{ eV cm}^{-3} \text{ s}^{-1} \\ L(e, \text{O}) &= 7.9 \times 10^{-19} n_e n_{\text{O}} \left(1 + 5.7 \times 10^{-4} T_e \right) T_e^{1/2} (T_e - T_n) \text{ eV cm}^{-3} \text{ s}^{-1} \\ L(e, \text{He}) &= 2.46 \times 10^{-17} n_e n_{\text{He}} T_e^{1/2} (T_e - T_n) \text{ eV cm}^{-3} \text{ s}^{-1} \\ L(e, \text{H}) &= 9.63 \times 10^{-16} n_e n_{\text{H}} \left(1 - 1.35 \times 10^{-4} T_e \right) T_e^{1/2} (T_e - T_n) \text{ eV cm}^{-3} \text{ s}^{-1} \end{aligned} \quad (2.47)$$

The species number densities above are expressed in units of cm^{-3} and the temperatures are in degrees Kelvin. Electron-ion elastic coulomb collisions typically transfer energy from the electron gas to the ion gas because the electron temperature is usually greater than or equal to the ion temperature. Consequently, the power density transfer from electrons to ions is usually a sink of electron thermal energy and a source of ion energy. Energy density transfer rates parameterized by electron gas temperature, ion temperature, and neutral species number density are expressed as electron energy density loss rates from *Schunk and Nagy* [1978]:

$$\begin{aligned}
 -\left. \frac{\delta E_i}{\delta t} \right|_{\text{elastic electron collisions}} &= \left. \frac{\delta E_e}{\delta t} \right|_{\text{elastic ion collisions}} \\
 &= 3.2 \times 10^{-8} n_e \frac{(T_e - T_i)}{T_e^{3/2}} (\ln \Lambda) (n_{\text{O}^+} + 4n_{\text{He}^+} + 16n_{\text{H}^+} + 0.5n_{\text{O}_2} + 0.53n_{\text{NO}^+}) \text{ eV cm}^{-3} \text{ s}^{-1}
 \end{aligned} \tag{2.48}$$

As with the earlier parameterization of collisional energy transfer rates, the species number density must be expressed in units of cm^{-3} ; the electron and composite ion gas temperatures are in degrees Kelvin. An exact expression for the so-called Coulomb logarithm $\ln \Lambda$ was derived by *Itikawa* [1971] but its approximate form is used in the model calculations, as provided by *Prölss* [2004]: $\ln \Lambda = \ln(12\pi\epsilon_0 k_B T_e l_D / e^2)$ where $l_D = \sqrt{\epsilon_0 k_B T_e / n_e}$ is the Debye length, ϵ_0 is the vacuum permittivity, and e is the electron charge. Near the critical layer of a typical daytime polar ionosphere, $\ln \Lambda \approx 14$.

Inelastic collisions with the major neutral species are a significant sink of electron thermal energy. Parameterizations for the following inelastic collision energy transfer rates from the relatively hot electron gas to the cool neutral gas are collected by *Schunk and Nagy* [1978] and are included in the model energy equation calculations: rotational excitation of molecular nitrogen and oxygen, vibrational excitation of molecular nitrogen and oxygen, and impact excitation of the electronic excited states of the major neutral species. The parameterization used in the model to describe the energy transfer rate due to impact excitation of the fine structure energy levels of atomic oxygen was taken from *Dalgarno and Degges* [1968]. The total electron thermal energy loss rate due to inelastic collisions with neutral species is the sum of the loss rates due to inelastic collisions with each neutral species:

$$\left. \frac{\delta E_e}{\delta t} \right|_{\text{inelastic neutral collisions}} = \left. \frac{\delta E_e}{\delta t} \right|_{\text{N}_2 \text{ and O}_2 \text{ rotational excitation}} + \left. \frac{\delta E_e}{\delta t} \right|_{\text{N}_2 \text{ and O}_2 \text{ vibrational excitation}} + \left. \frac{\delta E_e}{\delta t} \right|_{\text{O fine-structure excitation}} + \left. \frac{\delta E_e}{\delta t} \right|_{\text{N and O electronic excitation}} \quad (2.49)$$

The corresponding analytic expressions of the inelastic collision energy transfer rates, with the exception of the expression for the fine structure excitation of atomic oxygen from *Dalgarno and Degges* [1968], are reproduced here from *Schunk and Nagy* [1978]:

$$\begin{aligned} \left. \frac{\delta E_e}{\delta t} \right|_{\text{N}_2 \text{ rotational excitation}} &= 2.9 \times 10^{-14} n_e n_{\text{N}_2} (T_e - T_n) / T_e^{1/2} \text{ eV cm}^{-3} \text{ s}^{-1} \\ \left. \frac{\delta E_e}{\delta t} \right|_{\text{O}_2 \text{ rotational excitation}} &= 6.9 \times 10^{-14} n_e n_{\text{O}_2} (T_e - T_n) / T_e^{1/2} \text{ eV cm}^{-3} \text{ s}^{-1} \\ \left. \frac{\delta E_e}{\delta t} \right|_{\text{N}_2 \text{ vibrational excitation}} &= 2.99 \times 10^{-12} n_e n_{\text{N}_2} \exp\left(f \frac{T_e - 2000}{2000 T_e}\right) \left[\exp\left(-g \frac{T_e - T_n}{T_e T_n}\right) - 1 \right] \text{ eV cm}^{-3} \text{ s}^{-1} \\ \left. \frac{\delta E_e}{\delta t} \right|_{\text{O}_2 \text{ vibrational excitation}} &= 5.196 \times 10^{-13} n_e n_{\text{O}_2} \exp\left(h \frac{T_e - 700}{700 T_e}\right) \left[\exp\left(-2770 \frac{T_e - T_n}{T_e T_n}\right) - 1 \right] \text{ eV cm}^{-3} \text{ s}^{-1} \\ \left. \frac{\delta E_e}{\delta t} \right|_{\text{O fine-structure excitation}} &= 3.4 \times 10^{-12} n_e n_{\text{O}} (T_e - T_n) (1 - 7 \times 10^{-5} T_e) T_n^{-1} \text{ eV cm}^{-3} \text{ s}^{-1} \\ \left. \frac{\delta E_e}{\delta t} \right|_{\text{O-O}(\text{D}) \text{ electronic excitation}} &= 1.57 \times 10^{-12} n_e n_{\text{O}} \exp\left(d \frac{T_e - 3000}{3000 T_e}\right) \left[\exp\left(-22,713 \frac{T_e - T_n}{T_e T_n}\right) \right] \text{ eV cm}^{-3} \text{ s}^{-1} \\ \left. \frac{\delta E_e}{\delta t} \right|_{\text{N-N}(\text{D}) \text{ electronic excitation}} &= 3.4 \times 10^{-18} n_e n_{\text{N}} \Omega \exp(-2.74 \times 10^4 / T_e) / T_e^{1/2} \text{ eV cm}^{-3} \text{ s}^{-1} \end{aligned} \quad (2.50)$$

where

$$\begin{aligned} f &= 1.06 \times 10^4 + 7.51 \times 10^3 \tanh\left[1.10 \times 10^{-3} (T_e - 1800)\right] \\ g &= 3300 + 1233(T_e - 1000) - 2.056 \times 10^{-4} (T_e - 1000)(T_e - 4000) \\ h &= 3300 - 839 \sin\left[1.91 \times 10^{-4} (T_e - 2700)\right] \\ d &= 2.4 \times 10^4 + 0.3(T_e - 1500) - 1.947 \times 10^{-5} (T_e - 1500)(T_e - 4000) \\ \Omega &= T_e / (9190 + 0.45 T_e) \end{aligned} \quad (2.51)$$

Species number densities are expressed in units of cm^{-3} and temperatures are expressed in degrees Kelvin. Electron energy transfer rate parameterizations based on more recent data and more thorough calculations are available for inelastic electron collisions with major neutral species: vibrational excitation of molecular nitrogen [*Campbell et al.*, 2004; *Pavlov*, 1998b], vibrational

excitation of molecular oxygen [Pavlov, 1998a], and the fine structure excitation of atomic oxygen [Bell et al., 1998; Pavlov and Berrington, 1999]. These updated parameterizations provide alternative electron cooling rates that can be included in future model calculations.

2.1.3.2 Ion gas collisional energy transfer rates

While the principle heat source for the electron gas is collisions with energetic photoelectrons or precipitating auroral electrons, the primary ion gas heat sources include elastic collisions with the relatively hot thermal electrons and Joule heating from ion currents driven through the neutral gas. The currents result from imposed electric fields of magnetospheric origin. Ion gas collisional heat sinks include elastic collisions with the relatively cool neutral gas. As with the electron gas, collisional sources and sinks of ion gas energy are described through the collisional term on the right hand side of equation (2.38).

By the conservation of energy principle, the composite ion gas volume heating rate resulting from elastic collisions with hot thermal electrons is equal in magnitude to the electron volume cooling rate due to elastic collisions with the ions.

$$\left. \frac{\delta E_i}{\delta t} \right|_{\text{elastic electron collisions}} = - \left. \frac{\delta E_e}{\delta t} \right|_{\text{elastic ion collisions}} \quad (2.52)$$

Therefore the energy transfer rate parameterization given by Schunk and Nagy [1978], listed here as equation (2.48), can be used in the model equations as a source of ion heating. Note that a somewhat simpler expression for ion heating due to elastic collisions with electrons for a three-ion gas composed of $\{O^+, H^+, He^+\}$ ions is given by Banks [1967], but this simplified expression is not used in the SCIM calculations.

An electric field oriented perpendicular to the geomagnetic field is often present in the polar ionosphere and is maintained by solar wind convection across open magnetic field lines in the polar cap. The charge separation induced by the component of the solar wind force perpendicular to \mathbf{B}_0 creates an electric potential across the polar cap with a magnitude of tens of thousands of volts. This electric potential difference is not easily relaxed by charged particle drifts since the $\mathbf{E} \times \mathbf{B}_0$ drift direction is perpendicular to the applied electric field. As a consequence, the magnitude of the field does not vary significantly with altitude at F-region heights. The resulting

plasma convection in the polar cap is regularly observed by radar (e.g., see *Chisham et al.* [2007]) and has been reproduced with 3-D general circulation ionosphere models [*Kulchitsky et al.*, 2005; *Maurits*, 1996; *Roble and Ridley*, 1994; *Roble et al.*, 1988].

The perpendicular electric field parameter included in the self-consistent ionosphere model is used to simulate the effects of Joule, or resistive, heating of a slab-symmetric conducting ion gas moving through a ground-stationary neutral gas at a horizontal drift velocity $\mathbf{u}_i = \mathbf{E} \times \mathbf{B}_o / B_o^2$ relative to the earth co-rotating frame. Considering only the electric field component that is perpendicular to the geomagnetic field, and projecting the resulting ion drift velocity to the horizontal (tangent) plane, the convective drift velocity can be expressed simply as $u_{i\perp} = |\mathbf{E}_\perp| / |\mathbf{B}_o|$. Thus the perpendicular electric field can be inferred from ion drift velocity measured by ionosonde, radar, or satellite. Neglecting dynamo currents induced by horizontal neutral winds, a reasonable assumption that is appropriate for quiet geomagnetic conditions in the daytime polar ionosphere, the ion volume heating rate can be approximated by writing Ohm's law for the Pedersen conductivity and the perpendicular electric field [*Prölss*, 2004]

$$\left. \frac{\delta E_i}{\delta t} \right|_{\text{Joule heating}} \approx \sigma_p \mathbf{E}_\perp^2 \quad (2.53)$$

A simplified expression for the Pedersen conductivity σ_p that neglects the electron contribution and hence can be expressed as a sum of the individual ion species conductivities is provided by *Schunk and Nagy* [2000]

$$\sigma_p = \sum_i \sigma_i \frac{\nu_i^2}{\nu_i^2 + \omega_{ci}^2} \quad (2.54)$$

where the index i runs over the ion species, $\sigma_i = n_i e_i^2 / (m_i \nu_i)$ is the ion conductivity, $\nu_i = \sum_n \nu_{in}$ is the sum of over the collision frequency of ion species i with each major neutral species, and $\omega_{ci} = q_i |\mathbf{B}_o| / m_i$ is the ion cyclotron frequency. Plasma heating due to Joule dissipation from electron Hall currents is not considered in the polar ionosphere model.

Elastic collisions between the ions and the relatively cool neutral gas particles are the primary mechanism of energy transfer from the relatively warm ion gas to the cool neutral gas. As the number density of the neutral gas declines rapidly with height, the ion volume cooling rate

decreases with increasing altitude. Similar to the transfer of energy from the electron gas to the ion gas via elastic collisions, elastic collisions between the ions and the neutrals constitute a source of neutral gas energy that is equal in magnitude to the respective sink of ion gas energy. However, even at 1000 km altitude where the neutral gas density is at a minimum in the simulation domain, the neutral gas number density is still at least an order of magnitude larger than the number density of the ion gas—see e.g., Figure 2.1 and Figure 2.4—so warming of the neutral gas via elastic ion-neutral collisions is not considered in this model. (Neutral gas temperatures are given by the MSIS model.) Parameterizations of ion-neutral energy transfer rates due to elastic collisions between the various ion and neutral gas species were calculated by *Banks* [1966c]. The particular energy transfer rate parameterizations included in the model are:

$$\begin{aligned}
 \left. \frac{\delta E_i}{\delta t} \right|_{\text{O}^+ \text{ elastic collisions}} &= n_{\text{O}^+} (T_i - T_n) \cdot 10^{-14} \cdot \left(6.6n_{\text{N}_2} + 5.8n_{\text{O}_2} + 2.8n_{\text{He}} + 0.21n_{\text{O}}\sqrt{T_i + T_n} + 0.36n_{\text{H}}\sqrt{T_i} \right) \frac{\text{eV}}{\text{cm}^3 \text{ s}^1} \\
 \left. \frac{\delta E_i}{\delta t} \right|_{\text{H}^+ \text{ elastic collisions}} &= n_{\text{H}^+} (T_i - T_n) \cdot 10^{-14} \cdot \left(3.1n_{\text{N}_2} + 2.8n_{\text{O}_2} + 5.5n_{\text{He}} + 0.4n_{\text{O}}\sqrt{T_n} + 1.4n_{\text{H}}\sqrt{T_i + T_n} \right) \frac{\text{eV}}{\text{cm}^3 \text{ s}^1} \\
 \left. \frac{\delta E_i}{\delta t} \right|_{\text{NO}^+ \text{ elastic collisions}} &= n_{\text{NO}^+} (T_i - T_n) \cdot 10^{-14} \cdot \left(5.91n_{\text{N}_2} + 5.45n_{\text{O}_2} + 4.5n_{\text{O}} \right) \frac{\text{eV}}{\text{cm}^3 \text{ s}^1} \\
 \left. \frac{\delta E_i}{\delta t} \right|_{\text{O}_2^+ \text{ elastic collisions}} &= n_{\text{O}_2^+} (T_i - T_n) \cdot 10^{-14} \cdot \left(5.807n_{\text{N}_2} + 0.14n_{\text{O}_2}\sqrt{T_i + T_n} + 4.358n_{\text{O}} \right) \frac{\text{eV}}{\text{cm}^3 \text{ s}^1}
 \end{aligned} \tag{2.55}$$

The number densities in equation (2.55) are expressed in cm^{-3} , the temperatures are expressed in degrees Kelvin, and the total ion gas volume cooling rate due to elastic collisions is the sum of the individual ion cooling rates. An altitude profile of the ion gas Joule heating rate due to an imposed perpendicular electric field of magnitude 1 mV/m calculated in the self-consistent model of the daytime ionosphere is plotted in Figure 2.9 with the total ion gas heating rate that includes the warming effect of elastic collisions with thermal electrons.

Chapter 3. Solution of model equations

This chapter describes the spatial discretization method used by SCIM to numerically solve the system of model equations discussed in Chapter 2. The model equations are parabolic *partial differential equations* (PDEs) in one space variable. A numerical routine called *pdepe*, used for solving systems of parabolic and elliptic equations, is used by the self-consistent ionosphere model (SCIM) to solve the model equations and will be described in this chapter. Model equations described in Chapter 2 are rewritten in section 3.1.1 to conform to the *pdepe* numerical solver prerequisite requirements. Equation terms driven by empirical model software packages, such as the time-dependent altitude profiles of the neutral atmosphere species number densities, are also noted. Finally, the algorithmic structure of the SCIM computer program, including the simulation initialization procedure and the data structures of record are described in section 3.1.3. Numerical solutions of the model equation are validated in section 3.2 by comparing results derivable from electron density profiles calculated during a simulated diurnal ionosphere cycle with measurements and models of the ionosphere over interior Alaska. The numerical solution is found to be in agreement with observational measurements and an empirical ionosphere model.

3.1 Numerical solution

The coupled momentum (2.28) and energy (2.39) equations presented in sections 2.1.2 and 2.1.3, respectively, describe diffusive transport phenomena similar to those found in a variety of geophysical systems. In the self-consistent ionosphere model (SCIM) equations, the gas species number density and temperature variables are coupled through the collisional source terms, the diffusion coefficients, and the thermal conduction coefficients whose parameterizations result in a system of nonlinear partial differential equations. This system of equations is difficult to solve with standard numerical methods without introducing considerable software engineering complexity or neglecting important physics. SCIM solves the coupled momentum and energy equations with a flexible, accurate, and established numerical method that can be applied to a wider class of partial differential equations than other comparable methods.

For example, the finite-difference Crank-Nicolson scheme [Crank and Nicolson, 1996] can be used to solve the nonlinear system of model equations, first by replacing the coefficients that cause the system of equations to be classified as “nonlinear” with their Taylor expansions. However, such an approach can be problematic because the linearization procedure may discard important physics,

even if an impractically small simulation time step is used. Also, a fine-grained spatial grid must be used because diffusive transport dominates the solution behavior near the top of the simulation boundary while the balance of collisional sources and sinks dominates the solution behavior near the bottom simulation boundary. Furthermore, the transition altitude between the two solution behaviors varies with the ionospheric conditions that are attempted to be modeled. This transition height for the continuity-momentum equations may not even correspond to the transition-height for the continuity-energy equations. Consequently, attempts to split the simulation spatial domain according to solution behavior in order to reduce the number of spatial grid points, similar to the method of solution used by *Min* [1993], leads to somewhat arbitrary decisions of a “transition boundary altitude” based on solution behavior that must be inferred a priori.

An alternate method adapted by *Hastings and Roble* [1977] to solve a parabolic system of equations is based on Newton’s method and is well-suited for solving stiff equations. A differential equation is usually regarded as *stiff* if standard numerical solution methods will converge to rapidly varying solutions near the actual solution unless an extremely small time step is specified. They used this so-called *automatic technique of solving coupled vector systems of nonlinear equations* to solve coupled electron and ion-gas energy equations of the form

$$\frac{\partial}{\partial t} \begin{pmatrix} T_e \\ T_i \end{pmatrix} = \frac{\partial}{\partial s} \begin{pmatrix} \lambda_e & 0 \\ 0 & \lambda_i \end{pmatrix} \frac{\partial}{\partial s} \begin{pmatrix} T_e \\ T_i \end{pmatrix} + \begin{pmatrix} \delta T_e / \delta t \\ \delta T_i / \delta t \end{pmatrix} \quad (3.1)$$

in one spatial dimension parallel to the geomagnetic field. Both the thermal conduction coefficients and the net heating and cooling rates are functions of temperature and the distance s along the field line. The same numerical technique was used by *Roble and Hastings* [1977] to solve the system of electron and ion-species energy equations, coupled with a system of momentum equations similar to those derived in Chapter 2, for an ionosphere consisting of H^+ , He^+ , N^+ , and O^+ ions. Finally, *Roble* [1975] used this numerical method to solve the electron and ion-gas energy equations coupled with the ion species continuity-momentum equations for an ionosphere composed of NO^+ , O_2^+ , and O^+ ions.

Another technique for solving stiff parabolic system of equations, called *the method of lines* by *Madsen* [1975], was mentioned briefly by *Hastings and Roble* [1977] but not explored further. The method of lines refers to a method of reducing a system of partial differential equations into a system of coupled ordinary differential equations (ODEs) through spatial discretization and then

integrating the system of ODE's with a (possibly stiff) ODE numerical solver. The first general purpose software package that leverages modern ODE theory, e.g. that summarized by *Gear* [1971], to numerically solve an entire class of PDEs through the method of lines is provided by *Sincovec and Madsen* [1975]. Advantages of using this method-of-lines technique to solve parabolic and elliptic systems of equations are that it is accurate, robust, and stable. Most important to ionosphere modeling tasks is that the method of lines to solves systems of equations of the form

$$\frac{\partial n}{\partial t} = \frac{\partial}{\partial z} \left(D(n) \frac{\partial n}{\partial z} \right) + q(n) \quad (3.2)$$

after defining appropriate linear or nonlinear boundary equations. Implementing a numerical routine to solve system (3.2) that uses standard finite-difference techniques and can accommodate arbitrary and non-uniform spatial grids with an adaptive time-step is difficult and time consuming, but not necessary with the method of lines approach. Both implementations of the method of lines technique, referenced below, provide these features.

The SCIM software uses the method of lines to solve the full system of coupled electron and ion gas energy equations, and the ion and minor neutral species momentum equations. Specifically, the Skeel-Berzins method of spatial discretization was used to convert the coupled set of parabolic model equations into a system of coupled ODEs [*Skeel and Berzins*, 1990], and the system of ODEs was numerically integrated with an implementation of a backward difference formula from the Klopfenstein-Shampine family of numerical differentiation formulas [*Lawrence et al.*, 1999; *Shampine and Reichelt*, 1997]. Both the spatial discretization and time integration routines were performed with the software package *pdepe* provided in the Matlab problem-solving environment [*Mathworks*, 2007]. Alternate implementations of the method-of-lines technique via the Skeel-Berzins spatial discretization are available in the Do3P (parabolic equations) section of the NAG Fortran library.

In essence, the SCIM software repeatedly calls the *pdepe* function provided in Matlab to solve the following system of model equations that are expressed as an initial-boundary value problem of a system of one-dimensional PDEs that contain parameters driven by the output of empirical models and user-defined routines:

$$n_s(z, t_1) = \tilde{n}_s(z) \quad (3.3)$$

$$T_{\hat{s}}(z, t_1) = \tilde{T}_{\hat{s}}(z) \quad (3.4)$$

$$p(z, t; n_s, T_{\hat{s}}) + q(z, t) f\left(z, t; n_s, T_{\hat{s}}, \frac{\partial n_s}{\partial z}, \frac{\partial T_{\hat{s}}}{\partial z}\right) = 0 \quad (3.5)$$

$$c\left(z, t; n_s, T_{\hat{s}}, \frac{\partial n_s}{\partial z}, \frac{\partial T_{\hat{s}}}{\partial z}\right) \frac{\partial}{\partial t} \begin{pmatrix} n_s \\ T_{\hat{s}} \end{pmatrix} = z^{-m} \frac{\partial}{\partial z} \left[z^m f\left(z, t; n_s, T_{\hat{s}}, \frac{\partial n_s}{\partial z}, \frac{\partial T_{\hat{s}}}{\partial z}\right) \right] + s\left(z, t; n_s, T_{\hat{s}}, \frac{\partial n_s}{\partial z}, \frac{\partial T_{\hat{s}}}{\partial z}\right) \quad (3.6)$$

Equations (3.3) and (3.4) are initial conditions, equation (3.5) specifies the boundary conditions, and equation (3.6) is the partial differential vector equation. The index s ranges over the ion and neutral species referenced in the momentum equations and the index \hat{s} indicates the electron, ion, or neutral gas referenced in the energy equations. The index N is the total number of spatial grid points used to define a solution to the system of equations at some particular time. Solution variables $n_s(z, t)$ and $T_{\hat{s}}(z, t)$ are column vectors of size $|s|$ and $|\hat{s}|$, respectively, whose elements are functions of the temporal-spatial coordinates. The index $m \in \{0, 1, 2\}$ corresponds to slab, spherical, or cylindrical symmetry, respectively. Initial conditions on the solution variables at time $t = t_0$ are specified by number density and temperature profiles $\tilde{n}_s(z)$ and $\tilde{T}_{\hat{s}}(z)$ that are functions of the spatial coordinate. Boundary conditions satisfied by the solution variables are defined through user-specified functions $p(z, t; n_s, T_{\hat{s}})$ and $q(z, t)$ at the simulation boundaries $z \in \{z_1, z_N\}$ for all times in the interval $t \in [t_1, t_M]$. The boundary-condition function p is a column vector of size $|s| + |\hat{s}|$ and q is a diagonal square matrix; both are defined only at the lower and upper boundaries z_1 and z_N . The derivative coupling matrix c is a diagonal matrix and the local source function s is a column vector of size $|s| + |\hat{s}|$. A restriction on the elements of the diagonal matrices c and q is that the elements are either identically zero or nonzero.

Model boundary altitudes of 80 and 1040 km were used for the values z_1 and z_N , respectively. Dirichlet boundary conditions specified for the momentum and energy equations at the lower simulation boundary through the solver input functions p and q are defined as functions of empirical model output, as discussed in section 3.1.2. Neumann boundary conditions are specified at the upper simulation boundary and are set to user-defined parameterizations. The initial and final times t_1 and t_N passed in each call to the solver usually define an interval of 60 to 900 seconds, depending on the particular conditions modeled. Simulations lasting longer than 15

minutes of simulation time are performed by calling the solver function repeatedly in serial, using the final conditions of the previous call to determine the initial conditions of the next call. The wall time required to calculate a solution of the model equations does not depend strongly on the number of temporal points in the interval $[t_1, t_N]$ on which the final solution is defined so the temporal grid can be constructed somewhat arbitrarily, depending on the output required by the user.

A logarithmically spaced spatial grid with approximately 100 grid points was used in the SCIM simulations presented in this thesis. Grid points should be spaced such the vertical distance between adjacent points is about $\frac{1}{4}$ of the local scale height, the characteristic length scale over which ionosphere parameters like temperature and number density usually vary. Since the plasma scale height is altitude-dependent (and generally increases with altitude), grid points near the lower simulation boundary are more closely spaced than the grid points near the upper boundary. Note that the wall time required to solve the system of equations (3.6) depends strongly on the number of spatial grid points requested, decreasing the density of grid points near the upper boundary increases simulation efficiency without sacrificing model resolution. Therefore the spatial grid on which the final solution variables are defined is constructed such that the distance between grid points increases with altitude in order to economize the allocation of grid points throughout the simulation profile. The specific set of spatial grid points $\{z_k\}$ used in a typical run of the ionosphere model is defined by the equation

$$z_k = z_1 \left(\frac{z_N}{z_1} \right)^{\frac{k-1}{N-1}} \quad k = 1, 2, 3, \dots, N \quad (3.7)$$

Note that when the grid point positions are plotted on a logarithmic scale, the points are uniformly spaced. A refinement of this grid, described in section 4.2.4, is needed when simulating F -region HF pumping experiments described in later chapters because the region where the HF energy is converted to electron thermal energy may be as thin as 2 km of vertical thickness but the grid spacing near the critical layer is more than 5 km. The refined grid is constructed by defining a set of grid points using equation (3.7) and taking the set union with a uniform 500 m spaced grid that is 10 km tall and covers the HF interaction region. This refined grid was used even for simulation runs that did not incorporate HF-heating effects so that the resulting model ionosphere could be used as a set of initial conditions for subsequent numerical HF-heating experiments.

The SCIM ionosphere is assumed to be slab symmetric so $m = 0$ in the set of solver equations (3.6). For notational convenience, the arguments of the solver input functions c , f , and s , will be suppressed and the boundary condition vector equation will be separated into two scalar equations distinguished with the superscripts (1) and (N) to indicate the respective lower and upper boundaries. The solver equations can then be written in a compact form

$$n(z, t_1) = \tilde{n} \quad (3.8)$$

$$T(z, t_1) = \tilde{T} \quad (3.9)$$

$$p^{(1)} + q^{(1)} f = 0 \quad (3.10)$$

$$p^{(N)} + q^{(N)} f = 0 \quad (3.11)$$

$$c \frac{\partial}{\partial t} \begin{pmatrix} n \\ T \end{pmatrix} = \frac{\partial}{\partial z} f + s \quad (3.12)$$

Note that the species indices s and \hat{s} have been dropped from the solution variables n and T , as the vector nature of representing several species in the system of equations is implicit in this context. Section 3.1.1 describes how the numerical solver is applied to the model equations in Chapter 2 and how empirical model and user-specified parameterizations drive the simulation through the initial value functions \tilde{n} and \tilde{T} , the boundary condition functions p and q , the coupling matrix function c , the flux function f , and the source function s .

3.1.1 Solver equations

The system of fundamental equations written in sections 2.1.2 and 2.1.3 need to be written in the form of equations (3.6) or (3.12) before passing them to the pde numerical solver. For reference, the functions c , f , and s in equation (3.12) will be specified in this section for the model equations (2.32) and (2.39). Boundary conditions for each equation will be specified in the form of the functions p and q in equations (3.10) and (3.11). A procedure to set initial conditions (3.8) and (3.9) for each equation will be described. Knowledge of the actual physical conditions in the ionosphere that guide the choice of boundary or initial conditions is not generally known for most model domains; so the final solution should not be sensitive to the initial conditions chosen and it should vary in a straightforward way in response to variations in the boundary conditions. While the

solution to the model equations is driven to an extent by the boundary conditions used, this dependence can be exploited to estimate physical quantities like particle or heat flux through the top simulation boundary by fitting model-calculated density and temperature profiles to those measured by radar or rocket observations. Finally, the components of the model equations provided by empirical models need to be specified at arbitrary points in the space and time domain determined at model run-time, but some empirical model software (e.g., GLOW) need grid points specified at the time of software compilation. Interpolation is used to incorporate fixed-grid empirical models with the pdepe solver which demands values at arbitrary points in the space and time domain, not just at the grid points specified at run-time.

For notational convenience while putting the model equations into a form suitable for the numerical solver, each solution variable corresponding to electron, ion, and neutral species number density and temperature will be assigned to a numerical index defined in the following table.

$n_{\{n,i\}}$	$n_{\{n,i\}}$	$T_{\{n,i,e\}}$
$N_2 = 1$	$H^+ = 10$	$T_n = 20$
$O_2 = 2$	$He^+ = 11$	$T_i = 21$
$O = 3$	$O^+ (^4S) = 12$	$T_e = 22$
$H = 4$	$O^+ (^2D) = 13$	
$He = 5$	$O^+ (^2P) = 14$	
$N (^4S) = 6$	$O_2^+ = 15$	
$N (^2D) = 7$	$O_2^+ (a^4\Pi) = 16$	
$N (^2P) = 8$	$N^+ = 17$	
$NO = 9$	$N_2^+ = 18$	
	$NO^+ = 19$	

(3.13)

In the index notation defined in (3.13), the Mass Spectrometer and Incoherent Scatter (MSIS) data empirical model drives the neutral species number densities $n_{1...5}(z,t)$ and the neutral gas temperature T_{20} . Therefore the components of the solver functions f , s , and c in (3.12) associated with these quantities are

$$c_{1...5, 20} = 1 \quad f_{1...5, 20} = 0 \quad s_{1...5, 20} = 0 \quad (3.14)$$

Similarly, the boundary condition functions in (3.10) and (3.11) for these externally driven variables are

$$\begin{aligned}
P_{1\cdots 5}^{(1,N)}(t) &= n_{1\cdots 5}(z_{1,N}, t) - \tilde{n}_{1\cdots 20}(z_{1,N}) & q_{1\cdots 5}^{(1,N)}(t) &= 0 \\
P_{20}^{(1,N)}(t) &= T_{20}(z_{1,N}, t) - \tilde{T}_{20}(z_{1,N}) & q_{20}^{(1,N)}(t) &= 0
\end{aligned} \tag{3.15}$$

One consequence of the function definitions in equation (3.14) is that the particular solution variables referenced by the indices are independent of time and hence are identically equal to their respective initial conditions. The variables represented by indices 1...5 and 20 will be referred to as the *static* solution variables that are determined exclusively by the MSIS model. These variables are held constant over the time interval specified during each invocation of the pdepe solver function. The time-dependent state of the ionosphere is modeled by *dynamic* solution variables, represented by indices 6...19 and 21...22, are time dependent, and are outputs of the numerical solver. While it is possible to implement the ionosphere model such that the numerical solver extracts the required neutral species number densities and neutral gas temperature directly from the MSIS model without solving for solution variables that remain identically equal to their initial conditions, using the same data structure for both the static and dynamic solution variables was convenient for software development and data analysis.

The coupled continuity and momentum equations (2.32) of the dynamic solution variables, written in the form of equation (3.6) or (3.12), results in the following definitions of the numerical solver functions:

$$\begin{aligned}
c_s &= 1 & s &\in \{6, \dots, 19\} \\
f_s &= n_s w_N - D_n \left(\frac{\partial n_s}{\partial z} + \frac{n_s}{T_n} \frac{\partial T_n}{\partial z} - \frac{n_s}{H_s} \right) - K \left(\frac{\partial n_s}{\partial z} + \frac{n_s}{\langle H \rangle_n} + \frac{n_s}{T_n} \frac{\partial T_n}{\partial z} \right) & s &\in \{6, \dots, 9\} \\
f_s &= n_s w_N - (\sin^2 I) \left[D_s \left(\frac{\partial n_s}{\partial z} + \frac{n_s}{T_p} \frac{\partial T_p}{\partial z} - \frac{n_s}{H_s} \right) - K \left(\frac{\partial n_s}{\partial z} + \frac{n_s}{\langle H \rangle_n} + \frac{n_s}{T_n} \frac{\partial T_n}{\partial z} \right) \right] & s &\in \{10, \dots, 19\} \\
s_s &= \frac{\delta n_s}{\delta t} \Big|_{\text{driven production}} + \frac{\delta n_s}{\delta t} \Big|_{\text{chemical production and loss}} & s &\in \{6, \dots, 19\}
\end{aligned} \tag{3.16}$$

The bulk vertical velocity of the mean neutral gas $n_s w_N$ is difficult to measure in practice so it is assumed to be fixed at zero in the SCIM calculations described in this dissertation. The diffusion coefficients, scale heights, and the plasma temperature in (3.16) are replaced by expressions in terms of the solution variables as discussed in section 2.1.2 and are generally time-dependent.

In a slab-symmetric model ionosphere, the driven production term included in the source function s depends on the solar zenith angle, incident solar spectrum, and the number density altitude profile of the (static) major neutral species. Geophysical parameters, along with the static major neutral species densities and neutral gas temperature, are held fixed during the time span specified during each call to the numerical solver and consequently the driven production term in the source function s calculated with the GLOW model also remains fixed during each call to the solver. The chemical production and loss term in the source function is a dynamic quantity because it depends on dynamic density and temperature solution variables as well as static variables. Chemical production and loss rates that depend on the electron number density were calculated by substituting the sum of the initial-condition ion densities. The altitude-dependent dip angle I is calculated with the World Magnetic Model (WMM) [McLean *et al.*, 2004].

It is straightforward to write the coupled continuity and energy equations (2.39) for the electron and ion gases in the form of equation (3.6) or (3.12). The resulting solver input functions are:

$$\begin{aligned}
c_{20} &= 1 \\
c_s &= \frac{3}{2} n_e k_B & s \in \{21, 22\} \\
f_{20} &= 0 \\
f_s &= \lambda_s (\sin^2 I) \frac{\partial T_s}{\partial z} & s \in \{21, 22\} \\
s_{20} &= 0 \\
s_{21} &= \frac{\delta E_i}{\delta t} = \frac{\delta E_i}{\delta t} \Big|_{\text{Joule heating}} + \frac{\delta E_i}{\delta t} \Big|_{\text{elastic electron collisions}} + \frac{\delta E_i}{\delta t} \Big|_{\text{elastic neutral collisions}} \\
s_{22} &= \frac{\delta E_e}{\delta t} = \frac{\delta E_e}{\delta t} \Big|_{\text{supra-thermal collisions}} + \frac{\delta E_e}{\delta t} \Big|_{\text{elastic ion collisions}} + \frac{\delta E_e}{\delta t} \Big|_{\text{elastic neutral collisions}} + \frac{\delta E_e}{\delta t} \Big|_{\text{inelastic neutral collisions}}
\end{aligned} \tag{3.17}$$

As with the coupled continuity and momentum equations, the electron number density term is taken to be the sum of the ion species number densities. Parameterizations of the thermal conductivities of the ion and electron gases in terms of the solution variables are written in section 2.1.3; expressions of the energy transfer rates due to elastic and inelastic collisions between particles are provided for the electron gas in section 2.1.3.1 and for the ion gas in section 2.1.3.2. An arbitrarily defined altitude profile of the magnetospheric-imposed perpendicular electric field magnitude is used to calculate the ion gas Joule heating term via equations (2.53) and (2.54). Only altitude-independent perpendicular electric fields were used to drive the ion Joule-heating term in

the simulation results presented in this thesis. While the model does not enforce a limit to the value of $|E_{\perp}|$ applied to simulate ion Joule-heating effects, a practical upper limit is 40 mV/m since values larger than this may lead to ion drift speeds comparable to the mean neutral species thermal speed, resulting in a non-maxwellian ion velocity distributions. This violates the assumptions used to derive the model equations [Schunk and Nagy, 2000]. Joule heating of the ions is a static quantity that remains fixed during the time interval specified in each call to the numerical solver routine. The ion cooling rate due to elastic collisions with the neutral gas species, the electron cooling rate due to elastic collisions with the ion gas species (and the corresponding ion gas heating rate), and the electron cooling rate due to inelastic collisions with the neutral gas species are all dynamic quantities that can vary in time during each call to the numerical solver.

The volume heating rate of the electron gas due to elastic collisions with suprathermal photoelectrons or auroral electrons is calculated with the GLOW model. It uses the two-stream approximation [Nagy and Banks, 1970] to calculate the suprathermal electron energy distribution at each altitude bin, as well as the associated thermal electron heating rate due to collisions with the suprathermal electrons. Heating of the electron gas through collisions with suprathermal electrons is a static quantity in the ionosphere model. The two-stream electron transport calculation depends on assumed high-energy electron fluxes through the upper simulation boundary; the flux through the lower boundary is always assumed to be zero. Electron gas heating by collisions with photoelectrons originating in the geomagnetic conjugate profile is significant at low and mid latitudes [Narasinga Rao and Maier, 1970]. The model can include the effects of an arbitrary downward high-energy electron flux distribution (or a downward electron flux distribution equal to the upward photoelectron flux calculated with the two-stream code at the conjugate point) at the upper simulation boundary. However, downward energetic electron flux in the simulations described in this thesis is assumed to be zero.

Even though the final solution variables are specified only on predefined grid points -- on the set of points $\{z_i\}$ defined by equation (3.7), for example -- the numerical solver input functions in equation (3.6) need to be defined at arbitrary points between the simulation boundaries. While the MSIS and WMM empirical model software packages are local models in that the value of their output is determined by a single point in time and (parameter) space, the GLOW model output depends on an entire altitude profile. For instance, the photoionization rate of atomic oxygen at a given altitude depends on the fraction of high-energy photons incident from space that have been

absorbed at higher altitudes. That fraction, in turn, depends on the altitude profiles of the major neutral atmosphere number densities above the point of interest. Since the inputs to the solver functions in equation (3.6) do not include more than one point in space, entire altitude profiles of the photo-production and impact-production rates at the grid points are calculated with the GLOW model and neutral number density profiles that were calculated by the MSIS model at the beginning of each time interval. When the numerical solver needs neutral number density values and ion production rate values at altitudes that do not lie on the predefined grid points, the values are estimated by linear interpolation.

3.1.2 *Boundary conditions*

Although no distinct physical boundary exists at the lower or upper bounds of the simulation domain, specific boundary conditions still need to be assumed in order to numerically solve the model equations for a unique solution. In this section, Dirichlet or Neumann boundary conditions for the coupled continuity and momentum equations—(2.32) or (3.16)—and the coupled continuity and energy equations—(2.39) or (3.17)—will be specified; the choice of which type of boundary condition to use depends on the significance of particle or energy transport relative to the local production or loss at the particular boundary of interest. The purpose of this section is to express the boundary conditions of the model equations in the form of equation (3.5), as required by the pdepe numerical PDE solver.

At 100 km altitude, 20 km above the lower simulation boundary, the diffusion times as illustrated in Figure 2.7 are on the order of thousands of seconds so the continuity equation without the transport term (2.10) can be used to adequately describe the time evolution of the ion and minor neutral species number density. Particle transport across the lower simulation boundary can therefore be assumed to be insignificant over a typical simulation time step of one to tens of minutes in duration. Consequently, Dirichlet boundary conditions will be specified for the continuity-momentum equations at the lower simulation boundary. That is, the number density of each species must be pre-defined at 80 km altitude before each call to the numerical solver. For the driven neutral species, the number densities are determined directly with the MSIS model. The remaining ion and minor neutral species are assumed to be in photochemical equilibrium with the major neutral gas species $\{O, O_2, N_2, H, He\}$ that are driven by MSIS. In a typical simulation time step, usually about ten minutes or less in duration, the geophysical parameters including the solar zenith angle are held fixed so the photochemical equilibrium assumption implies that the time

derivative of the species number density on the left hand side of equation (2.10) is approximately zero and that the species number density can be written as the ratio of its volume production rate to its loss frequency.

$$n_s \approx P_s / L_s \quad (3.18)$$

Even if the geophysical parameters were not held fixed during each simulation time step, the approximation above is generally valid except during sunrise and sunset when the slant column optical thickness between the observation point and the sun changes rapidly with time. In terms of the numerical PDE solver functions p and q , the lower boundary condition of the coupled continuity and momentum equation for ion and minor neutral species is

$$p_s^{(i)} = n_s(z_1, t) - P_s(z_1, t) / L_s(z_1, t) \quad q_s^{(i)} = 0 \quad s \in \{6, \dots, 19\}. \quad (3.19)$$

The lower boundary conditions of the major neutral species number density are written in equation (3.15) and both are set to the initial conditions at the beginning of the time step as determined by MSIS.

Similarly, vertical thermal energy transport is negligible at the lower simulation boundary, as indicated in Figure 2.8 by the small values of thermal flux and thermal conductivity near 100 km altitude. Therefore Dirichlet boundary conditions for the ion and electron coupled energy and continuity equations (2.39) are enforced at the lower boundary by setting the electron and ion temperatures to the neutral gas temperature.

$$p_s^{(i)} = T_s(z_1, t) - T_{20}(z_1, t) \quad q_s^{(i)} = 0 \quad s \in \{21, 22\}. \quad (3.20)$$

The neutral gas temperature is driven by MSIS and its boundary condition is determined in equation (3.15) by its initial condition.

At 1000 km altitude, near the upper simulation boundary, particle transport affects the species number density in any given region more than the combined local production and loss values, as illustrated in Figure 2.3, Figure 2.4, and Figure 2.7. Therefore Neumann boundary conditions are applied to the coupled continuity-momentum equations (2.32) at the upper simulation boundary by specifying a number density flux through the boundary. Unfortunately, the value of this mass flux is not generally known at any given time and location so it must be estimated or used as a

parameter to fit simulation results with observational data. The magnitude and direction of the vertical or field-aligned plasma flux, also known as the “polar wind” [Axford, 1968], varies with altitude, latitude, universal time, ionospheric conditions, and ion species [Blilly *et al.*, 1992]. Near 1000 km altitude at 65 degrees north latitude during northern hemisphere summer, the field-aligned plasma velocity is typically around 100 m/s in the downward direction [Horvath and Lovell, 2009]. Assuming a reasonable plasma density of 10^{10} m^{-3} at 1000 km altitude, the resulting downward plasma flux is approximately $10^{12} \text{ m}^{-2}\text{s}^{-1}$ although note that the net plasma flux and velocity generally increases rapidly toward the upward direction from that altitude with increasing altitude [Ganguli, 1996]. The model ionosphere illustrated in Chapter 2 was calculated by assuming a zero net number density flux at the top boundary, in approximate agreement with the survey of in-situ ion-flux measurements at about 860 km altitude compiled by Horvath and Lovell [2009] from several months of Defense Meteorological Satellite (DMSF) satellite during the southern hemisphere winter.

The number density flux of a particular gas species can be expressed symbolically as the product of the species number density with the species mean drift velocity $n_s \mathbf{u}_s$ at the upper simulation boundary. In the SCIM software, the number density flux of at the top boundary is specified as a function of time for each species in the form of the functions p and q in equation (3.5). Since positive values of the function f_s in equation (3.16) corresponds to the upwards number density flux of species s , the functions p and q for a (downward) species flux of, say, $\phi_s = -10^{12} \text{ m}^{-2}\text{s}^{-1}$ are

$$p_s = \phi_s \quad q_s = -1 \quad s \in \{6, \dots, 19\}. \quad (3.21)$$

Note that the number density flux of each species through the top model boundary can be set individually, and are allowed to be time-dependent.

As with the species momentum-continuity equations, Neumann boundary conditions are used at the top simulation boundary for the ion and electron energy equations in equation (2.39) because the local heat sources and sinks, illustrated in Figure 2.9, decrease exponentially with increasing altitude above the f_oF_2 peak while the heat flux, illustrated in Figure 2.8, does not. Estimating a reasonable value of the thermal energy flux at the top simulation boundary is more problematic than setting the number density flux because a significant downward thermal energy flux may be required to bring mid- and high-latitude electron and ion gas temperature model profiles into

approximate agreement with corresponding observations [Rasmussen *et al.*, 1988; Roble, 1975]. Furthermore, the transport of thermal energy from the magnetosphere is likely latitude- and time-dependent even during quiet geomagnetic conditions. Typical downward electron heat flux values used in the model for quiet geomagnetic daytime conditions are between 10^{-6} and 10^{-5} W/m². In terms of the numerical solver boundary condition functions p and q in equation (3.5), an electron heat flux (downward) of $\phi_{22} = -10^{-5}$ W/m² through the top simulation boundary is specified as

$$p_{22} = \phi_{22} \quad q_{22} = -1 \quad (3.22)$$

since an upwards electron heat flux is represented by a positive value of f_{22} . Although the ion and electron heat fluxes through the top boundary can be set independent of each other and may be time dependent. However, only time-independent electron heat fluxes were used at the upper boundary for the simulation runs described in this thesis. Also, the boundary heat flux of the ion gas will always be set to zero. The assumed absence of thermal transport through the top simulation boundary via ion gas thermal conduction is justified in most circumstances because the ion thermal conductivity is significantly less than the electron thermal conductivity. Elastic (Coulomb) collisions with the magnetosphere-heated thermal electrons is the main source of ion heating in the topside ionosphere [Rees and Roble, 1975].

3.1.3 Simulation structure

This section describes the SCIM main program algorithm at the level of main subroutine calls. The general order of subroutine calls is described first, followed by a brief description of each main subroutine. For an initial set of user-defined geophysical conditions, the SCIM main program extracts data from the empirical model software, calls the numerical routine to solve the system of model equations, stores the resulting solution variables to non-volatile storage, and then finally updates the geophysical conditions according to the user-defined time step before repeating the process. The results of a specific model run for a representative set of geophysical initial conditions will be described in section 3.2 and validated by comparison with measurements and empirical ionosphere models.

The flow chart in Figure 3.1 illustrates the order of significant program steps in between calls to the numerical PDE solver routine during a typical model run. A model run is specified by a set of geophysical and simulation parameters, some of which may be time-dependent. Unless noted

otherwise, the following geophysical and simulation parameters are static and do not vary over the course of a single simulation: the geographic latitude and longitude, the initial simulation time, the (possibly time-dependent) time step for each call to the numerical PDE solver, the altitude range and grid points of interest, the time-dependent solar radio flux parameters F_{107} and F_{107a} , the time-dependent geomagnetic activity index A_p , the time-dependent mass and energy flux through the top simulation boundary, and the time-dependent energetic electron flux energy distribution at the top simulation boundary. Additionally, the empirical models MSIS and GLOW provide several additional parameters that can be modified depending on the particular conditions to be modeled. However these parameters are set to their default values in the simulation runs described here. If a multi-processor machine is available, then multiple copies of the simulation can be run simultaneously to perform a “parameter sweep” where the model dependence on a particular parameter can be explored without significantly increasing the total runtime of the simulation.

The SCIM main program calls a set of subroutines in a sequence that begins with setting initial conditions and ends with saving the results from the pdepe numerical solver. To define the initial conditions, SCIM retrieves the set of user-defined geophysical and simulation parameters, then a vertical profile of the geomagnetic field vector is calculated with WMM and an altitude-dependent gravitational acceleration is calculated with Newton’s law of gravitation. Then neutral gas species number density profiles are calculated with the MSIS model and ion-neutral collision frequencies are calculated as described in section 2.1.2. When the main program is at its initial time step, the ion density profiles may be loaded from nonvolatile storage along with ion and electron temperature profiles that were calculated in a previous simulation run for a “warm start.” Alternately, for a model “cold start” that does not use prior simulation data, the ion number density profiles are initialized to zero while the ion and electron temperature profiles are initialized to the MSIS neutral gas temperature. For subsequent time steps, the state of the number density and temperature profiles at the end of the previous time step are used to initialize the state at the beginning of the next time step. Then the GLOW model is used -- with the solar energy flux parameterization, major neutral gas number density profiles, and the energetic electron flux boundary condition -- to calculate altitude profiles of excitation rates, ionization rates, and the electron volume heat rate due to collisions with energetic electrons. The final software execution step in preparing the state of the ionosphere parameters before solving the model transport equations is to calculate an altitude profile of the ion volume heat rate due to Joule heating from an externally applied perpendicular electric field.

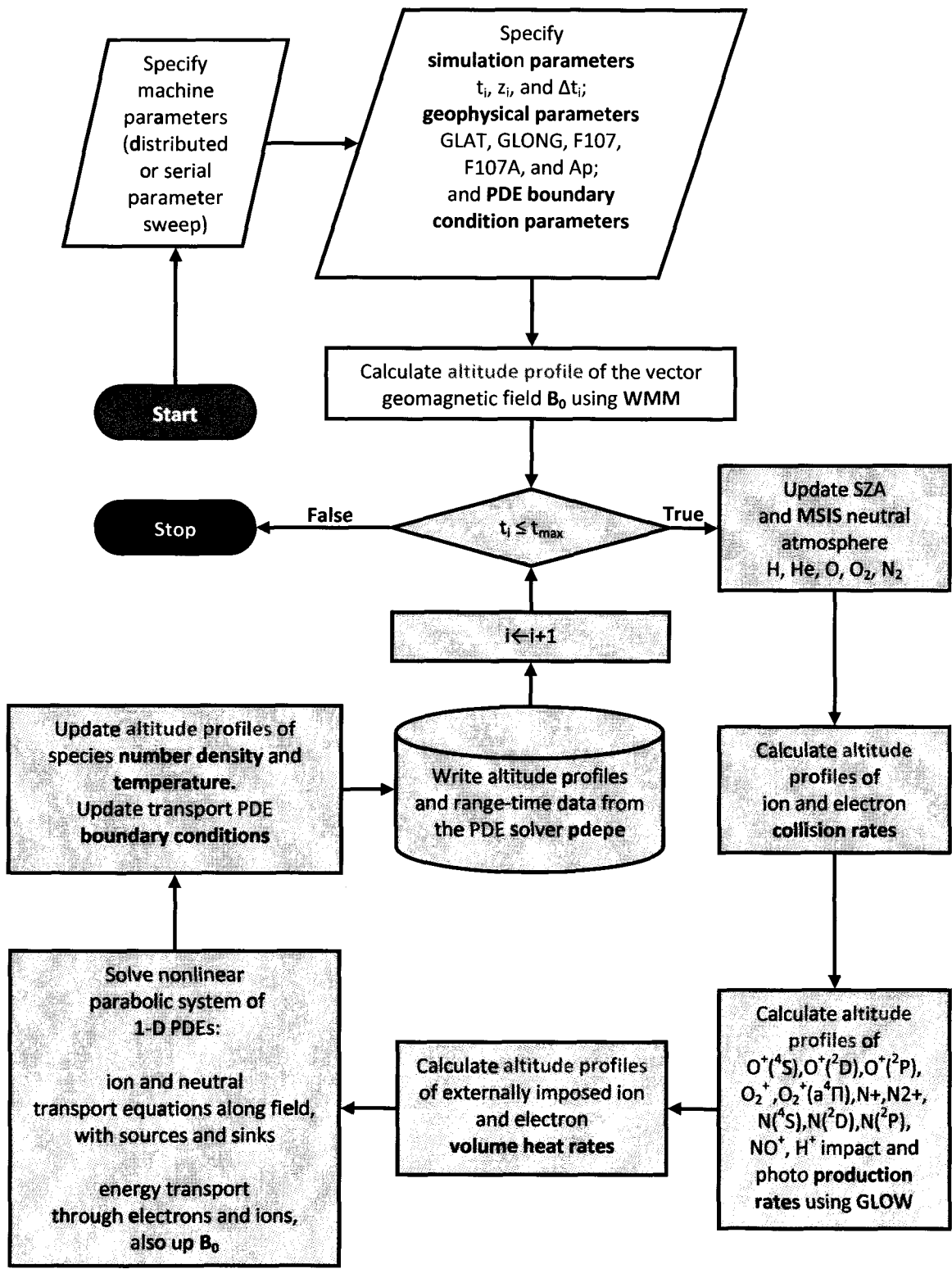


Figure 3.1 Flow-chart schematic of the 1-D Self-Consistent Ionosphere Model (SCIM).

The density and temperature initial conditions required by the numerical solver are set to the altitude profiles of the ion and neutral species number densities, as well as the neutral, ion, and electron gas temperatures at the beginning of the current time step. If there is little to no ionization because the simulation is at the beginning of a cold start, then only the coupled momentum-continuity equations are solved; the ion and electron gas temperatures are maintained at the neutral gas temperature for the duration of the time step. Once foF2 exceeds a pre-defined threshold, then the entire system of coupled continuity, momentum, and energy equations are solved as described above. At the i^{th} time step, the numerical PDE solver returns the values of the solution variables n_s and T_s evaluated on the specified spatial grid -- at the points given by equation (3.7) in this case -- at arbitrarily specified points in time between the times t_i and t_{i+1} . The values of the solution variables and all time-dependent parameters are saved to disk at the end of each time step, and then the current simulation time is advanced $t_i \leftarrow t_{i+1}$. Time-dependent geophysical parameters are updated including the neutral gas density and temperature variables driven by the MSIS empirical model. Finally, the initial values of the dynamic solution variables are set to the final values of the dynamic solution variables at the previous time step.

3.2 Model validation and sensitivity analysis

Simulation results from the self-consistent ionosphere model (SCIM) are compared in this section with output from the IRI empirical model, measurements of the vertical structure of the ionosphere made with the Poker Flat Incoherent Scatter Radar (PFISR), measurements of the critical frequency made with the HAARP ionosonde, and with measurements of the vertical total electron content made with the HAARP GPS receiver. Simulation results are found to be in agreement with measurements and empirical model output.

First, the model parameters listed in Table 3.1, including the particular time and location of the numerical and observational experiments, are summarized. Then the vertical structure of the ionosphere at a particular instant in time is validated in section 3.2.1 by comparing altitude profiles of the plasma composition and temperature calculated by the SCIM model with corresponding profiles calculated by the IRI model and measured with the Poker Flat ISR. The diurnal variation of the critical frequency calculated by the SCIM and IRI models is compared with the critical frequency measured by the HAARP ionosonde near Gakona, Alaska in section 3.2.2. Similarly, the diurnal variation of the total electron content as measured by a GPS receiver at the HAARP facility is compared with the TEC calculated by SCIM and the IRI model in section 3.2.3. And in section

3.2.4, the sensitivity of the SCIM-calculated electron column production rate and energy density to the diurnal variation in the neutral gas column number density is explored. Additional physical context for the observations made from Gakona and Poker Flat during the times of geophysical conditions similar to those simulated with the SCIM model is provided in Chapter 4 and Chapter 5.

The time evolution of the vertical structure of the polar ionosphere at a given location is, to a first approximation, largely determined by the vertical structure of the neutral atmosphere, the solar zenith angle, the intensity of the incident ionizing photons or energetic particles. The vertical structure of the neutral atmosphere is estimated with the MSIS model as a function of the geographic latitude, longitude, and time in addition to solar and geomagnetic indices. The photoionization rates are estimated with the GLOW model from the MSIS neutral atmosphere profile, the solar zenith angle, a parameterization of the incident photon flux energy distribution, and the electron temperature profile calculated by the SCIM model. Other geophysical parameters such as the magnitude of the perpendicular electric field or the neutral atmosphere wind velocity play an important but secondary role in determining the vertical structure of the ionosphere.

Table 3.1 Parameters used in a self-consistent ionosphere model (SCIM) model run configured to simulate the vertical structure and time-evolution of the natural ionosphere over Gakona, Alaska on 24 October 2008.

IONOSPHERE MODEL CONFIGURATION

Parameters	Values
<i>GEOSPATIAL</i>	
Initial time	20:00:00 24 October 2008 UT
Time period	24 hour diurnal cycle
Location	(62.4°, 214.9°) GEO
Time step	900 s
Vertical grid	$\left\{ 80 \left(\frac{1039}{80} \right)^{\frac{k-1}{99-1}} \mid k = 1, 2, \dots, 100 \right\}$ $\cup \left\{ 210 + \frac{1}{2}k \mid k = 0, 1, \dots, 20 \right\}$ km
<i>LOWER BOUNDARY</i>	
Altitude	80 km
Number density	Photochemical equilibrium
Temperature	MSIS neutral gas temperature
<i>UPPER BOUNDARY</i>	
Altitude	1040 km
Number density flux	0.0 m ⁻² s ⁻¹ (all species)
Thermal energy flux (electron gas)	2.0 μW m ⁻² ↓ (electron gas)
Thermal energy flux (ion gas)	0.0 μW m ⁻² (ion gas)
Non-thermal energetic electron flux	0.0 m ⁻² s ⁻¹ μW ⁻¹
<i>SOLAR AND GEOMAGNETIC INDICES</i>	
Ap	1
F107	68 × 10 ⁻²² W m ⁻² Hz ⁻¹
F107A	68 × 10 ⁻²² W m ⁻² Hz ⁻¹
<i>AMBIENT GEOPHYSICAL CONDITIONS</i>	
E _⊥	1.0 mV m ⁻¹
Solar parameterization	[Hinteregger and Fukui, 1981]
Horizontal neutral wind speed	0 m s ⁻¹

Parameters used to configure the SCIM model and drive the MSIS, GLOW, and WMM models—for the purpose of simulating the evolution of the ionosphere during a particular diurnal cycle—are compiled in Table 3.1. Each figure of a SCIM-calculated ionosphere over Gakona, Alaska on 24 October 2008 in Chapter 2, Chapter 3, and Chapter 4 was extracted from a single SCIM diurnal equilibrium run that was initialized using the parameters listed in the table. A *diurnal equilibrium run* refers here to a model run that is configured such that the universal time is advanced through one or more 24-hour cycles without advancing the model date. The PDE numerical solver was called 96 times in each diurnal cycle, once for each 15-minute (simulation time) intervals; variable quantities that are driven by empirical models, such as the major neutral gas composition or electron photoproduction rate, are held constant during each time interval. One practical motivation to build diurnal equilibrium model runs is to generate UT-indexed sets of suitable initial ionosphere model conditions for short duration but high time-resolution runs. Simulation runs described in Chapter 4 and Chapter 5 use 1-minute time steps and are initialized to model ionosphere conditions near those at 00:00 UT in the Gakona diurnal equilibrium run described here.

Next, the vertical structure and diurnal variation of the SCIM-calculated ionosphere will be validated by comparing SCIM results with IRI empirical model output and measurements made with the HAARP ionosonde. Static altitude profiles extracted from the model at a specific instant in time will be examined in section 3.2.1. The convergence of the model to a time-dependent equilibrium solution and the dependence of the solution on initial conditions will be discussed in sections 3.2.2 and 3.2.3. And finally, the sensitivity of ionosphere model quantities to tidal oscillations in the MSIS neutral atmosphere number density will be described in section 3.2.4.

3.2.1 Vertical structure of the ionosphere

The vertical structure of the SCIM-simulated ionosphere at a particular moment in time from the Gakona diurnal equilibrium run is validated here by first comparing model altitude profiles – of ion composition, electron density, ion gas temperature, and electron gas temperature – with the respective altitude profiles extracted from the IRI empirical model. Then the electron density profiles as calculated by SCIM and IRI are compared with an electron density profile measured with the HAARP ionosonde at Gakona, Alaska. IRI model temperature profiles, that initially appear to differ substantially from the SCIM-calculated electron and ion temperatures, are put into context

by comparing IRI model temperature profiles with temperature profiles measured by the incoherent-scatter radar at Poker Flat, Alaska.

The ion concentration as a function of altitude over Gakona at 01:00 UT is illustrated in Figure 3.2 for a SCIM-calculated ionosphere and as reported by the IRI empirical model. The electron density as a function of altitude, calculated as a sum of the ion number densities is also plotted in the figure for reference. The SCIM and IRI electron density profiles are generally in agreement but minor differences between the profiles exist. In particular, the IRI critical layer number density is approximately 40% less than the corresponding SCIM number density yet the topside number density reported by IRI is about 60% larger than the corresponding number density. Both differences can be readily explained by the significantly higher electron temperature reported by

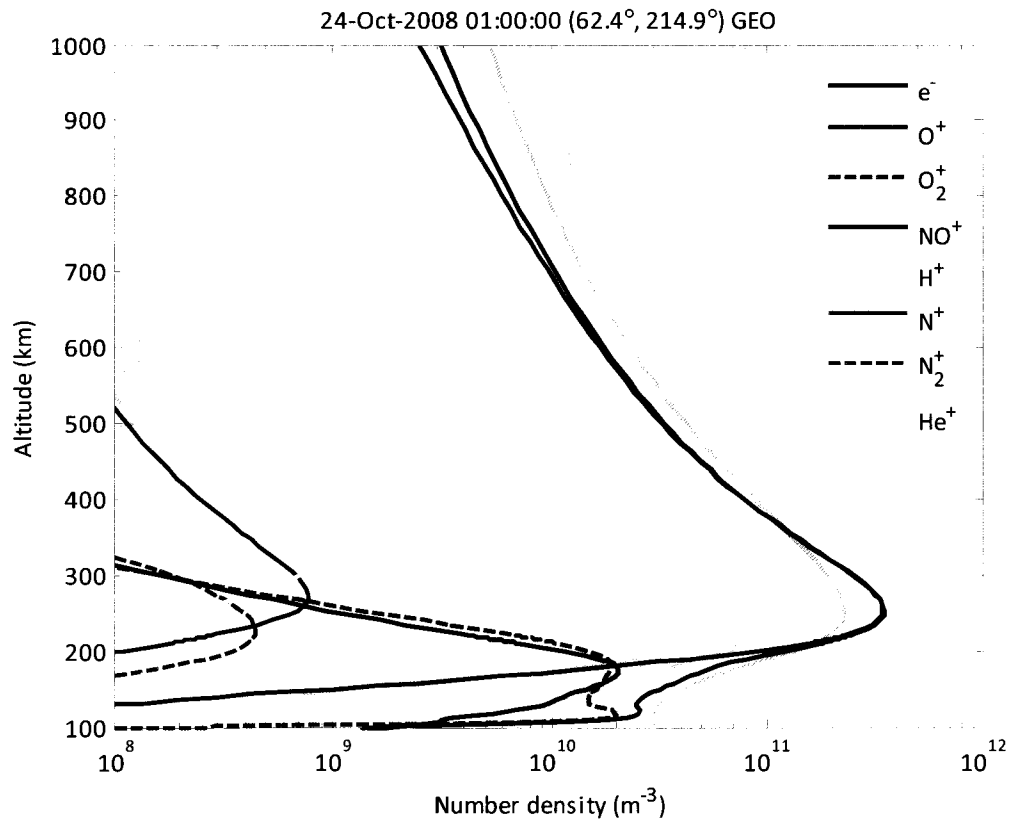


Figure 3.2 Altitude profiles of ion species number density modeled over Gakona, Alaska. The electron number density is calculated by the summing of the ion species number densities. Respective ion species number density profiles calculated with the IRI 2007 empirical model are printed in lighter colors relative to their respective SCIM-calculated profiles.

IRI over the corresponding SCIM profile, illustrated in Figure 3.3. The enhanced electron temperature reported by the IRI model would result in enhanced electron chemical loss rates near the critical layer where the neutral gas density is still relatively large. And in the topside ionosphere where chemical reactions with the neutral gas species are less important than transport effects for determining electron concentration, the enhanced electron temperatures result in enhanced electron number density because increasing the electron temperature decreases the vertical density gradient required to maintain a hydrostatic equilibrium of the electron-ion gas against gravity.

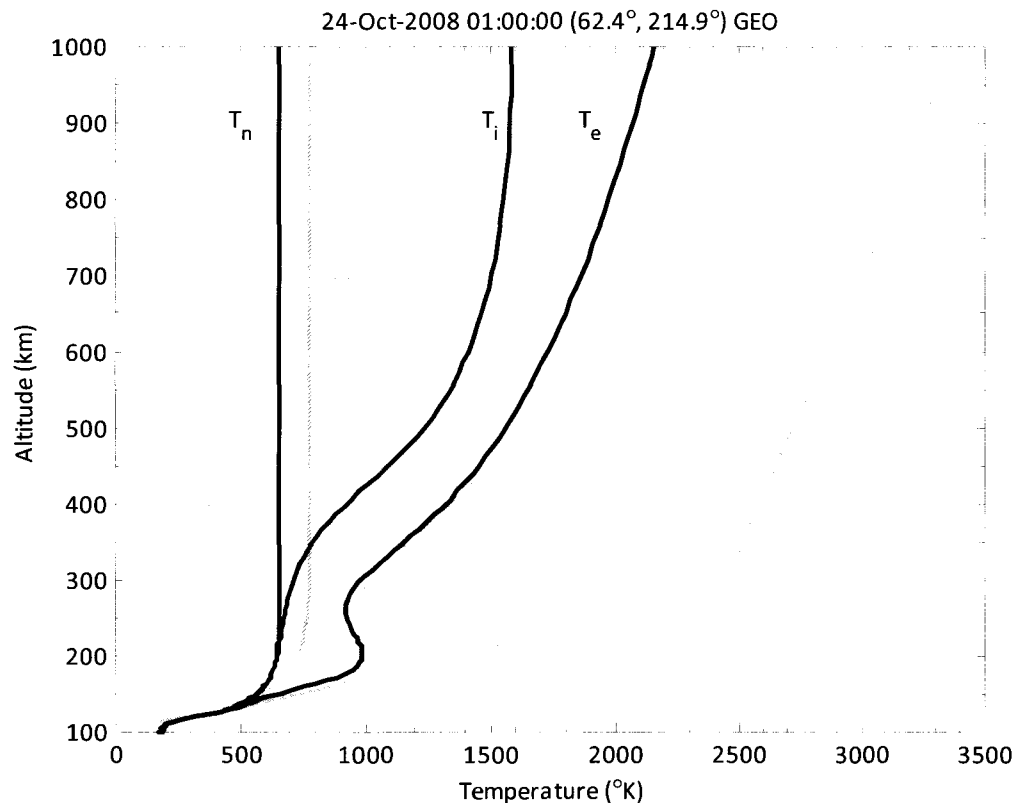


Figure 3.3 Altitude profiles of the electron temperature (dark red) and ion temperature (dark blue) modeled over Gakona, Alaska on 24 October 2008. The neutral gas temperature profile (black) is calculated with the MSIS empirical model. Corresponding temperature profiles (light colors) calculated with the IRI2007 model are provided for comparison.

Although the IRI electron temperature profile illustrated here is significantly larger than the SCIM electron temperature profile, the SCIM-calculated electron temperature – in the absence of transient heat sources like auroral precipitation or high-power HF radio waves – is significantly affected by the thermal heat flux specified at the top simulation boundary. The heat flux through the top model boundary maintains the positive electron temperature gradient with increasing altitude due to the lack of significant heat sinks and the large electron thermal conductivity in the topside ionosphere. Therefore, a larger choice of top boundary heat flux will bring the SCIM model electron temperature to good agreement with the IRI model value. Also note that the IRI model is based on data assimilation techniques applied to data that is more abundant in low to mid latitudes

so the model estimates at high latitudes must be evaluated critically. For example, the ion temperatures reported by IRI in Figure 3.3 are particularly suspect by virtue of the near-linear shape of the profile. Plasma temperature and number density profiles calculated by both models can be further validated through comparison with the respective profiles measured with incoherent scatter radar (ISR) during geophysical conditions similar to those input into the models.

The Advanced Modular Incoherent Scatter Radar (AMISR) at the Poker Flat Research Range (PF), Alaska -- about 330 kilometers north-northwest of Gakona (GAK) -- was operating shortly before the time interval simulated in the Gakona diurnal equilibrium run. Figure 3.4 shows a scatter plot of electron density, electron temperature, and ion temperature profiles inferred from long-pulse incoherent scatter radar measurements made during a 3-hour time interval. The density and temperature profiles, which are calculated from the incoherent scatter radar data, are time-integrated in 2-minute increments and a density or temperature measurement for each range bin at each 2-minute interval is plotted as a point. The magnitude of the variability with time in each measured quantity is compared in Figure 3.4 with the variation in profiles of those same quantities calculated by the SCIM and IRI models. The variation of plasma density and temperature observed over a 3-hour SCIM model run is due to the diurnal change in solar zenith angle and neutral gas number density. Variation in the profiles measured with ISR is due to measurement uncertainty and natural ionospheric variability.

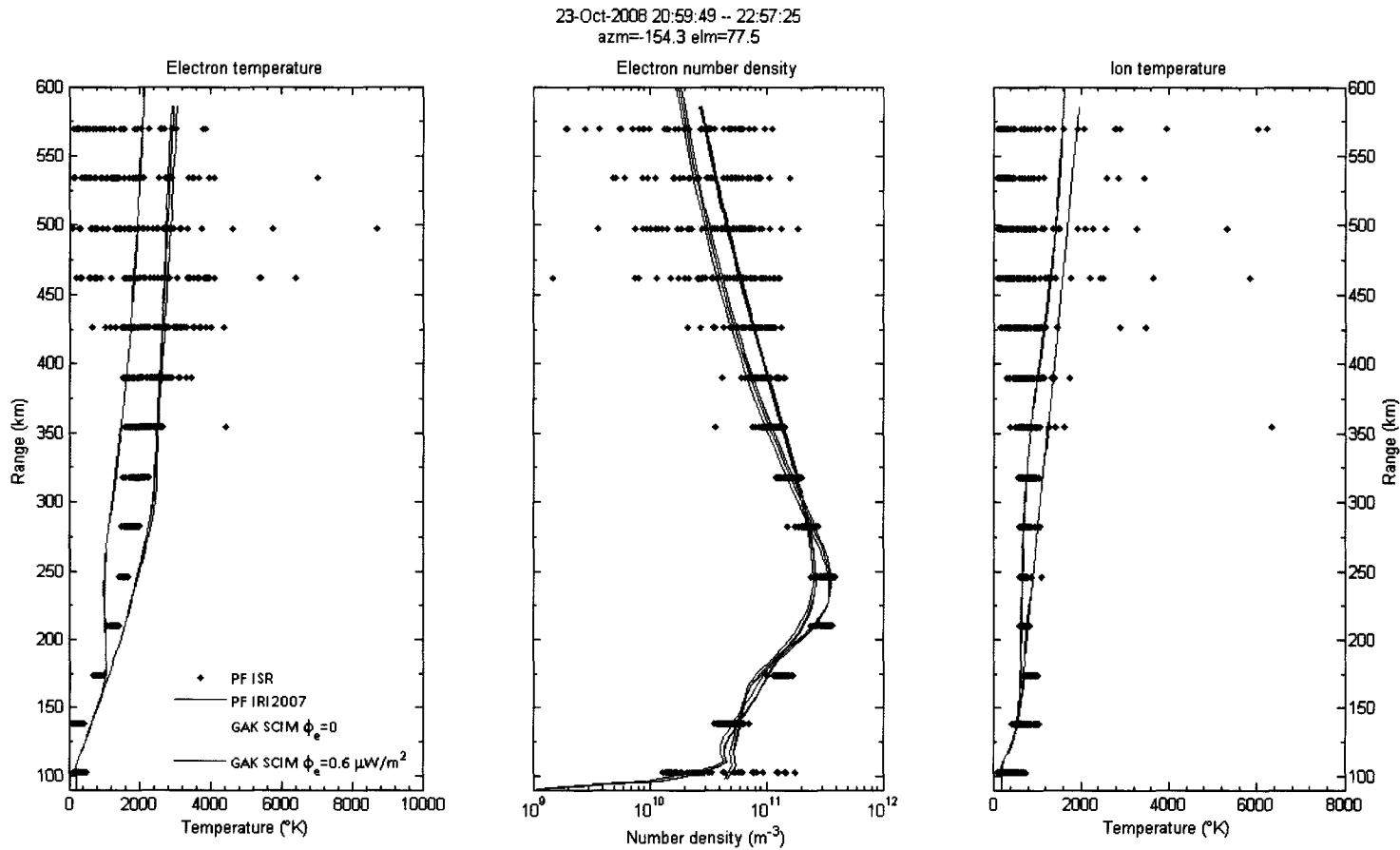


Figure 3.4 Scatter plot of ISR measurements of plasma density and temperature (black points) above Poker Flat, Alaska, on 23 October 2008. Altitude profiles of corresponding quantities at 21:00, 22:00, and 23:00 UT calculated with the IRI model (red lines) and the SCIM model (dark blue lines) are overlaid for comparison. The SCIM model output is from the Gakona diurnal equilibrium run defined by the parameters listed in Table 3.1 except that the electron gas heat flux at the top simulation boundary is set to $0.6 \mu\text{W}/\text{m}^2$ downward. Altitude profiles at 22:00 UT from a SCIM model run using a $0.0 \mu\text{W}/\text{m}^2$ electron gas thermal flux at the top simulation boundary (light blue lines) are also shown.

While the radar-measured and model-calculated profiles are in good agreement, both models calculate an ion temperature that is generally larger than the radar-measured ion temperature above 400 km range. Also, the shape of the electron temperature profile measured by radar is different from the shapes calculated by the models: the radar-measured electron temperature profile appears to have a local maximum at around 450 km range and a negative temperature gradient extending to the furthest measured temperature. The SCIM-calculated electron temperature profile has a local maximum at around 175 km, sustained by a source of energetic photoelectrons; a local minimum at 250 km, preserved by strong ion cooling near the critical layer; and a positive gradient at extending above 250 km that is maintained by the downward electron heat flux imposed at the upper simulation boundary. For comparison, density and temperature profiles are also plotted in Figure 3.4 that are calculated with SCIM for the case where the electron gas heat flux at the top simulation boundary is set to zero. Note that both the ion and electron temperatures in the topside ionosphere in the zero-heat flux model run are several hundred degrees Kelvin less than the profiles calculated for a $0.6 \mu\text{W}/\text{m}^2$ downward electron gas heat flux. Furthermore, the electron number density in the topside ionosphere is about $5 \times 10^9 \text{ m}^{-3}$ less than in the zero-heat flux model run. The cooler ion temperatures in the zero-heat flux model run are a direct result of the cooler electron temperatures since elastic collisions between the ions and electrons are the primary source of ion heating in the topside ionosphere. Consequently, the cooler plasma temperature inhibits ion ambipolar diffusion against the vertical electron number density gradient which results in a lower electron number density in the topside.

3.2.2 *Diurnal variation of daytime critical frequency*

In this section the SCIM-calculated diurnal variation of the ionosphere critical frequency foF_2 is compared to the time-variation of critical frequency as calculated by the IRI model and as measured by the HAARP ionosonde. As in the previous section, the time and location of the numerical simulation and observational measurements cover the day of 24 October 2008 at Gakona Alaska. The convergence time of the simulation foF_2 time series to a periodic diurnal cycle from an initial plasma-free ionosphere is also discussed.

Radio waves transmitted upward from earth—that are not first absorbed in the D or E region—may be reflected by the F-region ionosphere at the altitude where the electromagnetic wave frequency is approximately equal to the ionosphere plasma frequency. The radio wave is reflected at the first location along the wave path where the ionosphere plasma frequency

($f_p = \omega_p / 2\pi = \sqrt{n_e e^2 / m_e \epsilon_0} \approx 8.98 \sqrt{n_e \cdot 1(\text{m}^3)} \text{ s}^{-1}$, a function of electron number density only) is nearly equal to the electromagnetic wave frequency. The maximum plasma frequency above a location is called the ionosphere *critical frequency*. Alternately, the critical frequency is the greatest electromagnetic wave frequency that will be reflected by the ionosphere. An ionosonde measures the round-trip times of radio waves transmitted at various frequencies to estimate the electron number density below the critical layer as a function of height. The polarization and Doppler shift of the reflected waves are used by the HAARP ionosonde to infer additional characteristics, such as the bulk plasma velocity, of the ionosphere plasma above the transmitter.

The plasma frequency as a function of altitude, measured by the HAARP ionosonde at 01:00 UT, is plotted in Figure 3.5 as a solid black curve. Model calculations internal to the ionosonde device are used to infer the plasma frequency altitude profile above the ionosphere critical frequency; uncertainty in the plasma frequency measurements as estimated by the ionosonde is indicated by the gray lines surrounding the plasma frequency profile. Vertical profiles of ionosphere plasma frequency, as calculated by SCIM and the IRI model over Gakona at 01:00 UT, are plotted in Figure 3.5 as blue and cyan curves. The SCIM plasma frequency profile falls within the ionosonde-calculated uncertainty bounds below 450 km and the IRI model is below the lowest uncertainty bound near the critical layer, that is, the altitude region of maximum plasma frequency. Note that both SCIM and the IRI model overestimate the plasma frequency relative to that measured by the ionosonde in the E and F_1 regions but this may be due to the effects of HF ionosphere modification experiments, described in Chapter 4 and Chapter 5, that were operating during the duration of the ionosonde measurements that can enhance the ionosphere electron number density below the F_2 region [Ashrafi *et al.*, 2006; Perkins and Roble, 1978].

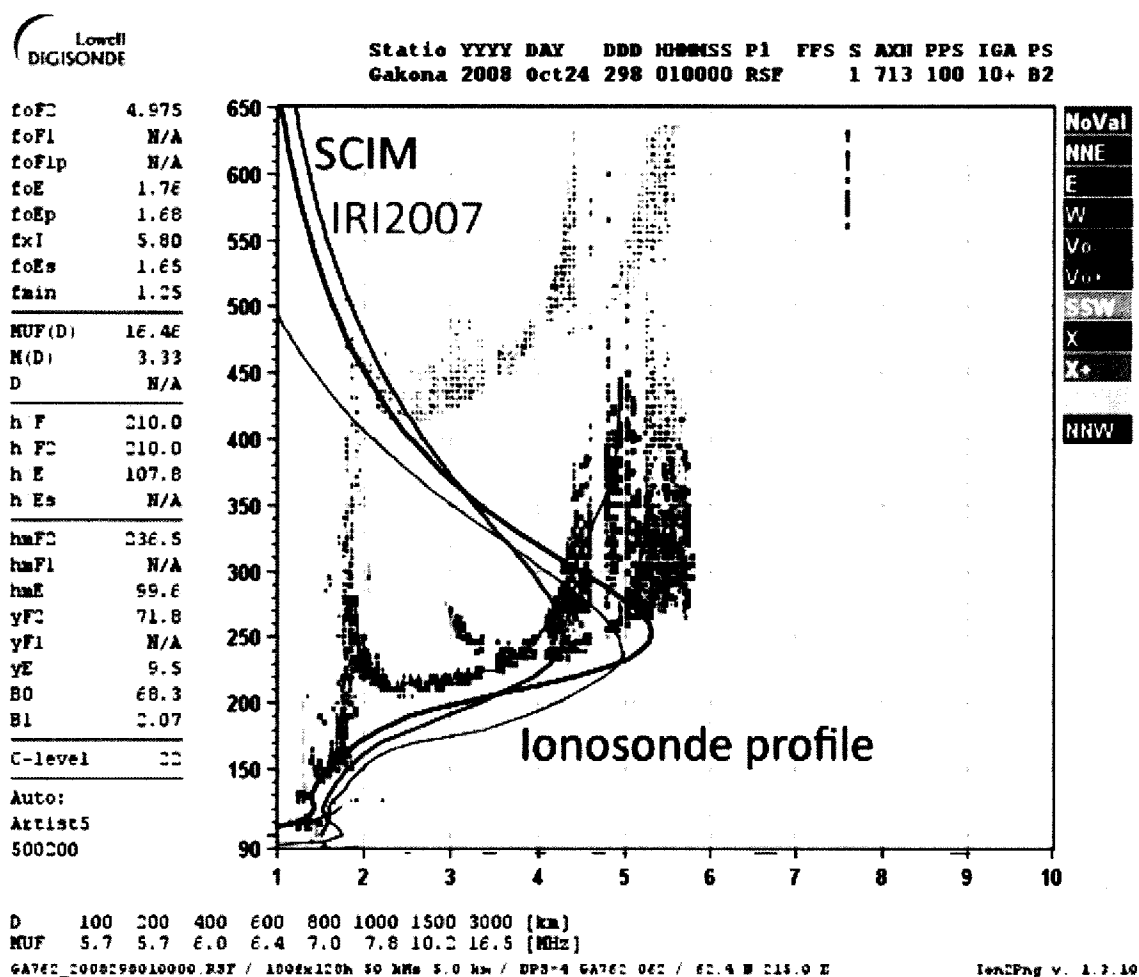


Figure 3.5 Altitude profile of plasma frequency (black) measured with the HAARP ionosonde over Gakona, Alaska at 01:00 UT (hh:mm) on 24 October, 2008. The corresponding frequency profiles calculated by the SCIM model (dark blue) and IRI 2007 empirical model (cyan) are overlaid on the ionogram for comparison. The horizontal axis is measured in MHz and the vertical axis is measured in vertical km.

Ionizing solar radiation incident on the major neutral gases in the thermosphere is the single largest source of electrons in the ionosphere. As the solar zenith angle varies in time, the optical thickness of the atmosphere between the point of observation and the sun varies, thus the number of ionizing photons incident from space at a particular location varies with time and consequently so does the electron photoproduction rate. The electron number density in the *F*-region critical layer depends strongly on the photoproduction rate, so a diurnal variation in the solar zenith angle is accompanied by a diurnal variation in the critical frequency. Figure 3.6 illustrates the SCIM-

calculated time-dependence of foF₂; the solar zenith angle at Gakona as a function of universal time also plotted reference. In particular, note that the foF₂ maxima lags the solar zenith angle minima by about 2 hours and the foF₂ minima lags the solar zenith angle maxima by about five hours. The time-dependence of foF₂ as calculated the IRI2007 model is also plotted in Figure 3.6 to provide validation of the SCIM calculations. The lower electron temperature calculated by SCIM relative to the temperature calculated by the IRI model near the time of SZA minimum is consistent with the larger SCIM-calculated critical frequency observed relative to IRI since the main loss process of the dominant ion, singularly charged atomic oxygen, is temperature dependent and its chemical reaction loss rate increases with increasing temperature [Rees, 1989]. Some of the difference between SCIM and IRI calculated critical frequencies when the SZA is greater than 100 degrees may be due to neglecting horizontal transport of plasma in the SCIM calculations. This could result in underestimating the nighttime critical frequency since convective transport of *F*-region plasma over the polar cap from the dayside ionosphere is a significant source of plasma in the evening sector of the nighttime *F*-region ionosphere [Maurits, 1996].

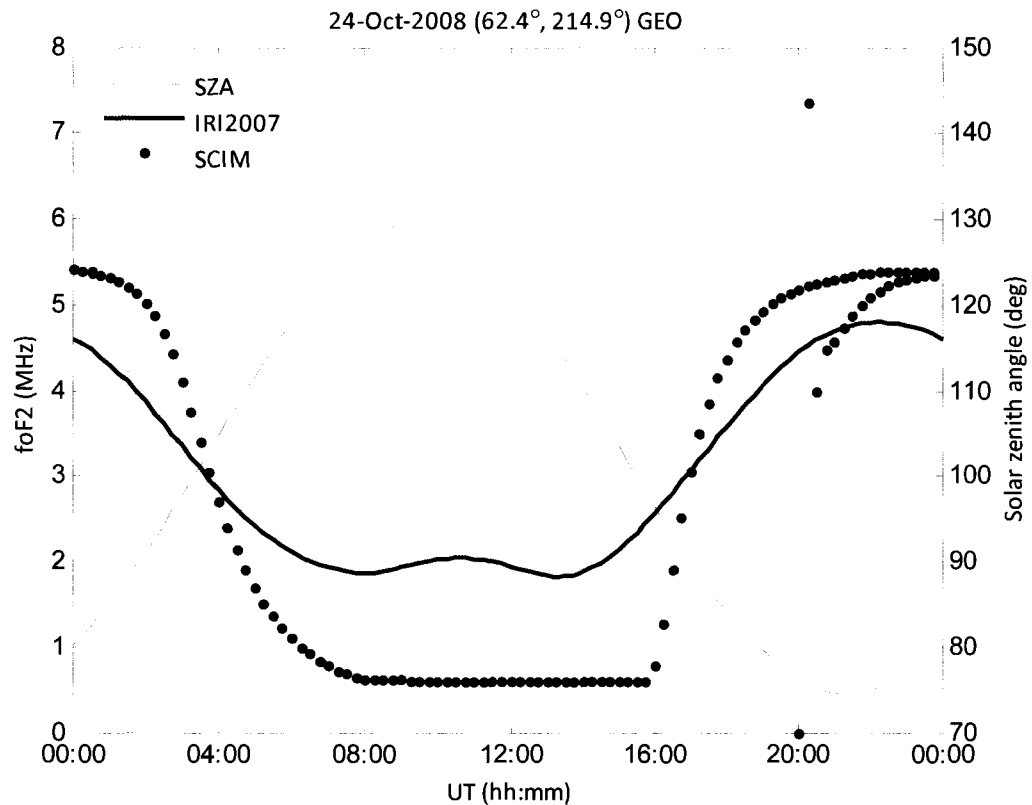


Figure 3.6 Time-dependence of the ionosphere critical frequency over Gakona, Alaska on 24 October 2008 calculated with the SCIM model (dotted) and with the IRI 2007 empirical model (solid). The solar zenith angle (dashed) is plotted against time for reference. Each dot in the SCIM model output represents 15 minutes of simulation time, the red dots correspond to the first 24 hours of simulation time following an initial ionization-free atmosphere at 20:00 UT and the overlaid black dots correspond to the following 48 hours of simulation time.

Each black and red point in Figure 3.6 corresponds to the critical frequency calculated during a single time step of the SCIM model. Each time step is separated by 900 seconds. A full 72 hours of simulation time is plotted in the figure, the first 24 hours of simulation calculations following a cold-start plasma-free ionosphere, are indicated by red points and the following 48 hours of calculations are indicated by overlaid black points. Note that the critical frequency calculated by the SCIM model immediately following the initial plasma-free atmosphere at 20:00 UT increases rapidly at about 2.5 MHz per hour for the first two hours, about the same rate of increase during the subsequent sunrise, before the rate of critical frequency increase begins to slow significantly. The model overshoots the critical frequency, relative to the IRI model output, in the second time

step at 20:15 UT due to an initial overproduction of molecular oxygen ions and nitric oxide ions: the electron number density used in the chemical production and loss calculation is based on the electron density in the previous time step and the dominant loss process of these ions is the (chemical) recombination with ambient electrons. After approximately four hours following the initial conditions, the critical frequency calculated by SCIM is almost exactly equal to the critical frequency calculated 24 hours later. Therefore the initial conditions of the model do not affect the critical frequency calculated from the SCIM diurnal equilibrium solution beyond four hours of simulation time.

The diurnal variation of foF_2 as calculated by SCIM and the IRI model is compared with the variation of foF_2 measured by the HAARP ionosonde in Figure 3.7. At each SCIM-model time step, the calculated critical frequency is represented in Figure 3.7 by a red point; exactly 24 hours of 900-second time steps are shown. Model calculations generally agree well with the ionosonde observations, the main differences between the model calculations and ionosonde observations are

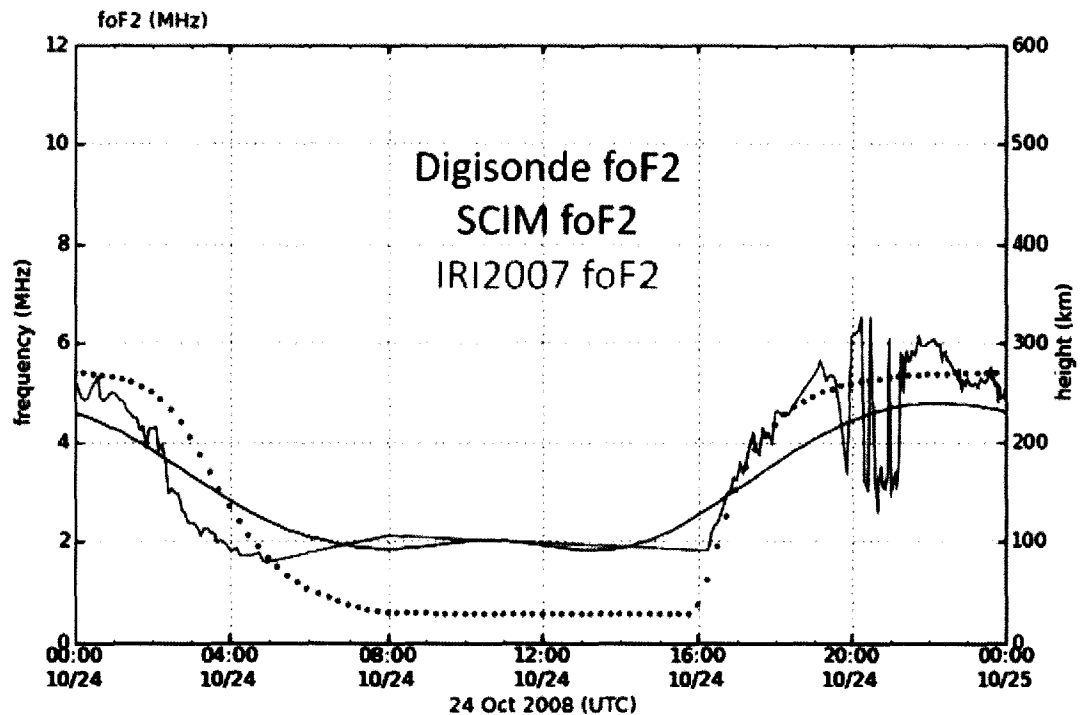


Figure 3.7 Ionosphere critical frequency vs. universal time over Gakona, Alaska on 24 October 2008 as measured by the HAARP ionosonde (blue), as calculated by the self-consistent ionosphere model (red), and as calculated by the IRI 2007 empirical model (cyan).

that SCIM underestimated the critical frequency at night and the IRI model underestimated the critical frequency during the day. Note that between approximately 04:00 and 16:00 UT, when the solar zenith angle is above 100 degrees and there is little direct photoproduction of electrons, the maximum electron density is not sufficient for the HAARP ionosonde to accurately measure the ionosphere critical frequency.

3.2.3 *Diurnal variation of daytime TEC*

In this section, the diurnal variation of the TEC calculated by SCIM is compared to the diurnal TEC variation calculated by IRI2007 and the “slab-equivalent” vertical TEC variation measured by the HAARP GPS receiver. The total electron content (TEC) is number of electrons per unit area (typically a square meter) above some reference location and can be calculated from model results by line-integrating the electron number density profile with respect to altitude. The time required for the SCIM-calculated TEC to converge to a periodic diurnal cycle will be discussed. As with the previous sections in this chapter, the numerical and observational experiments are focused on the day of 24 October 2008 at Gakona, Alaska near the HAARP facility.

The TEC below 1000 km altitude as a function of universal time calculated by SCIM and the IRI model is plotted in Figure 3.8. In the figure, a TEC calculation at a single SCIM time step is represented as a black or red point and the time-dependent TEC calculated by the IRI model is plotted as a blue line. Since the total electron production rate in the ionosphere, and thus the total electron content, is limited by the number of ionizing solar photons available in the *E* and *F* regions, the solar zenith angle as a function of universal time is also plotted in Figure 3.8 as a dashed black line for additional reference. The SCIM-calculated TEC increases at a rate of about 2×10^{16} electrons per square meter per hour during local morning as the solar zenith angle decreases below 100° elevation, near the angle at which the sun is observed to rise from a vantage point several hundred kilometers above the earth surface. The TEC then reaches a broad maximum of about $6 \times 10^{16} \text{ m}^{-2}$ lasting for approximately four hours and, after one or two hours beyond solar zenith, begins to decrease at nearly the same rate of morning TEC increase. Both SCIM and the IRI model-calculated TEC time series are in good agreement during the daytime, where the IRI estimate of maximum TEC is about 20% larger than and occurs less than one hour earlier than the maximum SCIM TEC estimate. As discussed earlier in section 3.2.2, SCIM neglects to include the horizontal transport of long-lifetime atomic oxygen ions from the dayside over the polar cap to the nighttime polar ionosphere so the SCIM-model estimate of TEC is less than the estimate by the IRI

model. Uncertainty in the chemical reaction rates used in the SCIM calculations will also cause uncertainty in the diurnal TEC calculations.

The first 24 hours of a SCIM-calculated TEC time series plotted in Figure 3.8 are represented as a sequence of red points and the overlaying black points represent the following 48 hours of TEC calculations, similar to critical frequency time series plotted in Figure 3.7. An initial overshoot of the SCIM-calculated TEC during the first time step following the initial cold-start plasma-free ionosphere condition at 20:00 UT is consistent with the overshoot of the critical frequency discussed in section 3.2.2. The cause of the overshoot was found to result from the model underestimating the electron loss rates during the first time step; that is, the electron loss rates are null in a plasma-free ionosphere. Unlike the monotonic convergence of the SCIM-calculated critical frequency, the calculated TEC oscillates around a path that converges to the diurnal TEC cycle during the first four hours of simulation time. After the model-calculated TEC converges to the diurnal cycle, no significant oscillations are apparent even during sunrise and sunset when the TEC time derivative is maximized. The TEC convergence time can be seen to be about one or two hours longer than the critical frequency f_oF_2 in Figure 3.7 and Figure 3.8. The reason may be due to transport of ions to the topside ionosphere that continues even after the electron number density and loss rates reach a maximum. In the topside ionosphere, chemical loss rates are smaller than the loss rates near the critical layer, indicating that transport effects play a more significant role in determining the TEC than the critical frequency.

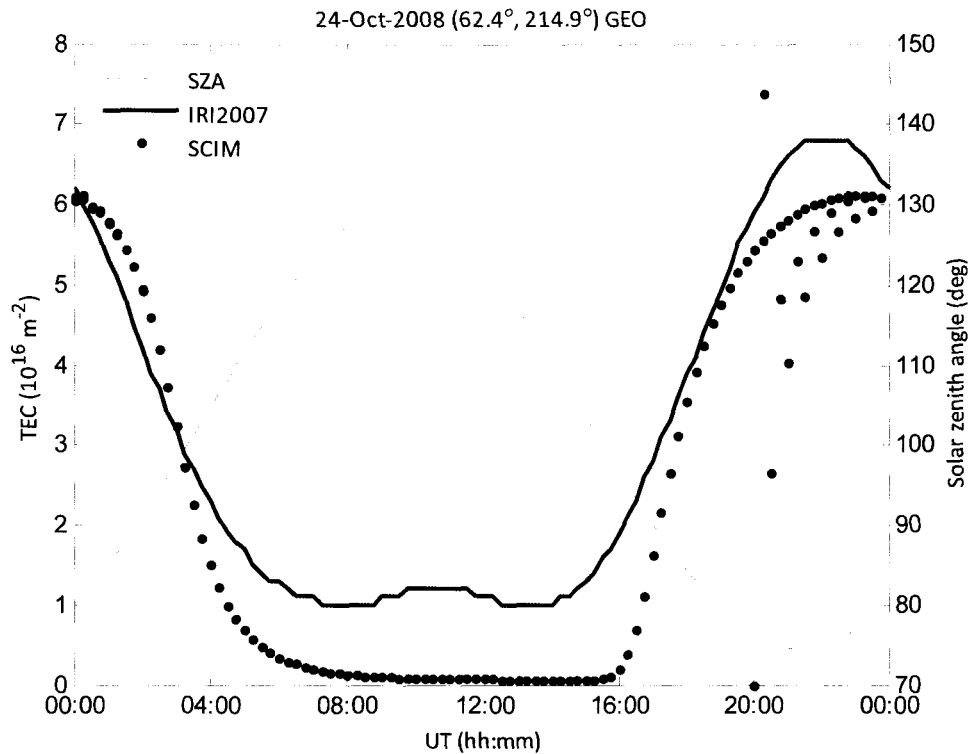


Figure 3.8 Time-dependence of the ionosphere total electron content (TEC) between 100 and 1000 km altitude over Gakona, Alaska on 24 October 2008 calculated with the SCIM model (dotted) and with the IRI 2007 empirical model (solid). The solar zenith angle (dashed) is plotted against time for reference. Each dot in the SCIM model output represents 15 minutes of simulation time, the red dots correspond to the first 24 hours of simulation time following an initial ionization-free atmosphere at 20:00 UT and the overlaid black dots correspond to the following 48 hours of simulation time

The SCIM TEC results were compared with the TEC measurements made with a GPS receiver at the HAARP facility. The time required for a radio wave to travel through the ionosphere depends on both the frequency of the wave and the total electron content along the path of the wave. GPS position measurements are sensitive to the time required for a radio signal to propagate from a GPS satellite to a GPS receiver, therefore temporal-spatial variations in the ionosphere electron density can cause variations in location determined by the GPS receiver. To partially compensate for ionospheric variability, GPS satellites broadcast simultaneously on two frequencies so that by measuring the difference in satellite-receiver distances calculated from each received frequency, the total electron content along the signal propagation path can be inferred and used to correct the range measurement. The measured *slant TEC*, or the total electron content per unit area between

an observer on earth and a line-of-sight GPS satellite, at a typical orbital of approximately 20,000 km altitude, is a useful byproduct of the GPS range correction technique. Multiple slant TEC measurements may be combined to infer the *slab equivalent TEC*, or the TEC directly above a point of observation, as shown in Figure 3.9 where the SCIM- and IRI model-calculated TEC has been superimposed on the instrument output. The slab-equivalent TEC measurement in Figure 3.9 is constructed with measurements from GPS satellites that passed 45° or more above the horizon from the location of the HAARP GPS receiver. A diurnal variation is clearly evident in the measured slab-equivalent TEC, obtaining a maximum value of approximately $9 \times 10^{16} \text{ m}^{-2}$ at about 21:00 UT and a minimum value of $2 \times 10^{16} \text{ m}^{-2}$ just before local sunrise at 16:00 UT.

Both the SCIM and IRI models significantly underestimate the TEC relative to the slab-equivalent TEC measured by the HAARP GPS receiver. One reason for this is that the TEC calculated by SCIM (red points) and the IRI2007 model (solid blue curve) accounts for only the electrons below 1000 km altitude while the measured slab-equivalent TEC is based on electron counts between the observer on the ground and GPS satellites in orbit at approximately 20,000 km altitude. The TEC below 2000 km, as calculated by the IRI model, is illustrated in Figure 3.9 (dashed blue curve) and is about $0.3 \times 10^{16} \text{ m}^{-2}$ greater than the TEC calculated by IRI below 1000 km at 21:00 UT. The total electron content between 2000 and 10,000 km altitude can be estimated by taking the average electron density of $1 \times 10^9 \text{ m}^{-3}$ at 10,000 km altitude above the polar cap, measured by radio plasma imager on the IMAGE spacecraft [Nsumei *et al.*, 2008] during quiet solar and geomagnetic conditions, and assume that this is representative electron density in the entire region so that an additional 0.8×10^{16} electrons per square meter are in the column between 2000 km and 10,000 km altitude. Therefore the TEC calculated by the IRI and SCIM models agree reasonably well with observed slab-equivalent TEC measured by GPS. SCIM significantly underestimated the local nighttime TEC relative to the value calculated by IRI and measured by the GPS receiver for reasons discussed earlier. However, note that SCIM reproduces the monotonic decrease in TEC from 06:00 to 16:00 UT observed in the GPS-measured TEC but not in the IRI-calculated TEC. Uncertainty in the photoproduction and chemical loss rates of atomic oxygen ions in the *F* region can also significantly affect the TEC calculated with a self-consistent model. The focus of this work is on the overall response of the model ionosphere to changes in the input drivers like the solar zenith angle or neutral gas number density.

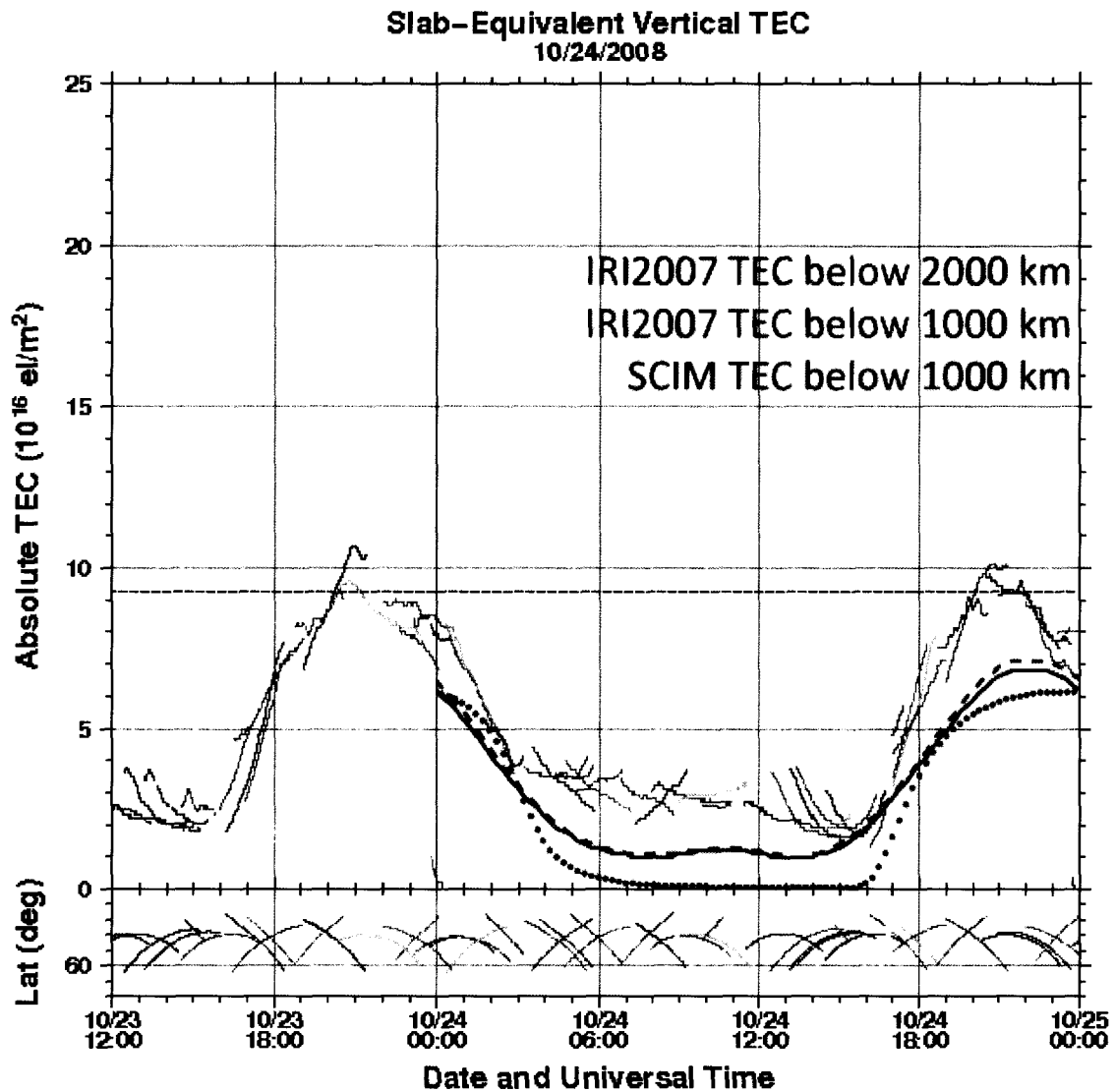


Figure 3.9 HAARP GPS measurements of the time-dependent total electron content (TEC) over Gakona, Alaska on 24 October 2008. The TEC below 1000 km as calculated by SCIM (red) and the IRI 2007 model (solid blue) are overlaid for comparison, the TEC below 2000 km as calculated by IRI (dashed blue) is also shown. Short line segments (various colors) represent the GPS data from individual satellite passes.

3.2.4 Model sensitivity to semidiurnal thermal tides

Ionizing solar photons directly incident on the neutral atmosphere is the dominant electron production mechanism in the daytime ionosphere. The photoionization rate at a given location in

a slab-symmetric ionosphere depends on the energy distribution of the incident radiation, the solar zenith angle, and the major neutral gas species number density altitude profiles. An intuitive validation of the SCIM calculations is presented in the form of a sensitivity analysis in this section. The time-dependent response of the modeled column-integrated electron photoproduction rate and the column-integrated electron thermal energy density to diurnal variation in the solar zenith angle and a tidal semidiurnal oscillation in the major neutral gas number density will be discussed.

To first order, the diurnal variation of the height-integrated—or *column*—photoproduction rate can be related directly to the solar zenith angle. The column neutral gas number density determines the diurnal variation of the column photoproduction rate to second order. Here, the column number density $\hat{\rho}$ of a volume number density ρ will refer to the line integral the volume number density along the zenith between 100 and 1000 km altitude: $\hat{\rho} = \int_{100 \text{ km}}^{1000 \text{ km}} \rho dz$. Figure 3.10 shows the electron column production rate (red points) and neutral gas column number density (blue points)—over Gakona, Alaska on 24 October 2008—as parametric functions of solar zenith angle. There is one point per 15-minute time step in each 24-hour diurnal equilibrium simulation cycle or orbit. Points calculated from time steps occurring between 00:00 – 11:45 UT are marked with x's and points falling between 12:00 – 23:45 are marked with o's. Universal time increases in the clockwise direction along the red curves and along the left half of the blue curves (UT increases in the counterclockwise direction on the right half of the blue curve). The neutral gas column number density is separated into two components: the dark blue curve shows the semidiurnal variation of the sum of molecular nitrogen and molecular oxygen column number densities, and the light blue curve shows the diurnal variation of the column number density of atomic oxygen, scaled by a factor of five for clarity.

Starting with the column electron photoproduction rate in Figure 3.10, notice that the rate increases approximately $5 \times 10^{12} \text{ m}^{-2} \text{ s}^{-1}$ for each degree decrease in the solar zenith angle for angles less than 95° , and is nearly identically zero for angles greater than that. Electron photoproduction maximizes at a rate of $90 \times 10^{12} \text{ m}^{-2} \text{ s}^{-1}$ at the time of minimum solar zenith angle when the elevation of the sun is 25° . The column electron production rate due to impact from energetic photoelectrons is slightly less than 50% of the photoproduction rate. During the daytime at fixed solar zenith angle, the column photo and impact electron production rates are smaller before local noon than it is after local noon. This difference in photoproduction rates at equal solar zenith

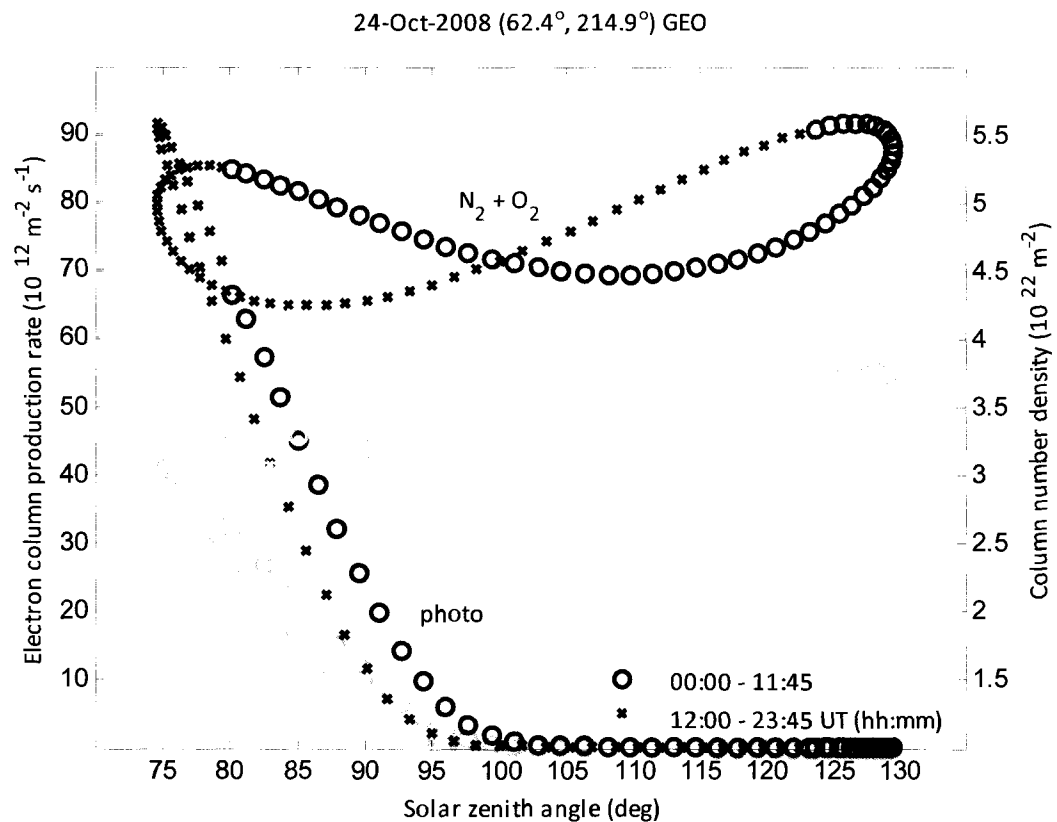


Figure 3.10 Height-integrated electron volume photoproduction rate (red) and height-integrated volume number density of molecular nitrogen and oxygen (blue) as parametric functions of solar zenith modeled over Gakona, Alaska on 24 October 2008. The electron column production rate due to impact with non-thermal photoelectrons (light red) and the column number density of atomic oxygen (light blue) are shown for comparison. The column number density of atomic oxygen has been scaled by a factor of five.

angles is due to the 12-hour semidiurnal variation in the lower thermosphere neutral gas number density, caused by periodic solar heating. These high-latitude semidiurnal *thermal tides* are reproduced in the MSIS model output [Sharp *et al.*, 1978] are apparent in the blue curves representing column number density. SCIM-calculated electron production rates are significantly affected by this tidal oscillation since the photon energy distribution is held constant throughout the simulation.

Figure 3.11 shows the height-integrated electron and ion volume heating rates and thermal energy densities as parametric functions of solar zenith angle at each simulation time step. As in Figure 3.10, universal time increases in the clockwise direction around each curve; noon and midnight (UT) can be identified by the time direction and transition in plotted point style. The electron column heat rate plotted in the figure is only the portion of the total electron heating rate that is due to elastic collisions with energetic photoelectrons. The ion column heat rate plotted in the figure is the portion of the total ion heating rate due to Joule heating by ion currents driven by the perpendicular electric parameter that describes the magnetosphere-imposed electric field. Column thermal energy densities were calculated by integrating the volume thermal energy densities

$$\epsilon_{\{e,i\}} = \frac{3}{2} k_B T_{\{e,i\}} n_{\{e,i\}}$$

of the ion and electron gases between 100 km to 1000 km altitude.

The column electron heat rate in Figure 3.11 clearly follows the column electron photoproduction rate in Figure 3.10 indicating that, as expected, the major heat source of the electron gas is collisions with energetic photoelectrons. In particular, note that the hysteresis in the column electron photoproduction rate caused by semidiurnal thermosphere tidal oscillation is also evident in the column electron gas heating rate. Although it is not visible in the plot because of the choice of vertical axis scale, thermal tide hysteresis also exists in the ion Joule-heating curve because the heating rate depends on the ratio of the ion-neutral collision frequency to the ion cyclotron frequency through the Pedersen conductivity term in equation (2.53); also the ion-neutral collision frequency is nearly proportional to the major neutral gas number density. Even on the vertical axis scale -- chosen here to emphasize that elastic collisions with thermal electrons, not Joule heating, is the major source of ion heating in the model -- it appears that ion Joule heating by electric fields ceases when the solar zenith angle increases to the point that there is no significant electron photoproduction. However, the electron photoproduction rate does contribute, albeit indirectly, to ion gas Joule heating because the ion Joule-heating rate described by equation (2.53) is proportional to electron density. As illustrated in section 3.2.2, the peak electron number density

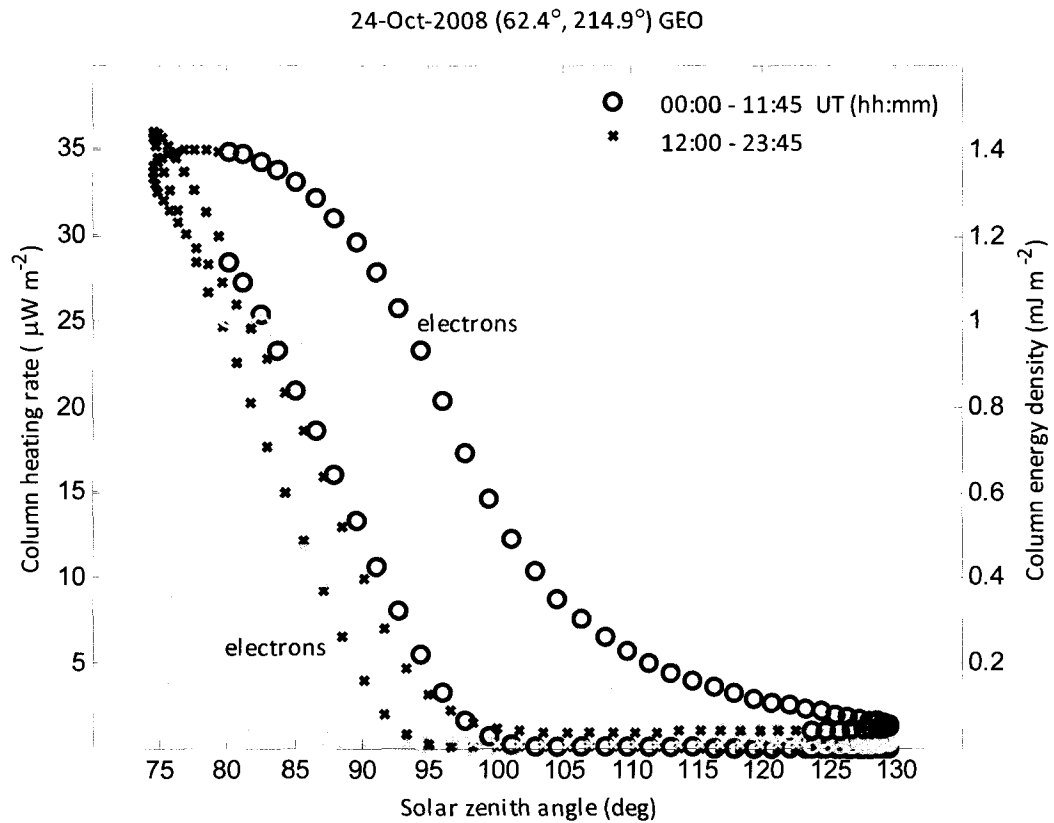


Figure 3.11 Parametric dependence on solar zenith angle of height-integrated electron volume heat transfer rate (dark red) due to collisions with energetic photoelectrons and of height-integrated electron volume thermal energy density. The column heat transfer rate of the ion gas (light red) due to Joule heat dissipation from a perpendicular magnetospheric electric field of 1.0 mV/m and the column thermal energy density of the ion gas (light blue) are provided for comparison.

decreases by an order of magnitude shortly after sunset so ion-Joule heating decreases after sunset as well even if the perpendicular electric field is held constant.

The large hysteresis in the height-integrated electron gas thermal energy density, however, is not caused by semidiurnal thermal tides; rather, it is due to the decreased rate of cooling through elastic collisions with the increasingly hot ion gas. Electron-ion energy transfer, described by equation (2.48), varies with the temperature difference between the electron and ion gases. Elastic collisions with the relatively cool major neutral gases is the major mechanism of ion gas cooling while elastic collisions with the ion gas is a significant mechanism of electron gas cooling. The ion

gas cools more slowly than the electron gas because the ion-neutral collision rate is much less than the electron-neutral collision rate and the ion-electron collision rate. Therefore, elevated ion gas temperatures diminish the electron gas cooling rate, and the ion gas temperature is slow to decrease relative to a gas of hot electrons amidst cool ions and neutrals. This is manifest in electron and ion column thermal energy densities that are over a factor of two larger in the local morning than local evening at times of equal solar zenith angle. Thermal tides are not responsible for this hysteresis effect because the neutral gas density increases from local morning to evening thereby increasing the ion and electron gas cooling rates via collisions with the major neutrals and actually diminishing the elevated column energy densities observed in the model calculations.

Chapter 4. Ascent of the ion line

Density irregularities in the polar ionosphere affect regional radio communication, radar observation, and positional triangulation by receiving radio transmissions from global positioning satellites. Powerful HF waves transmitted with O-mode polarization are known to cause significant plasma turbulence in the ionosphere in the vicinity where the plasma frequency is near the transmitted HF wave frequency [Dubois *et al.*, 1993; Rietveld *et al.*, 2000]. The HF-enhanced ion line (HFIL), measured with UHF or VHF radar during HF modification experiments may be used as a diagnostic to infer the magnitude and direction of HF-induced plasma density perturbations.

The HFIL is reported to descend during certain polar ionosphere HF-modification experiments, yet it is reported to ascend in other experiments. (See, e.g., Minkoff *et al.* [1974] for a report of the first HFIL observations made at the Platteville Atmospheric Observatory in Colorado.) HFIL measurements made during O-mode HF modification of the daytime polar ionosphere over the Tromsø HF heater in Norway showed that the HFIL descends by about 10 km during heating [Ashrafi *et al.*, 2006; Djuth *et al.*, 1994]. However, an ascending HFIL has been reported throughout HF pumping during HAARP experiments in Alaska [Kosch *et al.*, 2007]. And Djuth *et al.* [1994] observed “billowing altitude structures” with UHF and VHF radar following tens of seconds of HF pumping. Variations of the HFIL reflection height can be measured with high spatiotemporal resolution with a relatively modest radar and serve as a sensitive indicator of HF-induced plasma density perturbations that would otherwise be difficult or expensive to measure remotely with a large incoherent-scatter radar or in-situ with rockets or low-earth orbit satellites.

The objective in this chapter is to: 1. describe the time variation of the HF-enhanced ion-line reflection altitude during O-mode HF heating of the daytime polar ionosphere, 2. explain the physical mechanisms causing the variation, and 3. quantify the magnitude and direction of the underlying plasma density perturbations. UHF radar has been used to measure the HFIL altitude during HAARP experiments with a time resolution of 1 s, and the self-consistent ionosphere model (SCIM) simulated HF heating in a slab-symmetric ionosphere. Comparison of HFIL reflection height measurements and model results are used to infer the magnitude of the density perturbations observed in the experiments. Much of the observed time variation of HFIL altitude may be explained by altitude-dependent density perturbations that persist with altitude-dependent characteristic times.

4.1 Overview

Experiments described in this chapter use UHF radar to detect time-dependent plasma density perturbations caused by strong O-mode HF transmissions, and more generally, to test whether the field-aligned plasma intersecting the HF interaction region can be described by the one-dimensional slab-symmetric SCIM ionosphere.

Two experiments are reported here; both were conducted at the HAARP facility during early evening (daytime) ionosphere in late October. The first experiment was conducted on 24 October 2008 and the second experiment was on 21 October 2009. The objective of the first experiment was to detect density irregularities over a field-aligned range of 10 km surrounding the HF reflection height, that is, the ionospheric region where the plasma frequency equals the transmitted HF wave frequency. The spatial extent of HF-induced density modifications may be measured by transmitting a sequence of short HF pulses at various frequencies following ionosphere modification, so that each pulse reflects from a different altitude. Spatial resolution was sacrificed for increased time resolution in the second experiment, where a sequence of short HF pulses at a single frequency followed an ionosphere-modifying HF transmission. Nearly one year of exceptionally low solar activity separated the two experiments, so the geophysical conditions present during each experiment are very similar.

Prior experiments indicate that strong HF waves incident on the ionosphere can result in a rapid conversion of a portion of the HF electromagnetic energy to electron thermal energy near a layer where the electromagnetic wave and ionospheric plasma satisfy certain resonance conditions. The resulting change of electron temperature in the polar ionosphere can lead to detectable electron density perturbations [Frey *et al.*, 1984]. In the case of *O-mode* HF transmissions—that is, HF waves that are elliptically polarized such that the electric field rotation is in the opposite direction as the electron gyro-rotation about the geomagnetic field—the resonance conditions generally occur in the F region when the transmitter frequency is near the ionosphere critical frequency. Heat deposited in the F-region electron gas conducts along the geomagnetic field, increasing the electron temperature over a field-aligned distance of hundreds of kilometers within tens of seconds. Ambipolar diffusion and bimolecular chemical reactions with the neutral gases are important temperature-dependent processes governing the vertical distribution of ionospheric plasma. Modulation of the F-region electron temperature leads to modulation of ionospheric plasma and composition along the modification region. Since an upper bound on the power input

into the ionosphere via strong HF electromagnetic transmissions is known, controlled HF modification experiments provide a means to compare model predictions of the ionospheric response to modulated electron heating rates. Similarly, simulations of ionosphere heating experiments can be used to estimate the effective electron HF-heating rate by adjusting the model HF-heating rate until the temperature and density profiles of the model modified ionosphere agree with the observed profiles.

The time-variation of the HFIL reflection height, measured during short HF pulses, is used in both experiments to infer changes in the ionosphere electron density and to relate the measurements to physical mechanisms by through comparisons with the SCIM-calculated reflection heights during simulated heating experiments in a slab-symmetric ionosphere. SCIM simulations of HF heating a 2-km thick layer near the critical layer exhibit the following electron-density perturbations with altitude-dependent characteristic times: 1. A “notch” of plasma depletion in the electron density profile occurs in the 2-km layer where heat is deposited in the electron gas by the HF wave. The calculated density depletion of the notch relative to the surrounding plasma may be up to 20%. 2. Density enhancements occur in the E and lower F regions during HF heating that dissipate within tens of seconds after the HF electron heating ceases. 3. A density depletion of up to 20% in a 100 km layer surrounding the F2 peak forms within minutes of the start of HF heating and persists for up to tens of minutes following the end of HF heating. 4. Density enhancements of 20% occur in the topside ionosphere form, and are maintained at 20% over the equilibrium value, within minutes of continuous HF heating and persist for minutes following the end of heating.

4.2 Method

The Modular UHF Ionospheric Radar (MUIR) at the High-frequency Active Auroral Research Program (HAARP) facility near Gakona, Alaska can observe ion-line radar returns enhanced by strong HF transmissions broadcast by the HAARP IRI [Oyama *et al.*, 2006]. The HFIL return—observed with MUIR—is localized along the trajectory of the HF electromagnetic wave where the frequencies of the HF and UHF waves, as well as the electron number density and plasma temperature, satisfy the altitude-dependent HFIL matching condition, defined below. The range to the HFIL and its time variation can be used as diagnostics of the ambient ionospheric parameters and the response of those parameters to strong HF electromagnetic waves incident for several seconds or longer. The UHF-observed response of the HFIL to changes in the ionospheric plasma

can then be compared to changes in a simulated ionosphere by calculating the location where the HFIL resonance condition is satisfied by the model-calculated ionospheric parameters.

Calculations from the 1-D self-consistent ionosphere model (SCIM), a time-dependent Eulerian model ionosphere profile fixed in the co-rotating frame of Earth, are used to describe the altitude-variation of the HFIL reflection height and infer the magnitude of electron temperature and density perturbations necessary to produce the observed HFIL reflection height variation. A 1-D ionosphere model is limited in its reproduction of a physical ionosphere profile by the magnitude of ionospheric convection across the volume of interest; so convection velocities, observable with the HAARP ionosonde, during the experiments need to be low enough such that ionospheric takes several minutes to drift across the HF-interaction region over HAARP. Electric fields of magnetospheric origin drive convection of the ionospheric plasma across the UHF and HF interaction volumes that are typically fixed over the Earth surface. Therefore, quiet geophysical conditions characterized by low geomagnetic activity, a weak perpendicular electric field with small plasma convection velocities, and minimal incident energetic particle precipitation are desired in order to maximize the usefulness of any physical inferences derived from model calculations. In section 4.2.1, the geophysical conditions observed during the time of the experiment are described, and their relation to the geophysical parameters used to initialize the SCIM calculations will be discussed.

The response of the ionosphere to powerful HF transmissions depends on the HF frequency, HF wave polarization, the effective radiated power (ERP) and antenna gain of the HF transmitter, and the HF transmission modulation scheme. Long-duration HF transmissions lasting tens of seconds to several minutes are used to cause significant modifications to the ionosphere plasma temperature and density. Short-duration HF transmissions lasting between 1-100 ms are used to probe the ionosphere by creating plasma turbulence detectable with MUIR. The UHF radar can detect ion-acoustic (IA) plasma waves—enhanced by strong HF transmissions—at an altitude where the *matching condition* is satisfied. The matching condition relates the plasma density and temperature in the interaction region to the frequencies of the HF and UHF electromagnetic waves, the range to the HF-enhanced ion-line (HFIL) target detected with MUIR provides a diagnostic for identifying ionosphere plasma density and temperature changes. The sequences of modifying and probe HF transmissions broadcast by the HAARP IRI are described in section 4.2.2.

It is through the matching condition, discussed in section 4.2.3, that ionosphere simulation results are related to the UHF measurements of the range to the HFIL and its time variation. The spatial resolution of HFIL range measurements made with UHF radar depends on the radar transmitter frequency, the ERP and antenna gain of the transmit/receive antenna, the receiver bandwidth and sample rate, the UHF transmission modulation scheme, and the algorithm used to process the baseband signal of each received pulse recorded by the MUIR system. These parameters, as well as the details of the algorithm used to calculate the range to the HFIL return, are also described in section 4.2.3. Validation of the UHF observations is provided by calculating the range and frequency spectrum of a return from a satellite that passed through the radar volume during experiments performed at the HAARP facility in March 2009, and by range-time comparisons of HFIL intensity with HF-induced airglow measured along the HAARP geomagnetic field.

Customizations made to the SCIM software in order to simulate ionosphere HF-modification experiments will be discussed in section 4.2.4. Validation of the ionosphere HF-modification simulations will be provided by comparing the calculated time-dependent electron temperature response with published electron temperature observations made with incoherent-scatter radar during HF pumping of the high-latitude ionosphere. The matching condition parameterization used to describe location of the HFIL return observed with MUIR will be discussed. SCIM calculations are used to infer the magnitude of HF-induced ionospheric plasma perturbations from measurements of the HFIL range time-variation made with MUIR through the HFIL resonance parameterization.

4.2.1 Geophysical conditions

A quiet, undisturbed ionosphere is necessary to identify variations observed in the HFIL matching height resulting from heater-induced perturbations from those that are due to natural ionospheric variability. Small $\mathbf{E} \times \mathbf{B}$ drift ion-velocities are necessary to maximize the time that a particular field-aligned volume of plasma remains in the modifying HF transmission volume. Small ion drift velocities will also minimize the number of field-aligned irregularities that drift into the observation volume. Variations in the neutral wind velocity can cause variations in the field-aligned distribution of ionospheric plasma, so steady neutral winds are essential. All of these conditions that reduce uncertainty in the experiment measurements also make it easier to relate SCIM simulation results to the measurements. The experiments described in this chapter were conducted at the HAARP facility near Gakona, Alaska on 24 October 2008 and 21 October 2009,

Table 4.1 Geophysical conditions observed during ionosphere HF modification experiments at Gakona, Alaska.

Measurement	Units	Experiment	
		24 Oct 2008	21 Oct 2009
UT	hh:mm:ss	00:10:00 — 00:19:30	01:43:00 — 01:59:30
foF2	MHz	4.0 — 4.5	4.06
hmF2	km	200 — 250	220 — 270
Ion velocity N	m/s	40 — 60	75 — 125
Ion velocity E	m/s	-30 — 30	unknown
Ion velocity UP	m/s	0 — -10	0 - 20
Ap	nT	1	1
F10.7	$10^{-22} \text{ W m}^{-2} \text{ Hz}^{-1}$	66.7	70.6

times of exceptionally quiet geomagnetic activity. Measurements of ambient ionospheric and thermospheric conditions made during the experiments are compiled in Table 4.1 and will now be described.

Geomagnetic activity—as described by the Geomagnetic Averaged Planetary Index (Ap)—was very low during October 2008 as illustrated in Figure 4.1, and during much of the following year. On 24 October 2008 and 21 October 2009, the Ap index was approximately 1 nT. Enhancements of the Ap index are correlated with thermospheric density and temperature enhancements [Forbes *et al.*, 1993] and ionosphere foF2 enhancements [Lal, 1992]. Ap index effects are manifest in the SCIM calculations through neutral atmosphere density and relative composition changes as output by the MSIS model [Hedin, 1987]. Likewise, the 10.7 cm Solar Radio Flux (F107), also shown in Figure 4.1 for October 2008, was low. F107 effects are incorporated in the SCIM calculations through the Hintereger solar flux parameterization. As described by Hedin [1987], increasing the F107 index input to the model generally increases the MSIS neutral atmosphere density and temperature. The low values of the Ap and F107 indices recorded during both October experiments imply that thermosphere and ionosphere densities, temperatures, and winds over the HAARP facility were lower than typical values during the time of the experiments.

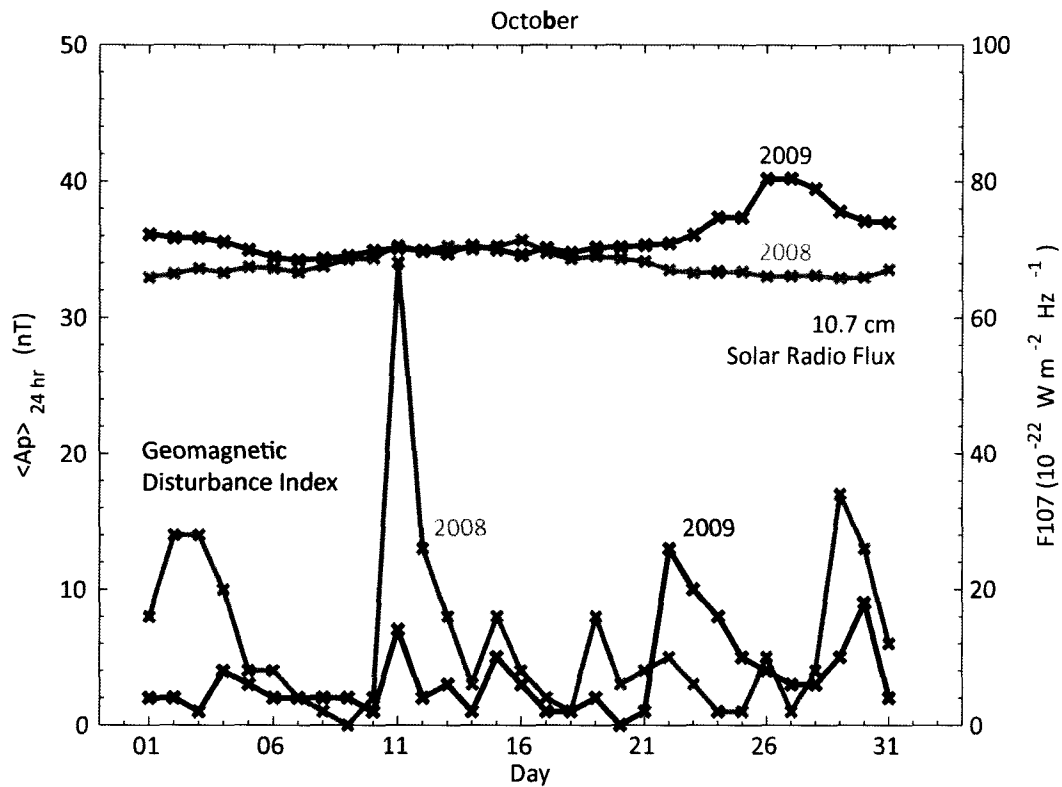


Figure 4.1 Daily measurements of the Geomagnetic Averaged Planetary Index (A_p) and the 10.7 cm Solar Radio Flux (F107) during October of 2008 and 2009.

The HAARP facility and the SCIM-simulated ionosphere co-rotate with Earth; but the ionosphere plasma may at times convect over the surface with a ground speed of hundreds of meters per second. Plasma $E \times B$ drift—driven by electric fields of magnetospheric origin—is the main mechanism of ionosphere plasma convection at high latitudes. Minimizing the ion drift velocity during experiments maximizes the time that a particular field-aligned volume of plasma will be affected by HF transmissions from the HAARP IRI. Ionosphere convection velocities over the HAARP facility during the experiments were measured with the HAARP ionosonde and are summarized in Table 4.1. Plasma horizontal drift speeds measured by the ionosonde were generally around 100 m/s for the duration of the experiments, and the vertical velocity was approximately 10 m/s in either direction. At transmit frequency of 3 MHz, the HAARP IRI HF half-power beam width is slightly greater than 50 km at a typical HF interaction height of 200 km, as described in

section 4.2.2, so a volume of plasma at the HF interaction height drifting at 100 m/s may spend up to 500 s in the most intense portion of the HF beam.

A Fabry-Perot interferometer—an instrument for measuring thermospheric neutral winds described by *Conde and Smith* [1995]—was operating at Poker Flat, Alaska on the night of the 24 October 2008 experiment and at the HAARP facility on the night of the 21 October 2009 experiment. HAARP experiments in Gakona were conducted during the late afternoon but the interferometer at Poker Flat is a night-time instrument needing dark conditions to observe the airglow of atomic oxygen, so the thermospheric wind velocity measurements beginning at about 03:00 UT are not directly relevant to the experiment. However, wind velocity measurements over Poker Flat, located further north and nearer to the auroral zone than Gakona, can help to confirm that the ionosphere and thermosphere over the HAARP facility were quiet and undisturbed during the experiments. Unfortunately, near-overcast conditions prevented reliable and continuous measurements of neutral wind velocity on the dates of the experiments. However, observations made on 21 October 2009 do show solar diurnal tidal winds, with no signature of ion drag forcing. The low value of the A_p geomagnetic disturbance index is also consistent with steady neutral wind conditions.

The thermospheric wind velocity, measured by observing the atomic oxygen red line emission above Poker Flat with the Scanning Doppler Imager (SDI), was observed to be steady -50 m/s in the magnetic zonal direction from 0400-1400 UT on 24 October 2008, with typical variations on the order of 10 m/s occurring over an hour. Low and steady neutral wind speeds are consistent with the low value of A_p recorded at the time and are considered favorable conditions for this experiment and model comparisons. SDI measurements of neutral wind velocity over Poker Flat are not available for 21 October 2009.

The diurnal variations of the measured foF₂ and vertical TEC above the HAARP facility on 24 October 2008 are plotted in Figure 3.6 and Figure 3.8, respectively. These measurements are representative of the respective measurements made during the second experiment on 21 October 2009. Electron density profiles derived from HAARP ionograms are presented in section 4.3.1.

A scatter plot of plasma density and temperature directly over Poker Flat measured with the Poker Flat ISR in the hours before the 24 October 2008 experiment is shown in Figure 3.4. As with the neutral wind measurements made with the Poker Flat Fabry-Perot interferometer on 24 October

2008, the thermosphere and ionosphere conditions over Poker Flat are assumed to be representative of conditions over the HAARP facility given the proximity of the two measurement locations, daytime conditions, and extremely low geomagnetic activity. No Poker Flat ISR measurements of the vertical ionosphere density or temperature profiles on 21 October 2009 are available.

The measurements described thus far describe conditions of the F-region ionosphere—the region of interest in this thesis—but the D and E regions are at times significant sinks of HF electromagnetic energy, thus preventing much of the transmissions from the HAARP IRI from reaching the F region. These absorption events are a common at high latitudes. The HAARP VHF riometer measured the total intensity of 30 MHz electromagnetic radiation incident from space that reached the ground during the October experiments. Times of enhanced D and E region absorption are indicated by a drop of RF power measured by the riometer from the power expected to be measured based on many prior measurements. Less than 0.3 dB signal absorption at 30 MHz was measured by the HAARP riometer on the 24 October for the duration of the experiment. No significant absorption was measured during the 21 October 2009 experiment. Absorption in the HF band at the frequencies transmitted by the HAARP IRI are typically much more significant than absorption at 30 MHz. However, the low VHF absorption measured during the experiments is consistent with favorable conditions for ionosphere HF modification experiments, as expected for a period of quiet solar and geomagnetic activity.

4.2.2 HF modulation for ionosphere modification and diagnostics

Powerful high-frequency (HF) electromagnetic waves transmitted toward the ionosphere can result in heating of the electron gas, as a fraction of the electromagnetic wave energy is converted into electron gas thermal energy when the frequency, polarization, and energy density of the electromagnetic is sufficient to excite one of several plasma resonances. HF transmitters used in ionospheric modification experiments typically generate from tens of megawatts to gigawatts of effective radiated power (ERP). The HF interaction region, where HF energy is converted into electron gas thermal energy, typically extends up to a few tens of kilometers below HF reflection height where the ionosphere plasma frequency is equal to the transmitted HF frequency. Various plasma resonances and instabilities may be excited in the interaction region, some of the plasma waves generated contribute to “heating” the ionosphere, and some greatly increase the local ionospheric reflectivity in certain radio bands. The focus of the experiments described in this

chapter is on the ionosphere response to O-mode HF pumping as measured by UHF radar observations of the HF-enhanced ion line (HFIL) matching height and interpreted through SCIM simulations of heating experiments. Ionosphere O-mode heating will be described briefly in this section, followed by descriptions of the two HF amplitude modulation schemes used to modify and probe the ionosphere above the HAARP facility in Gakona, Alaska during the experiments.

During powerful underdense O-mode HF transmissions, when HF waves with frequency near or below the ionosphere critical frequency and left hand circular polarization (in the northern magnetic hemisphere), are transmitted with megawatts or more of effective radiated power, a fraction of electromagnetic wave power may be lost to electrons along the HF wave ray path in the D and E regions before reaching the F region. The polarization of the transmitted wave relative to the direction of the geomagnetic field distinguishes O-mode from X-mode HF pumping. The names of the HF wave polarizations refer to the ordinary and extraordinary plasma waves that couple with the transmitted electromagnetic wave as it refracts below the reflection layer. O-mode HF waves are reflected by the ionospheric layer that has a plasma frequency equal to the HF wave frequency. At a point below this reflection layer, the HF wave refracts to the extent that the wave propagation vector is nearly perpendicular to the geomagnetic field, an alignment that is favorable for HF-plasma coupling, enhancing perpendicular wave modes like Langmuir and ion-acoustic waves in the ionospheric plasma [Rietveld *et al.*, 1993].

To a first approximation, one may assume that the electromagnetic wave energy is converted to plasma thermal energy through plasma waves near the reflection height. However, note that UHF radar measurements of the HFIL target range occur at the matching height, not the reflection height, as discussed in section 4.2.3. Also, evidence from multiple instruments indicates that the location of the most significant O-mode HF heating occur at the upper-hybrid (UH) resonance height (see Kosch *et al.* [2007], for examples of HF-enhanced airglow above the HAARP facility measured from Poker Flat). Both heights are less than 10 km below the O-mode HF reflection height during typical modification experiments, so the distinction will not be discussed further.

In addition to electron temperature enhancements, O-mode HF heating can cause local density irregularities. One form of density irregularity observed in the high-latitude ionosphere is electron density enhancements in the lower F region below the reflection height, inferred from the measured descent of the HFIL return [Ashrafi *et al.*, 2006] during modification experiments at the EISCAT HF facility. Density depletions occurring during HF modification experiments have also

been observed in the F-region high latitude ionosphere over the HAARP facility [Kosch *et al.*, 2007] and the EISCAT facility [Rietveld *et al.*, 2003]. Typical plasma depletions of up to 20% were observed. High-resolution ISR measurements of the mid-latitude ionosphere over Arecibo indicated HF-induced density depletions near 30% [Hansen *et al.*, 1992].

Density irregularities tend to align parallel to the geomagnetic field. In both experiments described below, the HAARP IRI transmitted amplitude modulated HF sequences in a single-array configuration parallel to the geomagnetic field, toward -155° azimuth and 75° elevation. MUIR was also directed in the field-aligned direction.

4.2.2.1 HF *heat-scan* modulation sequence

The HF modulation scheme used during the 24 October 2008 experiment—referred to here as the *heat-scan* sequence—consists of an HF “modification” sequence intended to cause significant ionosphere density and temperature irregularities alternated with a frequency-scanning “probe” sequence of low duty cycle pulses intended to be used in conjunction with UHF radar as a probe of the time-dependent state of the ionosphere at altitudes near the HF reflection height. Each *heat* sequence consisted of a continuous constant-amplitude wave at a fixed HF frequency. Each *scan* sequence consisted of ten 100 ms pulses with 1 s inter-pulse period (IPP); the frequency of each pulse was assigned to be a fixed increment above the frequency of the previous pulse such that the frequency distribution of the ten pulses was centered on the heat frequency. Figure 4.2 shows a frequency-time illustration of the typical heat-scan sequence centered on a heat frequency of 4.2 MHz, the first pulse in each scan sequence is at frequency 4.1 MHz and each subsequent pulse was transmitted at a frequency 22 kHz higher than the previous pulse. This particular heat-scan sequence was transmitted from 0020-0030 UT.

The frequency increments were chosen to cover the maximum altitude range possible within the technical limitations of the HAARP IRI and the frequency bands allocated by the FCC. Another consideration in constructing the heat-scan sequence is that the ionosphere plasma frequencies associated with adjacent pulses in the scan sequence need to be separated in altitude at least 1 km—or the approximate size of the MUIR range bins used during the experiment—for a total of about 10 km altitude range. The MUIR configuration used to measure the range to the HFIL return with up to 600 m range resolution is described in section 4.2.3. Variations of the heat-scan sequence of transmit frequencies are possible. For example, the heat frequency and center

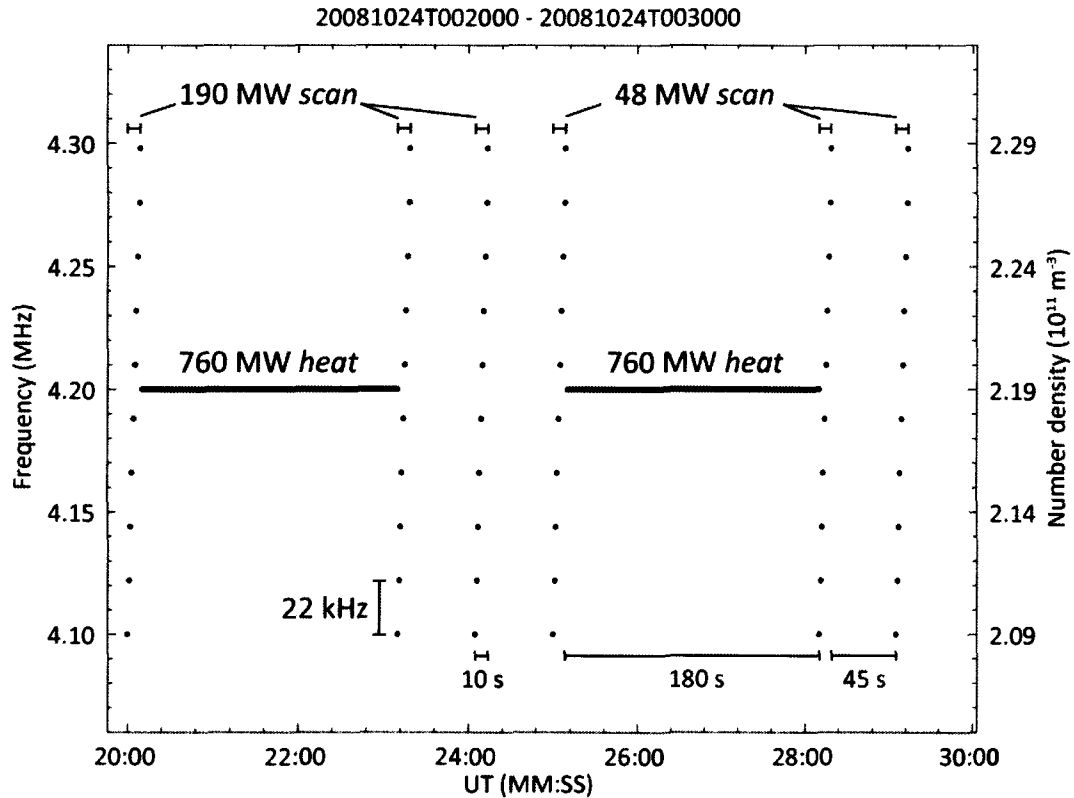


Figure 4.2 Frequency, time and effective radiated power (ERP) variation of the *heat-scan* HF on-off modulation sequence initiated at 0020 UT on 24 October 2007 by the HAARP IRI. Each discrete point represents a 10 ms pulse, and the inter-pulse period is 1 s. The horizontal lines represent times of continuous HF transmission. The right-side axis is labeled with the plasma frequency number densities corresponding to the labeled HF frequencies on the left-side axis.

frequency of the scan pulses can be varied in order to probe the ionosphere response to heating at different altitudes.

One restriction for heat-scan sequences transmitted by the HAARP IRI in single-array configuration is that any frequency changes outside of a 200 kHz window require at least 30 s pause in transmissions [McCarrick, 2008]. The HAARP frequency allocation also restricts the range of frequencies available. A consequence is that the size of the frequency increments in the scan sequence may vary with the heat frequency chosen.

4.2.2.2 HF heat-pulse modulation sequence

A *heat-pulse* HF modulation sequence was used in the 21 October 2009 experiment. The heat-pulse modulation sequence consists of a long duration constant-amplitude heat subsequence—similar to the heat subsequence in the heat-scan sequence from section 4.2.2.1—intended to cause significant and lasting ionospheric density modifications, alternated with a pulse sequence intended to probe the time-dependent ionospheric response to the heat sequence. However, unlike the scan subsequence in the heat-scan sequence, the pulse subsequence in the heat-pulse sequence is transmitted at a single HF frequency. This means that measurements of the spatial variation of UHF matching heights at multiple HF frequencies are sacrificed for increased time resolution measurements of the matching height at a single HF frequency.

The HAARP IRI was set to transmit parallel to the geomagnetic field in single-array configuration at a constant frequency. A probe sequence of short 1 ms pulses with 1 s an inter-pulse period (IPP) lasting for 2 or 5 min was alternated with heat sequences consisting of a continuous constant-amplitude wave lasting either 2 min or 5 min. The 21 October experiment described here took place from 01:43:00 to 01:59:30 UT. The HAARP IRI transmitted an amplitude modulated 4.04 MHz wave directed parallel to the geomagnetic field with O-mode polarization. A constant-amplitude HF heat sequence was transmitted from 01:49:00 to 01:54:00 UT. A HF probe sequence of 10 ms pulses with 1 s IPP was transmitted before and after the heat sequence, from 01:43:00 to 01:49:00 UT, and from 01:54:00 to 01:59:30 UT. The HF frequency of 4.04 MHz was chosen so that the reflection height was just below the critical layer at the time of the experiment as measured by the HAARP ionosonde. This experiment followed a similar experiment from 01:35:00 to 01:42:30 UT that transmitted at 5.04 MHz; the HF frequency was apparently above foF₂ at the time, because no HFIL was observed with MUIR.

4.2.3 Analysis of MUIR coded long-pulse data

The time-evolution of the vertical structure of the ionosphere in response to HF heating may be inferred from the time and HF-frequency dependence of the HFIL altitude as measured with UHF radar. Plasma density and temperature at the HFIL matching height are related to the HF pump frequency and the UHF radar frequency through the matching condition. During the October experiments at the HAARP facility in Gakona, the MUIR coded-long pulse operation mode *C996_4us* was used to measure the range to HFIL targets with approximately 0.6 km spatial and 1 s

time resolution. This section describes the phase and amplitude modulation of the UHF radar wave transmitted by MUIR in the *C996_4us* configuration. A *range-decoding algorithm* for transforming the baseband time series recorded by MUIR during each radar pulse to high-resolution signal-to-noise ratio vs. range data is described in detail below. Verification of the range-decoding algorithm applied to *C996_4us* data is provided by using the technique to measure the range to a polar-orbiting satellite and its radar line-of-sight approach velocity while it passed over Gakona and comparing the measurement to orbital data published in the North American Aerospace Defense Command (NORAD) and National Space Science Data Center (NSSDC) catalogs. The coded long-pulse technique for identifying the HFIL reflection height is validated for identifying regions of strong HF-induced plasma turbulence by comparing MUIR measurements of the HFIL reflection height with field-aligned measurements of HF-induced airglow during an March 2009 experiment reported by *Pedersen et al.* [2010].

Strong O-mode HF pumping of the ionosphere can excite the pondermotive plasma instabilities, with an F-region rise time of approximately 1 ms, that generate plasma waves detectable with UHF radars [*Stubbe, 1996*] like MUIR [*Oyama et al., 2006*]. Note that thermal instabilities at the upper-hybrid (UH) resonance altitude have rise times of seconds to minutes and may be excited during HF pumping in addition to the pondermotive instabilities. Two pondermotive plasma instabilities provide mechanisms that convert electromagnetic energy from the pump wave into plasma wave energy: the parametric decay instability (PDI) and the oscillating two-stream instability (OSTI). Both of these instabilities excite plasma waves with a wave number k that approximately satisfy the following expression given by *Stubbe [1996]*:

$$k^2 \approx \frac{2}{3} \omega_o \frac{\omega_o - \omega_p}{v_e^2} \quad (4.1)$$

where ω_o is the pump wave angular frequency, ω_p is the altitude-dependent plasma frequency, and v_e is the electron thermal velocity. UHF radar waves incident upon the turbulent region are back scattered when the plasma wave number satisfies the Bragg condition of the radar, that is, when the plasma wave number k is twice the radar wave number k_R . Substituting

$v_e = \sqrt{k_B T_e / m_e} \approx 3.89 \times 10^3 \sqrt{T_e}$ for the electron thermal velocity, $\omega_p = \sqrt{n_e e^2 / \epsilon_o m_e} \approx 57.2 \sqrt{n_e}$ for the plasma frequency, and $2k_R$ for k in equation (4.1) yields an expression that relates the pump and

radar wave frequencies to the electron temperature and number density in the spatial region containing the HFIL return

$$\begin{aligned} 0 &\approx \frac{m_e}{6k_B} \omega_o \frac{\omega_o - \sqrt{n_e e^2 / \epsilon_o m_e}}{T_e} - k_R^2 \\ &\approx 1.10 \times 10^{-8} \omega_o \frac{\omega_o - 57.1 \sqrt{n_e}}{T_e} - 87.38 \end{aligned} \quad (4.2)$$

Here, the MUIR wave number $k_R = 2\pi f_R / c \approx 9.35$ rad/m has been calculated from a transmit frequency $f_R = \omega_R / 2\pi = 446$ MHz. It is clear by inspection of equation (4.2) that time variations of electron number density and temperature will cause the altitude at which the radar observes HF-induced plasma waves to vary in time as well. Furthermore, due to the steep electron density vertical gradient below the critical layer, the ionosphere Bragg-scattering layer is thin. It may initially appear from equation (4.2) that it is impossible to ascribe variations in observed in the HFIL altitude to electron temperature variations or electron density variations. But knowledge of the direction of vertical density and temperature gradients along with typical electron temperature and density perturbation response times may sometimes be used—as described in section 4.4—to identify whether temperature perturbations or density perturbations are the major contributing factor (or factors) to observed HFIL altitude variations.

The MUIR *C996_4us* configuration used in the experiment described in this chapter specifies the transmit pulse duty cycle, sample rate, transmit frequency, phase code modulation baud and sequence, the pulse length and direction, and length of time that the receiver records data for each pulse. Table 4.2 contains a summary of the relevant parameters specified in the *C996_4us* configuration, as well as the symbols used in this chapter to refer to the parameters individually in the following description of the decoding and analysis algorithm.

Table 4.2 MUIR field-aligned coded long pulse configuration parameters and values. Symbolic names used in the text to refer to individual parameters are provided for reference.

MUIR *C996_4us* CONFIGURATION

Parameter	Value	Symbol
TX pulse width	996 μ s	t_{TX}
Inter-pulse period (IPP)	10 ms	t_{IPP}
TX code baud length	4 μ s	t_b
TX pulse width	249 bits	N_{TX}
Range bin length	600 m	Δr
RX sample rate	250 kHz	f_s
RX sample duration	8.8 ms	t_{RX}
RX sample width	1100 bits	N_{RX}
TX/RX direction	$(-155, 75)^\circ$ (field aligned)	\hat{r}

Without encoding the transmitted signal and decoding the received signal, it is difficult to distinguish the locations of two scatter-targets in the radar volume that are less than the pulse width of 996 μ s (150 km round-trip radar range) apart. If the only analysis objective is to determine the range to each target, then choosing a binary-phase modulation sequence that has an autocorrelation function resembling a Dirac δ -function allows range resolution significantly smaller than the transmitted pulse width. An algorithm that decodes the received baseband signal recorded during MUIR *C996_4us* long-pulse experiments to range-time data that resolves positions of multiple distinct scatter-targets with range resolution on the order of the baud length will now be described in detail. This *range-decoding* algorithm is a variation of the spectral method described by *Sulzer* [1986] and will convert the baseband time-series received during a single coded UHF pulse to high-resolution SNR vs. range data. It is useful for localizing point targets and volume targets with thin layers of enhanced UHF reflectivity.

It is first necessary to define a few mathematical representations of the data structures recorded by the MUIR system during an experiment in order to describe the *C996_4us* range-decoding algorithm. Let $\mathbf{z} \in \mathbb{C}^{N_{RX} \times 1}$ be a vector (array) containing the complex-valued baseband signal received from a single transmitted pulse p , let $\mathbf{t} \in \mathbb{R}^{N_{RX} \times 1}$ be a vector containing the time of each

by the convolution $*$ of the transmitted signal \mathbf{c} with the vector $\mathbf{1}(s_1, s_2, \dots, s_m) \in \mathbb{R}^{N_{\text{RX}} \times 1}$ whose values are 1 at the indices corresponding to the ranges of the scatters and are 0 everywhere else. As can be verified by examining the positive lags from the cross correlation of \mathbf{c} and $\mathbf{c} * \mathbf{1}(s_1, s_2, \dots, s_m)$, the cross-correlation function has sharp peaks at the lags corresponding to the locations of the targets s_i and is relatively flat everywhere else, unlike the result obtained for an un-coded pulse where the code \mathbf{c} is simply vector of ones followed by a vector of zeros. This means that bi-phase modulating the transmitted pulse with code (4.3) and cross-correlating the transmitted phase-code with the received signal is one possible procedure for obtaining the location of targets, improving the range-resolution from the pulse width to the width of one code baud.

The received baseband signal is complex-valued and scatter from turbulent ionospheric plasma can modify the complex-valued phase of the incident pulse so an extension of the procedure discussed above for decoding a signal received from idealized scatter targets is required. The spectral method described by *Sulzer* [1986] provides a method to calculate the range- and time-dependent power spectrum with high range resolution from coded long pulse radar data. This method may be extended by using the peak Fourier power in the frequency spectrum at a particular range bin to estimate the strength of the scattered power emitted from that particular range.

The following sequence of calculations in this *basic range-decoding algorithm* is performed once for each range bin index $i \in \{1, 2, \dots, N_{\text{RX}}\}$ of the received baseband signal from a single MUIR pulse. Each range-indexed result is the discrete frequency power spectrum of the scattered signal sampled from range bins r_i to $r_{i+N_{\text{TX}}}$:

1. Calculate the array-product of the code with the sample rate f_s baseband signal $\tilde{\mathbf{z}}_j = z_{i+j-1} \mathbf{c}_j$ for each $j \in \{1, 2, \dots, \min(N_{\text{TX}}, N_{\text{RX}} - i + 1)\}$ and store the result in the vector $\tilde{\mathbf{z}} \in \mathbb{C}^{N_{\text{TX}} \times 1}$, padding the end of the vector with zeros as necessary.
2. Calculate the Fast Fourier Transform (FFT) of the signal sample $\tilde{\mathbf{z}}$ and save the result to the vector $\tilde{\mathbf{Z}} = \text{FFT}(\tilde{\mathbf{z}}, f_s)$.
3. Calculate the power contained in each frequency bin j and store the resulting power spectrum $\Phi_{ij} = Z_j \bar{Z}_j$.
4. Set the *basic range-decoded power* $\zeta_i = \max_{1 \leq j \leq N_{\text{TX}}} (\Phi_{ij})$

The signal-to-noise ratio (SNR) may be improved further by integrating or filtering the power spectrum Φ_i or the maximum Fourier power ζ_i , that is, over multiple pulses.

Notice that measuring the range to the scattering target using the basic range-decoding algorithm depends on an application of Parseval's theorem: the total signal power contained in the frequency power spectrum Φ_i of range r_i is equal to the total signal power calculated from the time-series signal sample. Much of the total power contained in the power spectrum of a signal sample starting at a range bin containing a scattering target will be concentrated in only a few frequency bins if the signal is from a target rather than ambient noise. In the case of a hard scattering target like a satellite, the power will be concentrated in the frequency bin corresponding to the Doppler shift induced by the relative line-of-site velocity of the of the radar and the satellite. For ionospheric HFIL targets, the power will generally be concentrated in frequency bins within the approximate interval $[-10, 10]$ kHz. Therefore, ranges to various targets in the radar volume of a single pulse can be identified with high range resolution by setting each entry of a vector $\zeta \in \mathbb{R}^{N_{\text{rx}} \times 1}$ with the maximum value of the frequency power spectrum.

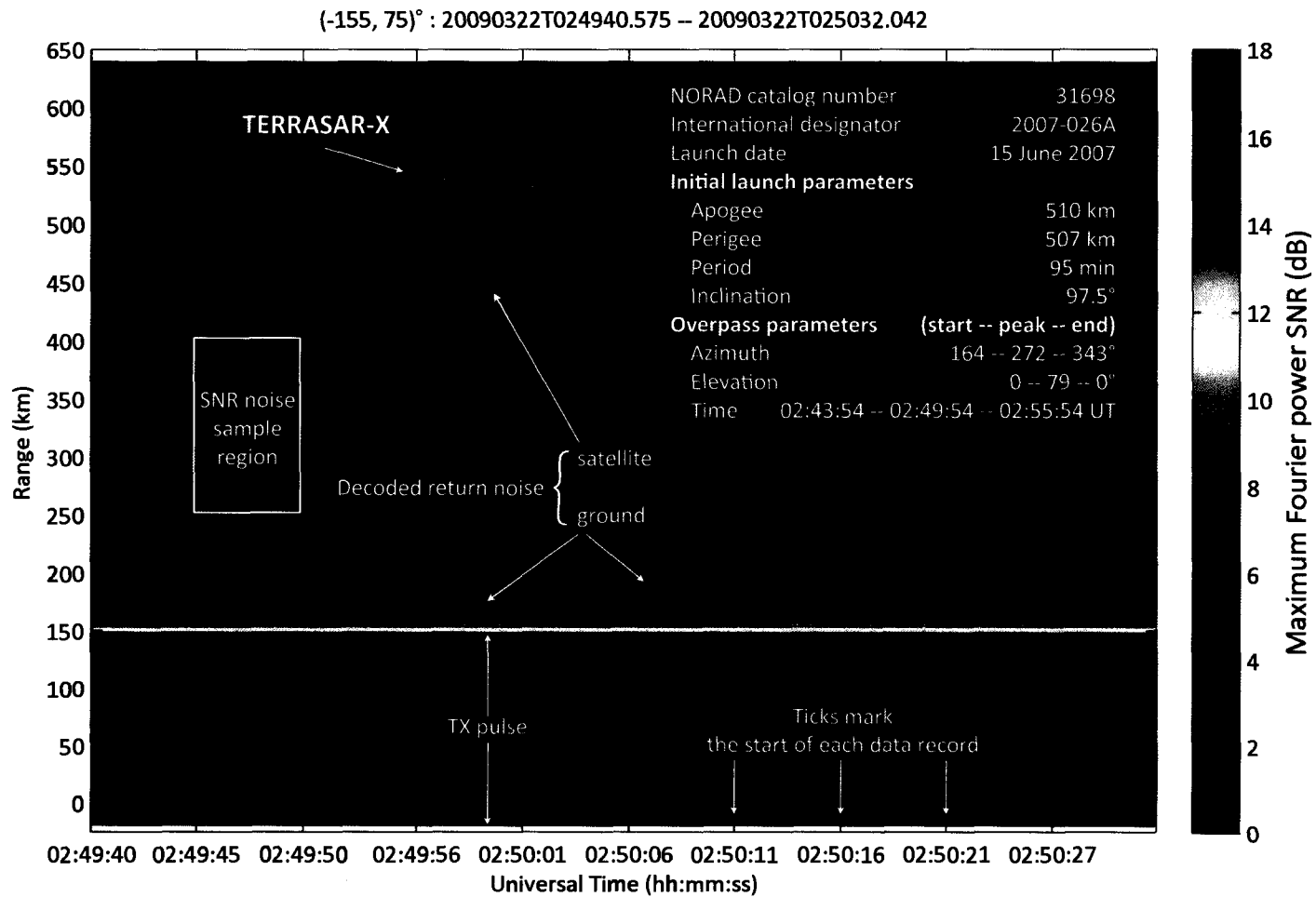


Figure 4.3 Annotated range-time diagram of the peak Fourier power signal-to-noise ratio (SNR) received measured with MUIR. The UHF radar was aligned parallel to the geomagnetic field and the Terra SAR-X satellite entered the radar volume at about 02:49:56 UT (hh:mm:ss).

Satellite transits over the HAARP facility provide observation of near-ideal point targets for the verification of MUIR pulse range-decoding algorithms. MUIR operated in the *C996_4us* configuration, directed parallel to the geomagnetic field, on 22 March 2009 at 0250 UT when a satellite entered the radar volume at approximately 535 km range. Figure 4.3 is a range-time diagram of 5000 coded pulses—about 50 seconds of data—processed with the range-decoded maximum power algorithm described above and post-processed with a 10 pulse (1 s) time-integration filter. The plotted power has been normalized by the mean decoded maximum power in the indicated region. A catalog of NORAD two-line element sets downloaded in June 2009 and freely available satellite pass prediction software [Vallado *et al.*, 2006] was used to identify the time-dependent positions of satellites in the catalog relative to the surface of Earth. The observed target return is likely the Terra SAR-X satellite (NORAD catalog number 31698). To an observer in Gakona on 22 March 2009, the satellite would rise above the horizon at 0244 UT and 164° azimuth up to a maximum elevation of 79° at 0250 UT and 272° azimuth, then set below the horizon at 0256 UT and 343° azimuth. The MUIR field-aligned 3-dB beam half-width is approximately 5° [Oyama *et al.*, 2006]; the azimuth and elevation (dip angle) of the geomagnetic field at Gakona are 205° and 75° , respectively. Note that Terra SAR-X was launched into a polar orbit—with an apogee of 510 km and a perigee of 507 km [NSSDC, 2009]—so the decrease in target range with time observed in Figure 4.3 is consistent with a south-to-north transit through a field-aligned radar beam.

A deficiency in basic range-decoding algorithm applied to MUIR long-pulse returns is apparent by inspection of Figure 4.3, or even by hand computations while applying the algorithm to a simple example. The decoded baseband sample at a particular range r_i includes returns from targets in the region $r_i \pm ct_{TX}/2c$, where c is the speed of light in a vacuum. A radar return from a point target appears as a ct_{TX} wide region in radar range of enhanced signal power—about 149 km wide in the *C996_4us* experiment mode—and may be from a satellite, a ground structure, or the pulse itself when the transmitter is engaged—features all present in Figure 4.3—as well as from multiple regions of turbulent plasma. While the basic range-decoding algorithm greatly increases the signal strength in baud-length regions containing scattering targets, it also increases the decoded signal (noise) in a much larger region: the diameter is twice the pulse-width and centered on target. Figure 4.4a and Figure 4.4b show the raw power $z_i \bar{z}_i$ and the basic range-decoded power ζ_i , respectively, vs. range r_i for two individual MUIR pulses, one transmitted shortly before, and one approximately 3 seconds after the Terra SAR-X satellite entered the radar volume. The 50 db power SNR step function beginning at range 0 km in Figure 4.4a is return from the transmitted pulse itself

and the noisy step function beginning at 538 km range is the return from the satellite. Since the range in the *C996_4us* configuration is determined by the time measured from the start of the transmitted pulse, the range to targets separated by round-trip distances greater than the pulse width can be easily inferred from the start of each step function observed in the raw power vs. range diagrams.

In the plots of basic range-decoded power vs. range, the decoded return noise surrounding each target may be reduced by normalizing the basic range decoded power ζ_i at each range bin r_i by the total power contained in the signal sample: $\zeta_i \leftarrow \zeta_i / \sum_j \Phi_{ij}$. Figure 4.4c shows a plot of this normalized decoded power. Return signal received from targets contained in the radar volume at ranges less than 200 km during *C996_4us* experiments is still contaminated, however, by the signal from the transmitted pulse itself and the ground clutter return signal. One technique for reducing contributions to the decoded power of ground clutter signal or other targets within a pulse width of range r_i is to mute the beginning of the phase modulation code c by setting the first N_{MUTE} elements of c to zero before applying the code-multiplication step. Then the decoded power spectrum Φ_i at range r_i contains signal contributions from elements $z_{i+N_{\text{MUTE}}}$ to $z_{i+N_{\text{TX}}-1}$ of the baseband signal vector, thereby eliminating the signal contribution contained in bins z_i to $z_{i+N_{\text{MUTE}}-1}$ from a target located in at or below range $r_{i+N_{\text{MUTE}}-N_{\text{TX}}}$. Of course, muting a fraction of the code in the manner described above incurs a cost: reducing the length of the sampled signal reduces the total signal power and length (or resolution) available to calculate the Fourier spectrum.

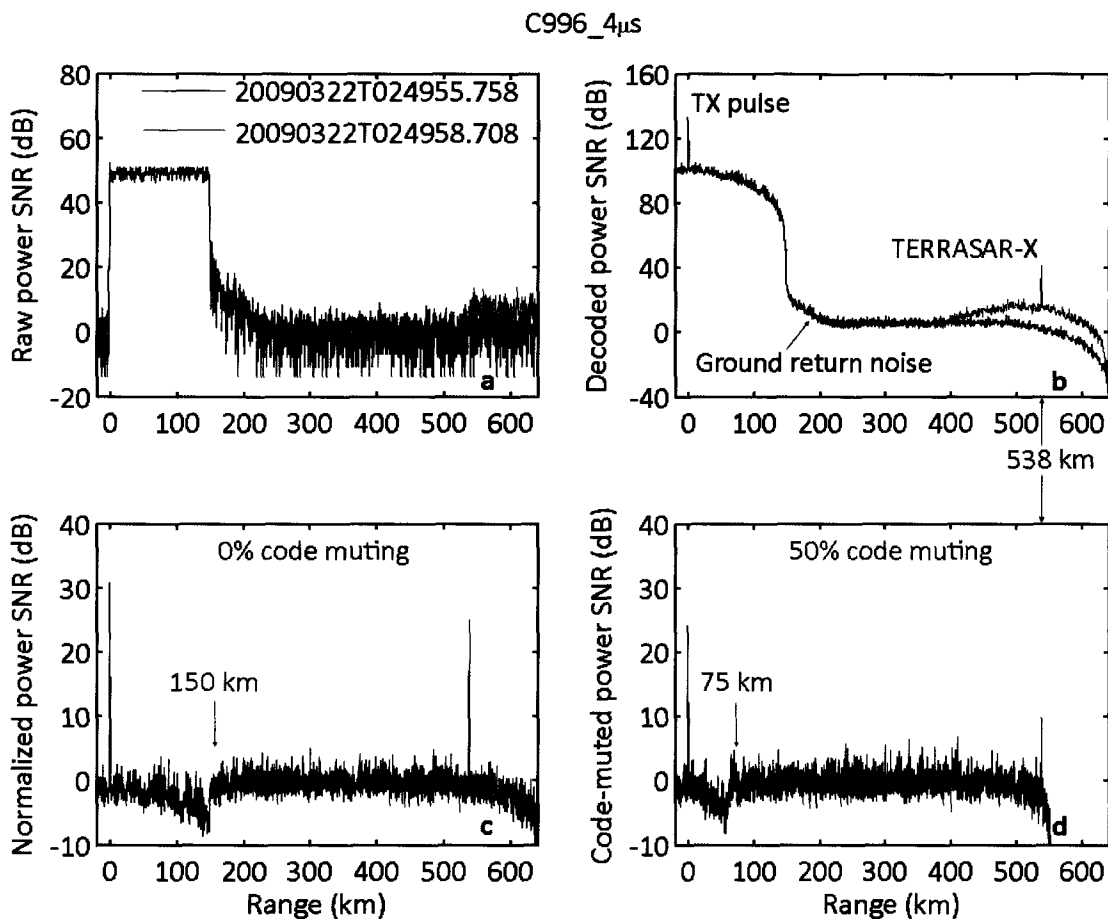


Figure 4.4 Received power signal-to-noise ratio (SNR) vs. time (range) of two MUIR 996 μ s pulses phase-modulated with a 4 μ s baud code. The two pulses were transmitted parallel to the geomagnetic field at approximately 02:49:55 (black) and 02:49:58 (red) UT (hh:mm:ss) on 22 March 2009. The TERRASAR-X satellite was traversing the radar pulse volume during the second UHF pulse, as indicated in Figure 4.3. SNR vs. range is plotted before (a) and after (b) applying the *basic range-decoding* algorithm. SNR vs. range calculated with the *normalized code-muted range-decoding* algorithm is plotted before (c) and after (d) muting the first half of the phase code used in the decoding algorithm, reducing the maximum signal range affected by ground clutter.

The basic range-decoding algorithm can be extended to include the independent *normalization* and *code-muting* options that may be selectively used to mitigate the interfering effects of ground clutter or signals from other nearby targets. In the *normalized and code muted range-decoding* algorithm described below, the code muting is applied to the first N_{MUTE} bits or elements of the code vector \mathbf{c} . The evaluations of the conditional statements in the algorithm correspond with the name of the algorithm. For example, in the *normalized and 50% code-muted range-decoding algorithm*—used to plot the SNR vs. range in Figure 4.4d—the first $\lfloor N_{\text{TX}}/2 \rfloor = 124$ bits of the code are muted before the multiplication step; then after calculating the spectrum, the maximum Fourier power $\zeta_i = \max_j (\Phi_{ij})$ is normalized by the total spectral power $\sum_j \Phi_{ij}$ calculated at range r_i .

As in the basic range-decoding algorithm described above, the following sequence of calculations is performed once for each range bin index $i \in \{1, 2, \dots, N_{\text{RX}}\}$ to calculate the frequency spectrum of the signal scattered from range bin r_i during a single MUIR pulse. The spectrum is calculated by sampling and decoding the baseband signal z_i through $z_{i+N_{\text{TX}}}$. Each range-indexed result is the discrete frequency power spectrum of the scattered signal sampled from range bins r_i to $r_{i+N_{\text{TX}}}$.

1. If *code muting* is set to N_{MUTE} bits, then mute the code vector \mathbf{c} by setting the first N_{MUTE} bits of \mathbf{c} to zero.
2. Calculate the array-product of the code with the sample rate f_s baseband signal $\tilde{\mathbf{z}}_j = z_{i+j-1} \mathbf{c}_j$ for each $j \in \{1, 2, \dots, \min(N_{\text{TX}}, N_{\text{RX}} - i + 1)\}$ and store the result in the vector $\tilde{\mathbf{z}} \in \mathbb{C}^{N_{\text{TX}} \times 1}$, padding the end of the vector with zeros as necessary.
3. Calculate the Fast Fourier Transform (FFT) of the signal sample $\tilde{\mathbf{z}}$ and save the result to the vector $\tilde{\mathbf{Z}} = \text{FFT}(\tilde{\mathbf{z}}, f_s)$.
4. Calculate the power contained in each frequency bin j and store the resulting power spectrum $\Phi_{ij} = Z_j \bar{Z}_j$.
5. If *spectrum normalization* is set, then replace the power spectrum with the normalized spectrum $\Phi_i \leftarrow \Phi_i / \sum_j \Phi_{ij}$.
6. Set the *range-decoded power* $\zeta_i = \max_{1 \leq j \leq N_{\text{TX}}} (\Phi_{ij})$

As with the basic range-decoding algorithm, the SNR of the normalized code muted power in this algorithm may be improved by integrating or filtering the decoded power in time. While no filtering or time integration is necessary to measure the range to target in many *C996_4us* experiments, a one-second moving window average is sufficient to considerably reduce noise in the measurement and enhance image quality without significantly sacrificing temporal resolution.

Observe in Figure 4.4b-d that the measured range to the satellite target is not affected by applying the code muting or normalization steps in the above algorithm. The code muting step decreases the amplitude of the maximum Fourier power at the target range, and it decreases the maximum range affected by ground clutter. The spectrum normalization step reduces the amplitude of the noise surrounding the target. Figure 4.4c shows the normalized range-decoded power of the two MUIR pulses and Figure 4.4d shows the code-muted and normalized range-decoded power, where the first $N_{TX}/2$ elements of c have been set to zero before multiplying by the code. Inspection of Figure 4.3 and Figure 4.4, combined with available satellite orbit data, verifies that applying the range-decoding method, described above, to MUIR *C996_4us* coded long pulse data yields the range to a satellite (point) target with a precision of one baud length and an accuracy that is consistent with satellite orbit calculations based on NORAD two-line element sets and the initial launch parameters recorded in the National Space Science Data Catalog (NSSDC) Master Catalog of Spacecraft.

The accuracy of the frequency spectrum Φ_i calculated in the range-decoding algorithms from the decoded signal at each range bin r_i remains to be verified. A polar-orbiting satellite in a near-circular orbit like TerraSar-X will approach—or retreat from—a high-latitude radar like MUIR looking parallel to the local geomagnetic field with a line-of sight (LOS) velocity that is a significant fraction of the linear velocity tangent to the satellite orbit. The field-aligned LOS velocity of TerraSar-X during the Gakona overpass will first be calculated from the Doppler shift observed in the spectrum of the particular decoded pulses illustrated in Figure 4.4, and then it will be compared with the LOS velocity measured from the slope of the trajectory connecting two points at each end of the satellite return illustrated in Figure 4.3.

An observer on the polar-orbiting TerraSar-X satellite, traveling approximately from the south to the north during the Gakona overpass, would detect UHF electromagnetic radiation emitted from MUIR while the satellite is contained within the radar volume. The frequency of the UHF signal received by the detector on the satellite is Doppler shifted a few kilohertz higher than the

frequency of the UHF signal transmitted by MUIR due to the relative LOS motion between the observer and the source. If f is the frequency of the UHF wave transmitted by the source, and the detector is moving toward the source with LOS velocity u , then the frequency f' of the UHF wave detected by an observer on the satellite is given by the formula

$$f' = \left(1 \pm \frac{u}{c}\right) f. \quad (4.5)$$

Where the sign is determined by whether the detector is moving toward or away from the wave source. During the satellite overpass, the electrons in the metal structure of the satellite re-radiate an electromagnetic wave with frequency equal to the observed UHF wave. This scattered UHF wave has frequency f' in the reference frame of the satellite, but when the wave is received by the MUIR detector in the co-rotating frame of Earth, the wave has frequency f'' since the satellite (now the wave source) is moving relative to the radar (now the detector). The frequency f'' of the scattered UHF wave received by MUIR after scattering from the satellite can be written using equation (4.5) in terms of the original transmitted frequency, the LOS velocity, and the speed of light.

$$\begin{aligned} f'' &= \left(1 \pm \frac{u}{c}\right) f' \\ &= \left(1 \pm \frac{u}{c}\right)^2 f \end{aligned} \quad (4.6)$$

Let $\Delta f = f'' - f$ be the difference between the frequency of the satellite-scattered wave received by MUIR and the frequency of the original transmitted wave in the MUIR reference frame. Solving equation (4.6) for the LOS velocity u of the satellite relative to MUIR, yields an expression in terms of the frequency difference and the speed of light.

$$u = \pm \left(\sqrt{\frac{\Delta f}{f} + 1} - 1 \right) c \quad (4.7)$$

If $\Delta f = 0$ then the LOS velocity is zero as expected. Assuming a circular orbit at 510 km altitude and an orbital period of 95 min, the linear velocity of the TerraSar-X satellite is 7600 m/s. Radar scatter from a target with this LOS velocity approaching MUIR would be observed to be Doppler shifted by 23 kHz. An estimate for the Doppler shift expected to be observed during an a TerraSar-

X pass over Gakona can be calculated by slightly simplifying the geometry of the TerraSar-X orbit relative to the geomagnetic field line out of Gakona. Assume the orbit is aligned with the northward component of the geomagnetic field, then the field-aligned component of the linear velocity $v = 7600$ m/s is $u = v \cos \gamma \approx 2000$ m/s where $\gamma = 15^\circ$ is the geomagnetic dip angle at Gakona. This LOS velocity estimate is an upper bound on the actual satellite velocity because no inclination of the orbit relative to magnetic north was assumed; it corresponds to a MUIR-observed Doppler shift of 5.9 kHz via equation (4.7).

To measure the frequency Doppler shift of the signal scattered from the satellite target, the baseband signal from a single pulse at 02:49:58 UT—when the satellite was contained in the radar

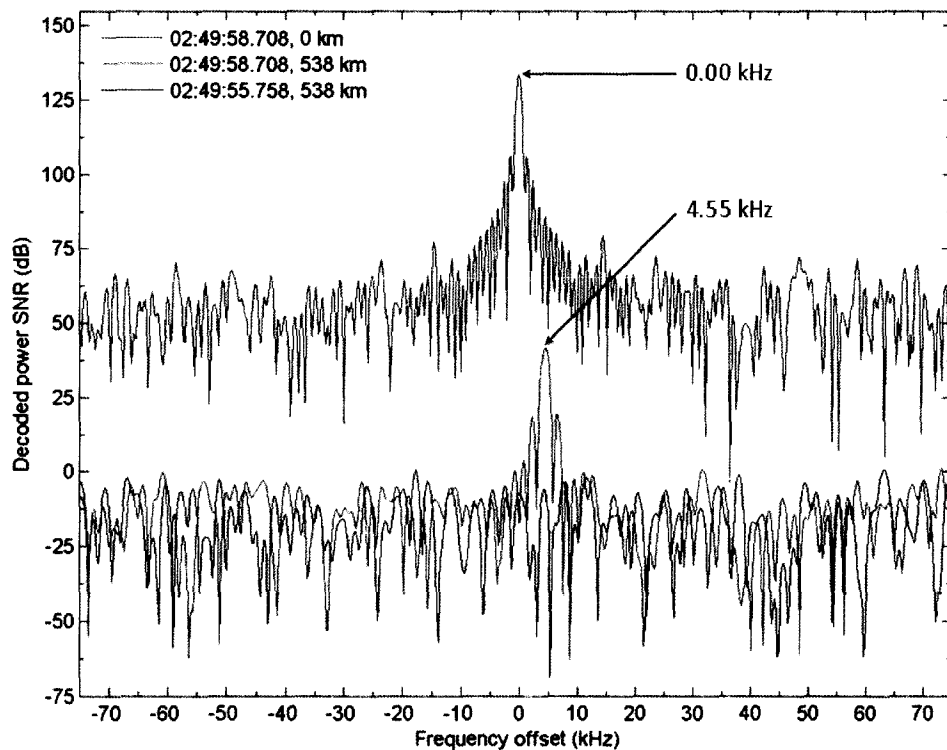


Figure 4.5 1100-point Fast Fourier Transform (FFT) power spectrum of samples taken from the two decoded MUIR pulses illustrated in Figure 4.4 (b). The pulse samples begin at (blue) 0 km range and (red, black) 538 km range; the sample rate is 250,000 kHz. Samples starting at 538 km range are 684 μ s in duration and the sample starting at 0 km range is 996 μ s in duration. A positive frequency offset indicates that the relative distance between the receiver and the reflecting target is decreasing.

volume—was sampled from the satellite range out to the shorter of the pulse width or the end of the recorded baseband signal for that pulse. After array-multiplying the sample vector by the phase code vector as described in the basic range-decoding algorithm, the discrete Fast Fourier Transform (FFT) function provided in the Matlab environment [Mathworks, 2007] was used to calculate the 1100-point Fourier power spectrum. Since the baseband signal return from the satellite at $r_{929} = 537$ km range extends beyond the $r_{1100} = 639$ km maximum recorded range in the MUIR C996_4us configuration, the 171-point sample was padded with zeros before calculating the FFT. The Fourier power spectrum of a sample from the decoded signal recorded during a pulse at 02:49:55 UT, just before the satellite entered the radar volume, was also calculated from the same sample ranges used in calculating the spectrum of the satellite return. The spectra from both pulses at the satellite target range are plotted in Figure 4.5 on a frequency axis with values that are expressed as offsets relative to the transmit frequency 446 MHz of MUIR. The spectrum of the 02:49:58 UT pulse sampled from range bins $r_{33} = 0$ km to $r_{282} = 149$ km is also plotted in the figure, verifying that the combined transmitted signal and ground clutter return are not Doppler shifted in frequency. The Doppler frequency shift observed in Figure 4.5 from the satellite return at 537 km is 4.55 kHz, consistent with a target that is moving toward the radar with a LOS velocity of 1500 m/s and with the velocity upper bound estimated from the initial orbit parameters of the satellite.

4.2.4 *Ionosphere HF modification simulation*

This section describes the 1-D ionosphere HF heating simulation constructed with the self-consistent ionosphere model (SCIM). The model equations were modified by incorporating an HF-heat source term in the electron energy equation and solved on a refined spatial grid of 0.5 km resolution near the HF-reflection height chosen at the time of the heating experiment. The HF-electron heat source term simulates the heating of ionospheric plasma surrounding the reflection height due to HF-driven plasma turbulence. A refined grid is necessary since a sharp gradient in any of the model equation terms ideally needs to span at least four grid points, and the plasma frequency just below the critical layer can vary by about 100 kHz over an altitude range of about 2 km.

The HF-electron heat source term is specified here to be a step function: *on* at locations where the plasma frequency is in some small interval surrounding the chosen HF frequency below the reflection height and *off* everywhere else. Consequently, the refined grid also needs to cover at least 10 km in altitude since one observed and modeled effect of increasing the F-region electron

temperature is to increase the plasma density below the critical layer—and move the reflection height about 10 km downward at the heating rates used in these simulations—through decreased recombination of NO^+ and O_2^+ . The most significant limitation of the HF heating simulation is that the electron energy equation heat source term at any particular grid point can only depend on the solution variables at that point, meaning that a self-consistent calculation of the heating rate at an altitude that incorporates the absorption of HF energy at lower altitudes must be done in between calls to the pdepe numerical solver, that is at a fairly coarse time resolution. Finally, the specific ionosphere HF-heating numerical experiments described in this chapter were initialized from ionosphere profiles calculated during the diurnal equilibrium simulation described in section 3.2.

HF heating of the ionosphere can be modeled with the one-dimensional set of time-dependent equations described in Chapter 2 by introducing an additional heat source term $\delta E_s/\delta t$ in the electron gas energy equation (2.39):

$$\frac{3}{2} n_s k_B \frac{\partial T_s}{\partial t} - \sin I \frac{\partial}{\partial z} \left(\lambda_s \sin I \frac{\partial T_s}{\partial z} \right) = \frac{\delta E_s}{\delta t} \quad (4.8)$$

The electron gas energy equation is fully coupled to the ion momentum equations in the SCIM calculations so HF-induced density effects can be modeled by incorporating HF-enhanced electron volume heat rates in the electron energy equation. The HF-enhanced electron heat source term in equation (4.8) can be approximated with the source term s in the pdepe solver equation (3.6):

$$c \left(z, t; n_s, T_s, \frac{\partial n_s}{\partial z}, \frac{\partial T_s}{\partial z} \right) \frac{\partial}{\partial t} \left(\frac{n_s}{T_s} \right) = z^{-m} \frac{\partial}{\partial z} \left[z^m f \left(z, t; n_s, T_s, \frac{\partial n_s}{\partial z}, \frac{\partial T_s}{\partial z} \right) \right] + s \left(z, t; n_s, T_s, \frac{\partial n_s}{\partial z}, \frac{\partial T_s}{\partial z} \right) \quad (4.9)$$

In the electron gas energy equation, written in the form required for the pdepe parabolic solver in equation (3.17), the heat source term s_{22} for the electron gas now includes a term describing the heating from interaction with strong HF transmissions:

$$s_{22} = \frac{\delta E_e}{\delta t} = \frac{\delta E_e}{\delta t} \Big|_{\text{supra-thermal collisions}} + \frac{\delta E_e}{\delta t} \Big|_{\text{elastic ion collisions}} + \frac{\delta E_e}{\delta t} \Big|_{\text{elastic neutral collisions}} + \frac{\delta E_e}{\delta t} \Big|_{\text{inelastic neutral collisions}} + \frac{\delta E_e}{\delta t} \Big|_{\text{HF interaction}} \quad (4.10)$$

Absorption of HF wave energy below the interaction region will not be considered here so the total amount of HF energy available for F region heating will assumed to be independent of conditions

above and below the chosen interaction region. Furthermore, the heat source term will be assumed to be independent of the constituent gas temperatures.

As a first approximation, HF-electron heating is set at the altitude where the HF O-mode index of refraction is zero. That is, the simulated heating occurs near the region below the critical layer where the plasma frequency is equal to the HF wave frequency $f_{\text{HF}} = \omega_p(n_e(z,t))/2\pi$. The interaction region is defined to be where the ionosphere plasma frequency is in a neighborhood of the heater frequency $\omega_p \in 2\pi \cdot [f_{\text{HF}} - \frac{1}{2}\Delta f_{\text{HF}}, f_{\text{HF}} + \frac{1}{2}\Delta f_{\text{HF}}]$ where $\Delta f_{\text{HF}} = 120$ kHz. For $f_{\text{HF}} = 4.2$ MHz, the electron number density associated with this plasma frequency neighborhood is $n_e \in (2.19 \pm 0.06) \times 10^{11} \text{ m}^{-3}$ and covers an altitude range of about 217 ± 1 km of the simulated ionosphere, illustrated in Figure 3.2, over Gakona at 0100 UT on 24 October 2008.

About four spatial grid points are needed to represent a steep gradient or peak in any of the model equations to be solved numerically. The sharp peak anticipated in the electron heat source term due to simulated HF heating over an altitude range of 2 km requires 0.5 km grid spacing near the chosen HF reflection height. Since the pdepe numerical solver can incorporate a non-uniform spatial grid, one possible grid refinement procedure to cover a 10-km region that also leaves the total number of grid points fixed is to remove the highest 20 grid points from the set defined by equation (3.7)—using the parameters $z_1 = 80$ km, $z_N = 1800$ km, and $N = 120$ to specify base grid—and then take the union with the 20-point set $\{210 + \frac{1}{2}k \mid k = 0, 1, \dots, 20\}$ km. The altitude range of the fine grid spacing was chosen to be just below the F-region peak layer above Gakona at the time of the heating simulation, around midnight UTC, so the entire simulation grid consists of the points

$$\left\{ 80 \left(\frac{1032}{80} \right)^{\frac{k-1}{99}} \mid k = 1, 2, \dots, 100 \right\} \cup \left\{ 210 + \frac{1}{2}k \mid k = 0, 1, \dots, 20 \right\} \text{ km.} \quad (4.11)$$

Note that the diurnal equilibrium simulation described in section 3.2 also used this refined HF-heating simulation grid, so that the solution results from the diurnal equilibrium run could be used as initial conditions for the HF-heating run.

A simulation time step of 60 seconds was used in the HF heating numerical experiments, to allow for easier control of start and stop times of the HF heat source and subsequent data analysis. This

time step is much shorter than the 15 minute step used in the diurnal equilibrium simulation, chosen for computational and data storage economy.

The HF-electron interaction volume heating rate in the electron energy equation is defined here to be a step function in time and space: HF heating turns *on* in the region below the F-region peak where the plasma frequency was within 4.20 ± 0.06 MHz and remains *off* everywhere else. In each experiment, the HF heating rate was set to values between 0 and 94 nW/m^3 . The maximum heating rate was chosen by taking 12.5% of the peak 4.2 MHz HF energy flux at 200 km altitude, calculated to be $1.5 \times 10^6 \text{ nW/m}^2$ by the online HAARP IRI Transmitter Performance Calculator¹, and assuming that the HF energy is absorbed uniformly over an altitude region of 2 km resulting in a maximum HF-electron volume heating rate of 94 nW/m^3 .

There is a limitation to the SCIM HF-heating simulations, beyond the tendency for the fixed-plasma frequency HF interaction region to descend beyond the grid refinement region, that is: the electron energy equation (4.9) source function $s(z, t; n_s, T_s, \partial n_s / \partial z, \partial T_s / \partial z)$ evaluated at any particular point (z, t) may only be a function of the solution variables at that point. In other words, the effects of D- and E-region absorption of HF wave energy on F-region heating cannot be included within a single call to the pdepe solver, only between calls to the pdepe solver, because the simulated electron gas volume heating rate at a point can only depend on the local plasma frequency, and not on conditions below this point. One possible workaround to this limitation is to incorporate absorption effects between calls to the pdepe solver. This “leapfrog” process is also used to calculate the photoionization rate profiles since the absorption of solar ionizing radiation by the neutral atmosphere at altitudes above the point of interest can only be calculated self-consistently in between calls to pdepe. However, the two major factors driving the photoionization rate for given solar conditions, the neutral gas density and the solar zenith angle, change over time scales of several minutes while the drivers of HF energy flux absorption in the D and E regions change over time scales of several seconds or less. Therefore, the simulation time step of 60 seconds used in the heating simulations described in this chapter would likely need to be much

¹ <http://www.haarp.alaska.edu/haarp/calciri.html>

shorter, by at least one or two orders of magnitude, in order to self-consistently calculate HF energy flux absorption along the HF ray path.

4.3 Results

This section describes results from the 24 October 2008 and 21 October 2009 experiments at the HAARP facility, and SCIM results from the 24 October 2008 simulated heating experiment. These experiment and model results show that the HFIL reflection height descends during continuous HF-electron heating transmissions, consistent with prior work, and then subsequently ascends above its original altitude during low duty-cycle (minimum electron heating) pulsed transmissions. New observational results are also presented from an experiment conducted on 17 March 2010, showing that the descending HF-induced 557.7 nm airglow, reported by *Pedersen et al.* [2010], is localized at and above the HFIL reflection height measured with MUIR.

4.3.1 HFIL returns observed with UHF radar

Figure 4.6 and Figure 4.7 show range-time diagrams of received radar power signal-to-noise ratio measured with MUIR during the 24 October 2008 and 21 October 2009 experiments, respectively. The normalized phase-decoding algorithm described in section 4.2.3 calculated the signal intensity in each image; the mean noise level was determined from 5 s of signal between 200 and 300 km range, recorded during the 21 October experiment while the IRI was turned off. MUIR records the C996_4us baseband data from each pulse in *frames* of 500 consecutive pulses each. The frames are separated in time, typically between 0.1 and 0.8 seconds, by hardware design. Range-time diagrams in the figures are illustrated such that each frame is plotted immediately adjacent to the surrounding frames, without leaving a gap for the dropped pulses. Consequently, HF pulses from HAARP transmitted with a regular IPP occasionally appear with irregular spacing. This effect is apparent in Figure 4.6, where tick marks have been placed on the time axis to coincide with the MUIR frame boundaries, and the pair of HF pulses occurring on either side of the frame boundary are noticeably closer than any other pairs of successive HF pulses.

In the absence of HFIL returns, ground clutter is evident in the range-time diagrams as an elevated signal confined to one or two range bins that do not vary with time. HFIL signal is evident by an approximately 2-km thick return that varies with time in both intensity and range. The ground clutter signal intensity decreases at times of strong HFIL signal because the spectrum calculation at

the range bin containing the ground clutter includes samples at higher range bins that are affected by the turbulent plasma, thus affecting the maximum normalized frequency-component amplitude corresponding to the ground clutter (the 0 Hz component) that determines the range-time diagram signal intensity. Ground clutter signal intensity is comparable to the HFIL signals in either experiment using the normalized decoding algorithm. The SNR of the ground clutter and HFIL returns in the 21 October experiment are comparable.

24 October 2008 Experiment

The range-time diagram in Figure 4.6 covers the 4.2 MHz *heat-scan* sequence from 0020 to 0030 UT. The previous heat-scan sequence at 3.26 MHz did produce measureable HFIL returns. Hence, the prior experiment may have preconditioned the ionosphere and affected the measurements in the subsequent experiment. But this is not likely because ionosphere modification at a plasma frequency well under the critical frequency (and far below the critical layer), is seen in Chapter 5 to be an unlikely means of causing critical layer and topside modifications that persist for more than several seconds.

Reflection heights of the *scan* pulses ascend with time, indicating that the surrounding ionosphere density increases with (field-aligned) altitude, as expected, since the HF frequency increases by 22 kHz with each pulse in the 10-pulse sequence. The HFIL reflection height descends about 15 km during the first 90 s of each 180 s *heat* sequence before disappearing and beginning to descend again from near its original range between 190 and 195 km. Reflection heights from the following three *scan* sequences, each separated by 45 s of HF-off, ascend 5 km from their respective heights preceding the *heat* sequence. Figure 4.6 shows reflection heights of HF pulses from the four *scan* sequences. Reflection heights are determined by identifying the (range, time) bin with maximum decoded signal intensity. Uncertainty in the reflection height is bounded by the width of the HFIL return, by about 2 km.

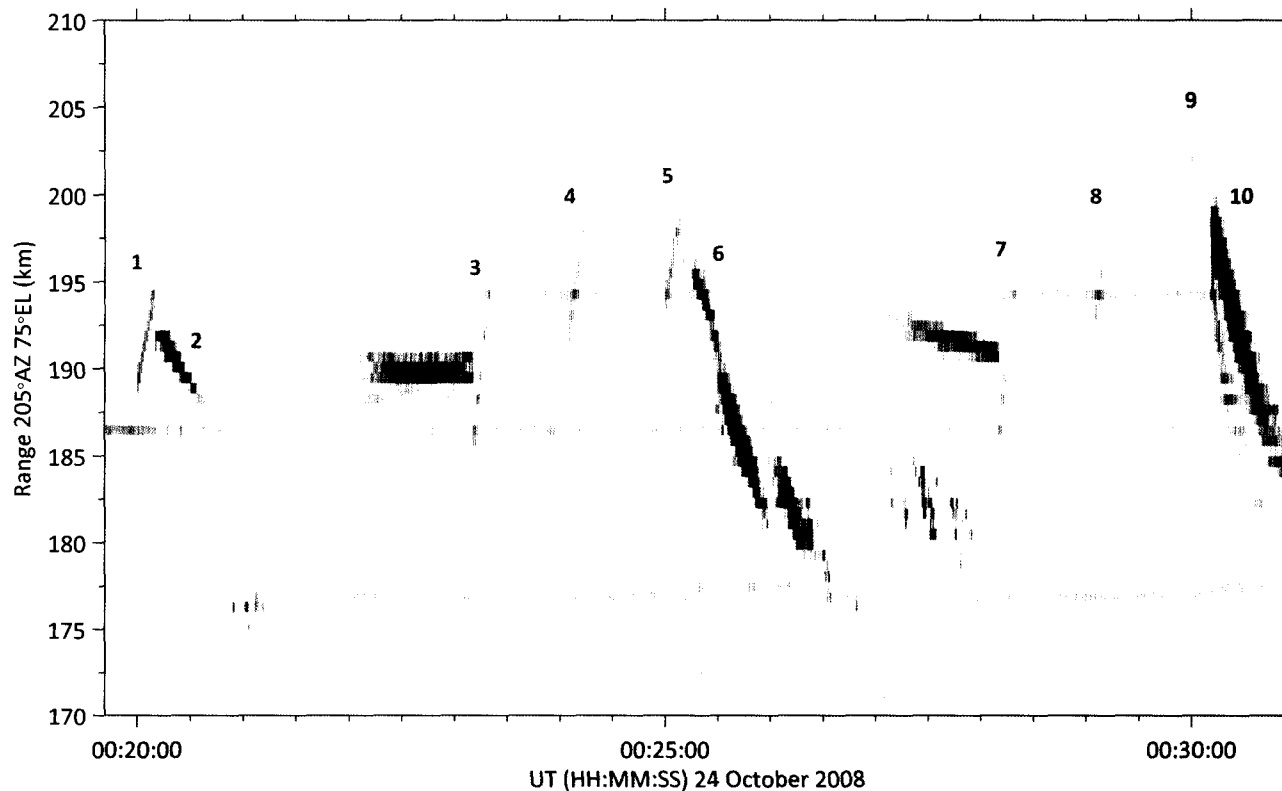


Figure 4.6 Range-time diagram of HAARP HF-enhanced ion-line (HFIL) power measured with MUIR during a HAARP *heat-scan* cycle on 24 October 2008. Both HAARP and MUIR are directed toward magnetic zenith. HAARP transmitted three 4.2 MHz *scan* sequences (1, 3, and 4) consisting of ten 1 ms pulses with 1 s IPP at 25% power starting at 4.1 MHz and ending at 4.298 MHz in 22 KHz steps, three 4.2 MHz scan sequences (5, 7, and 8) at 12.5% power, and two 4.2 MHz *heat* sequences (2, 6) at full power. The 4.55 MHz scan sequence (9) transmits 4.5 to 4.6 MHz in 11 kHz steps and with the 4.55 MHz heat sequence (10) is part of the next heat-scan cycle. MUIR transmitted a 996 μ s phase-coded pulse with 4 μ s baud length. Range to the HFIL return was calculated using the normalized decoding algorithm after muting the first 109 bits of the code. The HFIL during scan sequences (3-5) following the first heat sequence (2) are several kilometers further from MUIR than the original HFIL scan

21 October 2009 Experiment

Figure 4.7 shows the 4.06 MHz *heat-probe* sequence from 01:43:00 to 01:59:30 hours UT during the 21 October 2009 experiment. The previous 5.04 MHz *heat-probe* sequence from 01:35:00 to 01:42:30 UT did not result in measurable HFIL returns, yielding an upper bound on the ionosphere critical frequency in the direction of the geomagnetic field. HFIL reflection heights of the probe pulses are clustered around 192 km range during the first 6-minute *probe* sequence, with no discernable downward or upward trend. During the 5 min *heat* sequence, the HFIL descends 25 km in 45 s, disappears, then reappears at its original altitude and descends to 170 km range in 80 s, nearly repeating the initial cycle twice by the time *heat* sequence ends at 01:54:00 UT. The HFIL ascends to 205 km range and the signal intensity decreases throughout the following 5.5-minute *probe* sequence. A descending HFIL return starting at 177 km range and 02:00:00 UT is apparent following the second visible *probe* sequence. This return resulted from an unrelated experiment beginning at 0156 UT, transmitting 2.75 MHz O-mode waves at full power parallel to the geomagnetic field.

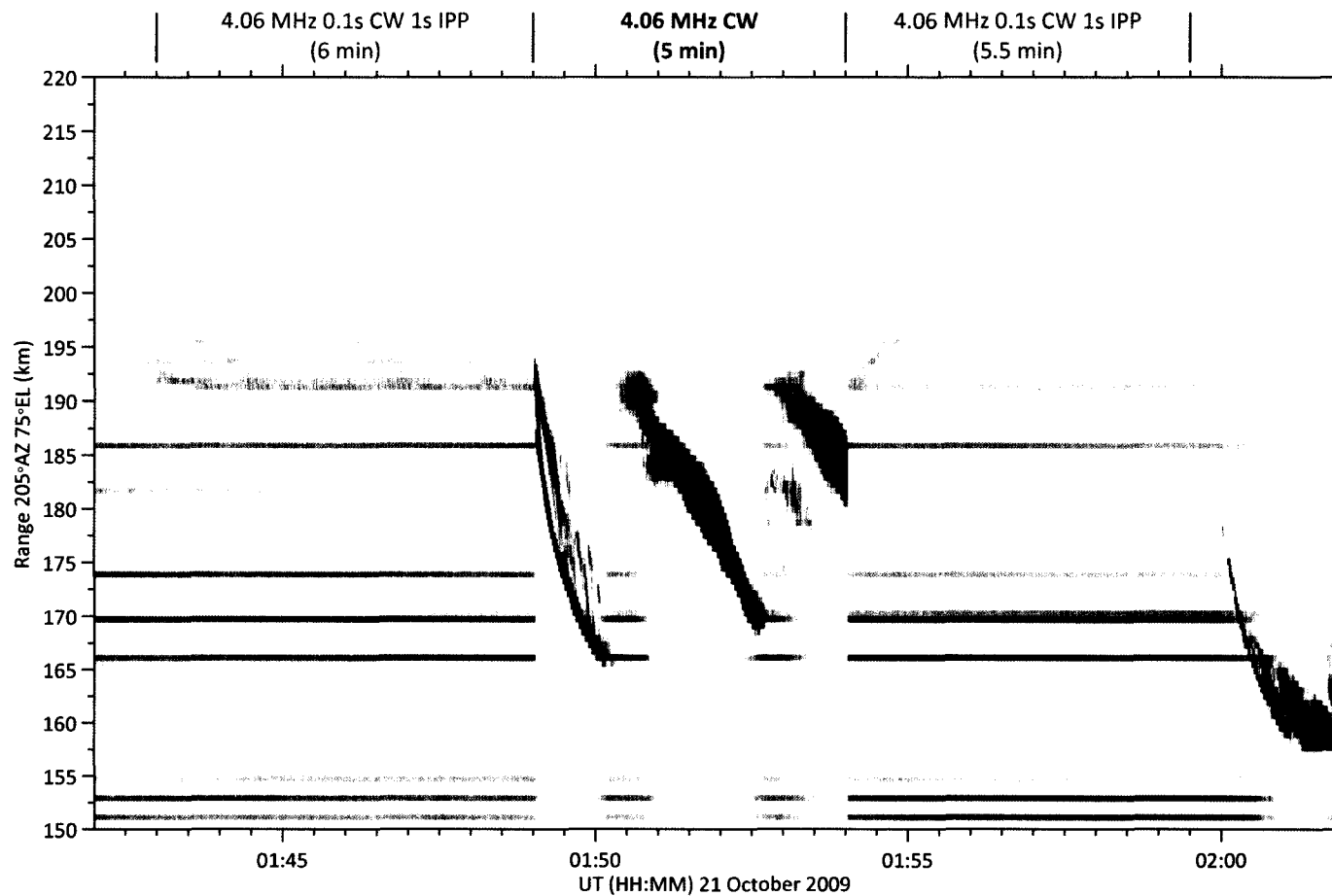


Figure 4.7 Range-time diagram of HAARP HF-enhanced ion-line (HFIL) power measured with MUIR during a HAARP *heat-pulse* cycle on 21 October 2009. Both HAARP and MUIR are directed toward magnetic zenith. HAARP transmitted a 6 min *pulse* sequence consisting of 0.1 s pulses with 1 s IPP at 4.06 MHz from 01:43:00 to 01:49:00 hours UT, a 5 min heat sequence at 4.06 MHz from 01:49:00 to 01:54:00 UT, and a 5.5 min 4.06 MHz pulse sequence from 01:54:00 to 01:59:30 UT. The HFIL return beginning at 02:00:00 UT is from an unrelated subsequent experiment, where HAARP transmitted O-mode at 2.75 MHz along the geomagnetic field. MUIR data processing parameters are the same as those used for Figure 4.6.

An ionogram recorded at the beginning of the 21 October 2009 experiment is shown in Figure 4.8. The ionosonde is co-located with the HAARP IRI. The 4.06 MHz HF frequency used during the experiment is indicated on the ionogram with its corresponding reflection height of 195 km. Compare this to the reflection height of 186 km measured with MUIR during the initial pulse sequence depicted in Figure 4.7, after correcting for the non-zero dip angle of the geomagnetic field. Reflection heights measured by MUIR during the *scan* and *heat* sequences are in good agreement with the heights measured with the ionosonde, given the ionosonde profile uncertainty and the 1 km difference between the matching and reflection heights (illustrated in Figure 4.13).

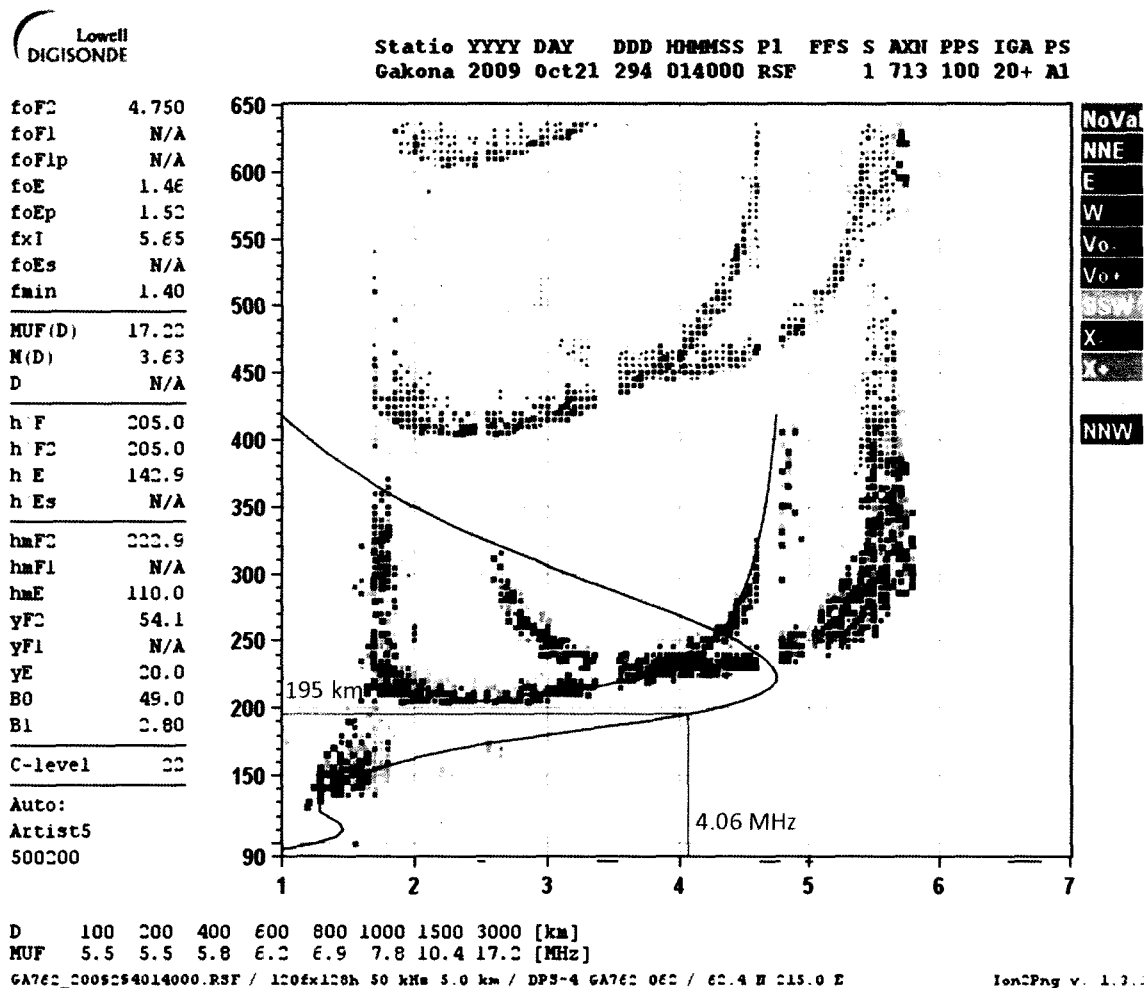


Figure 4.8 Inverted ionogram showing the O-mode 4.06 MHz reflection height over HAARP before the start of the 21 Oct 2009 experiment. The horizontal axis is expressed in MHz and the vertical axis is expressed in km.

4.3.2 Time-dependent simulation of HFIL reflection height

The diurnal equilibrium run discussed in section 3.2 provided initial conditions for the heating experiment simulation starting at 0000 hours UT model time. Simulated HF electron heating at regions below the electron density peak where the plasma frequency was within 4.20 ± 0.06 MHz began at 00:00:00 UT and ended at 00:05:00 UT. While the two experiments described above occur at local times separated by about an hour and a half and HF frequencies separated by nearly 200 kHz, the simulated 4.2 MHz HF heating experiment starting at 00:00:00 UT is used to interpret the results of both experiments. Figure 4.9 Altitude profile of the electron plasma frequency in a diurnal-equilibrium model ionosphere on 24 October 2008 over Gakona, Alaska. shows the electron density profile, expressed as a plasma frequency profile calculated with SCIM at 00:00:00 and 00:01:00 UT. While the electron density in the does decrease appreciably as the sun further below the horizon, especially in E and lower F regions, the difference does not alter the main results or conclusions.

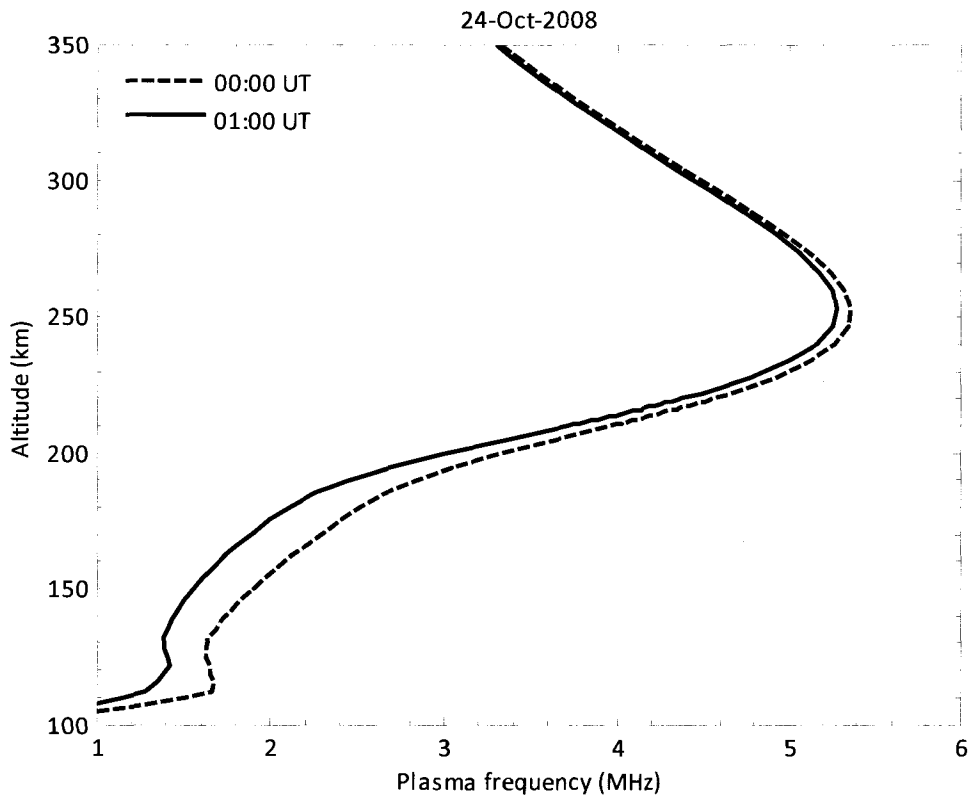


Figure 4.9 Altitude profile of the electron plasma frequency in a diurnal-equilibrium model ionosphere on 24 October 2008 over Gakona, Alaska.

Let $z_{\text{HF}}(t, \omega_{\text{HF}})$ denote the time-dependent altitude in the HF-interaction region where the ionosphere plasma frequency equals the transmitted HF frequency:

$$\omega_p(t, z_{\text{HF}})/2\pi = \omega_{\text{HF}}(t)/2\pi = 4.20 \text{ MHz} \quad (4.12)$$

Unless stated otherwise, all results and figures describe a heating simulation where the electron heat source term in equation (4.8) was stepped up by 94.6 nW/m^3 in the HF-interaction region. The HF-interaction region z_{HF} in the heating simulation descended from 213 km to 208 km altitude during the five minutes of heating, but it remained at a nearly constant altitude during the respective time period in the control simulation.

Figure 4.10 shows calculated altitude profiles of electron and ion temperature immediately before and after simulated HF heating, and 5 min beyond the end heating. Profiles of electron number density are plotted in Figure 4.11. The HF-interaction region at the altitude z_{HF} is a clearly identifiable peak in the plot of electron temperature, and also to a lesser extent as a notch in the plot of electron number density.

At 00:01 UT, just one minute after starting the HF heat source, the perturbed electron temperature $T_e' = T_e + \delta T_e$ in the HF-interaction region surrounding z_{HF} is nearly 3000 K warmer than the ambient electron temperature, and is just 200 K cooler than its final perturbed temperature 00:05 UT when the heat source is removed. In the lower F region, extending nearly 50 km below z_{HF} , the temperature perturbation δT_e decreases linearly with decreasing altitude until $T_e' \approx T_e$ at 150 km altitude. The electron temperature perturbation is nearly altitude-independent above 300 km, resulting in a near-isothermal electron temperature profile in the modified topside ionosphere. Electron temperature in the modified topside ionosphere continues to rise while the HF heat source is active.

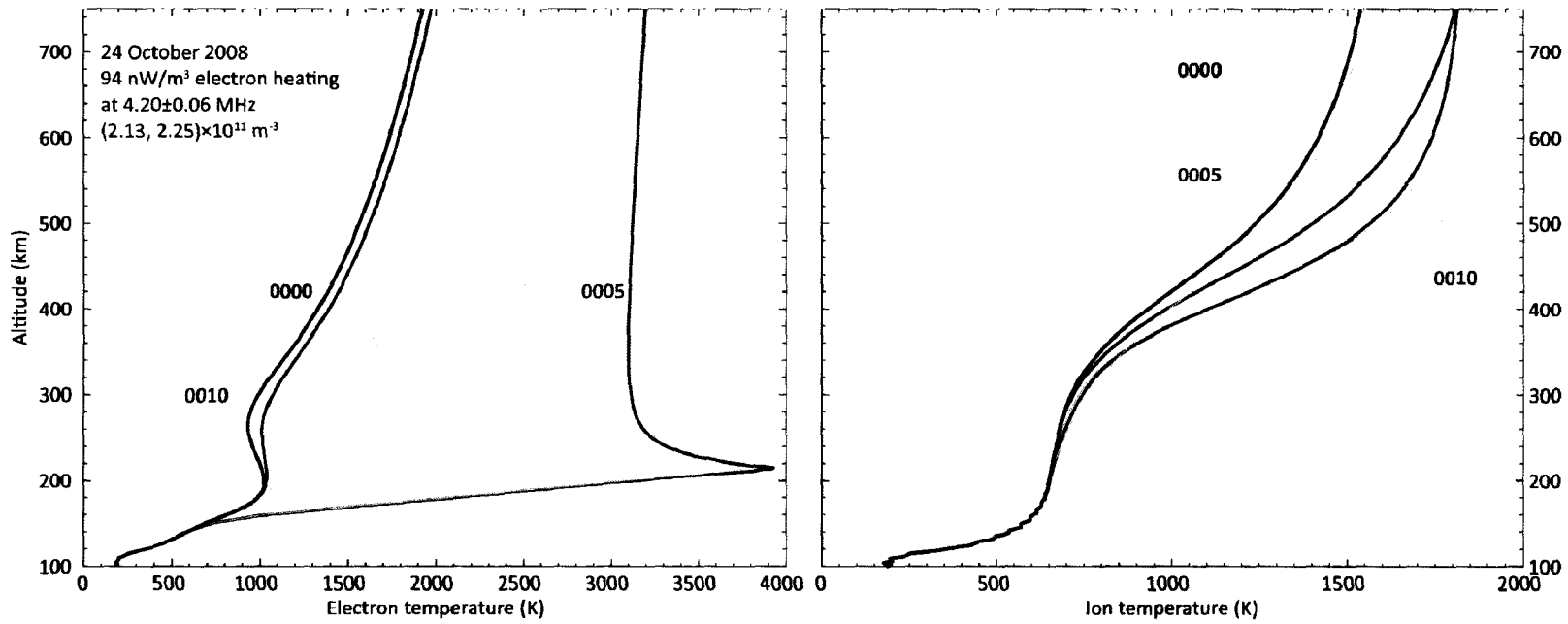


Figure 4.10 Electron (left) and ion (right) temperature profiles calculated at the times (hhmm UT) indicated by the labels. Simulated HF-electron heating began at 0000 UT and ended at 0005 UT.

The sharply peaked electron temperature profile evident in Figure 4.10 indicates that the pressure gradient force, i.e. a net force per unit volume, may be significant. If that force is large enough, then the hydrostatic equilibrium assumption used in deriving the model continuity-momentum equation may be invalidated because the electrostatic-coupled ion gas may be accelerated with a magnitude significantly larger than other acting forces like gravity. A quick estimate of this vertical ion acceleration during the first minute of heating, may be calculated by taking the electron pressure gradient force (density) and dividing by the mass density of a unit volume of atomic oxygen ions $f_{\nabla p_e} / \rho_{O^+}$. Relevant values may be read from the temperature profile in Figure 4.10 and the equilibrium density profile in Figure 4.11:

$$\begin{aligned}
 f_{\nabla p_e} &= -\frac{\partial p_e}{\partial z} \\
 &\approx -n_e k_B \frac{\Delta T_e}{\Delta z} \\
 &\approx -(2.4 \times 10^{11} \text{ m}^{-3})(1.4 \times 10^{-23} \text{ m}^2 \text{ kg s}^{-2} \text{ K}) \left(\frac{3900 - 2900 \text{ K}}{220 - 240 \text{ km}} \right) \\
 &\approx 1.7 \times 10^{-13} \text{ N m}^{-3}
 \end{aligned} \tag{4.13}$$

The atomic oxygen ion mass density ρ_{O^+} at 220 km altitude is approximately $5.5 \times 10^{-15} \text{ kg m}^{-3}$, so the resulting acceleration is 30 m/s^2 upwards, or about three times the gravitational acceleration value. This indicates that hydrostatic equilibrium is not a good assumption near the interaction region, and that the model is likely underestimating ion flux near the region. However, the role of transport in determining plasma density in the lower F region is secondary to chemical production and loss, so model results are not likely to be significantly affected. Also note that the gradient just above the local electron temperature maximum decreases considerably by about half after five minutes of heating.

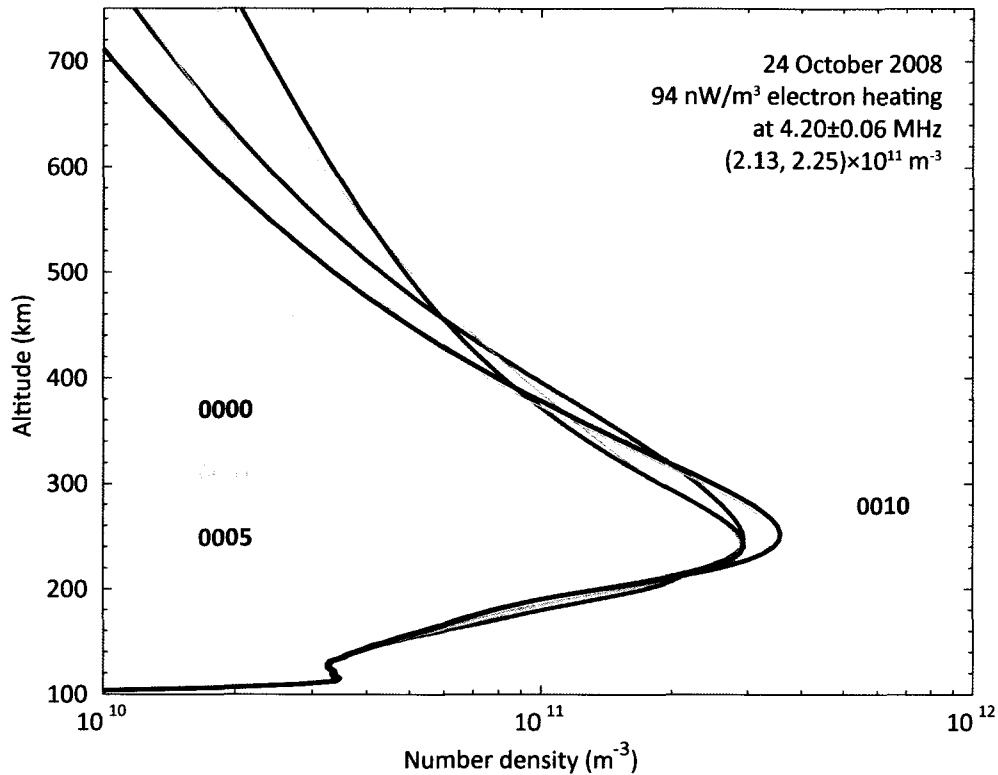


Figure 4.11 Electron number density profiles calculated at the times (hhmm UT) indicated by the labels. Simulated HF-electron heating began at 0000 UT and ended at 0005 UT. The approximate electron number density with a plasma frequency equal to the chosen HF heating frequency is identified with a dashed line.

Ion temperatures respond to HF electron heating on a much slower time scale than the electrons. No ion temperature perturbation occurs at or below the interaction region containing z_{HF} . Nearly 300 km above z_{HF} in the topside ionosphere, the nearly-isothermal ion temperature profile above 500 km steadily increases from 1500 to 1800 K for the duration of HF heating.

The electron density below z_{HF} , illustrated in Figure 4.11, increases 25% to a new equilibrium value less than 1 min after starting the HF heat source. This density increase results in a descending z_{HF} to a new equilibrium value. The electron density increase ensues from the decrease of NO^+ and O_2^+ recombination rates with increasing temperature, since ion number density in the lower F region is determined by the ratio of production and loss rates. (Recall equation (3.18) and see

section 4.4 below.) Above z_{HF} and extending to 350 km altitude, about 100 km above the F-region peak layer, the electron density decreases by 25% by the end HF heating. This result is in agreement with EISCAT ISR measurements of HF-induced density depletions [Rietveld *et al.*, 2003]. The modeled depletion of the critical layer n_e results in a 500 kHz reduction in foF2, or nearly 10%. The topside altitude profile of (logarithmic) electron density gradient decreases during heating but otherwise remains independent of altitude. At 600 km altitude n_e' is 60% larger than n_e .

At 00:05 UT when the simulated HF-electron heat source ceases, the electron volume heating rate returns immediately to the ambient value driven by collisions with non-thermal photoelectrons and secondary electrons. Then the elevated plasma temperature and density profiles begin to relax. The characteristic times for the plasma density and temperature to return to their equilibrium temperature depends on altitude, similar to the characteristic times of ionosphere plasma density and temperature perturbations during HF heating.

The peak electron temperature immediately surrounding interaction region z_{HF} rapidly returns from 390% to 110% of its initial equilibrium value of 1000 K in the first minute following the end of HF heating. Electron temperature responds over a longer characteristic time of in the topside ionosphere than at or below the critical layer. At 450 km in the topside ionosphere, T_e' falls from 3100 to 2000 K, or from 200% to 130% of the equilibrium $T_e = 2000$ K. And at 5 min following the end of HF heating at 0010 hours, the topside T_e' above 250 km is maintained at 5-10% above T_e by the more slowly cooling ions.

Ion temperatures above 450 km continue to increase, up to 1850 K at 700 km, in the first minute after removal of the HF-electron heat source. T_i' subsequently begins to drop, but at 700 km T_i' remains 250 K or 17% above T_i at 0010 hours. The possible influence of persistently elevated ion temperatures on TEC measurements following certain HF modification experiments is discussed in Chapter 5.

Following the rapid drop of T_e' below the critical layer height z_mF2 , the enhanced electron density in the lower F region—resulting from elevated concentrations of O_2^+ and NO^+ —returns to its equilibrium value in the first minute after removal of the HF heat source. The critical layer

electron density, depleted by 25% during heating begins to rebuild by photoionization and recovers almost completely by 0010 UT. Topside n_e' is still elevated 33% above n_e at 0010 UT above 350 km even though the perturbed vertical (logarithmic) density gradient $\partial(\ln n_e')/\partial z$ has relaxed to the equilibrium gradient. See sections 4.3.2 and 5.2. (Figure 5.7 in particular) for a more detailed illustration of the spatiotemporal evolution of the electron density perturbation.

4.3.3 Concurrent radar and optical measurements of HF-induced structures

A recent experiment reported by *Pedersen et al.* [2010] showed remarkable HF-enhanced 557.6 nm “green line” airglow emissions that descended by more than 50 km. The MUIR radar was operating in ion-line mode during the experiment. This section shows the first direct comparison of the spatiotemporal evolution—along the geomagnetic field line—of the HFIL reflection altitude with the HF-enhanced 557.7 nm airglow emission altitude.

The HAARP ionosonde may be used to quickly validate the MUIR coded long-pulse range-decoding algorithm for identifying the range to HFIL targets by comparing the observed HFIL range to the range expected by locating the HF reflection height by inspection of the inverted ionogram profile. Poor spatiotemporal resolution of the ionosonde limits the precision of the validation, however. Instead of validating the HFIL measurements via ionosonde data, HF-induced airglow intensity along the HAARP field line measured during a typical on-off O-mode transmission cycle are compared with MUIR measurements of the HFIL field-aligned range in Figure 4.12. Powerful HF transmissions can produce artificial airglow at low, mid, and high latitudes. (See *Gustavsson et al.* [2001], e.g., for further references.) The two most common HF-enhanced emissions at 557.7 and 630.0 nm from the $O(^1D)$ and $O(^1S)$ excited states, respectively, with characteristic radiative lifetimes 0.7 and 107 s. (Although $O(^1S)$ is usually quenched in about 30 s at 225 km altitude [*Gustavsson et al.*, 2002].) The thermal velocity of atomic oxygen at 1000 K is about 720 m/s; so given an optical imager of sufficient sensitivity and resolution to measure the 557.7 nm airglow, the HF-interaction region may be localized with high spatiotemporal resolution of approximately 500 m every 1 s. This is nearly the same resolution as the MUIR coded long-pulse measurements processed with the range-decoding algorithm.

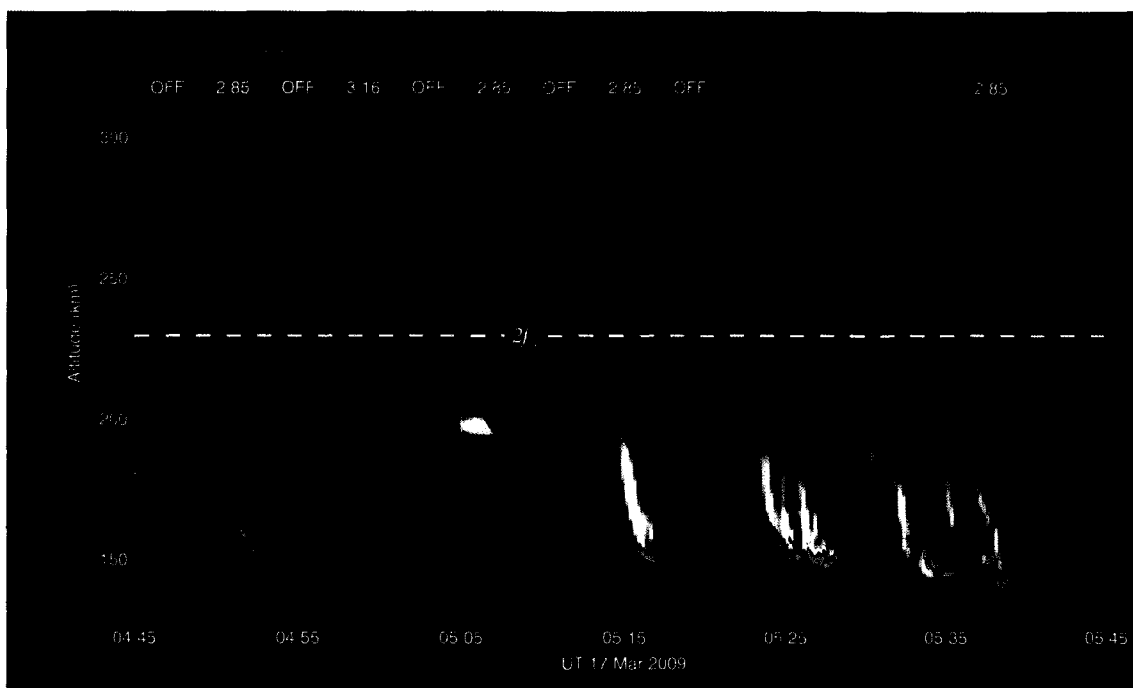


Figure 4.12 Field-aligned 557.7 nm airglow (white) and HF-enhanced ion-line intensity (copper) over the HAARP facility in Gakona, Alaska. A wide-field imager located about 160 km north of HAARP measured the airglow intensity. MUIR, collocated with HAARP, measured ion-line intensity. The IRI transmitted O-mode polarized waves parallel to the geomagnetic field at either 2.85 or 3.16 MHz in an on-off cycle, as indicated by the vertical lines. White horizontal specks are stars passing through the imager field of view. Copper horizontal lines are artifacts from ground-clutter returns from MUIR side-lobes. The dashed horizontal line indicates the altitude where the second electron gyroharmonic is 2.85 MHz, as calculated by an empirical model of the geomagnetic field. See *Pedersen et al. [2010]* for the original airglow image and description, including an overlay of reflection heights measured with the HAARP ionosonde.

Measurements of the time- and range-dependent HFIL intensity are compared with corresponding 557.7 nm airglow measurements in Figure 4.12, validating the MUIR coded long-pulse range-decoding algorithm for identifying time-dependent field-aligned altitudes of strong HF-induced plasma turbulence. The location of strongest HFIL intensity, measured up the field line by MUIR, corresponds with the field-aligned green-line airglow observed from the remote site north of HAARP.

The HAARP IRI transmitted an on-off cycle during the time period depicted in the figure, using O-mode polarization at either 2.85 or 3.16 MHz in the field-aligned direction from 04:45 to 05:45 UT. A wide-field imager stationed at a remote location, about 160 km N of the HAARP facility in

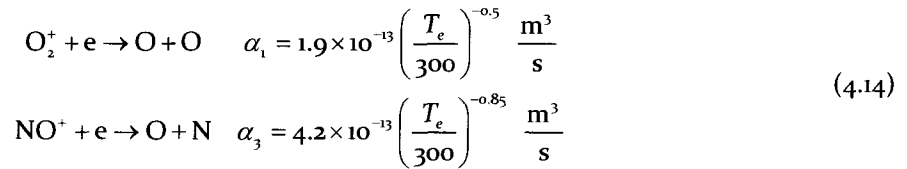
Gakona, measured the 557.7 nm airglow along the HAARP geomagnetic field line. MUIR, co-located with the HAARP IRI, measured the ion-line channel in the field-aligned direction using the C996_4us mode. Figure 4.12 is modified from the original image by *Pedersen et al.* [2010] by the overlaying the decoded backscatter intensity from MUIR. The field-aligned altitude-time HFIL intensity is calculated with the normalized range-decoded algorithm with 100 pulse (about 1 s) integration. Radar signal-to-noise ratio increases toward lighter colors. The first 109 of 249 bits of the code were muted during the 100-point Fast Fourier Transform processing to reduce significant ground clutter effects amongst the strong HFIL returns occurring near 150 km altitude.

Returns from the HFIL are clearly bounded between the lower limit and middle of the airglow intensity even as the airglow descends from 200 to 150 km altitude. The agreement between the time-dependent 1-D airglow and radar measurements during the 1-hour experiment of multiple HF pumping cycles validates the MUIR normalized range-decoding algorithm for identifying regions of significant energy dissipation. With the exception of the 2.85 MHz transmission starting at 0505 UT, the HFIL appears between 210 and 220 km range, descends 30 km in about 2 min before disappearing and then immediately reappearing with greater intensity at 160 km range, finally descending another 10 km in the remaining 2 min. HFIL signals from the long-duration 2.85 MHz transmission starting at 0521 UT appear to follow the sequence described, terminate at 0528 UT, and begin again just below 200 km range. The repeating descending HFIL signal has been observed in several prior experiments at HAARP that use full power, field-aligned, O-mode transmissions that last longer than a few minutes. See, for example Figure 4.6 and Figure 4.7.

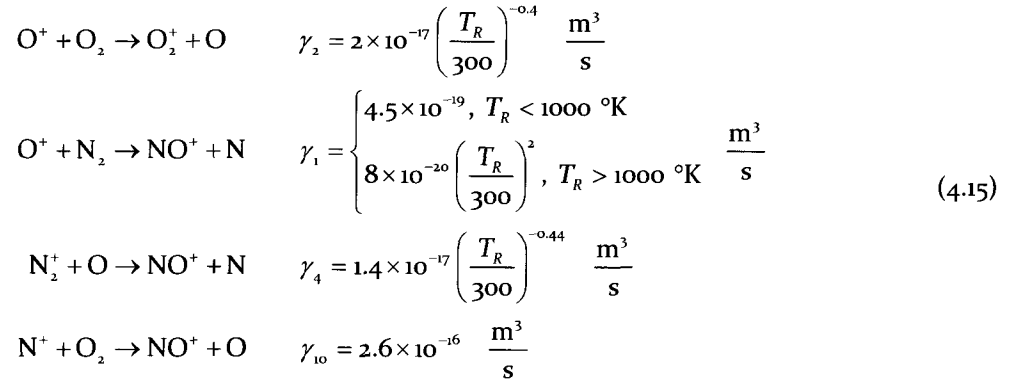
4.4 Discussion

High power HF waves incident upon the ionosphere induce plasma instabilities, where the HF wave decays into electrostatic plasma waves that quickly dissipate heat into the electron gas. MUIR detects HF-induced electron Langmuir and ion-acoustic waves in the ionosphere where the wave number of the driven waves satisfies the matching condition determined by the plasma number density, temperature, HF frequency, and radar frequency. Heat from the dissipating electrostatic waves conducts through the electron gas along the geomagnetic field, resulting in large-scale density and temperature perturbations. One manifestation of these density changes is a change in the HFIL and HF reflection heights. Figure 4.13 shows the HFIL reflection height of the perturbed ionosphere, calculated for the 446 MHz transmit frequency of MUIR, as a function of time against the HFIL reflection height of a control ionosphere.

Since ground-based HF heating occurs below the F-region peak density layer where the electron temperature and density generally increase monotonically with height, the observed descent of the HFIL during heating may be attributed—through the matching condition—to an increase of electron temperature and density. The electron density in the lower F region is governed primarily by the balance between the local production and loss rates of the major constituent ions O^+ , O_2^+ , and NO^+ . Downward transport of O^+ along the geomagnetic field line is a significant temperature-dependent source of O^+ in the lower F region below the peak photoproduction layer apparent in Figure 2.3, but is not appreciably affected by HF heating of the electrons because the rate of O^+ loss increases exponentially with the molecular neutral number density with decreasing height. Electron recombination is the most significant loss mechanism of O_2^+ and NO^+ :



The reaction rates decrease with increasing electron temperature, but the photoproduction O_2^+ and the major chemical production sources of O_2^+ and NO^+ , described by the reactions



(from Rees [1989]) are not affected by electron temperature variations. A net increase of electron number density results in the lower F region. Note that the O_2^+ and NO^+ loss rates vary directly with the electron number density, limiting the density enhancements caused by a heating a layer of electrons with constant power input. The production of the other significant F-region ion, O^+ , by

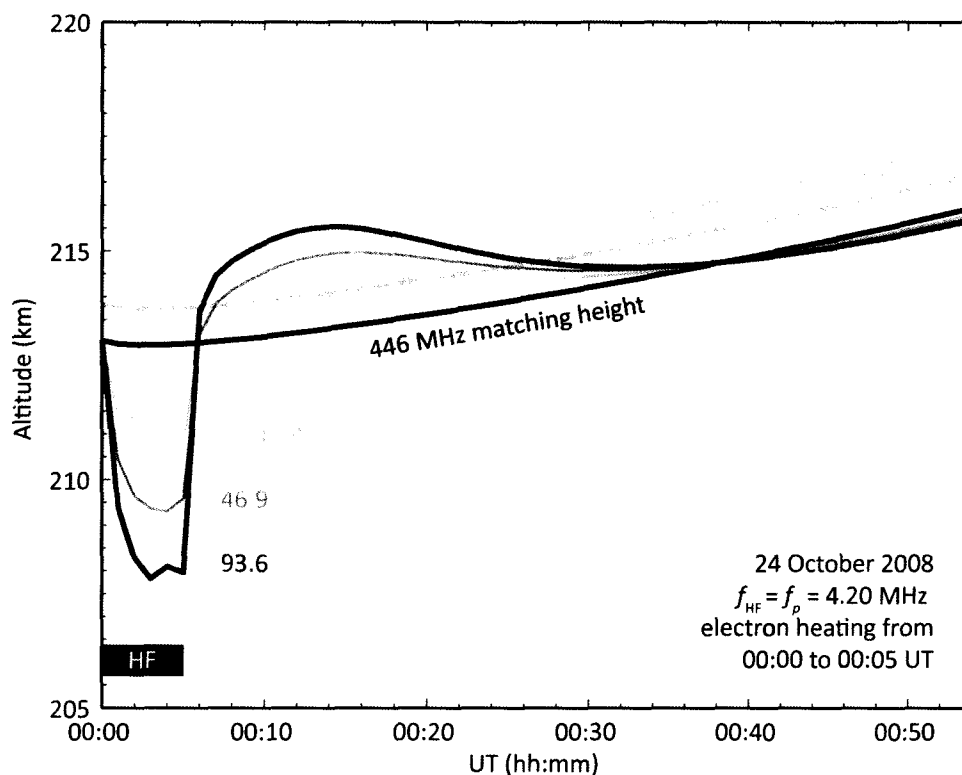


Figure 4.13 Time-variation of the 446 MHz HFIL matching height (red) perturbed by 5 min of simulated HF pumping at the bottom-side 4.2 MHz layer and (black) without HF pumping. Sensitivity of the matching height perturbation to the HF-electron volume heating rate is indicated by shading. The red bar indicates the time range of simulated HF-heating. A dashed line shows the natural variation of the 4.2 MHz reflection height.

photoionization of atomic oxygen and the destruction of O^+ by reacting with O_2 and N_2 are unaffected by changes in the electron temperature.

The modeled HFIL descent illustrated in Figure 4.13 is less in magnitude than the descent measured with MUIR shown in Figure 4.6 and Figure 4.7. Some difference is to be expected on account of differences between model used in the heating simulation and the natural ionosphere. Specifically, one potentially major uncertainty affecting the HFIL reflection height is the difference between the time-dependent modeled one-dimensional HF-interaction electron heat source and the observed three-dimensional HF-plasma interaction effects (e.g., see Kosch *et al.* [2007]) that presumably correspond with the physical HF-electron heating mechanism. The difference in time scales of the

simulated HFIL descent and subsequent ascent perturbations may be understood by the different characteristic times that the lower and upper F regions reach equilibrium following the introduction or removal of the HF-electron heating disturbance.

Near z_{mF2} , depleted by 25% during 5 min of 94 nW/m^3 simulated HF-electron heating, O^+ is the dominant ion constituent. Divergent transport loss in the natural ionosphere is comparable in magnitude to the local production and loss rates. The rates of major O^+ reactions with O_2 and N_2 are independent of electron temperature, but not independent of ion temperature. However, the model ion temperature is not significantly modified near the critical layer. Therefore HF-enhanced electron temperature depletes O^+ ions near the critical layer through enhancement of the ambipolar diffusion coefficient, encapsulating the physical role of divergent transport in the hydrostatic balance of the field-aligned ionosphere profile. The F-region peak density reaches a new equilibrium when the topside density gradient relaxes to the point that divergent transport loss once again balances the local photoionization production rate.

The repeated cycle of HFIL descents detected with MUIR that are apparent in Figure 4.7 during the 21 October experiment is difficult to explain with a slab-symmetric ionosphere model and one-dimensional radar observations. Complicated HF ray path geometry is three dimensional and time dependent during each experiment, since the index of refraction determining the HF band depends on the electron number density along the path. Multidimensional HFIL imaging and modeling is likely needed to interpret the measurements of repeated HFIL descents during a single period of HF modification.

Immediately following the end of the HF heating period at 00:05 UT, the electron temperature below the critical layer returns to its ambient temperature by 00:06 UT, along with the O_2^+ and NO^+ recombination rate coefficients. The characteristic times of ionization build-up due to local production τ_p and loss τ_l may be estimated by neglecting the divergent transport term in the continuity equation and linearizing the remaining terms with respect to their corresponding characteristic times:

$$\begin{aligned} \frac{\partial n_i}{\partial t} &= \left| \frac{\partial n_i}{\partial t} \right|_{\text{chemical production}} - \left| \frac{\partial n_i}{\partial t} \right|_{\text{chemical loss}} \\ &\approx \frac{n_i}{\tau_p} - \frac{n_i}{\tau_l} \end{aligned} \quad (4.16)$$

Following a sudden imbalance of, for example, when the loss term suddenly exceeds the production term as when HF-enhanced electron temperatures in the lower F region abruptly drop by 1000 K, the relevant recovery time constant is approximately:

$$\tau_l \approx \left(\frac{1}{n} \left| \frac{\partial n_i}{\partial t} \right| \right)^{-1} = n_i \left| \frac{\partial n_i}{\partial t} \right|_{\text{chemical loss}}^{-1} \approx L_i^{-1} \quad (4.17)$$

A loss constant value of $L_i = 2 \times 10^{-2} \text{ s}^{-1}$ can be estimated from the combined O_2^+ and NO^+ loss constants in Figure 2.4 at 175 km altitude. The resulting characteristic time $\tau_l \approx 50 \text{ s}$ to lose the HF-enhanced ionization build-up estimated from model calculations are in agreement with the observed return of the descended HFIL back to its original altitude—and beyond—within seconds of switching from the continuous HF-heat sequence to the low duty-cycle HF-probe pulses in each experiment.

Once the enhanced O_2^+ and NO^+ plasma in the lower F region recombine with the cooled electrons, the HF and HFIL reflection heights ascend closer to the O^+ peak layer. Although the electron temperature surrounding the (modified) zmfz quickly (10s of seconds) returns to the ambient temperature, the peak plasma frequency foFz recovers on a time scale an order of magnitude larger. The peak layer plasma at 250 km altitude is maintained by local photoproduction and field-aligned transport from the peak photoproduction layer, about 50 km below. Plasma production exceeds losses in the critical layer until the density gradient and plasma density increase to the point that divergent transport with the molecular neutral gases increase to balance the relatively time-independent photoproduction rate.

Consequently the F-region plasma density at the production layer peak, just below the critical layer, may remain depleted for up to tens of minutes and the HFIL will remain at an altitude above its prior unmodified equilibrium. The time constant to rebuild plasma density immediately below the critical layer may be estimated from the production time constant, assuming the losses due to transport and chemical loss immediately following a disruption to equilibrium:

$$\tau_p \approx \left(\frac{1}{n} \left| \frac{\partial n_i}{\partial t} \right| \right)^{-1} = n_i \left| \frac{\partial n_i}{\partial t} \right|_{\text{chemical production}}^{-1} \approx \frac{n_i}{P_i}. \quad (4.18)$$

Taking $n_i = 10^{11} \text{ m}^{-3}$ to be the atomic oxygen ion density and $P_i = 2 \times 10^8 \text{ m}^{-3}\text{s}^{-1}$ to be the atomic oxygen photoproduction rate, both estimated from Figure 3.2 and Figure 4.14 at 190 km altitude, yields a production time constant of 500 s. This is approximately the time that an ascended HFIL may be expected to remain significantly above its equilibrium altitude, in absence of horizontal convective disturbance or transport. Unfortunately, this is a few minutes longer than the time to the next modifying HF-heat sequence, so the HFIL return to equilibrium altitude was not observed.

The production time constant just below the critical layer depends on the O^+ photoproduction rate, which varies with solar zenith angle, the intensity of the ionizing solar radiation, and the semi-diurnally varying neutral gas density profile. The total photoproduction rate profile calculated at 0000 and 0100 UT is plotted in Figure 4.14. Additional experiments performed at different times of day and geophysical conditions may show a measureable difference in the HFIL recovery time following modifying HF-heat sequences of several minutes duration.

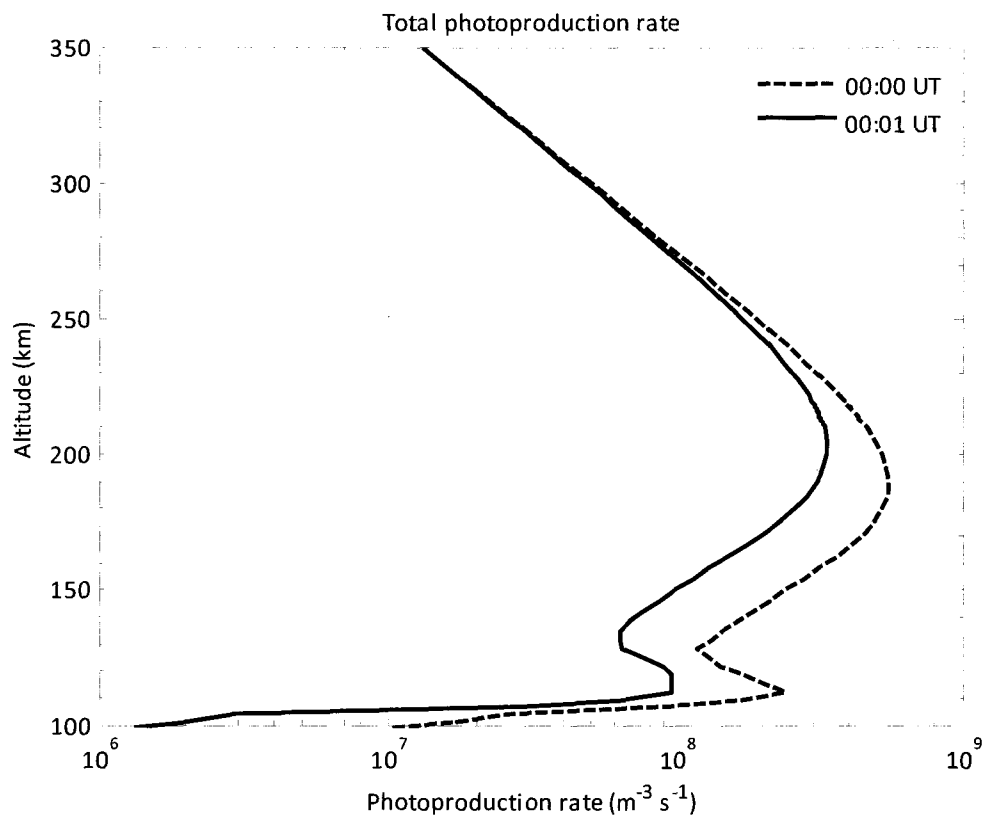


Figure 4.14 Altitude profile of the electron photoproduction rate in a diurnal-equilibrium model ionosphere on 24 October 2008 over Gakona, Alaska.

One important conclusion may be drawn from the model and experiment results is that ionosphere preconditioning may be a significant contaminating factor in high-latitude O-mode HF-modification experiments during quiet geophysical conditions similar to those described here. That is, many experiments use HF on-off cycles with periods of five minutes or less, and thus fail to let the ionosphere fully relax between subsequent HF transmissions. The experiment described by *Pedersen et al.* [2010], for example, is a high-latitude O-mode HF-modification experiment occurring in the late local evening uses a 50% duty cycle of 4 min “on-off,” specifically to “allow recovery from artificially induced effects.” The duty cycle described likely leaves an insufficient amount of time to allow the ionosphere to relax completely.

Chapter 5. Total electron content enhancements

HF heating the polar ionosphere was shown in Chapter 4 to cause field-aligned electron density modifications, consistent with previous incoherent scatter radar (ISR) measurements [*Rietveld et al., 2003*], high-latitude ionosphere model results [*Mingaleva and Mingalev, 1997; 2008; Perrine et al., 2006*], and mid-latitude ionosphere model results [*Perkins and Roble, 1978*]. This chapter explores the effect of these density perturbations on the slant total electron content (TEC) in the direction of the geomagnetic field. Recent experiments at HAARP, described below, have demonstrated that O-mode heating below foF₂ lead to increases of slant TEC of up to .5 TECu (10^{16} m^{-2}).

Self-consistent ionosphere model (SCIM) results presented in section 5.2 demonstrate that bottomside electron heating, similar to heating that occurs during powerful HF transmissions, causes TEC increases that occur (and later dissipate) on a fast time scale of several seconds, and also causes increases that occur and dissipate on a time scale of several minutes. The fast TEC increase is found to result from the electron density increase in the lower F region resulting from electron temperature enhancements that decrease the recombination rates of the molecular ions O₂⁺ and NO⁺, while the slow TEC increase is a consequence of the topside electron density enhancement that results from enhanced temperature-dependent ambipolar diffusion.

The discussion in section 5.3 presents measurements of enhanced ion flux and density in the topside ionosphere over HAARP during O-mode heating made by a Defense Meteorological Satellite Program (DMSP) satellite provide additional evidence for the topside density enhancement and mechanism inferred from model calculations. ISR measurements made during polar ionosphere HF modification experiments at Tromsø also show HF-induced ion flux consistent with SCIM calculations and the DMSP measurements. SCIM calculations show that the topside density enhancement depends strongly on the altitude where the electron gas is heated, determined primarily by the HF frequency chosen and the ambient electron density profile. Heating at higher altitudes, where the electron temperature and associated electron thermal conductivity is greater, leads to much more efficient heating of the electrons in the topside ionosphere. Consequently, topside electron heating leads to ion heating, field-aligned ion flux, and atomic ion density enhancements.

This chapter is a joint work with James Secan of NorthWest Research Associates, Inc., who is a co-author. He also analyzed the GPS data to provide measurements of TEC, and the DMSP scintillation data to provide in-situ measurements of ion density and velocity. F. J. Rich of the Massachusetts Institute of Technology Lincoln Laboratories and K. R. Martin of the Institute for Scientific Research at Boston College provided raw data from DMSP satellite passes from the Air Force Research Laboratory database. Significant portions of this chapter appeared in previous conference presentations [*Fallen and Watkins, 2008; Fallen et al., 2009; Secan et al., 2008*].

5.1 Experiment

The slant TEC data shown in Figure 5.1 were taken during an experiment run at the HAARP facility on 14 October 2009. The data are TEC derived from the observed differential group delay (DGD) and differential carrier phase (DCP) on signals from the GPS PRN 30 satellite using an Ashtech Z-FX receiver located at the HAARP facility. The solid vertical lines in the plot indicate times that the HAARP heater was turned on, and dashed vertical lines indicate when the heater was turned off. In both examples shown, the heater was configured with a narrow beam using a frequency of 4.80 MHz in O-mode CW operation. The frequency was selected to be slightly below the critical frequency of the F₂ layer (f_oF_2), which was estimated from soundings taken by the Digisonde located at the HAARP facility. Figure 5.2 is a sample sounding, taken during the “off” period between the two TEC enhancements in Figure 5.1. This shows that f_oF_2 was just over 4.8 MHz at that time, as it was for the majority of the period shown in Figure 5.1.

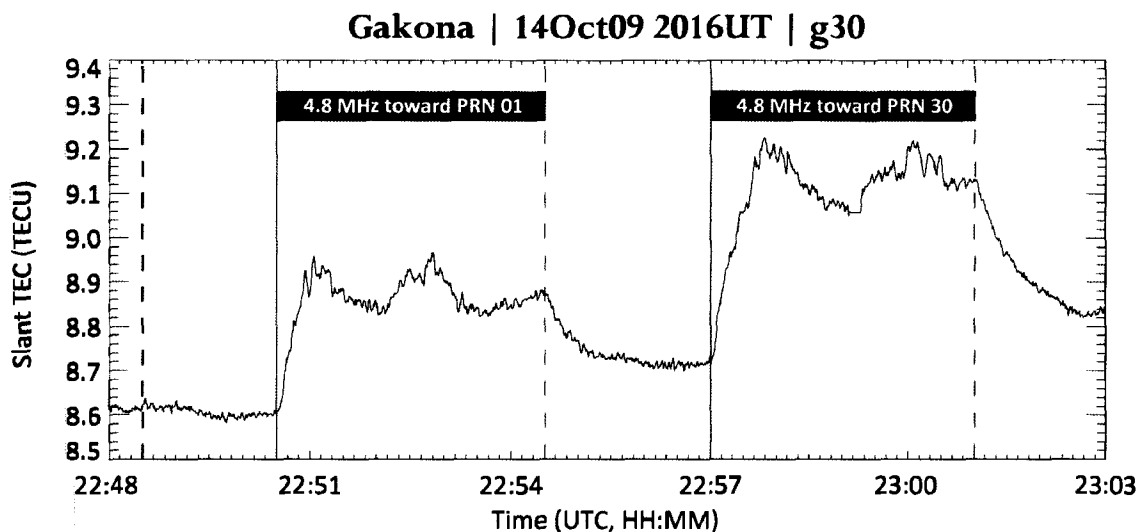


Figure 5.1 Slant TEC data derived from GPS signals from satellite PRN 30 observed from a GPS receiver at the HAARP facility during two heater “on” periods.

The beam was pointed in different directions during the two “on” periods. During the first period, it was aimed at the GPS PRN 01 satellite which was trailing PRN 30 in near-common orbits. During the second period, the beam was aimed at PRN 30. Figure 5.3 shows the geometry for this experiment. The black, red, and blue line indicates the 220-km Ionospheric Penetration Point (IPP) track for the ray path between PRN 30 and the GPS receiver at HAARP, and the blue and red ovals indicate the 3dB (inner oval) and 6dB (outer oval) contours for the IRI beams for the first (blue) and second (red) “on” periods. The wide blue and red sections of the IPP track indicate the satellite locations during the two “on” periods, with the colors corresponding to the colors used in the beam-contour lines. The red star near Copper Center indicates the location of the geomagnetic zenith as viewed from HAARP at 220-km altitude. Note that the ray path to PRN 30 was between the 3dB and 6dB contours during the first “on” period, and completely within the 3dB contour during the second “on” period.

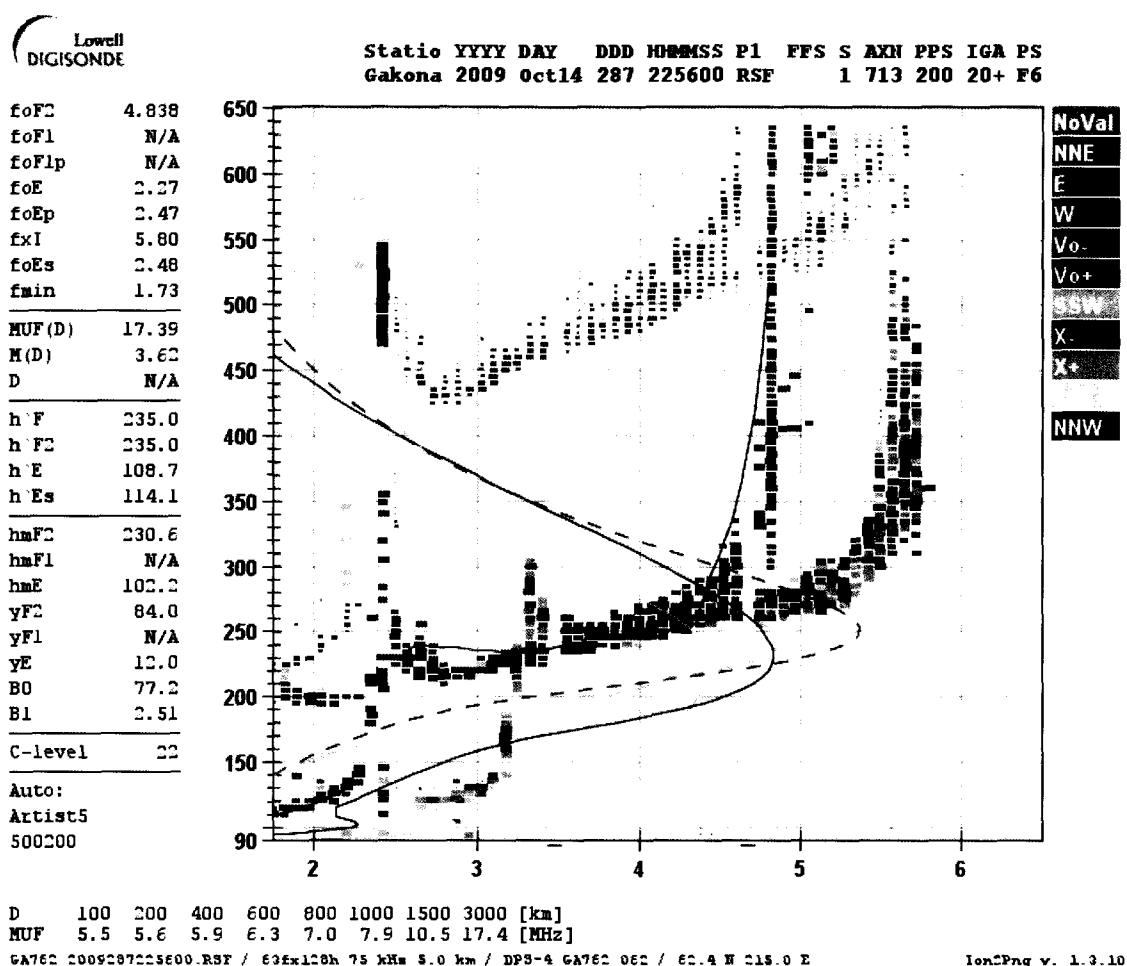


Figure 5.2 Ionogram taken with the HAARP Digisonde during the heater “off” period in the center of Figure 5.1. The solid black lines on this figure are estimates of the O-mode trace derived from the sounding and the true-height profile (with uncertainty estimates) derived from the estimated trace. This analysis was made using the University of Massachusetts at Lowell ARTIST program (Version 5). Plasma frequency profiles for 00:00 (black dashed line) and 00:05 UTC (orange dashed line) calculated in the 94 nW/m^3 simulated heating experiment described in sections 4.2.4 and 5.2 are overlaid for comparison.

The apparent “overshoot” phenomenon evident at the start of both enhancements in Figure 5.1, and the peaks seen in both enhanced segments, have been seen in other experiments of this type. It appears that these phenomena occur when the heating is very close to, but not above f_oF_2 . We have also observed that the maximum TEC enhancements, to date on order 0.5 to 0.6 TEC units, also occur when the heater frequency is just slightly below f_oF_2 .

October 2009 HAARP BRIOCHE Campaign

(IPP Track : 220km)

Station: Gakona

Satellite: PRN 30

IRI Frequency: 4.8 MHz

Date/time: 14Oct09 / 21:42 UT

(220km beam locations)

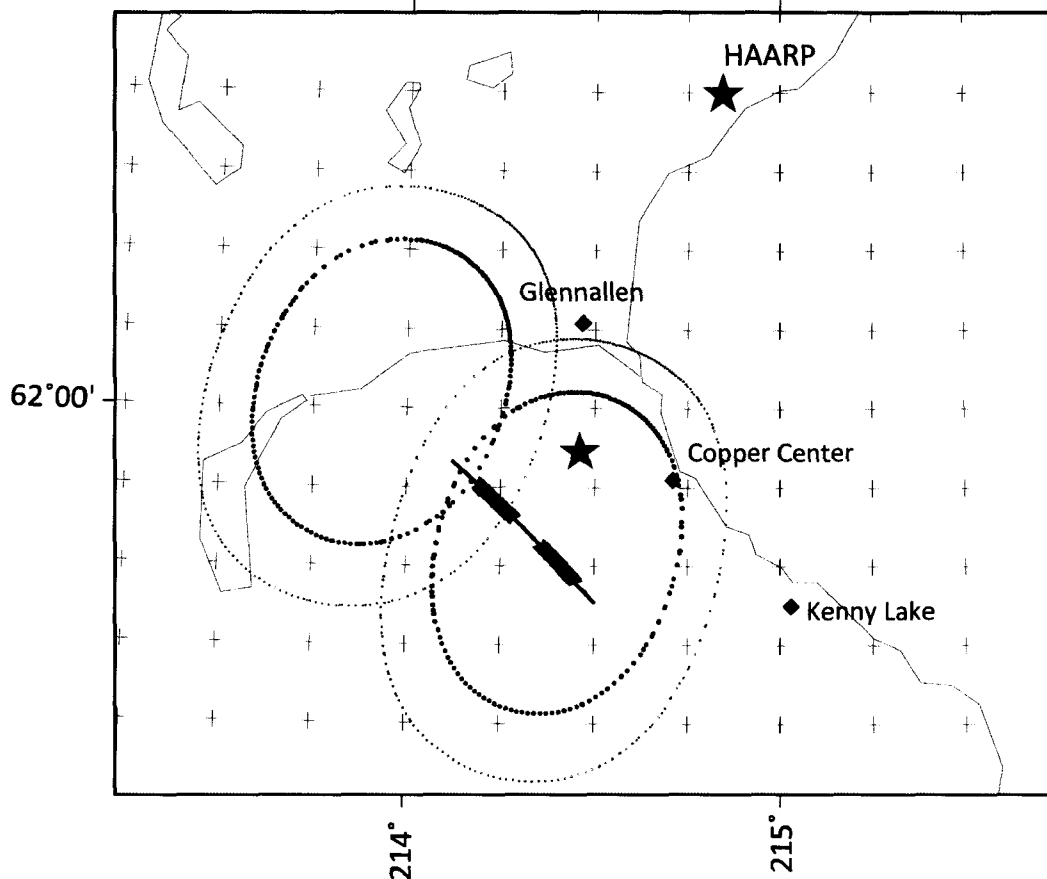


Figure 5.3 Map of the area near the HAARP facility showing the geometry of the experiment. Blue and red ovals indicate the 3dB (inner) and 6dB (outer) contours for the IRI beams for the first (blue) and second (red) "on" periods. The blue and red sections of the Ionospheric Penetration Point (IPP) track indicate the satellite PRN 30 locations corresponding to the two "on" periods.

5.2 Model

The one-dimensional self-consistent ionosphere model (SCIM) was used to simulate the plasma density and temperature response to HF heating of the F-region electron gas. The model assumes a slab symmetric ionosphere; neutral gas density and temperature are driven by the NRLMSISE-00 (MSIS) model; the input solar flux and resulting photo-production rates are driven by the Hinteregger parameterization of $F_{10.7}$ index of solar radio flux through the GLOW 0.97 model. Ion

transport equations include a coupled continuity-momentum equation—with hydrostatic approximation—for each ion species including H^+ , O^+ , O_2^+ , N^+ , N_2^+ , and NO^+ . Energy equations for the electron and mean ion gas temperature are fully coupled to the ion transport equations through temperature and density dependent chemical reaction rates, heating and cooling rates, diffusion coefficients, and thermal conduction coefficients.

The fully coupled set of nonlinear model equations is solved on a non-uniform spatial grid, over a simulation time period of 1 min, with a parabolic and elliptic partial differential equation solver based on the Skeel-Berzins method [Skeel and Berzins, 1990]. The driving geophysical parameters are then updated and the equations are solved again, using the final simulation state in the previous time step as initial conditions for the next 1 min step. To minimize effects of externally imposed initial conditions on model results, the simulation is initialized with a plasma-free neutral atmosphere and run through a complete diurnal cycle before starting the HF heating experiment. The heating simulations begin at 00:00 UTC when the HF heat source is activated by stepping up the electron volume heating rate by either 0, 23, 47, or 94 nW/m³ at the altitude grid points below the critical layer where the plasma frequency is within 4.20 ± 0.06 MHz. At 00:05 UTC, the heating rate is stepped down to its ambient value determined primarily by the model photoelectron flux. To calculate TEC as a function of time, the solution sequence of electron density profiles was integrated with respect to altitude between 100 and 1000 km.

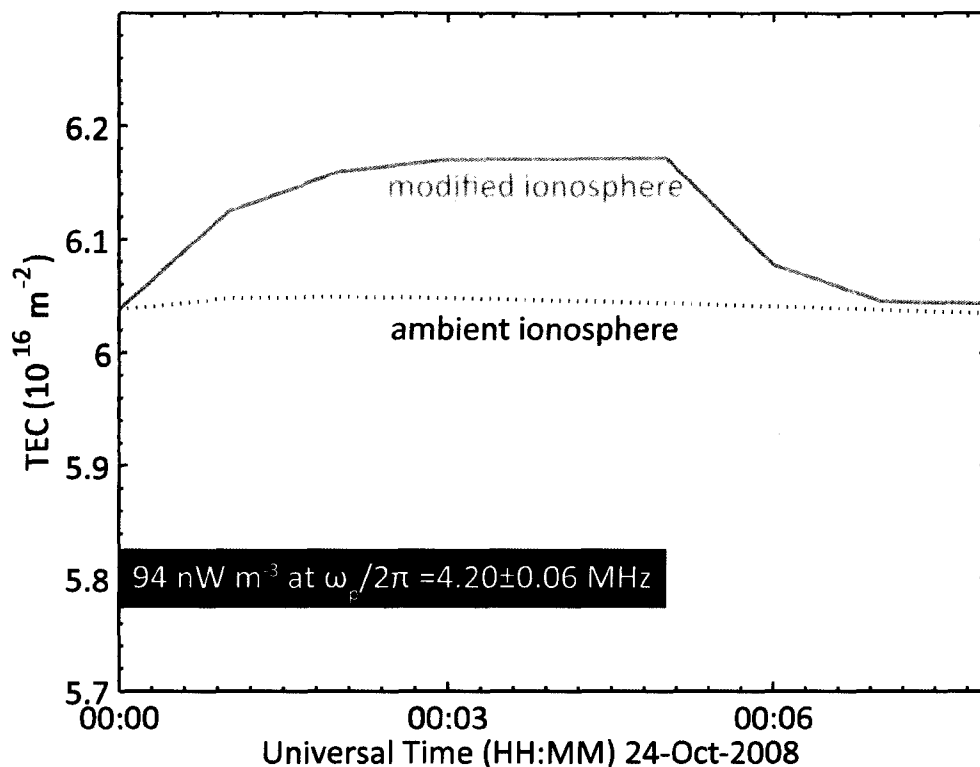


Figure 5.4 Time variation of TEC calculated below 1000km with the self consistent ionosphere model during and immediately following five minutes of simulated HF heating. The simulation time step is 1 min.

Figure 5.4 shows TEC enhancements greater than 0.1 TECu are generated and maintained within 3 min of 94 nW/m^3 of electron heating at the 4.2 MHz layer and much of but not the entire enhancement dissipates within 3 min of HF “off.” A secondary TEC enhancement beginning about 5 min after HF “off” and persisting for at least 60 min can be seen in Figure 5.5. The total TEC enhancement is the sum of a transient density enhancement in the lower F region, a depletion of the plasma density peak, and a persistent density enhancement in the topside ionosphere. This can be seen in the cumulative TEC profiles $c\text{TEC}(z,t) = \int_{100 \text{ km}}^z n_e(z',t) dz'$ in Figure 5.6a, the self-consistent model profiles in Figure 5.2, and the altitude-time contours of density perturbation in Figure 5.7.

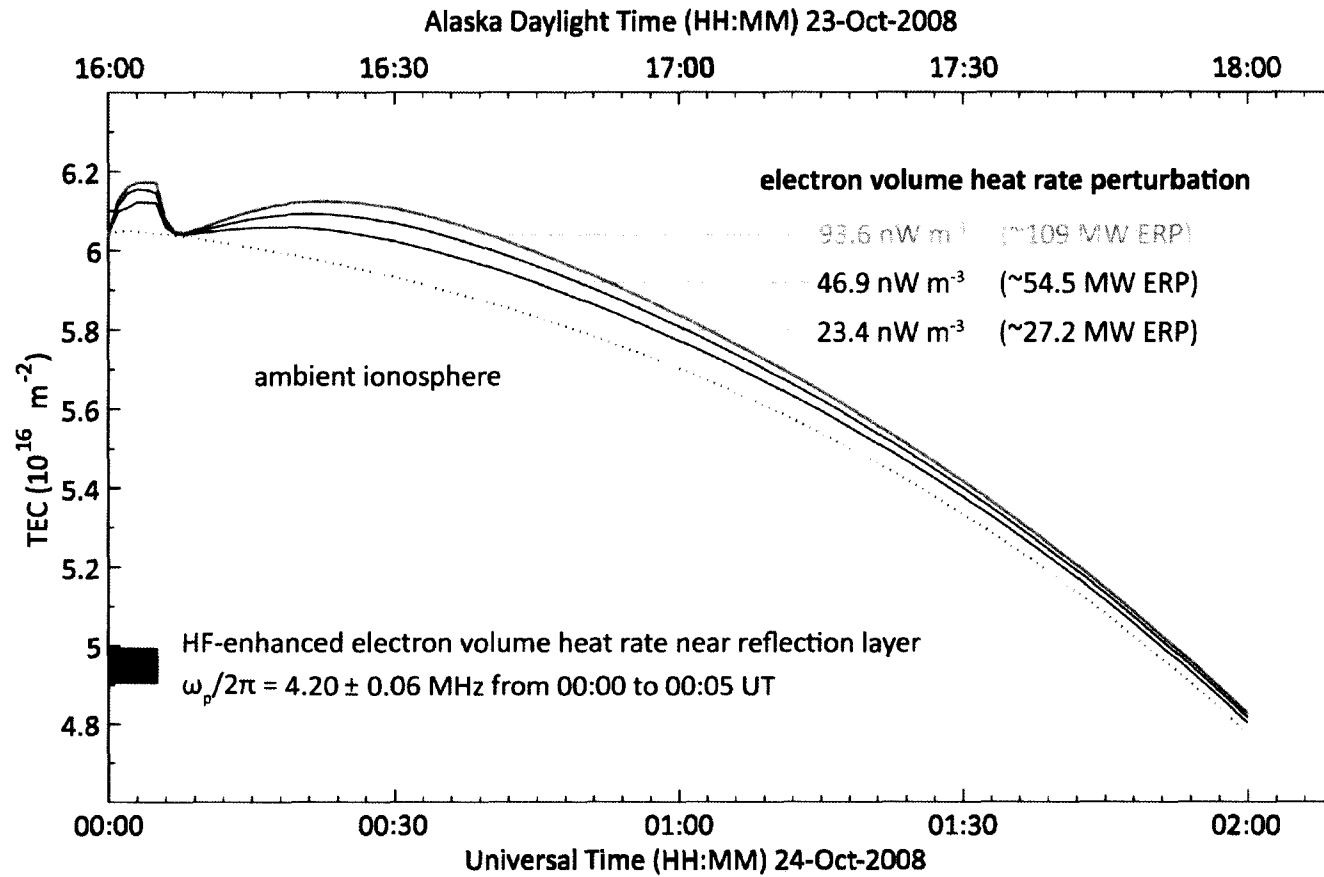


Figure 5.5 Temporal evolution of TEC calculated below 1000 km altitude during simulated HF heating experiments. Contours indicate sensitivity of the TEC perturbation to variations in the HF-to-electron power transfer density during five minutes of heating starting at 00:00 UTC.

Ion and electron temperature profiles in Figure 5.6b show that within 1 min following HF “on”, the calculated electron temperature reaches 3900 K in the HF interaction region, nearly 3000 K greater than the ambient electron temperature. Heat then conducts along the geomagnetic field resulting in widespread electron temperature enhancements exceeding 1000 K above 150 km altitude. Below the density peak, the $O^+ + N_2 \rightarrow NO^+ + N$ and $O^+ + O_2 \rightarrow O_2^+ + O$ reaction rates are not directly affected by an increase of T_e while the rates of recombination via the reactions $O_2^+ + e \rightarrow O + O$ and $NO^+ + e \rightarrow N + O$ decrease, resulting in a net plasma density enhancement. The ambipolar diffusion coefficient, governing the vertical electron density gradient above the density peak, is proportional to the plasma temperature so the topside density gradient decreases with increasing temperature. Atomic oxygen ions from the peak layer are lifted upwards by diffusion where the chemical loss rates decrease exponentially with height, resulting in an overall TEC enhancement.

The temperature profiles in Figure 5.6b also show that within 1 min following “HF off” the electron temperature near the HF interaction region relaxes to the ambient temperature, and the molecular O_2^+ and NO^+ ions below the critical layer return to their unperturbed equilibrium concentrations. The topside O^+ ions, which have been heated through elastic collisions with the hot electrons, remain 100 K warmer than the ambient ion temperature for 20 min due to the relatively low ion-neutral collision frequency. Thus, the elevated plasma temperature, now due to warm ions rather than hot electrons, continues to lift plasma from the critical layer as it is replenished by photoproduction with a characteristic time of about 10 min.

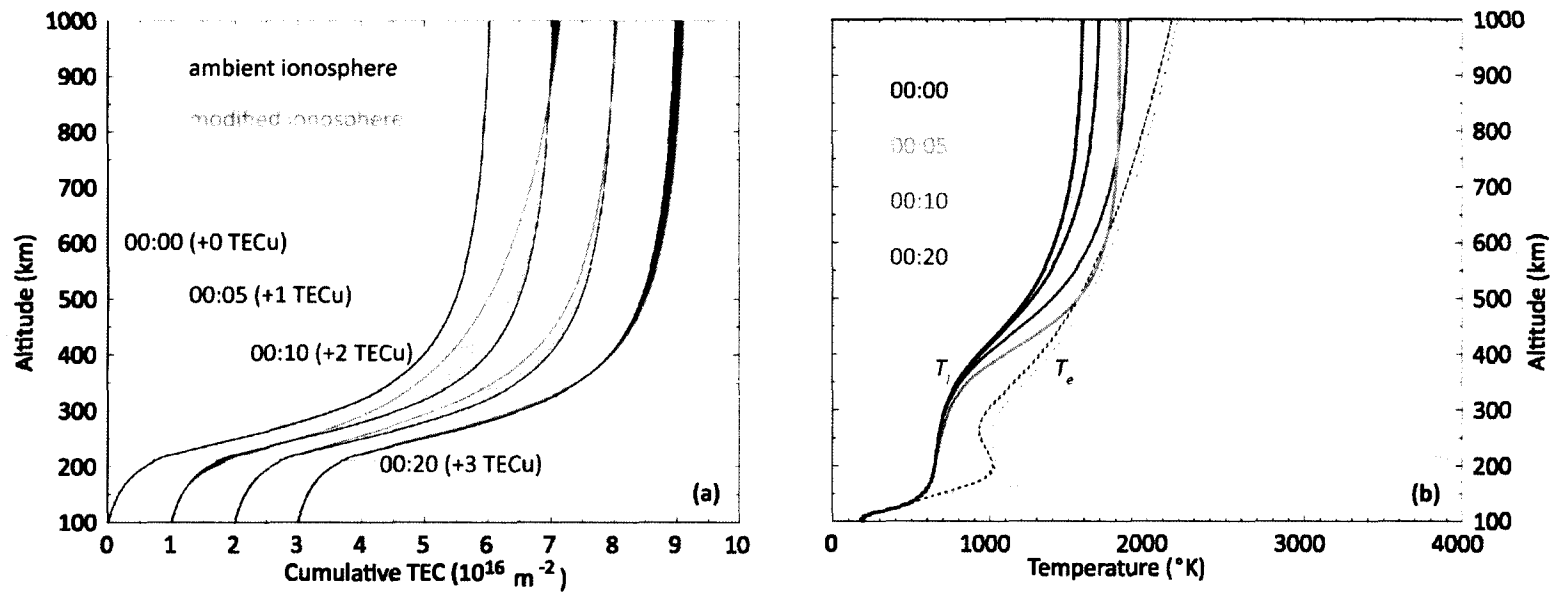


Figure 5.6 (a) Cumulative TEC calculated at selected times during a simulated HF heating experiment. (b) Differences between the ambient and modified profiles are emphasized with shading. Ion and electron temperature profiles at selected times. In each figure, the ambient electron volume heating rate was increased by 94 nW/m^3 from 00:00 to 00:05 UTC at the 4.2 MHz layer.

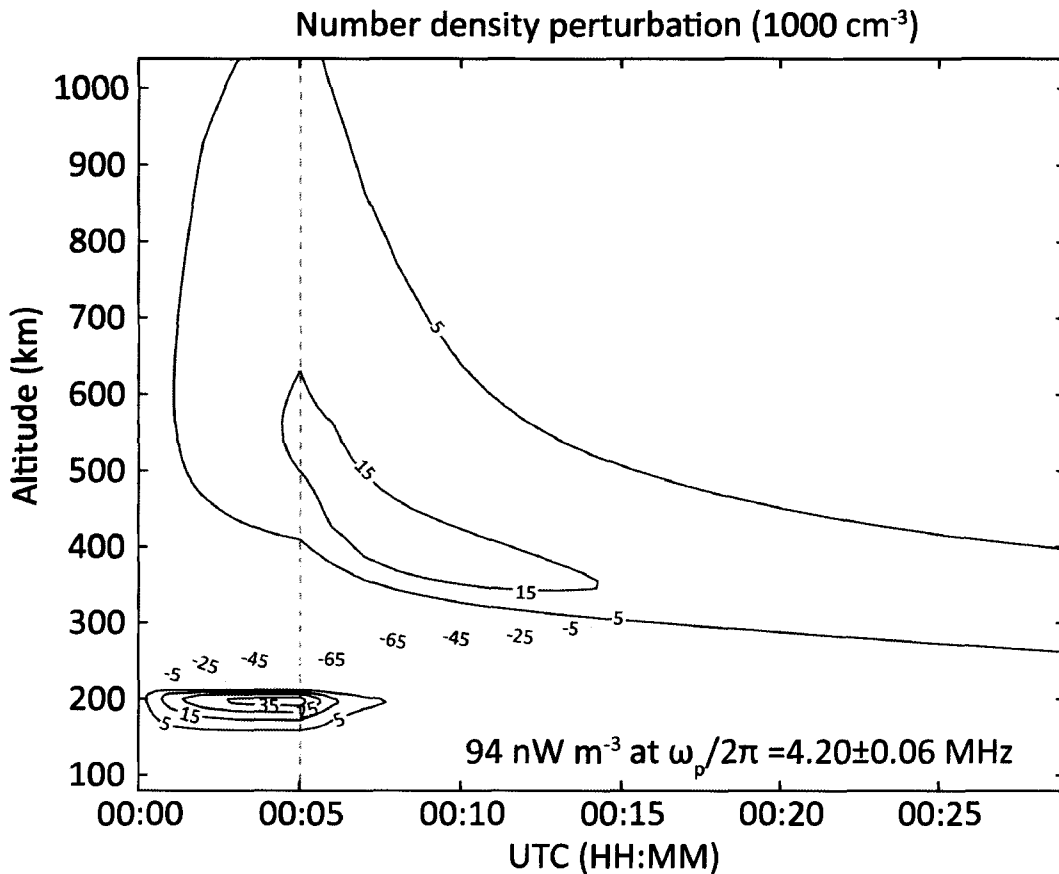


Figure 5.7 Time and altitude variation of the electron number density perturbation from ambient values. Purple contours indicate regions of density enhancement with intervals of $10,000 \text{ cm}^{-3}$ and yellow contours indicate regions of density depletion with intervals of $20,000 \text{ cm}^{-3}$.

In summary, short-time transient TEC enhancements induced by HF heating result from enhanced O_2^+ and NO^+ concentrations below the density peak; and long-time persistent TEC enhancements result from enhanced O^+ ion concentrations in the topside ionosphere. If the persistent enhancement apparent in the TEC measurements illustrated in Figure 5.1 is a result of HF-induced ionosphere modifications, then electron density enhancements in E and lower F regions are neither large enough nor persistent enough to account for the observed effect.

5.3 Discussion on HF-induced topside ionosphere perturbations

The slow-time scale TEC enhancements discussed in section 5.2 above were conjectured from SCIM calculations to result from elevated electron temperatures extending hundreds of kilometers along

the geomagnetic field. The heated plasma subsequently lifts atomic oxygen ions via ambipolar diffusion from the critical layer to the topside ionosphere where the influence of the neutral gases by way of chemical reactions and momentum-transferring collisions decreases exponentially to have negligible influence over time scales of minutes. However it is not clear that field-aligned transport of plasma thermal energy or ion mass over hundreds of kilometers, inferred from a one-dimensional slab-symmetric fluid model of the ionosphere, necessarily applies to HF heating experiments in the natural ionosphere. The natural ionosphere is influenced by neutral winds, and by horizontal plasma convection driven by perpendicular electric fields of magnetospheric origin. Moreover, the waning influence of the neutral gas with height means that a collisional description of the ionosphere is less applicable far above the F-region peak than below, thus weakening the usefulness of the Maxwellian plasma assumption and strengthening the need to incorporate temperature anisotropic effects [Blelly and Schunk, 1993].

ISR measurements made during heating experiments at Tromsø provide evidence of field-aligned electron temperature perturbations extending hundreds of kilometers co-occurring with upward field-aligned transport of topside plasma during O-mode heating experiments [Rietveld *et al.*, 2003]. In section 5.3.1 below we provide in-situ evidence of topside ion density enhancements and upwards ion flux induced by O-mode HF heating over HAARP. Section 5.3.2 then describes the modeled dependence of HF-induced topside ionosphere perturbations on the altitude of HF-interaction altitude.

5.3.1 *In-situ measurements of topside density enhancements*

An O-mode heating experiment run during the February-March 2008 HAARP campaign made use of the in-situ instrumentation on the Defense Meteorological Satellite Program (DMSP) F15 satellite. The measurements show ion density and cross-track ion velocities along the orbit a DMSP F15 satellite pass at a nominal altitude of 840 km near the HAARP facility around 0400 UT on 25 February 2008. Ion density and flux data from the in-situ instruments shown in Figure 5.8 exhibits what may well be a density structure generated either directly or as a by-product of HAARP IRI heating in the F region near 220 km altitude. Figure 5.9 shows the geometry of this pass and of the IRI heated region as configured during the time of the pass. The left panel shows the track of the DMSP pass as mapped down field lines (using IGRF 2005) to an altitude of 220 km, the nominal F2-layer altitude at the time, and the 3dB and 6dB contours of the IRI beam at that same altitude. The heavy line indicates the location of the structure apparent in Figure 5.8. The

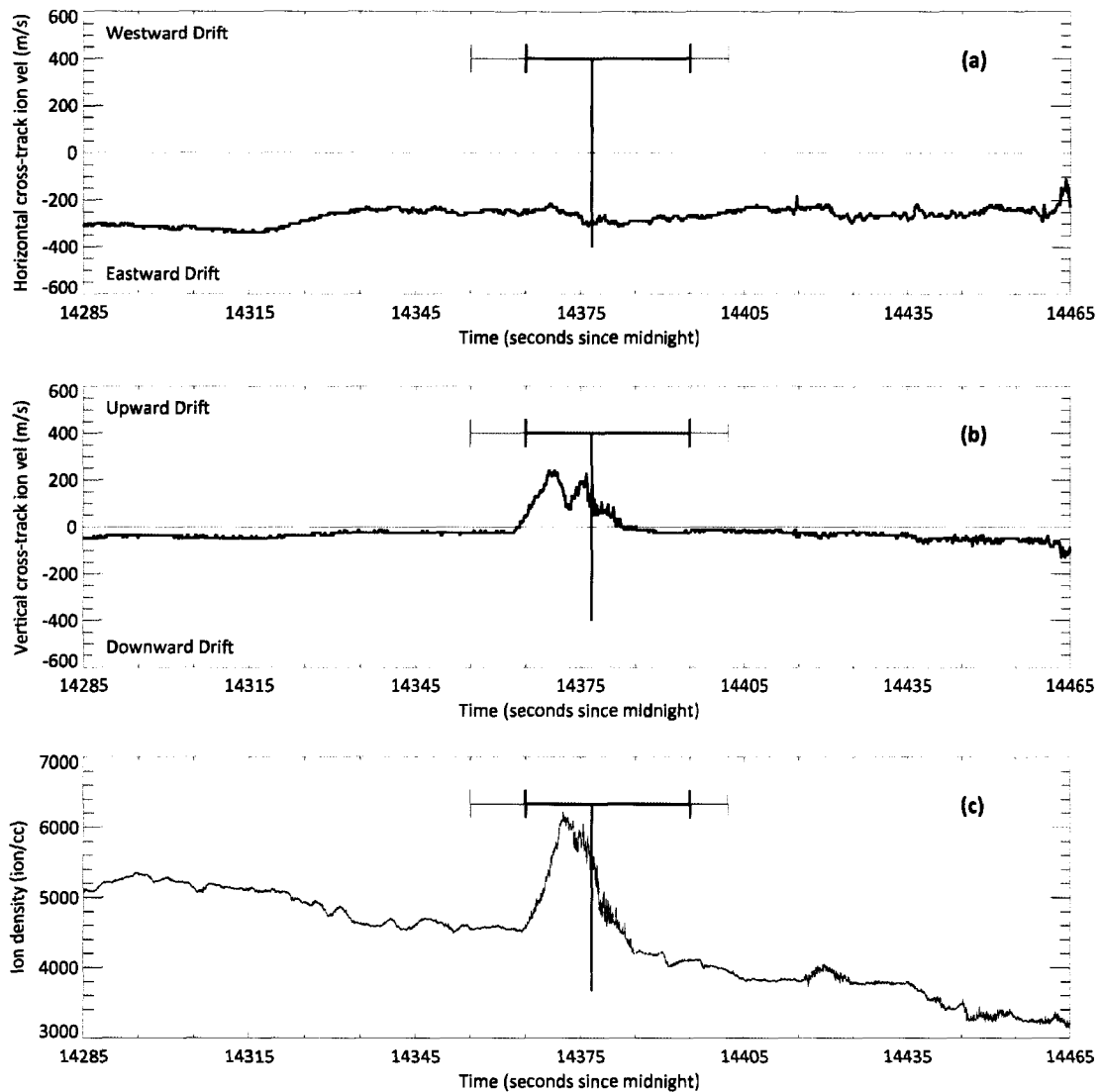


Figure 5.8 DMSP SSIES data from F15 (25 February 2008) showing (a) horizontal and (b) vertical cross-track velocities from the DMSP Drift Meter instrument; (c) total ion density data from the DMSP Scintillation Meter instrument. The vertical line indicates the nearest approach to HAARP. The horizontal lines indicate times when the satellite was within the 3dB (heavy lines) and 6dB (light lines) contours of the IRI beam at 840km altitude.

right panel in Figure 5.9 shows the satellite track at the nominal satellite altitude of 840 km and the IRI beam contours at that altitude.

Note that there is no auroral precipitation at the time of the experiment. The HAARP magnetometer and riometer traces for 25 February, along with ionosonde profiles at the time of the

experiment, all indicate relatively quiet geomagnetic conditions with low HF absorption in the D and E regions. The structure of interest is in the center of the ion density and vertical flux plots in Figure 5.8. This is an isolated structure in the pass data, and while it does not map directly into the F-layer region in which the IRI is interacting, it is very near to it. One tentative hypothesis, consistent with model calculations described in sections 4.2.4 above and 5.3.2 below, is that the larger-scale structure is plasma that has been moved along the field line from lower altitudes, with the finer-scale structure evident in the top panel of Figure 5.8 more probably generated locally due to plasma instabilities, possibly a gradient-drift instability. The fact that one side of the larger structure has this finer-scale structure than the other is symptomatic of this type of instability in isolated density structures. The coincident onset of upwelling evident in the vertical drift velocity (bottom panel of Figure 5.8) is consistent with this picture.

A critical question that must be answered in order to confirm that this structure is: How was the structure formed by IRI interactions? Options include (1) direct interaction between the IRI beam and the ionosphere either at or very near the satellite altitude, (2) bulk transport of plasma from below the satellite altitude along the field line by either changing the thermal equilibrium along the field line or by changing the electric field structure along the field line, or (3) some form of interference with the DMSP instruments by direct interaction with the IRI RF signal. Option 3 is unlikely but should be verified with engineers familiar with the DMSP instrumentation. The fact that the horizontal cross-track velocity shows no feature where the vertical velocity shows a sharp change in the flow characteristics, and the uneven distribution of the effect on all parameters in the region where the satellite is within the IRI beam at 840km altitude, are strong circumstantial evidence that this is unlikely to be interference. Similarly, option 1 is unlikely, unless a plausible mechanism for the interaction can be proposed given the IRI operating mode and frequency. One observation against this option is that there was no change in the electron temperature as measured by the Langmuir probe on DMSP (data not shown). One of the scenarios in option 2 is the most likely, although whether the plasma transport is due to thermal or electro-dynamical processes is not clear at present. But model calculations described above imply that thermal processes are sufficient to account for the observed density perturbation. Model calculations of the ion vertical flux perturbation, and its dependence on HF-interaction altitude is discussed in section 5.3.2.

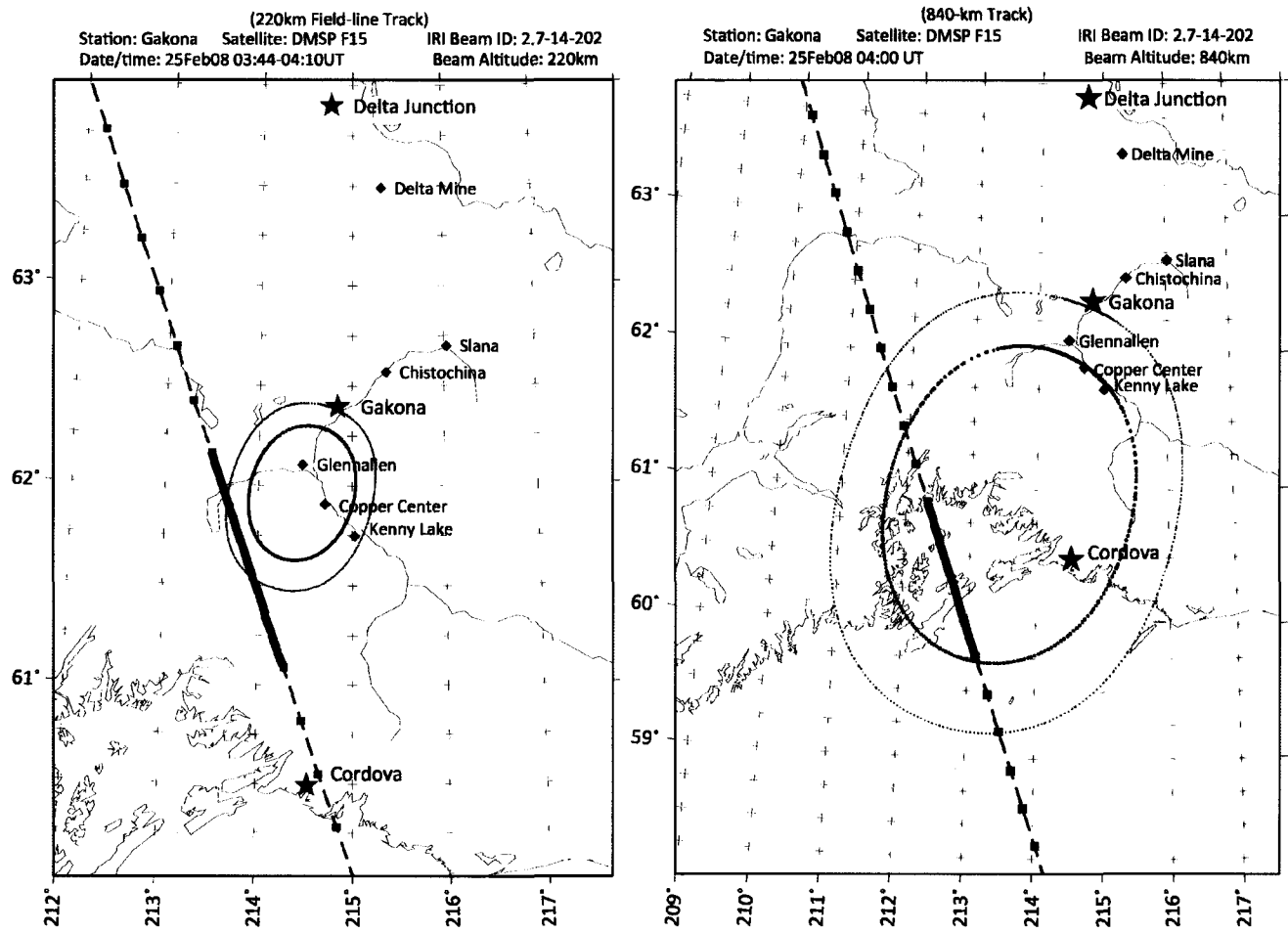


Figure 5.9 DMSP F15 orbit track and IRI heater beam contours (3dB and 6dB). Left panel is at 200km altitude (DMSP locations traced down field lines); right panel is at 840km altitude. Heavy line along DMSP track indicated region of enhanced ion density. Tic marks along the track are at 5 second intervals.

If the plasma in the feature has been transported from lower in the ionosphere without any loss, if we assume a scale height of about 150 km and a Chapman profile the density difference between the center of this feature and the density away from the feature would correspond with densities roughly 50 km below the satellite altitude when outside the feature. With an upward velocity of 200 m/s, that infers a transit time on the order of four minutes, which is well within the time frame of the experiment and the model density profiles in Figure 4.11.

5.3.2 HF frequency dependence of ionosphere modification

The self-consistent ionosphere model (SCIM) is a multi-species ion model of the one-dimensional ionosphere used in section 4.2.4 to simulate heating of the ionospheric electron gas at a *fixed plasma frequency*. SCIM is used in this section to simulate the ionosphere temperature, composition, and vertical mean velocity response to electron heating at a *fixed altitude*. The magnitude and altitude-extent of the resulting temperature, density, and vertical ion velocity perturbations are found to depend strongly on the altitude of enhanced electron heating. Electron thermal conductivity increases with electron temperature at increasing altitude, and efficient cooling through collisions with the neutrals decreases with increasing altitude. Consequently, long-lasting topside ionosphere perturbations are much more pronounced if heating can be maintained at or just below foF₂ (which was shown in section 4.3 and 5.2 to drop by as much as 500 kHz during heating).

Recall that one-dimensional continuity and hydrostatic momentum equations are used to describe the number density $n_i(z, t)$ of each ion species i :

$$i \in \{H^+, O^+(^4S), O^+(^2D), O^+(^2P), O_2^+, O_2^+(a^4\Pi), N^+, N_2^+, NO^+\}, \quad (5.1)$$

as a function of altitude and time. Energy equations are used to describe the electron T_e and ion T_i temperatures. Temperature- and density-dependent sinks and sources of energy and mass are included in the model equations. The method-of-lines [Skeel and Berzins, 1990] is used to solve the fully coupled set of continuity, momentum, and energy equations with parameterized boundary conditions and HF volume heating rates. Electron temperature, ion temperature, and ion densities are held fixed at the lower simulation boundary at 80 km. A downward electron heat flux of 0.01 mW/m², zero ion heat flux, and zero ion mass flux are fixed at the upper simulation boundary at 1100 km. A perpendicular electric field of 10 mV/m was applied at all altitudes to simulate Joule-

heating of the ions. Major neutral atmosphere constituent densities are parameterized functions of altitude and time using the MSIS00 model; ionization and geophysical heating rates are parameterized functions of time, altitude, and constituent densities by the GLOW model; and the World Magnetic Model (WMM) is used to parameterize the geomagnetic field strength and direction. Quiet solar and geomagnetic indices $F_{10.7}$ and AP recorded at the beginning of August 2007 were used to initialize the MSIS, GLOW, and WMM models at the latitude and longitude of Gakona, Alaska.

The simulation was initialized from a zero-ionization state by running for 48 simulation hours with diurnal variation. HF heating followed initialization by adding a uniform volume heating rate $Q_{\text{HF}}(t, z) = \phi_{\text{HF}} / \delta z_{\text{HF}}$ with a 15-minute on-off time variation starting at 0215 UT over an altitude of range of $\delta z_{\text{HF}} = 2$ km, centered at the chosen interaction altitude of either $z_{\text{HF}} = 200$ km or $z_{\text{HF}} = 250$ km. At 0215 UT, the interaction altitudes correspond to an ionosphere plasma frequency of approximately 4.0 and 7.0 MHz, respectively. The simulated F-region peak layer height $h_m F_2$ at 0215 UT is about 250 km altitude. The electromagnetic energy flux density ϕ_{HF} was chosen by taking the output of the HAARP HF Transmitter Performance Calculator² at the chosen interaction altitude, assuming full array availability and a transmitter frequency of 2.8 MHz. HF-electron volume heating rates were calculated from these values by distributing the power flux evenly over a 2 km layer. Note that the effective radiated power (ERP) of the IRI at HAARP generally increases with the chosen HF frequency, and that these calculations assume 100% absorption of available HF power by the thermal electron gas.

Each model run used a spatial grid constructed by taking the union of a logarithmic spaced altitude grid with a uniform 0.5 km grid extending ± 5 km from the chosen interaction altitude. At 250 km altitude, the non-refined logarithmic grid spacing is about 7 km. The four simulation runs were computed concurrently with Matlab 7.4 [Mathworks, 2007] on one Sun Fire X2200 node of the 415-node *Midnight Sun* Microsystems Opteron cluster maintained at the Arctic Region Supercomputing Center. One advantage of the method-of-lines numerical solver used here [Skeel and Berzins, 1990] is that an explicit discretization of time is not necessary as it is with standard finite-difference techniques [Hastings and Roble, 1977]. Consequently, negligible performance penalty is incurred by

² <http://www.haarp.alaska.edu/haarp/calcul.html>

resolving fast-time processes, such as the electron temperature response to HF heating, simultaneously with slow-time processes such as the ion temperature response, as illustrated in Figure 5.10.

Rapid (<10 sec) increases of the electron temperature near the HF interaction altitude and slow increases (<100 sec) at altitudes up to the simulation boundary of 1100 km are noted. The magnitude and time variations are in approximate agreement with observations [*Gustavsson et al.*, 2005]. Corresponding increases in the ion temperature are also noted at upper F-region heights. Topside ion temperature enhancements, resulting from elastic collisions with heated electrons, depend in magnitude on the topside electron temperature enhancement. Heating the electrons at higher altitudes—up to the F-region critical layer—allows more heat to conduct to the topside, thereby heating the ions more efficiently than heating the electrons further below the critical layer.

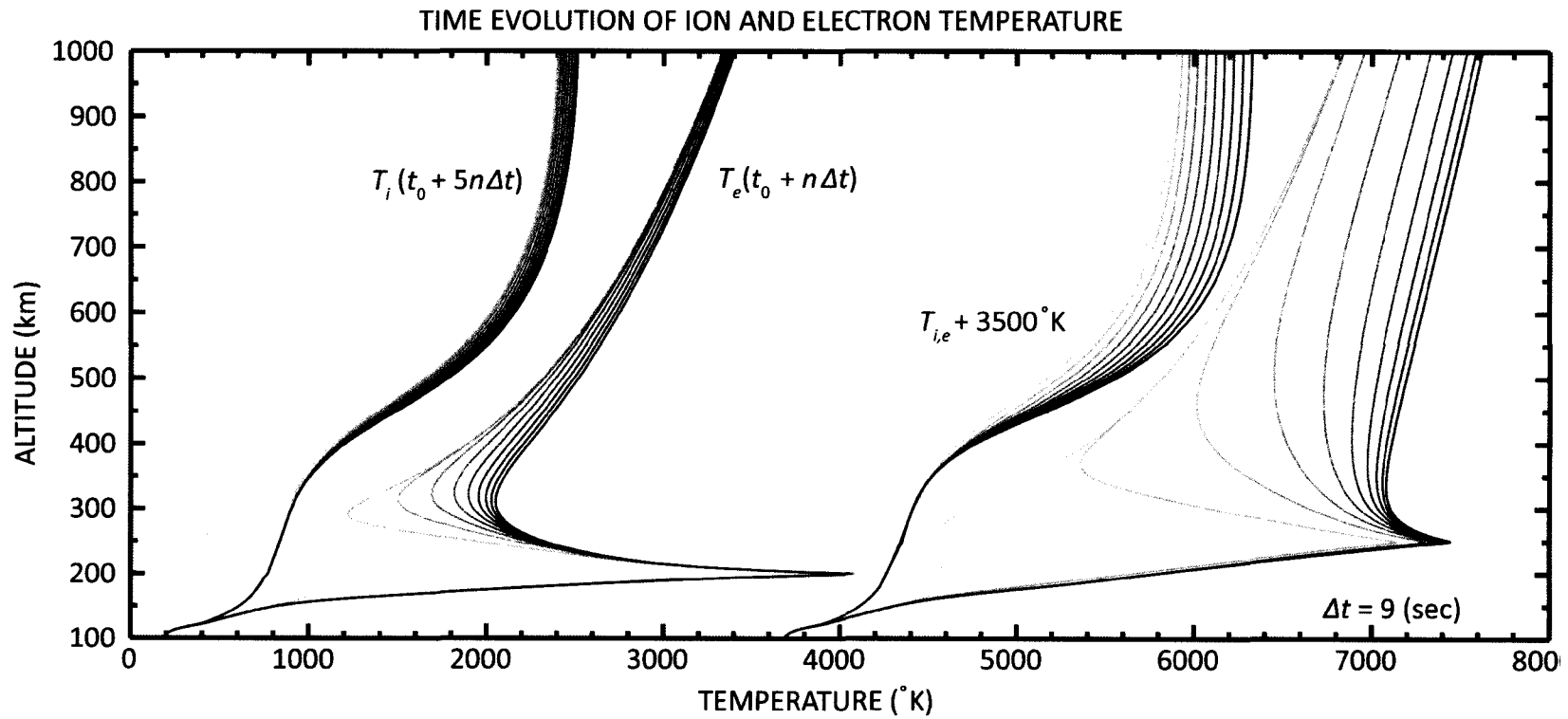


Figure 5.10 Time evolution of ion and electron temperature. Ion and electron temperature variation following HF heating at 200 km (left) and 250 km (right) altitudes. The ion temperature profiles are separated by 90 sec intervals and the electron temperature profiles are at 9 sec intervals; time increases in the direction of darker shading.

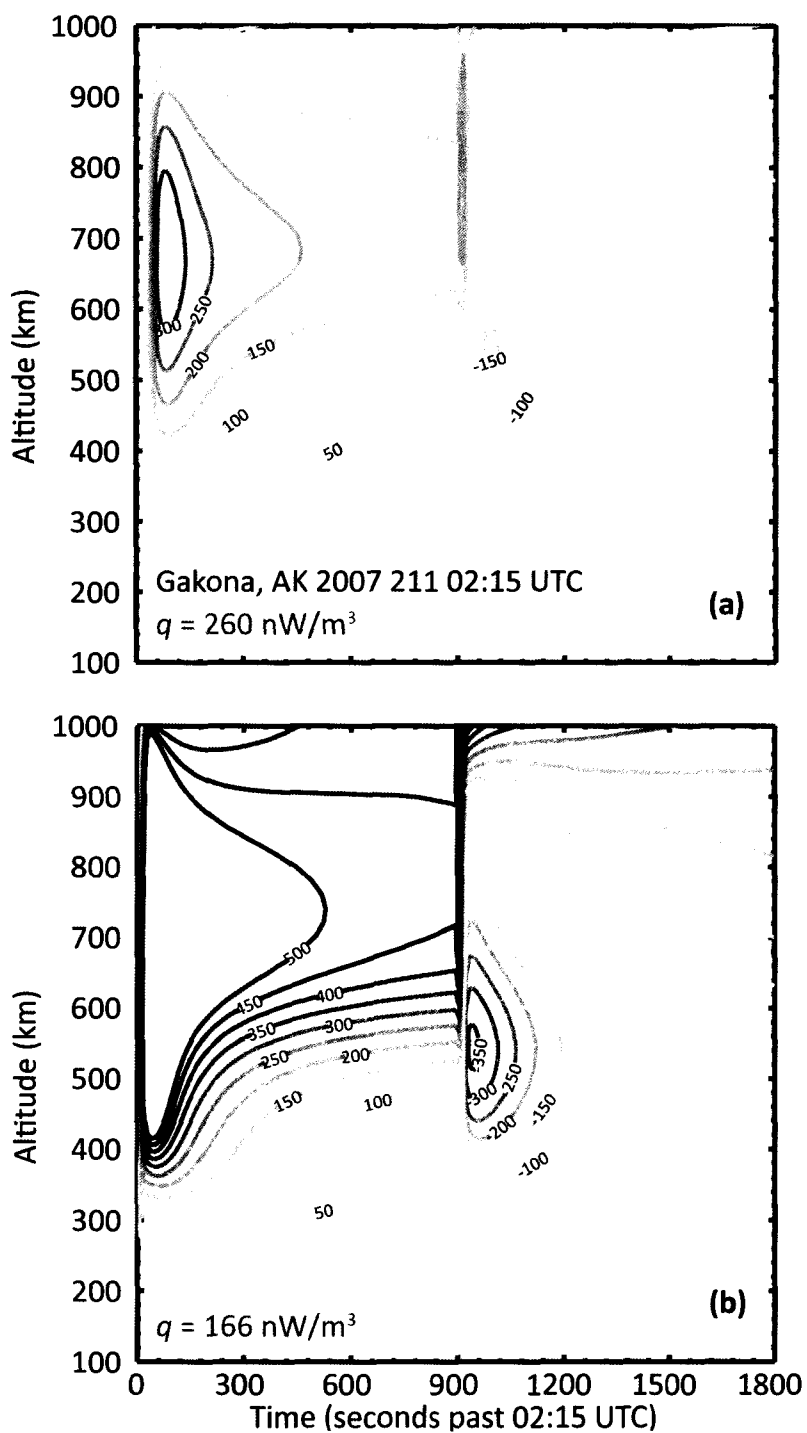


Figure 5.11 Time evolution of the major ion ambipolar diffusion velocity relative to the unperturbed velocity. Change in the vertical component of geomagnetic field-aligned O^+ ambipolar diffusion velocity during HF heating at 200 km (a) and 250 km (b) altitude. The simulated HF heater is switched on at time $t = 0$ s and switched off at $t = 900$ s.

Enhanced plasma temperatures in the topside ionosphere enhances the ambipolar diffusion velocity of the ions, shown in Figure 5.11, as the ions redistribute along the geomagnetic field line to reach equilibrium at the new plasma temperature. Mean ion diffusion velocities are enhanced by up to 300 m/sec upwards with heating at 200 km, consistent with observations [Rietveld *et al.*, 2003] and prior predictions [Mingaleva and Mingalev, 1997]. Increasing the heating altitude to 250 km, at the critical layer height, drastically increases the diffusion velocity even if the actual volume heating rate decreases with the square of the distance from the IRI.

Figure 5.12 shows that the peak F-region density decreases by about 20% in response to heating at 200 km and 35% at 250 km, while the that the upper F-region density increases by 50% during 200 km heating and approximately doubles during 250 km heating. The values at 200 km heating are in apparent agreement with reported observations [Rietveld *et al.*, 2003] and the DMSP measurements described in section 5.3.1 above. Density enhancements in the lower F region clearly result from increased concentrations of the molecular ions O_2^+ and NO^+ but the topside density enhancements consist entirely of atomic oxygen ions. Also note the depletion “notch” of about 10 km thickness in the density profile. It occurs at the HF interaction altitude, similar to ISR measurements at Arecibo during HF experiments reported by Hansen *et al.* [1992].

The difference in perturbed temperature profiles—resulting from HF-electron heating at different altitudes below the critical layer—causes significant changes in plasma density at the critical layer and in the topside ionosphere. This effect has two important consequences for HF modification experiments: (1) Heating at or just below foF2 maximizes the density perturbations surrounding the critical layer and in the topside ionosphere. This affects in-situ observations at satellite orbital altitudes and measurements like slant TEC that are sensitive to the plasma density over a wide range of altitudes. (2) Density modifications caused by heating at or just below foF2 are likely self-limiting as foF2 drops below the chosen IRI frequency due to HF-induced density depletion of the critical layer.

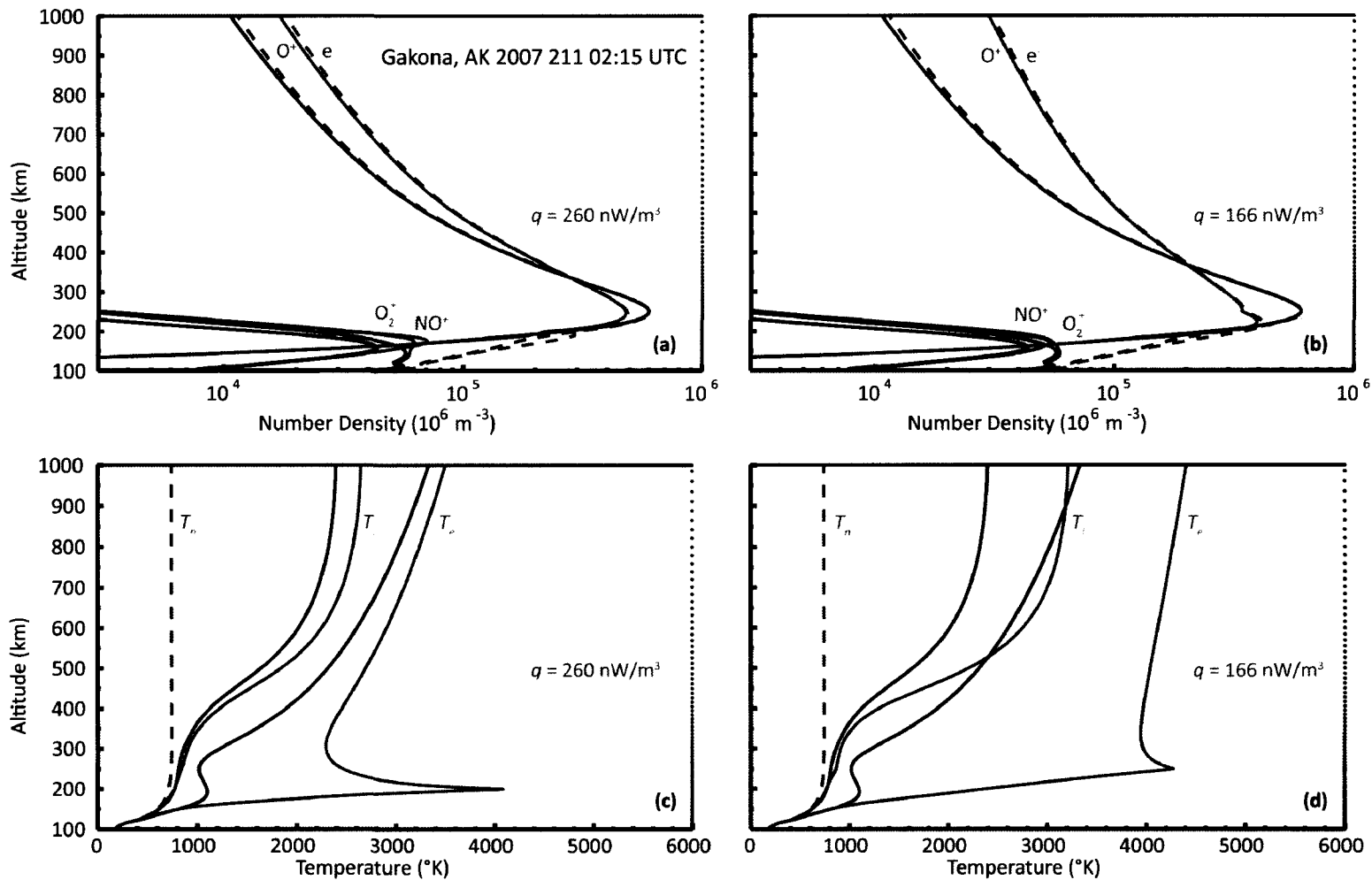


Figure 5.12 Ion and electron density and temperature profiles before (black curves) and after (red curves) 900 sec of simulated HF heating at (a, c) 200 km altitude and (b, d) 250 km altitude. The neutral gas temperature profiles in (c) and (d) are provided by the NRLMSIS00 model.

Chapter 6. Summary and conclusions

Powerful high frequency (HF) radio waves incident on the polar F-region ionosphere can couple to naturally occurring plasma waves through various instabilities, heating electrons in a thin layer near the HF-reflection height. The power transfer rate available through O-mode transmissions from the ionospheric research instrument (IRI) at the High-frequency Active Auroral Research Program (HAARP) is approaching that available from the most significant source of ionosphere heating—the Sun, in the form of hard ultraviolet electromagnetic radiation. Within tens of seconds after starting HF transmissions, thermal energy conducts from the heated electrons, thousands of degrees warmer near the HF reflection height, along the geomagnetic field: down tens of kilometers to the lower F region and up hundreds of kilometers to the topside ionosphere, raising the surrounding electron temperature by hundreds of degrees. The subsequent ion response, coupled to the electrons through the electrostatic force, occurs on altitude-dependent time scales of minutes to tens of minutes. In the lower F region where ion transport is negligible, artificially enhanced electron temperatures will lead to a net plasma density enhancement as the molecular ion recombination rates decrease with increasing electron temperature. In the 100 km layer surrounding the F-region peak, ion density decreases as atomic oxygen ions are transported along the field line through temperature-enhanced ambipolar diffusion. As the altitude gradient of plasma density declines, the topside plasma density increases. Field-aligned HF-induced plasma density modifications during tens or hundreds of seconds of O-mode heating are manifest in radar measurements of the reflection height, global positioning system (GPS) measurements of slant total electron content (TEC), and in-situ scintillation measurements of ion density. Perturbations observed in reflection height or TEC measurements have characteristic times that may be used to identify the responsible ionosphere layer, and hence, the relevant governing physical processes.

Aeronomy—the physics of the upper atmosphere—is complex, with many competing forces and sources or sinks of energy that wax or wane in relative influence according to altitude. Most of these quantities are known only through empirical curves, and fit to difficult laboratory or in-situ measurements. Equations relating the forces and energies are numerous, coupled, and impossible to solve analytically so either significant simplifications must be made ad-hoc, or solved numerically with a computer. Numerical routines capable of solving the set of equations accurately and using a reasonable amount of programmer or computer time are themselves the product of an active area of research and are still evolving. Computer models of the ionosphere help to clarify

relevant physical processes and make predictions that otherwise would be difficult or expensive, if possible at all, with traditional analytic calculations.

This thesis describes a fluid model of the high-latitude ionosphere that reproduces the well-known vertical structure of daytime ionosphere plasma composition and temperature during quiet geophysical conditions. The diurnal variation of the self-consistent ionosphere model (SCIM) critical frequency f_oF_2 agrees with the variation measured by an ionosonde at Gakona, Alaska. SCIM solves the set of hydrostatic ion momentum-continuity equations, coupled with average ion and electron gas energy equations, with a numerical solver based on the “method of lines.” The method of lines was originally discussed in an aeronomy context by *Hastings and Roble* [1977] as possible method for solving a system of equations determining the vertical structure of the ionosphere that allows the programmer to express the numerical equations in a form close to the mathematical form presented in the literature, reducing program complexity and the potential for mistakes. Computational resources at that time were not sufficient to implement the technique in a detailed ionosphere model, however, and this thesis is the first reported use of the Skeel-Berzins method of lines [*Skeel and Berzins*, 1990] to solve the continuity, energy, and momentum equations describing the vertical ionosphere structure.

The Ionosphere Research Instrument (IRI) and Modular UHF Ionosphere Radar (MUIR) at the High Frequency Active Auroral Research Program (HAARP) facility in Gakona, Alaska provides a means to do active experiments, rather than passive observational experiments, with the polar ionosphere. Transmitting powerful HF electromagnetic waves into the ionosphere with the IRI both heats the electron gas by way of enhancing electrostatic plasma waves along the HF ray path at a layer near where the ionosphere plasma frequency equals the transmitted HF frequency. Heat from HF-induced electron temperature perturbations quickly conducts along the geomagnetic field, causing plasma temperature and density perturbations extending hundreds of kilometers from the HF-interaction region. Additionally, the HF-enhanced plasma waves are detectable with UHF radar, providing a diagnostic for the resulting changes in the ionosphere profile.

A coded long-pulse technique was developed for this thesis, based on the technique described by *Sulzer* [1986], to locate the HF-enhanced ion line (HFIL) with high spatiotemporal resolution using MUIR. The HFIL—a signature of HF-driven ion-acoustic waves—occurs at the so-called matching-height, a function of electron temperature, electron density, HF-heater frequency, and UHF (or VHF) radar frequency. As the electron temperature and density vary above HAARP, by way of

artificial or natural changes in the ionosphere, the matching height will vary. This variation in the matching height provides a means to compare observations made with MUIR. An experiment technique was developed for this thesis that uses a sequence short HF pulses from the IRI to probe the ionosphere without significantly heating, alternated with long continuous IRI transmissions designed to heat the ionosphere and subsequently modify the field-aligned density profile, so that the natural ionosphere response could be compared with SCIM-calculated ionosphere response.

6.1 Conclusions

This thesis describes a time-dependent polar ionosphere model that can reproduce the daytime diurnal variation of the ionosphere, as measured by ionosondes or incoherent scatter radars, and this model was used to simulate the ionosphere response to high-power radio modification experiments. The response of the vertical ionosphere structure to HF-induced electron heating, as defined by changes in the altitude profiles of plasma composition and temperatures, depends on the altitude observed as well as the altitude of electron heating. Typical number density perturbations calculated in the simulations at affected altitudes are approximately 20-50% of the ambient values. Main findings from model calculations of the ionosphere response to heating, with supporting measurements, are as follows:

1. The electron temperature immediately surrounding the HF-heating layer reaches a sharply-peaked steady state less than a minute after the start of heating. The power transfer into the electron gas, in the form of heat, is thermally conducted up and down the geomagnetic field away from the HF-interaction region, and is also transferred into the ion and neutral gases through elastic and inelastic collisions with the heated electrons. The 2 km thin-layer heating assumption used in the model is supported by observations of the HFIL signals that are about 2 km thick on the range-time diagrams.
2. Electron number-density enhancements in the lower F-region ionosphere, typically extending at least 50 km below the HF-heating layer, also reaches a steady state within about one minute after heating begins. The plasma-density enhancement, composed of the molecular ion species O_2^+ and NO^+ , forms in response to the decrease in chemical loss rates of those ions with the increase of electron temperature. The number density of molecular ion species in the lower F-region is largely determined by the balance of its chemical loss rate (the product of the species number density and its chemical loss

frequency) to its production rate. Since the production rate is nearly independent of electron temperature, the molecular ion density must increase to balance the decrease in the loss frequency. Given the large collision frequency in the lower ionosphere relative to the upper ionosphere, both the build-up of ionization during heating and the subsequent recovery following heating are rapid and have a characteristic time of about one minute. The lower F-region plasma density and electron temperature enhancements are apparent in radar measurements as a descent of the HFIL reflection height during continuous HF heating. TEC measurements also show buildup and decay of ionization with a characteristic time of about one minute.

3. Depletion of plasma density surrounding the F-region peak density layer during HF-heating results from enhanced upward ambipolar diffusion of atomic oxygen ions along the geomagnetic field. Increasing the ambipolar diffusion coefficient, proportional to the plasma temperature (the mean of the ion and electron temperatures), decreases the vertical number density gradient of the ion species by redistributing plasma from the F-region peak layer to the topside ionosphere. Recovery of the peak density layer occurs on a time scale of up to tens of minutes, and is limited by the photoproduction rate. This depletion is observed with UHF radar as an ascent of the HFIL during low duty-cycle pulsed HF transmissions: the lower F-region density enhancement decays, leaving the HF transmissions to reflect from higher altitudes near the now-depleted F region peak.
4. Plasma density enhancements in the topside ionosphere are composed of atomic oxygen ions, lifted from the F-region peak layer during first minutes of heating, and persist for several minutes following the end of heating. The topside density enhancement is maintained in part by the significantly smaller O^+ chemical loss rates found at high altitudes, a consequence of the exponential decrease of molecular neutral gas density with increasing height. Model predictions were verified by measurements of enhanced ion density and upward velocity made in-situ at 840 km altitude during a DMSP satellite flyby of the HAARP field-aligned interaction region.
5. TEC enhancements, i.e. changes in electron column density calculated by taking the (Reimman-Stieltjes) integral of plasma density along a vertical or slanted ray, are generated by HF heating. Modeled TEC enhancements of about a tenth of a TEC unit buildup during

two minutes of heating and are the net result of the altitude-dependent density irregularities catalogued above. Measurements of slant-TEC enhancements of several tenths of a TEC unit taken along the line-of-sight that intersected the HF interaction region between the HAARP facility and a GPS satellite confirm model calculations.

Finally, one conclusion with practical relevance to planners of ionosphere modification experiments is supported by both the model calculations and radar observations presented in this thesis. Reports of O-mode HF-modification experiments sometimes describe results collected over several cycles of a typical five-minute 'on' five-minute 'off' HF sequence, while explicitly noting that the five-minute 'off' period allows the ionosphere to recover completely, thus eliminating ionosphere preconditioning from affecting subsequent cycles. The ascending HF-enhanced ion-line observed with MUIR shows that five minutes 'off' is not likely to be a sufficient recovery time, and follows from the HF-induced depletion of the F_2 -region peak layer plasma. In the absence of strong horizontal convection, this depletion above the heating facility persists for at least several minutes, and possibly up to several tens of minutes. The recovery time of the F_2 region is limited by the photoproduction rate, so experiments near dawn or dusk may be particularly affected by preconditioning.

6.2 Future work

Results of this thesis may be expanded toward three main fronts: model, observation, and computational.

Model calculation of airglow volume emission rates, based on the collision frequency between thermal electrons and, for example, atomic oxygen atoms or atomic nitrogen ions, could be compared with measurements of HF-induced airglow emissions. Heating rates and electron temperatures during radio modification experiments could be estimated by adjusting the model heating rate so that both the airglow emissions and HFIL decent agree with observations. Additionally, several improvements could be made to the model that would widen the range of physical conditions it could simulate. For example: include neutral wind effects in the continuity-momentum equation, include advection heat flows in the energy equations, and eliminate the top boundary of the model extending the simulation domain along the geomagnetic field line to the conjugate point.

The pulse-to-pulse steering capability of MUIR allows new observations of the two- and three-dimensional structure of HFIL signals. Time-evolving volume-images of HFIL structures then could be compared with volume-images of HF-enhanced airglow, constructed from multiple measurement directions. The aspect-angle dependence of HFIL reflection height variations provides another means of validating model calculations, since the transmitted HF energy flux (and hence the electron volume heating rate) also depend on the aspect angle. New TEC measurements using GPS receivers stationed nearby the HAARP facility such that the line-of-sight (LOS) slant columns of total electron content intersect different regions of the modified ionosphere is another method to explore its spatiotemporal evolution. For example, model calculations discussed earlier suggest that a decrease in TEC will be measured along a LOS slant-TEC intersecting the just the F₂ peak layer over HAARP.

Model calculations progress in approximately real-time on commodity workstations available in 2010. That is, 24 hours of computer time are needed to calculate 24 hours of simulated diurnal ionosphere quantities. Multi-core workstations allow multiple copies of the model to run simultaneously without significant performance penalty to, for example, study model dependence on a particular parameter. Additional work that improves initialization and data-analysis workflow is needed before parameter sweeps can be usefully extended to tens of parameter values or more. Significant high-performance computing (HPC) resources were required to process the MUIR phase-coded long-pulse data into high-resolution range-time HFIL images. Improvements to calculation efficiency are needed to allow more detailed “quick looks” of MUIR data at the HAARP facility during experiment campaigns, as well as for large-scale off-site processing of archival data.

Appendix. Distributed radar data processing on *Midnight*

This appendix describes parallel programming techniques with the Matlab problem-solving environment that were used to run modular UHF ionosphere radar (MUIR) coded long-pulse range-decoding software the *Midnight* supercomputer. *Midnight* is a Sun Microsystems Opteron cluster maintained at the Arctic Region Supercomputing Center. A brief overview of Matlab is provided, followed by a review of software extensions to Matlab that allow it to take advantage of machines with multiple-processors with or without distributed memory. The Mathworks Parallel Toolbox with the Matlab Distributed Computing Server was installed and configured to run on *Midnight* cluster, a shared resource at the Arctic Region Supercomputing Center (ARSC), such that Matlab and its associated parallel tools run in a “virtual cluster” allocated to the user by the batch queue system. After describing the configuration process unique to *Midnight*, a case study is presented that explores the computational performance and scalability of a simulated radar-processing task using parallel Matlab on *Midnight*. The computational performance is found to scale to the maximum number of Matlab licenses available at ARSC despite the small amount of programmer effort required to parallelize the example code.

A.1 Matlab overview

It has been observed that the most significant time or financial expense of solving problems with computational resources is the cost of the programmers hired to cast the problems in terms amenable to machine computation. The Matlab® problem-solving environment—developed and licensed by The Mathworks, Inc.—provides high-level programming tools for applying mathematical operations to multidimensional arrays of data that are independent of the underlying hardware, operating system, and installed mathematical library implementations. By separating the programmer from attending to the time-consuming tasks unique to each system that serve only to distract from the problem at hand, Matlab can dramatically reduce the total cost of a problem solution.

The Matlab application provides the user with a comprehensive integrated development environment (IDE), similar to Microsoft Visual Studio®, tightly coupled to the Matlab Runtime Environment (MRE), similar to the Java Runtime Environment (JRE). Together, the Matlab IDE and MRE compose a rich—and platform independent—virtual machine for writing and executing Matlab programs. A Matlab user may interact with the MRE directly through the “Desktop”

workspace that interactively evaluates Matlab programming commands as they are entered, or interact with the MRE indirectly through by evaluating Matlab programs (called *m-files*) written with the Matlab editor or some other text editor. Other tools in the IDE include a debugger, profiler, and source code checker.

Although a graphical display is necessary to use the Matlab desktop environment, the MRE may be used to execute Matlab programs even when a graphical display is not available. The Matlab workspace also may be accessed through a text-based terminal. For a Matlab session run on a remote machine, the graphical desktop is light enough to run responsively through an X server and a typical LAN connection. For instance, the interactive Matlab plotting tools are regularly used at the Arctic Region Supercomputing Center (ARSC) to interactively image, zoom, pan, and crop very large (>4 GB) ionosphere radar data from a Linux workstation connected to a Matlab session running on a large-memory node of the *Midnight* cluster.

Since Matlab is a full-featured and general—albeit very high level—programming language, there is no “typical” Matlab program. However, the Matlab language allows easy and natural expression of mathematical operations on multidimensional structured data that often require at least an order of magnitude more lines of code to accomplish with a traditional HPC programming language like Fortran or C. As an interpreted environment, Matlab programs that have been translated line-by-line from Fortran or C program source code will generally perform at least an order of magnitude worse than the original compiled program, despite the Matlab “Just-In-Time” (JIT) compilation technology³. But given that many Matlab commands are, at their core, little more than user-friendly interfaces to compiled BLAS⁴, LAPACK⁵, and now ScaLAPACK⁶ library functions; the performance of a Matlab program written using multidimensional vectors in place of nested loops,

³ <http://www.mathworks.com/products/matlab/description2.html>

⁴ <http://www.netlib.org/blas/>

⁵ <http://www.netlib.org/lapack/>

⁶ <http://www.netlib.org/scalapack/>

for example, may approach or even exceed the performance of a similar Fortran or C program written by an inexperienced HPC programmer.

Many *toolboxes*, packages of specialized libraries and programs, are available for license from The Mathworks or third parties. Two of these toolboxes provide parallel capability to Matlab programs are reviewed briefly in the next section.

A.2 Review of Parallel Matlab implementations

With the increasing availability of computational clusters and multiprocessor-core workstations, there is expanding interest in extending the cost-saving tools provided by Matlab to cover the complicated and difficult, i.e. expensive, realm of programming parallel machines. This section provides a brief overview of several techniques available to leverage a shared distributed resource through the Matlab problem-solving environment, concluding with two products offered by The Mathworks that allow Matlab to access parallel capabilities of clusters and multiprocessor core workstations. The computational performance and scalability of selected applications of these two products, as implemented on the *Midnight* multiuser cluster maintained at the Arctic Region Supercomputing Center (ARSC), will be discussed below.

A.2.1 Multithreaded Matlab

Selected functions included with Matlab® versions 7.4 and later—such as the matrix and array multiplication functions—are capable of leveraging multiple processors or processor cores in a single workstation or node through the use of multithreading. This capability may be enabled through the “preferences” menu of the Matlab desktop interface or through the *maxNumCompThreads* Matlab command. A list of Matlab functions that may take advantage of multithreaded computation is provided in the Mathworks online support documents⁷.

The *multithreadedcomputations* demonstration script included with Matlab 7.4, for example, calculates the speedup of various multithreaded functions using a user-specified number of threads. The speedup results plotted in Figure A.1 were calculated by the demonstration script, run

⁷ <http://www.mathworks.com/support/solutions/en/data/1-4PG4AN/?solution=1-4PG4AN>

in turn on two different types of compute nodes in *Midnight*—a 4-core Sun Fire X2200 and a 16-core Sun Fire X4600—using either 4 or 8 threads. The performance speedup of multithreaded computations on each compute node is comparable to the speedup attainable using a high-end workstation.

The Matlab computational performance and scalability case studies discussed below are focused on multi-node computations with the Mathworks Parallel Computing Toolbox™ (PCT) and Matlab Distributed Computation Server™ (MDCS). Since the multithreaded computation feature reviewed here is disabled when using multiple Matlab workers via the MPT or MDCS, due to the potential for conflicts and confusion when allocating Matlab workers and processors for optimal performance, the Matlab functions called from the code examples in the following performance and scalability study are single-threaded.

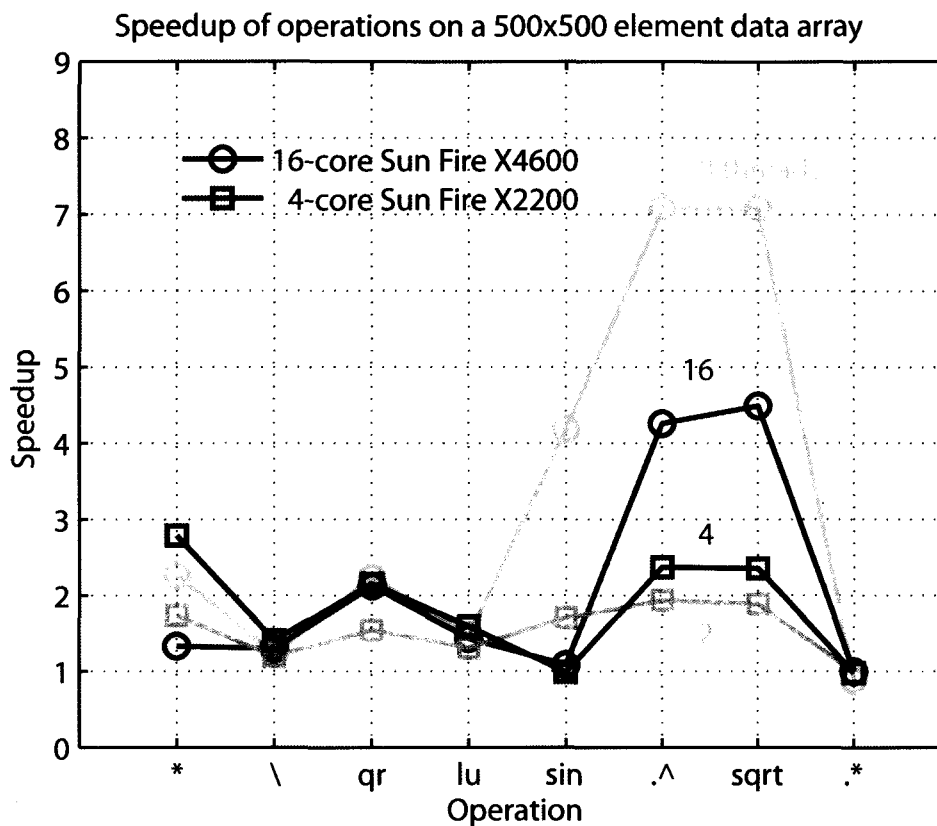


Figure A.1 Speedup of selected multithreaded Matlab functions using two processor cores per thread (light colors) and one core per thread (dark colors) measured on a 4-core (blue squares) and on a 16-core (red circles) *midnight* compute node. The functions operated on one, or two if necessary, 500x500 element arrays of real numbers.

A.2.2 MIT-LL MatlabMPI and pMatlab Parallel Toolbox

MatlabMPI⁸ and the pMatlab Parallel Toolbox⁹ are two technologies created and maintained at the Massachusetts Institute of Technology (MIT) Lincoln Laboratory (LL) that bring parallel capability

⁸ <http://www.ll.mit.edu/MatlabMPI>

⁹ <http://www.ll.mit.edu/pMatlab>

to Matlab running on multiprocessor machines with shared or distributed memory. The pMatlab Parallel Toolbox provides high-level parallel capabilities similar in some respects to the capabilities provided by the Mathworks Parallel Computing Toolbox™ and Matlab Distributed Computing Server™. The pMatlab library is built with low MPI-level parallel functions provided by the MatlabMPI software. Both MatlabMPI and pMatlab are released under a free software license; a single-user Matlab license is required for each active node using these technologies.

A.2.3 Shell-script distributed Matlab

Multiple copies of Matlab may be run simultaneously on a multi-processor core machine to complete distributed or data-parallel tasks faster than one copy of Matlab could complete the sub-tasks in serial. A distributed task refers here to a task that can be divided into smaller, independent, sub-tasks and solved independently, possibly in parallel and in arbitrary order, without requiring communication between sub-tasks. Matlab licenses are debited by machine (or node), so multiple copies of Matlab can run simultaneously on a single node under a single Matlab license.

Shell scripts may be used to automate the workflow process of launching multiple Matlab instances, assigning tasks to each Matlab instance, and gathering the results from each instance. One particularly effective technique uses multiple copies of Matlab¹⁰ with bash or ksh *here-documents* to perform a parameter sweep¹¹ on the *Midnight* cluster and is described in the ARSC HPC Users' Newsletter. While the number of Matlab instances that may be used efficiently with this technique is typically limited to the number of processor cores available in a single node, jobs submitted to *Midnight* may take advantage of the PBS Pro *job array* feature to launch even more copies of Matlab spanning over multiple nodes (at the expense of an additional Matlab license for each node). Note that the multithreaded computation feature needs to be manually disabled for optimal performance while running multiple copies of Matlab per node.

¹⁰ <http://www.arsc.edu/support/news/HPCnews/HPCnews378.shtml#article3>

¹¹ <http://www.arsc.edu/support/news/HPCnews/HPCnews381.shtml#article2>

The self-consistent ionosphere model (SCIM) was parallelized by way of Matlab shell-script parallelization. Although the bulk of simulation wall time is spent within the `pdepe` function call that solves the system of parabolic model equations, there is no available parallelization of this method at present. On a single processor core of *Midnight*, the implementation of SCIM described in this thesis runs in approximately real time. That is, one hour of simulation time requires approximately one hour of wall time. However, if multiple processor cores are available, as on the 4-core “4-way” or 16-core “16-way” nodes of *Midnight*, then shell script parallelization may be used to perform a SCIM parameter sweep where multiple copies of SCIM are run on a single node simultaneously, each copy differs by a single parameter. For example, Figure 4.13 was produced by a parameter sweep of four HF-electron volume heating rates running on a *Midnight* X2200 node. SCIM parallel parameter sweeps were also used to find appropriate driving parameters like the top boundary flux that in order to match SCIM ionosphere profiles to target ambient conditions determined by ionosonde or incoherent scatter radar (ISR).

A.2.4 Parallel Computing Toolbox and Matlab Distributed Computing Server

The Mathworks, Inc. Parallel Computing Toolbox™ (PCT) provides high-level Matlab and Simulink® programming tools, including language constructs for parallel processing on a multicore or multiprocessor workstation similar to a single compute node of the *Midnight* cluster. A copy of Matlab or Simulink is required to use the features of the PCT. The Matlab Distributed Computing Server (MDCS) extends the functionality of Matlab with the Parallel Computing Toolbox to multi-workstation clusters or to multiple nodes of *Midnight*. Figure A.2 illustrates the workflow relationship between the PCT and the MDCS¹². Similar to how the Matlab problem-solving environment separates the user interface to matrix computations and manipulations from hardware- and library software-dependent tasks like memory management and data input/output; the PCT separates the interface to task- and data-parallel operations from hardware- and network-dependent tasks¹³. One end result of such a separation is the dramatic reduction in programmer time necessary to implement a new parallel application or to convert an existing serial application to a parallel application.

¹² <http://www.mathworks.com/support/solutions/en/data/1-3U7V8R/index.html?product=DW>

¹³ <http://www.mathworks.com/products/parallel-computing/>

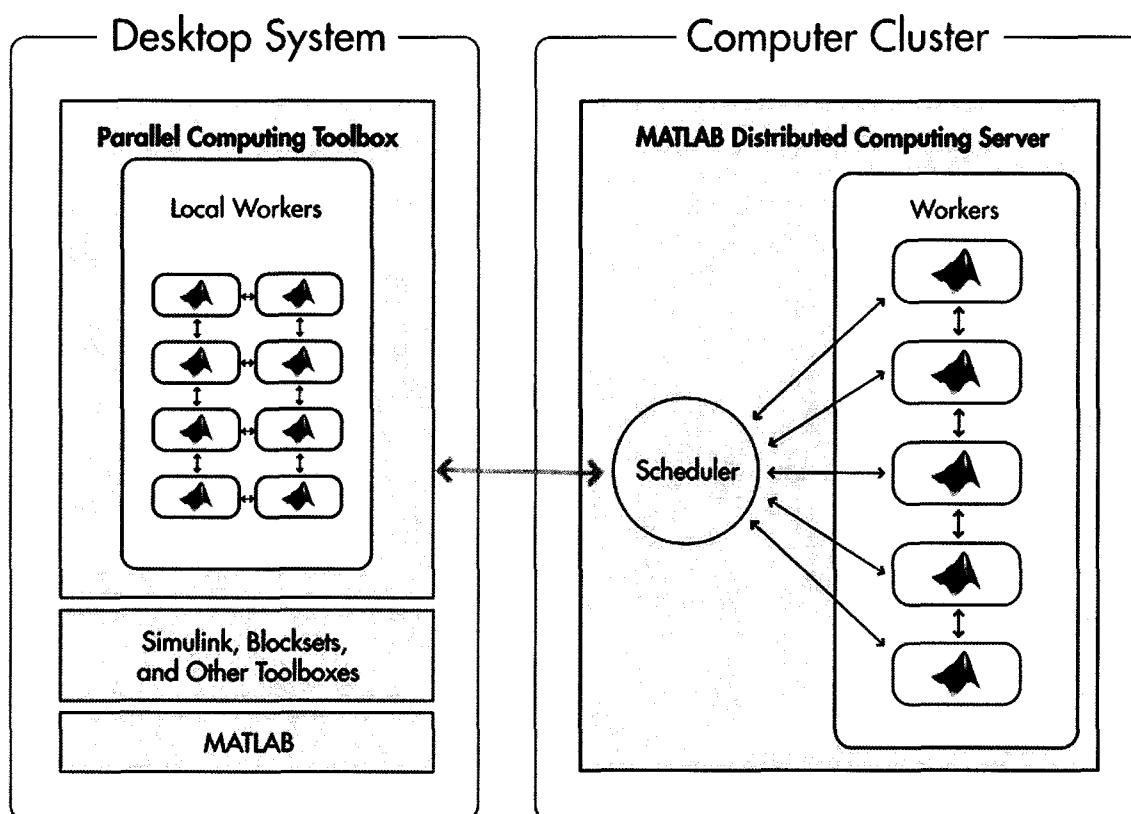


Figure A.2 Workflow overview of the Parallel Computing Toolbox and Matlab Distributed Computing Server. As configured on ARSC's *Midnight* cluster, both the Matlab client with the Parallel Computing Toolbox (PCT) and the Matlab Distributed Computing Server (MDCS) are started and run on a set of compute nodes allocated to a user by the PBS Pro batch queue system. The image is provided by The Mathworks.

Both the Parallel Computing Toolbox and Distributed Computing Server products require licensing from the Mathworks in addition to the license that grants interactive use of Matlab by one or more users. The cost of the parallel computing license depends on, among other factors, the number of independent Matlab "labs" or "workers" available simultaneously. Each Matlab worker is the software equivalent of a processor core in the sense that it is capable of executing tasks and accessing or storing information independent of the other Matlab workers, yet is controlled by a central managing process. The number of Matlab workers used simultaneously is typically less than the number of processor cores available in the cluster. Just as the cost of a single-user Matlab license is likely to be significant relative to the acquisition cost of the destination workstation, the

cost of fully licensing the Parallel Computing Toolbox and Matlab Distributed Computing Server on a cluster is significant relative to the acquisition cost of the cluster.

For example, an academic license to use MDCS with 32 workers, the number of workers licensed at ARSC, costs \$5,500 or \$172 per worker annually¹⁴. The academic license cost per worker decreases to \$115 for a 256-worker license, for a total annual cost of \$29,500. Note that the license for the MDCS and its respective workers is purchased in addition to the prerequisite Matlab and PCT licenses, both of which are licensed per user. This means that a user who starts a *matlabpool* (discussed below) containing, for instance, 8 out of the 32 available Matlab workers at ARSC, then the remaining 24 worker licenses are inaccessible to other users if, as is the case at ARSC, there is only one single-user PCT license available.

A further note on Mathworks licensing is appropriate. A Matlab user at ARSC is subject to Matlab licenses purchased by one or more of the following organizations: the Arctic Region Supercomputing Center, the University of Alaska, and the Department of Defense (DoD). While the set of ARSC users is broad and diverse in terms of organizational affiliation, the group of users authorized according to affiliation by each license from The Mathworks may be correspondingly narrow. For example, an ARSC user may be a student or staff member of the University of Alaska Fairbanks (or both), a student of the University of Alaska Anchorage, a visiting student from an outside university that is part of a summer research internship program, or a remote DoD user. At the time of writing in early January 2010, the University of Alaska Matlab site license used on the local ARSC workstations is plentiful and allows only University of Alaska Fairbanks users, while the multi-center Matlab license used on *Midnight* purchased by the DoD is scarce but allows any ARSC user to launch Matlab. Licensing for various combinations of the dozens of available Matlab toolboxes is similarly difficult for the user, since each license includes distinct but intersecting sets of toolboxes.

One additional and significant complication is that the license negotiated and purchased by the University of Alaska, the interpretation of the license, and attempts to remain in compliance with the license change annually, usually sometime between January and February. From the perspective of an ARSC user of parallel Matlab, this means that the Matlab resource incorporated in

¹⁴ The MathWorks *Products and Prices for North America Academic Users*, January 2010.

current projects must be assumed to expire at the end of each year, and that the project implementation may need to be adjusted, perhaps significantly, in order to remain compatible with the new licensing scheme. On *Midnight* during 2009 for example, the Matlab Distributed Computing Server is authorized by a license purchased by ARSC, but the Matlab and Parallel Toolbox products are authorized by a license purchased by the Department of Defense. But in 2008, Matlab and the Parallel Toolbox were authorized by the University of Alaska license, and will be again in 2010. Each January is a period that this author discovers so-called *license instabilities* in existing code.

A.3 Multi-node parallel Matlab on ARSC's *Midnight* via the PCT and MDCS

Matlab, the Parallel Computing Toolbox, and the Matlab Distributed Computing Server are installed on the *Midnight* cluster at the Arctic Region Supercomputing Center. This section describes the parallel Matlab initialization procedure unique to *Midnight* in detail, and then two case studies are presented that explore the performance and scalability of certain parallel tasks.

A.3.1 Initialization and configuration on *Midnight*

A procedure for starting a parallel Matlab session using the MDCS and spanning one or more *Midnight* compute nodes will be described. This procedure is a culmination of techniques developed for various Matlab projects to leveraged high-performance computing (HPC) resources at the Arctic Region Supercomputing Center (ARSC) from 2006 to 2009 and applies to Matlab version 7.6.0 running on the SUSE Linux operating system. Only the *interactive* mode of operation on *Midnight*, allowing a user to interactively execute Matlab commands on the allocated compute nodes, will be discussed in detail here. Since the parallel Matlab *batch* mode on *Midnight* is initialized with a similar sequence of steps, this focus will not suffer much loss of generality.

It is assumed that the user has connected to a *Midnight* login node from an *xterm* window using *ssh* with *X* forwarding enabled (i.e. using the command `ssh -X`). While a working *X*-server on the client workstation is not required to use the parallel Matlab interactive mode, it does allow the use of the Matlab graphical desktop that includes useful parallel configuration, coding, and debugging tools. The Matlab graphical configuration tools for parallel operation are particularly useful, but these tools are only needed to complete the initial configuration by each user the first time parallel

Matlab is run. Or if a standard configuration file is provided by ARSC consultants, then the configuration step may be skipped entirely.

The user connects to the Matlab client running on a *Midnight* compute node from a workstation through a terminal window or through an X session. Additionally, it is assumed here that the user has access to a *Midnight* filesystem like *WORKDIR* or *HOME* that is shared across the login and compute nodes, and is used by the Matlab workers to store logging or other information. Both the PCT and MDCS components shown in Figure A.2 reside on a set of compute nodes allocated to a specific user through the PBS Pro batch queue system. That is, the PBS Pro batch scheduler allocates a “cluster” of compute nodes to the user, and the user must initialize and set up their own personal Matlab PCT and MDCS environment. This process is automated with shell scripts, and with proper licensing, allows multiple users to simultaneously run their own private “Matlab cluster” on *Midnight*. Note that typical PCT and MDCS installation examples described by The Mathworks assume that the Matlab cluster may be shared by multiple users. However, security and data integrity requirements at ARSC do not easily facilitate this sort of use case on HPC-allocated shared resources.

The steps required to establish a parallel Matlab interactive mode session—with the graphical desktop interface enabled and displayed on the user workstation—using a cluster of *Midnight* compute are now described.

First, the *tunnelx* script is launched by the user in the X-forwarded *ssh* session connected to a *Midnight* login node. The *tunnelx* program allows an X-client on a compute node to connect with an X-server on the user workstation, allowing a Matlab instance on a compute node to display the graphical desktop on the user workstation. Then a script is submitted to the PBS Pro batch queuing system that requests a set of *Midnight* compute nodes to be allocated for the job, and then executes other initialization commands on the *manager* node (the remaining nodes in the allocation will be referred to here as the *worker* nodes). This batch script performs the following tasks, described below.

From the shell running on the manager node allocated by PBS, the *tunnelx* program is launched again and the configuration parameters for the local session are loaded with the shell *source* command. An environment variable containing the address to the Matlab license server is set on

the manager node and the *module* command is used at this point to set other environment variables to paths containing the Matlab installation files.

Next, the Matlab Distributed Computing Engine (MDCE, part of the MDCS package) is started on the manager node and each of the worker nodes. Then the Mathworks *jobmanager* scheduler is started on the manager node and Matlab workers are started on both the manager and worker nodes. Usually up to one Matlab worker is started for each processor core in the allocated nodes, up to the number of available Matlab worker licenses. (In an approximate sense, *jobmanager* scheduler is to the Matlab workers as the PBS batch queuing system is to the Midnight compute nodes in that the *jobmanager* scheduler allocates Matlab workers to a queue of Matlab tasks that have been submitted by one or more users.)

Finally, an instance of Matlab is started on the manager node that provides the user interface or executes the main user program. The Matlab graphical desktop environment is displayed on the user workstation. If the parallel environment has not been configured yet, then the graphical configuration tool under the Matlab desktop *parallel* menu may be used to set the default job manager to the Mathworks job manager. Diagnostic tests are also available in the Matlab parallel configuration tool that will validate the operation of the parallel Matlab features and save *configuration* file. The configuration step may be skipped if the new configuration is set as default and the associated configuration file is in the Matlab path (or the Matlab launch directory) the next time a parallel Matlab session is started.

At this point the user may, for example, start a *matlabpool* or *pmode* session from the Matlab workspace. A *matlabpool* is a set of Matlab labs¹⁵ (referred to here as *workers*) reserved by the user to complete concurrent tasks through programming constructs like the *parfor* parallel for loop or *spmd* single program multiple data block. A *pmode* session is a parallel-mode session where the user may send commands to each worker and view the results interactively; in particular, *pmode* sessions allow the interactive use of distributed matrices. The performance and scalability of the

¹⁵ The Parallel Computing Toolbox glossary defines a lab as follows: “When workers start, they work independently by default. They can then connect to each other and work together as peers, and are then referred to as labs.”

parfor loop connected to a *matlabpool* and non-trivial distributed matrix computations in a pmode session will be explored through the following case studies.

A.3.2 Case study: Modular UHF ionosphere radar data processing

This section describes parallelization and performance speedup of the Modular UHF ionosphere radar (MUIR) range-decoding algorithm using the Matlab *parfor* loop on *Midnight*. The Matlab Distributed Computing Server (MDCS) was used to task-parallelize the calculation over multiple compute nodes.

A.3.2.1 Objective

A compute cluster may be leveraged to complete a task by using more processors to decrease the time required to complete that task. The purpose of this study is to estimate the scalability and speedup of an example concurrent processing task using the *matlabpool parfor* feature in the PCT that may use up to a license-limited 32 workers distributed over multiple compute nodes of *Midnight* with the MDCS. The processing power dependence on the number of Matlab workers running per available processor core is also explored.

A.3.2.2 DCT and MDCE configuration

The numerical experiments described below were conducted with Matlab 7.6 and the PCT running on a *Midnight* compute node with the graphical display tunneled to a Linux workstation at ARSC. A *matlabpool* of 32 workers was started, along with the MDCS on multiple compute nodes allocated by the PBS Pro batch queuing system. Tasks were allocated to the workers from the Matlab graphical desktop session through the Mathworks *jobmanager* scheduler. The total time to open a parallel Matlab session distributed over multiple *Midnight* nodes from the instant PBS runs the Matlab and MDCE start script is under five minutes and depends on the number of workers requested. Once the parallel Matlab session is started on *Midnight*, less than 30 seconds are needed to open a *matlabpool* session of 32 workers. A detailed overview of the steps taken to establish a parallel Matlab interactive mode session on a cluster of *Midnight* compute nodes is described in section A.3.1 above.

A.3.2.3 Numerical experiment

The Arctic Region Supercomputing Center (ARSC) on the University of Alaska Fairbanks (UAF) campus is located near significant sources of experimental and operational data from two AMISR-class¹⁶ phased-array radars built to observe the polar ionosphere: the Poker Flat Incoherent Scatter Radar (PFISR) and the AMISR-prototype Modular UHF Ionosphere Radar (MUIR). MUIR is a diagnostic tool used to observe ionosphere interactions with powerful high-frequency (HF) electromagnetic waves transmitted by the co-located Ionospheric Research Instrument (IRI) at the High-Frequency Active Auroral Research Program (HAARP) facility at Gakona, Alaska. A typical HAARP experiment campaign may generate 40 hours of MUIR data, stored in HDF5¹⁷ files each containing about one minute of baseband data. The file storage size needed for an experiment depends on the mode of radar operation used, but the data collection rate typically ranges from 3 to 18 Gigabytes per hour. Real-time views of MUIR data as it is being collected are limited, so most data is processed following the conclusion of the campaign. Data is transferred manually on portable hard drives from the HAARP site to individual researchers or to UAF for later processing.

High-resolution range-time diagrams of received signal intensity are useful imaging products that may be derived from MUIR coded-pulse data files with Matlab. The computational cost per image depends on the radar sample rate, range of interest, inter-pulse period, and the radar pulse-code baud length. One *Midnight* X2200 compute node generally needs at least 1000 seconds of wall time to process one minute of typical MUIR coded-pulse data using single-threaded processing. This wall-time cost, and the speedup attainable by parallel processing, may be estimated by straightforward variations of the following Matlab script:

¹⁶ AMISR is the *Advanced Modular Incoherent Scatter Radar*

¹⁷ Hierarchical Data Format. See <http://www.hdfgroup.org/HDF5/>

```

nlabs = 32; nrows = 1024; ncols = 6000;
% simulation data (random complex matrix)
x = complex(randn(nrows, ncols), randn(nrows, ncols));
% pre-allocate memory for the (discarded) processed data
X = zeros(size(x));

matlabpool(nlabs)
parfor k = 0:nrows-1
    % Calculate the decoded power spectrum at the kth sample bin of
    % the radar pulses (columns of x), and store the result in the
    % matrix X.
    X = abs(fft(x .* circshift(x,k), nrows, 1)).^2;
end
matlabpool close

```

The script simulates the task of range-time processing one minute of coded-long pulse baseband data output from a typical MUIR experiment. Note that the example source code has already been parallelized for the PCT by opening a *matlabpool* session with 32 workers and replacing the *for* loop with a *parfor* loop. Parallelization through the *parfor* loop is possible in this case because each loop calculation may be completed independent of the results from any other iteration. When the program is started and a *matlabpool* of 32 workers has been successfully opened, the iteration tasks will be distributed to the workers and the tasks will be completed concurrently and in arbitrary order.

Inside the k^{th} loop iteration, a compound operation first phase-decodes the return and then calculates the signal return spectrum scattered from the k^{th} range bin. This calculation operates along the columns of the matrix x , that contain the simulated baseband signals received from the MUIR pulses. Only the execution time is of interest in this investigation, so the calculated spectrum is discarded.

Two different classes of measurements were performed: 1. The wall time required to perform the compound calculation inside the loop was measured while the worker density of ρ workers per processor core was varied over the two classes of *Midnight* nodes. 2. The total wall time required to complete all loop iterations was measured while the number of workers in the *matlabpool* was varied up to the maximum number of workers currently allowed by the Mathworks Distributed Computing Server license at ARSC.

A.3.2.4 Results and discussion

The Matlab *tic* and *toc* functions (not shown in the source code example) were used in the first experiment to measure the wall time required to complete an iteration of the example loop. Figure A.3 shows the probability distribution function of wall time required to complete the loop calculation. Dark-color curves are probability distribution functions of the calculation time required to process a 1024-by-6000 element matrix of complex numbers with normally distributed Cartesian coordinates; each color represents the distribution function estimated from the wall-time frequency distribution recorded for a particular (ρ , node class) pair. The frequency distribution used 100 logarithmically-sized bins.

Most MUIR coded long pulse experiments produce one 1100-by-5000 element complex matrix for each 52 seconds of operation. A new class of MUIR coded long pulse experiments are in development use a faster sample rate and larger maximum range, producing one 8100-by-5000 element complex matrix for each 52 seconds of data. The light colored curves are the respective probability distribution functions associated with 8192-by-6000 element random complex matrices. Mean loop execution wall times calculated in each experiment are indicated on the plot.

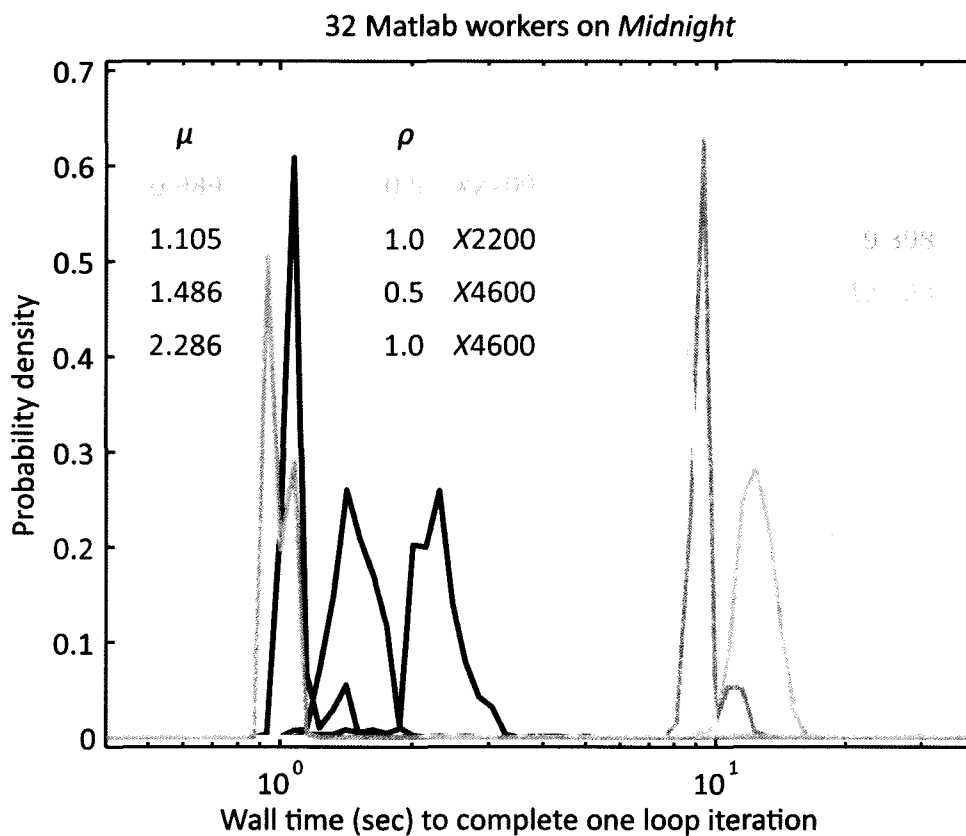


Figure A.3 Probability distribution function of wall time (sec) required to complete one *parfor* loop iteration in the example program code. Colors represent experiments performed on combinations of *Midnight* node class (4-core X2200 nodes vs. 16-core X4600 nodes) and worker density of ρ workers per processor core. Dark (light) colored curves are probability distribution functions of the time to calculate the decoded signal spectrum at a single row—representing a single radar range bin—estimated for random 1024-by-6000 element (8192-by-6000 element) complex matrices. The mean time μ to complete a loop iteration in each experiment is reported in seconds.

The mean times listed in Figure A.3 suggest that each minute-length file of standard coded long pulse data will require about 1000 seconds total compute time to process the 1000 loop iterations, using a set 4-core Sun Fire X2200 nodes. Allocating one or two processor cores per Matlab worker in the *matlabpool* does not appreciably alter the mean loop execution time. Twice as much compute time is required to process the same file using a set of 16-core Sun Fire X4600 nodes; and while decreasing the worker-to-core ratio to 50% does improve performance somewhat, the performance as measured by wall time is still less than that of the X2200 nodes. Finally, observe that processing each minute-length file from the new high-resolution extended-range coded long

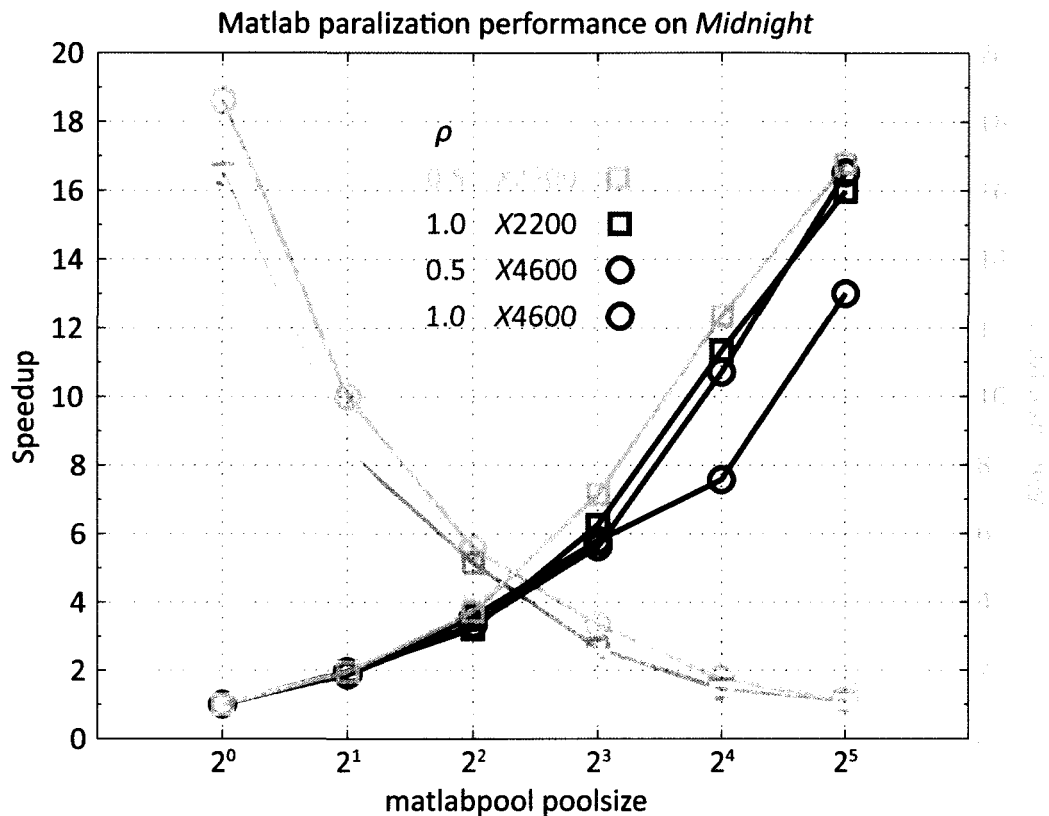


Figure A.4 Speedup factor (dark colors) of the total execution wall time to complete all loop iterations concurrently over an experiment run using *parfor* with a single-worker *matlabpool* on the indicated *Midnight* compute node type. Each color represents a sequence of experiment trials where the number of workers in the *matlabpool* was doubled after processing each 1024-by-6000 element random complex matrix while the number of workers per available processor core was maintained below the indicated threshold ρ . The total execution wall time (light colors) required to process each matrix does not include the time to start Matlab or to connect to the *matlabpool*.

pulse experiments will require nearly 10,000 seconds of total compute time (several hours), or well beyond what is practical to accomplish on a single workstation given the typical time available between experiment campaigns.

The total wall time required to complete all 1024 loop iterations using the example code on a 1024-by-6000 element random complex matrix was measured in the second experiment. Each experiment run consisted of a sequence of trials performed on one of the two classes of *Midnight* compute nodes, where the time to process the random matrix was measured while the size of the *matlabpool*—measured by number of available Matlab workers, or the variable *nlab*s in the source

code example—was doubled after each trial from 1 to 32, the maximum number of workers allowed by the MDCS license at ARSC. Four experiment runs were performed; each run is represented by a different color in Figure A.4. Both Matlab and the MDCS were initialized, as described in section A.3.1 above, on a set of nodes allocated by the PBS Pro batch allocation system. The number of nodes requested in each run was determined by the number of processor cores in one node of the chosen node class, and by a chosen threshold of maximum workers per processor core. Thresholds of 0.5 and 1.0 workers per processor core were chosen runs on the 4-core X2200 and 16-core X4600 compute nodes.

The dark-color curves in Figure A.4 show the speedup factor of the total wall time required to complete all loop iterations in the example code with a *parfor* loop connected to a *matlabpool*. The speedup for a particular run using a *matlabpool* with *n* workers is defined in each experiment run as the ratio of the wall time required to complete the iterations using a *matlabpool* with one worker to the time required using *n* workers. Observe that the scalability of speedup is relatively consistent for each (ρ , node type) combination. The two X4600 nodes each running 8 workers or more fared the worst relative to the other allocations running the same number of workers. A maximum speedup factor of 16 was obtained with eight X2200 nodes running 32 workers. Light-colored curves show the total calculation wall time as a function of the number of workers in the *matlabpool*. Across all (ρ , node type) combinations, the total wall time required to process each matrix—representing about a minute of radar data—decreases from about 17 minutes using one worker to just over one minute when using 32 workers.

This parallelization, obtained by simply replacing the *for* loop with a *parfor* loop in the example code, reduces the time-consuming task of processing standard coded long pulse data from MUIR to a task that can nearly be completed in real time. Completing a similar processing task on the new high-resolution and extended-range MUIR data is simply not practical on a standard workstation and must be completed on a parallel machine. The features in the PCT and MDCS make parallelization of the existing MUIR data processing code a near-trivial task for the programmer.

References

- Allen, M., et al. (1981), Vertical transport and photochemistry in the terrestrial mesosphere and lower thermosphere (50-120 km), *J. Geophys. Res.*, **86**.
- Anicich, V. G. (1993), Evaluated bimolecular ion-molecule gas phase kinetics of positive ions for use in modeling planetary atmospheres, cometary comae, and interstellar clouds, *J. Phys. Chem. Ref. Data*, **22**(6), 1469-1569.
- Ashrafi, M., et al. (2006), Heater-induced altitude descent of the EISCAT UHF ion line enhancements: Observations and modelling, *Adv. Space Res.*, **38**(11), 2645-2652.
- Ashrafi, M., et al. (2007), Spatiotemporal evolution of radio wave pump-induced ionospheric phenomena near the fourth electron gyroharmonic, *J. Geophys. Res.*, **112**(A5), 1-11.
- Axford, W. I. (1968), The polar wind and the terrestrial helium budget, *J. Geophys. Res.*, **73**.
- Bailey, S., et al. (2002), A model of nitric oxide in the lower thermosphere, *J. Geophys. Res.*, **107**(0).
- Bailey, V. A., and L. Goldstein (1958), Control of the ionosphere by means of radio waves, *J. Atmos. Terr. Phys.*, **12**, 216.
- Banks, P. (1966a), Charged Particle Temperatures and Electron Thermal Conductivity in the Upper Atmosphere, *Ann. Geophys.*, **22**, 557-587.
- Banks, P. (1966b), Collision frequencies and energy transfer, electrons., *Planet. Space Sci.*, **14**, 1085-1103.
- Banks, P. (1966c), Collision frequencies and energy transfer, ions., *Planet. Space Sci.*, **14**(1105-1122).
- Banks, P., and A. Nagy (1970), Concerning the influence of elastic scattering upon photoelectron transport and escape, *J. Geophys. Res.*, **75**(10), 1902-1910.
- Banks, P., and C. Chappell (1974), A new model for the interaction of auroral electrons with the atmosphere: Spectral degradation, backscatter, optical emission, and ionization, *J. Geophys. Res.*, **79**(10), 1459-1470.
- Banks, P. M. (1967), Ion temperature in the upper atmosphere, *J. Geophys. Res.*, **72**, 15323-15330.
- Bell, K. L., et al. (1998), Electron impact excitation of the ground-state ³P fine-structure levels in atomic oxygen, *Mon. Not. R. Astron. Soc.*, **293**, L83-L87.
- Bilitza, D., and B. W. Reinisch (2008), International Reference Ionosphere 2007: Improvements and new parameters, *Adv. Space Res.*, **42**(4), 1845-1950.

- Blelly, P.-L., et al. (1992), Observations of the structure and vertical transport of the polar upper ionosphere with the EISCAT VHF radar 1. Is EISCAT able to determine O^+ and H^+ polar wind characteristic? A simulation study., *Ann. Geophys.*, 10, 367-374.
- Blelly, P.-L., and R. W. Schunk (1993), A comparative study of the time-dependent standard 8-, 13-, and 16-moment transport formulations of the polar wind, *Ann. Geophys.*, 11, 443-469.
- Blelly, P.-L., et al. (1996), Numerical modeling of intermittent ion outflow events above EISCAT, *J. Atmos. Terr. Phys.*, 58(1-4), 273-285.
- Burgers, J. M. (1969), *Flow Equations for Composite Gasses*, Academic Press, New York.
- Campbell, L., et al. (2004), Production of vibrationally excited N_2 by electron impact, *Planet. Space Sci.*, 52(9), 815-822.
- Chamberlain, D. M., and J. W. Hunten (1989), *Theory of Planetary Atmospheres : An Introduction to Their Physics and Chemistry*, Academic Press.
- Cheung, P. Y., et al. (2001), High-power high-frequency-induced Langmuir turbulence in the smooth ionosphere at Arecibo. II. Low duty cycle, altitude-resolved, observations, *Phys. Plasmas*, 8(3), 802.
- Chisham, G., et al. (2007), A decade of the super dual auroral radar network (SuperDARN): Scientific achievements, new techniques and future directions, *Surv. Geophys.*, 28(1), 33-109.
- Colegrove, F. D., et al. (1965), Eddy diffusion and oxygen transport in the lower thermosphere, *J. Geophys. Res.*, 70.
- Colegrove, F. D., et al. (1966), Atmospheric composition in the lower thermosphere, *J. Geophys. Res.*, 71.
- Conde, M., and R. Smith (1995), Mapping Thermospheric Winds in the Auroral Zone, *Geophys. Res. Lett.*, 22(22), 3019-3022.
- Crank, J., and P. Nicolson (1996), A practical method for numerical evaluation of solutions of partial differential equations of the heat-conduction type, *Adv. Comput. Math.*, 6(1), 207-226.
- Dalgarno, A., and T. C. Degges (1968), Electron cooling in the upper atmosphere, *Planet. Space Sci.*, 16(1), 125-127.
- Dickinson, R. E., et al. (1981), A three-dimensional general circulation model of the thermosphere, *J. Geophys. Res.*, 86(A3), 1499-1512.
- Dickinson, R. E. (1984), Infrared radiative cooling in the mesosphere and lower thermosphere, *J. Atmos. Terr. Phys.*, 46(11), 995-1008.

- Diloy, P.-Y., et al. (1996a), Is EISCAT able to determine H^+ temperature and velocity? Numerical simulation, *J. Atmos. Terr. Phys.*, 58(1-4), 287-295.
- Diloy, P.-Y., et al. (1996b), A numerical model of the ionosphere, including the E-region above EISCAT, *Ann. Geophys.*, 14, 191-200.
- Djuth, F. T., et al. (1987), Large F-region electron-temperature enhancements generated by high-power HF radio waves, *Geophys. Res. Lett.*, 14.
- Djuth, F. T., et al. (1990), High resolution observations of HF-induced plasma waves in the ionosphere, *Geophys. Res. Lett.*, 17.
- Djuth, F. T., et al. (1994), Altitude characteristics of plasma turbulence excited with the tromsø superheater, *J. Geophys. Res.*, 99.
- Dougherty, J. P., and D. T. Farley (1960), A theory of incoherent scattering of radio waves by a plasma, *Proc. Roy. Soc.*, A259, 79-99.
- Dubois, D., et al. (1993), Space and time distribution of HF excited Langmuir turbulence in the ionosphere: Comparison of theory and experiment, *J. Geophys. Res.*, 98(A10), 17543-17567.
- DuBois, D. F., et al. (2001), High-power high-frequency-induced Langmuir turbulence in the smooth ionosphere at Arecibo. I. Theoretical predictions for altitude-resolved plasma line radar spectra, *Phys. Plasmas*, 8(3), 791.
- Duncan, L. M., and W. E. Gordon (1982), Ionospheric modification by high power radio waves, *J. Atmos. Solar Terr. Phys.*, 44(12), 1009-1017.
- Evans, J. (1969), Theory and practice of ionosphere study by Thompson scatter radar, *Proceedings of the IEEE*, 57(4), 496-530.
- Fallen, C. T., and B. J. Watkins (2008), Ambipolar Flux Response to Localized Electron Heating in the Polar F region, in *RF Ionospheric Interactions Workshop*, edited, Boulder, Colorado.
- Fallen, C. T., et al. (2009), TEC enhancements during HF ionospheric heating experiments, *AGU Fall Meeting Abstracts*, 13, 1468.
- Farley, D. T. (1963), Artificial Heating of the Electrons in the F Region of the Ionosphere, *J. Geophys. Res.*, 68(2), 401-413.
- Fejer, J. A. (1960), Scattering of radio waves by an ionized gas in thermal equilibrium, *Can. J. Phys.*, 38, 1114-1133.
- Forbes, J. M., et al. (1993), Magnetic activity dependence of high-latitude thermospheric winds and densities below 200 km, *J. Geophys. Res.*, 98.

- Frey, A., et al. (1984), First experimental evidence of HF produced electron density irregularities in the polar ionosphere; Diagnosed by UHF radio star scintillations, *Geophys. Res. Lett.*, 11.
- Ganguli, S. B. (1996), The polar wind, *Rev. Geophys.*, 34.
- Gear, C. W. (1971), *Numerical Initial Value Problems in Ordinary Differential Equations*, 253 pp., Prentice Hall PTR.
- Gerard, J. C., and D. Rusch (1979), The auroral ionosphere: Comparison of a time-dependent model with composition measurements, *J. Geophys. Res.*, 84(A8), 4335-4340.
- Gunther, S. (2006), *The History of Sounding Rockets and Their Contribution to European Space Research*, European Space Agency.
- Gustavsson, B., et al. (2001), First tomographic estimate of volume distribution of HF-pump enhanced airglow emission, *J. Geophys. Res.*, 106(A12), 29105-29123.
- Gustavsson, B., et al. (2002), Nearly simultaneous images of HF-pump enhanced airglow at 6300 Å and 5577 Å, *Geophys. Res. Lett.*, 29.
- Gustavsson, B., et al. (2005), The electron energy distribution during HF pumping, a picture painted with all colors, *Ann. Geophys.*, 23, 1747-1754.
- Gustavsson, B., et al. (2006), Electron gyroharmonic effects in ionization and electron acceleration during High-Frequency pumping in the ionosphere, *Phys. Rev. Lett.*, 97(19), 195002-195004.
- Hagfors, T. (1961), Density Fluctuations in a Plasma in a Magnetic Field, with Applications to the Ionosphere, *J. Geophys. Res.*, 66(6), 1699-1712.
- Hagfors, T., et al. (1983), Observations of enhanced plasma lines by EISCAT during heating experiments, *Radio Sci.*, 18(6), 861-866.
- Hansen, J. D., et al. (1992), Large-scale HF-induced ionospheric modifications: Experiments, *J. Geophys. Res.*, 97.
- Hastings, J. T., and R. G. Roble (1977), An automatic technique for solving coupled vector systems of non-linear parabolic partial differential equations in one space dimension, *Planet. Space Sci.*, 25, 209-215.
- Hedin, A. E. (1979), Neutral thermospheric composition and thermal structure, *Rev. Geophys.*, 17.
- Hedin, A. E. (1983), A revised thermospheric model based on mass spectrometer and incoherent scatter data: MSIS-83, *J. Geophys. Res.*, 88.
- Hedin, A. E. (1987), MSIS-86 thermospheric model, *J. Geophys. Res.*, 92.

- Hedin, A. E., et al. (1996), Empirical wind model for the upper, middle and lower atmosphere, *J. Atmos. Terr. Phys.*, 58(13), 1421-1447.
- Hinteregger, H. E., and K. Fukui (1981), Observational, reference and model data on solar EUV, from measurements on AE-E, *Geophys. Res. Lett.*, 8(11), 1147-1150.
- Hoegy, W. R. (1983), Thermal electron heating rate: A derivation, *J. Geophys. Res.*, 89.
- Horvath, I., and B. C. Lovell (2009), Distinctive plasma density features of the topside ionosphere and their electrodynamics investigated during southern winter, *J. Geophys. Res.*, 114.
- Huba, J., et al. (2000), Sami2 is Another Model of the Ionosphere (SAMI2): A new low-latitude ionosphere model, *J. Geophys. Res.*, 105(A10), 23035-23053.
- Itikawa, Y. (1971), Effective collision frequency of electrons in atmospheric gases, *Planet. Space Sci.*, 19(8), 993-1007.
- Jahn, R. G. (1962), Microwave probing of ionized-gas flows, *Phys. Fluids*, 5(6).
- Kellogg, W. W. (1964), Pollution of the upper atmosphere by rockets, *Space Science Reviews*, 3, 275-316.
- Kosch, M. J., et al. (2004), Novel artificial optical annular structures in the high latitude ionosphere over EISCAT, *Geophys. Res. Lett.*, 31.
- Kosch, M. J., et al. (2007), Temporal evolution of pump beam self-focusing at the High-Frequency Active Auroral Research Program, *J. Geophys. Res.*, 112(A8), 1-9.
- Kulchitsky, A., et al. (2005), $E \times B$ drift simulation in an Eulerian ionospheric model using the total variation diminishing numerical scheme, *J. Geophys. Res.*, 110(A9).
- Lai, S. T. (1999), Guide to reference and standard ionosphere models, American Institute of Aeronautics and Astronautics (AIAA).
- Lal, C. (1992), Global F_2 layer ionization and geomagnetic activity, *J. Geophys. Res.*, 97.
- Lawrence, F. S., et al. (1999), Solving index-1 DAEs in MATLAB and Simulink, *SIAM Rev.*, 41(3), 538-552.
- Lettau, H. (1951), Diffusion in the Upper Atmosphere, in *Compendium of Meteorology*, edited by T. F. Malone, American Meteorological Society, Boston.
- Lummerzheim, D., and J. Liliensten (1994), Electron transport and energy degradation in the ionosphere: evaluation of the numerical solution, comparison with laboratory experiments and auroral observations, *Ann. Geophys.*, 12, 1039-1051.

- Madsen, N. K. (1975), The method of lines for the numerical solution of partial differential equations, *SIGNUM Newsl.*, 10(4), 5-7.
- Mantas, G., et al. (1981), Thermal Response of the F Region Ionosphere in Artificial Modification Experiments by HF Radio Waves, *J. Geophys. Res.*, 86(A2), 561-574.
- Mathworks, I. (2007), Matlab 7.4, edited, Natick.
- Maurits, S. (1996), Modeling of the Polar Ionosphere in the Inertial Corotating Frame, Ph.D. Thesis thesis, University of Alaska, Fairbanks.
- Maurits, S., et al. (2000), WWW-based visualization of the real time run of a space weather forecasting model, in *Eurograph-IEEE VisSym'00 (Visualization Symposium'2000)*, edited, Amsterdam.
- McCarrick, M. (2008), Personal Communication, edited by C. T. Fallen.
- McLean, S., et al. (2004), US/UK World Magnetic Model for 2005-2010.
- Meltz, G., and R. LeLevier (1970), Heating the F region by deviative absorption of radio waves, *J. Geophys. Res.*, 75(31), 6406-6416.
- Meltz, G., et al. (1974), Ionospheric heating by powerful radio waves, *Radio Sci.*, 9(11), 1049-1063.
- Meltz, G., and F. Perkins (1974), Ionospheric modification theory: past, present, and future, *Radio Sci.*, 9(11), 885-888.
- Min, Q.-L. (1993), A self-consistent time varying auroral model, University of Alaska, Fairbanks.
- Min, Q.-L., et al. (1993), Effects of a parallel electric field and the geomagnetic field in the topside ionosphere on auroral and photoelectron energy distributions, *J. Geophys. Res.*, 98(A11), 19223-19234.
- Mingaleva, G. I., and V. S. Mingalev (1997), Response of the convecting high-latitude F layer to a powerful HF wave, *Ann. Geophys.*, 15(10), 1291-1300.
- Mingaleva, G. I., and V. S. Mingalev (2008), A model study of the large-scale modification of the nocturnal middle-latitude F layer by powerful HF waves with different powers edited, p. 3, Polar Geophysical Institute, Apatity, Russia.
- Minkoff, J., et al. (1974), Radio frequency scattering from a heated ionospheric volume, 1, VHF/UHF field-aligned and plasma-line backscatter measurements, *Radio Sci.*, 9.
- Nagy, A., and P. Banks (1970), Photoelectron fuxes in the ionosphere, *J. Geophys. Res.*, 75(31), 6260-6270.

- Narasinga Rao, B. C., and E. J. R. Maier (1970), Photoelectron flux and protonospheric heating during the conjugate point sunrise, *J. Geophys. Res.*, 75.
- Nisbet, J. (1967), Neutral atmospheric temperatures from incoherent scatter observations, *J. Atmos. Sci.*, 24, 586.
- NSSDC (2009), NSSDC Master Catalog: Spacecraft, edited by N. S. S. D. Center, National Aeronautics and Space Administration.
- Nsume, P. A., et al. (2008), Polar cap electron density distribution from IMAGE radio plasma imager measurements: Empirical model with the effects of solar illumination and geomagnetic activity, *J. Geophys. Res.*, 113.
- Oyama, S.-i., et al. (2006), Persistent enhancement of the HF pump-induced plasma line measured with a UHF diagnostic radar at HAARP, *J. Geophys. Res.*, 111(A06309).
- Pavlov, A. V. (1998a), New electron energy transfer and cooling rates by excitation of O₂, *Ann. Geophys.*, 16(8), 1007-1013.
- Pavlov, A. V. (1998b), New electron energy transfer rates for vibrational excitation of N₂, *Ann. Geophys.*, 16(2), 176-182.
- Pavlov, A. V., and K. A. Berrington (1999), Cooling rate of thermal electrons by electron impact excitation of fine structure levels of atomic oxygen, *Ann. Geophys.*, 17(7), 919-924.
- Pavlov, A. V., et al. (2000), Comparison of the measured and modelled electron densities and temperatures in the ionosphere and plasmasphere during 20/30 January, 1993, *Ann. Geophys.*, 18(10), 1257-1272.
- Pedersen, T., et al. (2010), Creation of artificial ionospheric layers using high-power HF waves, *Geophys. Res. Lett.*, 37.
- Pedersen, T. R., and H. C. Carlson (2001), First observations of HF heater-produced airglow at the High Frequency Active Auroral Research Program facility: Thermal excitation and spatial structuring, *Radio Sci.*, 36.
- Pedersen, T. R., et al. (2003), Magnetic zenith enhancement of HF radio-induced airglow production at HAARP, *Geophys. Res. Lett.*, 30(4), 18-11 - 18-14.
- Perkins, F. W., et al. (1974), Parametric instabilities and ionospheric modification, *J. Geophys. Res.*, 79(10), 1478-1496.
- Perkins, F. W., and R. G. Roble (1978), Ionospheric heating by radio waves: Predictions for Arecibo and the satellite power station, *J. Geophys. Res.*, 83(A4), 1611-1624.
- Perrine, R. P., et al. (2006), An interhemispheric model of artificial ionospheric ducts, *Radio Sci.*, 41(4), 1-13.

- Picone, J. M., et al. (2002), NRLMSISE-00 empirical model of the atmosphere: Statistical comparisons and scientific issues, *J. Geophys. Res.*, 107(A12), SIA 15-11-SIA 15-16.
- Prölss, G. W. (2004), *Physics of the Earth's Space Environment*, Springer-Verlag, Berlin.
- Rasmussen, C. E., and R. W. Schunk (1988), A photochemical equilibrium model for ionospheric conductivity, *J. Geophys. Res.*, 93(A9), 9831-9840.
- Rasmussen, C. E., et al. (1988), Comparison of simultaneous Chatanika and Millstone Hill temperature measurements with ionospheric model predictions, *J. Geophys. Res.*, 93, 1922-1932.
- Rees, M. H., and J. Walker (1968), Ion and electron heating by auroral electric fields, *Ann. Geophys.*, 24, 193-199.
- Rees, M. H., and R. Roble (1975), Observations and theory of the formation of stable auroral red arcs, *Rev. Geophys. Space Phys.*, 13(1), 201-242.
- Rees, M. H. (1989), *Physics and Chemistry of the Upper Atmosphere*, Cambridge University Press.
- Rietveld, M., et al. (1993), Introduction to ionospheric heating at Tromsø (Norway). Part 1: Experimental overview, *J. Atmos. Terr. Phys.*, 55(4/5), 557-599.
- Rietveld, M., et al. (2000), Measurements of HF-enhanced plasma and ion lines at EISCAT with high-altitude resolution, *J. Geophys. Res.*, 105(A4), 7429-7439.
- Rietveld, M. T., et al. (2003), Ionospheric electron heating, optical emissions, and striations induced by powerful HF radio waves at high latitudes: Aspect angle dependence, *J. Geophys. Res.*, 108(A4).
- Robineau, A., et al. (1996), Time-dependent models of the auroral ionosphere above EISCAT, *J. Atmos. Terr. Phys.*, 58(1-4), 257-271.
- Robinson, T. R. (1989), The heating of the high latitude ionosphere by high power radio waves, *Physics Reports*, 179(2-3), 79-209.
- Roble, R., and E. C. Ridley (1994), A thermosphere-ionosphere-mesosphere-electrodynamics general circulation model (time-GCM): Equinox solar cycle minimum simulations (30-500 km), *Geophys. Res. Lett.*, 21(6), 417-420.
- Roble, R. G. (1975), The calculated and observed diurnal variation of the ionosphere over Millstone Hill on 23-24 March 1970, *Planet. Space Sci.*, 23, 1017-1033.
- Roble, R. G., and J. T. Hastings (1977), Thermal response properties of the Earth's ionospheric plasma, *Planet. Space Sci.*, 25, 217-231.

- Roble, R. G., and E. C. Ridley (1987), An auroral model for the NCAR Thermospheric General Circulation Model (TGCM), *Ann. Geophys.*
- Roble, R. G., et al. (1987), On the global mean structure of the thermosphere, *J. Geophys. Res.*, 92(A8), 8745-8758.
- Roble, R. G., et al. (1988), A coupled thermosphere/ionosphere general circulation model, *Geophys. Res. Lett.*, 15.
- Salpeter, E. E. (1960), Scattering of radio waves by electrons above the ionosphere, *J. Geophys. Res.*, 65(6), 1851-1852.
- Salpeter, E. E. (1963), Density fluctuations in a nonequilibrium plasma, *J. Geophys. Res.*, 68(5), 1321-1333.
- Schunk, R. W., and J. Walker (1970a), Thermal diffusion in the F₂ region of the ionosphere, *Planet. Space Sci.*, 18, 535-557.
- Schunk, R. W., and J. Walker (1970b), Minor ion diffusion in the F₂ region of the ionosphere, *Planet. Space Sci.*, 18, 1319-1334.
- Schunk, R. W., and J. Walker (1970c), Transport properties of the ionospheric electron gas, *Planet. Space Sci.*, 18, 1535-1550.
- Schunk, R. W., and J. Walker (1971), Transport processes in the E region of the ionosphere, *J. Geophys. Res.*, 76(25), 6159-6171.
- Schunk, R. W., and J. Walker (1973), Theoretical ion densities in the lower ionosphere, *Planet. Space Sci.*, 21, 1875-1896.
- Schunk, R. W. (1975), Transport equations for aeronomy, *Planet. Space Sci.*, 23, 437-485.
- Schunk, R. W. (1977), Mathematical structure of transport equations for multispecies flows, *Rev. Geophys. Space Phys.*, 15(4), 429-445.
- Schunk, R. W., and A. Nagy (1978), Electron temperatures in the F region of the ionosphere: Theory and observations, *Rev. Geophys. Space Phys.*, 16(3), 355-399.
- Schunk, R. W. (1988), A mathematical model of the middle and high latitude ionosphere, *PAGEOPH*, 127(2,3), 255-303.
- Schunk, R. W., and A. Nagy (2000), *Ionospheres: physics, plasma physics, and chemistry*, Cambridge University Press, Cambridge.
- Secan, J. A., et al. (2008), Evidence of IRI-Generated Structure at 840km Altitude in DMSP In-Situ Observations, in *RF Ionospheric Interactions Workshop*, edited, Boulder, Colorado.

- Shampine, L. F., and M. W. Reichelt (1997), The MATLAB ODE Suite, *SIAM J. Sci. Comp.*, 18(1), 1.
- Sharp, L. R., et al. (1978), The Altitude Dependence of the Local Time Variation of Thermospheric Density, *Geophys. Res. Lett.*, 5.
- Shkarofsky, I. P. (1961), Values of the transport coefficients in a plasma for any degree of ionization based on a Maxwellian distribution, *Can. J. Phys.*, 39, 1619.
- Sincovec, R. F., and N. K. Madsen (1975), Software for nonlinear partial differential equations, *ACM Trans. Math. Softw.*, 1(3), 232-260.
- Skeel, R. D., and M. Berzins (1990), A method for the spatial discretization of parabolic equations in one space variable, *SIAM J. Sci. Stat. Comput.*, 11(1), 1-32.
- Solomon, S., et al. (1988), The auroral 6300 Å emission: Observations and modeling, *J. Geophys. Res.*, 93(A9), 9867-9882.
- Solomon, S., and V. J. Abreu (1989), The 630 nm dayglow, *J. Geophys. Res.*, 94(A6), 6817-6824.
- Solomon, S., and L. Qian (2005), Solar extreme-ultraviolet irradiance for general circulation models, *J. Geophys. Res.*, 110(A10306).
- Stamnes, K., and M. H. Rees (1983), Heating of thermal ionospheric electrons by suprathermal electrons, *Geophys. Res. Lett.*, 10(4), 309-312.
- Strobel, D. F., et al. (1970), Production and diffusion of nitric oxide, *J. Geophys. Res.*, 75(22), 4307-4321.
- Stubbe, P. (1970), Simultaneous solution of the time dependent coupled continuity equations, heat conduction equations, and equations of motion for a system consisting of a neutral gas, an electron gas, and a four component ion gas, *J. Atmos. Terr. Phys.*, 32, 865-903.
- Stubbe, P. (1996), The ionosphere as a plasma laboratory, in *Modern Ionospheric Science*, edited by H. Kohl, et al., European Geophysical Society.
- Sulzer, M. P. (1986), A radar technique for high range resolution incoherent scatter autocorrelation function measurements utilizing the full average power of klystron radars, *Radio Sci.*, 21(6), 1033-1040.
- Swartz, W., and J. Nisbet (1971), Diurnal variation of the neutral temperature profile at Arecibo from incoherent scatter measurements and its relevance to the 1400-hour density maximum, *J. Geophys. Res.*, 76(1), 185-196.
- Swartz, W. E., et al. (1971), Analytic expression for the energy-transfer rate from photoelectrons to thermal-electrons, *J. Geophys. Res.*, 76.
- Tellegen, B. D. H. (1933), Interactions between radio waves?, *Nature*, 840.

- Vallado, D. A., et al. (2006), Revisiting spacetrack report #3, in *AIAA/AAS Astrodynamics Specialist Conference*, edited, Keystone, CO.
- Watkins, B. J., and P. Banks (1974), A preliminary study of high-latitude thermospheric temperatures from incoherent scatter radar observations, *J. Geophys. Res.*, 79(34), 5307-5310.
- Whitten, R. C., and I. G. Poppoff (1971), *Fundamentals of Aeronomy*, 446 pp., J. Wiley, New York,.
- Wong, A. Y., et al. (1990), High-power radiating facility at the HIPAS Observatory, *Radio Sci.*, 25(6), 1269-1282.
- Woods, T. N., et al. (2005), Solar EUV Experiment (SEE): Mission overview and first results, *J. Geophys. Res.*, 110(A1), A01312.