MIDLATITUDE THERMOSPHERIC WIND AND TEMPERATURE: NETWORKED FABRY-PEROT INTERFEROMETER OBSERVATIONS AND RADIATIVE TRANSFER MODELING

BY

BRIAN JOSEPH HARDING

DISSERTATION

Submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy in Electrical and Computer Engineering
in the Graduate College of the
University of Illinois at Urbana-Champaign, 2017

Urbana, Illinois

Doctoral Committee:

Professor Jonathan Makela, Chair
Professor Erhan Kudeki
Assistant Professor Lara Waldrop
Assistant Professor Davide Curreli

ABSTRACT

This dissertation presents studies of the midlatitude and low-latitude thermosphere, primarily using networks of Fabry-Perot interferometers (FPIs). First, we describe an algorithm which estimates thermospheric line-of-sight wind and temperature from raw FPI data. This new algorithm has the advantage over previous work in that it provides accurate temperature estimates and uncertainties. We then present a novel regularization-based technique to estimate the thermospheric wind field from an FPI network's line-of-sight wind measurements. This technique makes no explicit assumptions about the functional form of the wind field, and instead lets the data inform the shape. We apply this technique to study the wind dynamics associated with the equatorial midnight temperature maximum, finding direct evidence of a converging wind field during its development. Next, we apply this technique to study the midlatitude thermospheric response to the geomagnetic storm of 01–02 Oct 2013. Though the horizontal wind and temperature measurements corroborate previous observations and theory in a broad sense, the downward vertical winds measured by six independent FPIs are unreasonably large (>100 m/s) and sustained (5 hours). A superposed epoch analysis of 15 different storms shows that such downward winds are commonly measured during the main phase. Using radiative transfer modeling, we show that these vertical winds are not real, and instead are artifacts of the scattering of airglow radiation in the lower atmosphere. This is the likely explanation for the large midlatitude vertical winds and horizontal convergences previously reported in the literature. We also show that some of the vertical winds repeatedly observed at equatorial latitudes may be explained as artifacts of atmospheric scattering. These results suggest that the effects of the lower atmosphere should be accounted for in any quantitative ground-based airglow measurement.

ACKNOWLEDGMENTS

First and most importantly, I would like to thank my adviser, Jonathan Makela, for his unbounded support and his unique ability to simultaneously care deeply about my future and give me the freedom to find my own path. I truly cannot imagine a better mentor, and I will miss working with him. In addition to the members of my committee, I would also like to thank Farzad Kamalabadi, Gary Swenson, John Meriwether, Koki Chau, and Marco Milla for mentorship.

This work is primarily focused on data from the NATION and RENOIR FPI networks, and it would not have been possible without the efforts of many scientists, engineers, and assistants to plan, fund, design, deploy, maintain, and operate these systems. Specific thanks go to John Meriwether, Jonathan Makela, Dan Fisher, Aaron Ridley, Greg Earle, Nathaniel Frissell, Marco Ciocca, Michael Castellez, Ricardo Buriti, Peter Sherwood, Sam Sanders, Rafael Mesquita, and the machine shop at Clemson University. Jeff Baumgardner, Carlos Martinis, Cristiano Wrasse, John Noto, and Robert Kerr provided independent data for this work and assisted with its interpretation. Grateful acknowledgment goes to Stan Solomon for providing the initial motivation for the work in Chapter 5 and to Jianqi Qin for useful conversations regarding radiative transfer.

Thanks to Tim Duly for spearheading the development of pyglow, a Python package that wraps several upper atmosphere climatological models and is used extensively in this work. The pyglow package is open-sourced and available at `https://github.com/timduly4/pyglow/`. I am grateful to Pablo Reyes for the thankless job of building and administering the server used to store and analyze our data. Thanks to Nancy Morris for administrative support. To my fellow graduate students, too numerous to list, thanks for all of your help in both the academic and personal realms. Dan, thanks for keeping the office fun.

None of the work during my 6 years in graduate school would have been possible without the love and support from my mom, dad, and sister during the previous 22. I cannot thank them enough. Finally, thanks to my partner Eileen for her love, encouragement, and understanding.

I acknowledge financial support by the Jack S. Kilby Fellowship, the Carver Fellowship, the National Science Foundation Graduate Research Fellowship under Grant No. DGE-1144245, the National Science Foundation under Grant Nos. NSF AGS-1452291 and AGS-1138998, and NASA under Grant No. NNG12FA45C. Any opinion, findings, and conclusions or recommendations expressed in this material are those of the author and do not necessarily reflect the views of the National Science Foundation.

TABLE OF CONTENTS

CHAPTER 1

INTRODUCTION

1.1 Earth's Thermosphere

Earth's atmosphere is described in terms of five regions, distinguished by the vertical variations of temperature, as shown in Figure 1.1. In the lowest region, the troposphere, the atmosphere cools with increasing altitude, until the absorption of ultraviolet solar radiation by ozone becomes important, and the temperature increases. The stratosphere and mesosphere span the region of significant ozone absorption. Above the mesosphere, the absorption of solar extreme ultraviolet (EUV) radiation causes the temperature to increase again, until asymptoting to a constant value. This region is the thermosphere, the focus of this dissertation. Above the thermosphere is the exosphere, which is usually defined to be the region where the neutral density becomes so low that the atmosphere no longer acts like a fluid, and kinetic equations are necessary. Although it is somewhat artificial to define exact regional boundaries, the thermosphere is usually defined to span from about 85 km to 500 km in altitude.

In the thermosphere, the primary constituents are atomic oxygen (O, dominant above ∼200 km) and molecular nitrogen (N_2, dominant below ∼200 km). Molecular oxygen (O_2), helium (He), and argon (Ar) are minor constituents. For the purposes of this work, which focuses on altitudes above 200 km, we assume that the thermosphere can be represented as a single fluid of atomic oxygen. It follows the ideal gas law:

$$p = k\frac{\rho}{M}T \tag{1.1}$$

where p is the pressure (Pa), k is the Boltzmann constant (1.38×10^{-23}

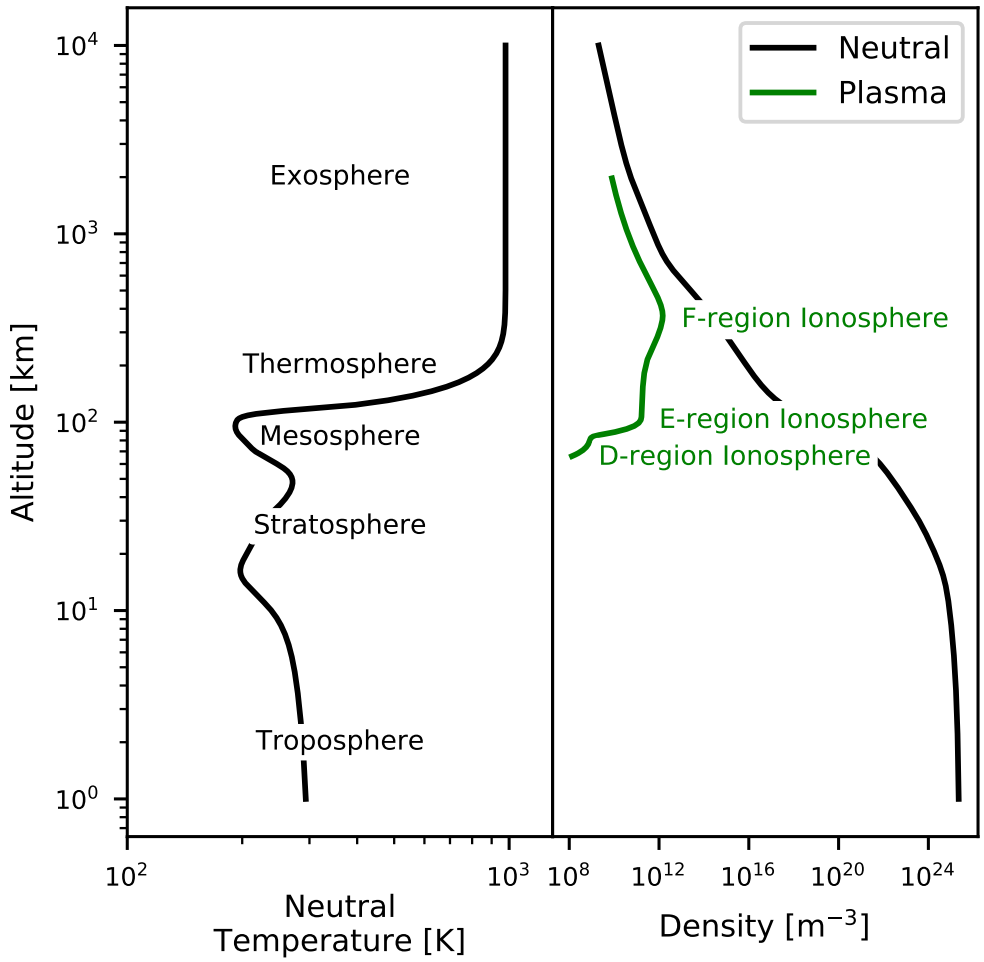

Figure 1.1: Layers of Earth's atmosphere. Neutral temperature and density are from the NRL-MSISE-00 model [*Picone et al.*, 2002]. Plasma density is the electron density from the International Reference Ionosphere (IRI) model [*Bilitza and Reinisch*, 2008]. These profiles are typical for daytime solar medium conditions.

m^2 kg s^{-2} K^{-1}), ρ is the mass density (kg m^{-3}), M is the mean atomic mass ($16 \times 1.67 \times 10^{-27}$ kg for atomic oxygen), and T is the temperature (K).

A common approximation for planetary atmospheres is hydrostatic equilibrium, where the downward force of gravity is exactly balanced by the upward pressure gradient force:

$$\frac{dp}{dz} = -\rho g \tag{1.2}$$

where z is altitude and g is gravitational acceleration. Integrating this equation and using the ideal gas law, we see that hydrostatic equilibrium implies

an exponential pressure profile:

$$p(z) = p_0 \exp\left(-\frac{z - z_0}{H}\right) \tag{1.3}$$

where the scale height, H, is related to the temperature:

$$H = \frac{kT}{Mg} \tag{1.4}$$

and is about 50 km in the upper thermosphere. Hydrostatic equilibrium is a good approximation on large spatial and temporal scales but is expected to break down when thermospheric forcing is rapid or local [*Deng et al.*, 2008; *Yigit and Ridley*, 2011].

The upper thermosphere is of particular interest because it contains the highest density plasma in geospace. Solar radiation ionizes thermospheric particles, populating the ionosphere, which is also shown in Figure 1.1. During the day, three layers of the ionosphere are evident, known as the D region, E region, and F region, though the boundaries between these regions are not well defined. The F region is dominated by O^+, while molecular ions dominate the lower regions. During the night, the D and E regions almost completely recombine. The thermosphere is only partially ionized; the neutral density is typically 2–3 orders of magnitude larger than the plasma density.

The change in the orientation of Earth's dipole magnetic field causes significantly different behavior from high to low latitudes. For example, the auroral region is connected via magnetic field lines to the magnetosphere's plasma sheet, a region of reconnection and particle acceleration. In this work, we focus on midlatitudes, which are connected to plasma populations trapped in the inner magnetosphere. We also show some examples from equatorial latitudes, which do not directly interact with the magnetosphere.

In addition to its intrinsic scientific value, investigations of the thermosphere-ionosphere system have societal importance. This region of the atmosphere contains satellites in low-Earth orbit, and its variability affects space-to-ground and over-the-horizon communication and navigation systems. The thermosphere-ionosphere system regulates energy transfer from the space to the atmosphere and vice versa, and it causes variations in satellite drag. As society becomes more reliant on space-based assets,

it becomes more important to understand the medium in which they operate. One of the main barriers to better predictions of the dynamics of the thermosphere-ionosphere system is an insufficient characterization of the dynamics of the neutral fluid and how it interacts with the colocated plasma fluid [e.g., *Schunk and Sojka*, 1996; *Colerico et al.*, 2006; *Fuller-Rowell*, 2011]. This interaction is complex, as the presence of plasma adds Maxwell's equations to the fluid equations that must be considered. Useful introductions to the physics of thermosphere-ionosphere system are given by *Rishbeth and Garriott* [1969], *Rees* [1989], *Baumjohann and Treumann* [1997], *Prölss* [2004], and *Schunk and Nagy* [2009].

1.2 Thermospheric Wind

The wind is the mean velocity of particles within a small parcel of air. The thermospheric wind vector is typically defined as (u, v, w), where u is the zonal component (positive eastward), v is the meridional component (positive northward), and w is the vertical component (positive upward). In this work, when we use the term "thermospheric wind," we are referring to the wind at F-region altitudes ($\sim$200–500 km), since these are the altitudes where the wind directly interacts with the densest plasma reservoir. Lower thermospheric winds are also important, as they span the E-region, where the conductivity is largest, and thus the electrodynamic forcing is greatest, at least during the day. However, much less is known about lower thermospheric winds, and they are outside the scope of this work.

Counterintuitively, a consequence of the low neutral density in the upper thermosphere is that the viscosity is high. The large mean free path between collisions means that small-scale shears cannot be maintained, since the particles creating the shear do not remain localized. In terms of the equation of motion, the viscosity term becomes dominant in the upper thermosphere, which requires the altitude profile of the horizontal wind to be constant above a certain altitude [*Rishbeth*, 1972]. In this work, we assume the horizontal wind is constant over the altitudes spanned by the 630.0-nm emission (described below), roughly 200–300 km. This assumption is not always justified, but due to observational constraints, there is no alternative.

This caveat must be kept in mind when interpreting thermospheric wind measurements from ground-based interferometers.

The basic circulation in the thermosphere has been understood for many years, and the descriptions in the reviews by *Rishbeth* [1972], *Roble* [1983], and *Killeen* [1987] are still remarkably valid today. The primary driver of the wind is the pressure gradient caused by in situ solar heating, which produces a predominant flow from day to night and from the summer hemisphere to the winter hemisphere. However, there are numerous other drivers which superimpose temporal and spatial variability onto this general circulation, and this variability in turn affects the ionosphere. Through collisions with ions (and, to a much lesser degree, electrons), the wind directly forces the ionosphere, and it also drives currents and electric fields [e.g., *Heelis*, 2004]. The day-to-day variability of the thermospheric wind is thought to contribute to the day-to-day variability of ionospheric instabilities that can adversely affect trans-ionospheric radio wave propagation [*Kudeki et al.*, 2007]. The thermospheric wind has a considerable impact on the Joule heating rate, one of the major processes by which energy is transferred from Earth's magnetosphere to its atmosphere [e.g., *Aruliah et al.*, 2004]. During geomagnetic storms, discussed in Section 1.4, the thermospheric wind can act as a conduit for the redistribution of energy and composition from the auroral region to the global thermosphere-ionosphere system [*Fuller-Rowell*, 2011]. The wind is a key component of large-scale tidal waves and small-scale gravity waves, which carry energy from the lower atmosphere to the upper atmosphere and impose the variability of tropospheric weather on the thermosphere-ionosphere system [e.g., *England*, 2011; *Liu et al.*, 2014].

Given this spatiotemporal variability, observations are paramount. However, measuring the thermospheric wind is difficult, as the altitudes are too high for in situ measurements by aircraft or balloons. In situ satellite measurements can be made by tracking variations in the orbital drag [e.g., *Häusler and Lühr*, 2009; *Xiong et al.*, 2015], but these are expensive, and it can be difficult to distinguish spatial and temporal variations on a moving platform. Moreover, such observations are usually limited to high altitudes; as the orbital altitude drops to F-peak altitudes and below, atmospheric drag causes re-entry within weeks. In situ measurements are also possible from sounding rockets [e.g., *Larsen*, 2002], but these are also expensive for the limited spatial and temporal information they provide. Because of these lim-

itations, observational studies of the thermosphere rely heavily on remote sensing. Incoherent scatter radar remote sensing techniques have been used to estimate the wind [*Salah and Holt*, 1974; *Hysell et al.*, 2014]; however, these measurements are limited to several locations near incoherent scatter radars, and they are indirect, based on plasma measurements.

The only direct remote sensing technique for thermospheric wind, and the one used in this work, utilizes the 630.0-nm airglow emission. Airglow is the emission of light caused by natural chemical reactions in the atmosphere. The 630.0-nm emission usually spans ∼200–300 km in altitude, with a peak emission around 250 km, though it is variable. The presence of a wind causes a Doppler shift of the emission which can be measured from the ground or from space. The chemical processes leading to the 630.0-nm emission and its spectrum are described in Section 1.3.

Fabry-Perot interferometers (FPIs), or other interferometric instruments such as Michelson interferometers, are used to measure the spectrum of the 630.0-nm emission, and thus estimate the wind and temperature. Figure 1.2 depicts the geometry of such an observation using a ground-based FPI with a narrow field of view. The Doppler shift of the observed spectrum is a proxy for the line-of-sight component of the thermospheric wind, under the assumption that it is constant along the path where the line of sight intersects the emission region. Additionally, the line width of the emission provides information on the thermospheric temperature, which is important given its relationship to pressure gradients, scale heights, composition, and chemical reaction rates. Space-based observations using the 630.0-nm emission are also common [e.g., *Killeen and Roble*, 1988; *Shepherd et al.*, 2012]. FPIs are described in detail in Chapter 2.

In this work, we are particularly interested in the vertical component of the wind. Although often much smaller than horizontal winds, vertical winds have a larger impact on the dynamics, chemistry, and electrodynamics of the thermosphere-ionosphere system since vertical gradients of composition, density, and temperature are larger than horizontal gradients, so even small vertical winds cause large advective transport and adiabatic temperature changes.

Over 24 hours, there exists a small ∼1 m/s vertical wind associated with the rise and fall of constant-pressure levels as the thermosphere heats and cools. This is the so-called "barometric" or "breathing" component of the

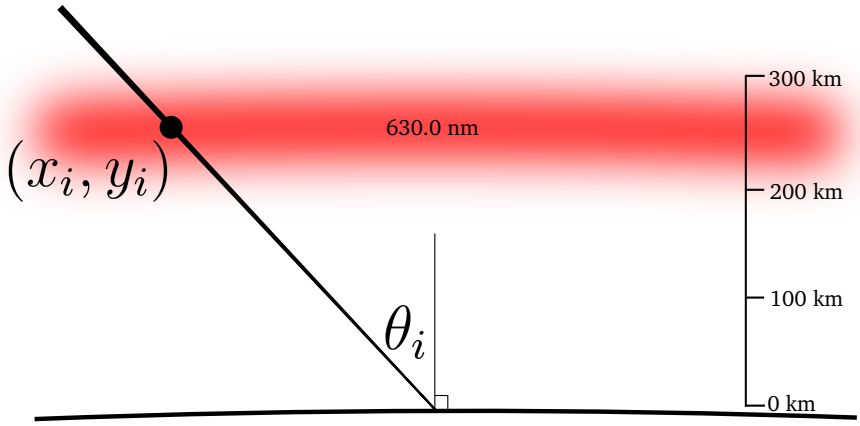

Figure 1.2: The geometry of a ground-based FPI observation of the 630.0-nm airglow layer, shown to scale. The zenith angle is θ_i, and the measurement location is (x_i, y_i), approximated to be the peak altitude of the emission. The azimuth angle, ϕ_i, is not shown.

vertical wind [*Smith*, 1998]. Larger vertical winds can exist via the "divergent" component. These vertical winds drive, or are driven by, divergent horizontal wind flows. Under the assumption of hydrostatic equilibrium and incompressible flow, the relationship between the vertical wind and the divergence of the horizontal wind is linear, where the constant of proportionality is the neutral scale height, H. For the thermosphere, this has come to be known as the Burnside relation [*Burnside et al.*, 1981]:

$$w = H\left(\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y}\right). \tag{1.5}$$

Under periods of strong, rapid, and/or small-scale forcing, the assumption of hydrostatic equilibrium can break down, allowing for even larger vertical winds. The Global Ionosphere Thermosphere Model (GITM) [*Deng et al.*, 2008; *Yigit and Ridley*, 2011] was the first to include non-hydrostatic effects. The *Deng et al.* [2008] study suggested that vertical winds can reach 150 m/s at an altitude of 300 km, though this effect lasts only several minutes. However, 50 m/s vertical winds lasted for an hour in their simulation.

For many years, the vertical winds measured by FPIs have been surprisingly large and difficult to reconcile with theory. In the polar cap, vertical winds of a few tens of m/s are common, and large upwellings of 150 m/s have been seen [*Rees et al.*, 1984; *Smith and Hernandez*, 1995; *Guo and McEwen*, 2003; *Ronksley*, 2016]. The majority of studies reporting vertical winds have

been conducted in the auroral region where vertical winds of up to 50 m/s are common [*Wardill and Jacka*, 1986; *Crickmore et al.*, 1991; *Crickmore*, 1993; *Conde and Dyson*, 1995; *Aruliah and Rees*, 1995; *Ishii et al.*, 1999, 2001; *Greet et al.*, 2002; *Aruliah et al.*, 2005; *Anderson et al.*, 2012a; *Ronksley*, 2016]. Many studies have found that an unreasonable scale height is needed to explain these vertical winds in terms of the Burnside relation [e.g., *Anderson et al.*, 2011]. Larger vertical winds of 100 m/s or more have also been reported [*Rees et al.*, 1984; *Price et al.*, 1995; *Innis et al.*, 1996, 1997, 1999; *Anderson et al.*, 2011]. The general trend emerging from these studies is an upward wind just poleward of the aurora and a downward wind equatorward of it. The observational evidence from midlatitudes is more sparse. Quiet-time vertical winds of 20–40 m/s have been seen by *Hernandez* [1982] and *Biondi* [1984], and 50–150 m/s downward winds were reported by *Sipler et al.* [1995] during a couple geomagnetic storms. At low and equatorial latitudes, vertical winds are much smaller (usually within measurement error), but some studies show statistically significant winds of 20 m/s or more [*Biondi and Sipler*, 1985; *Fisher et al.*, 2015].

Measurements of vertical winds using other techniques have corroborated the existence of significant vertical winds but are generally more conservative. The first direct measurement of the thermospheric vertical wind was by *Rieger* [1974], who used a series of barium releases from sounding rocket launches at various latitudes. They found that vertical winds were generally small ($\sim$15 m/s) but one data point in the auroral region reached 42 m/s. A more recent rocket experiment by *Wescott et al.* [2006] in the auroral region failed to find vertical winds larger than 20 m/s.

Space-based in situ mass spectrometers have also provided a useful independent vertical wind data source, but the measurements are sensitive to knowledge of spacecraft orientation and velocity. Atmospheric Explorer-C measured vertical winds at all latitudes of a few m/s, but at mid and high latitudes, waves with a peak-to-peak amplitude of up to 100 m/s were sometimes observed [*Spencer et al.*, 1976]. Initial results from Dynamics Explorer-2 were similar, with 10–20 m/s winds at lower latitudes and larger winds (up to 175 m/s in one case) in the auroral region [*Spencer et al.*, 1982]. A long-term statistical study of the Dynamics Explorer-2 data by *Innis and Conde* [2002] revealed that vertical wind variability was less than 30 m/s for 99% of all the midlatitude samples and 68% of the polar cap samples taken when

the AE index was high. *Raghavarao et al.* [1993] found examples of 10–40 m/s downwelling at equatorial latitudes. Reviews of thermospheric vertical wind measurements are given by *Smith* [1998] and *Larsen and Meriwether* [2012].

1.3 The 630.0-nm Airglow Emission

The nighttime 630.0-nm emission is a representation of the slow release of energy built up by solar photoionization during the day. This emission arises when oxygen atoms in the excited $^1D_2$ state (often simplified to ^1D) relax to the ground $^3P_2$ state:

$$O\left(^1D_2\right) \rightarrow O\left(^3P_2\right) + h\nu_{\lambda=630.0\ \mathrm{nm}}. \tag{1.6}$$

A more accurate emission wavelength is 630.0304 nm [*Osterbrock et al.*, 1996]; here we simply use 630.0 nm. Transition to the $^3P_1$ state produces an emission at 636.4 nm which is about three times dimmer than 630.0 nm, and the transition to the $^3P_0$ state produces an extremely dim emission at 639.2 nm [*Rishbeth and Garriott*, 1969].

$O(^1D)$ atoms can be created by a number of processes. At night, in the absence of auroral processes, the dominant source is dissociative recombination of O_2^+ ions:

$$O_2^+ + e \rightarrow \begin{cases} O\left(^3P\right) + O\left(^3P\right) \\ O\left(^3P\right) + O\left(^1D\right) \\ O\left(^1D\right) + O\left(^1S\right) \end{cases} \tag{1.7}$$

where the O_2^+ ions are produced from a charge exchange reaction between O_2, a minor species in the thermosphere, and O^+, the dominant ion in the F-region ionosphere:

$$O^+ + O_2 \rightarrow O + O_2^+. \tag{1.8}$$

A detailed model for the 630.0-nm airglow intensity caused by dissociative recombination of O_2^+ is given by *Link and Cogger* [1988, 1989].

Other important sources of $O(^1D)$ include impact on O by electrons that were energized by auroral electron and proton precipitation, impact on O by electrons in the high-energy tail of the thermal distribution, and relaxation

from the ^{1}S state. Comprehensive reviews of the nighttime 630.0-nm emission are given by *Torr and Torr* [1982] and *Solomon et al.* [1988].

The dayglow is less studied because ground-based observations are contaminated by the solar spectrum. However, space-based observations exist [e.g., *Zhang and Shepherd*, 2005]. The processes leading to the 630.0-nm dayglow are reviewed by *Solomon and Abreu* [1989] and *Thirupathaiah and Singh* [2014]. As difficult as it is to measure the 630.0-nm dayglow intensity, it is more difficult still to resolve the spectrum. Thus, in this work, we focus on winds and temperatures from nightglow measurements. However, recent advances in instrumentation and analysis are a promising step toward routine daytime wind observations [*Gerrard and Meriwether*, 2011].

The Einstein coefficient for (1.6) is (110 sec)$^{-1}$, which means that, once O(^{1}D) is created (by any process), the time before it emits a photon follows an exponential distribution with a mean of 110 sec. In this work, and in the aeronomy literature at large, a common approximation is that this is much longer than the O-O (and, at lower altitudes, the O-N_2) collision frequency, so a newly created O(^{1}D) is able to thermalize with the ambient neutral population before emitting a photon. Under this approximation, the spectrum of the O(^{1}D) emission is a proxy for the velocity distribution of the ambient neutral population. This allows us to use the Doppler shift of the emission to estimate the wind. We can also use the line width to estimate the temperature. Figure 1.3 shows the spectrum of the 630.0-nm emission for various combinations of line-of-sight wind and temperature.

The thermalization approximation has been investigated by numerous researchers with regard to the temperature measurement. The potential problem is that dissociative recombination is exothermic, with an excess energy of 7 eV, much of which goes to kinetic energy of the two resultant oxygen atoms. Since O(^{1}D) has such a large initial velocity, it may take many collisions to thermalize. *Yee* [1988] used a two-population model to explain the discrepancy between the 630.0-nm temperatures measured by the FPI on the Dynamics Explorer-2 spacecraft and the MSIS-83 model [*Hedin*, 1983]. In the model, he found a 90 K difference between the thermospheric temperature and the effective temperature indicated by the 630.0-nm line width. A Monte Carlo model developed by *Shematovich et al.* [1999] also found such a discrepancy, especially at high altitudes. *Hubert et al.* [2001] compared this model with re-analyzed Dynamics Explorer-2 data, finding that the discrep-

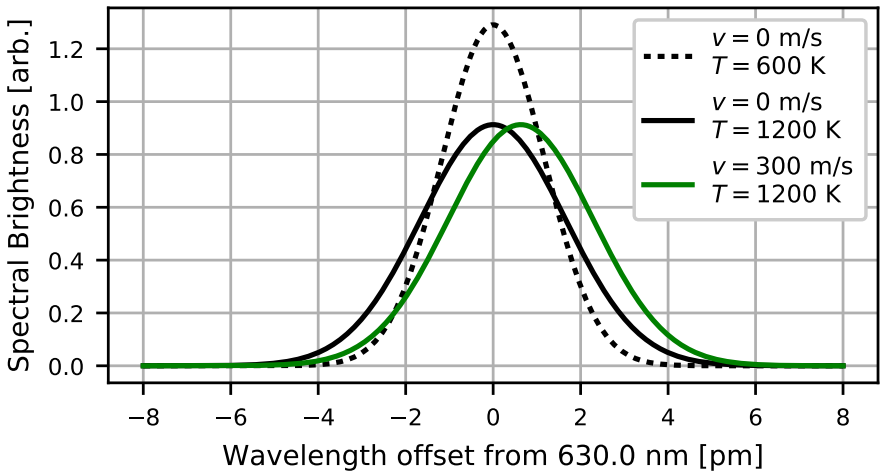

Figure 1.3: Example 630.0-nm spectra for various combinations of line-of-sight wind, v, and temperature, T. v is defined as positive away from the observer. The intensity is constant.

ancies of 100–150 K were likely explained by incomplete thermalization of O(^{1}D). *Kharchenko et al.* [2005] directly solved the Boltzmann equation in the lower thermosphere (80–200 km). Even though they find the thermalization time of O(^{1}D) at 200 km to be 1.1 sec (100 times faster than the emission rate), the fraction of "hot" (i.e., non-thermalized) O(^{1}D) is 5% of the ambient O population. *Sipler and Biondi* [2003] used the *Shematovich et al.* [1999] and *Hubert et al.* [2001] results to estimate that the effect on ground-based FPI temperature measurements is only 17 K.

To our knowledge, there has been no quantification of the effect on wind measurements, which could potentially be significant if (1) there is a large difference in the ion and neutral velocities, (2) O(^{1}D) is sourced by more energetic processes than dissociative recombination, (3) the 630.0-nm emission occurs at high altitude (e.g., in SAR arcs or when the F-region is elevated), or (4) a space-based limb scan is made at high tangent altitude. We also note that wind and temperature estimates using the 557.7-nm line from the ^{1}S–^{1}D transition (∼90-105 km at night) could also be affected. Although the collision frequency is much higher in the lower atmosphere, the radiative lifetime of the ^{1}S state is commensurably lower.

1.4 Geomagnetic Storms

Most of the data shown in this work were acquired during geomagnetic storms, periods of strong and rapid forcing which are often used to test theories and validate models. A geomagnetic storm is a depression in the strength of Earth's magnetic field, a largely historical definition. Today, we know that the magnetic field variations are but one effect of a complex chain of energy transfer from the sun to the upper atmosphere.

The energy driving geomagnetic storms ultimately comes from the sun. Coronal mass ejections or shocks in the solar wind impact and transfer energy into Earth's magnetosphere. The ionosphere-thermosphere receives part of this energy via Joule heating (electric fields of magnetospheric origin) and particle precipitation (the loss of energetic ring current ions and electrons to the atmosphere). Over the course of several days, this energy is re-radiated back into space, mostly in the form of infrared radiation by CO_2 and NO.

The thermospheric heating by magnetospheric inputs in the auroral regions leads to neutral upwelling, a pressure bulge at the poles, and equatorward winds. This new circulation pattern is communicated to mid and low latitudes via a large-scale gravity wave (sometimes called a traveling atmospheric disturbance), which arrives at midlatitudes within a couple hours of storm onset, causing disturbances in the dynamo electric field, equilibrium ionospheric height, and chemical recombination rates, among others. All of these have drastic effects on the ionosphere with a complex dependence upon local time, season, and solar cycle.

Useful introductory reviews of the thermospheric and ionospheric processes associated with geomagnetic storms are given by *Buonsanto* [1999] and *Mendillo* [2006]. The thermospheric wind is a key component of the response of the global thermosphere-ionosphere storm response [*Rishbeth et al.*, 1987; *Rishbeth*, 1991; *Hocke and Schlegel*, 1996; *Prölss*, 1997; *Fuller-Rowell and Codrescu*, 1997; *Fuller-Rowell*, 2011]. Much of the observational evidence that has led to this understanding has come from studies of the 630.0-nm emission by FPIs or other interferometric instruments, as reviewed by *Meriwether* [2008].

1.5 Organization of the Dissertation

The remainder of this dissertation is organized approximately chronologically, following our use of networked FPIs to study the thermospheric wind, with focus on the midlatitude storm response and interpretation of large vertical wind measurements. We cover instrumentation, data analysis, interpretation, and physical modeling. Section 2.2, Chapter 3, and Chapter 4 are based upon previously published work [*Harding et al.*, 2014, 2015; *Makela et al.*, 2014], while the work in Chapter 5 has been submitted for publication [*Harding et al.*, Submitted].

In Chapter 2, we describe instrumentation. In addition to general FPI fundamentals, we outline the design of the eight FPIs used in this work and their deployment to midlatitudes and equatorial latitudes. We also discuss operations. In the second half of Chapter 2, we describe data analysis. Specifically, we present the algorithm which estimates the line-of-sight wind and temperature from raw FPI data. Chapter 3 describes the higher-level algorithm which takes the line-of-sight wind data from many FPIs and generates a map of the regional wind field using statistical estimation techniques. This technique is used in Chapter 4 in a case study of the 02 Oct 2013 geomagnetic storm. A surprising feature in the FPI data from this storm is the appearance of a large 100 m/s apparent downward vertical wind lasting for several hours. We also summarize data from 15 storms, showing that these vertical winds are not uncommon. However, they are much too large to be believable, and we conclude they must be an artifact. In Chapter 5, we develop a radiative transfer model which demonstrates that these large vertical wind measurements can be explained as an artifact of the scattering of airglow radiation in the lower atmosphere, obviating the previous explanation of O^+ precipitation. We also apply our model to equatorial latitudes, showing that the vertical winds reported by *Fisher et al.* [2015] may be an artifact of scattering as well.

CHAPTER 2

INSTRUMENTATION AND DATA REDUCTION

In this chapter, we begin in Section 2.1 by describing the basics of the Fabry-Perot interferometer (FPI), including instrumentation, operation, and distribution. In Section 2.2, we describe the algorithm that we developed to analyze the raw data and generate estimates of the line-of-sight wind and temperature in the thermosphere. Section 2.2 is based off of the work published by *Harding et al.* [2014].

2.1 The Fabry-Perot Interferometer

2.1.1 Theory

A schematic representation of an FPI is shown in Figure 2.1. The most important component of the FPI is the etalon, two parallel plates of partially reflective glass separated by a distance t. Light incident upon the etalon from above will undergo multiple reflections between the plates. At each reflection, some light is transmitted. The transmitted rays have different phase shifts due to the different distances traveled, so they interfere either constructively, destructively, or somewhere in-between, depending on the incidence angle. When light is incident from a cone of angles, this creates the circular interference pattern characteristic of FPI data. The lens projects this image onto the charge-coupled device (CCD), where it is recorded.

The interference pattern caused by monochromatic light is described by an Airy function:

$$A(\theta, \lambda) = \frac{I}{1 + \frac{4R}{(1-R)^2} \sin^2\left(\frac{2\pi n t}{\lambda} \cos\theta\right)} \tag{2.1}$$

where

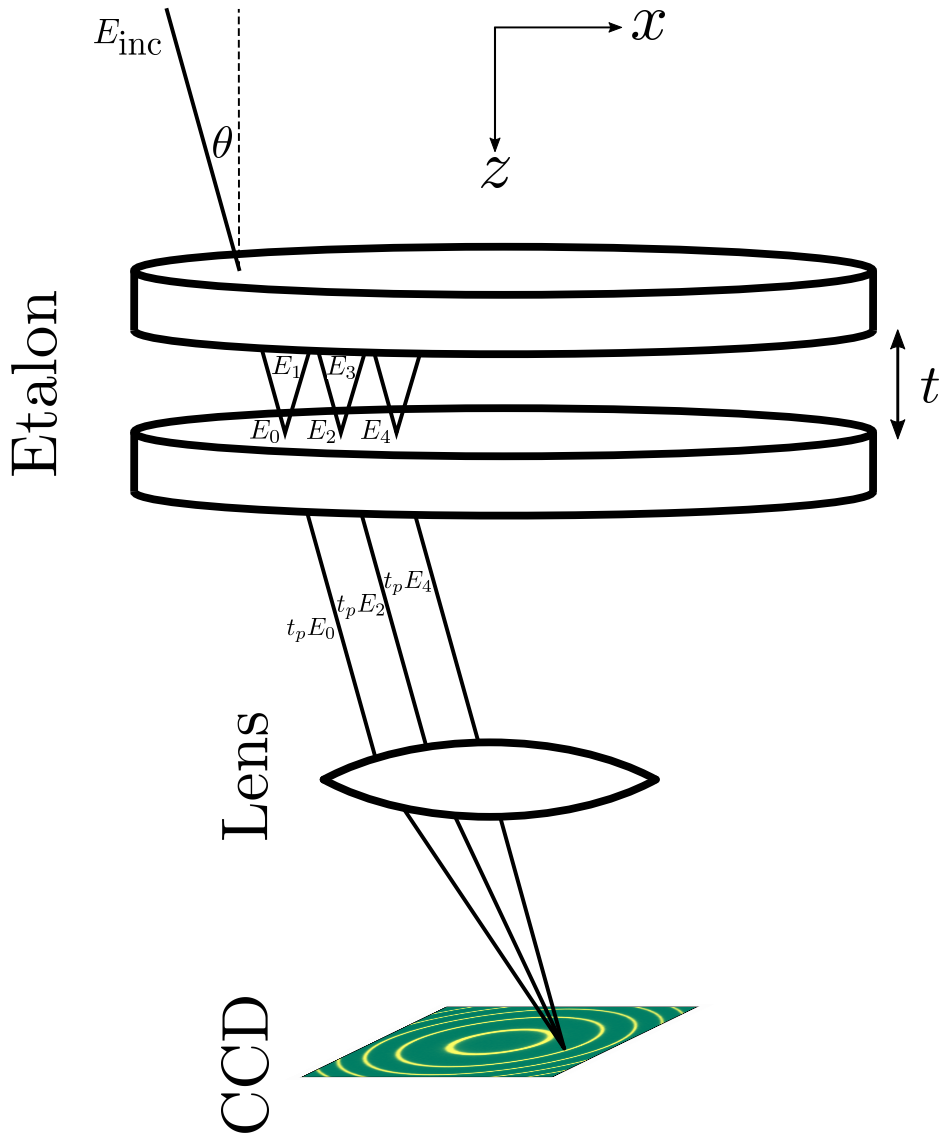

Figure 2.1: Schematic representation of light propagating through a Fabry-Perot interferometer. Not to scale.

I is the intensity of incident light,

R is the reflectivity of the etalon plates,

n is the index of refraction between the plates,

t is the distance between the plates,

λ is the free-space wavelength of incident light, and

θ is the incidence angle.

It is difficult to find the derivation of the Airy function in the literature, so we provide it here. As shown in Figure 2.1, let z denote the optical axis (positive downward), and let x denote a direction parallel to the etalon plates. We do not consider the third dimension due to cylindrical symmetry. The wavenumber, k, of the incident monochromatic light is given by

$$k = \frac{2\pi n}{\lambda}. \tag{2.2}$$

We first consider the incident light as an electromagnetic plane wave and then transition to a geometrical optics approximation. The incident electric field can be written in phasor form as

$$E_{\text{inc}} e^{-j(kz\cos\theta + kx\sin\theta)} \tag{2.3}$$

where E_{inc} is a scalar field strength. The phase term varies with z and x. The z-variation will contribute to the interference of the various reflected rays to create the Airy pattern. The x-variation will contribute to the location on the CCD to which the rays will be focused. While we could carry all the x-related terms through the derivation, they would be canceled by the phase variation introduced by the lens. From a geometrical optics point of view, this is identical to saying that the lens will focus all light rays incident from a given angle onto one point on the CCD [*Blahut*, 2004]. Thus, we drop the x-variations in the phase term and write the incident field as

$$E_{\text{inc}} e^{-jkz\cos\theta}. \tag{2.4}$$

From the phase term, we see that the phase accumulation from one traversal of the gap, which we denote ϕ, is

$$\begin{aligned} \phi &= kt\cos\theta \\ &= \frac{2\pi nt}{\lambda}\cos\theta. \end{aligned} \tag{2.5}$$

We define the transmission coefficient of an etalon plate as t_p and the reflection coefficient as r_p, so that the transmissivity and reflectivity are given by

$$\begin{aligned} T &= t_p^2 \\ R &= r_p^2. \end{aligned} \tag{2.6}$$

It is assumed that the two plates are identical.

Now taking a geometrical optics approach, we track the phase of the wave after each reflection of the light ray, as shown in Figure 2.1. After the incident ray is transmitted through the first plate, and just before interacting with the second plate, we can write the electric field as

$$E_0 = t_p E_{\text{inc}} \tag{2.7}$$

where we have arbitrarily chosen $z = 0$ as the top of the bottom plate so that the phase is zero. This ray is partially reflected and partially transmitted. The transmitted ray is simply $t_p E_0$ (ignoring the common phase associated with the path from the plate to the CCD). The reflected ray, evaluated just before interacting with the top plate again, is

$$\begin{aligned} E_1 &= r_p \, E_0 \, e^{-j\phi} \\ &= r_p \, t_p E_{\text{inc}} \, e^{-j\phi}. \end{aligned} \tag{2.8}$$

Likewise, the subsequent reflected rays are

$$\begin{aligned} E_2 &= r_p^2 \, t_p E_{\text{inc}} \, e^{-j2\phi} \\ E_3 &= r_p^3 \, t_p E_{\text{inc}} \, e^{-j3\phi} \\ E_4 &= r_p^4 \, t_p E_{\text{inc}} \, e^{-j4\phi} \\ &\vdots \end{aligned} \tag{2.9}$$

The interference pattern arises from the coherent sum of all the transmitted rays:

$$\begin{aligned} E_{\text{tot}} &= t_p E_0 + t_p E_2 + t_p E_4 + \dots \\ &= t_p^2 E_{\text{inc}} \left(1 + r_p^2 e^{-j2\phi} + r_p^4 e^{-j4\phi} + \dots\right) \\ &= T E_{\text{inc}} \left(1 + R e^{-j2\phi} + R^2 e^{-j4\phi} + \dots\right) \\ &= \frac{T E_{\text{inc}}}{1 - R e^{-j2\phi}} \end{aligned} \tag{2.10}$$

where in the last step, the Taylor series expansion for $\frac{1}{1-x}$ was used. The intensity of the interference pattern is proportional to the squared magnitude

of the electric field:

$$\begin{aligned}
|E_{\text{tot}}|^2 &= \frac{T^2\,|E_{\text{inc}}|^2}{(1-Re^{-j2\phi})(1-Re^{j2\phi})} \\
&= \frac{T^2\,|E_{\text{inc}}|^2}{1+R^2-2R\cos 2\phi} \\
&= \frac{T^2\,|E_{\text{inc}}|^2}{1+R^2-2R(1-2\sin^2\phi)} \\
&= \frac{T^2\,|E_{\text{inc}}|^2}{(1-R)^2+4R\sin^2\phi} \\
&= \frac{|E_{\text{inc}}|^2}{1+\frac{4R}{(1-R)^2}\sin^2\phi} \\
&= \frac{|E_{\text{inc}}|^2}{1+\frac{4R}{(1-R)^2}\sin^2\left(\frac{2\pi nt}{\lambda}\cos\theta\right)}
\end{aligned} \tag{2.11}$$

where in the second-to-last line, we used the fact that $R+T=1$ by energy conservation, and in the last line, we used (2.5). If we define $|E_{\text{inc}}|^2$ as I, the source intensity (ignoring constants), we arrive at (2.1), the Airy function.

The most important term in the Airy function is the phase term, $\frac{2\pi nt}{\lambda}\cos\theta$. A maximum in the Airy function occurs where this term is a multiple of π. The change in λ required to change the phase term by π is known as the free spectral range, though this term is often used in the looser sense, also applying to changes in t, θ, or some transformations thereof. Since the gap is much larger than a wavelength, the phase term is quite large, for near-vertical incidence. As a result, any small relative perturbation to n, t, λ, or θ will result in a large change in the phase term (relative to π). This allows for highly accurate monitoring of the tiny Doppler shifts associated with the thermospheric wind. However, in order to find the maximum of the Airy function, we have to measure at least one period. This can be done by sweeping over n (e.g., changing the pressure of the gas between the plates), by sweeping t (e.g., using a piezoelectric spacer material), or, as is done by most modern FPIs in the visible regime, by measuring many different θ simultaneously using a CCD. The systems used here are designed to gather several periods in order to improve the signal-to-noise ratio. One peak of the Airy function, as seen by sweeping over one of these independent variables, is referred to as a fringe.

For analyzing data taken by a CCD, it is often more convenient to use r,

the fringe radius, defined as the distance of the CCD pixel from the optical axis, instead of θ, as the independent variable. The mapping between these two variables for narrow-field optical systems is given simply by

$$\theta = \tan^{-1}(\alpha r) \tag{2.12}$$

where α is known as the magnification, and is given by the ratio of the pixel size to the focal length, assuming r is specified in units of pixels. For systems with more complicated lens assemblies or wider fields of view, a more complicated mapping between θ and r is needed [*Conde*, 2002]. From here on, we will refer to the Airy function as $A(r, \lambda)$ instead of $A(\theta, \lambda)$.

The Airy function describes the response to monochromatic light. Because Maxwell's equations are linear, and all the optical components have linear responses, superposition holds, and we can write the response to an arbitrary spectrum as a Fredholm integral of the first kind. If we let $Y(\lambda)$ describe the spectrum of light at the aperture of the etalon, then the recorded response, $S(r)$, usually specified in units of CCD counts, is given by:

$$S(r) = \int_0^\infty A(r, \lambda) Y(\lambda)\, d\lambda. \tag{2.13}$$

This form offers an intuitive interpretation of FPI data. The measurement can be thought of a weighted average of Airy functions, where the weights are determined by the source spectrum. It is common to see this referred to as a convolution in the literature, but since the shape of A can vary with r, we cannot make this approximation.

More details on the theory of Fabry-Perot interferometers are described by *Hernandez* [1988]. A description of how this theory is used to analyze FPI data is given in Section 2.2.

2.1.2 System Design

In addition to the etalon, lens, and CCD, the FPI systems include a SkyScanner, an interference filter, a laser, and various other supporting systems and sensors. The system design, known as MiniME, is documented in detail by *Makela et al.* [2011], *Makela et al.* [2012], and *Fisher* [2013], but the basic

outline is described here, as are relevant changes and lessons learned since 2013.

The SkyScanner, shown in Figure 2.2, consists of two mirrors which can be independently rotated to control the azimuth and zenith angles of the line of sight. The SkyScanner is usually mounted to the roof of a building or portable trailer, protected by a plastic dome. As to be expected from a component with moving parts, the SkyScanner has been the primary source of maintenance problems for the entire FPI system. Regular cleaning and greasing of the bearings and bearing tracks are required to reduce motor failures, the primary contributor to instrument downtime. A recommendation for an improved design is to eliminate the SkyScanner, perhaps using a more complex optical design to observe multiple directions simultaneously on different parts of the CCD.

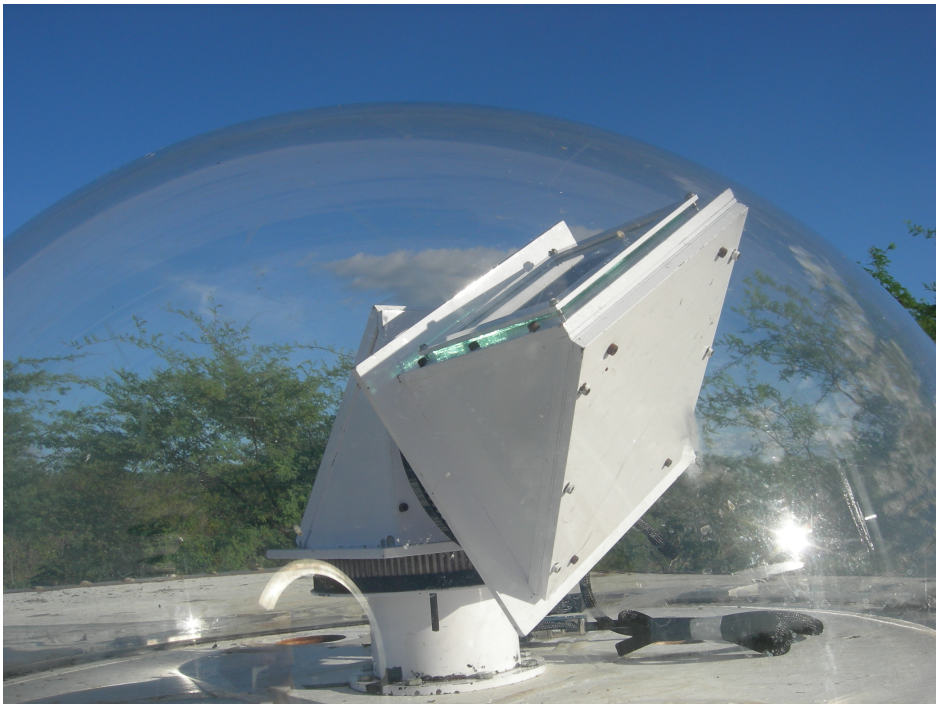

Figure 2.2: The SkyScanner, comprising two independently controlled mirrors to steer the line of sight. Adapted from *Fisher* [2013].

Light entering the aperture of the SkyScanner is redirected down through the roof, where the rest of the optical components are mounted on a frame suspended from the ceiling. This frame is shown in Figure 2.3. The first optical component is the interference filter. This filter rejects light with wavelengths outside of a narrow spectral window centered at 630.0 nm with a

∼1-nm full-width at half-maximum. This spectral window is narrow in order to limit the influence from the background continuum, which contributes to shot noise, and to suppress other atmospheric emission lines. However, there are two hydroxyl (OH) transitions that emit at wavelengths close enough to 630.0 nm to be of concern: one at 629.7903 nm and a doublet at ∼630.7 nm (630.6869 and 630.6981 nm) [*Osterbrock et al.*, 1996]. In Section 2.2.4, we discuss the details of how we handle the presence of these emission lines in our analysis algorithms. In the future, a narrower passband could be used to filter the nearby OH lines, provided an acceptable transmissivity at 630.0 nm can still be obtained at reasonable cost.

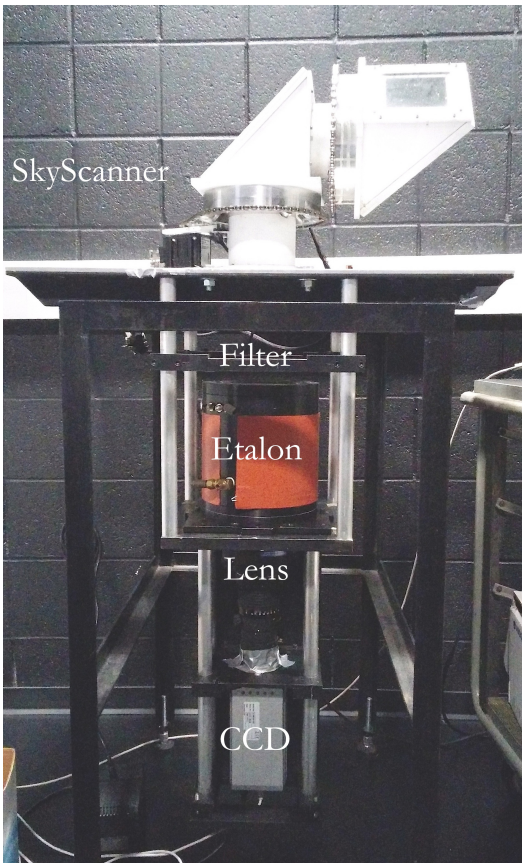

Figure 2.3: The MiniME FPI frame with all optical components included, except the laser and laser diffuser box. When installed, the system is hung from a roof and shrouded.

After passing through the interference filter, light enters the etalon, where, as described in Section 2.1.1, an interference pattern is created which is imaged onto the CCD by the objective lens. The etalon and the CCD are the two most expensive components of the FPI. The cost of the etalon is

proportional to its size, and therefore, its size usually determines the system aperture size. The driver of the cost of the CCD is the need for low read and dark noise. The CCD is equipped with a thermoelectric cooler, which minimizes dark noise by cooling the detector to as low as -80 °C.

The size of the CCD and the focal length of the lens set the field of view, which is 1.8° for the MiniME FPI system. This field of view intersects a region with a ∼10 km horizontal extent in the thermosphere, which is smaller than the scale sizes of interest in this dissertation, and also likely smaller than the minimum scale size permitted in the viscosity-dominated thermosphere. Thus, in this work we treat the FPI measurement as a point measurement. For other systems with larger fields of view, this assumption cannot be made [e.g., *Anderson et al.*, 2009].

Variations in the ambient temperature surrounding the etalon cause the spacer material between the plates to expand and contract. Although the spacer material was chosen to have a small coefficient of thermal expansion, even small variations in t can have large impacts on the estimated wind. For example, for MiniME (t = 15 mm), if the gap changes by 1 nm, and this change is not corrected, it will manifest as a 20 m/s error in the estimated wind.

In order to track these changes in t, in addition to possible changes in n or drifts in the relative positions of optical components, we employ a frequency-stabilized laser to act as a calibration source. Since the source spectrum of this laser, $Y(\lambda)$, is a delta function, the measurement represents $A(r, \lambda)$ at the wavelength of the laser. Combining this with the theoretical λ-dependence of A, we can fully characterize $A(r, \lambda)$. The algorithm is discussed in more detail in Section 2.2. We use a helium-neon laser at 632.8 nm, which is close enough to 630.0 nm to give us an accurate characterization of A. Even though this is outside the passband of the interference filter, the laser is bright enough to overcome the small transmissivity of the filter at 632.8 nm, resulting in a signal that is much stronger than airglow. A diffuser is used to fill the field of view with light from the laser. Throughout the night, the SkyScanner routinely rotates to observe the diffused laser. When the SkyScanner is observing the sky, a shutter mounted on the laser is closed to eliminate contamination by the laser.

Although the laser is advertised to be stable to within ± 1 MHz (equivalent to ± 0.6 m/s) over an 8-hour time period, on occasion we have had reason to

suspect that the laser was drifting. For example, some nights exhibit a 100 m/s linear trend in the estimated vertical wind over the night. Frequency-stabilized lasers are expensive (∼$4000), but are not a dominant cost, so for future deployments, it may be useful to employ two independent lasers for cross-calibration and redundancy.

Even though changes in the etalon gap caused by ambient temperature variations can be tracked, it is still desirable to minimize them. If the gap changes so much that the phase term varies by π (one period of the Airy function), an ambiguity can arise. Also, if the gap changes on time scales shorter than cadence of calibration images, the changes will not be adequately corrected for. Further, rapid heating or cooling can warp the etalon plates and distort the Airy function. Although we initially attempted to control the etalon temperature with a simple proportional-integral-differential (PID) control system using a temperature sensor and heating pad, we ultimately found that it caused a larger problem than it solved; the heating from the pad was too rapid. Now, we rely on the heating and air conditioning system of the room containing the instrument. In the future, it may be worth revisiting the environmental control of the etalon housing.

To monitor viewing conditions, we utilize a Boltwood cloud sensor mounted to the roof. The cloud sensor compares the infrared sky temperature with the ambient temperature. It measures the sky temperature in the 6- to 14-μm band over a 120° field of view using a thermopile. When the difference between the sky and ambient temperatures is large, the sky is clear, and when the difference is small, it is cloudy. Although Boltwood recommends 25 °C as a threshold, experience has shown that a smaller number is more appropriate. Also, the threshold seems to be change, possibly depending on temperature or humidity. For example, in the summer, the temperature difference rarely reaches 25 °C, even on clear nights. The appropriate threshold also seems to change from site to site. A more detailed study with an independent source of cloud cover data is required to resolve this issue.

2.1.3 Operations

All of the components and sensors described above are autonomously controlled and monitored by a local computer. The system initializes each night

when the solar zenith angle reaches 98°, about a half-hour after sunset on the ground. After taking dark and bias images, normal operations begin, which are described by *Makela et al.* [2012] and *Fisher* [2013]. In this work we are concerned with cardinal mode, not common-volume mode, observations. In common-volume mode, multiple FPIs work together to obtain simultaneous observations of the same 250-km altitude source region from different vantage points. With different line-of-sight projections of the same wind vector, the wind vector can be estimated. Common-volume mode provides two or three components of the wind at a single point, depending on whether the point is observed by two or three FPIs.

In cardinal mode, each FPI acts independently, cycling through a predetermined observation routine, consisting of consecutive measurements looking toward the zenith, north, south, east, and west. The zenith angle of the latter four measurements is 45°. Cardinal mode provides wider spatial coverage, but each sample only provides one of the three components of the wind vector. However, due to the high viscosity of the thermosphere, we know that the wind varies smoothly in latitude and longitude, so we can employ statistical estimation techniques such as that described in Chapter 3 to estimate the latitude- and longitude-dependence of all three components of the wind.

At each location in the measurement cycle, the SkyScanner pauses and the CCD begins an exposure. Although the exposure time was initially set to a constant 180 sec, in August 2013 we began automatically adjusting it based on the airglow brightness in order to achieve a constant signal-to-noise ratio. The desired signal-to-noise ratio corresponds to statistical uncertainties in the wind and temperature of about 5 m/s and 20 K. Just after sunset and just before sunrise, this can sometimes be achieved with exposure times as short as 20 sec. In the middle of the night, longer exposures are required, sometimes as long as 480 sec, which is the maximum we allow. Thus, the sample rate for a particular direction can vary from about 2 to 45 minutes.

Laser calibration exposures are interspersed throughout the measurement cycle. Initially, this occurred once per measurement cycle, but in mid-2014, we increased the cadence, taking laser calibration exposures before every sky exposure. This reduced the uncertainty associated with the long intervals between calibrations, which is important when the etalon gap is changing rapidly. Before making this change, we would often find that the calibration

uncertainty exceeded the statistical uncertainty, which meant it was worth a small sacrifice in the sample rate to achieve large improvements in total uncertainty.

Data are gathered until the solar zenith angle reaches 98° again, about a half-hour before sunrise. At this point, if the internet connection is stable, data are sent to the server at the University of Illinois, where they are permanently stored and analyzed using the algorithm described in Section 2.2. The resulting wind and temperature estimates are posted online (`http://airglow.ece.illinois.edu`) and sent to the CEDAR Madrigal Archive (`http://cedar.openmadrigal.org/`), where they can be downloaded by the public.

2.1.4 Distribution

In recent years there has been rapid growth in the distribution of FPIs around the world. In this work, we focus on data from two networks: the Remote Equatorial Nighttime Observatory for Ionospheric Regions (RENOIR) [*Fisher*, 2013; *Makela et al.*, 2013] and the North American Thermosphere Ionosphere Observing Network (NATION) [*Makela et al.*, 2012]. RENOIR includes two FPIs, denoted CAR and CAJ, located in Cajazeiras and Cariri, Brazil. NATION includes five FPIs, located at the Urbana Atmospheric Observatory in Illinois (UAO), the Pisgah Astronomical Research Institute in North Carolina (PAR), Eastern Kentucky University (EKU), Peach Mountain near Ann Arbor, Michigan (ANN), and at a site near Virginia Polytechnic Institute (VTI). In Chapter 5, we also utilize data from an FPI operated by the Millstone Hill Observatory, which we denote MH. Figure 2.4 shows the geographical locations of these eight FPIs, and Table 2.1 provides further details.

The MH FPI is operated independently by Computational Physics, Inc., and its data are analyzed with a different algorithm than the one described later in this dissertation. The optical design is slightly different than the NATION and RENOIR instruments. The etalon gap is 10.5 mm, and a telecentric configuration is used for the filter, which provides a more consistent filter response across the field of view but causes more flat-field artifacts

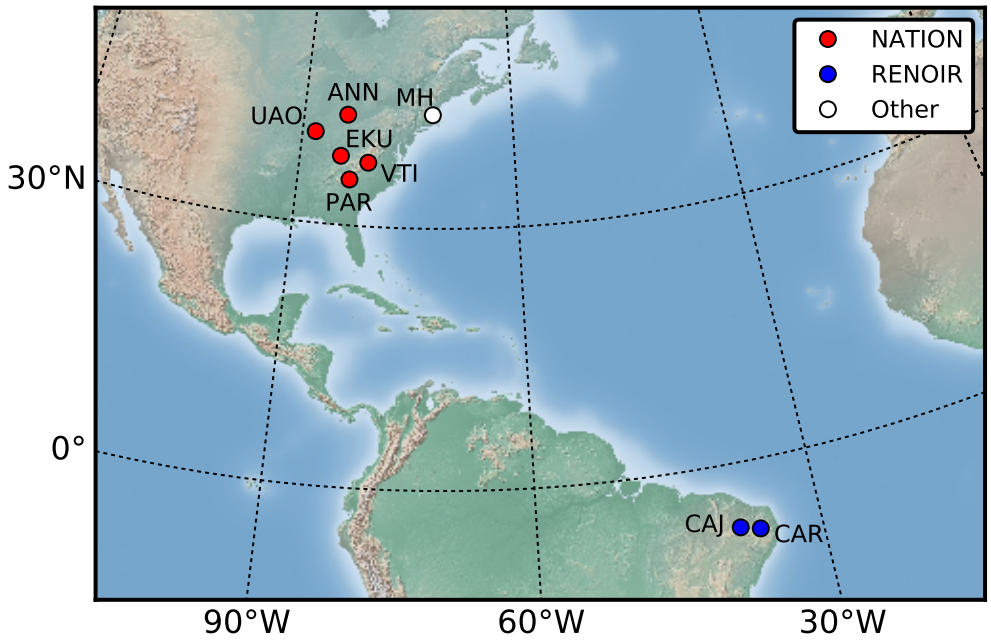

Figure 2.4: Locations, site codes, and network names for the FPIs used in this work.

Table 2.1: FPI Sites Used in This Work

Site Code	Network	Lat.[a]	Long.[a]	Data Availability
CAR	RENOIR	-7.38°	-36.52°	2010–Present
CAJ	RENOIR	-6.88°	-38.56°	2009–2014
UAO	NATION	40.17°	-88.16°	2012–Present
PAR	NATION	35.20°	-82.85°	2011–Present
EKU	NATION	37.75°	-84.29°	2012–2015
ANN	NATION	42.27°	-83.75°	2012–2015
VTI	NATION	37.21°	-80.42°	2013–2015
MH	Other	42.62°	-71.49°	2009–Present

[a]Latitude and longitude are given in geographic coordinates.

which need to be removed. Details of the MH instrument and processing are given by *Kerr* [2014].

2.2 Line-of-Sight Wind and Temperature Estimation

In this section, we describe the analysis technique which estimates the line-of-sight wind and temperature from a raw FPI image. This is largely based on

the work described in *Harding et al.* [2014], but more recent details are also provided, particularly with regard to quality control, Doppler referencing, and uncertainty estimation.

2.2.1 Background

Fundamentally, the analysis of FPI images involves tracking the fringe locations and widths, which provide the wind and temperature, respectively. The crucial challenge is to accurately correct for instrumental effects, namely the temporal drift of the fringes and the broadening caused by the Airy function. Usually this is achieved by making observations of a monochromatic, extended source at a known wavelength, such as that provided by a frequency-stabilized laser observed through a diffusing sphere.

Historically, many methods have been used. *Biondi et al.* [1995] used simple Gaussians to fit both the calibration and airglow fringes. Although any symmetric function works fairly well to estimate the fringe center, a Gaussian assumption produces large errors in the temperature estimate, because neither the calibration fringe shape (theoretically an Airy function) nor airglow fringe shape (theoretically a Gaussian blurred with an Airy function) are well-approximated by Gaussians. *Brum et al.* [2012] tried Gaussian, Lorentzian, and Airy functions to fit their data, and found that the temperature estimates can vary by 120 K. This systematic error is significant, about an order of magnitude larger than the statistical uncertainty for the instruments used here. *Brum et al.* [2012] suggest that FPI temperatures must be normalized to an independent reference such as incoherent scatter radar (using the assumption that the neutral temperature and the ion temperature are equal, which is likely satisfied at night). However, we believe that accurate temperature estimates can be achieved directly from the FPI data with a suitable algorithm.

Accurate temperature estimation is complicated by nonidealities in the optics. Practical etalons are not perfectly flat; rough surfaces will broaden and distort the Airy function. Other imperfections can also contribute to the deviation of the instrument function from a true Airy function: lens defects, vignetting, a finite aperture, and imperfect focusing, to name a few.

Hays and Roble [1971a] attempt to characterize many of these nonidealities analytically, but full characterization requires a data-driven approach.

One popular and elegant solution to this problem is to avoid making any particular assumptions about the shape of the fringes, and instead decompose the fringes into a Fourier series, keeping only the first dozen or so terms, as described by *Killeen and Hays* [1984]. The blurring of the sky spectrum by the instrument function in the Fourier domain can be performed analytically, simplifying the analysis, and by keeping only the low-order Fourier terms, noise is somewhat suppressed. However, *Makela et al.* [2011] show that at high thermospheric temperatures, the temperature estimate can be biased, which is likely due to the fact that the Fourier method requires the fringes to be isolated and analyzed independently, so errors can arise when the spectrum is broader than a single period of the Airy function, causing adjacent orders to overlap. This method also makes it difficult to accurately estimate uncertainty.

Unfortunately, the FPI community remains fractured with regard to analysis techniques. Some use Fourier-based techniques, some fit Gaussians or other functions, and some use a mixture of techniques: Gaussian fits for wind estimates and Fourier methods for temperature [e.g., *Shiokawa et al.*, 2001]. This growing list of often ill-documented techniques is a barrier to entry for students, a complication for the reproducibility of research, and a confounding factor for comparative studies. At the risk of adding to this list, here we propose a new technique that poses the problem as a model fit and accounts for optical defects. A significant benefit of this technique is that it analyzes all fringes at once, implicitly accounting for order overlap, and it provides accurate uncertainties. This algorithm has demonstrably better performance, is generally applicable, and is easily implementable as a nonlinear least-squares problem. In the following, we present our model for the calibration and airglow fringes, describe the inversion, and test it using Monte Carlo simulations. We also provide details on post-processing and quality control.

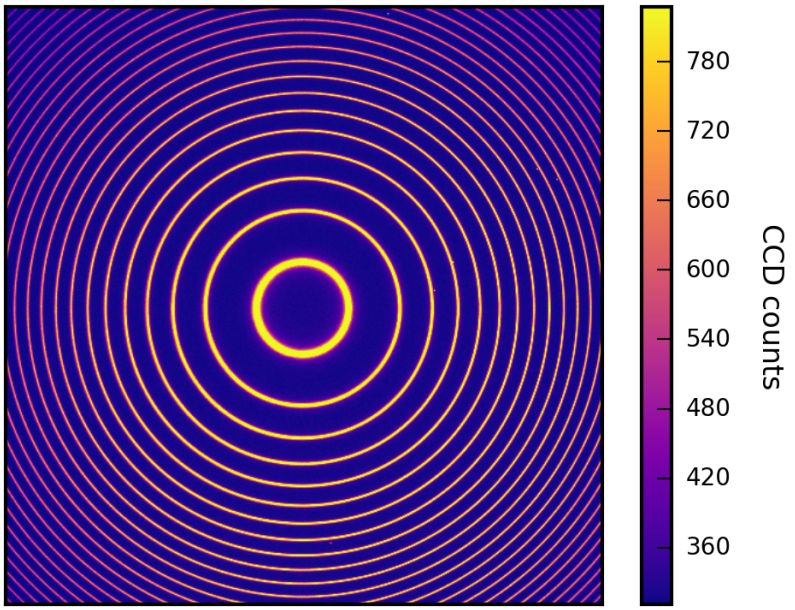

Figure 2.5: Example laser calibration image taken by the UAO FPI.

2.2.2 Model

An example of a laser calibration image is shown in Figure 2.5, and an example of an airglow image is shown in Figure 2.6. The units of these raw data are CCD counts per pixel, proportional to the photon flux on the pixel plus a constant offset, the CCD bias. In the following we describe the models used to fit these data.

Laser Fringe Model

As described in Section 2.1.1, the ideal shape of the instrument function is an Airy function. We use the laser calibration image to determine the fringe center and bin the data as a function of r, in order to convert the 2D image to a 1D fringe pattern. This process, commonly known as annular summation, is described in more detail in Section 2.2.3.

Figure 2.7 shows the 1D laser fringe pattern, compared to the best-fit Airy function. The data are shown as a function of r^2 so that the fringes appear equally spaced. Also shown is the residual, the difference between the data and the Airy model. It is clear that there are at least two important nonidealities causing deviations from the Airy model. First, as r increases,

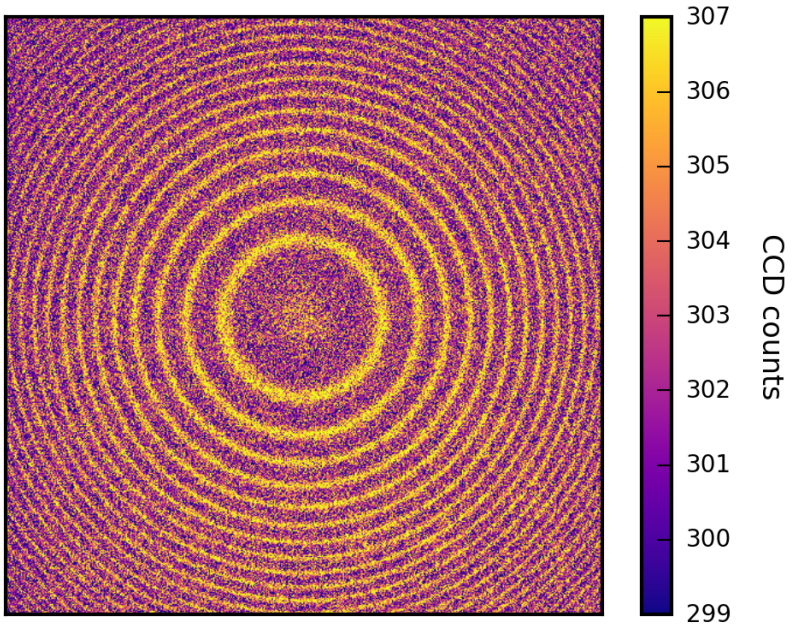

Figure 2.6: Example airglow image taken by the UAO FPI.

there is a decrease in optical throughput caused by vignetting and absorption. Second, the fringes appear to be shorter and wider than the Airy function, especially as r increases. This is likely due to the optical defects described above.

In order to account for these nonidealities, we add more parameters to the instrument function. First, we add parameters to describe the falloff in throughput, replacing the constant term I in (2.1) with a quadratic term in r:

$$I(r) = I_0 \left(1 + I_1 \left(\frac{r}{r_{\max}}\right) + I_2 \left(\frac{r}{r_{\max}}\right)^2\right) \tag{2.14}$$

where $r_{\max}$, a constant, is the maximum value of r used in the analysis. The new parameters, I_1 and I_2, control the shape of the throughput falloff with r. Although a higher-order fit could theoretically better fit the true falloff, we are often fitting to a small number of peaks (nine in Figure 2.7). Fitting a large polynomial to a small number of points can be unstable. We found a quadratic fit to be sufficient. A more sophisticated approach could be employed in the future.

Second, we add parameters which describe imperfect focusing and optical defects leading to broadening of the fringes. We introduce a point-spread

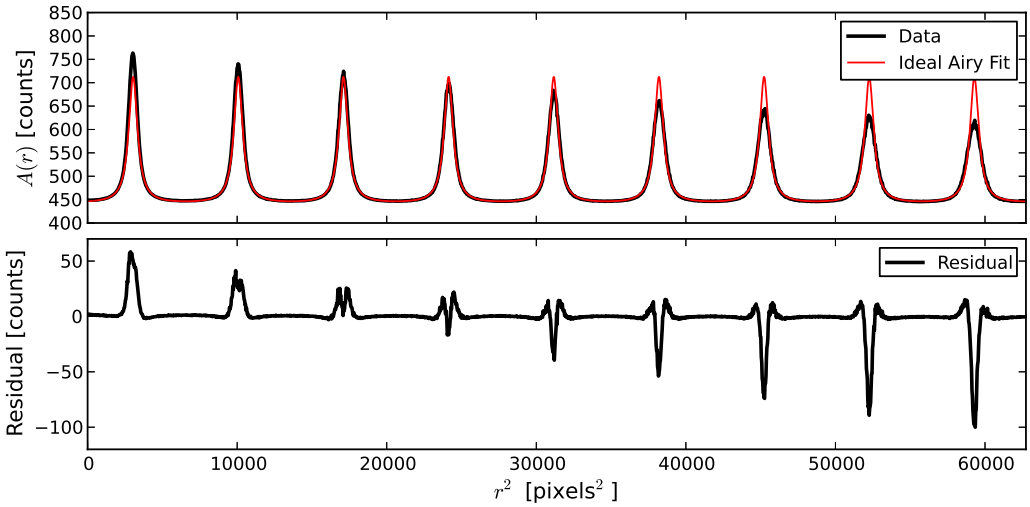

Figure 2.7: (Top) The laser fringe pattern resulting from annular summation of the raw laser image, compared to the best-fit ideal Airy model. (Bottom) Difference between data and model. Adapted from *Harding et al.* [2014].

function, $b(s, r)$, which blurs the Airy function via a Fredholm integral, to create a modified Airy function, denoted $\tilde{A}(r, \lambda)$:

$$\tilde{A}(r, \lambda) = \int_0^{r_{\max}} b(s, r)\, A(s, \lambda)\, ds. \tag{2.15}$$

The dummy variable s is in the same coordinates as r. Each pixel of the modified Airy function is a weighted average of the neighborhood pixels of the ideal Airy function, where the weights are determined by the shape of the point-spread function. In the case that the point-spread function $b(s, r)$ only depends on $s - r$, then this simplifies to a simple convolution. However, since in practice the blurring gets more severe with increasing r, we cannot make this assumption. We model the point-spread function as a normalized Gaussian centered on the pixel r:

$$b(s, r) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(s - r)^2}{\sigma^2}\right). \tag{2.16}$$

The only free parameter is σ, the width of the blurring kernel. If σ were constant, then the outer fringes would be blurred more than the inner fringes, since the outer fringes are closer together. While this matches the observed blurring behavior qualitatively, we found that we could obtain a better quan-

titative characterization of the instrument function by allowing the width of the blurring kernel to vary from the middle to the edge of the CCD, in order to mimic the variation in the path lengths through the optics for higher incidence angles. We re-write (2.16) to explicitly include the r-dependence of σ:

$$b(s,r) = \frac{1}{\sqrt{2\pi\sigma(r)^2}} \exp\left(-\frac{(s-r)^2}{\sigma(r)^2}\right) \tag{2.17}$$

where we use a first-order Fourier series to parameterize $\sigma(r)$:

$$\sigma(r) = \sigma_0 + \sigma_1 \sin\left(\pi \frac{r}{r_{\max}}\right) + \sigma_2 \cos\left(\pi \frac{r}{r_{\max}}\right). \tag{2.18}$$

The free parameters are σ_0, σ_1, and σ_2. We chose a Fourier series instead of a polynomial because in practice the fit was more stable, as Fourier coefficients are orthogonal, while polynomial coefficients are not. However, in practice σ_1 and σ_2 tend to be much smaller than σ_0, so other parameterizations of σ could also be suitable, provided the parameters can be reliably estimated from the laser calibration fringes.

The final modification has to do with a detail of the CCD. Because of dark current and intentional offsets in the CCD, the recorded CCD counts has additive terms, which we lump into one parameter, the CCD bias, B_{CCD}.

With these modifications, we update our observation model, (2.13):

$$S(r) = \int_0^\infty \tilde{A}(r,\lambda) Y(\lambda)\, d\lambda + B_{\text{CCD}}. \tag{2.19}$$

Putting everything together, (2.1), (2.12), (2.14), (2.15), (2.17), (2.18), and (2.19) describe a forward model for the laser calibration fringes, where the input spectrum $Y(\lambda)$ is taken to be a delta function at 632.8 nm, and the measurement is $S(r)$. For convenience, we simplify and collect the equations describing this model here:

$$
\begin{aligned}
S(r) &= \int_0^{r_{\max}} b(s,r)\, A(s,\lambda_{\text{laser}})\, ds \;\; + B_{\text{CCD}} \\
A(r,\lambda) &= I_0 \, \frac{1 + I_1\left(\frac{r}{r_{\max}}\right) + I_2\left(\frac{r}{r_{\max}}\right)^2}{1 + \frac{4R}{(1-R)^2}\sin^2\!\left(\frac{2\pi n t}{\lambda}\cos\tan^{-1}(\alpha r)\right)} \\
b(s,r) &= \frac{1}{\sqrt{2\pi\sigma(r)^2}} \exp\left(-\frac{(s-r)^2}{\sigma(r)^2}\right) \\
\sigma(r) &= \sigma_0 + \sigma_1 \sin\left(\pi\frac{r}{r_{\max}}\right) + \sigma_2 \cos\left(\pi\frac{r}{r_{\max}}\right).
\end{aligned}
\tag{2.20}
$$

Table 2.2 summarizes the parameters of this model, specifies which are fixed and which are varied to fit the data, and lists their typical values for the MiniME FPI system. In Section 2.2.3, we describe the details of the inversion which estimates the free parameters.

Table 2.2: Model Parameters for Laser Calibration Fringes

Param.	**Description**	**Fixed/Varied**	**Typical Value**
λ_{laser}	Laser Wavelength	Fixed	632.8 nm
n	Index of Refraction	Fixed	1
$r_{\max}$	Maximum Radius	Fixed	256 pixels
R	Etalon Reflectivity	Varied	0.8
t	Etalon Gap	Varied	15 mm
α	Magnification	Varied	8.5×10^{-5}
I_0	Intensity at Center	Varied	1000 counts
I_1	Linear Intensity Falloff	Varied	-0.1
I_2	Quadratic Intensity Falloff	Varied	0.005
σ_0	Average Blur Size	Varied	0.8 pixels
σ_1	Sin-variation of Blur Size	Varied	0.1 pixels
σ_2	Cos-variation of Blur Size	Varied	-0.05 pixels
B_{CCD}	CCD Bias	Varied	300 counts

Using this model to fit the example data from Figure 2.7 produces a smaller residual, as shown in Figure 2.8. Evidently, (2.20) can provide a more accurate characterization of the instrument function than the ideal Airy function alone. However, it is not perfect. The residual, while smaller, is not entirely noise-like, exhibiting features indicating systematic errors in the shape of the fringes. The reduced χ^2 of the fit is 2.6; it should be near 1 for a residual consisting purely of Gaussian white noise. The imperfections of our model are amplified by the high signal-to-noise ratio (SNR) of the laser calibration

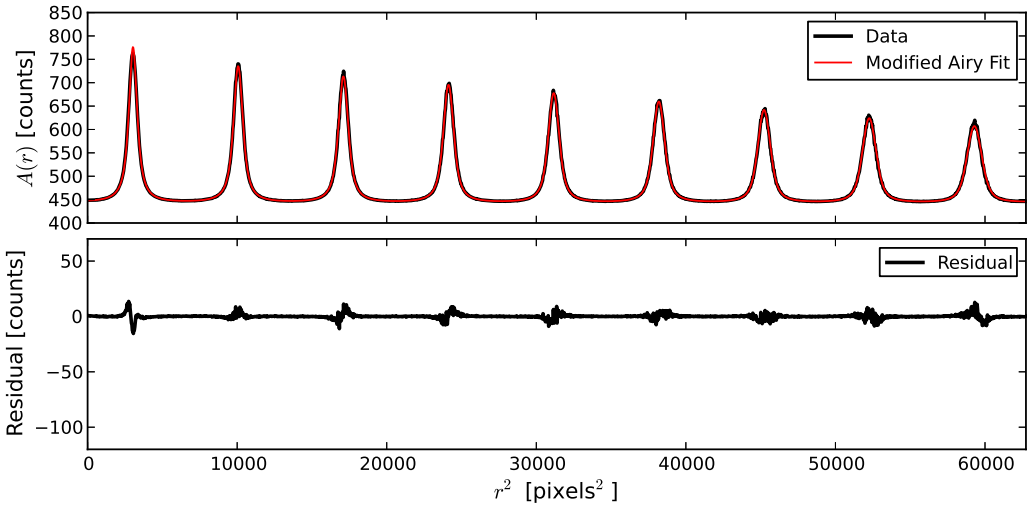

Figure 2.8: Same as Figure 2.7, except using a modified Airy model, as described in the text and summarized in (2.20). Adapted from *Harding et al.* [2014].

images; since the noise is so low, small imperfections of our model have a large impact on χ^2.

An improvement to this model would be a valuable topic for future work. As detector technology improves, and statistical errors in the airglow images become smaller, we can benefit more from improved models which reduce systematic errors.

Airglow Fringe Model

Having described the model for the laser calibration fringes, we now describe the model for the airglow fringes. The same equations apply, except that instead of a delta function, the source spectrum is a Gaussian function with a width determined by the thermospheric temperature and a Doppler shift determined by the line-of-sight thermospheric wind:

$$Y(\lambda) = Y_{\text{bg}} + \frac{Y_{\text{line}}}{\sigma_\lambda} \exp\left[-\frac{1}{2}\left(\frac{\lambda - \lambda_c}{\sigma_\lambda}\right)^2\right] \tag{2.21}$$

where Y_{bg} is the contribution from the spectral background, assumed to be independent of λ, Y_{line} is proportional to the observed brightness of the emission line, λ_c is the center wavelength of the line, and σ_λ is the linewidth. The

center wavelength is related to the rest wavelength, $\lambda_0 = 630.0$ nm, via the Doppler shift equation:

$$\lambda_c = \lambda_0 \left(1 + \frac{v}{c}\right) \tag{2.22}$$

where v is the line-of-sight thermospheric wind velocity, defined as positive away from the instrument, and c is the speed of light. The linewidth is related to the temperature:

$$\sigma_\lambda = \frac{\lambda_0}{c}\sqrt{\frac{kT}{m}} \tag{2.23}$$

where k is the Boltzmann constant, T is the temperature of the emitting species, and m is its mass.

The previous three equations, along with the instrumental equations used above ((2.1), (2.12), (2.14), (2.15), (2.17), (2.18), and (2.19)), describe the forward model for the airglow fringes. These equations are summarized here:

$$\begin{aligned}
S(r) &= \int_0^\infty \tilde{A}(r,\lambda)\, Y(\lambda)\, d\lambda + B_{\text{CCD}} \\
Y(\lambda) &= Y_{\text{bg}} + \frac{Y_{\text{line}}}{\sigma_\lambda}\, \exp\left[-\frac{1}{2}\left(\frac{\lambda - \lambda_c}{\sigma_\lambda}\right)^2\right] \\
\tilde{A}(r,\lambda) &= \int_0^{r_{\max}} b(s,r)\, A(s,\lambda)\, ds \\
A(r,\lambda) &= I_0\, \frac{1 + I_1\left(\frac{r}{r_{\max}}\right) + I_2\left(\frac{r}{r_{\max}}\right)^2}{1 + \frac{4R}{(1-R)^2}\sin^2\left(\frac{2\pi n t}{\lambda}\cos\tan^{-1}(\alpha r)\right)} \\
b(s,r) &= \frac{1}{\sqrt{2\pi\sigma(r)^2}}\exp\left(-\frac{(s-r)^2}{\sigma(r)^2}\right) \\
\sigma(r) &= \sigma_0 + \sigma_1 \sin\left(\pi\frac{r}{r_{\max}}\right) + \sigma_2 \cos\left(\pi\frac{r}{r_{\max}}\right).
\end{aligned} \tag{2.24}$$

The free parameters of this model are summarized in Table 2.3. These parameters are varied in order to fit the airglow fringes. All other parameters (i.e., those in Table 2.2) are fixed at the values determined in the analysis of the laser images. Note that B_{CCD} is estimated separately for the laser calibration and airglow fringes, due to the different exposure times.

An example fit of this model to actual data is shown in Figure 2.9. Al-

Table 2.3: Model Parameters for Airglow Fringes

Parameter	Description
λ_c	Line Center
σ_λ	Linewidth
Y_{bg}	Background Emission
Y_{line}	Line Brightness
B_{CCD}	CCD Bias

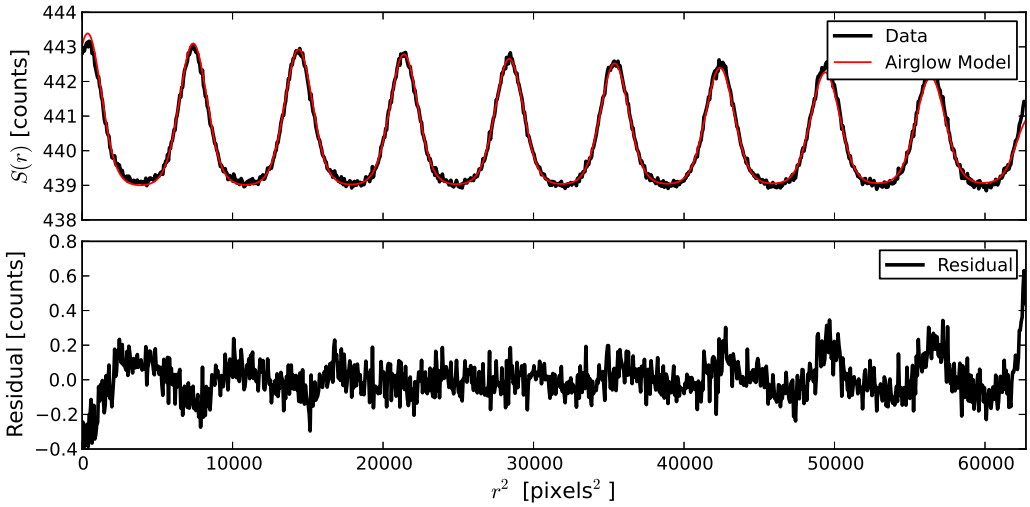

Figure 2.9: (Top) An example airglow fringe pattern along with the best fit of the model in (2.24). (Bottom) Difference between data and model. Adapted from *Harding et al.* [2014].

though the fit is good, some systematic errors are evident; the residual is not purely white noise. The primary systematic error is the imperfect characterization of the throughput attenuation. The inner fringes are estimated slightly too large, and the outer fringes slightly too small. A secondary systematic error arises from edge effects of the blurring filter. To account for this, we increase the uncertainties of the edge samples by a factor of 10 before inverting.

2.2.3 Inversion

Here we describe the details of the reduction of the raw data to a 1D fringe pattern, and the nonlinear least-squares fits which estimate parameters from the laser calibration and airglow fringe patterns.

Even though it is theoretically possible to fit the models above directly to the 2D raw data, any algorithm to do so would have to explicitly account for azimuthal variations in the illumination, which are averaged over by annular summation. More importantly, fitting directly to the 2D image requires computing the model at every pixel, which is prohibitively expensive.

We perform annular summation as described by *Makela et al.* [2011]. Briefly, we first determine the center of the 2D fringe pattern by thresholding the laser image at the 70^{th} percentile level, identifying the pixels composing contiguous regions as individual rings, fitting a circle to the points in each ring, and taking the median center value over all the rings. Practically, we have found that it is important to *fit* a circle to each ring, not *bound* each ring, which is sometimes done, but is much more susceptible to noise. There are also multiple ways to fit a circle; we have found that the geometric method produces smaller errors than the algebraic method [*Gander et al.*, 1994], especially in cases with azimuthal variation in the laser illumination. For MiniME, this process produces an estimate of the center location with an uncertainty of a few hundredths of a pixel, which is much less than 0.3 pixels, which *Armstrong* [2008] finds is required in order to achieve a temperature error of less than 1%. The annular summation process assumes that the rings are perfectly circular. For MiniME, a narrow-field system, we find that the ring ellipticity is greater than 0.999. *Conde* [2002] finds that for FPIs with all-sky optics, the elliptical distortion needs to be accounted for, especially since winds are estimated separately for different azimuthal sections of the ring.

After finding the center pixel, we sort the pixels by distance from the center and organize them into R radial bins. The bins are chosen such that there is an approximately equal number of pixels in each bin (requiring bin widths proportional to $\sqrt{r}$). The MiniME system produces images of size 512×512, so we typically use $R = 500$ radial bins. The choice of R is not important; our simulations (not shown) suggest that as long as R is greater than 200, but small enough that each bin has several pixels, the estimates are not sensitive to R.

Within the i^{th} bin, we define r_i as the mean radius of the bin, and $N_{\text{laser}}(r_i)$ and $N_{\text{airglow}}(r_i)$ as the mean counts of all the pixels in that bin, for the laser fringes and airglow fringes, respectively. To minimize the influence of hot pixels, dead pixels, and cosmic rays, we ignore all pixels that differ by more

than three standard deviations from the mean in each bin before averaging. Since there are many pixels in each bin, we can also bootstrap an uncertainty estimate for $N_{\text{laser}}(r_i)$ and $N_{\text{airglow}}(r_i)$ by dividing the standard deviation of pixel counts by the square root of the number of pixels in each bin. These uncertainties are denoted $\sigma_{N_{\text{laser}}}(r_i)$ and $\sigma_{N_{\text{airglow}}}(r_i)$.

The first step is to estimate the instrument parameters from $N_{\text{laser}}(r_i)$, which we will use in the second step to estimate airglow parameters from $N_{\text{airglow}}(r_i)$. For the first step, we pose a standard least-squares problem, finding the parameters that minimize χ^2:

$$\chi^2(\mathbf{p}_{\text{instr}}) = \sum_{i=0}^{R-1} \left(\frac{N_{\text{laser}}(r_i) - S(r_i, \mathbf{p}_{\text{instr}})}{\sigma_{N_{\text{laser}}}(r_i)} \right)^2 \tag{2.25}$$

where $\mathbf{p}_{\text{instr}}$ denotes the vector of free instrument parameters from Table 2.2, and $S(r_i, \mathbf{p}_{\text{instr}})$ is from (2.20), though we have made the dependence on $\mathbf{p}_{\text{instr}}$ explicit here.

In order to evaluate (2.20), we must numerically evaluate an integral. We do this with a simple Riemann sum over the radial bins:

$$\begin{aligned} S(r) &= \int_0^{r_{\max}} b(s, r)\, A(s, \lambda_{\text{laser}})\, ds \; + B_{\text{CCD}} \\ &\approx \sum_{j=0}^{R-1} \; b(s_j, r)\, A(s_j, \lambda_{\text{laser}})\, \Delta s_j \; + B_{\text{CCD}}. \end{aligned} \tag{2.26}$$

In practice, the blurring kernel, b, is narrow enough that not all terms in the sum need be retained, and Δs_j may be taken as constant.

Various algorithms exist to solve nonlinear least-squares problems. Our code uses the `lmfit` Python package, which provides a high-level interface to the `scipy.optimize` implementation of the Levenberg-Marquardt algorithm. This algorithm finds locally optimum parameters and calculates their statistical uncertainties.

Because of the highly nonlinear behavior of the Airy function with respect to many of its parameters, the choice of minimization algorithm is not nearly as important as the choice of the initial guess that the algorithm uses. Some of these parameters, such as I_0, I_1, and B can be easily estimated directly from the data. For other parameters which are known to be small, we have found that small initial guesses suffice ($\sigma_0 = 0.5$ pixels, $I_2 = \sigma_1 = \sigma_2 = 0$).

Some parameters, such as α and R, are known to high accuracy based on the instrument design. For the etalon gap, t, we can take advantage of the fact that the Airy function is periodic in t, and we know the approximate value from the design. Using a brute-force grid search, we try 50 different values of t spanning one period to find the value that correlates most with the data. It is important to center this search window on the estimate of t obtained from the last laser calibration image, or else the estimate may jump by a period over the night, as the etalon drifts. Except for this detail, the inversion is performed independently from image-to-image.

Even with these good initial guesses, we have found that immediately running the Levenberg-Marquardt algorithm to optimize all parameters occasionally causes the algorithm to get stuck in a local minimum of χ^2 that is not the desired global minimum. Although this is rare, we have over 3 million FPI images in our database (as of January 2017), so it is essential that the analysis algorithm is robust with no human oversight. A more conservative approach is to run Levenberg-Marquardt in stages, optimizing a few parameters at a time, while keeping the others fixed at their previously determined values. Although ad-hoc, the sequence of stages we use has undergone extensive testing and appears to ensure convergence. The final step is an optimization over all the parameters. The cost of this overly conservative approach is runtime; however, even on a typical desktop computer, it runs in a few seconds per image, an order of magnitude faster than the exposure time, which means that real-time operation is a possibility. In practice, we gather a full night of data and process it all at once in the morning. For an unknown reason, the estimation of the blur parameters (σ_0, σ_1, and σ_2) for some FPI systems is unstable, and produces artifacts, especially in the temperature estimates. Although not a problem for any of the FPIs used in this work, it is a problem for the FPI in Morocco, for which we fix these values at predetermined values instead of estimating them.

Having estimated the instrument parameters, we move on to the second step: estimating the airglow parameters from the airglow images. This occurs in an analogous way to the laser images. First, the images are converted to 1D fringe patterns using the center pixel determined from the laser image.

We then minimize a χ^2 function, as above:

$$\chi^2(\mathbf{p}_{\text{airglow}}) = \sum_{i=0}^{R-1} \left(\frac{N_{\text{airglow}}(r_i) - S(r_i, \mathbf{p}_{\text{airglow}})}{\sigma_{N_{\text{airglow}}}(r_i)} \right)^2 \tag{2.27}$$

where $S(r_i, \mathbf{p}_{\text{airglow}})$ is from (2.24), and $\mathbf{p}_{\text{airglow}}$ is the vector of airglow parameters in Table 2.3. All of the other parameters in (2.24) are held fixed at the values determined from the laser image.

As above, evaluation of (2.24) requires numerical evaluation of integral equations with a Riemann sum. The blurring integral is handled analogously to (2.26):

$$\begin{aligned} \tilde{A}(r, \lambda) &= \int_0^{r_{\max}} b(s, r)\, A(s, \lambda)\, ds \\ &\approx \sum_{j=0}^{R-1} b(s_j, r)\, A(s_j, \lambda)\, \Delta s_j \,. \end{aligned} \tag{2.28}$$

The spectral integral is complicated by the fact that its domain is infinite. However, due to the properties of the optical filter used, we know $Y(\lambda)$ has limited support. Moreover, because of our assumption that the background continuum, Y_{bg}, is independent of λ, it is not necessary to integrate far into the tails of the Gaussian spectrum. We thus restrict the domain of integration to five periods of the Airy function surrounding 630.0 nm, and the spectral integral is approximated by

$$\begin{aligned} S(r) &= \int_0^{\infty} \tilde{A}(r, \lambda)\, Y(\lambda)\, d\lambda + B_{\text{CCD}} \\ &\approx \sum_{l=0}^{L-1} \tilde{A}(r, \lambda_l)\, Y(\lambda_l)\, \Delta\lambda \; + B_{\text{CCD}} \end{aligned} \tag{2.29}$$

where we discretize the spectrum into L equally sized bins of width $\Delta\lambda$. In practice, we use $L = 301$. Since we must evaluate (2.29) at every radial bin, r_j, every time the forward model is evaluated, and since $\tilde{A}(r, \lambda)$ does not depend on $\mathbf{p}_{\text{airglow}}$, we can improve runtime by precomputing and storing the values of $\tilde{A}(r_j, \lambda_j)$ in a matrix. Thus, (2.29) can be written as a matrix-vector product:

$$\mathbf{s} = \tilde{\mathrm{A}}\mathbf{y} + B_{\text{CCD}}\mathbf{1} \tag{2.30}$$

where $\mathbf{1}$ is a vector of length R consisting of all ones, and the other variables

are

$$\mathbf{s} = \begin{bmatrix} S(r_0) \\ S(r_1) \\ \vdots \\ S(r_{R-1}) \end{bmatrix}$$

$$\tilde{\mathrm{A}} = \begin{bmatrix} \tilde{A}(r_0, \lambda_0) & \tilde{A}(r_0, \lambda_1) & \dots & \tilde{A}(r_0, \lambda_{L-1}) \\ \tilde{A}(r_1, \lambda_0) & \tilde{A}(r_1, \lambda_1) & \dots & \tilde{A}(r_1, \lambda_{L-1}) \\ \vdots & \vdots & \ddots & \vdots \\ \tilde{A}(r_{R-1}, \lambda_0) & \tilde{A}(r_{R-1}, \lambda_1) & \dots & \tilde{A}(r_{R-1}, \lambda_{L-1}) \end{bmatrix} \Delta\lambda \tag{2.31}$$

$$\mathbf{y} = \begin{bmatrix} Y(\lambda_0) \\ Y(\lambda_1) \\ \vdots \\ Y(\lambda_{L-1}) \end{bmatrix}.$$

As with the laser inversion, we must carefully choose good initial guesses for the airglow parameters. Y_{line} and B_{CCD} can be estimated directly from the data, and λ_c can easily be estimated using a grid search identical to the one described above for t. We have found that a simple guess of $Y_{\text{bg}} = 0$ works well, as does setting σ_λ so that $T = 1000$ K.

As above, we use the Levenberg-Marquardt algorithm to optimize the parameters in stages to ensure convergence. After convergence, we convert λ_c and σ_λ to line-of-sight wind and temperature using (2.22) and (2.23). Because of the uncertainties with absolute calibration of the FPI, we do not attempt to provide Y_{line} in absolute units such as Rayleighs.

Because the laser calibration images are not taken at exactly the same time as the airglow images, we must account for the fact that $\mathbf{p}_{\text{instr}}$ changes with time. The center pixel varies quite slowly, if at all, and the uncertainty with its estimate is larger than its image-to-image variation. Thus, we fit a polynomial over time to a full night of x- and y-components of the center pixel, and use this to determine the center for annular summation. The degree of the polynomial adapts to the number of samples to minimize over- or under-fitting. To avoid extrapolation, we only process airglow images that are straddled by valid, successfully processed laser images.

Many data quality problems manifest at this stage in the analysis. For

example, if there are any intermittent contaminant light sources, if the CCD cooler fails, or if the SkyScanner fails to look directly at the laser diffusion box, then we see large variations in the estimated center pixel. If the standard deviation of the center pixel is greater than 1.5 pixels, we remove all points greater than two standard deviations away from the mean.

The other instrument parameters exhibit larger variations in time than their estimation uncertanties, so we use linear interpolation in time. This captures variability more accurately than a polynomial or spline fit. Additionally, with linear interpolation, the effect of outliers is contained to one or two samples around the outlier, whereas with a full-night fit, a single outlier can affect the entire night of data.

The laser calibration signal is orders of magnitude brighter than the airglow signal, so we may therefore treat $\mathbf{p}_{\text{instr}}$ as having negligible uncertainty in the airglow inversion even though it is an estimated quantity. The only parameter for which we consider its calibration uncertainty is the etalon gap, t, due to its sensitive impact on the wind estimate. There are two components of the calibration uncertainty of t: (1) the uncertainty associated with the Levenberg-Marquardt fit of the laser fringes (which is largely driven by model uncertainty and not noise), denoted $\sigma_{t,\,\text{fit}}$, and (2) the uncertainty associated with the linear interpolation in time, denoted $\sigma_{t,\,\text{interp}}$. The latter uncertainty is large when the etalon gap is changing rapidly in a way that is not captured by a linear interpolation between consecutive laser images. This is difficult to characterize. We use an ad-hoc technique, which computes the difference between the estimated t of laser image n and the expected t based on linear interpolation of laser images $n-1$ and $n+1$. The difference provides a basic characterization of the error associated with linear interpolation. We assume that the uncertainties are uncorrelated and therefore add in quadrature to determine the total uncertainty of λ_c:

$$\sigma_{\lambda_c}^{\text{total}} = \sqrt{\sigma_{\lambda_c}^2 + \frac{\lambda_0^2}{t^2}\left(\sigma_{t,\,\text{fit}}^2 + \sigma_{t,\,\text{interp}}^2\right)}. \tag{2.32}$$

In optimized observations, the calibration uncertainties are often negligible compared to σ_{λ_c}. However, for historical data when the laser cadence was slower, or for cases when the laser is dimmer than usual, these two factors can be important, and sometimes even dominate, so we include them in the standard analysis routine.

It is worth noting that the two "background" parameters in the airglow fringe model, Y_{bg} and B_{CCD}, have different physical sources but similar signatures in the data. They both contribute to the level of background counts upon which the airglow fringes are superimposed. Expanding (2.24), we see that the contribution of Y_{bg} to the measured signal $S(r)$ occurs via the term $Y_{\text{bg}} \int_0^\infty \tilde{A}(r, \lambda)\, d\lambda$, where the integral can be thought of as the sensitivity of the system at radius r. If this term varies with r, which is usually the case due to vignetting and different path lengths through the optics, then a large Y_{bg} causes a slope with r (often negative) to be superimposed on the airglow fringe pattern, while B_{CCD} is a constant. This allows us to distinguish between these two parameters in the inversion. However, although $\int_0^\infty \tilde{A}(r, \lambda)\, d\lambda$ has a slope with r, the slope sits on a large background offset, which looks the same as the constant offset applied by B_{CCD}. This means Y_{bg} and B_{CCD} are somewhat degenerate. Thus, Y_{bg} estimates should not be treated as quantitatively useful data. In fact, it often is estimated to be negative. This is physically unrealistic but allows us to better fit the fringe pattern and achieve better estimates of λ_c, σ_λ, and Y_{line}. Estimates of Y_{bg} can still be useful qualitatively, as they can corroborate cloud sensor measurements, indicate when moonlight may be affecting our data, and suggest the presence of contaminant stray light sources. Ideally, this problem should be fixed by determining B_{CCD} independently from the laser calibration images, the dark images, and the bias images, instead of being estimated as part of the airglow fringe analysis. However, all attempts to do so have failed, possibly due to the dependence of B_{CCD} on ambient temperature.

2.2.4 Post-Processing and Quality Control

After iterating over all laser and airglow images and estimating their parameters, there are several post-processing and quality control steps of importance.

The most critical step is Doppler referencing. Because neither the laser wavelength nor the rest wavelength of the 630.0-nm emission is known to sufficient accuracy, the wavelength corresponding to zero wind is not known. Doppler referencing is also important in order to correct for errors in specifying the nominal t or n. The result of this is that there is a constant offset

added to the estimated winds. In the literature there have been two methods used, which we refer to as "zenith reference" and "laser reference."

Some of the earliest observations [e.g., *Hays and Roble*, 1971b] did not make use of an independent stable-wavelength source like a laser or 630.0-nm lamp. Instead, they made the assumption that the vertical wind was zero, or at least, negligibly small compared to the uncertainties in their measurements. They used the zenith-looking observations to establish the wavelength shift corresponding to zero wind, and applied this to other, obliquely looking measurements. We use the zenith reference method when laser images are not available.

Even when laser images are available to track relative changes in the Doppler reference, there remains a constant offset to be estimated. This is usually established by assuming that, while the vertical wind is not instantaneously zero, the mean vertical wind over the night is approximately zero [e.g., *Hernandez*, 1982]. *Aruliah and Rees* [1995] find that this assumption is good to within 5–10 m/s, even in the auroral region during active periods. However, oftentimes the vertical wind can appear to obtain sustained positive or negative values, which can occur either due to actual vertical winds in the auroral region [*Larsen and Meriwether*, 2012], or due to contamination by OH lines or atmospheric scatter, as discussed in Chapter 5. In these cases, the mean vertical wind can deviate from zero by many tens of m/s, which renders it unacceptable for use as a Doppler reference.

In order to handle these cases, we propose a method which uses a different measure of central tendency. The mean is the statistic which minimizes the 2-norm of the error:

$$\text{mean}(\mathbf{x}) = \underset{a}{\text{argmin}} \sum_{i=0}^{N-1} |x_i - a|^2 \tag{2.33}$$

where $\mathbf{x}$ is a vector of length N. In this form, it is clear why the mean is sensitive to outliers caused by OH or large vertical winds: outliers are amplified by the squared term. We thus use a general p-norm instead of the 2-norm:

$$\underset{a}{\text{argmin}} \sum_{i=0}^{N-1} |x_i - a|^p. \tag{2.34}$$

Choosing $p = 1$ results in the median of the data, which is more resistant

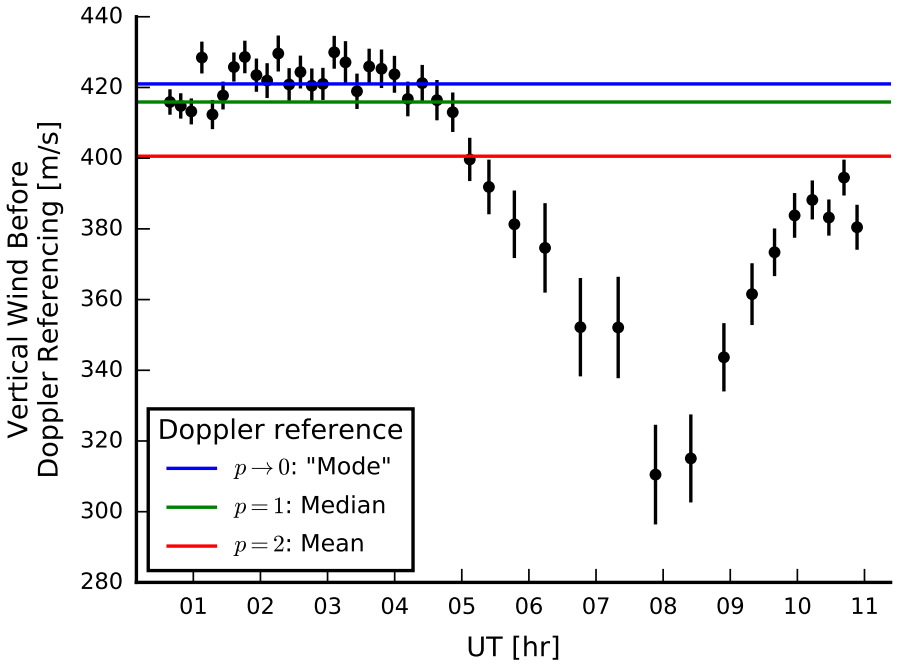

Figure 2.10: Three ways of establishing the constant Doppler reference, using different values of p in (2.34), corresponding to the mean, median, and (approximately) the mode of the data. Data are taken from day 264 of year 2014 at UAO, and shows a particularly extreme example of the effects of OH contamination. The laser has already been used to remove relative variations in the line center.

to outliers than the mean. As p approaches zero, this statistic approaches the most common value, the mode, which we have found to be a more reliable indicator of zero vertical wind in most cases. If the estimated vertical wind remains at approximately the same value for a long period of time, that value is likely to be zero wind. Since this process is needed only once per night of data, we are not concerned with runtime and thus minimize (2.34) with a brute force search. Numerically, (2.34) is ill-defined at $p = 0$, and with real data, there are no repeated values, so it is not useful to calculate the mode directly. Instead, we use a small value, $p = 0.1$. In our testing, the result is not sensitive to p as long as it is small. Figure 2.10 shows an example night of UAO data that is severely affected by OH contamination, which we verified by a visual inspection of the fringes, where a second peak can be clearly seen (not shown). This causes an artificial negative bias on the vertical wind data in the later hours of the night. The different measures of central

tendency are also shown, where the mode appears to provide the most reliable Doppler reference based on the consistency of early-night vertical winds. This mode statistic can fail if vertical winds are sustained for long periods at a constant non-zero value or are consistently fluctuating. For manual analysis of individual nights or other specialized studies, various alternate strategies for establishing the Doppler reference can be used, but for our standard data product, we use the mode.

In theory, we could use many nights or years of data to get a better estimate of the Doppler reference. In practice, we establish the Doppler reference independently for each night, since there is no guarantee that the laser will lock to the same wavelength each night. For some of our instruments, the Doppler reference is consistent (e.g., the standard deviation of the CAJ Doppler reference in 2013 was only 7 m/s). However, for others, such as UAO, there appears to be a 50 m/s variation throughout the year, possibly due to the variation of n with temperature and humidity. Additionally, the Doppler reference can change any time modifications are made to the instrument, which are often not reliably recorded.

Automated quality control for FPIs is difficult, and our algorithm is constantly evolving. We attempt to summarize the myriad of quality issues into a single quality flag, which takes a value of either 0, 1, or 2. Samples for which we are confident in their high quality are given a flag of 0. Samples which may have an issue, or may be slightly corrupted but still useful, are given a flag of 1. For many scientific studies, these could be used, but they should be treated with caution. Samples which are most likely bad are given a flag of 2. There is a separate flag for the wind and temperature estimates.

The current quality control algorithm is based on information about laser drift, airglow brightness, cloud sensor readings, CCD cooling status, and the result of the Levenberg-Marquardt fit. If drift in the laser wavelength is suspected (i.e., if the estimated vertical wind at the beginning and end of the night differ by more than 30 m/s, and the laser brightness changes by more than 20% – an indicator that the laser's frequency lock has failed), the wind flag is set to 1. If zenith reference is used for the Doppler reference, both the wind and temperature flag are set to 1. If the cloud sensor's sky and ambient temperature difference is less than 22 °C, it is somewhat cloudy, and both flags are set to 1. If less than 10 °C, it is almost certainly overcast, and the wind flag is set to 2. Temperature data, while noisier, are not necessarily

biased by the clouds, since the Doppler width of the emission is much larger than the Doppler shifts caused by the wind. The temperature flag remains at 1. If the cloud sensor was not installed or not working, both flags are set to 1. If either the wind or temperature fit uncertainty is greater than 100 m/s or 100 K, both flags are set to 2. Experience has shown that in these cases, it is likely that the fit is not reliable, and the actual error is not well captured by the uncertainty. In some cases, the flag is set manually to account for a known problem, such as the small amount of light contamination at the CAJ site in mid-2013 causing 50–100 K errors in the temperature estimate.

The final factor in the quality-control algorithm is contamination by OH emission lines. These lines have been known to be a problem since the beginning of thermospheric wind and temperature observations using FPIs [e.g., *Armstrong*, 1969], and their effects on FPI measurements are discussed at length by *Hernandez* [1974], *Burnside et al.* [1977], and *Ronksley* [2016]. Two OH lines are near enough to 630.0 nm to be of possible concern: the $P_1(3)$ line at ∼630.7 nm and the $P_2(3)$ line at ∼629.8 nm. These emissions are dim; *Hernandez* [1974] reports 2–4 R for the 629.8-nm emission and 5–8 R for the 630.7-nm emission. The 630.7-nm emission is deep in the wing of our filters, so its brightness is reduced, but the 629.8-nm emission is closer to the middle of the passband. When the 630.0-nm emission is bright, these lines do not cause large errors. However, the 630.0-nm emission is quite variable, and the OH emission can become the largest source of error when the 630.0-nm emission becomes dim, which occurs when the electron density is low or when the F region is lifted to high altitudes.

Although these lines are well-isolated from 630.0 nm, the periodic nature of the Airy function aliases them onto the 630.0-nm peak, where they could overlap it, lie on the wing, or lie between peaks, depending on the etalon gap. The effect on wind and temperature estimates depends on the instrument design, the filter characteristics, and to a lesser degree, the fitting algorithm. We could attempt to estimate the impact of OH using knowledge of our etalon and filter design, as well as an educated guess about the relative brightness of the OH and 630.0 lines. However, given the uncertainties associated with this, we use a data-driven approach.

We assume that the OH line has constant brightness. Despite the fact that the OH brightness is known to vary [*Hernandez*, 1974], and also that its measured brightness will vary as the filter's spectral response shifts with

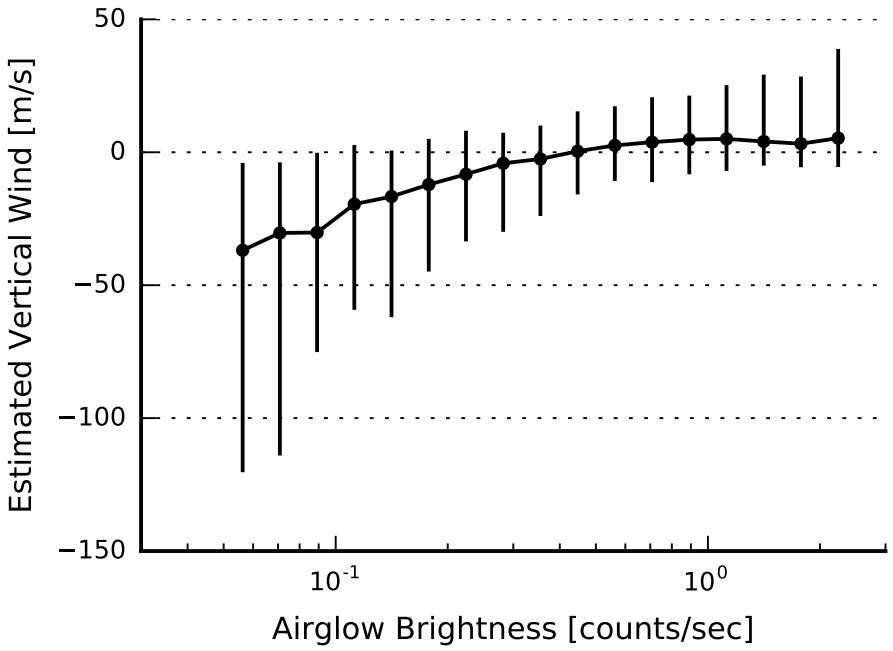

Figure 2.11: All available vertical wind data from UAO that passed the quality check, binned as a function of airglow brightness. The dots are the median vertical wind in each bin, and the bars indicate the intervals in which the middle 68% of samples lie.

temperature, these variations are smaller than the variation in 630.0-nm brightness. We thus treat the OH effect as purely dependent on 630.0-nm brightness.

For each instrument, we collect all available nights of vertical wind data that pass the following quality checks: (1) 80% of points have a wind quality flag of 0 or 1; (2) the cloud sensor and laser are both operating; (3) the geomagnetic activity is low (Kp ≤ 3); (4) the night is within 5 days of new moon; and (5) at least five vertical wind samples are available. We then bin the data by the estimated airglow brightness, Y_{line}, and compute the median vertical wind in each bin. An example for UAO is shown in Figure 2.11. Clearly, below a certain threshold brightness (about 0.4 counts/sec for UAO), there is a decrease in the estimated vertical wind with decreasing brightness. We attribute this to contamination by the OH emission. Indeed, manual inspection of UAO's fringes at very low brightness (not shown) indicates the presence of a second peak on the blue side of the 630.0-nm peak. This effect varies considerably from instrument to instrument. For example, CAR

exhibits almost no OH effect, and for MOR, the effect appears to be a slight upward wind. This difference may due to different filter properties; perhaps the 630.7-nm OH emission has an effect at these sites. It is also possible that the etalon gap is slightly different, causing the peak to alias to a different location on the 630.0-nm fringe. It is also possible, though less likely, that laser light is somehow contaminating the sky exposure.

In order to inform the user when OH may be contaminating the data, we raise the quality flag to 1 when the emission brightness drops below an instrument-dependent threshold. This threshold is chosen using the instrument-specific median line similar to that shown in Figure 2.11. Where the line deviates by 5 m/s from the average of the high-brightness samples, we choose the corresponding brightness as the threshold. For UAO, this threshold is 0.45 counts/sec.

An interesting feature of Figure 2.11 is that the brightest samples have a median vertical wind of 3–5 m/s. Since we expect the brightest samples to be the most reliable, and we also expect that the median vertical wind is actually zero, this suggests that OH is doing more than affecting individual samples; it is also biasing the Doppler reference low on average. If the Doppler reference were always set perfectly, the line should be shifted down so that these samples are centered on zero.

Instead of merely flagging its effect, we have spent considerable effort attempting to improve the fitting algorithm to account for or remove OH. We have tried introducing an additional Gaussian to $Y(\lambda)$, and fitting both the OH and 630.0 lines simultaneously, but a double-Gaussian fit is too unreliable for an automated algorithm. We have also tried simpler approaches such as treating the position of the OH line as known, treating the entire line shape as known, and increasing the uncertainties in radial bins where OH is known to be present. While potentially useful for specialized studies, none of these algorithms has produced results satisfying enough for general use. This is a worthy topic for future consideration.

Until now, we have discussed the generation of *line-of-sight* wind estimates. Many methods have been used to convert obliquely looking line-of-sight winds into horizontal winds. For example, *Shiokawa et al.* [2012] use the assumption that the horizontal wind is constant across the ∼500 km separating the north-looking and south-looking measurements. Under this assumption, and the assumption of zero vertical wind, the difference between these two

line-of-sight measurements provides the meridional wind. The zonal wind is determined analogously from the east- and west-looking measurements. *Makela et al.* [2012] do away with the assumption of zero vertical wind, since it can be tracked. They instead assume that the overhead vertical wind is the same as the vertical wind ∼250 km away at the pierce point of the oblique measurement, remove the vertical-wind component from the oblique measurement, and extract the horizontal component. This method allows independent estimates of the meridional wind at the north and south pierce points, for instance. However, all of these techniques make some assumptions about the spatial extent of the vertical, horizontal, or both components of the wind.

When multiple FPIs are used in tandem, some spatial ambiguities can be resolved. *Fisher* [2013] describes the common-volume technique, which can unambiguously resolve two or three components of the wind at a specific point. An alternative technique is a wind field fitting algorithm, which uses line-of-sight winds at various locations from multiple FPIs to fit a two-dimensional, latitude- and longitude-dependent distribution of zonal, meridional, and vertical winds. An algorithm for this purpose is the topic of Chapter 3.

2.2.5 Monte Carlo Simulations

In this section, we report the results of three Monte Carlo simulations which test the inversion presented above. The first tests the accuracy of uncertainty estimates, and the second and third test for biases over inputs (wind and temperature) and SNR, respectively.

For each Monte Carlo trial, we start with a "true" wind and temperature, and simulate a laser calibration image and airglow image using (2.20) and (2.24), with instrument parameters set to values seen in practice. The integrals are evaluated with three times higher resolution during image generation than in the model for inversion. Complications such as OH, instrument drift, Doppler referencing, and the calibration uncertainty in (2.32) are ignored. These simulations are intended to test the model fit only. Noise is

added according to the SNR, which we define simply as

$$\text{SNR} \equiv \frac{\Delta S}{\sigma_n} \tag{2.35}$$

where ΔS is the difference between the maximum and minimum $S(r)$ measured in CCD counts, and σ_n is the standard deviation of the Gaussian white noise in the same units.

We apply the algorithm described in Section 2.2.3 to the simulated laser and airglow images to obtain estimates of the wind and temperature, which are compared with the "true" values originally used.

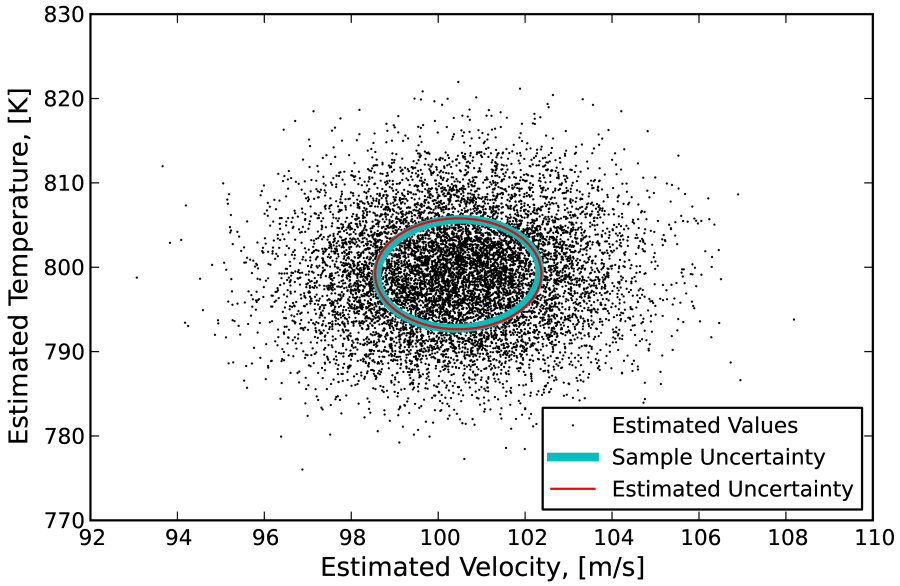

Figure 2.12: Estimated wind and temperature for 10^4 simulated FPI inversions with the true wind and temperature set to 100 m/s and 800 K. The error ellipses represent the sample covariance matrix (blue) and the uncertainties reported by the inversion routine (red). Adapted from *Harding et al.* [2014].

Simulation 1: Uncertainty Estimates

For this simulation, we use a constant wind (v = 100 m/s), temperature (T = 800 K), and noise level (SNR = 5). We run 10^4 trials using different noise realizations, and compare the reported uncertainties with the standard

deviation of the wind and temperature estimates. If the inversion's uncertainty estimates are accurate, these should match exactly.

The results are shown in Figure 2.12. Each dot represents the estimates from a single trial. The blue ellipse represents the sample covariance ellipse of all 10^4 points, and the red ellipse represents the uncertainty reported by the inversion, averaged over the 10^4 inversions. The coincidence of these two ellipses shows that the uncertainties are accurately estimated. Further, since 68% of the dots lie within the red ellipse, it also shows that our errors follow Gaussian statistics, as expected. Of note is that the inversion reports uncertainties in wind and temperature which are assumed to be uncorrelated. Evidently, this is a good assumption. This simulation also shows that the noise added to the laser images has a negligible effect on the uncertainty in the airglow parameter estimates.

Simulation 2: Bias over Wind and Temperature

Having characterized statistical errors, we now characterize systematic errors. In this simulation, we keep SNR = 5, but vary the true wind and temperature from -300 to 300 m/s and 300 K to 1500 K, respectively. For each of the 10^3 trials, we randomly pick a wind and temperature from a uniform distribution, simulate the images, perform the inversion, and compute the estimation error. In Figure 2.13 we show the wind and temperature error as a function of true wind and temperature.

For the velocity estimates, there is a systematic 0.4 m/s bias of unknown origin. However, it is negligible compared to the error in the knowledge of the 630.0-nm emission wavelength and the laser wavelength, and it is removed in the Doppler-referencing process. The velocity error does not appear to depend on velocity or temperature.

For the temperature, we find no statistically significant bias. This is particularly important when compared to the Fourier-based algorithm tested by *Makela et al.* [2011], which exhibits a positive temperature bias at high temperature, unless the fringe wings are ignored, which increases the effective SNR. Our method accurately accounts for the overlapping fringes at high temperature, yielding no bias. The temperature estimation variance increases with temperature, which is expected. When the temperature is high,

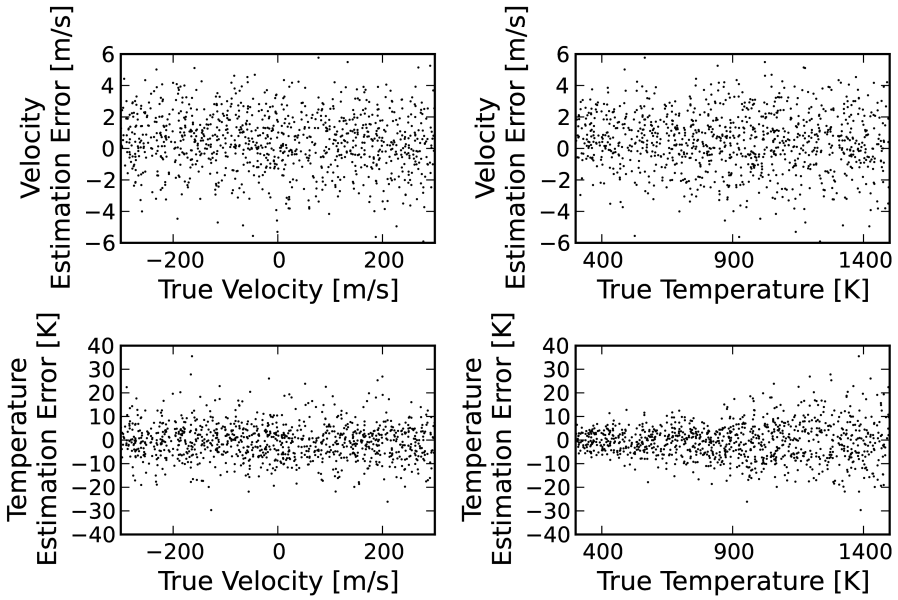

Figure 2.13: Errors in wind and velocity estimates for 10^3 simulated FPI inversions using various winds and temperatures. Adapted from *Harding et al.* [2014].

the signal is spread over more radial bins and thus the total noise power, integrated over the fringe, is larger.

Simulation 3: Bias over SNR

For the last simulation, we test for biases over SNR. We use a randomized wind and temperature, as above, and a randomized SNR, varying from 0.5 to 5. The results of 10^4 trials are shown as a function of SNR in Figure 2.14. No biases are seen in the wind and temperature over the SNR ranges tested here, which roughly match those we see in practice. However, it should be noted that these were simulated and inverted with the same model. For instruments where the instrument function deviates significantly from (2.20), or for cases with spectral contamination such as OH in the airglow images, we expect to see biases at low SNR.

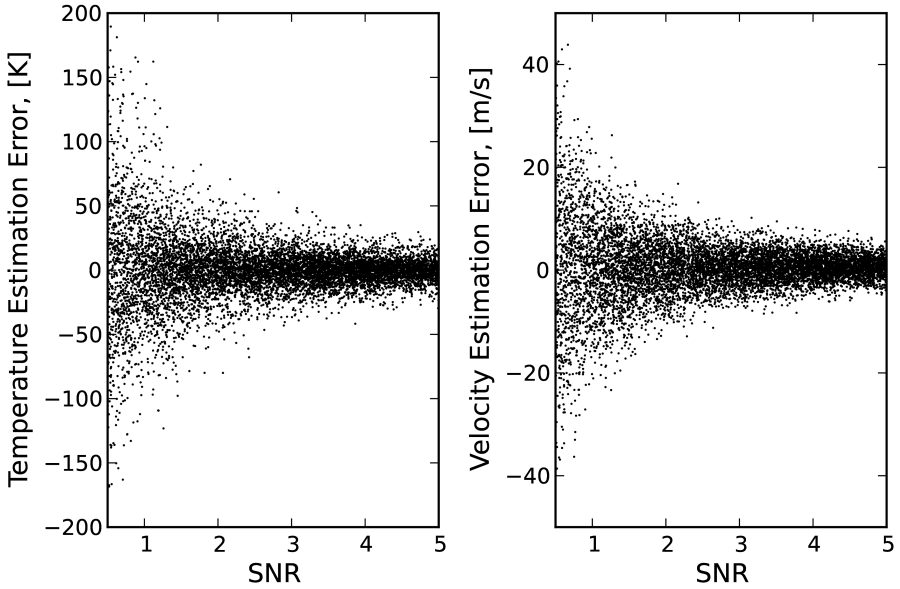

Figure 2.14: Errors in wind and temperature estimates for 10^4 simulated FPI inversions with various noise levels, characterized by the SNR. Adapted from *Harding et al.* [2014].

CHAPTER 3

WIND FIELD ESTIMATION

In this chapter, we describe a statistical estimation technique which uses line-of-sight wind measurements to estimate not just the three-component wind vector, but also how it varies over latitude and longitude, which we hereafter refer to as a wind field. Consecutive applications of this technique provide the time-dependence of the wind field. This chapter is based on work described by *Harding et al.* [2015].

3.1 Introduction

For many years after the first FPI measurements, FPIs were mostly distributed and operated individually, and were able to provide detailed information on the time-dependence of the wind at one location. Although a single FPI can provide some information on the latitude- and longitude-dependence of the wind by looking at different zenith and azimuth angles, there is always an ambiguity between spatial variations of the vertical wind and spatial variations of the horizontal wind. Furthermore, since FPIs are only sensitive to the line-of-sight component, the variation of the wind transverse to the line-of-sight is unknown, which means that a single station can only obtain information on the divergence of the wind field. No information on the curl (commonly referred to as the vorticity in the atmospheric sciences) is contained in the measurements, regardless of the sampling density and strategy. This means that even for all-sky versions of FPIs which sample many points in the sky simultaneously, known as Scanning Doppler Imagers (SDIs) [*Anderson et al.*, 2012b], additional assumptions need to be made to derive a wind field. For example, it is often assumed that the vertical wind is uniform across the field of view and that zonal gradients in the meridional wind can be determined from temporal changes in the meridional wind un-

der the assumption that the station is rotating under a meridional wind field that is stationary in universal time [*Burnside et al.*, 1981; *Conde and Smith*, 1998; *Anderson et al.*, 2012b].

Recently, researchers have begun to combine multiple instruments together to perform networked observations, as described in a review by *Meriwether* [2006]. With multiple observations of the airglow layer from different ground stations, the ambiguities described above can be resolved. Examples of these networks include an FPI and SDI in Antarctica [*Anderson et al.*, 2011], two SDIs in Alaska [*Anderson et al.*, 2012c], three FPIs in Peru [*Meriwether et al.*, 2008; *Shiokawa et al.*, 2012], three FPIs in Scandinavia [*Aruliah et al.*, 2005], two FPIs in the northeastern United States [*Sipler et al.*, 1995], the RENOIR network comprising two FPIs in Brazil, and the NATION network comprising five FPIs in the eastern half of the United States. The latter two are described in Chapter 2. Given this recent growth, there is a need to develop advanced techniques to combine data from multiple FPIs to estimate the regional wind field.

We denote the wind with u, v, and w, representing the zonal, meridional, and vertical components. The problem of wind field estimation is to determine $u(x, y)$, $v(x, y)$, and possibly also $w(x, y)$ from the measurements, where x is positive eastward, y is positive northward, the altitude is a constant, typically 250 km, and the domain spans the combined field of view of all instruments being used. To determine the time-dependence of the wind field, the estimation problem is repeated at different times throughout the night.

Much of the previous work on wind field estimation has assumed that $u(x, y)$ and $v(x, y)$ take a predetermined functional form, such as polynomials [*Burnside et al.*, 1981; *Conde and Smith*, 1998; *Greet et al.*, 1999; *Meriwether et al.*, 2008; *Anderson et al.*, 2011]. These authors expanded u and v in a Taylor series, truncated at the first or second order, and chose the coefficients that most closely matched the data, an idea originally put forth by *Browning and Wexler* [1968] to estimate wind gradients in the lower atmosphere from Doppler radar. This method resolves the fundamental problem of wind field estimation, which is that the data alone are too sparse and incomplete to uniquely determine a wind field. However, the limitation of this technique is that the wind field is constrained to take a polynomial form, which

can smooth over structures in the real wind field and potentially introduce artificial structure where measurements are sparse or noisy.

In this work, we take an inverse-theoretic approach, finding the wind field that agrees with the measurements to within their uncertainties, while also producing the least "structured" wind field possible, where we define "structure" using gradient and curvature operators, not an a priori declaration of the wind field's functional form. Inverse theory is widely used in geophysics [*Aster et al.*, 2013; *Menke*, 1989], but its introduction to aeronomy has been slow. There are a few examples of inverse theory applied to radio observations of ionospheric parameters [*Hysell and Chau*, 2006; *Hysell*, 2007; *Nicolls et al.*, 2014], and similar estimation techniques are use in magnetosphere-ionosphere coupling problems in the polar regions [e.g., *Cosgrove et al.*, 2014; *Cousins et al.*, 2015; *McGranaghan et al.*, 2016]. For estimation of the neutral atmosphere, and specifically the wind, inverse theoretic inversions have been used experimentally by the team analyzing Wind Imaging Interferometer (WINDII) data on the Upper Atmosphere Research Satellite [*Rochon*, 2000], and by *Hysell et al.* [2014] to estimate lower thermospheric wind profiles using Arecibo data.

In the remainder of this chapter, we present the technique in Section 3.2, followed by a discussion of the choice of the structure operators in Sections 3.3 and 3.4, using NATION data as an example. Section 3.5 characterizes the error of the wind field estimate. To test this method, we apply it to synthetic data in Section 3.6. Finally, in Section 3.7, we use this technique on RENOIR data to study the wind dynamics associated with the midnight temperature maximum.

3.2 Basic Formulation

A line-of-sight wind measurement is depicted in Figure 1.2. We assume a network of instruments provides M line-of-sight wind measurements, and we consider the i^{th} measurement, which has a zenith angle θ_i (defined as 0 for a zenith-looking measurement, increasing downward), and azimuth angle ϕ_i (defined as 0 for a northward-looking measurement and 90° for an eastward-looking measurement). The observation equation for the i^{th} line-of-sight wind measurement, d_i, defined as positive away from the instrument, can be

written as

$$d_i = u(x_i, y_i) \sin\theta_i \sin\phi_i + v(x_i, y_i) \sin\theta_i \cos\phi_i + w(x_i, y_i) \cos\theta_i \tag{3.1}$$

where (x_i, y_i) denotes the location (in units of kilometers) where the line of sight intersects the airglow layer at 250-km altitude, which can be calculated using a spherical-Earth approximation:

$$\gamma_i = \sin^{-1}\left(\frac{R_E}{R_E + h} \sin\theta_i\right) \tag{3.2}$$

$$x_i = x_i^0 + (R_E + h)\,(\theta_i - \gamma_i) \sin\phi_i \tag{3.3}$$

$$y_i = y_i^0 + (R_E + h)\,(\theta_i - \gamma_i) \cos\phi_i \tag{3.4}$$

where (x_i^0, y_i^0) is the location of the instrument (typically defined so that the center of the network is zero), R_E is the radius of the Earth (6371 km), h is the assumed emission altitude (250 km), and the intermediate variable, γ_i, is the local zenith angle of the line-of-sight at its intersection with the emission layer. It is more correct to use γ_i instead of θ_i in (3.1) in order to correct the projection for the slight curvature of the Earth. However, the difference is negligible for $\theta_i < 70°$, so we ignore it since high-zenith-angle observations are not made due to the presence of obstacles near the horizon and large atmospheric absorption and scattering. There is a similar, yet smaller, modification for ϕ_i at (x_i, y_i) which is ignored. Although not explicitly notated, (3.1) is valid for a single time, and it is assumed that all d_i have been interpolated to this time. It is imagined that multiple instruments will be used, and each will have multiple different directions contributing to the M total measurements.

To formulate the problem to estimate the wind field, we first discretize the domain into a high-resolution grid of $N \times N$ pixels. As long as N is chosen large enough so that the pixel spacing is much smaller than the measurement spacing, the results are not sensitive to N. In this work we use $N = 40$. The estimation problem is simply to find the $3N^2$ values which define u, v, and w at every pixel.

A single measurement, d_i, contains information from only one pixel, but the measurement contains contributions from u, v, and w, which are impossible to distinguish with one measurement alone. Any wind field estimation

technique must use multiple measurements to simultaneously interpolate between measurements and untangle the contributions to each measurement. We thus consider all measurements together by stacking M instances of (3.1) into a single matrix equation:

$$\mathbf{d} = \mathrm{A}\mathbf{u} \tag{3.5}$$

where the column vector $\mathbf{d}$ contains the measurements, the column vector $\mathbf{u}$ contains the unknown zonal, meridional, and vertical wind at every pixel, and the sparse, $M \times 3N^2$ matrix A (not to be confused with A from Chapter 2) implements (3.1):

$$\mathbf{d} = \begin{bmatrix} d_0 \\ d_1 \\ \vdots \\ d_{M-1} \end{bmatrix} \tag{3.6}$$

$$\mathbf{u} = \begin{bmatrix} u_0 \\ u_1 \\ \vdots \\ u_{N^2-1} \\ v_0 \\ v_1 \\ \vdots \\ v_{N^2-1} \\ w_0 \\ w_1 \\ \vdots \\ w_{N^2-1} \end{bmatrix} \tag{3.7}$$

$$\mathrm{A} = \begin{bmatrix} 0 & 0 & s\theta_0 s\phi_0 & 0 & \dots & 0 & 0 & s\theta_0 c\phi_0 & 0 & \dots & 0 & 0 & c\theta_0 & 0 \\ s\theta_1 s\phi_1 & 0 & 0 & 0 & \dots & s\theta_1 c\phi_1 & 0 & 0 & 0 & \dots & c\theta_1 & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \end{bmatrix} \tag{3.8}$$

where $c \equiv \cos$ and $s \equiv \sin$. Since each d_i depends on the values from only one pixel, the only columns of A with nonzero entries are those corresponding to the pixel containing (x_i, y_i) for each row. Since most pixels are unobserved,

there are many columns containing only zeros. To avoid storing and computing with many zeros unnecessarily, we store A using the `scipy.sparse` Python library and perform all subsequent computations using its sparse matrix routines.

In all practical situations, $M \ll 3N^2$, so A is "short and fat," and the problem is extremely underdetermined, meaning there is no unique solution based purely on the data. Of the infinite number of wind fields that could explain the data, we must choose one by incorporating prior information. We formulate the problem as one of constrained optimization, choosing the "smoothest" wind field that matches the data to within their uncertainty:

$$\begin{aligned} &\text{minimize} \quad r(\mathbf{u}) \\ &\text{such that} \quad \|\Sigma^{-\frac{1}{2}}(\mathrm{A}\mathbf{u} - \mathbf{d})\|_2^2 \leq \epsilon. \end{aligned} \tag{3.9}$$

The function $r(\mathbf{u})$ measures the "roughness" (i.e., inverse smoothness) of the wind field $\mathbf{u}$ and will be discussed further in Sections 3.3 and 3.4. The operator $\|\cdot\|_2$ is the vector 2-norm, and ϵ is a tuning parameter. The $M \times M$ matrix Σ is the measurement covariance matrix. Multiplying by $\Sigma^{-\frac{1}{2}}$ whitens the errors so that the vector $\Sigma^{-\frac{1}{2}}(\mathrm{A}\mathbf{u} - \mathbf{d})$ consists of uncorrelated random variables with unit variance. Under the assumption that the errors are distributed as Gaussians, then $\|\Sigma^{-\frac{1}{2}}(\mathrm{A}\mathbf{u} - \mathbf{d})\|_2^2$ follows a χ^2 distribution with M degrees of freedom.

In practice, we make the assumption that errors are uncorrelated and Σ is diagonal. This is reasonable when statistical errors dominate. However, when systematic errors such as OH contamination or rapid etalon fluctuations dominate, we may expect errors from a single instrument to be significantly non-Gaussian and correlated. In this case, the reconstructed wind field could contain artifacts of systematic errors, such as a divergence centered over a single instrument. We populate the diagonals of Σ with the variances (squared uncertainties) from the output of the analysis algorithm described in Chapter 2.

Hysell et al. [2014] notes that optimization problems like that in (3.9) can be thought of as an implementation of Occam's Razor. Of all possible solutions that fit the data to some desired level of fidelity (as set by ϵ), we choose the simplest. The choice of ϵ significantly affects the reconstruction. If ϵ is too small, the reconstructed wind field will match the data exactly, and

it is susceptible to artifacts caused by noise. Choosing ϵ too large will result in a wind field that is too smooth and does not take full advantage of the information contained in the data. We choose to set ϵ equal to the expectation of $\|\Sigma^{-\frac{1}{2}}(\mathrm{A}\mathbf{u} - \mathbf{d})\|_2^2$, which is M. This has the intuitive interpretation of requiring the wind field to match the measurements to within one standard error, in an average sense.

3.3 Curvature Regularization

All that remains is to choose the roughness metric, $r(\mathbf{u})$. We desire a metric that satisfies two criteria. First, it should match our intuitive sense of the complexity or unlikeliness of wind fields, and second, it should allow for efficient solution of (3.9). To satisfy the second criterion, we choose a metric that can be expressed as the norm of a linear operator on $\mathbf{u}$:

$$r(\mathbf{u}) = \|\mathrm{C}\mathbf{u}\|_2^2 \tag{3.10}$$

where C is a matrix of our choosing which measures the roughness at every pixel, and we seek to minimize the total roughness, added in quadrature over all the pixels. An obvious choice for C, and one that has been used in aeronomical inverse problems by *Hysell et al.* [2014], *Nicolls et al.* [2014], and *Cosgrove et al.* [2014], is the curvature operator. In the simplest case of a scalar function of one dimension, the curvature operator can be written as

$$C_{1\mathrm{D}} = \begin{bmatrix} -1 & 2 & -1 & 0 & 0 & \dots & 0 & 0 & 0 \\ 0 & -1 & 2 & -1 & 0 & \dots & 0 & 0 & 0 \\ 0 & 0 & -1 & 2 & -1 & \dots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & 0 & \dots & -1 & 2 & -1 \end{bmatrix}. \tag{3.11}$$

In our case, we have three scalar functions of two dimensions each, so we construct C as a block matrix which measures all second derivatives of u, v,

and w:

$$\mathrm{C} = \begin{bmatrix} \mathrm{C}^u_{xx} \\ \mathrm{C}^u_{xy} \\ \mathrm{C}^u_{yx} \\ \mathrm{C}^u_{yy} \\ \mathrm{C}^v_{xx} \\ \mathrm{C}^v_{xy} \\ \mathrm{C}^v_{yx} \\ \mathrm{C}^v_{yy} \\ \mathrm{C}^w_{xx} \\ \mathrm{C}^w_{xy} \\ \mathrm{C}^w_{yx} \\ \mathrm{C}^w_{yy} \end{bmatrix} \tag{3.12}$$

where the notation C^u_{xx} describes a matrix that measures $\frac{\partial^2 u}{\partial x^2}$ at each pixel:

$$\mathrm{C}^u_{xx} = \frac{1}{(\Delta x)^2} \begin{bmatrix} -1 & 2 & -1 & 0 & \dots & 0 & 0 & 0 & 0 & 0 & \dots & 0 & 0 & \dots \\ 0 & -1 & 2 & -1 & \dots & 0 & 0 & 0 & 0 & 0 & \dots & 0 & 0 & \dots \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots \\ 0 & 0 & 0 & 0 & \dots & -1 & 2 & -1 & 0 & 0 & \dots & 0 & 0 & \dots \end{bmatrix} \tag{3.13}$$

and the other matrices are constructed analogously. Mathematically, this choice of C penalizes the squared Frobenius norms of the Hessian matrices of u, v, and w, added in quadrature over all the pixels and over u, v, and w. More intuitively, this choice of C penalizes any non-planar features of u, v, or w. As with A, C can be stored and used efficiently using sparse matrix routines.

With a simple application of Lagrange multipliers, (3.9) and (3.10) can be converted from a constrained to an unconstrained formulation:

$$\underset{\mathbf{u}}{\text{minimize}} \qquad \|\Sigma^{-\frac{1}{2}}(\mathrm{A}\mathbf{u} - \mathbf{d})\|_2^2 + \lambda_0 \|\mathrm{C}\mathbf{u}\|_2^2. \tag{3.14}$$

Now the problem takes the form of Tikhonov regularization [*Aster et al.*, 2013]. The solution balances the importance of two terms: the data cost on the left, which is small when the wind field matches the data, and the curvature cost on the right, which is small when the wind field is smooth. The balance between these two terms is set by the Lagrange mul-

tiplier λ_0, which is known as a regularization parameter in the context of Tikhonov regularization. It must be chosen so that the solution, $\mathbf{u}^*$, satisfies $\|\Sigma^{-\frac{1}{2}}(\mathrm{A}\mathbf{u}^*-\mathbf{d})\|_2^2 = \epsilon$. The solution to (3.14) can be found analytically [*Aster et al.*, 2013], which is the primary reason we chose a linear operator for $r(\mathbf{u})$:

$$\mathbf{u}^* = \left[\mathrm{A}^{\mathrm{T}}\Sigma^{-1}\mathrm{A} + \lambda_0\mathrm{C}^{\mathrm{T}}\mathrm{C}\right]^{-1}\mathrm{A}^{\mathrm{T}}\Sigma^{-1}\mathbf{d}. \tag{3.15}$$

We find the correct value for λ_0 using a bisection search, since the data cost increases monotonically with λ_0. First, we must find two values of λ_0 which bound the solution. This is done by starting with an arbitrarily small value (10^{-5} in practice), and evaluating the solution. This is a wind field that matches the data almost exactly, but has high curvature cost. By factors of 10, we increase λ_0 until the data cost exceeds ϵ. Having found two values of λ_0 that straddle the solution, we converge to the optimal λ_0 using the bisection method.

In converting (3.9) to (3.14), we have assumed that the solution to (3.9) lies on the boundary of the set of solutions satisfying the inequality, i.e., that the solution to (3.9) satisfies $\|\Sigma^{-\frac{1}{2}}(\mathrm{A}\mathbf{u}^*-\mathbf{d})\|_2^2 = \epsilon$. This is usually true. The only situation in which the solution could lie in the interior ($\|\Sigma^{-\frac{1}{2}}(\mathrm{A}\mathbf{u}^*-\mathbf{d})\|_2^2 < \epsilon$) is when better agreement with the data causes no additional roughness. This occurs when the wind field is truly planar, or when the data are not of sufficient density or quality to distinguish it from a plane. Although rare, we handle this case by initially evaluating (3.15) using an extremely large value of λ_0 (10^8 in practice, which is 10 orders of magnitude larger than typical values), which recovers a solution with negligibly small curvature, essentially equivalent to fitting planes to u, v, and w. If this solution has a data cost less than ϵ, then it is the solution to (3.9). If not, the solution to (3.9) is on the boundary of the feasible set, and the optimal value of λ_0 is found using the algorithm described above.

We apply this technique to NATION data from 15 June 2014. The line-of-sight wind data are linearly interpolated to the reconstruction time of 04:35 UT. On this night, the five instruments composing NATION were all operational with high data quality, and the measurements which contributed to the interpolation were all taken within 7.5 minutes of 04:35 UT, which was a period of clear skies and low geomagnetic activity.

Figure 3.1 shows the estimated wind field. The color represents the vertical

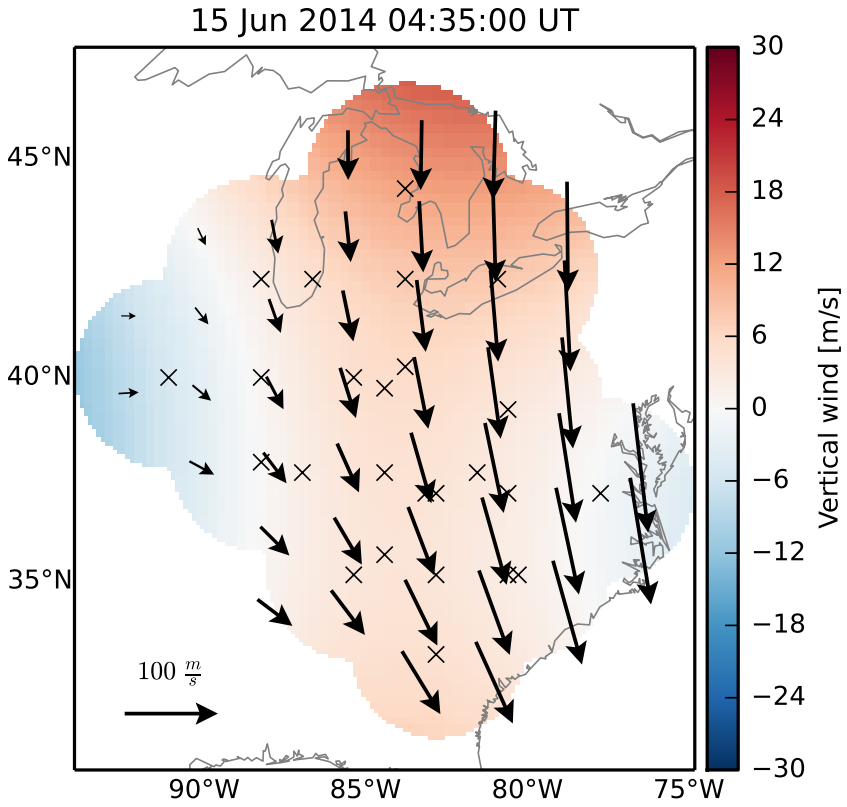

Figure 3.1: Thermospheric wind field over the eastern United States, estimated from NATION data using curvature regularization. (Color) vertical wind, (arrows) horizontal wind, (crosses) measurement locations. Adapted from *Harding et al.* [2015].

wind estimate, and the arrows represent the horizontal wind estimate. The small crosses indicate the location of the measurements' intersections with the airglow layer at 250 km. Thin gray lines are coastlines, shown for reference. Although we reconstruct the wind field over the entire rectangular region, we only display the pixels within 250 km of a measurement location, to limit extrapolation. In the future, we could improve the algorithm so that these pixels are the only ones that are included in $\mathbf{u}$.

The estimated vertical wind varies from -11 to +18 m/s. The meridional wind varies mostly zonally, from 0 m/s in the west to -120 m/s in the east. The zonal wind varies mostly meridionally, from 40 m/s in the south to 0 m/s in the east. Clearly, the variations in the wind are dominated by shear

and not stretching. This behavior is expected from physical reasoning; diverging or converging winds require a source or sink of momentum or energy, typically from heating and/or vertical winds. However, as discussed above, FPIs are much more sensitive to divergent features than shear features, so it is somewhat suspicious that the features most apparent in the reconstruction are those least observable. We are left to wonder whether these features are truly demanded by the data, or whether they are artificially created by the reconstruction algorithm.

One problem is that even a zero-curvature wind field can have large divergences and shears. Intuitively, we consider divergences and shears to be structures which should be required by the data rather than being arbitrarily imposed. Thus, we generalize our notion of "roughness" from second-order derivatives to include first-order derivatives as well, which we refer to collectively as gradients.

3.4 Curvature and Gradient Regularization

In this section, we use a different roughness metric, which includes both gradient and curvature terms:

$$r(\mathbf{u}) = \|\mathrm{G}\mathbf{u}\|_2^2 + \lambda_1 \|\mathrm{C}\mathbf{u}\|_2^2 \tag{3.16}$$

where the matrix G measures gradients analogously to C, and λ_1 is a tuning parameter, the value of which will be discussed later. Like C, G is constructed as a block matrix:

$$\mathrm{G} = \begin{bmatrix} \mathrm{G}_x^u \\ \mathrm{G}_y^u \\ \mathrm{G}_x^v \\ \mathrm{G}_y^v \\ \mathrm{G}_x^w \\ \mathrm{G}_y^w \end{bmatrix} \tag{3.17}$$

where the notation G_x^u describes a matrix that measures $\frac{\partial u}{\partial x}$ at each pixel:

$$G_x^u = \frac{1}{\Delta x}\begin{bmatrix} 1 & -1 & 0 & \dots & 0 & 0 & 0 & 0 & \dots & 0 & 0 & \dots \\ 0 & 1 & -1 & \dots & 0 & 0 & 0 & 0 & \dots & 0 & 0 & \dots \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots \\ 0 & 0 & 0 & \dots & 1 & -1 & 0 & 0 & \dots & 0 & 0 & \dots \end{bmatrix} \tag{3.18}$$

and the other gradient matrices are constructed analogously. Mathematically, this choice of G penalizes the magnitudes of the gradients of the functions u, v, and w.

With the addition of the gradient penalty, it may not be obvious why the curvature penalty is still needed. After all, any region that has large curvature also has a large gradient. However, the use of a purely gradient-based penalty term leads to reconstructions of functions that are pointy, much like a circus tent stretched over poles. The curvature term penalizes this pointy behavior.

The unconstrained formulation with this new roughness metric is a direct generalization of (3.14):

$$\underset{\mathbf{u}}{\text{minimize}} \qquad \|\Sigma^{-\frac{1}{2}}(\mathrm{A}\mathbf{u} - \mathbf{d})\|_2^2 + \lambda_0 \left(\|\mathrm{G}\mathbf{u}\|_2^2 + \lambda_1 \|\mathrm{C}\mathbf{u}\|_2^2\right) \tag{3.19}$$

and its solution is

$$\mathbf{u}^* = \left[\mathrm{A}^{\mathrm{T}}\Sigma^{-1}\mathrm{A} + \lambda_0 \left(\mathrm{G}^{\mathrm{T}}\mathrm{G} + \lambda_1 \mathrm{C}^{\mathrm{T}}\mathrm{C}\right)\right]^{-1} \mathrm{A}^{\mathrm{T}}\Sigma^{-1}\mathbf{d} \tag{3.20}$$

where, for a constant λ_1, λ_0 is chosen as described in the previous section.

The seemingly arbitrary choice of λ_1 is a disadvantage of adding the gradient penalty. A λ_1 that is too large will not penalize gradients enough, and it will have the problems discussed in Section 3.3. A λ_1 that is too small will not penalize curvatures enough, and the solution will exhibit pointy behavior. The optimal λ_1 is somewhere in the middle, but to our knowledge there does not appear to be any theoretical justification for choosing any particular value. One popular method for choosing tuning parameters is Generalized Cross Validation [*Aster et al.*, 2013], but for this problem it failed to produce a reliable minimum. Another popular method is the L-curve criterion [*Hansen and O'Leary*, 1993], but this also failed because the plot of gradient cost against curvature cost is not L-shaped in general.

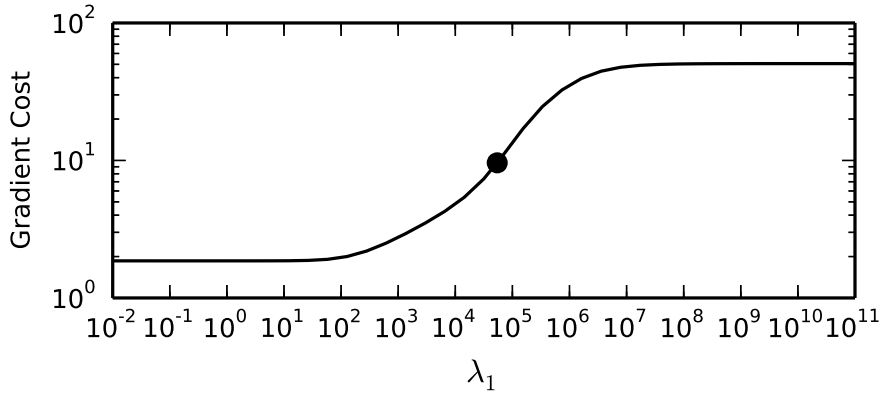

Figure 3.2: The gradient cost of the solution to (3.19) for different values of the tuning parameter λ_1. The algorithm sets the desired gradient cost equal to the geometric mean of its minimum and maximum value, shown as a circle. Adapted from *Harding et al.* [2015].

We developed an ad hoc criterion from our observation that the dependence of the gradient cost, $\|\mathrm{G}\mathbf{u}^*\|$, on λ_1 reliably looks like that shown in Figure 3.2. The trace in Figure 3.2 was generated by solving (3.19) for various values of λ_1, using the same data as used for Figure 3.1. As λ_1 grows large, the gradient cost asymptotes to its maximum value. As λ_1 grows small, the gradient cost achieves a minimum value. Our algorithm aims for a gradient cost that is the geometric mean of these two extremes, shown as a circle in Figure 3.2. This is found by a bisection search identical to the one used for λ_0. Within each iteration of this outer search for λ_1 is an inner search for λ_0. We have found that this choice of λ_1 robustly yields wind fields that are smooth.

We note that *Nicolls et al.* [2014] describes a similar algorithm for estimating ion drift patterns (or equivalent electric fields) from coherent radar measurements. They also use a mixture of gradient and curvature penalties, but they impose the curvature penalty inside the convex hull of their measurement locations and impose the gradient penalty outside, to limit extrapolations while keeping a pure curvature penalty in the interior. In contrast, we treat every pixel identically. While the *Nicolls et al.* [2014] approach eliminates the need to choose λ_1 explicitly, they still must have some scaling between gradient and curvature penalties, which they implicitly set to 1.

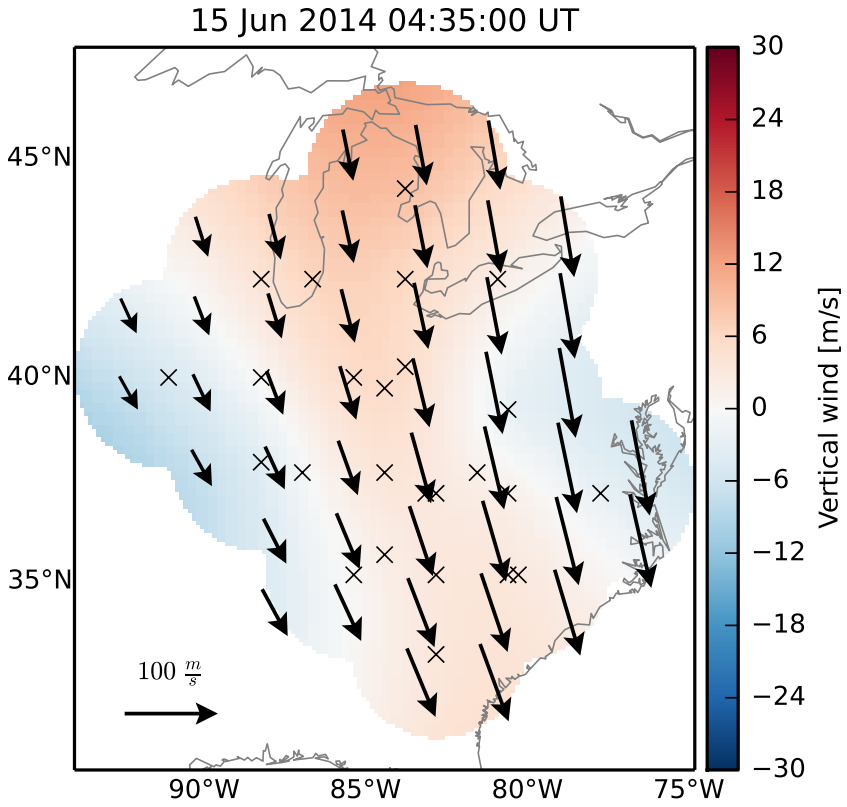

Figure 3.3: Same as Figure 3.1, but with an added penalty on gradients, which recovers a smoother wind field. Adapted from *Harding et al.* [2015].

An example reconstruction resulting from this technique is shown in Figure 3.3, which uses the same data as in Figure 3.1. As in Figure 3.1, Figure 3.3 displays a wind field with more shear than divergence. However, the gradients are reduced. The estimated vertical wind is smaller, varying from -9 to +12 m/s. The zonal wind still varies mostly meridionally but only from 30 to 12 m/s (instead of 40 to 0 m/s), and the meridional wind varies zonally from -40 to -100 m/s (instead of 0 to -120 m/s). Both Figures 3.1 and 3.3 display wind fields that agree equally well with the data, but Figure 3.3 displays a wind field that is smoother, in the intuitive sense of the word.

When the FPI network is small, such as RENOIR, or when some of the sites are clouded over or experiencing maintenance issues, the vertical wind estimation can become ill-posed. Especially for RENOIR, the two sites are clustered near the middle of the network's field of view, so the only two direct

vertical wind measurements are closely spaced. The estimation of the vertical wind field is analogous to balancing a plate on two fingers near the center of the plate. The data contain little information on vertical wind variations perpendicular to the inter-site line, and variations parallel to the line suffer from extrapolation errors. While it is true that some information regarding the vertical wind is contained in the oblique measurements near the edge of the field of view, this information is entangled with information on horizontal wind variations, which are difficult to untangle with only two sites (i.e., 10 measurements).

For these cases, we have developed an option to neglect variations in the vertical wind field, and instead assume it is constant with respect to x and y. The constant value is determined from the average of the direct, zenith-looking vertical wind measurements, weighted by their uncertainties. This new constrained inversion problem can be transformed into a problem of the same form as the one described above. To account for the fact that the vertical wind is being treated as known, we must move it from the right-hand side of (3.5) to the left-hand side, which can be done by simply re-defining the data vector as

$$\mathbf{d}' = \mathbf{d} - \mathrm{A}_w \mathbf{w} \tag{3.21}$$

where $\mathbf{d}'$ is the new data vector, corrected for the contribution of the known vertical wind to each measurement, A_w is the right third of the matrix A, and $\mathbf{w}$ is a vector of length N^2 where every element is the same and is equal to the known vertical wind. After discarding all zenith-looking measurements, we simply remove the right third of A, C, and G and the bottom third of the unknown $\mathbf{u}$ vector. After these modifications, the same equations apply, and the same code can be used. We use this modification for the RENOIR wind fields shown in Section 3.7.

3.4.1 Uniqueness

An advantage of adding the gradient penalty is that uniqueness is guaranteed. Without it, more than one plane could fit the data when uncertainties are large, which means that (3.9) has no unique solution, and the matrix in (3.15) has no inverse.

In this section we sketch a proof that (3.19) always has a unique solution,

except in degenerate cases. First, we rearrange (3.19) into a stacked matrix form:

$$\underset{\mathbf{u}}{\text{minimize}} \quad \left\| \begin{bmatrix} \Sigma^{-\frac{1}{2}}\mathrm{A} \\ \lambda_0 \mathrm{G} \\ \lambda_0\lambda_1 \mathrm{C} \end{bmatrix} \mathbf{u} - \begin{bmatrix} \Sigma^{-\frac{1}{2}}\mathbf{d} \\ 0 \\ 0 \end{bmatrix} \right\|_2^2 . \tag{3.22}$$

This is a convex optimization problem; its minimum is guaranteed to be unique so long as the stacked matrix has full column rank, which is true unless there is a vector that is in the null space of all three matrices: A, G and C. Let us denote the null space of X with $\mathcal{N}(\mathrm{X})$. Formally, the condition for uniqueness is

$$\mathcal{N}(\mathrm{A}) \cap \mathcal{N}(\mathrm{G}) \cap \mathcal{N}(\mathrm{C}) = \emptyset \tag{3.23}$$

which we can simplify to

$$\mathcal{N}(\mathrm{A}) \cap \mathcal{N}(\mathrm{G}) = \emptyset \tag{3.24}$$

because any wind field with zero gradient also has zero curvature, i.e., $\mathcal{N}(\mathrm{G}) \subset \mathcal{N}(\mathrm{C})$.

The null space of G is easy to characterize; it is spanned by three vectors: one with constant u (and $v = w = 0$), one with constant v (and $u = w = 0$) and one with constant w (and $u = v = 0$). If the network makes at least three measurements with lines of sight that are not co-planar, these three projections of the wind field can be used to estimate the constant u, v, and w. This means that the vectors in the null space of G are not in the null space of A, except in degenerate cases with purely co-planar lines of sight. Since we only run this algorithm when we have more than three measurements, and it is never the case that the lines of sight are co-planar, then the solution is unique in all practical cases of interest. If it were not, the matrix in (3.20) would be singular, and the routine would crash.

3.5 Error Analysis

In addition to estimating the wind field, our algorithm provides the error of that estimate. The term "error" can have different meanings in different contexts, and unfortunately it is not always clear what definition an author

has in mind when including error bars on data. This becomes even more ambiguous when an inversion is used on data.

In this work, we refer to the error as

$$\mathbf{e} = \mathbf{u}^* - \mathbf{u}^{\text{true}} \tag{3.25}$$

where $\mathbf{u}^*$ is the estimated wind field and $\mathbf{u}^{\text{true}}$ is the unknown true wind field. Note that because they depend on noisy data, $\mathbf{u}^*$ and $\mathbf{e}$ are both random vectors.

Error is often specified with its first and second central moments. The first central moment is often referred to as the bias, or systematic error:

$$\text{bias} \equiv \mathbb{E}[\mathbf{e}]\,. \tag{3.26}$$

One can imagine repeating the same observations multiple times, where the only change is the realization of noise. If we averaged all of the resulting estimated wind fields together, we would obtain a wind field that is smoother than the true wind field. This difference is the bias. Regularized estimates have bias by construction. It is deliberately imposed in order to reduce the effect of noise and make the problem well-posed. In many geophysical applications, the bias has the effect of smoothing, or reducing the resolution. Measurements are often too sparse and noisy to contain information on small-scale structure, so we should not attempt to reconstruct this structure.

Quantifying the bias comprehensively is impossible; it depends on $\mathbf{u}^{\text{true}}$, which is unknown. *Aster et al.* [2013] and *Menke* [1989] discuss techniques for characterizing the data and model resolution, but these are difficult to apply to our problem due to the point-like nature of the measurements. Some notion of resolution could be gleaned simply from the inter-sample spacing, but this does not account for the effect of noise: noisier data will be more highly regularized and thus have lower resolution. A relative measure of bias is λ_0. When it is large, the wind field is highly smoothed. We have not yet been able to quantify bias sufficiently well, so we do not attempt to report it. A few examples of the bias are shown in the simulations in Section 3.6.

The second central moment of the error is the variance, the square-root of which is often referred to as statistical error or uncertainty. For the random

vector $\mathbf{e}$, the relevant quantity is the covariance matrix:

$$\Sigma_{\mathbf{e}} \equiv \mathbb{E}\left[\mathbf{e}\mathbf{e}^{\mathrm{T}}\right]. \tag{3.27}$$

One major difference which distinguishes different methods of estimating the covariance matrix is whether $\mathbf{u}^{\text{true}}$ is treated as a random variable. On the one hand, $\mathbf{u}^{\text{true}}$ is not random; it has some value. On the other hand, we do not know this value, and when we incorporate prior information to estimate it, we are essentially specifying a probability density function on $\mathbf{u}^{\text{true}}$. The first philosophy leads to a formulation in terms of Tikhonov regularization, and the second philosophy leads to a formulation in terms of Bayes theorem, but no matter the philosophy, the solution (either least-squares or maximum a posteriori under the assumption of a Gaussian prior) is the same, (3.20), if the parameters are chosen identically (see Chapter 11 of *Aster et al.* [2013] for further discussion).

Even though these two philosophies agree on the solution, they disagree on the covariance matrix, since the Bayesian approach must also account for the variance of $\mathbf{u}^{\text{true}}$. From the Tikhonov perspective, the only contributor to the variance of $\mathbf{u}^*$ is noise on $\mathbf{d}$ that is propagated through (3.20), which is easily derived by taking the covariance matrix of both sides. From the Bayesian perspective, the variance of $\mathbf{u}^*$ is the posterior covariance matrix:

$$\Sigma_{\mathbf{u}^*} = \left[\mathrm{A}^{\mathrm{T}}\Sigma^{-1}\mathrm{A} + \lambda_0\left(\mathrm{G}^{\mathrm{T}}\mathrm{G} + \lambda_1\mathrm{C}^{\mathrm{T}}\mathrm{C}\right)\right]^{-1}. \tag{3.28}$$

We have found that the posterior covariance gives slightly higher error in practice. For this reason, and to follow *Nicolls et al.* [2014], we use it in this work. The uncertainty of the wind field is extracted from the square root of the diagonal elements of $\Sigma_{\mathbf{u}^*}$. Because it involves a matrix inversion, the computation of the uncertainty increases the algorithm's runtime by an order of magnitude, so is usually omitted in practice.

Figure 3.4 shows the uncertainty of the wind field in Figure 3.3. Instead of showing the individual uncertainties of u, v, and w, we show the magnitude of the uncertainty, $\sqrt{\sigma_u^2 + \sigma_v^2 + \sigma_w^2}$, where σ_u, σ_v, σ_w are the uncertainties in u, v, and w, respectively. The horizontal uncertanties, σ_u and σ_v, are of similar magnitude, while σ_w is generally about 30% smaller, owing to the fact that it is derived directly from a measurement. The uncertainty increases from the

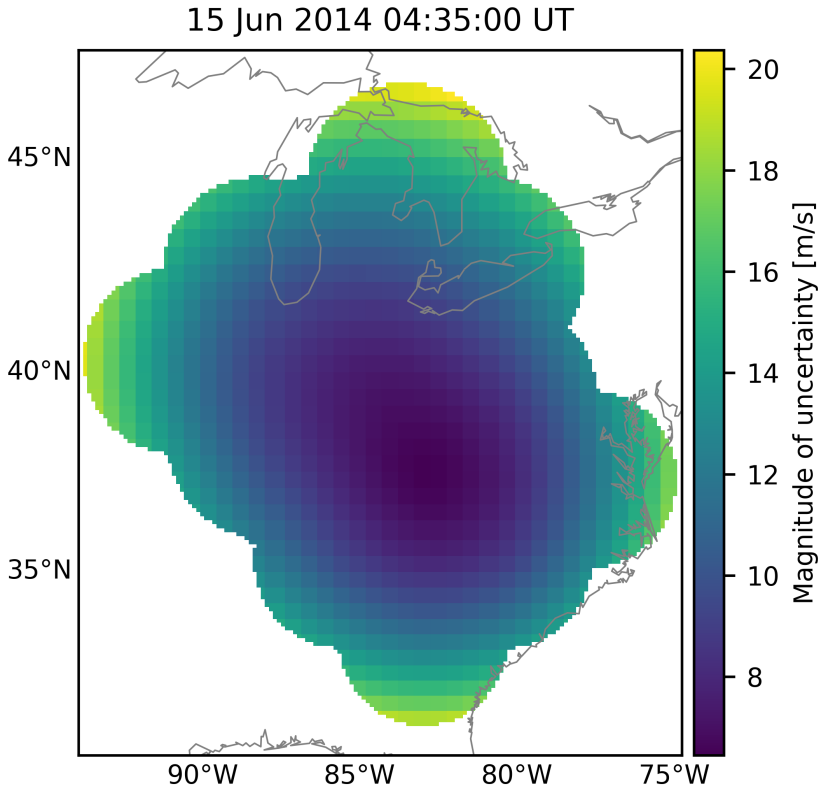

Figure 3.4: The uncertainty (square root of posterior covariance) of the wind field in Figure 3.3, displayed as the total magnitude of the uncertainty, $\sqrt{\sigma_u^2 + \sigma_v^2 + \sigma_w^2}$.

middle of the domain (∼6 m/s) to the edge of the domain (∼20 m/s) where it is less constrained by the data, as expected. The average uncertainty of the line-of-sight data that contributed to this inversion is 8.4 m/s (root mean square).

3.6 Simulation Results

In this section, we test the wind field algorithm with synthetic data generated from two wind field models. In the first, we use the Horizontal Wind Model 2014 (HWM14) [*Drob et al.*, 2015], which assumes the vertical wind is zero. We evaluate this model at an altitude of 250 km and a time of 04:35 UT

on 15 Jun 2014 to correspond to the data in Figure 3.3. We use (3.1) to simulate the data. The instrument locations and lines of sight are chosen to be identical to those used in Figure 3.3. Independent, identically distributed Gaussian noise is added to the measurements with variance σ^2, which takes various values as discussed below. In the inversion, it is assumed that σ^2 is known exactly. The formulation with gradient and curvature penalties is used.

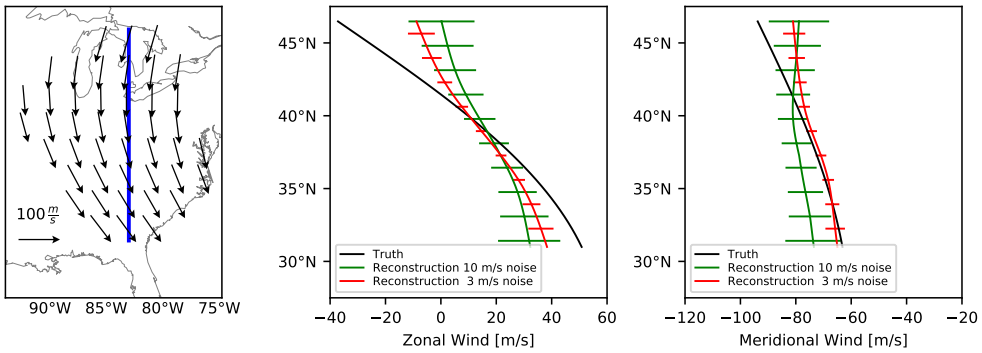

Figure 3.5: Simulation results using the NATION network. (Left) Truth model (HWM14) used to simulate the data, along with the latitudinal cut shown in the next two panels. (Middle) Reconstructed zonal wind for various noise levels, compared to the truth. (Right) Same as middle, for the meridional wind. Adapted from *Harding et al.* [2015] with modifications.

Figure 3.5 shows the results of the simulation using HWM14. The left panel shows the true wind field, and the right two panels show the reconstructed wind along a cut of constant longitude (i.e., the cut shown in blue in the left panel), for two different values of σ, 10 and 3 m/s. The middle panel shows the zonal wind and the right panel shows the meridional wind. A cut of true wind field is also shown for reference in these two panels.

Qualitatively, the agreement between the true and reconstructed wind fields provides confidence that the algorithm is correctly minimizing (3.19). Quantitatively, we see good agreement near the center of the field of view, and disagreement in the north and south, where the measurements are essentially extrapolated. As desired, the reconstructed wind field is smoother than the true wind field. Apparently, there is no evidence in the data for the large departure of the zonal wind to -50 m/s in the north. For both the meridional and zonal wind, the lower-noise data (red line) create a wind field more closely matching the truth than the higher-noise data (green line). As

expected, the algorithm is able to better capture meridional variations of the meridional wind (i.e., a divergence feature) than meridional variations of the zonal wind (i.e., a shear feature).

The error bars on the estimate increase as the data get noisier, but they do not include the truth, especially for the zonal wind. This is a representation of the smoothing bias introduced by regularization. The bias component of the total error in the 3 m/s case is evidently larger than the statistical component of the total error.

For our second synthetic wind field, we introduce an artificially large zonal shear, in order to test the algorithm's response to a wind field which is not smooth. In this case, we expect the bias to be larger. The zonal wind varies from -50 m/s north of 38 °N to +50 m/s south of 38 °N. The reconstructed wind field is shown in Figure 3.6 in the same format as Figure 3.5. For both noise levels, the general morphology of the shear is captured, but it is smoothed over several degrees of latitude, exhibiting the smoothing bias introduced by the limited resolution of the NATION network. The smoothing is more severe in the 10 m/s case, as expected. The meridional wind is captured well.

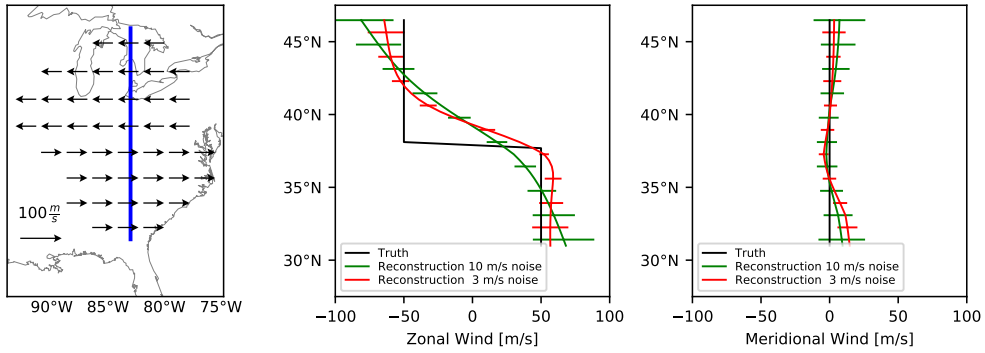

Figure 3.6: Same as Figure 3.5, except for a true wind field with a large shear. Adapted from *Harding et al.* [2015] with modifications.

3.7 Case Study: Wind Dynamics and the Equatorial Midnight Temperature Maximum

In this section, we apply the wind field estimation technique described above to actual data from the RENOIR network in Brazil. Specifically, we study wind dynamics associated with the midnight temperature maximum (MTM), which is a recurring feature of the equatorial and low-latitude thermosphere, characterized by a 25–200 K increase in temperature accompanied with an increase in the airglow brightness, observed to propagate poleward from the equator and occur over a period of 1–2 hours near local midnight [*Spencer et al.*, 1979; *Herrero and Spencer*, 1982; *Meriwether et al.*, 2008]. It is believed that this feature is driven by a converging wind field, possibly caused by the upward propagation of the terdiurnal tide [*Akmaev et al.*, 2010].

Using our wind field estimation technique, we report direct evidence of converging winds during MTM development over Brazil. As mentioned above, for RENOIR we exercise the option to set the vertical wind to a constant, in order to stabilize the horizontal wind estimation problem. Figure 3.7 shows the estimated wind field at four times near local midnight during the night of 01–02 Feb 2013. Both CAR and CAJ data were used, and all cardinal directions were used except the east, which was looking too close to the Moon to provide reliable measurements. The local time (LT) in Brazil is about 3 hours earlier than the universal time (LT = UT – 3).

At 01:30 UT, the flow is mostly eastward and uniform, though there is a slight convergence. This is also the case for the wind fields before 01:30 UT (not shown). At 02:15 UT, a strong convergence develops, carried by both the zonal and meridional components. By 03:00 UT, the strong convergence has abated, the magnitude of the wind has decreased, the flow has begun to turn northward, and the main structures in the flow are shears. By 03:45 UT, the flow has become northeastward and has a slight divergence. The estimated vertical wind is less than 5 m/s at all four times.

We quantify the divergence by approximating it with a finite difference. The horizontal divergence is given by $\left(\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y}\right)$ which we calculate at the centroid of the reconstructed wind field with a 1° finite difference. For this analysis, we estimate the wind field at 5 minute intervals and calculate its divergence at every time step, shown in Figure 3.8 as a black line. A slight, persistent convergence is evident in the first half of the night. Beginning

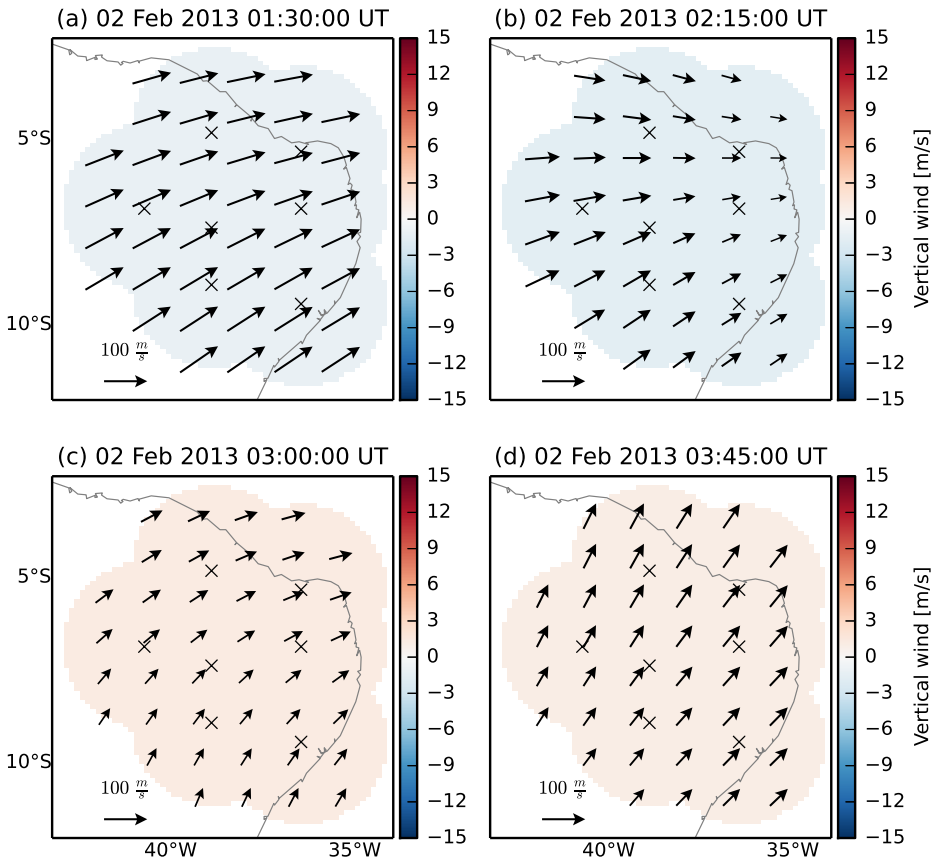

Figure 3.7: Wind field reconstruction using RENOIR data from 02 Feb 2013 at four different times displaying the passage of the wind signature associated with the MTM. The local time is 3 hours earlier than the universal time. Adapted from *Harding et al.* [2015].

around 01:45 UT, a large convergence is seen, almost 0.2 m/s per km. This convergence abruptly ends around 02:45 UT, after which time the wind field becomes slightly divergent for the next couple hours before returning to a slight convergence around 05:00 UT. It is interesting that these estimates indicate a net convergence over the night, which is presumably caused by the predominant day-to-night flow of the thermosphere.

Also shown in Figure 3.8 is the raw, zenith-looking vertical wind measurement from CAR, in red. Measurements after 06:00 UT had sporadic cloud cover and are omitted. We compare the vertical wind with the horizontal

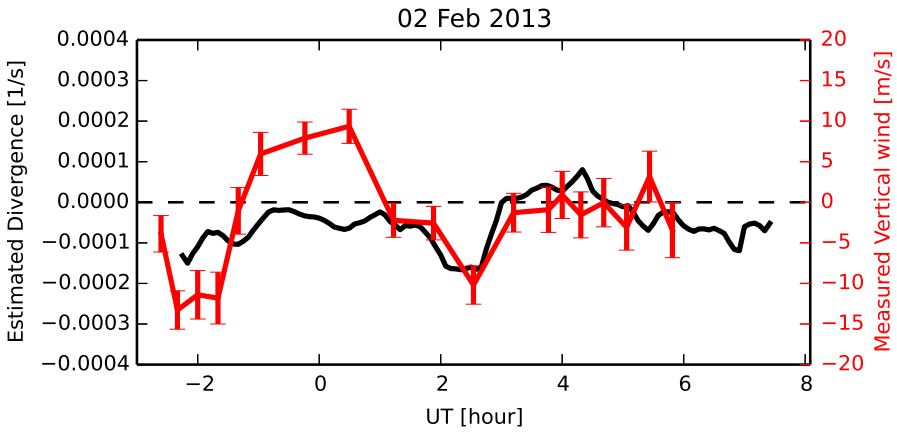

Figure 3.8: (Black, left axis) Divergence calculated from the estimated wind field. (Red, right axis) Raw, zenith-looking, line-of-sight wind measurement from CAR, i.e., the vertical wind. The scaling between the two axes is the scale height, $H = 50$ km. Adapted from *Harding et al.* [2015].

divergence in terms of (1.5), the Burnside relation, which states that under assumptions of hydrostatic equilibrium and incompressible flow, the vertical wind is the scale height, H, times the horizontal divergence. In light of this, $H = 50$ km was chosen as the constant of proportionality for the two axes in Figure 3.8. The agreement between the vertical wind and horizontal divergence is striking, and it provides a physics-based self-consistency check on the quality of the data and estimation algorithm, as the Burnside relation is not built in to the estimation. Three vertical wind samples near 00:00 UT disagree, which might be due to the sparse sampling of the instrument function during this time. Especially notable is the downward wind of 10 m/s during the period of strong convergence, although this only occurs during one vertical wind sample from CAR.

In Figure 3.9, we show the raw temperatures measured by CAR on the same night, in all five directions. We omit measurements during sporadic cloud cover (after 06:00 UT) and with uncertainties larger than 20 K. Temperatures gradually cool throughout the night from 1000 K to 750 K, but superimposed on this trend is a large temperature increase – the classical MTM. In Figure 3.9 we also display the divergence from Figure 3.8. As soon as the winds begin to converge at 01:45 UT, the temperatures begin to increase. When the divergence turns abruptly positive at 02:45 UT, the temperatures begin to decrease again. This correspondence between the con-

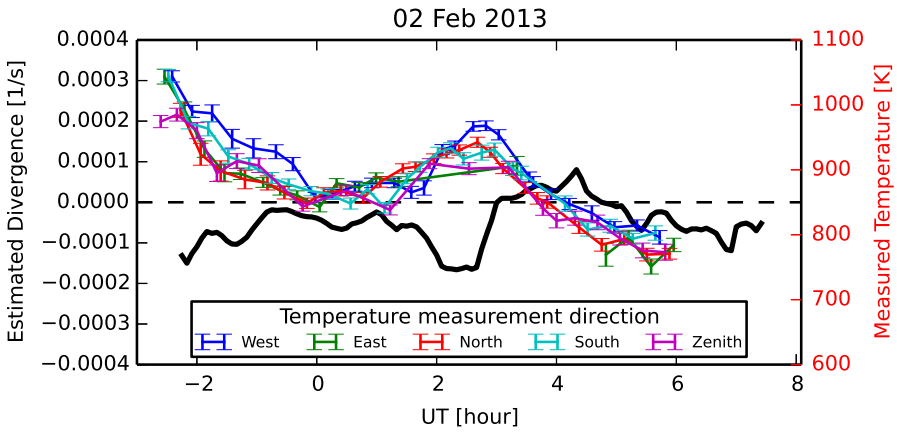

Figure 3.9: (Black, left axis) Same divergence as in Figure 3.8. (Colors, right axis) CAR temperature measurement observing in cardinal directions. Adapted from *Harding et al.* [2015].

vergence and the time-derivative of temperature matches physical expectations about compressional heating, as we briefly show next with a simplified model.

Let p and T denote the pressure and temperature, and let γ be the adiabatic index, which is $\frac{5}{3}$ for atomic oxygen. We assume adiabatic heating:

$$p^{1-\gamma}T^{\gamma} = \text{constant}. \tag{3.29}$$

Taking the time derivative and setting it equal to zero, it can be shown that the rate of change of pressure and temperature are related by

$$\frac{\partial T}{\partial t} = T\left(\frac{\gamma-1}{\gamma}\right)\frac{1}{p}\frac{\partial p}{\partial t} \tag{3.30}$$

where t is time. To relate the rate of change of pressure to the vertical wind, and thus the horizontal convergence, we use the hydrostatic assumption, which implies an exponential pressure profile, as discussed in Chapter 1:

$$p(z) = p_0 \exp\left(-\frac{z-z_0}{H}\right) \tag{3.31}$$

where p_0 is the pressure at an arbitrary reference altitude, z_0. Taking the time derivative, and noting that since we are following a parcel of air, $\frac{\partial z}{\partial t} \equiv w$,

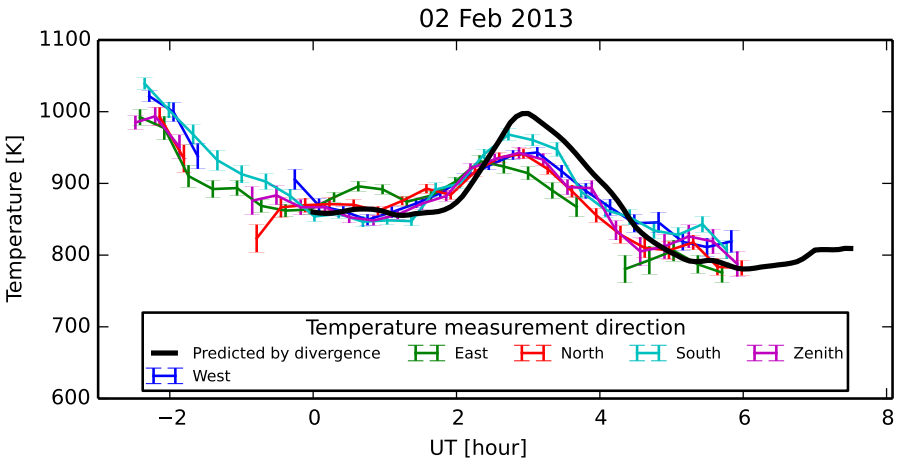

Figure 3.10: Modeled temperature variations driven by the measured divergence, compared to the measured temperatures at CAR.

it can be shown that

$$\frac{1}{p}\frac{\partial p}{\partial t} = -\frac{w}{H} \tag{3.32}$$

$$= -\left(\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y}\right) \tag{3.33}$$

where the second line uses the Burnside relation. Substituting this into (3.30) yields a direct relationship between temperature changes and horizontal wind divergence, which we can use to compare with the data:

$$\frac{\partial T}{\partial t} = -T\left(\frac{\gamma - 1}{\gamma}\right)\left(\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y}\right). \tag{3.34}$$

We discretize (3.34) on a 5-minute cadence and use a simple zero-order forward integrator to model the time-dependence of T from the time-dependence of the estimated horizontal divergence. We use the initial condition that $T = 860$ K at 00:00 UT, which was chosen to match the data. Of course, (3.34) only accounts for heating and cooling caused by convergence and divergence of the horizontal wind. It omits the radiative and chemical processes contributing to the nightly cooling of the thermosphere. For the purposes of this simplified model, we superimpose a linear cooling trend of 500 K every 8 hours, which was chosen to bring the modeled and measured temperatures into agreement at the end of the night.

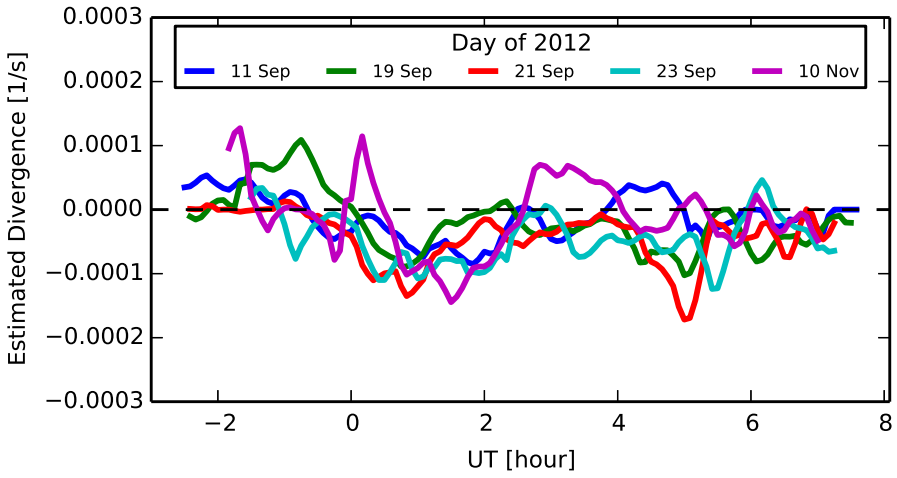

Figure 3.11: Divergence calculated from the estimated wind field for five nights in 2012. A convergence lasting 1–2 hours is seen around the same time every night. Adapted from *Harding et al.* [2015].

The results are shown in Figure 3.10, where the modeled temperature is shown in black, and the measured temperatures are the same as in Figure 3.9. In light of the simplifying assumptions used to derive (3.34), the agreement between the data and model is striking, and it provides another physics-based self-consistency check on the wind field estimation.

So far we have shown data from a single night. In Figure 3.11, we show five more nights of estimated divergences from RENOIR data. These five nights were chosen because they were the only nights in 2012 that had high-quality data from both instruments and used cardinal mode, not common-volume mode. During much of 2012, one instrument had an unreliable laser calibration source, so for the sake of conservatism, these nights are omitted.

The strength and duration of the convergence varies from night to night, but it is always present between about 00:00 UT and 02:00 UT. On all five nights, visual inspection of the temperature data (not shown) indicated the presence of the MTM. We show one additional night of temperature data in Figure 3.12 (from 11 Sep 2012; blue line in Figure 3.11), with the estimated convergence. This is an example of a night with a small MTM, though this feature may be arguably classified as a plateau rather than a maximum. The convergence of the winds on this night is smaller – about half of that in Figure 3.9, suggesting that nights with smaller wind convergence have smaller MTMs.

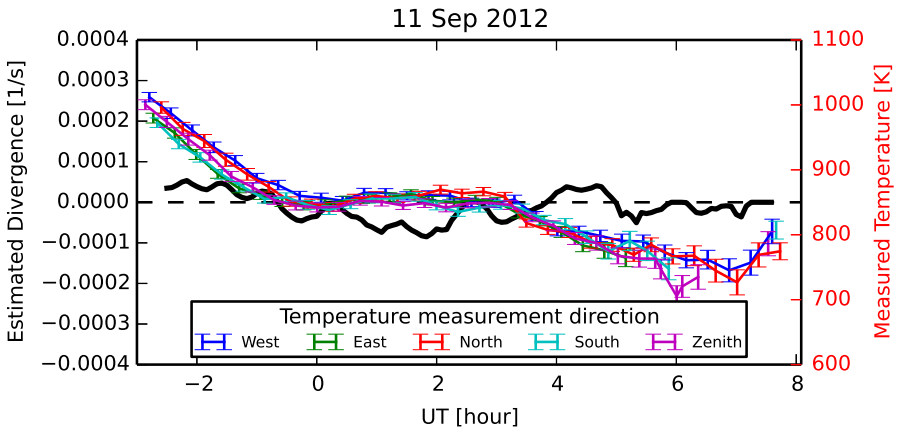

Figure 3.12: Same as Figure 3.9 for a night with a weaker MTM. The wind convergence is also weak. Adapted from *Harding et al.* [2015].

3.8 Conclusion and Future Work

In this chapter we have introduced a technique to estimate the regional x- and y-dependence of the zonal, meridional, and vertical wind, known as a wind field, from its line-of-sight projections. This technique fits a smooth wind field to the line-of-sight measurements, where the notion of smoothness is defined as having minimal gradients and curvatures. The balance between gradient and curvature penalties is set by an ad hoc criterion. The closeness of fit to the measurements is constrained by the measurement uncertainty. The primary advantage of this technique over previous work is that it does not make any explicit assumptions about the functional form of the wind field, and instead lets the data inform the shape.

This technique was applied to study the relationship between the thermospheric wind and temperature during the passage of the MTM in Brazil. Using RENOIR data, we found direct evidence of converging winds during the temperature increase characteristic of the MTM. This was followed by diverging winds as the temperature fell. Other examples show that weaker converging winds produce weaker MTMs, and that the converging winds are persistent feature, occurring in all five nights we investigated in 2012.

There are several possible improvements to this technique. First, to limit the effect of pixels outside the field of view, it would be better to remove them entirely from $\mathbf{u}$, instead of estimating and then ignoring them. This

would require a more complicated indexing scheme and a more sophisticated generation of G and C.

Another improvement is to incorporate the time variable in the inversion. However, this would add computational complexity, require an additional tuning parameter to balance between spatial and temporal structures, and spread the effect of short-lived instrumental fluctuations to the entire night.

It may be possible to use the temperature measurements in the wind inversion. Including a physical constraint into the inversion, such as (3.34), may be able to aid in the estimation of divergence. However, this could be complicated by other sources of temperature changes such as advection and Joule heating.

Our definition of smoothness includes both gradients and curvatures, requiring a parameter, λ_1, to balance between them. The choice of λ_1 is ad-hoc and computationally expensive to calculate. There is potentially an alternative definition of $r(\mathbf{u})$ that matches our intuitive sense of smoothness and is computationally feasible.

One idea is to use models or data to generate a prior distribution for the wind field, and treat the inversion as truly Bayesian. However, there is some benefit to performing the inversion free from constraints created from our current physical understanding, as it allows us to independently verify self-consistency and discover new relationships. It would be worthwhile to try the empirical orthogonal function analysis that is beginning to be used for magnetosphere-ionosphere coupling problems [e.g., *Cousins et al.*, 2015].

With an estimation procedure in hand, it may be possible to optimize the network's observation strategy to minimize bias and uncertainty. Preliminary work has suggested that cardinal mode is not exactly optimal but is close.

Finally, this technique is general and applies to any system making radial velocity measurements. For example, meteor radar systems make mesospheric and lower thermospheric wind measurements by tracking the Doppler shift of the plasma trails of meteors. As these radars begin to be networked together to make coordinated measurements, advantages can be gained by using regularization techniques such as the one described here. We have begun preliminary work applying this algorithm to the meteor radar system described by *Stober and Chau* [2015] with encouraging results. One challenge is runtime; a high-resolution meteor field requires a high-resolution grid, and we must perform the inversion at each altitude. It may also be useful to

generalize from a 2D to 3D inversion, removing the need to artificially bin meteors in altitude. However, the added computational complexity is likely prohibitive.

CHAPTER 4

MIDLATITUDE STORMTIME OBSERVATIONS

In this chapter, we report FPI observations during geomagnetic storms at midlatitude. In Section 4.1, we perform a detailed case study of the 02 Oct 2013 storm using NATION data as well as data from the FPI and all-sky camera at Millstone Hill. These observations show a large equatorward increase in the horizontal wind and an increase in temperature, which are consistent with previous observational results and theory. However, a surprising feature of these observations is the appearance of a large 100 m/s apparent downward wind lasting for several hours. This section is based on work reported by *Makela et al.* [2014].

In Section 4.2, we summarize NATION vertical wind observations from multiple storms over a 5-year period. Significant vertical winds appear during 14 of the 15 storms studied, indicating that the puzzling vertical winds seen during the 02 Oct 2013 storm are not uncommon. This section is based on work reported by *Harding et al.* [Submitted]. The rest of this chapter and Chapter 5 are concerned with explaining these observations.

In Section 4.3, we describe the hypothesis put forward by *Makela et al.* [2014], which implicates contamination by precipitating energetic O^+ particles from the ring current. The observational and theoretical evidence for and against this hypothesis are discussed.

4.1 The Geomagnetic Storm of 02 Oct 2013

On 29 Sep 2013, a coronal mass ejection occurred on the Sun, which passed Earth in the early hours of 02 Oct 2013, initiating a geomagnetic storm. Figure 4.1 shows the Dst and Kp indices during this storm [*Data Analysis Center for Geomagnetism and Space Magnetism, Kyoto University*, 2017; *National Centers for Environmental Information*, 2017]. The Dst index, a

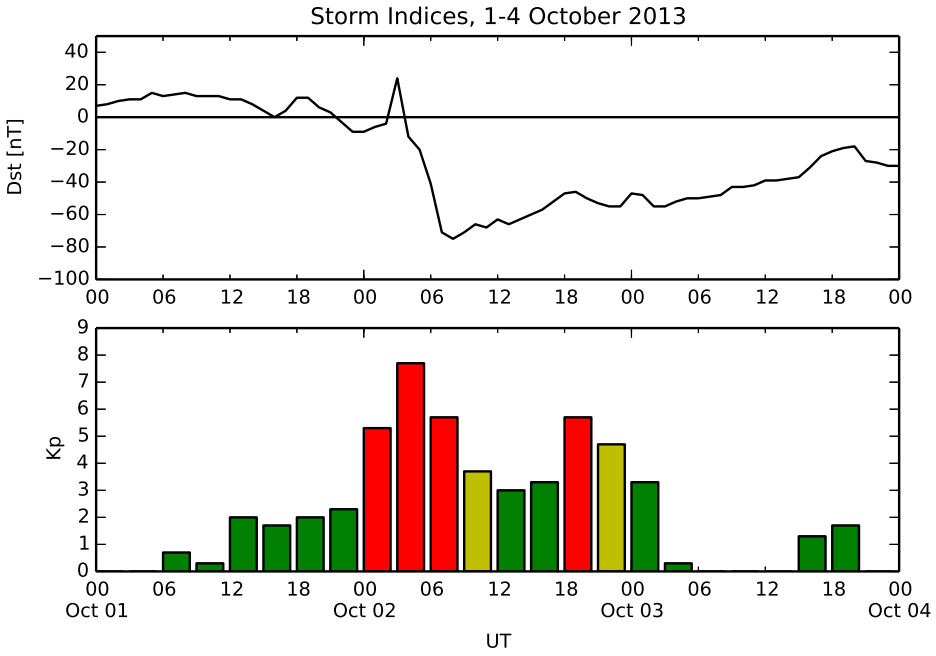

Figure 4.1: Dst (top) and Kp (bottom) indices for 01–04 Oct 2013. Adapted from *Makela et al.* [2014].

measure of ring current intensity, abruptly rose to 24 nT at 03:00 UT on 02 Oct 2013, a classical storm sudden commencement. It then dropped over the next few hours to a minimum of -75 nT at 08:00 UT. This is the main phase of the storm. The recovery phase began at 08:00 UT and lasted for several days. The Kp index, a global measure of geomagnetic activity, reached a maximum of 8 between 03:00 and 06:00 UT.

All five NATION FPIs were operating during this night, with measurements starting just before 00:00 UT and ending just after 11:00 UT on 02 Oct 2013. Except for brief periods of intermittent clouds, the data quality was excellent. The cloud sensor reading indicated cloud cover during 00:00–02:00 UT at PAR, 10:00–11:00 UT at EKU, 09:00–11:00 UT at UAO, and 08:00–10:00 UT at ANN. VTI did not have a cloud sensor at the time, but weather reports indicated clear skies until 07:30 UT [*Weather Underground*, 2013], and the data appear to be of high quality. The timing of the storm was ideal for NATION to observe the entire main phase of the storm. Because of the clear skies, optimal timing, and lack of maintenance issues, this storm is arguably the best-observed in the NATION database.

4.1.1 Horizontal Winds

In Figure 4.2, we show the horizontal wind fields estimated using data from all five FPIs and the algorithm described in Chapter 3, evaluated at a 30-minute cadence. Note that due to the large velocities on this night, the scale has been changed from Chapter 3.

Before 03:00 UT, the wind field is relatively smooth and the wind speed is small. The meridional wind is equatorward with a velocity of about 20 m/s. Initially, the zonal wind is 75 m/s eastward, gradually slowing to zero by about 02:30 UT. These pre-storm data are in general agreement with the quiet-time Horizontal Wind Model 2014 (HWM14) [*Drob et al.*, 2015], except in HWM14 the zonal winds are about 50 m/s larger.

Beginning abruptly at 03:00 UT, large deviations from HWM14 occur, and more structure in the wind field is evident. The zonal wind turns predominantly westward around 03:00 UT, reaching -75 m/s at the center of the field of view. It remains westward until about 06:00 UT, when it returns to predominantly eastward (0–50 m/s in the center of the field of view). From 03:00 UT to 08:00 UT, the zonal wind is 25–75 m/s more westward in the southern half of the field of view than in the northern half.

The meridional wind experiences a large equatorward surge starting around 03:00 UT. This disturbance is first seen in the north, and it propagates to the south. The meridional wind becomes more equatorward (i.e., negative) until 06:30 UT, when it reaches its minimum, -450 m/s in the center of the field of view. At about 07:30 UT, the meridional wind turns northward in the south, and this northward disturbance propagates northward. By 08:45 UT, the meridional wind is northward over the entire field of view, reaching a maximum value of about 175 m/s. During almost the entire period from 03:00 UT onward, the meridional wind is convergent, with winds in the south more positive than winds in the north by as much as 100 m/s.

Of particular importance is the unusually strong convergence seen at 04:45 UT, concentrated in the south. This convergence is so strong that the estimated wind at the southern edge of the field of view briefly turns northward, while the meridional wind is 200 m/s southward at the center of the field of view. Via the Burnside relation, (1.5), this implies a vertical wind of about -15 m/s.

Broadly, these horizontal wind measurements are in agreement with the

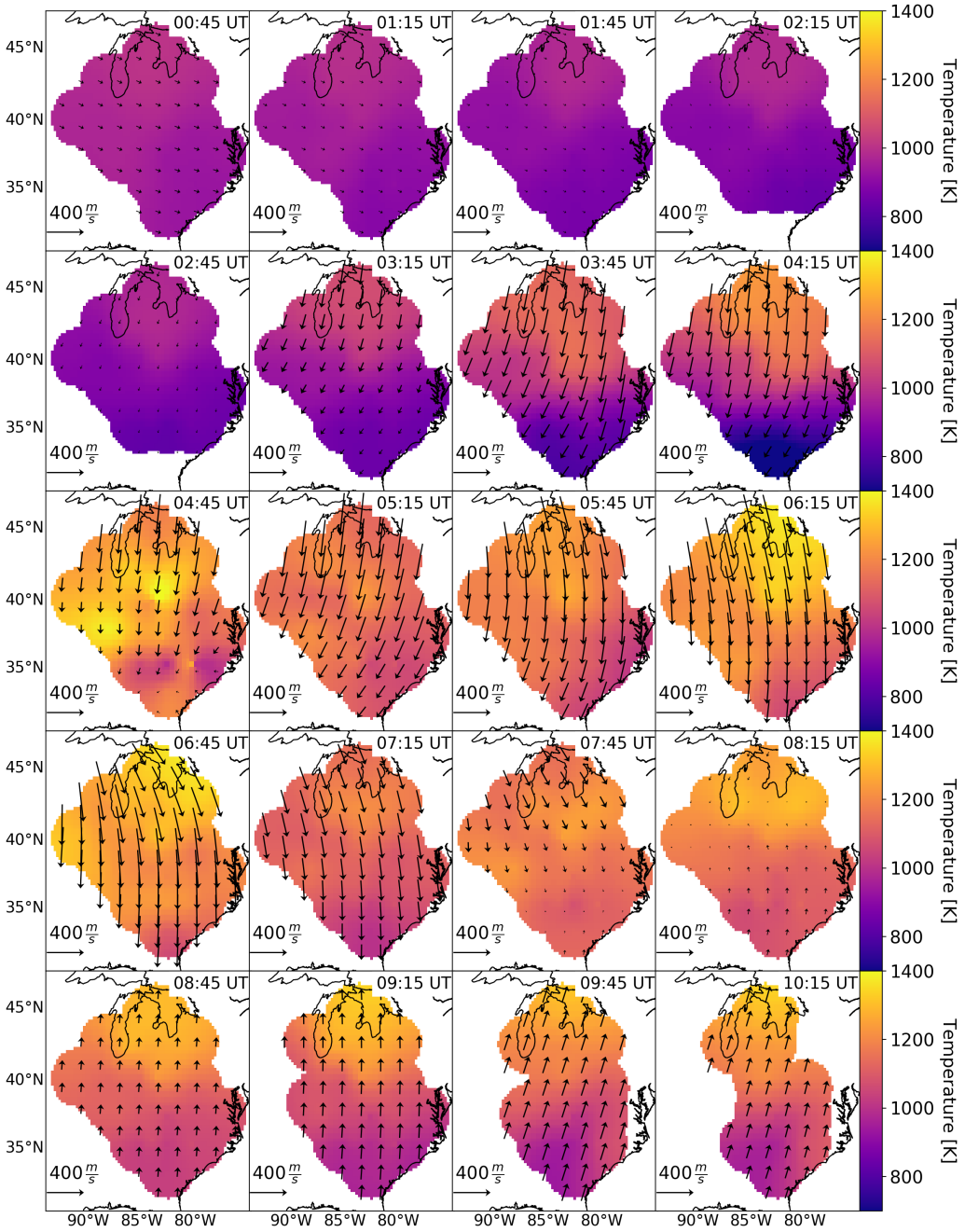

Figure 4.2: Horizontal wind and temperature estimated from NATION data for 02 Oct 2013.

current understanding of the thermospheric response to geomagnetic activity [*Rishbeth et al.*, 1987; *Prölss*, 1997; *Fuller-Rowell and Codrescu*, 1997; *Meriwether et al.*, 2008]. Namely, rapid heating in the auroral region, caused by particle precipitation and Joule heating, launches a large-scale gravity wave (sometimes called a traveling atmospheric disturbance), which propa-

gates from high to low latitudes. The passage of this disturbance from the northern auroral region brings a new circulation pattern to the thermosphere, consisting of strong southward winds moving air away from the auroral pressure bulge. The northward turning of the wind beginning at 07:30 UT is consistent with the arrival of the analogous disturbance from the southern auroral region. Qualitatively, the convergence of the winds (and thus, a downward vertical wind) is consistent with the expected new circulation, in which the downwelling feeds a return flow in the lower thermosphere. Quantitatively, it is surprising that the convergence is so large and sustained, but previous FPIs have also seen large convergences [e.g., *Hernandez and Roble*, 1976; *Hernandez*, 1982; *Biondi and Sipler*, 1985].

The westward turning of the zonal wind could in principle be caused by two mechanisms: the Coriolis force or the ion drag force. During storms, and especially when the interplanetary magnetic field turns southward, the high-latitude ion convection pattern expands equatorward, impinging on midlatitudes. This causes an ion drag force that is westward in the dusk sector and eastward in the dawn sector, which is consistent with the behavior of the zonal wind shown in Figure 4.2. However, ion drag cannot explain why the wind is more westward in the south. Ion drag would have a larger effect at higher latitude. This feature suggests the influence of the Coriolis force, which causes the wind to turn to the right in the northern hemisphere. However, the Coriolis force cannot explain why the zonal wind returns to eastward while the meridional wind is still southward at 06:15 UT. Evidently, both forces are significant.

4.1.2 Temperatures

Also shown in Figure 4.2 is the estimated temperature distribution. Although not explicitly mentioned in Chapter 3, it is trivial to modify the wind field estimation algorithm to estimate temperature. For the case of wind estimation, each measurement is only one of the three components of the wind vector. For temperature, the measurement is direct. Thus, the "temperature field" estimation problem is simply one of interpolation, for which there are many possible solutions. For the sake of consistency and simplicity, we use the same algorithm used for the winds.

Before 03:00 UT, the temperatures exhibit the normal cooling pattern, starting around 1000 K and cooling by 50–150 K. From 03:00 UT until 04:45 UT, temperatures in the north rise to 1200 K while temperatures in the south continue to drop. At 04:45 UT, a rapid temperature increase occurs across the field of view. The temperature distribution is complex, showing structuring on 100-km scales. After 04:45 UT, the temperature distribution becomes smoother, fluctuating between 1050 and 1200 K at the center of the field of view, but maintaining a 50–200 K spatial gradient from the south to the north.

The heating of the thermosphere during storms is well-documented, arising from particle precipitation, Joule heating, and advection from strong heating in the auroral region, and our measurements are generally consistent with this understanding. However, the structure in the temperature distribution at 04:45 UT is surprising, as no localized heat sources are known, except perhaps stable auroral red (SAR) arcs [*Kozyra et al.*, 1997], which could explain the latitudinal structuring, but not the longitudinal structuring.

It is noteworthy that the structuring occurs where two different instruments are looking at nearly the same location (e.g., UAO looking south with EKU looking west, and PAR looking east with VTI looking south). In Figure 4.3, we compare the raw temperature measurements from ANN looking south and EKU looking north. Assuming the airglow layer is at 250 km altitude, these measurement locations differ by 54 km horizontally, so we would expect their temperature measurements to be similar, and on a typical quiet night they are, to within a couple tens of Kelvin. On this night, they differ systematically. The south-looking measurement is hotter than the north-looking measurement by 93 K on average during the storm period, and it is almost 400 K different at 04:45 UT. For the three other near-coincident measurement locations on this night (VTI south with PAR east, UAO south with EKU west, and EKU south with PAR west), the south-looking measurement is similarly hotter than the others (not shown).

Figure 4.3 suggests that the temperature measurement is aspect-sensitive, which means that the emission spectrum is anisotropic. Under the assumption that the 630.0-nm emission spectrum represents the velocity distribution of neutral oxygen atoms, this implies a significant departure of the thermosphere from thermodynamic equilibrium, which is unlikely.

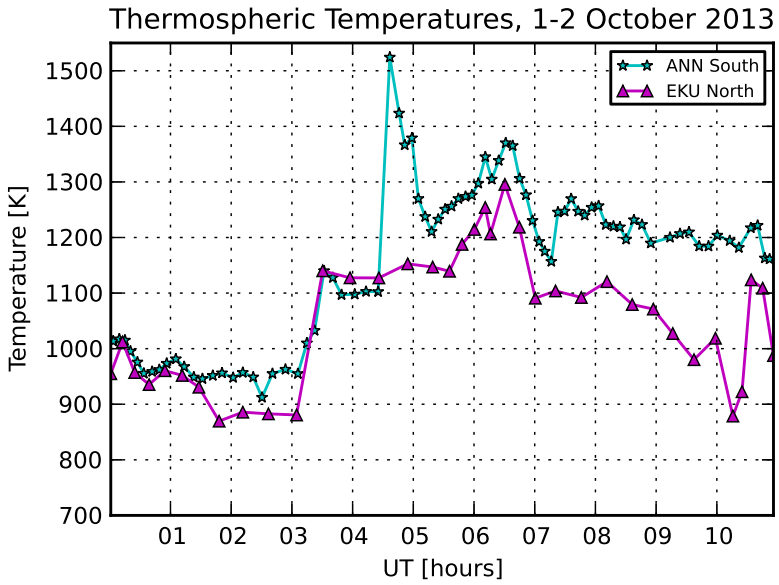

Figure 4.3: Comparison of the thermospheric temperature measured from ANN looking south and EKU looking north. The intersections of these lines of sight at 250-km altitude differ by 54 km. Adapted from *Makela et al.* [2014].

4.1.3 Vertical Winds

Until now we have only discussed the horizontal winds and temperatures. Here, we show the vertical winds. Instead of showing the estimated vertical wind fields, we will show the raw zenith-looking measurements, which are direct measurements of the vertical wind. Figure 4.4 shows a stack plot of the vertical winds measured from the five NATION FPIs and also the MH FPI. The data have been shifted up or down proportional to the latitude of the site for ease of comparison. The horizontal line indicates zero for each site.

For ANN, a custom Doppler referencing technique is required in order to correct for the drift of the laser wavelength, which superimposes a linear trend on the winds of ∼50 m/s over the night. To remove this drift, we assume that the pre-storm (00:00–02:00 UT) and post-storm (09:00–12:00 UT) vertical winds are zero on average, and that the Doppler reference varies linearly with time. The lasers at the other sites were stable.

High-quality vertical wind measurements during quiet nights are typically less than 10 m/s. However, on this night, a large downward wind is seen at

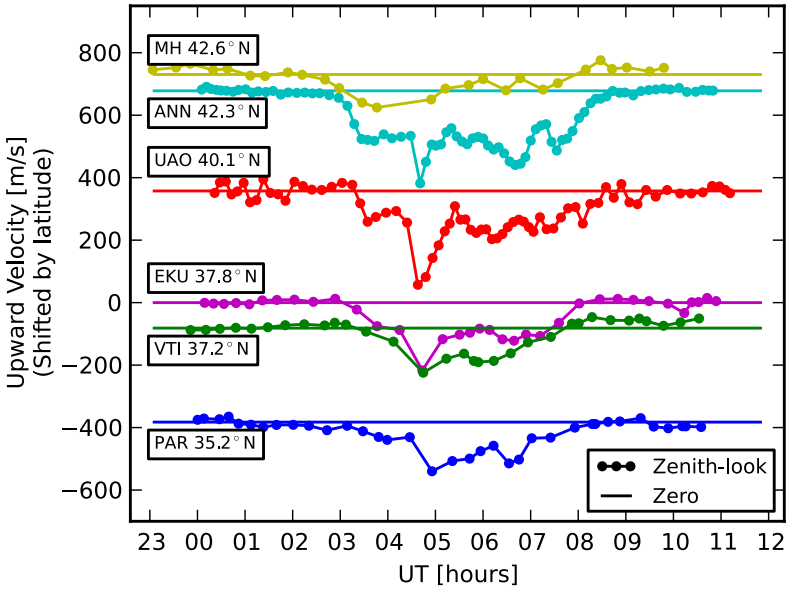

Figure 4.4: Vertical winds measured by the five NATION FPIs plus the MH FPI. The zero reference lines have been shifted proportional to the latitude of the site. Statistical uncertainties are less than 5 m/s. Adapted from *Makela et al.* [2014] with modifications.

all sites. The vertical wind is strongest in the north (e.g., -150 m/s at ANN) and weaker at the southern sites. It lasts about 5 hours, appearing first in the north around 03:00 UT, and subsiding last in the north, around 08:30 UT. Superimposed on this sustained downward wind is a strong downward surge seen during 04:30–05:00 UT. It is seen at all sites, except MH, which was not taking data during 04:00–05:00 UT to perform flat field calibrations. This strong downward wind is coincident with the temperature increase, temperature structuring, and strong horizontal convergence seen above. The strength of this surge is strongly latitude-dependent: ANN observes a -300 m/s vertical wind, while PAR observes -130 m/s.

These large sustained vertical winds are about a factor of 10 stronger than the vertical wind inferred from the horizontal convergence, which at its strongest at 04:45 UT is only about -15 m/s. At times, even the sign of the Burnside vertical wind and measured vertical wind are different (not shown), which has been observed before in FPI studies [*Smith and Hernandez*, 1995]. A key assumption in deriving the Burnside relation is hydrostatic equilibrium, so these results seem to indicate a departure from this equilibrium. The large (<-100 m/s), long-lasting (5 hours), wide-spread (10 degrees of latitude)

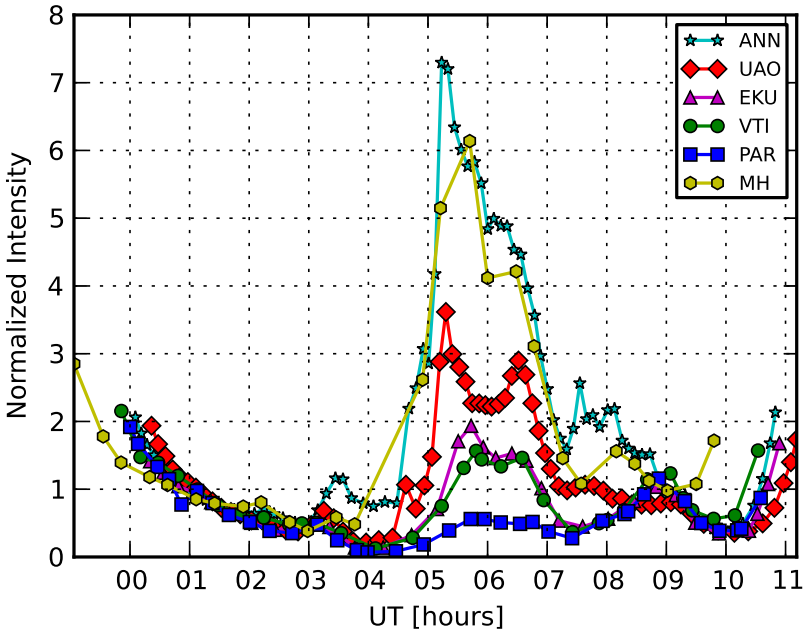

Figure 4.5: Zenith-looking intensities (or brightnesses) measured from each FPI, on an arbitrary scale. The 00:00–01:00 UT period was used for cross-calibration. Adapted from *Makela et al.* [2014].

vertical winds seen during this storm should rapidly deplete the thermosphere and cause extraordinary heating, if they were real. While recent models have shown that departures from hydrostatic equilibrium can lead to large 100 m/s vertical winds [e.g., *Deng et al.*, 2008], these are short-lived and cannot explain the sustained vertical winds seen here.

4.1.4 Airglow Brightness

Figure 4.5 shows the normalized airglow brightness measured by the FPIs at each site, looking toward zenith. Since the FPIs are not absolutely calibrated, we have calibrated them relative to each other by assuming they observe the same average brightness from 01:00 to 02:00 UT. On a typical quiet night, the brightness starts high and gradually decreases until just before dawn, when it increases again. Until 03:00 UT on this night, the behavior is typical. At 03:00 UT, coincident with the equatorward surge in the horizontal wind, the temperature increase, and the downward wind, we observe a small increase in brightness at all sites. Beginning at around 04:45 UT, coincident

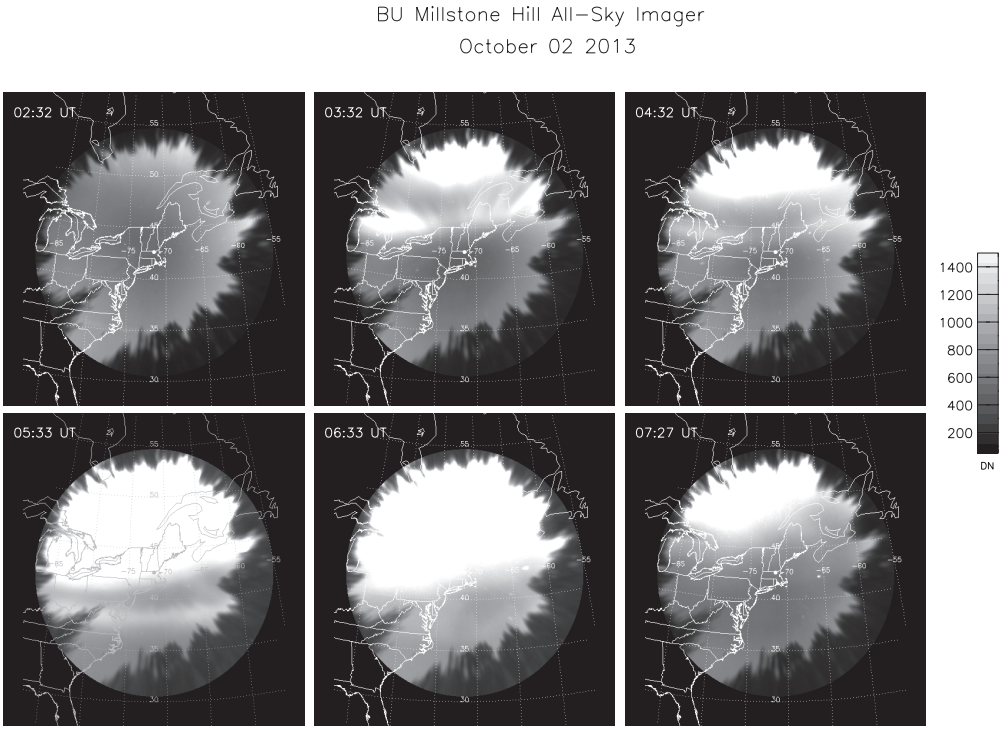

Figure 4.6: All-sky 630.0-nm images from the Boston University imager, projected onto a shell at 400-km altitude. Adapted from *Makela et al.* [2014].

with the large downward surge in the vertical winds, the brightness increases dramatically by almost a factor of 10. During the entire period from 03:00 to 08:30 UT, there is a large northward gradient in brightness. A secondary peak in the brightness is seen around 09:00 UT at the southern sites (EKU, VTI, and PAR), coincident with the northward turning of the meridional wind. This indicates that this secondary peak is caused by increased dissociative recombination, as the northward wind drags ions down the magnetic field line. The earlier, larger peak must be due to some other source than dissociative recombination.

To provide spatial context for interpreting the FPI measurements, we show data from the Boston University Imaging Group's all-sky imager, hereafter referred to as the BU imager, in Figure 4.6. The BU imager is colocated with the MH FPI. The instrument and calibration are described by *Baumgardner et al.* [2008] and references therein. A subsample of images from 02:30 to 07:30 UT are shown at an hourly cadence. Pixels above 10° elevation are

mapped to a geographic grid assuming a thin-shell emission at 400 km, which was chosen because of its relevance for SAR arc studies. The dark protrusions at some azimuth angles are trees blocking the field of view.

In the first image at 02:32 UT, the brightness is relatively constant, arising purely from dissociative recombination. In subsequent images, two other contributions to the 630.0-nm emission are seen: the bright diffuse red aurora in the north, and a stable auroral red (SAR) arc [*Kozyra et al.*, 1997], a thin strip of brightness separated from the diffuse red aurora by a few degrees of latitude. Sometime between 04:32 and 05:33 UT, the SAR arc and the equatorward edge of the diffuse red aurora pass over MH. This is consistent with the large increase in brightness seen by the FPIs starting at 04:45 UT.

The coincidence at 04:45 UT between the arrival of the diffuse red aurora and the apparent departures from thermodynamic and hydrostatic equilibrium in the wind and temperature data suggest that the FPI data must be somehow contaminated by this new emission source during this period. Specifically, we question whether the 630.0-nm emission spectrum is representative of the velocity distribution of atomic oxygen under storm conditions at midlatitudes. While it is well accepted that the majority of the $O(^1D)$ particles resulting from dissociative recombination will thermalize with the ambient oxygen gas before emitting, it is less clear that the energetic $O(^1D)$ particles created in the diffuse red aurora will thermalize. The characteristic energy of the precipitating ion population that causes the diffuse red aurora is measured in keV, while the initial energy of $O(^1D)$ resulting from dissociative recombination is only a few eV. Due to their high initial speeds, it stands to reason that the $O(^1D)$ in the diffuse red aurora might emit a photon before thermalization is complete, invalidating the assumption that the emission represents the velocity distribution of the neutral oxygen gas. Before discussing this hypothesis further in Section 4.3, we perform a statistical study in the next section of the storms in the NATION database to determine whether the downward apparent winds seen during the 02 Oct 2013 storm are common.

4.2 Observational Statistics: 2011–2016

In this section we summarize NATION vertical wind observations during geomagnetic storms in the 5-year period starting 01 Sep 2011 and ending 01 Sep 2016. It is difficult to establish the Doppler reference in an automated way when apparent vertical winds are large and sustained. For this analysis, instead of assuming the mean, median, or mode of the vertical wind is zero, we assume that the average of the first five vertical wind samples after sunset is zero.

Out of the 5-year period, we select storm periods of interest using the following criteria:

1. The Dst decreases by more than 50 nT in a 6-hour time window.
2. The minimum value of Dst is less than or equal to -50 nT during this window.
3. The center of the time window occurs between 00:00 and 07:00 UT.
4. At least one of the FPIs was operating with excellent data quality (as defined below).

The combination of these four criteria ensures that only large storms are included and that NATION is well-positioned in local time to observe the entire main phase at night. The data are determined to be of excellent quality if they satisfy the following criteria:

1. There are no known maintenance issues (e.g., laser operating, CCD cooling to -60 °C).
2. The wind uncertainty is less than 50 m/s.
3. If a cloud sensor is present, the reading indicates a difference in the sky temperature and ambient temperature of more than 10 °C, and if it is not present, all data are used (in order to maximize the available data for analysis).
4. The airglow brightness is above the threshold determined by the OH contamination analysis described in Section 2.2.4.

As long as 75% of the samples from one instrument during one night pass the quality check, the samples are included in the analysis, though the bad samples are excluded. One data set (14–15 Jul 2012) was manually excluded because no cloud sensor was present, but it was clear from the rapid fluctuations in brightness measurements that clouds were present.

In total, 15 storms were identified, with an average of 1.9 instruments available per storm. Using these data, we perform a superposed epoch analysis. For each storm, we identify the middle of the main phase as a reference time (called the "epoch" time) and plot vertical wind data relative to the epoch time. Specifically, we define the epoch time as the midpoint between the time of most negative and most positive curvature in Dst, which correspond to storm sudden commencement and to the end of the main phase, respectively. We also require the end of the main phase to be a local minimum in Dst. Although seemingly unnecessarily complex, this definition was the only one we found to provide a robust and useful epoch time for the wide variety of storms studied here, which include storms without a classical storm sudden commencement and storms that comprise multiple substorms.

The results of the superposed epoch analysis are shown in Figure 4.7. The top panel displays the Dst index for each of the 15 storms, relative to the epoch time. The bottom panel shows the vertical wind data from all available FPIs for each storm. The solid line is an average of the vertical wind data binned into 30-minute intervals. From this analysis, it is clear that there is a persistent downward vertical wind seen by NATION during the main phase of storms, between -2 and +4 hours since epoch. There is a second decrease seen at +7 hours since epoch, corresponding to the storms that have a second drop in Dst after the initial drop.

Although the pattern is clear in the average sense, there is a significant amount of variability about the average. The primary cause of this is storm-to-storm variability. For some storms, such as the 01–02 Oct 2013 storm, the apparent vertical winds are large (-150 m/s). For others, they are smaller. Individual analysis of each storm (not shown) indicates that for all storms except one (08–09 Oct 2013), large apparent vertical winds (less than -50 m/s) were seen.

Another contributor to the variability is storm timing. Even though we organize the events by Dst, the complex timing of the chain of events from ring current intensification to wind disturbances depends on inner magnetospheric

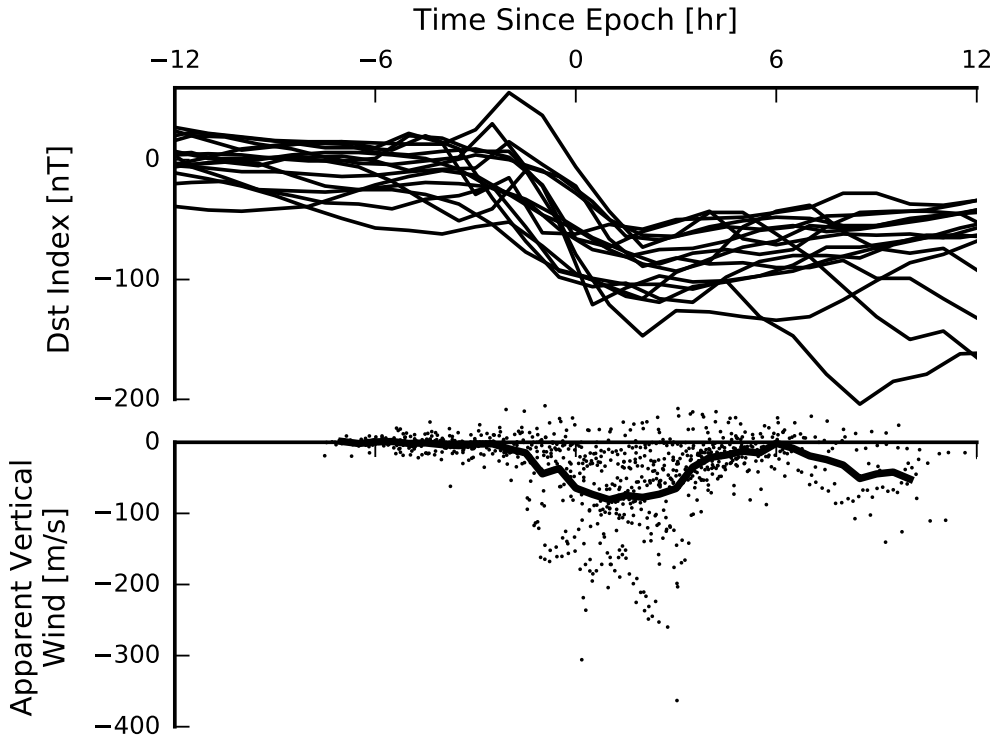

Figure 4.7: A superposed epoch analysis of 15 storms from the NATION database. The epoch is defined as the middle of the storm main phase. The Dst index (top) and all vertical wind samples from all available FPIs (bottom) are plotted relative to the epoch time. The solid line is the average of the FPI data in 30-minute bins. Adapted from *Harding et al.* [Submitted].

conditions, the interplay between the magnetosphere and ionosphere at auroral latitudes, and the prevailing conditions of the thermosphere, which can depend on season as well as previous storm activity. Because of this, some vertical wind events last 6 hours, while others last as short as 2 and can occur at different epoch times. Another contributor to the variability is the latitudinal spread of the NATION FPIs. Often, the higher latitude sites experience stronger apparent vertical winds than the lower latitude sites, and sometimes at slightly different times.

However, despite the variability, it is clear that apparent downward winds are a common stormtime feature, observed across multiple NATION instruments spanning several degrees of latitude and longitude. In the next section, we describe one hypothesis to explain these observations.

4.3 The O^+ Precipitation Hypothesis

For almost 50 years, FPIs have been used to monitor thermospheric wind and temperature, under the assumption that the 630.0-nm emission spectrum represents the velocity distribution of atomic oxygen, the dominant species in the thermosphere. NATION observations call this assumption into question at midlatitudes during storms, suggesting the presence of some type of contamination that leads to an anomalous Doppler shift of the 630.0-nm emission.

To elucidate the possible sources of contamination, we briefly summarize the sources of $O(^1D)$, the species which emits a photon at 630.0 nm. As described in Chapter 1, the primary source of $O(^1D)$ during quiet times is dissociative recombination of O_2^+ [*Link and Cogger*, 1988, 1989]:

$$O_2^+ + e \rightarrow O + O\left(^1D\right). \tag{4.1}$$

During geomagnetically active times, there are alternative sources of $O(^1D)$, detailed comprehensively by *Solomon et al.* [1988]. Here, we describe the most important. First, electron precipitation in the auroral region, and proton precipitation over a wider latitude range, produce secondary electrons which are efficient at exciting $O(^1D)$ [*Lummerzheim et al.*, 2001]:

$$O + e^* \rightarrow O\left(^1D\right) + e^* \tag{4.2}$$

where the * superscript denotes a fast, suprathermal species.

Second, ionospheric electrons that are in the high-energy tail of the thermal distribution can excite oxygen atoms, since the excitation energy of $O(^1D)$ is only 1.97 eV. This becomes a significant source of $O(^1D)$ when the electron temperature is higher than $\sim$2500 K, which produces SAR arcs [*Kozyra et al.*, 1997]:

$$O + e \rightarrow O\left(^1D\right) + e. \tag{4.3}$$

However, neither of these sources of $O(^1D)$ is able to produce an anomalous Doppler shift. The SAR arc source has no preferential downward motion. While there is a preferential downward motion to precipitating protons and electrons, they do not directly excite oxygen; rather, the secondary electrons do. The low-energy ($\sim$20 eV) secondary electrons are most efficient at excit-

ing O(^{1}D), so momentum transfer to O(^{1}D) from electrons is expected to be low given their relatively low mass and energy.

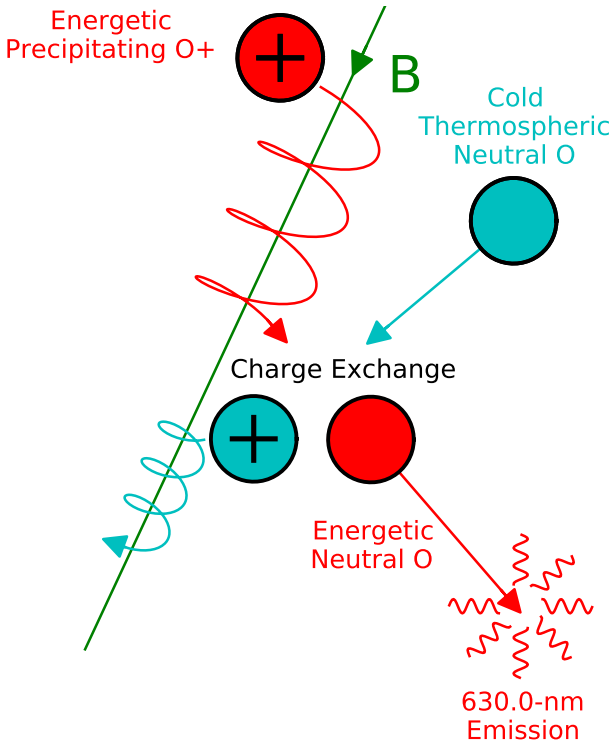

Figure 4.8: Schematic representation of the O^+ precipitation hypothesis. Precipitating O^+ particles charge exchange with neutral thermospheric particles, producing fast O with a preferential downward motion. Some of these fast O, or the secondary fast O from subsequent collisions, will be excited to the O(^{1}D) state. This emission will have a Doppler shift which will appear to indicate a downward vertical wind.

The only source of O(^{1}D) that could potentially produce a preferential downward motion is the precipitation of O^+ particles. When these energetic O^+ particles precipitate into the thermosphere from the ring current, they collide with ambient oxygen atoms. The most likely reaction resulting from this collision is charge exchange [*Kozyra et al.*, 1982]:

$$O^{+*} + O \rightarrow O^* + O^+ \tag{4.4}$$

which creates a shower of energetic neutral oxygen atoms. These atoms

collide with ambient oxygen atoms, transferring momentum and occasionally exciting them to O(^{1}D):

$$O^* + O \rightarrow O^* + O^*(^1D)\,. \tag{4.5}$$

These O(^{1}D) atoms then emit a photon at 630.0 nm before fully thermalizing. This mechanism is shown schematically in Figure 4.8.

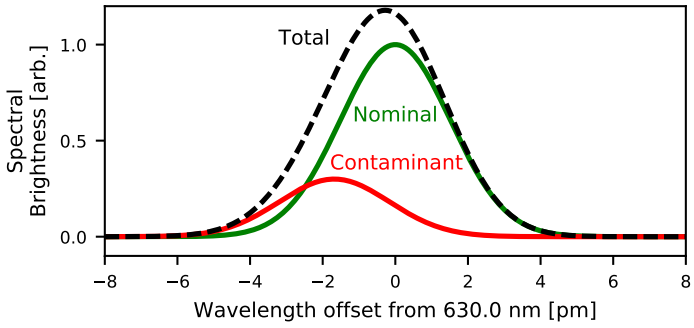

Figure 4.9: Hypothesized contamination of the 630.0-nm spectrum causing apparent downward winds. The total spectrum measured by a zenith-looking instrument consists of a combination of the nominal spectrum, which has no Doppler shift, and the contaminant spectrum, which has a Doppler shift. The combined spectrum is then falsely interpreted to represent a downward vertical wind.

The hypothesized spectrum observed by a zenith-looking FPI is shown in Figure 4.9. The total spectrum consists of two parts. The nominal spectrum (shown in green) arises from O(^{1}D) excited by dissociative recombination, electron precipitation, ion precipitation, SAR arcs, and other mechanisms which produce a thermalized emission. Because there is no actual vertical wind, this component of the spectrum has no Doppler shift. The second component of the spectrum (shown in red) is hypothesized to arise from O^+ precipitation. These O(^{1}D) atoms are not able to thermalize completely, and thus have a preferential downward motion when they emit, resulting in a negative Doppler shift. This yields a total spectrum that appears to represent a downward wind. Although the Doppler shift in Figure 4.9 appears small, it represents a wind of -130 m/s. A similar mechanism has been shown to produce Doppler shifts in hydrogen emissions during proton precipitation [*Lummerzheim and Galand*, 2001; *Galand and Chakrabarti*, 2006]. While protons are the primary constituent of the ring current during quiet

times, it has become widely recognized that O^+ can become a significant and potentially dominant constituent during geomagnetically active times [*Keika et al.*, 2013, and references therein].

The effect of precipitating O^+ on the production of energetic neutral oxygen atoms has been studied extensively [*Torr et al.*, 1974; *Torr and Torr*, 1979; *Torr et al.*, 1982; *Kozyra et al.*, 1982; *Ishimoto et al.*, 1986; *Bisikalo et al.*, 1995; *Goldstein and McComas*, 2013]. However, the effect on the 630.0-nm emission has only been addressed by *Ishimoto et al.* [1994], who use a two-stream approximation to model the altitude-dependent energy spectra of O and N_2 resulting from a given incident O^+ spectrum at the top of the atmosphere. The model accounts for a number of physical mechanisms, the most important of which are charge exchange, momentum transfer, and ionization. Using the incident O^+ spectrum measured by *Shelley et al.* [1972], the model predicts a small but potentially significant amount of 630.0-nm emission, with a Doppler shift which the authors considered negligible but is actually quite significant for the FPI community (equivalent to a 500–7000 m/s downward wind). This has never been verified experimentally. Additionally, the authors did not account for other sources of $O(^1D)$ such as dissociative recombination when calculating the apparent Doppler shift.

It is also worth mentioning that instead of occurring in the thermosphere, it is possible that the charge exchange occurs in the exosphere, where the collision would likely be with neutral hydrogen instead of neutral oxygen. This mechanism is a significant contributor to ring current loss and could contribute to the penetration of fast atoms to low latitudes [*Tinsley*, 1981]. It has also been implicated in the production of ultraviolet airglow emissions, but no connection with red emission has been made [*Stephan et al.*, 2000, 2004; *DeMajistre et al.*, 2005; *Zhang et al.*, 2006].

4.3.1 Supporting Arguments

The O^+ precipitation hypothesis is consistent with many aspects of the observations presented above. First, it explains why the apparent vertical winds are downward instead of upward: the precipitating O^+ particles have a preferential downward motion. It also explains the reduction in the vertical wind

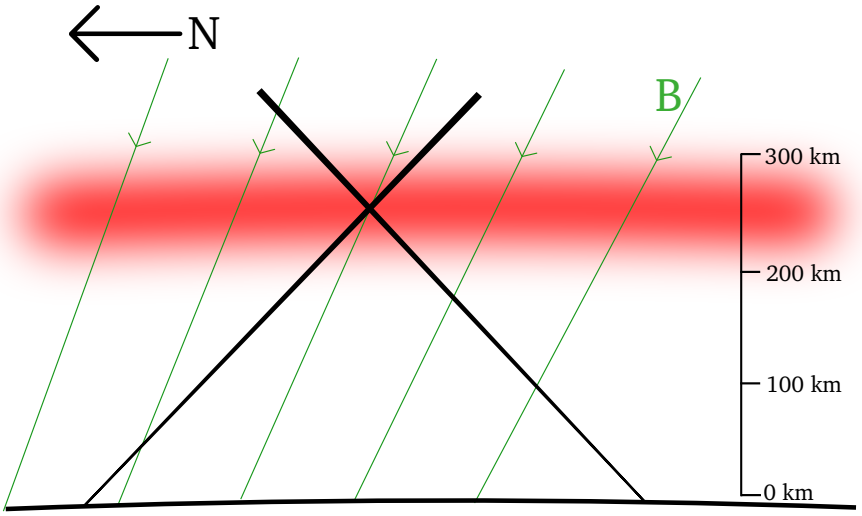

Figure 4.10: Schematic of the near-common-volume measurements presented in Figure 4.3. North is to the left. Green lines represent the magnetic field. Two lines of sight are shown, representing ANN looking south and EKU looking north.

with decreasing latitude, as the flux of O^+ particles has been measured to have a similar reduction [*Sharp et al.*, 1976].

The coincidence between the initial airglow brightness enhancement and downward vertical winds at 03:00 UT on 02 Oct 2013 suggests a connection between an additional source of $O(^1D)$ and the wind contaminant. The fact that the airglow enhancements and vertical winds both subside around 08:00 UT strengthens this argument, as does the simultaneous large enhancement seen in both quantities at 04:45 UT.

The O^+ precipitation hypothesis is also consistent with the apparent anisotropy in the temperature measurements shown above in Figure 4.3. In Figure 4.10, we show the geometry of this near-common-volume observation with the magnetic field lines for reference. The effect of O^+ streaming down the magnetic field lines will cause a velocity distribution of O which is wider along the magnetic field than across it. A south-looking line of sight is approximately parallel to the magnetic field, so the spectrum observed would be wider than a north-looking spectrum. This explains why the ANN FPI looking south observes a higher apparent temperature than the EKU FPI looking north, even though they are observing roughly the same location.

It is important to point out that precipitating O^+ would contaminate the horizontal wind measurements in addition to the vertical winds, as the

oblique lines of sight would also observe an anomalous Doppler shift, which would depend on the angle between the line of sight and the magnetic field. In particular, south-looking lines of sight would observe artificial blue shifts, and north-looking lines of sight, which are roughly perpendicular to the field, would not be affected as much, though the effect is complicated as it depends on the degree of thermalization of the $O(^1D)$ particles. The excess blue shift in the south would result in an apparent convergence of the horizontal wind field, which is consistent with the observations, especially during the strong enhancement at 04:45 UT.

It is likely that previous observational studies have also been subject to contamination. For example, the large convergence that has been reported at midlatitudes during storms by *Hernandez and Roble* [1976], *Hernandez et al.* [1982], and *Biondi and Sipler* [1985] appears to be analogous to the convergence observed by NATION. These studies were performed with a single FPI. With NATION, we are able to compare wind and temperature measurements from different sites in order to recognize the presence of contamination in a way that would have been impossible from a single site.

Though controversial, several previous studies have provided evidence of the existence of suprathermal oxygen atoms in the thermosphere, even during quiet times, though these are thought to be produced by the exothermic dissociative recombination of O_2^+ and N_2^+ [*Hickey et al.*, 1995], not ion precipitation. The observational evidence for fast oxygen (or, as it is called during quiet times, "hot" oxygen) uses various techniques: post-sunset observations of the 732.0- and 733.0-nm oxygen ion lines from the ground [*Yee et al.*, 1980], 630.0-nm intensity profiles from the Visible Airglow Experiment on the Atmosphere Explorer-3 satellite [*Schmitt et al.*, 1981], 630.0-nm temperature measurements from the FPI on the Dynamics Explorer satellite [*Hubert et al.*, 2001], and in situ measurements of fast oxygen from the low energy neutral atom (LENA) instrument on the IMAGE spacecraft [*Wilson and Moore*, 2005]. This observational evidence is corroborated by modeling which indicates that hot oxygen should cause a 25–75 K positive bias on estimated temperatures from $O(^1D)$ atoms above 200 km [*Shematovich et al.*, 1999; *Sipler and Biondi*, 2003; *Kharchenko et al.*, 2005]. The hypothesis presented here adds O^+ precipitation to dissociative recombination as a source of fast O, and claims that it can have an effect on wind measurements as well as temperature.

4.3.2 Refuting Arguments

Theoretically, the strongest argument against the O^+ precipitation hypothesis is the inefficiency with which O^+ precipitation causes 630.0-nm emission. Due to the large resonance charge exchange cross section, the incoming beam of O^+ gets rapidly converted to neutral O. The vast majority of this energy goes to heating the thermosphere via momentum transfer collisions, not electronic excitation or ionization [*Rees*, 1989]. Nevertheless, the model by *Ishimoto et al.* [1994] suggests that the 630.0-nm vertical brightness can be as large as 153 R, which is about an order of magnitude larger than dissociative recombination. However, many of the inputs to the *Ishimoto et al.* [1994] model are uncertain (e.g., emission cross sections, elastic scattering cross sections, and incident spectra). Other model runs by *Ishimoto et al.* [1994] with different inputs predicted much less emission, some below 1 R. Because this emission is expected to be co-located or near the SAR arc emission and proton-precipitation emission, it is likely negligible. However, the emission is proportional to the flux, so it is conceivable that this effect could be significant in cases where the flux is much larger than the strongest flux used by *Ishimoto et al.* [1994], or in cases where SAR arc emission is weak and the precipitation is dominated by O^+ .

Observationally, there are a few inconsistencies as well. The proposed spectral contamination should cause an extra bump in the measured spectrum, but no statistically significant departures from Gaussianity were found in the data, despite an exhaustive effort to find them. However, this is expected for two reasons. First, the fast $O(^1D)$ are much faster than the wind, but since they are partially thermalized, their velocity is not large relative to the thermal width of the emission. Thus, it is difficult to distinguish the total spectrum from a true Gaussian (see Figure 4.9). Second, due to noise and the broadening effect of the Airy function, our ability to resolve subtle spectral features is limited.

In Figure 4.11 we show data from an independently run FPI located at the Arecibo Observatory (latitude 18.34°, longitude -66.75°), provided by J. Noto [personal communication]. We have established the Doppler reference so that the mean vertical wind in the 00:00 to 02:00 UT pre-storm period is zero. A clear upward vertical wind is seen starting at around 04:30 UT, coincident with the large surge seen in the NATION data. This upward

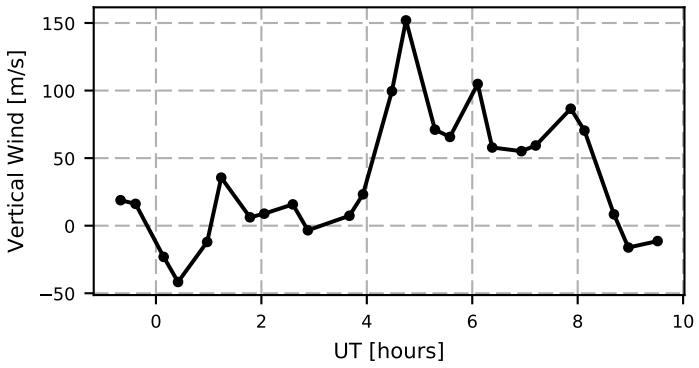

Figure 4.11: Apparent vertical wind measured by the FPI at the Arecibo Observatory. Data are courtesy of John Noto.

vertical wind peaks at 150 m/s, and then fluctuates between 50 and 100 m/s until subsiding around 09:00 UT.

This observation of an apparent upward wind is strong evidence against the O^+ precipitation hypothesis. Even though the *Ishimoto et al.* [1994] model has shown that in the case of high-energy (> 50 keV) precipitation, an upward Doppler shift is possible (due to the fact that the main contributor to the emission in this case are the backsplashed O atoms), the emission brightness is small and negligible compared to other sources of 630.0-nm emission. One caveat is that the emission over Arecibo becomes quite dim at 04:30 UT, so the measured vertical winds may be an artifact of OH. However if the Arecibo data are trustworthy, they indicate that a new contamination explanation is required, which can explain both positive and negative Doppler shifts.

As mentioned above, the O^+ precipitation hypothesis is consistent with the temperature anisotropy seen in Figure 4.3: temperatures measured looking nearly parallel to the magnetic field (south) are larger than temperatures measured perpendicular to it (north). However, the magnetic dip angle over UAO is about 67°, so the zenith-looking direction is as close to parallel as the south-looking direction. Thus, we would expect to measure hot temperatures in the zenith as well as south directions. Unfortunately, there are no common-volume locations we can use to test this, but instead we can use UAO's zenith direction, which is separated from the ANN-EKU common-volume point by 4 degrees of longitude but is approximately the same latitude, as shown in Figure 4.12. We expect most of the variation in airglow brightness, wind,

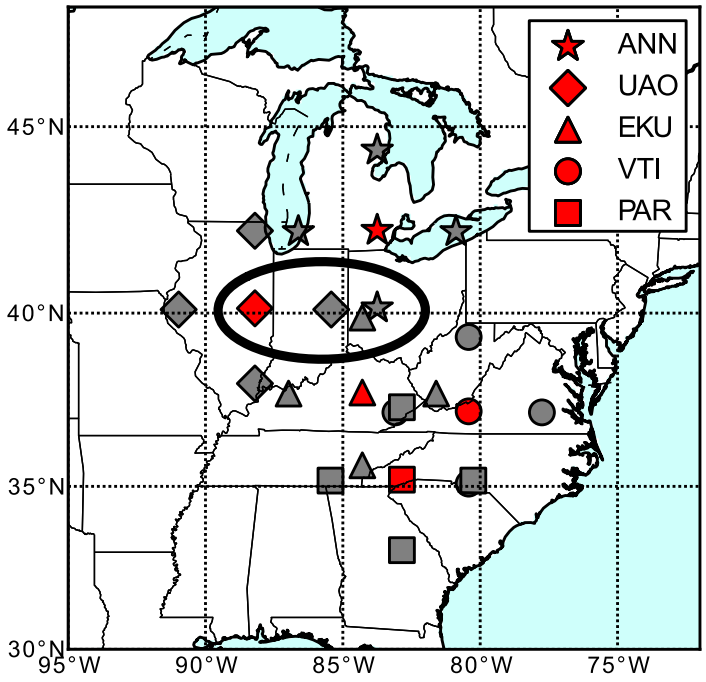

Figure 4.12: (Red) Locations of NATION sites. (Gray and Red) Intersections of lines of sight with the assumed emission altitude of 250 km. The oval includes UAO zenith, ANN south, and EKU north, which are used in Figure 4.13. UAO east is also circled but not included in Figure 4.13.

and temperature to be with latitude, not longitude, so this is a suitable approximation.

In Figure 4.13, we update Figure 4.3 to include the UAO zenith-looking direction. The UAO zenith temperatures are cooler than either the ANN south or EKU north temperatures for most of the night. During the large airglow enhancement and apparent downward wind surge from 04:30 UT to 05:00 UT, the UAO temperature briefly surpasses the EKU temperature, but it never exceeds the ANN temperature. After about 07:30 UT, the UAO and EKU temperatures are no longer significantly different, but the ANN temperature remains higher. These data are inconsistent with the O^+ precipitation hypothesis, but the 4 degrees of longitudinal separation is a caveat.

To elucidate these inconsistencies in the temperature measurement, in Jan-

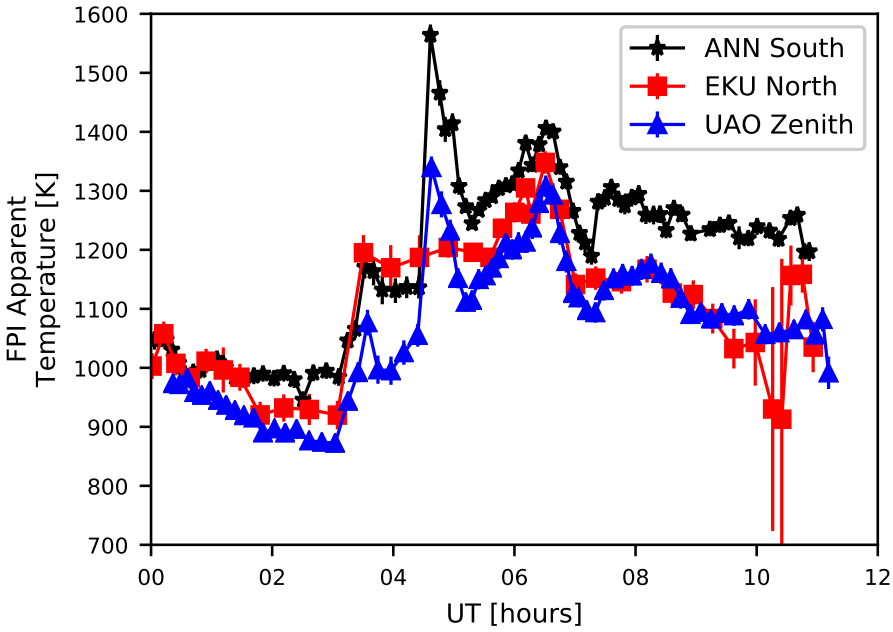

Figure 4.13: Same as Figure 4.3, but the UAO zenith-looking direction is included.

uary 2015 we added an additional direction to the cardinal cycle (zenith, north, south, east, west). This new direction is pointed antiparallel to the magnetic field. According to the O^+ precipitation hypothesis, this direction should measure the largest temperature (assuming the actual temperature field is constant). During the large geomagnetic storm of 17 Mar 2015, two of our instruments (UAO and PAR) were operating. We show the temperatures measured by UAO in Figure 4.14. Coincident with the arrival of a brightness enhancement and downward vertical winds at around 06:00 UT (not shown), the temperatures increase from 800 K to about 1200 K. After this time, the south-looking direction becomes consistently hotter than the other directions, until 10:00 UT, when the temperatures return to being roughly equal in all directions. This observation appears to refute our hypothesis. The caveat with this analysis is that since we are not using a common-volume point, and instead using data from a single site looking in different directions, we are potentially conflating angular and spatial information. However, it is hard to imagine that the actual thermospheric temperature is larger in the south, as it is farther from the region of intense auroral heating.

The last piece of evidence against the O^+ precipitation hypothesis is the

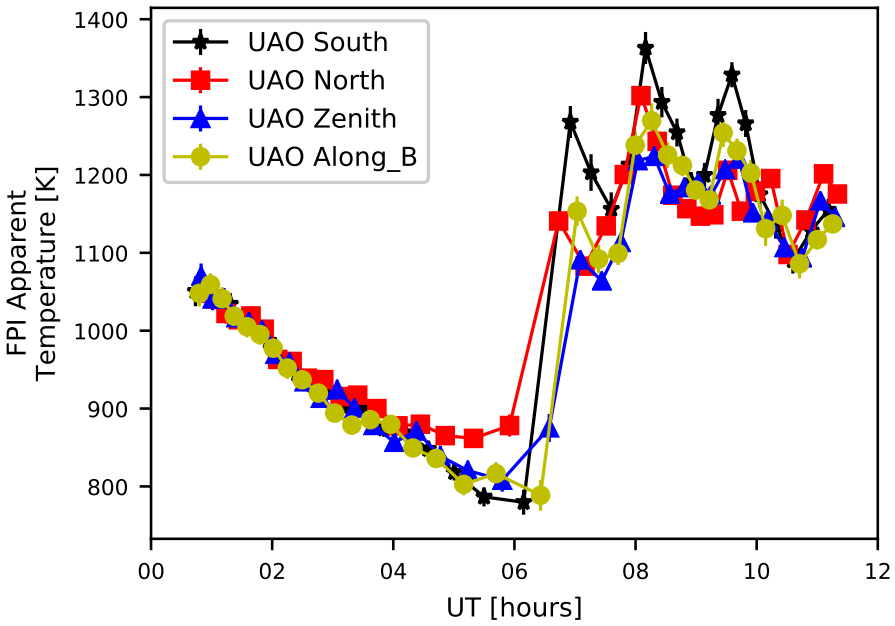

Figure 4.14: Apparent thermospheric temperatures measured by the UAO FPI during the 17 Mar 2015 storm.

outlier storm mentioned in Section 4.2: 08–09 Oct 2013. This was the only storm without a significant downward apparent wind. A large brightness enhancement occurred, so one would expect O^+ contamination to be occurring, if the hypothesis is correct. The only other feature of the observations distinguishing this storm from the others that we can identify is that it lacks a strong equatorward wind. In Chapter 5, we propose an alternative hypothesis, which provides a connection to the equatorward wind and is able to produce both downward and upward apparent vertical winds.

CHAPTER 5

ATMOSPHERIC SCATTERING EFFECTS

In this chapter we evaluate the hypothesis that atmospheric scattering can contaminate FPI measurements and cause artificial vertical wind observations. Specifically, molecules and aerosols in the troposphere can scatter photons into the field of view from outside of it, contaminating the wind and temperature measurements. *Abreu et al.* [1983] studied this effect for the auroral region, concluding that the large convergence often seen near auroral arcs could be an artifact of atmospheric scattering, but the topic has not been revisited, and it has not been applied in the context of midlatitude or low-latitude winds.

We argue that now is the time to do so, for four reasons. First, the widespread deployment of FPIs in recent years (and their networking) has led to the discovery of new phenomena, such as large, widespread vertical winds, that may be explained by scattering. Second, the literature has forgotten or neglected this effect. Except for a recent citation in association with the present work [*Miladinovich et al.*, 2016], the *Abreu et al.* [1983] result has not been mentioned in the literature in over 20 years, since *Price et al.* [1995]. Third, the advances in computing power in the past 34 years allows us to solve the full radiative transfer equations instead of using approximations. As discussed below, we can solve for the full angular distribution of scattered light instead of using a two-term Legendre polynomial expansion of the zenith-angle distribution and neglecting azimuthal dependence, as is done by *Abreu et al.* [1983]. It is likely these simplifications which led *Abreu et al.* [1983] to falsely conclude that temperature measurements are not affected by atmospheric scattering, a critical point that led to the dismissal of this effect in the citing literature. Fourth, we now have a much better characterization of the radiative characteristics of tropospheric aerosols, the primary contributor to scatter, thanks to the growth of monitoring networks such as the Aerosol Robotic Network (AERONET) [*Holben et al.*, 2001].

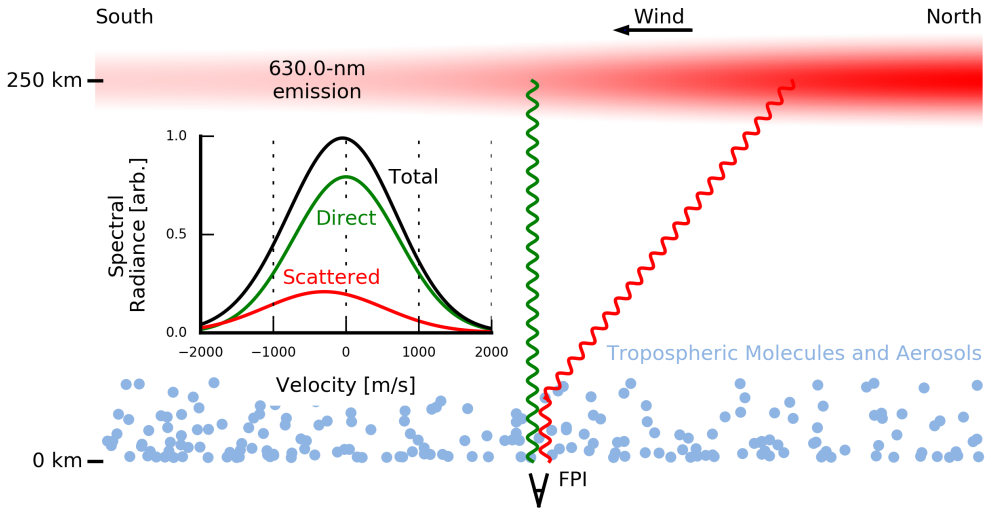

Figure 5.1: Airglow radiation can scatter in the lower atmosphere and cause errors in the wind measured by a ground-based FPI. This example depicts stormtime conditions at midlatitude. Green represents radiation that is not scattered, while red represents the scattering of light from outside the field of view into the field of view. Only one red ray is drawn, but light is incident from all directions in reality. Not to scale. The inset displays an example spectrum measured by a vertically looking FPI. The x-axis is equivalent Doppler velocity instead of wavelength. The total spectrum is affected by the Doppler shift of the scattered light, and is falsely interpreted to represent a vertical wind. Adapted from *Harding et al.* [Submitted].

Although the atmosphere is mostly transparent at 630.0 nm, it is not perfectly transparent. Under typical conditions, 10% of the incident photons collide with molecules and aerosols (i.e., particulates such as dust or smoke) in the lower atmosphere before reaching the ground [*Bodhaine et al.*, 1999; *Holben et al.*, 2001]. However, the aerosol content varies significantly, depending on location, season, and humidity, among others.

The atmospheric scattering hypothesis is shown schematically in Figure 5.1. Stormtime conditions at midlatitudes are shown: an auroral airglow enhancement exists in the north, and the wind is southward. Most of the photons observed by the vertically looking FPI originated within the field of view, directly above the FPI. The path of these photons is indicated in green. However, due to scattering, there is a component of the observation

that arises from photons outside the field of view. An example path is shown in red. Because the airglow is brightest in the north, the majority of these contaminating photons originate in the north, where they were imbued with a blueshift due to the southward wind. When these photons scatter into the field of view of the FPI, their blueshift is preserved, and they appear to indicate a downward vertical wind. The total spectrum observed by the FPI is shown in the inset. Even though the direct photons dominate the spectrum, even a small amount of scattering can cause a significant Doppler shift (as shown, equivalent to a -50 m/s wind).

It should be mentioned that although the description above focused on the specific case of a vertical wind measurement from a narrow-field FPI at mid-latitudes during storms, atmospheric scattering can be significant in many other situations. It affects wide-field FPIs, Michelson interferometers, all-sky imagers – any ground-based optical instrument. In addition to the vertical wind, horizontal wind and temperature measurements are also potentially contaminated. Scattering can be effective anywhere there is a bright airglow region outside the field of view co-located with a large line-of-sight wind. This can occur in the auroral region and near the edge of the equatorial arcs, for example, not just in the midlatitudes during stormtime. Figure 5.1 depicts single scattering, but in reality, scattered photons can (and if scattered horizontally, almost certainly will) scatter again. Thus, even for a case of near-transparency, it is important to consider multiple scattering.

In Section 5.1, we discuss relevant background concepts of the theory of radiative transfer, which describes the transmission of light in a medium which scatters, absorbs, and emits. The numerical model we developed to calculate the scattered light field is described in detail in Section 5.2. In Section 5.3, we model the stray light susceptibility of the FPI, which is a key component of the observation model, while Section 5.4 generalizes the model to include wavelength dependence. In Sections 5.5 and 5.6, we apply this model to data from midlatitudes and equatorial latitudes, respectively. Finally, Section 5.7 discusses the correction of FPI data for the presence of scattering.

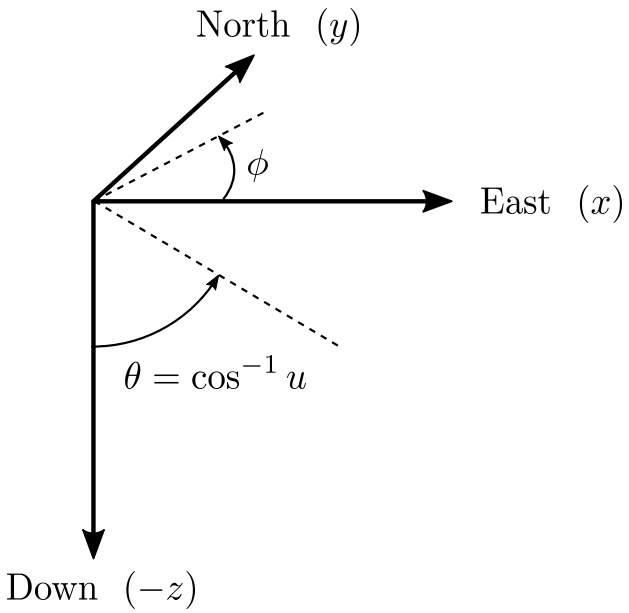

Figure 5.2: Coordinate system and angular variables, u and ϕ, used in this chapter.

5.1 Radiative Transfer Theory

Radiative transfer theory is described extensively in the literature. The development in this section is based largely on the works by *Chandrasekhar* [1960], *Sobolev* [1975] and *Hansen and Travis* [1974].

The fundamental quantities describing the light field in a medium are the specific intensity, I, and the emission coefficient, ε. Both quantities are defined as a function of wavelength, angle, and location in space. In this work, we ignore the wavelength dependence because we only consider elastic scattering, which preserves wavelength. Thus, we avoid the complexities associated with frequency redistribution, and all problems we solve are monochromatic. We also ignore the effects of polarization since airglow radiation is unpolarized, though we note that scattering (especially Rayleigh scattering) can be polarizing. According to *Hansen and Travis* [1974], ignoring polarization effects for sunlight-scattering problems introduces an error less than 10% for Rayleigh scattering and less than 1% for Mie scattering.

The angular coordinates typically used in radiative transfer theory are u and ϕ, shown in Figure 5.2. The direction cosine of the zenith angle, u is defined so that down $(-z)$ is +1 and up $(+z)$ is -1. East is x and north is y. The azimuth angle, ϕ, is defined using the mathematical convention (degrees

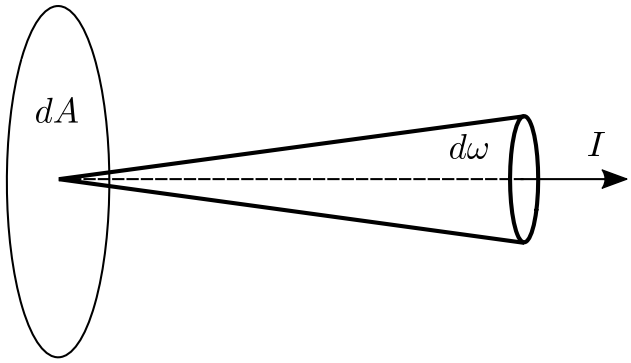

Figure 5.3: The specific intensity, I, of radiation incident upon a differential area element, dA, in a cone of differential solid angle, $d\omega$.

north of east), rather than the typical geophysical convention (degrees east of north, as in Chapters 2 and 3).

The specific intensity, I, often just referred to as the intensity, describes the differential energy carried by a light ray. Consider Figure 5.3. The energy carried by $I(u, \phi)$ through a differential area element, dA, in a differential solid angle, $d\omega$, surrounding the direction (u, ϕ), per unit time, dt, in a small frequency interval, $d\nu$, is

$$I(u, \phi)\, dA\, d\omega\, d\nu\, dt. \tag{5.1}$$

The emission coefficient, ε, is somewhat simpler to understand. It describes the differential energy of sources of radiation. The energy emitted by the medium described by $\varepsilon(u, \phi)$ within a small volume, dV, into a differential solid angle, $d\omega$, surrounding the direction (u, ϕ), per unit time, dt, in a small frequency interval, $d\nu$, is

$$\varepsilon(u, \phi)\, dV\, d\omega\, d\nu\, dt. \tag{5.2}$$

If $I(u, \phi)$ and $\varepsilon(u, \phi)$ are known at every location in the medium, then any desired radiative quantity can be calculated (e.g., flux, energy density, or transmissivity). In free space, the intensity remains constant along rays:

$$\frac{dI(u, \phi)}{ds} = 0 \tag{5.3}$$

where s parameterizes the ray path. When the propagation medium emits, scatters, or absorbs, the intensity can change along a ray. The equation that

describes this change is the equation of transfer:

$$\frac{dI(u,\phi)}{ds} = -\chi I + \varepsilon \tag{5.4}$$

where χ is the extinction coefficient. Extinction has two components, absorption and scattering:

$$\chi = \alpha + \sigma. \tag{5.5}$$

The absorption coefficient, α, describes the rate at which radiation is lost along a ray. The scattering coefficient, σ, also contributes to loss of radiation along a particular direction, but it causes additional radiation to be added into a different direction. In this way, ε has two sources: one from the medium's internal emission, and one from scattered light. The contribution to the emission coefficient at a given point from scattered light alone is given by an integral over all the rays incident on that point:

$$\varepsilon(u,\phi) = \frac{\sigma}{4\pi}\int_0^{2\pi}\int_{-1}^{1} P\left(u,u',\phi,\phi'\right) I\left(u',\phi'\right) du' d\phi' \tag{5.6}$$

where $P\left(u,u',\phi,\phi'\right)$ is known as the scattering phase function. It is a probability density function over u and ϕ, characterizing the probability of a photon scattering into the direction (u,ϕ) given that a scattering event occurred and the photon was originally traveling in the direction (u',ϕ'). It is defined so its integral over all directions (u,ϕ) is 4π. Rayleigh scattering is nearly isotropic, while Mie scattering off of aerosols is very highly forward-scattering. The specification of $P\left(u,u',\phi,\phi'\right)$ is discussed in more detail in Section 5.5.

It is convenient in radiative transfer problems to work with the ratio of the emission coefficient to extinction coefficient, which has the same units as I. This quantity is known as the source function, J:

$$J(u,\phi) = \frac{\varepsilon(u,\phi)}{\chi}. \tag{5.7}$$

Rewriting (5.4) and (5.6) in terms of J yields the two fundamental equations of radiative transfer, which must be satisfied at every location in the solution domain:

$$\frac{dI(u,\phi)}{\chi ds} = -I + J \tag{5.8}$$

$$J(u,\phi) = \frac{\tilde{\omega}}{4\pi} \int_0^{2\pi} \int_{-1}^{1} P\left(u, u', \phi, \phi'\right) I\left(u', \phi'\right) du' d\phi' \tag{5.9}$$

where we have defined the single-scattering albedo, $\tilde{\omega}$ as the fraction of photon collision events which are scattering events, not absorption events:

$$\tilde{\omega} = \frac{\sigma}{\chi}. \tag{5.10}$$

For the remainder of this work we will assume $\tilde{\omega} = 1$. In reality, based on a survey of AERONET measurements, we have found that $\tilde{\omega}$ varies from about 0.75 to 1 at 630.0 nm. However, our results are not significantly sensitive to $\tilde{\omega}$ since we are interested in the ratio of scattered to direct light, not the absolute radiance.

These equations do not include source terms. Sources can either be included as boundary conditions for the differential equation for I, or as an extra term in the equation for J.

For a specific case of a plane-parallel atmosphere illuminated from above, the problem can be simplified. This simplification assumes that I and J can be approximated as depending solely on altitude, not latitude and longitude. This is reasonable because due to the exponential atmospheric density profile, the majority of scattering occurs in the first few km above the ground, where the mean free path of a photon is much smaller than the horizontal variations of the scattered light field. We verified this by calculating the source function for singly scattered photons over two dimensions (altitude and latitude, since the primary horizontal variation is with latitude, not longitude). Over the simulation domain (10 horizontal optical depths at the ground), it varied only $\pm 6\%$ horizontally.

Of course, the solutions for I and J will vary significantly across the globe, but to solve for I and J at a given latitude and longitude, we can safely assume local spherical symmetry for the scattered light field. However, the latitude- and longitude-dependence of the initial airglow source is of utmost importance to include.

The details of the plane-parallel simplification are not shown here. Briefly, this simplification comprises defining altitude in terms of optical depth, and

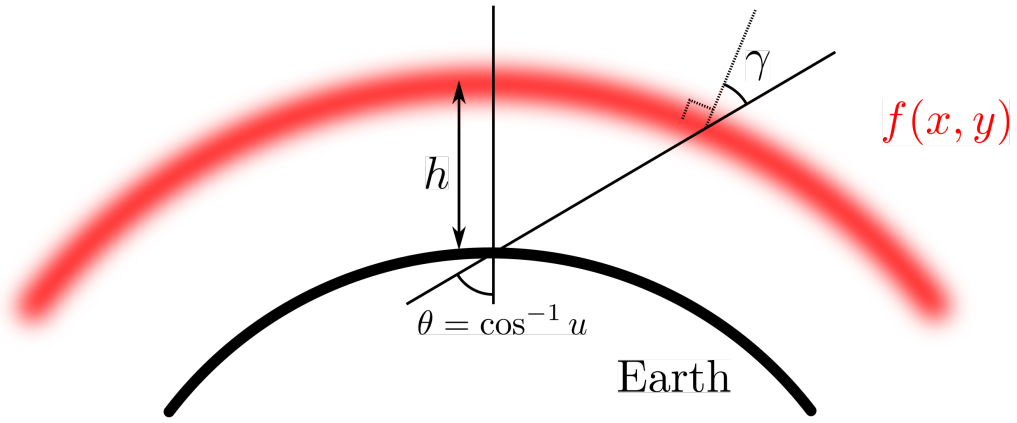

Figure 5.4: Geometry used for the calculation of $J_0(\tau, u, \phi)$. Not to scale.

separating the direct and scattered components of I and J. The optical depth, τ, at the altitude z, is defined as the integral of the extinction coefficient:

$$\tau(z) = \int_z^\infty \chi(z')\, dz'. \tag{5.11}$$

The value of τ runs from 0 at the fictional "top of the atmosphere" to τ_0, known as the optical thickness, at the ground.

The variables I and J, while defined above to include all radiation, are now understood to represent only those photons that have scattered at least once. The direct, unscattered radiation from the source emission will be handled separately. Thus, the plane-parallel radiative transfer equations to be solved are

$$u\frac{dI(\tau, u, \phi)}{d\tau} = -I(\tau, u, \phi) + J(\tau, u, \phi) \tag{5.12}$$

$$J(\tau, u, \phi) = \frac{\tilde{\omega}}{4\pi}\int_0^{2\pi}\int_{-1}^{1} P\left(u, u', \phi, \phi'\right) I\left(\tau, u', \phi'\right) du' d\phi' + J_0(\tau, u, \phi) \tag{5.13}$$

where we have included the source term in J, which accounts for the direct, unscattered radiation. This general source term, $J_0(\tau, u, \phi)$ takes the place of the specific source term usually used in the literature to represent a point source such as the sun (e.g., Equation 3.36 from *Hansen and Travis* [1974], which we note is missing the factor of $\frac{\tilde{\omega}}{4\pi}$).

To compute $J_0(\tau, u, \phi)$, we must know the distribution of source airglow brightness with latitude and longitude. We approximate the airglow layer as a thin shell of isotropic emission at $z = h \equiv 250$ km, as shown in Figure 5.4. Its vertical column brightness (i.e., the intensity in the downward direction,

evaluated just below the layer), is described by a function $f(x, y)$, which is an input to our model. It may be specified in arbitrary units since our analysis does not depend on absolute calibration. The location variables x and y are specified in km. In this work, we will determine $f(x, y)$ using an all-sky camera, as described below. Once $f(x, y)$ is known, $J_0(\tau, u, \phi)$ can be calculated with an integral analogous to that in (5.13):

$$J_0(\tau, u, \phi) = \frac{\tilde{\omega}}{4\pi} \int_0^{2\pi} \int_{-1}^{1} P\left(u, u', \phi, \phi'\right) I_0\left(\tau, u', \phi'\right) du' d\phi' \tag{5.14}$$

where $I_0(\tau, u, \phi)$ is the intensity caused by unscattered rays directly from the source. To relate this to $f(x, y)$, we must consider extinction and geometry. Since most of the scattering occurs at the lowest altitudes, for extinction we can use the plane-parallel approximation, with an attenuation factor $\exp\left(-\frac{\tau}{u}\right)$. A geometrical factor must be included to account for the fact that off-zenith rays have a longer path length through the emitting layer and are thus brighter, a phenomenon known in the aeronomy literature as the "van Rhijn effect." This extra term is the secant of the ray's zenith angle where it intersects with the emitting layer, an angle we define as γ (see Figure 5.4). For large zenith angles, this is not quite the same as θ due to the curvature of the earth. To calculate γ and the location (x, y) where the ray intersects the emitting layer, we use spherical-Earth geometry:

$$\gamma = \sin^{-1}\left(\frac{R_E}{R_E + h}\sqrt{1 - u^2}\right) \tag{5.15}$$

$$\rho = (R_E + h)\left(\cos^{-1} u - \gamma\right) \tag{5.16}$$

$$x = -\rho \cos\phi \tag{5.17}$$

$$y = -\rho \sin\phi. \tag{5.18}$$

These equations are analogous to (3.2)−(3.4) except for the different convention for the angular variables.

With extinction and geometry, we write the source intensity as

$$I_0(\tau, u, \phi) = \sec\gamma \, \exp\left(-\frac{\tau}{u}\right) f\big(x(u, \phi), y(u, \phi)\big). \tag{5.19}$$

Plugging this into (5.14) yields the desired expression for J_0 in terms of

$f(x, y)$:

$$J_0(\tau, u, \phi) = \frac{\tilde{\omega}}{4\pi} \int_0^{2\pi} \int_0^1 P\left(u, u', \phi, \phi'\right) \sec\gamma' \exp\left(-\frac{\tau}{u'}\right) f(x', y')\, du' d\phi' \,. \tag{5.20}$$

All variables of integration are primed. Although not explicitly denoted, x' and y' are functions of u' and ϕ'.

We could in principle solve the integro-differential equation set represented by (5.12) and (5.13) directly, but it is more convenient to first solve the differential equation in order to convert to a system of integral equations. The boundary conditions for (5.12) are simple; there is no scattered light arriving from above the atmosphere or from below the ground:

$$\begin{aligned} I(0, u, \phi) &= 0 \qquad \text{for } u > 0 \\ I(\tau_0, u, \phi) &= 0 \qquad \text{for } u < 0. \end{aligned} \tag{5.21}$$

We note that *Abreu et al.* [1983] used a nonzero ground albedo, in which case the second boundary condition should be

$$I(\tau_0, u, \phi) = \alpha_g\, I(\tau_0, -u, \phi) \tag{5.22}$$

where α_g is the ground albedo, which can be as large as 0.8 during snow cover. However, we performed a sensitivity study of the *Abreu et al.* [1983] model to α_g and found negligible sensitivity, so we simply set it to zero here.

With these boundary conditions, the solution of the differential equation for I is

$$I(\tau, u, \phi) = \begin{cases} \displaystyle\int_0^{\tau} \frac{1}{u} J(\tau', u, \phi) \exp\left[\frac{-(\tau - \tau')}{u}\right] d\tau & \text{if } u > 0 \\ \displaystyle -\int_{\tau}^{\tau_0} \frac{1}{u} J(\tau', u, \phi) \exp\left[\frac{-(\tau - \tau')}{u}\right] d\tau & \text{if } u < 0. \end{cases} \tag{5.23}$$

Taken together, (5.23) and (5.13) compose the final system of plane-parallel integral equations. In this form, the challenge of radiative transfer problems is evident: I depends on J, and J depends on I. In Section 5.2, we describe the algorithm used to solve these equations.

5.2 Numerical Solution

To solve the radiative transfer equations, we use the method of successive orders of scattering, described by *Hansen and Travis* [1974]. We first discretize I on a grid of M optical depths, N zenith angles, and R azimuth angles. We have found that we can achieve a model accuracy equivalent to a 5 m/s apparent vertical wind by using $M = N = R = 10$. Higher accuracies can be achieved with a denser grid, but runtime increases as the square of M, N, and R. The grid spacing is given by

$$\Delta\tau = \frac{\tau_0}{M-1} \qquad \Delta u = \frac{2}{N} \qquad \Delta\phi = \frac{2\pi}{R}. \tag{5.24}$$

The values of I on this grid are stored in a three-dimensional array, $\mathbf{I}$. J is stored similarly in $\mathbf{J}$, which is defined on the same angular grid, except in order to simplify the evaluation of (5.23), it is defined at the midpoints of the optical depth grid on which I is defined:

$$\mathbf{I}_{m,n,r} \equiv I\Big(\tau_m, \quad u_n, \, \phi_r\Big) \quad \forall m \in [0, M-1], n \in [0, N-1], r \in [0, R-1] \tag{5.25}$$

$$\mathbf{J}_{m,n,r} \equiv J\Big(\tau_m + \frac{\Delta\tau}{2}, \quad u_n, \, \phi_r\Big) \quad \forall m \in [0, M-2], n \in [0, N-1], r \in [0, R-1] \tag{5.26}$$

where the grid points are

$$\tau_m = m\Delta\tau \qquad u_n = (n + 0.5)\,\Delta u \qquad \phi_r = (r + 0.5)\,\Delta\phi. \tag{5.27}$$

As an alternative to using a grid to discretize the angular dimensions, it is common to use a truncated Legendre polynomial expansion [e.g., *Chandrasekhar*, 1960; *Sobolev*, 1975]. Although the scattered light distribution is never exactly isotropic, it is often smooth, so the use of smooth basis functions could offer a significant improvement in runtime, but we have not tried implementing it.

The method of successive orders of scattering is appropriate for problems where most of the photons are scattered a small number of times, i.e., $\tau \lesssim 1$, which is certainly true in our case. It first computes the contribution from photons scattered once, then computes the fraction of those photons that

are scattered again, and so on. Defining the superscript k as the iteration number, and using a Riemann-sum approximation to the integrals in (5.23) and (5.13), one iteration is

$$\mathbf{I}^k_{m,n,r} = \begin{cases} \displaystyle\sum_{m'=0}^{m-1} \frac{1}{u_n} \mathbf{J}^k_{m',n,r} \exp\left[-\frac{\Delta\tau}{u_n}\left(m - m' - \frac{1}{2}\right)\right]\Delta\tau & \text{if } u_n > 0 \\ \displaystyle -\sum_{m'=m}^{M-2} \frac{1}{u_n} \mathbf{J}^k_{m',n,r} \exp\left[-\frac{\Delta\tau}{u_n}\left(m - m' - \frac{1}{2}\right)\right]\Delta\tau & \text{if } u_n < 0 \end{cases} \tag{5.28}$$

$$\mathbf{J}^{k+1}_{m,n,r} = \frac{\tilde{\omega}}{4\pi} \sum_{r'=0}^{R-1} \sum_{n'=0}^{N-1} P(u_n, u_{n'}, \phi_r, \phi_{r'}) \left[\frac{\mathbf{I}^k_{m,n',r'} + \mathbf{I}^k_{m+1,n',r'}}{2}\right] \Delta u \Delta\phi. \tag{5.29}$$

The initial condition for this iteration is set by evaluating (5.20), where we again use a Riemann-sum approximation for the integrals, but use a higher angular resolution $\left(\tilde{N} = \tilde{R} = 20\right)$ because the brightness distribution $f(x, y)$ has more small-scale structure than the scattered light field:

$$\mathbf{J}^0_{m,n,r} = \frac{\tilde{\omega}}{4\pi} \sum_{r'=0}^{\tilde{R}-1} \sum_{n'=0}^{\tilde{N}-1} P(u_n, u_{n'}, \phi_r, \phi_{r'}) \sec\gamma' \exp\left(-\frac{\tau_m + \frac{\Delta\tau}{2}}{u_{n'}}\right) f(x', y') \Delta u \Delta\phi. \tag{5.30}$$

We iterate until the relative change in the solution is less than 10^{-4}, which usually occurs after only a few iterations. After calculating the intensity for all scattering orders, the total intensity is computed by summing over the orders:

$$\mathbf{I}_{m,n,r} = \sum_{k=0}^{K-1} \mathbf{I}^k_{m,n,r} \tag{5.31}$$

where K is the total number of iterations. $\mathbf{J}$ can be computed similarly, but it is not needed for subsequent calculations.

$\mathbf{I}$ provides the scattered component of the brightness that would be measured by an instrument at any altitude, looking in any direction. To model an ideal narrow-field FPI measurement, we must simply evaluate $\mathbf{I}$ at $\tau = \tau_0$

(the ground), and at the direction (u_0, ϕ_0) which is antiparallel to the boresight of the FPI. We define this measured brightness as g_{sc}:

$$g_{\text{sc}} = I(\tau_0, u_0, \phi_0). \tag{5.32}$$

If there is not a grid point at this exact angle, two-dimensional spline interpolation is used.

We must also consider the contribution from photons directly from the source that never scatter, defined as g_{dir}. This is computed analogously to (5.19):

$$g_{\text{dir}} = \sec\gamma_0 \exp\left(-\frac{\tau_0}{u_0}\right) f\big(x(u_0, \phi_0), y(u_0, \phi_0)\big). \tag{5.33}$$

However, real FPIs are not ideal narrow-field instruments. They have finite fields of view, and, more importantly, do not have perfect rejection of light from outside their field of view. This extra contribution from outside the field of view is known as "stray light," and it represents another source of contamination that must be accounted for.

5.3 Stray Light Model

In this section we describe our stray light model, which generalizes (5.32) and (5.33) to realistic FPIs. First, we define an observation function, $s(\theta)$, which is the fraction of photons incident on the aperture from an angle θ which is detected by the instrument, where θ is the angle from boresight. We assume that this is not dependent on the azimuth angle. It is normalized to 1 inside the field of view (out to $\theta \approx 1^\circ$), and because of the reflection of light off of the protective dome and entrance optics, it is non-zero outside the field of view.

To measure $s(\theta)$, we briefly re-located the UAO FPI to a darkroom at the University of Illinois. By recording the brightness measured from a point source at various angles outside the field of view, and normalizing by the brightness measured looking directly at the point source, we build $s(\theta)$, shown in Figure 5.5. A fifth-order polynomial is fit to the logarithm of the data in order to interpolate between the measurements. Outside the field of view, $s(\theta)$ drops quickly to 10^{-3} at 10° and falls to about 2×10^{-5} at angles larger than 45°. Lab testing indicated that about half of the stray light is due to

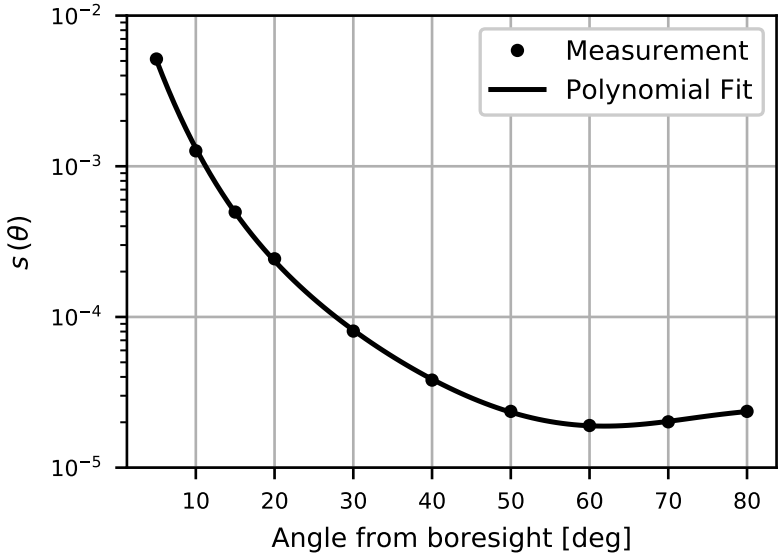

Figure 5.5: Lab measurements of the stray light observation function, $s(\theta)$, with a fifth-order polynomial fit. The component inside the field of view (equal to 1) is omitted.

reflections off the dome, and the other half is due to reflections off of the aperture window on the SkyScanner. After a few years in the field, these can become quite dirty.

Although $s(\theta)$ appears small, its effects are significant once integrated over the entire sky. By including stray light in the observation model, the apparent vertical wind is increased by ∼30–50%; thus, it is a secondary effect, but it is important to include. Although only measured for UAO, we assume that $s(\theta)$ also applies to the MH and CAR FPIs used in the case studies below. This assumption is a significant source of uncertainty in our model, but it is necessary due to the difficulty of measuring $s(\theta)$ for MH (which is less mobile) and CAR (which is in Brazil).

The generalization of (5.32) and (5.33) is achieved by integrating over all angles, weighting the incident intensity by $s(\theta)$:

$$g_{\text{sc}} = \int_0^{2\pi} \int_{-1}^{1} s(\theta)\, I(\tau_0, u, \phi)\, du\, d\phi \tag{5.34}$$

$$g_{\text{dir}} = \int_0^{2\pi} \int_{-1}^{1} s(\theta)\, \sec\gamma\, \exp\left(-\frac{\tau_0}{u}\right) f(x, y)\, du\, d\phi \tag{5.35}$$

where θ is the angle between (u, ϕ) and (u_0, ϕ_0):

$$\theta = \cos^{-1}\left(uu_0 + \sqrt{(1-u^2)(1-u_0^2)}\cos\left(\phi - \phi_0\right)\right). \tag{5.36}$$

These integrals are carried out with Riemann-sum approximations, as above.

So far, we have described how our model computes the scattered and direct brightness measured by an FPI, given the source brightness distribution, $f(x, y)$, the stray light observation function, $s(\theta)$, and radiative properties of the atmosphere, τ_0 and P. However, we have only described a monochromatic problem. To calculate the errors in apparent wind and temperature measured by an FPI, we must compute the spectrum of scattered light.

5.4 Computation of Spectrum

Fortunately, since Rayleigh and Mie scattering preserve wavelength, we can solve the spectral problem by solving a series of monochromatic problems. This assumes the motion of scattering particles can be neglected, which is always valid for aerosols and is a suitable approximation for molecules since the wind and temperature in the troposphere is smaller than in the thermosphere.

We discretize the spectrum into $L = 20$ wavelength bins and run the model described above for each bin. The only quantity that changes with wavelength is $f(x, y)$. While τ_0 and P technically depend on wavelength, the wavelength band of interest is only a few pm, which is negligible. The wavelength-dependence of $f(x, y)$ is a Gaussian, with a width based on the temperature and a Doppler shift based on the line-of-sight wind at each location (see (2.21)–(2.23)). The line-of-sight wind is the dot product of the true wind vector with the unit vector from the emission location (x, y, h) to the FPI location $(0, 0, 0)$. It is assumed that the horizontal wind and temperature are constant over the domain, and that the actual vertical wind is zero. The wind and temperature are an input to our model.

Once the model has been run for each wavelength bin, we add the direct and scattered spectra to obtain the total measured spectrum. By fitting a Gaussian, we determine the modeled apparent line-of-sight wind and temper-

ature. The apparent vertical wind is calculated by using a vertically directed boresight (i.e., $u_0 = 1$).

5.5 Application to Midlatitudes: 02 Oct 2013 Storm

In this section we apply our radiative transfer model to the 02 Oct 2013 geomagnetic storm, when large apparent vertical winds were observed by the midlatitude NATION and MH FPIs, as described in Section 4.1. We use our model to test whether the apparent vertical winds seen by the MH FPI can be explained by atmospheric scattering. We choose the MH FPI because it has a co-located all-sky camera operated by the Boston University Imaging Group.

We begin by describing how we specify the inputs to the model, namely, the brightness distribution, $f(x, y)$, the radiative characteristics of the atmosphere, τ_0 and P, and the horizontal wind and temperature.

5.5.1 Brightness, $f(x, y)$

The Boston University imager takes 630.0-nm all-sky images at a 400-sec cadence. Calibration of the raw data is described in the Appendix of *Baumgardner et al.* [2008]; briefly, the calibration routine corrects for vignetting, spectral background, and flat fielding. After calibration, the data provide a map of the measured brightness across the sky, integrated over the 630.0-nm line shape. However, to relate this to $f(x, y)$, we must account for atmospheric extinction, scattering, and geometry.

The measured intensity at the pixel sensitive to the light ray propagating in direction (u, ϕ) can be written as a combination of the scattered intensity and direct intensity. If we let $I_{\text{meas}}(u, \phi)$ denote the measured intensity, this is given by

$$I_{\text{meas}}(u, \phi) = I(\tau_0, u, \phi) + \sec\gamma \, \exp\left(-\frac{\tau_0}{u}\right) \, f(x, y). \tag{5.37}$$

This equation cannot be solved directly for $f(x, y)$, since the scattered intensity, $I(\tau_0, u, \phi)$ implicitly depends on $f(x, y)$ through the radiative transfer

model. However, an iterative approach can be used:

$$f^{k+1}(x,y) = \frac{I_{\text{meas}}(u,\phi) - I^k(\tau_0,u,\phi)}{\sec\gamma \, \exp\left(-\frac{\tau_0}{u}\right)} \tag{5.38}$$

where to compute the scattered intensity at iteration k, $I^k(\tau_0,u,\phi)$, we run the monochromatic radiative transfer model using the brightness distribution at iteration k, $f^k(x,y)$. The initial step is $I^k(\tau_0,u,\phi) = 0$, and the process converges to a relative change of less than 10^{-4} after several iterations. This radiative transfer iteration can be thought of as an extra calibration step, which simultaneously accounts for atmospheric effects and the van-Rhijn effect.

This process seems somewhat paradoxical, since the whole point of finding $f(x,y)$ is to solve the radiative transfer problem, yet we need to solve a radiative transfer problem in order to determine $f(x,y)$. However, it is possible, since we are first solving a monochromatic problem to determine $f(x,y)$, then using $f(x,y)$ to solve the spectral problem.

The usable field of view of all-sky cameras, while assumed above to cover all (u,ϕ), is often blocked by trees or buildings in practice, and if not, the measurements near the horizon are often affected by interfering light sources. For these missing values in $f(x,y)$, we use a zero-order hold extrapolation, replacing them with the nearest pixels. This adds a small amount of uncertainty. Bilinear interpolation is used to determine $f(x,y)$ at an arbitrary location between pixels. Finally, we note that $I_{\text{meas}}(\tau_0,u,\phi)$ can be provided in arbitrary units, since we are only concerned with the ratio of scattered and direct light.

Stray Light

In addition to affecting FPI measurements, stray light also affects all-sky imager measurements, and we have found that it is important to include a stray light correction on the raw all-sky data, before the corrections mentioned above.

Consider Figure 5.6, which shows data from the all-sky imager without a stray light correction. The top panel is an example image from 07:34 UT on 02 Oct 2013, which shows a brightness enhancement in the north, which is at

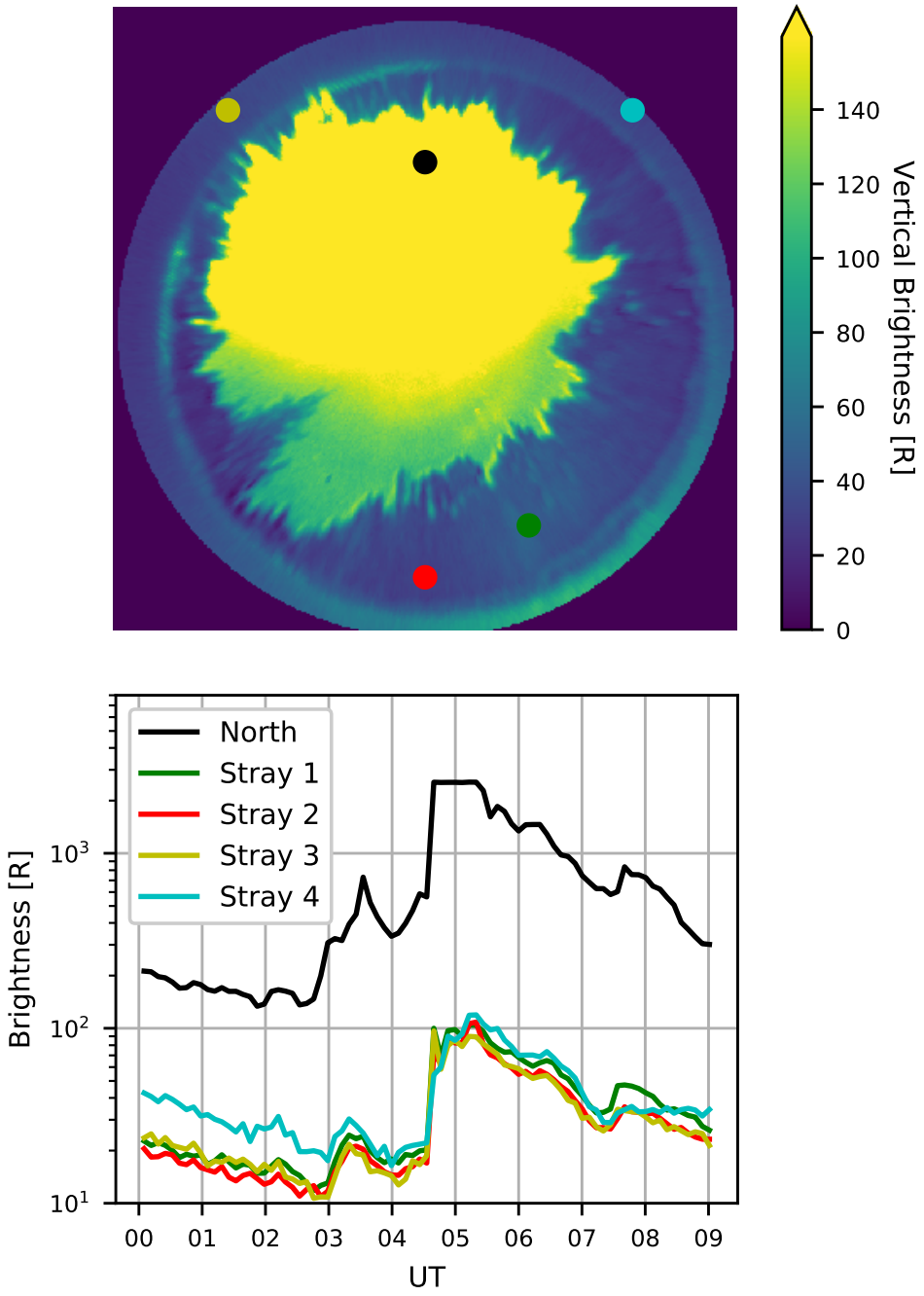

Figure 5.6: (Top) Raw all-sky image from the Boston University imager on 07:34 UT on 02 Oct 2013. Up is north, and the data are shown as measured on the pixel grid. The color scale has been zoomed to show stray light at the edges. Pixels below 15° elevation have been cropped. (Bottom) The time history of the brightness measured at the five pixels indicated by colored dots in the top panel.

the top of the image. Trees obstruct the field of view at low elevation angles, and pixels at even lower elevation angles ($< 15°$) have been cropped. The color scale has been constrained to show stray light clearly. The colored lines in the lower panel show the measured brightness at five locations indicated by the colored dots in the top panel. The black line indicates the brightness measured to the north, and the other colors are the measured brightness at pixels outside the field of view, some looking at trees and others looking even lower. The fact that these pixels correlate so well with the north pixel is an indication that, like with the FPI, light is scattering off of the dome and elsewhere in the optics, causing an artificial brightening in the dark regions of the sky.

To correct for this, we assume that this stray light causes a spatially uniform, but time-dependent increase in measured brightness, which we determine from the average of the four colored points indicated Figure 5.6. We subtract this value from the entire image before performing the correction in (5.38). We note that the stray light correction adds a considerable amount of uncertainty, since we expect that stray light is not uniform, but rather contains a significant specular component. It is difficult to characterize stray light susceptibility in the lab, and it is more difficult to correct for it. Though the procedure used here is not perfect, we believe it is a valid first approximation. That being said, it is the largest source of uncertainty in the model. Including this stray light correction increases the modeled apparent vertical wind by 30–80%.

After all corrections have been applied, we have a brightness distribution which is suitable for use in the radiative transfer model. An example $f(x, y)$ is shown in Figure 5.7, which is valid for 05:06 UT. The diffuse red aurora is seen in the north, and a thin SAR arc is seen at 40 °N. This figure also shows the mask used to crop out trees. As mentioned above, $f(x, y)$ is extrapolated to fill in these gaps.

5.5.2 Optical Thickness, τ_0, and Scattering Phase Function, P

The optical thickness, τ_0, has contributions from Mie scattering off of aerosols and Rayleigh scattering off of N_2 and O_2 molecules:

$$\tau_0 = \tau_0^{\text{aero}} + \tau_0^{\text{mol}}. \tag{5.39}$$

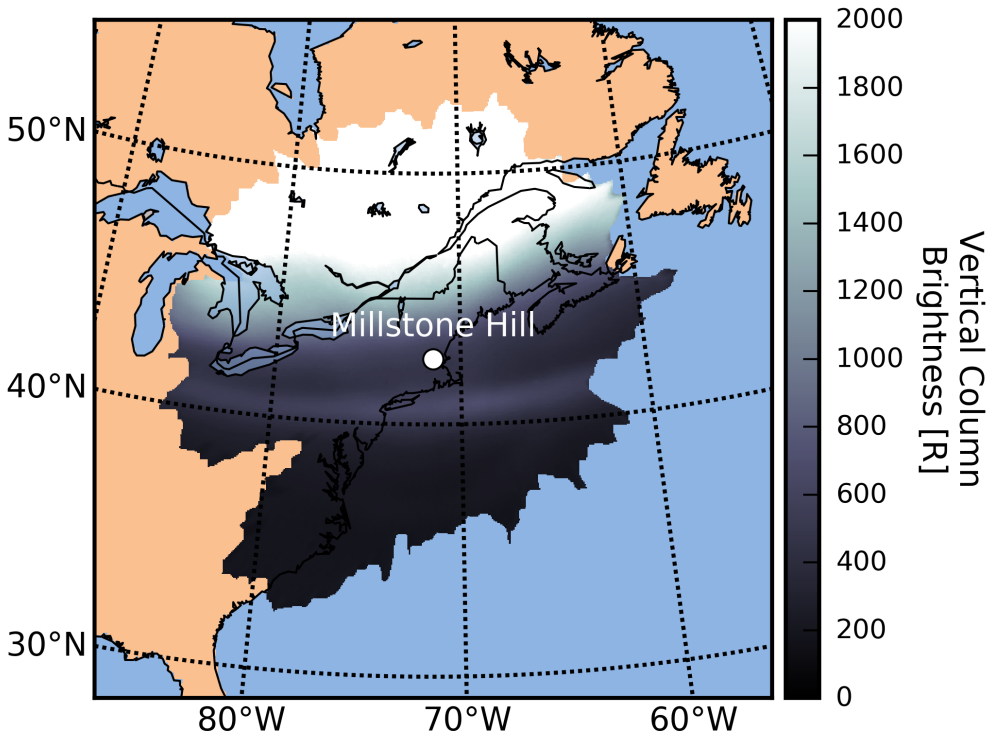

Figure 5.7: Calculated source airglow brightness distribution $f(x, y)$ at 05:06 UT on 02 Oct 2013, including the atmospheric and stray light corrections. The irregular cropping is due to trees blocking the field of view. Adapted from *Harding et al.* [Submitted].

We ignore a small ozone contribution, which mostly contributes to absorption, not scattering.

To specify the aerosol component, we use measurements from AERONET [*Holben et al.*, 2001], which estimates daytime τ_0^{aero} at multiple wavelengths and hundreds of locations around the Earth. There are two AERONET sites near Millstone Hill: Billerica, 20 km away, and Harvard_Forest, 60 km away. Figure 5.8 shows the AERONET optical thickness interpolated to 630.0-nm for 01–02 Oct 2013. It varies significantly from day to day and even hour to hour. Unfortunately, no measurements are available at night. We average the pre-sunset value on 01 Oct and the post-sunrise value on 02 Oct, and average between the two sites, obtaining $\tau_0^{\text{aero}} = 0.067$, which is assumed to be constant over the night. Given its daytime variability, this is a non-negligible source of uncertainty in our model. The Rayleigh component is

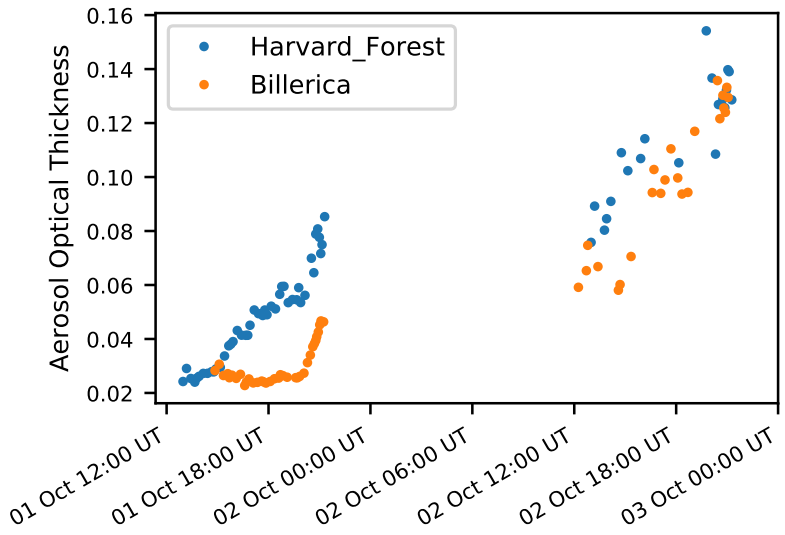

Figure 5.8: Aerosol optical thickness measured by two AERONET sites during the daytime hours of 01 and 02 Oct 2013, interpolated to 630.0 nm.

given by *Bodhaine et al.* [1999] as $\tau_0^{\text{mol}} = 0.0558$, yielding a total optical thickness of $\tau_0 = 0.123$. *Abreu et al.* [1983] used $\tau_0 = 0.27$.

AERONET also measures the aerosol scattering phase function. Like above, the total scattering phase function has two components:

$$P(\theta) = \frac{\tau_0^{\text{aero}} P^{\text{aero}}(\theta) + \tau_0^{\text{mol}} P^{\text{mol}}(\theta)}{\tau_0}. \tag{5.40}$$

This assumes that P is independent of altitude, which is reasonable since most of the scattering occurs near the ground, and aerosols and molecules both have exponential height distributions, though aerosols fall off quicker. Note that P is specified here in terms of θ, the scalar scattering angle, where above, it is specified in terms of the full incident and scattered directions (u, u', ϕ, ϕ'). The connection is simply geometrical and is given by (5.36).

Since our analysis is not particularly sensitive to the small daily changes in P^{aero}, we simply use the AERONET estimate from 21:00 UT on 02 Oct at 670 nm. P^{mol} takes the well-known form for Rayleigh scatter:

$$P^{\text{mol}}(\theta) = \frac{3}{4}\left(1 + \cos^2\theta\right). \tag{5.41}$$

Figure 5.9 shows the total scattering phase function, its two components, and the scattering phase function used by *Abreu et al.* [1983]. The scattering

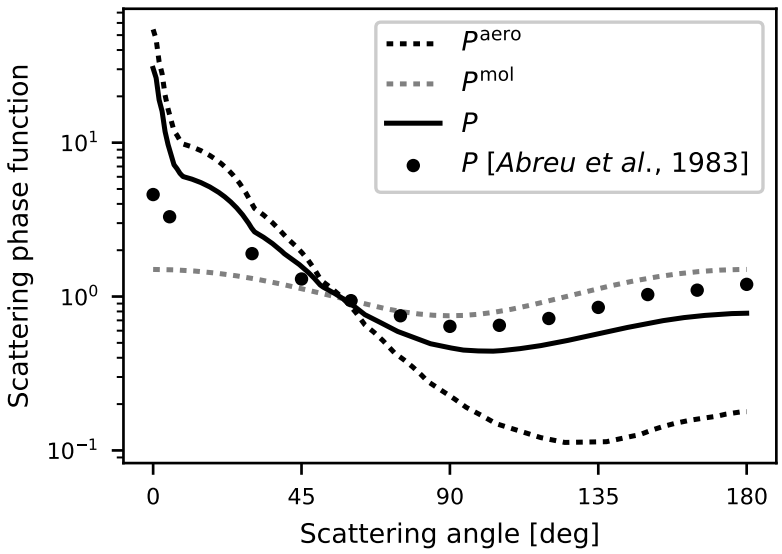

Figure 5.9: Scattering phase function, P, with its two components from molecular and aerosol scattering, P^{mol} and P^{aero}. The scattering phase function used by *Abreu et al.* [1983] is also shown.

phase function we use is significantly more forward-peaked than that used by *Abreu et al.* [1983]. Using the *Abreu et al.* [1983] phase function instead of the AERONET-derived phase function changes the resulting apparent vertical wind by 5 m/s.

5.5.3 Horizontal Wind and Temperature

The final inputs are the horizontal wind and temperature, which determine the line shape of $f(x, y)$. Unfortunately, the only available measurements of these are from FPI data, which have errors from atmospheric scattering. However, when the FPI is observing the brightest region of the sky, the ratio of scattered to direct light is low, and the measurement is minimally contaminated. This means that for this specific case (where there is a bright region in the north) we can use the north-looking measurement to determine the meridional wind and temperature, assuming they are spatially constant and there is no actual vertical wind. Since the east and west regions are dimmer than the north, the zonal wind estimate is contaminated by scattering. However, since the brightness gradient is almost purely meridional, our model results are not sensitive to the zonal wind input, so we set it to zero.

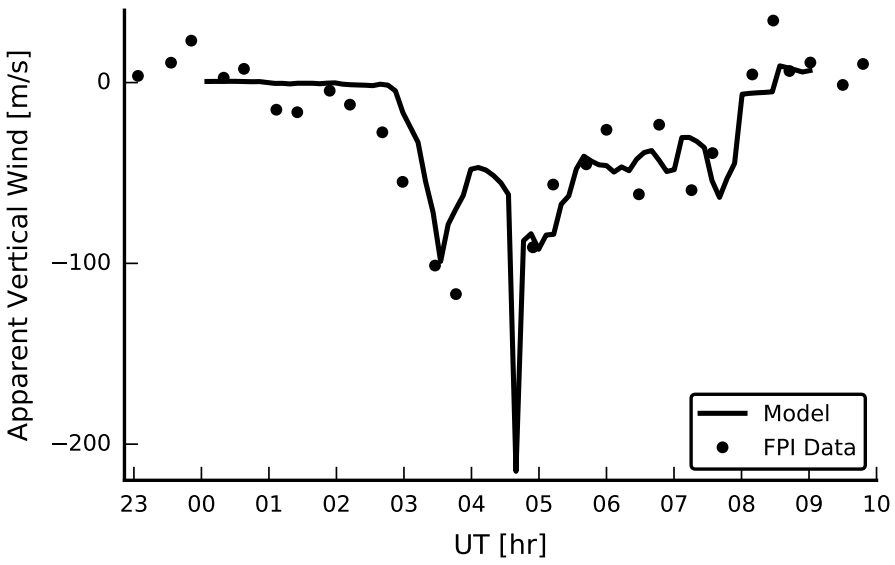

Figure 5.10: The modeled apparent vertical wind using the Boston University all-sky camera to determine the brightness distribution and the FPI to determine the meridional wind and temperature, compared to the measured apparent vertical wind by the FPI during the geomagnetic storm of 01–02 Oct 2013. Adapted from *Harding et al.* [Submitted].

5.5.4 Errors in Vertical Wind

The Boston University all-sky imager recorded eighty-one 630.0-nm images on the night of 01–02 Oct 2013. For each image, we determine $f(x, y)$ as described above, and interpolate the meridional wind and temperature from the FPI to the time of the all-sky image. With these inputs, we run the radiative transfer model and compute the apparent vertical wind. In Figure 5.10, we compare this modeled vertical wind with the apparent vertical wind measured by the MH FPI (which was previously shown in Figure 4.4). The statistical uncertainty of the FPI data is not shown since it is only 1–2 m/s.

Before 02:30 UT, both the measured and modeled vertical winds are near zero. Around 02:30–03:00 UT, a drop in the measured wind is followed shortly by a drop in the modeled wind. The reason for the time difference is unknown, but may be attributable to OH contamination since the dimmest airglow of the entire night occurs during 02:30–03:30 UT (see Figure 4.5). Between 03:00 and 08:00 UT, both the measured and modeled vertical wind are significant, fluctuating between about -50 and -100 m/s with good agreement. At 04:45 UT, for the duration of a single all-sky image, the modeled

vertical wind drops to -215 m/s. As mentioned in Chapter 4, no data are available from the MH FPI from 04:00 to 05:00 UT due to routine flat field calibration. However, as seen in Figure 4.4, all five NATION FPIs saw a similar downward surge at this time. The NATION surge lasted longer, likely because NATION is farther south than MH, and the bright region stayed well north of zenith for the NATION instruments. For the MH instrument, the bright region impinged quickly upon the zenith measurement, minimizing the period of large error. At 08:00 UT, both the modeled and measured vertical return to values near zero.

This correspondence between the data and the model strongly suggests that atmospheric scattering fully explains the large apparent vertical winds seen at midlatitudes during storms. Whether the oft-observed large vertical winds in the auroral region may be explained by atmospheric scattering is an open question. Since this region has stronger and more rapid forcing from Joule heating and particle precipitation, larger actual vertical winds are expected. However, given the large airglow gradients and horizontal winds, scattering is expected to be more effective as well. Some observational features, such as the nighttime upward (downward) winds seen poleward (equatorward) of the auroral arcs, seem to be obvious candidates for a scattering explanation.

5.5.5 Errors in Horizontal Wind and Temperature

In addition to errors in vertical wind estimates, atmospheric scattering causes errors in horizontal wind and temperature estimates, as we show with another model run, this time using an FPI looking southward with a zenith angle of 45° $\left(u_0 = \frac{\sqrt{2}}{2},\, \phi_0 = 90^\circ\right)$. This measurement is observing the darkest region of the sky and is thus significantly affected by scattering. The apparent line-of-sight wind and temperature are compared to the input meridional wind and temperature, and the error is computed, shown in Figure 5.11. During 03:00–08:00 UT, the line-of-sight wind has an error of 100–400 m/s in a direction toward the FPI, and the temperature is artificially high by 75–200 K. Since the north-looking measurement has minimal error, this would appear to indicate a large convergence in the horizontal wind field and an

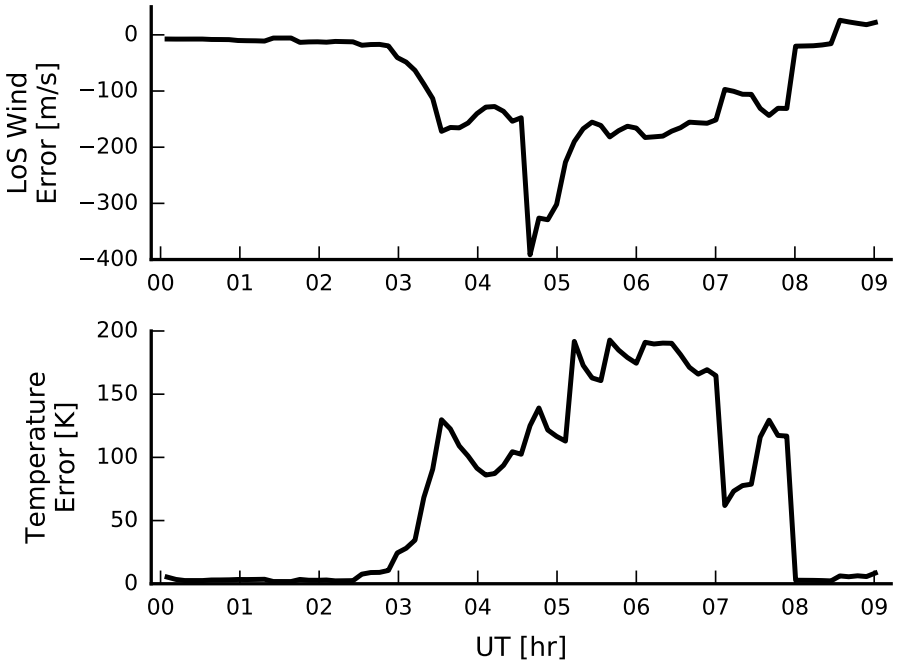

Figure 5.11: Modeled error in the line-of-sight wind and temperature measured by an FPI looking south with a 45° zenith angle during the geomagnetic storm of 01–02 Oct 2013. Line of sight is defined positive away from the FPI. Adapted from *Harding et al.* [Submitted].

apparent southward gradient in the temperature over the FPI, when no such features truly exist.

Quantitatively comparing this model result with data is complicated by the assumption of a spatially constant wind and temperature field. Although this is a suitable assumption for computing the scattered light spectrum, and for computing the apparent vertical wind, it is not necessarily reasonable for comparing the direct spectrum from the north-looking and the south-looking measurements. These points are separated by 500 km, so it is expected that the true wind and temperature will vary over this distance. A further complication is that the north-looking measurement has a data gap from 03:50 to 06:30 UT.

With these caveats, we show the measured line-of-sight wind and temperature from the MH FPI in Figure 5.12. If the wind field were constant, the true vertical wind were zero, and atmospheric scatter were negligible, the two lines in the top panel would lie on top of each other. Note that we have negated the north-looking line-of-sight wind, as it should have an opposite sign to

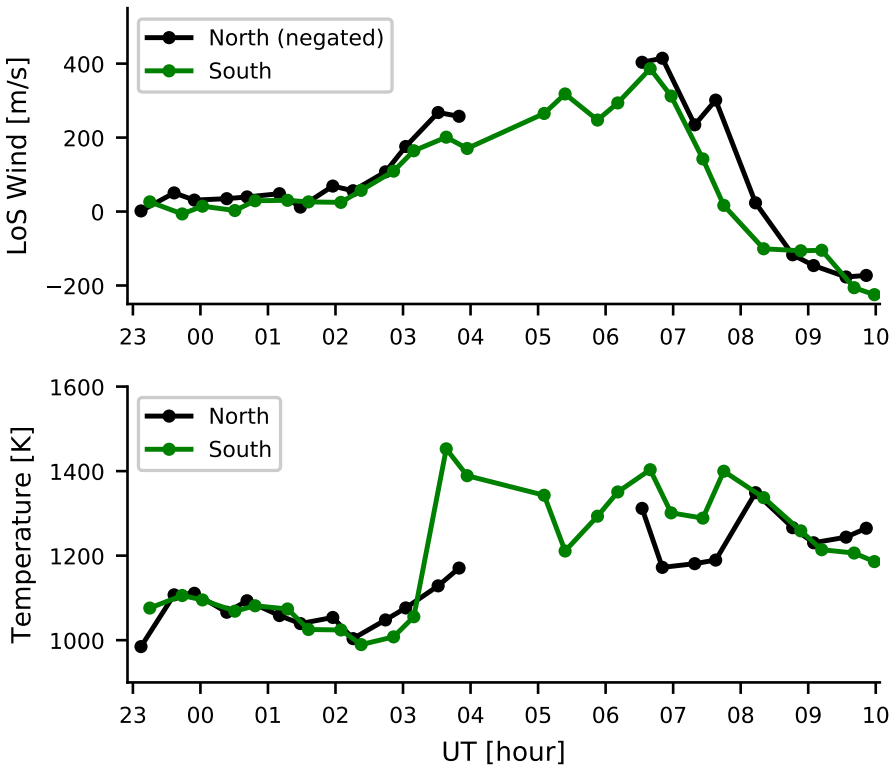

Figure 5.12: Measured line-of-sight wind and temperature by the MH FPI during the geomagnetic storm of 01–02 Oct 2013. Line of sight is defined positive away from the FPI.

the south-looking measurement. The two measurements differ by about 100 m/s during 03:00–08:30 UT. If it is assumed that the north-looking measurement accurately represents the true meridional wind, this roughly matches the model result in Figure 5.11. The large spike seen at 04:45 UT in Figure 5.11 occurred during the FPI's calibration routine (as described in Section 4.1), so it is not captured in the data.

The bottom panel of Figure 5.12 shows the temperature measurements from the MH FPI. Under the assumption that the actual temperature field is constant, these results indicate temperature errors reached 300 K at 03:45 UT, about three times larger than the model result in Figure 5.11. During 06:30–08:00 UT, the measured north-south difference of 100–200 K roughly matches the model error. A more quantitative comparison is complicated by the data gap and the large temperature gradients that are expected, but in

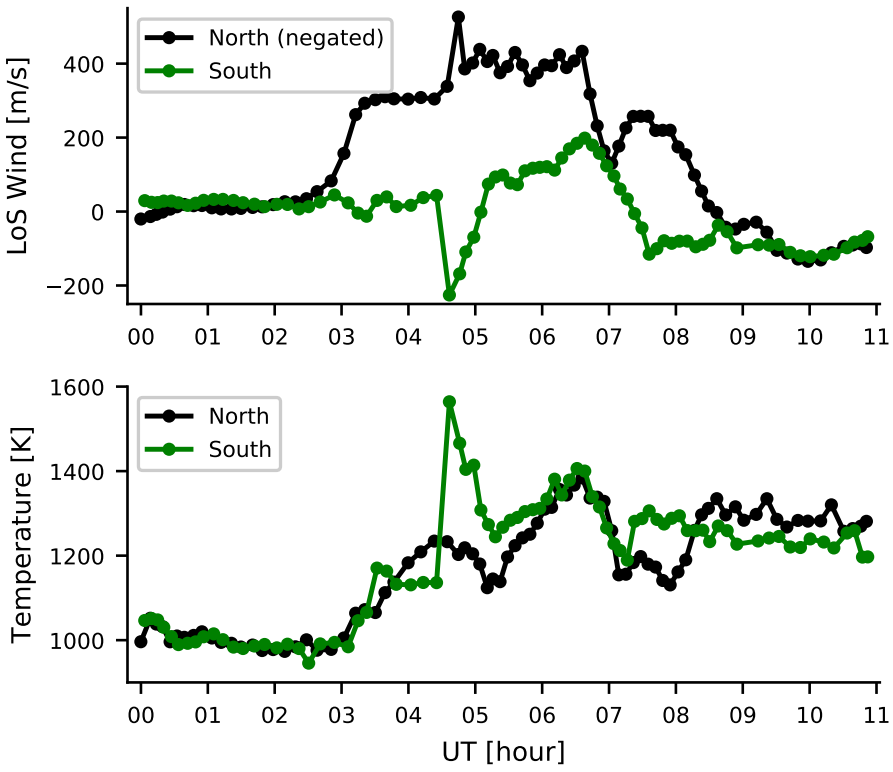

Figure 5.13: Measured line-of-sight wind and temperature by the ANN FPI during the geomagnetic storm of 01–02 Oct 2013. Line of sight is defined positive away from the FPI.

general, the data corroborate the model prediction of hotter temperatures in the south.

The large spike seen at 04:45 UT, while not seen at MH, is seen at ANN, which is located ∼1000 km to the west of MH. The ANN data have a higher cadence and no data gaps. Since it is at the same latitude, we expect it to experience similar conditions to MH, though the wind, temperature, airglow gradients, and optical thickness could be different. Data from ANN are shown in Figure 5.13 in the same format as Figure 5.12. The wind data clearly show an apparent convergence during 03:00–08:30 UT (except a brief period around 07:00 UT), and a spike at 04:45 UT. This convergence is much larger than at MH, presumably either due to the larger airglow gradients or optical thickness. The temperature data show an apparent southward gradient during 04:45–06:00 UT and 07:15–08:15 UT, but, as above, the

interpretation of apparent temperature gradients is complicated by the large actual temperature gradients that are expected.

The apparent wind convergence caused by scattering is likely the explanation for the horizontal convergence seen in Figure 4.2 and the large convergences often reported at midlatitudes during storms [e.g., *Hernandez and Roble*, 1976; *Hernandez et al.*, 1982; *Biondi and Sipler*, 1985]. It also may explain some convergences reported in the auroral region near the auroral arcs [e.g., *Anderson et al.*, 2011]. However, there are many artifacts which can cause apparent convergences, such as inaccurate Doppler referencing, hydroxyl contamination, instrumental or laser calibration fluctuations, and actual vertical winds that are not properly handled in the analysis.

The temperature result disagrees with *Abreu et al.* [1983], who claim that temperature measurements are unaffected by scattering. According to our model, this is an erroneous claim. Unfortunately, it has been used by many authors to argue that their results are unaffected by scatter. A common argument is that since temperatures measured in different directions are different, atmospheric scattering effects must not be important [e.g., *Hernandez et al.*, 1982; *Sica*, 1984; *Price et al.*, 1995]. The work shown here invalidates this argument, and brings into question previous reports of horizontal convergences, temperatures, and vertical winds in the midlatitude thermosphere. For some studies mentioned here, for example *Hernandez et al.* [1982], the FPIs were located at high altitude where atmospheric scattering effects would be expected to be less effective. However, stray light alone can cause apparent vertical winds and convergences.

This result elucidates the difficulty of interpreting FPI measurements, since the errors introduced by atmospheric scattering can appear to be physically self-consistent. The bright airglow enhancement in the north is an indicator of particle precipitation, and it is also a region of auroral currents and high-speed ion flows. The heating from these sources would be expected to cause vertical winds. Downward winds should be associated with a converging wind field (via the Burnside relation) and a temperature increase. Furthermore, instruments farther from the aurora would experience reduced effects. All of these physically justifiable features are seen in our data from the 02 Oct 2013 storm, yet they are all likely artifacts of atmospheric scattering.

The reason that atmospheric scattering is so significant during stormtime is the presence of a large horizontal wind and a large brightness gradient. Dur-

ing quiet times, the wind is smaller and the brightness gradient is usually negligible, at least at midlatitudes. We performed an additional model calculation identical to the one shown in Figure 5.11, except we set $f(x, y) = 1$, in order to eliminate the brightness gradient. The results, shown in Figure 5.14, indicate that the measured temperature is artificially high by several tens of K during the period of largest horizontal winds (∼600 m/s), and the error is only a few K when winds are small early in the night. The line-of-sight wind error is roughly proportional to the horizontal wind, introducing a ∼13% bias towards zero. Of course, when the horizontal wind is zero, the scattered spectrum is identical to the direct spectrum, so the wind and temperature error are both zero. Although this simulation used a south-looking measurement, the north-looking measurement would be identically affected. These results suggest that even under quiet conditions, atmospheric scattering effects should be considered.

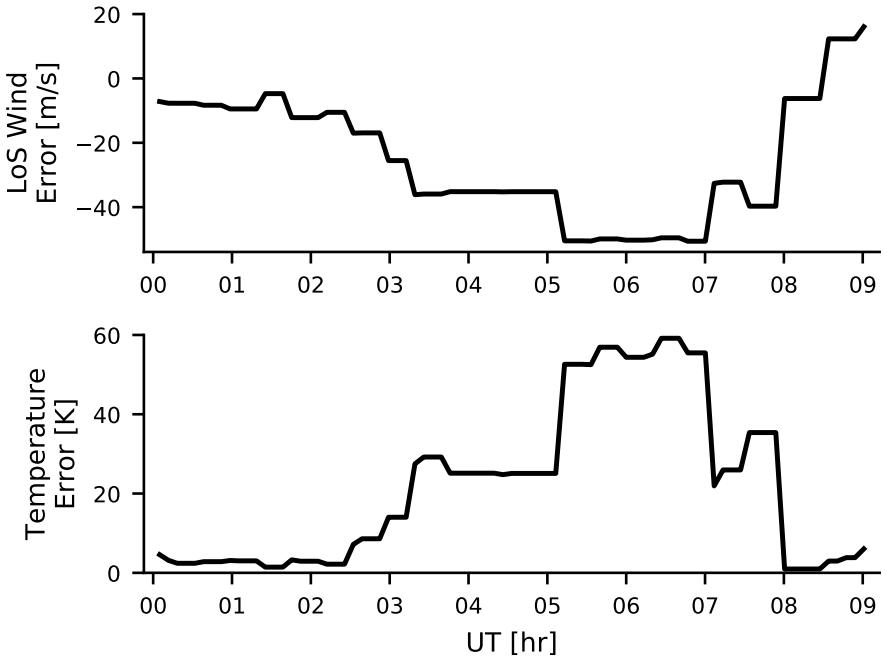

Figure 5.14: Same as Figure 5.11 except with $f(x, y) = 1$ (i.e., all-sky imager data are not used, and airglow brightness is constant).

5.6 Application to Equatorial Latitudes

In this section, we apply our radiative transfer model to equatorial latitudes. In particular, we study the vertical winds reported by *Fisher et al.* [2015], who performed a climatological study of 5 years of data from the CAR and CAJ FPIs in Brazil. Near the December solstice, they observe a consistent 20 m/s downward wind in the first hour after sunset. Near the March equinox, there is also a post-sunset downward wind in the first hour, but it is not as strong, and it is followed in the second hour by a 10 m/s upward wind. Pre-sunrise vertical winds are also reported, but here we focus on the post-sunset period. At all other times of the night, and in different seasons, the vertical winds are statistically near zero.

We run the radiative transfer model for 21 nights in the 2013–2015 period, spanning multiple seasons to determine if the seasonal dependence of the apparent vertical winds caused by atmospheric scattering matches the seasonal dependence reported by *Fisher et al.* [2015].

5.6.1 Inputs

To evaluate the hypothesis of whether the *Fisher et al.* [2015] vertical winds can be explained as artifacts of atmospheric scattering, we use data from an all-sky imager that is co-located with the CAR FPI. This all-sky camera belongs to the Brazilian National Institute for Space Research (INPE) and is operated by the Federal University of Campina Grande (UFCG). The airglow imager includes the appropriate lenses to form an all-sky telecentric image near the 3-inch diameter filter plane, and to re-image achromatically onto the CCD chip. The primary (front) lens in the 3-inch system is a 24 mm, f/4.0 medium-format achromatic fisheye lens, while the final lens in front of the CCD is an ultra-fast f/0.95 Canon 50 mm objective. The OI 630.0-nm filter has a bandwidth of 2.0 nm and is mounted in the filter wheel. The CCD is equipped with a back-illuminated e2v CCD47-10 chip with 1024 $\times$ 1024 pixels, 13.3-μm square pixels with a total area of 13.3 mm $\times$ 13.3 mm. In order to enhance the signal-to-noise ratio, the images were binned down to a resolution of 512 $\times$ 512 pixels. The CCD has a high quantum efficiency in the visible (greater than 95% at 630.0 nm) and NIR spectrum and a low dark current (4×10^{-4} electrons/pixel/sec at -70 °C).

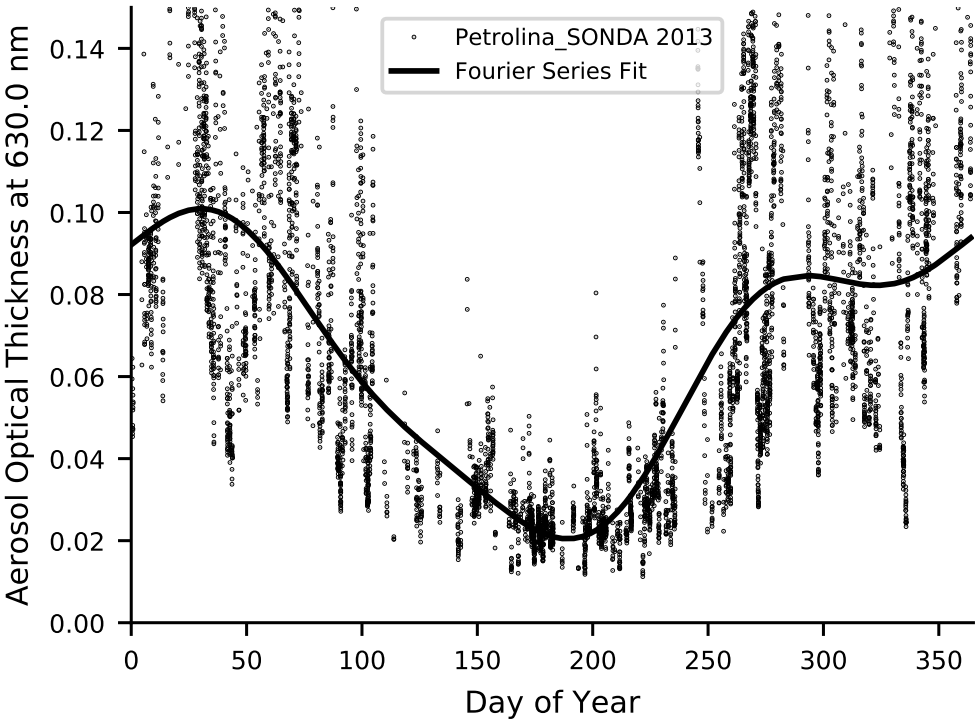

Figure 5.15: Aerosol optical thickness data from the Petrolina_SONDA AERONET site during 2013, interpolated to 630.0 nm. Optical thicknesses larger than 0.15 have been omitted. A third-order Fourier series fit is used to determine the seasonal dependence of τ_0^{aero} for use in the radiative transfer model.

The CCD has a thermoelectric cooling system reaching temperatures around -70 °C. We correct the raw data for stray light as described above. No flat-field calibration data are available, so we perform a rudimentary flat-field correction using the known solid angle of each pixel and projected aperture area. The uncertainty with this correction is significant, so the magnitude of the modeled apparent vertical wind should be treated with caution. However, the sign is trustworthy, as are relative variations in the magnitude from night to night.

Unfortunately, the nearest AERONET site (Petrolina_SONDA) is 500 km away from CAR. This adds significant uncertainty to our model result, but, as above, although the magnitude of the apparent vertical wind is uncertain, the sign is correct. In Figure 5.15, we show measured τ_0^{aero} from Petrolina_SONDA during 2013. Optical thicknesses larger than 0.15 (8% of the dataset) have been omitted since they seem unreasonably large; they

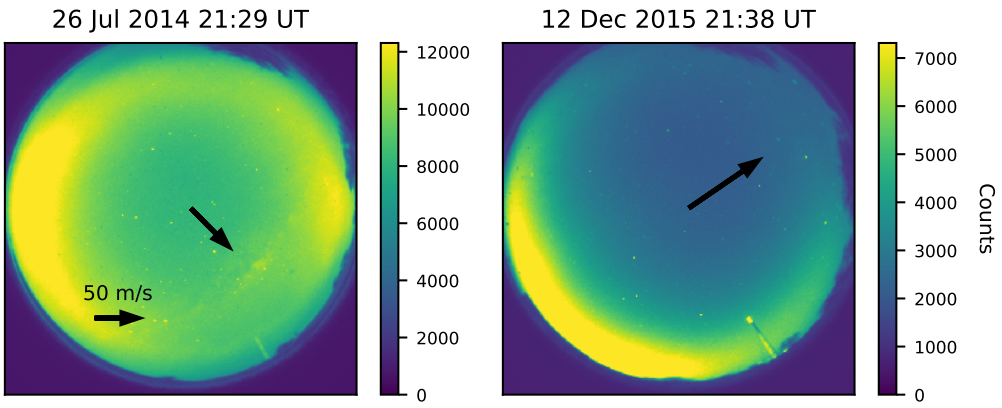

Figure 5.16: Raw all-sky images from the Cariri imager, shown as measured on the pixel grid. North is up and east is right. The arrow in the center indicates the thermospheric wind direction and speed from the HWM14 model. Two post-sunset examples are shown: local winter on the left (when measured vertical winds are near zero) and local summer on the right (when $\sim$20 m/s downward vertical winds are commonly measured).

may have been corrupted by intermittent clouds, for example. The aerosol content is much larger during Sep–Mar season than Mar–Sep. We fit a third-order Fourier series to obtain a seasonal dependence to be used for modeling an arbitrary date. The last full year of operation for Petrolina_SONDA was 2013. We assume that the seasonal dependence determined from 2013 also applies to 2014 and 2015. We use the same scattering phase function as above, since our analysis is not particularly sensitive to small changes in P.

For the midlatitude analysis in Section 5.5, we determined the horizontal wind and temperature using the north-looking FPI measurement. For the equatorial case, we cannot do this, since the airglow gradient is not always purely meridional or zonal, and it varies significantly from season to season and night to night. Moreover, the brightest regions of the sky are often at larger zenith angles than those measured by the FPI. Thus, instead of using FPI data to determine the wind and temperature, we use climatological models. For the wind, we use the Horizontal Wind Model 2014 (HWM14) [*Drob et al.*, 2015]. This model cannot be used for the midlatitude analysis since it does not capture stormtime dynamics. For the temperature, we use the NRLMSISE-00 model [*Picone et al.*, 2002].

In Figure 5.16, we show the HWM14 wind vectors superimposed on two example raw all-sky images from the INPE imager. The left image is near

the June solstice (when vertical winds are statistically near zero) and the right image is near the December solstice (when post-sunset downward vertical winds are measured). The airglow gradient in this case is caused by a combination of the gradient at sunset and the equatorial ionization arcs. In December, the airglow gradient is larger, the wind is faster, the aerosol content is higher, and the wind vector is antiparallel to the airglow gradient. This suggests that the downward winds in December could be explained as atmospheric scattering artifacts.

5.6.2 Model Results

To quantitatively assess this hypothesis, we run the model for select nights during the 2013–2015 period, for which the CAR FPI has higher-quality data than previous years. We selected these nights to satisfy the following criteria:

1. Both the FPI and all-sky imager were operating during the first 2 hours after sunset with no maintenance problems.

2. No clouds were present in a visual inspection of the all-sky images.

3. The Moon was down for the first 2 hours after sunset.

These criteria were met by 21 nights. For each night, we run the radiative transfer model using the inputs described above. We use a larger FPI field of view (2.5° instead of 1.8° as in the previous section) to account for the different FPI design. We average the modeled apparent vertical wind during the first hour of each night, and compare it to the average of the FPI data during this same interval (weighted by the statistical uncertainties). We also compute the average for the second hour.

The results are shown in Figure 5.17, which contains two points for each night: one each for the first and second hour. If the model perfectly matched the data, all points would like on the dashed line, which has a slope of 1. The color of each point represents the average airglow brightness measured by the FPI.

Except for the samples with the lowest brightness, the agreement is good. It is noteworthy that both positive and negative values of the vertical wind seem to be correctly predicted by our model. If the vertical winds reported by *Fisher et al.* [2015] are caused by some kind of spectral contamination or

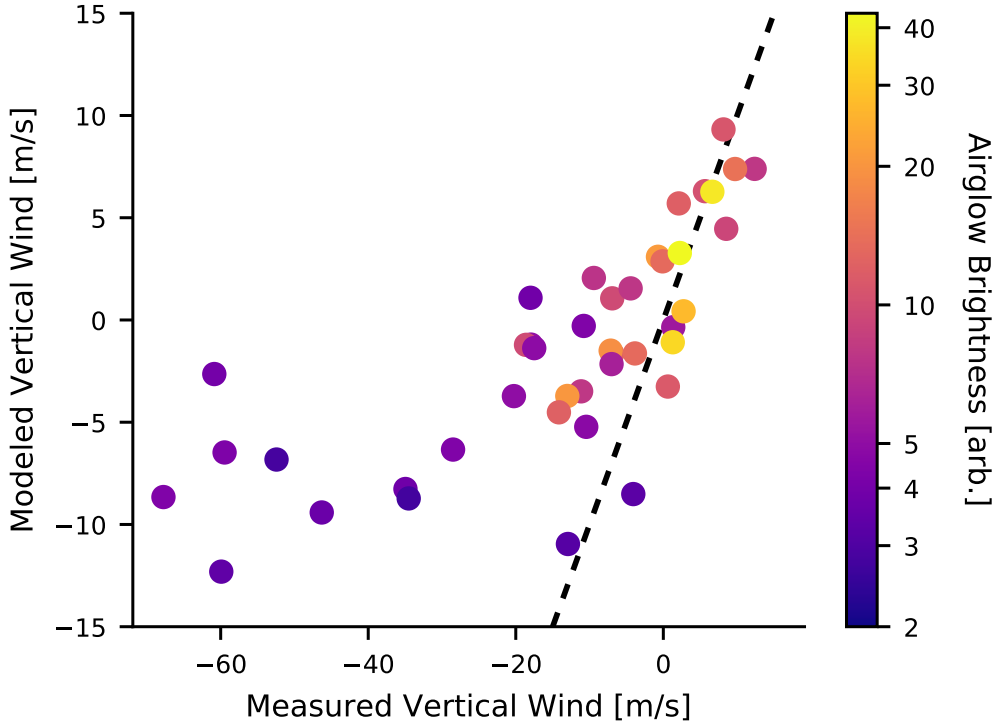

Figure 5.17: The modeled apparent vertical wind (using data from the Cariri all-sky imager) is compared to the measured apparent vertical wind from the CAR FPI for 21 nights during 2013–2015. Two points are plotted per night, averages during the first hour and second hour after sunset. The color of each point is the average airglow brightness measured by the FPI during that hour. The dashed line has a slope of 1, shown for reference. Adapted from *Harding et al.* [Submitted], except modified to include the seasonal dependence of τ_0^{aero}.

systematic error, we would expect purely positive or negative errors. Moreover, if the vertical winds are real, we would not expect an atmospheric scattering model to correctly predict their sign. The fact that the model predicts the sign, more so than the magnitude (which is uncertain), is strong evidence that atmospheric scattering is a significant contributing artifact to the vertical winds reported by *Fisher et al.* [2015].

The low-brightness outliers consist of measured vertical winds that are quite large, almost -70 m/s, The modeled vertical winds, while also negative, are never less than -13 m/s. This difference could be explained by the large uncertainty in the calibration of the all-sky images, the specification of the optical thickness, and the uncertainty of HWM14. However, a more likely

explanation is hydroxyl contamination, though we have been unable to prove this.

5.7 Scattering Correction

We have described a radiative transfer model which, when compared with vertical wind data from midlatitudes and equatorial latitudes, strongly suggests that these vertical winds are artifacts of atmospheric scattering. An obvious question is whether the FPI data can be corrected for the effects of scattering. In theory, the answer is yes. We need to simply model the scattered spectrum and subtract it from the measured spectrum. *Abreu* [1985] describes such a correction that does not use an all-sky camera, but assumes that airglow variations are purely meridional, and the FPI makes a dense north-south scan of measurements.

In practice, the uncertainties with measuring $s(\theta)$, removing stray light from the all-sky data, calibrating the all-sky imager, cross-calibrating intensities measured by the FPI and all-sky imager, and specifying τ_0^{aero} combine to yield apparent vertical wind uncertainties than can reach 100%. Thus, the correction can add considerable uncertainty. Moreover, because it involves a subtraction of two spectra, it multiplies the statistical uncertainty of the original data. For these reasons, we have been unable to obtain meaningful results on real data.

CHAPTER 6

CONCLUSION

This dissertation has described the analysis and interpretation of midlatitude and low-latitude thermospheric wind and temperature data from FPIs, with a focus on geomagnetic storms. We described the design and deployment of eight FPIs to the eastern United States and Brazil. Except for the MH FPI, which is independently operated, these FPIs compose the networks known as NATION and RENOIR. We detailed a new nonlinear-regression-based technique to analyze raw FPI data and generate estimates of line-of-sight wind and temperature. Advantages of this technique include easy implementation, accurate uncertainty estimates, and, in contrast to Fourier-based techniques which do not handle adjacent fringe overlap, accurate temperature estimates. We presented a Monte Carlo simulation which proved the efficacy of this method.

A novel regularization-based technique was presented to estimate the regional (latitude- and longitude-dependent) thermospheric wind vector from a network's line-of-sight wind measurements. An advantage of this technique over previous wind field estimation algorithms is that is does not make any explicit assumptions about the functional form of the wind field, and instead finds the "smoothest" wind field that agrees with the measured line-of-sight winds to within their uncertainty. We defined a "smooth" wind field to be one with minimal gradients and curvatures. This technique was applied to study the thermospheric dynamics associated with the midnight temperature maximum (MTM) using RENOIR data. We showed direct evidence of a converging wind field and a downward vertical wind during the period of heating associated with the MTM. These features were found to be physically self-consistent when compared in the context of the Burnside relation and a simple model of adiabatic heating.

Using these algorithms on NATION data, we performed a case study of the 01–02 Oct 2013 geomagnetic storm. The data showed an equatorward surge

in the horizontal wind and a temperature increase, which were consistent with previous theory and observations. However, large, 100 m/s downward vertical winds were also observed across the eastern United States, lasting several hours. ANN measured the largest downward wind at 300 m/s. We performed a superposed epoch analysis of 15 different storms observed by NATION which showed that downward winds in the main phase of geomagnetic storms are commonly measured.

Although large vertical winds have been seen by FPIs before, the physically unreasonable magnitude, duration, and extent of these observations, combined with the apparent anisotropy in the temperature observations, led us to question the long-held assumption that the Doppler shift and linewidth of the 630.0-nm represent the thermospheric wind and temperature. Two possible explanations were considered: contamination by the non-thermalized emission caused by O^+ precipitation, and contamination by the scattering of airglow radiation in the troposphere. The hypothesis of O^+ precipitation was ruled out using data from the Arecibo FPI (which measured an upward wind), a new observation strategy which began in 2015 (which ruled out a connection with the magnetic field), and the lack of downward wind seen during the 08–09 Oct 2013 storm (which suggested a connection with the equatorward wind, a connection that is explained by atmospheric scattering).

The atmospheric scattering hypothesis was quantitatively evaluated with a radiative transfer model. This model uses optical thickness data from AERONET and brightness data from an all-sky camera to determine the direct and scattered 630.0-nm spectra observed by a co-located FPI. With these spectra, the line-of-sight wind and temperature errors caused by atmospheric scattering were evaluated and compared to data from the 01–02 Oct 2013 storm. The predicted vertical wind error matched the measured vertical wind well, suggesting that atmospheric scattering is fully able to explain the large vertical winds measured at midlatitudes during geomagnetic storms. We also applied our model to low latitudes, showing that the climatological vertical winds reported by *Fisher et al.* [2015] are likely also artifacts of atmospheric scattering.

In some sense, the atmospheric scattering hypothesis is preferable to the O^+ precipitation hypothesis. Atmospheric scattering only affects ground-based measurements, so space- and balloon-based measurements are unaf-

fected. Additionally, given accurate knowledge of the brightness distribution, calibration, and the radiative characteristics of the atmosphere, the effects of atmospheric scattering can in principle be removed from the data, though in practice, this is sensitive to calibration accuracy.

REFERENCES

Abreu, V. (1985), Atmospheric scattering effects on ground-based Fabry-Perot measurements of thermospheric winds: An inversion technique, *Planetary and Space Science*, *33*(9), 1049–1055, doi:10.1016/0032-0633(85)90023-6.

Abreu, V., G. A. Schmitt, P. Hays, J. W. Meriwether, C. Tepley, and L. L. Cogger (1983), Atmospheric scattering effects on ground-based measurements of thermospheric winds, *Planetary and Space Science*, *31*(3), 303–310, doi:10.1016/0032-0633(83)90080-6.

Akmaev, R. A., F. Wu, T. J. Fuller-Rowell, H. Wang, and M. D. Iredell (2010), Midnight density and temperature maxima, and thermospheric dynamics in whole atmosphere model simulations, *Journal of Geophysical Research: Space Physics*, *115*(A8), doi:10.1029/2010JA015651.

Anderson, C., M. G. Conde, P. Dyson, T. Davies, and M. J. Kosch (2009), Thermospheric winds and temperatures above Mawson, Antarctica, observed with an all-sky imaging, Fabry-Perot spectrometer, *Annales Geophysicae*, *27*(5), 2225–2235, doi:10.5194/angeo-27-2225-2009.

Anderson, C., T. Davies, M. G. Conde, P. Dyson, and M. J. Kosch (2011), Spatial sampling of the thermospheric vertical wind field at auroral latitudes, *Journal of Geophysical Research*, *116*(A6), A06,320, doi:10.1029/2011JA016485.

Anderson, C., M. G. Conde, and M. G. McHarg (2012a), Neutral thermospheric dynamics observed with two scanning Doppler imagers: 2. Vertical winds, *Journal of Geophysical Research*, *117*(A3), A03,305, doi:10.1029/2011JA017157.

Anderson, C., M. G. Conde, and M. G. McHarg (2012b), Neutral thermospheric dynamics observed with two scanning Doppler imagers: 1. Monostatic and bistatic winds, *Journal of Geophysical Research*, *117*(A3), A03,304, doi:10.1029/2011JA017041.

Anderson, C., M. G. Conde, and M. G. McHarg (2012c), Neutral thermospheric dynamics observed with two scanning Doppler imagers: 3. Horizontal wind gradients, *Journal of Geophysical Research*, *117*(A5), A05,311, doi:10.1029/2011JA017471.

Armstrong, E. (1969), Doppler shifts in the wavelength of the OI λ6300 line in the night airglow, *Planetary and Space Science*, *17*(5), 957–974, doi:10.1016/0032-0633(69)90101-9.

Armstrong, S. (2008), Fabry-Perot data analysis and simulation for the RENOIR observatories, Master's thesis, University of Illinois at Urbana-Champaign.

Aruliah, A. L., and D. Rees (1995), The trouble with thermospheric vertical winds: geomagnetic, seasonal and solar cycle dependence at high latitudes, *Journal of Atmospheric and Terrestrial Physics*, *57*(6), 597–609, doi:10.1016/0021-9169(94)00100-3.

Aruliah, A. L., E. M. Griffin, I. McWhirter, A. D. Aylward, E. A. K. Ford, A. Charalambous, M. J. Kosch, C. J. Davis, and V. S. C. Howells (2004), First tristatic studies of meso-scale ion-neutral dynamics and energetics in the high-latitude upper atmosphere using collocated FPIs and EISCAT radar, *Geophys. Res. Lett.*, *31*(3), L03,802, doi:10.1029/2003gl018469.

Aruliah, A. L., E. M. Griffin, A. D. Aylward, E. A. K. Ford, M. J. Kosch, C. J. Davis, V. S. C. Howells, S. E. Pryse, H. R. Middleton, and J. Jussila (2005), First direct evidence of meso-scale variability on ion-neutral dynamics using co-located tristatic FPIs and EISCAT radar in Northern Scandinavia, *Annales Geophysicae*, *23*(1), 147–162, doi:10.5194/angeo-23-147-2005.

Aster, R., B. Borchers, and C. Thurber (2013), *Parameter Estimation and Inverse Problems*, Academic Press.

Baumgardner, J. L., J. Wroten, J. Semeter, J. Kozyra, M. Buonsanto, P. Erickson, and M. Mendillo (2008), A very bright SAR arc: Implications for extreme magnetosphere-ionosphere coupling, *Annales Geophysicae*, *25*(12), 2593–2608, doi:10.5194/angeo-25-2593-2007.

Baumjohann, W., and R. Treumann (1997), *Basic Space Plasma Physics*, Imperial College Press.

Bilitza, D., and B. W. Reinisch (2008), International reference ionosphere 2007: Improvements and new parameters, *Advances in Space Research*, *42*(4), 599–609, doi:10.1016/j.asr.2007.07.048.

Biondi, M. A. (1984), Measured vertical motion and converging and diverging horizontal flow of the midlatitude thermosphere, *Geophysical Research Letters*, *11*(1), 84–87, doi:10.1029/GL011i001p00084.

Biondi, M. A., and D. P. Sipler (1985), Horizontal and vertical winds and temperatures in the equatorial thermosphere: Measurements from Natal, Brazil during August-September 1982, *Planetary and Space Science*, *33*(7), 817–823, doi:10.1016/0032-0633(85)90035-2.

Biondi, M. A., D. P. Sipler, M. E. Zipf, and J. L. Baumgardner (1995), All-sky Doppler interferometer for thermospheric dynamics studies, *Applied Optics*, *34*(10), 1646–54.

Bisikalo, D. V., V. I. Shematovich, and J. C. Gérard (1995), A kinetic model of the formation of the hot oxygen geocorona: 2. Influence of O+ ion precipitation, *Journal of Geophysical Research*, *100*(A3), 3715, doi:10.1029/94JA03196.

Blahut, R. (2004), *Theory of Remote Image Formation*, Cambridge University Press.

Bodhaine, B., N. Wood, E. Dutton, and J. Slusser (1999), On Rayleigh optical depth calculations, *Journal of Atmospheric and Oceanic Technology*, *16*, 1854–1861, doi:10.1175/1520-0426(1999)016⟨1854:ORODC⟩2.0.CO;2.

Browning, K. A., and R. Wexler (1968), The determination of kinematic properties of a wind field using Doppler radar, *Journal of Applied Meteorology*, *7*(1), 105–113, doi:10.1175/1520-0450(1968)007⟨0105:TDOKPO⟩2.0.CO;2.

Brum, C. G. M., C. A. Tepley, J. T. Fentzke, E. Robles, P. T. dos Santos, and S. A. Gonzalez (2012), Long-term changes in the thermospheric neutral winds over Arecibo: Climatology based on over three decades of Fabry-Perot observations, *Journal of Geophysical Research*, *117*, A00H14, doi: 10.1029/2011JA016458.

Buonsanto, M. (1999), Ionospheric storms – A review, *Space Science Reviews*, *88*(3), 563–601, doi:10.1023/A:1005107532631.

Burnside, R. G., J. W. Meriwether, and M. Torr (1977), Contamination of ground-based measurements of OI (6300 Å) and NI (5200 Å) airglow by OH emissions, *Planetary and Space Science*, *25*, 985–988.

Burnside, R. G., F. A. Herrero, J. W. Meriwether, and J. C. G. Walker (1981), Optical observations of thermospheric dynamics at Arecibo, *Journal of Geophysical Research*, *86*(A7), 5532, doi:10.1029/JA086iA07p05532.

Chandrasekhar, S. (1960), *Radiative Transfer*, Dover books on advanced mathematics, Dover Publications.

Colerico, M. J., M. Mendillo, C. G. Fesen, and J. W. Meriwether (2006), Comparative investigations of equatorial electrodynamics and low-to-mid latitude coupling of the thermosphere-ionosphere system, *Annales Geophysicae*, *24*(2), 503–513, doi:10.5194/angeo-24-503-2006.

Conde, M. (2002), Deriving wavelength spectra from fringe images from a fixed-gap single-etalon Fabry-Perot spectrometer, *Applied Optics*, *41*(14), 2672–2678, doi:10.1364/AO.41.002672.

Conde, M. G., and P. Dyson (1995), Thermospheric vertical winds above Mawson, Antarctica, *Journal of Atmospheric and Terrestrial Physics*, *57*(6), 589–596, doi:10.1016/0021-9169(94)00099-A.

Conde, M. G., and R. W. Smith (1998), Spatial structure in the thermospheric horizontal wind above Poker Flat, Alaska, during solar minimum, *Journal of Geophysical Research*, *103*(A5), 9449, doi:10.1029/97JA03331.

Cosgrove, R. B., H. Bahcivan, S. Chen, R. J. Strangeway, J. Ortega, M. Alhassan, Y. Xu, M. V. Welie, J. Rehberger, S. Musielak, and N. Cahill (2014), Empirical model of Poynting flux derived from FAST data and a

cusp signature, *Journal of Geophysical Research: Space Physics*, *119*(1), 411–430, doi:10.1002/2013JA019105.

Cousins, E. D. P., T. Matsuo, and A. D. Richmond (2015), Mapping high-latitude ionospheric electrodynamics with SuperDARN and AMPERE, *Journal of Geophysical Research A: Space Physics*, *120*(7), 5854–5870, doi:10.1002/2014JA020463.

Crickmore, R., J. Dudeney, and A. Rodger (1991), Vertical thermospheric winds at the equatorward edge of the auroral oval, *Journal of Atmospheric and Terrestrial Physics*, *53*(6-7), 485–492, doi:10.1016/0021-9169(91)90076-J.

Crickmore, R. I. (1993), A comparison between vertical winds and divergence in the high-latitude thermosphere, *Annales Geophysicae*, *11*, 728–733.

Data Analysis Center for Geomagnetism and Space Magnetism, Kyoto University (2017), World Data Center for Geomagnetism, Kyoto, `http://wdc.kugi.kyoto-u.ac.jp/`.

DeMajistre, R., P. C. Brandt, T. J. Immel, J. Yee, A. Dalgarno, L. J. Paxton, and V. Kharchenko (2005), Storm-time enhancement of mid-latitude ultraviolet emissions due to energetic neutral atom precipitation, *Geophysical Research Letters*, *32*(15), L15,105, doi:10.1029/2005GL023059.

Deng, Y., A. D. Richmond, A. J. Ridley, and H.-L. Liu (2008), Assessment of the non-hydrostatic effect on the upper atmosphere using a general circulation model (GCM), *Geophysical Research Letters*, *35*(1), L01,104, doi:10.1029/2007GL032182.

Drob, D. P., J. T. Emmert, J. W. Meriwether, J. J. Makela, E. Doornbos, M. G. Conde, G. Hernandez, J. Noto, K. A. Zawdie, S. E. McDonald, J. D. Huba, and J. H. Klenzing (2015), An update to the Horizontal Wind Model (HWM): The quiet time thermosphere, *Earth and Space Science*, *2*(7), 301–319, doi:10.1002/2014EA000089.

England, S. L. (2011), A review of the effects of non-migrating atmospheric tides on the Earth's low-latitude ionosphere, *Space Science Reviews*, *168*(1-4), 211–236, doi:10.1007/s11214-011-9842-4.

Fisher, D. J. (2013), Three-dimensional wind measurements and modeling using a bi-static Fabry-Perot interferometer system in Brazil, Master's thesis, University of Illinois at Urbana-Champaign.

Fisher, D. J., J. J. Makela, J. W. Meriwether, R. A. Buriti, Z. Benkhaldoun, M. Kaab, and A. Lagheryeb (2015), Climatologies of nighttime thermospheric winds and temperatures from Fabry-Perot interferometer measurements: From solar minimum to solar maximum, *Journal of Geophysical Research A: Space Physics*, *120*(8), 6679–6693, doi:10.1002/2015JA021170.

Fuller-Rowell, T. J. (2011), Storm-time response of the thermosphere-ionosphere system, in *Aeronomy of the Earth's Atmosphere and Ionosphere*, edited by M. A. Abdu and D. Pancheva, chap. 32, pp. 419–435, Springer Netherlands, Dordrecht, doi:10.1007/978-94-007-0326-1.

Fuller-Rowell, T. J., and M. Codrescu (1997), How does the thermosphere and ionosphere react to a geomagnetic storm?, *Geophysical Monograph Series*, doi:10.1029/GM098p0203.

Galand, M., and S. Chakrabarti (2006), Proton aurora observed from the ground, *Journal of Atmospheric and Solar-Terrestrial Physics*, *68*(13), 1488–1501, doi:10.1016/j.jastp.2005.04.013.

Gander, W., G. H. Golub, and R. Strebel (1994), Least-squares fitting of circles and ellipses, *BIT*, *34*(4), 558–578, doi:10.1007/BF01934268.

Gerrard, A. J., and J. W. Meriwether (2011), Initial daytime and nighttime SOFDI observations of thermospheric winds from Fabry-Perot Doppler shift measurements of the 630-nm OI line-shape profile, *Annales Geophysicae*, *29*(9), 1529–1536, doi:10.5194/angeo-29-1529-2011.

Goldstein, J., and D. J. McComas (2013), Five years of stereo magnetospheric imaging by TWINS, *Space Science Reviews*, *180*(1-4), 39–70, doi:10.1007/s11214-013-0012-8.

Greet, P., M. G. Conde, and P. Dyson (1999), Thermospheric wind field over Mawson and Davis, Antarctica; simultaneous observations by two Fabry-Perot spectrometers of λ630 nm emission, *Journal of Atmospheric and Solar-Terrestrial Physics*, *61*.

Greet, P. A., J. L. Innis, and P. L. Dyson (2002), Thermospheric vertical winds in the auroral oval/polar cap region, *Ann. Geophys.*, *20*(12), 1987–2001, doi:10.5194/angeo-20-1987-2002.

Guo, W., and D. McEwen (2003), Vertical winds in the central polar cap, *Geophys. Res. Lett.*, *30*(14), 1–4, doi:10.1029/2003GL017124.

Hansen, J. E., and L. D. Travis (1974), Light scattering in planetary atmospheres, *Space Science Reviews*, *16*(4), 527–610, doi:10.1007/BF00168069.

Hansen, P., and D. O'Leary (1993), The use of the L-curve in the regularization of discrete ill-posed problems, *SIAM Journal on Scientific Computing*, *14*(6), 1487–1503.

Harding, B. J., T. W. Gehrels, and J. J. Makela (2014), Nonlinear regression method for estimating neutral wind and temperature from Fabry-Perot interferometer data, *Applied Optics*, *53*(4), 666, doi:10.1364/AO.53.000666.

Harding, B. J., J. J. Makela, and J. W. Meriwether (2015), Estimation of mesoscale thermospheric wind structure using a network of interferometers, *Journal of Geophysical Research: Space Physics*, *120*(5), 3928–3940, doi:10.1002/2015JA021025.

Harding, B. J., J. J. Makela, J. Qin, D. J. Fisher, C. R. Martinis, J. Noto, and C. M. Wrasse (Submitted), Atmospheric scattering effects on ground-based measurements of thermospheric vertical wind, horizontal wind, and temperature, *Journal of Geophysical Research: Space Physics*.

Häusler, K., and H. Lühr (2009), Nonmigrating tidal signals in the upper thermospheric zonal wind at equatorial latitudes as observed by CHAMP, *Annales Geophysicae*, *27*(7), 2643–2652, doi:10.5194/angeo-27-2643-2009.

Hays, P. B., and R. G. Roble (1971a), A technique for recovering Doppler line profiles from Fabry-Perot interferometer fringes of very low intensity, *Applied Optics*, *10*(1), 193–200.

Hays, P. B., and R. G. Roble (1971b), Direct observations of thermospheric winds during geomagnetic storms, *Journal of Geophysical Research*, *76*(22), 5316–5321, doi:10.1029/JA076i022p05316.

Hedin, A. E. (1983), A revised thermospheric model based on mass spectrometer and incoherent scatter data: MSIS-83, *Journal of Geophysical Research*, *88*(A12), 10,170–10,188, doi:10.1029/JA088iA12p10170.

Heelis, R. (2004), Electrodynamics in the low and middle latitude ionosphere: A tutorial, *Journal of Atmospheric and Solar-Terrestrial Physics*, *66*(10), 825–838, doi:10.1016/j.jastp.2004.01.034.

Hernandez, G. (1974), Contamination of the OI(3P2-1D2) emission line by the (9-3) band of OH X2II in high-resolution measurements of the night sky, *Journal of Geophysical Research*, *79*(7), 1119–1123, doi:10.1029/JA079i007p01119.

Hernandez, G. (1982), Vertical motions of the neutral thermosphere at midlatitude, *Geophysical Research Letters*, *9*(5), 555–557, doi:10.1029/GL009i005p00555.

Hernandez, G. (1988), *Fabry-Perot Interferometers*, Cambridge Studies in Modern Optics, Cambridge University Press.

Hernandez, G., and R. G. Roble (1976), Direct measurements of nighttime thermospheric winds and temperatures, 2. Geomagnetic storms, *Journal of Geophysical Research*, *81*(28), 5173–5181, doi:10.1029/JA081i028p05173.

Hernandez, G., R. G. Roble, E. C. Ridley, and J. H. Allen (1982), Thermospheric response observed over Fritz Peak, Colorado, during two large geomagnetic storms near solar cycle maximum, *Journal of Geophysical Research*, *87*(A11), 9181, doi:10.1029/JA087iA11p09181.

Herrero, F. A., and N. W. Spencer (1982), On the horizontal distribution of the equatorial thermospheric midnight temperature maximum and its seasonal variation, *Geophysical Research Letters*, *9*(10), 1179–1182, doi:10.1029/GL009i010p01179.

Hickey, M. P., P. G. Richards, and D. G. Torr (1995), New sources for the hot oxygen geocorona: Solar cycle, seasonal, latitudinal, and diurnal variations, *Journal of Geophysical Research*, *100*(A9), 17,377, doi:10.1029/95JA00895.

Hocke, K., and K. Schlegel (1996), A review of atmospheric gravity waves and travelling ionospheric disturbances: 1982-1995, *Annales Geophysicae*, *14*(9), 917, doi:10.1007/s005850050357.

Holben, B. N., D. Tanré, A. Smirnov, T. F. Eck, I. Slutsker, N. Abuhassan, W. W. Newcomb, J. S. Schafer, B. Chatenet, F. Lavenu, Y. J. Kaufman, J. V. Castle, A. Setzer, B. Markham, D. Clark, R. Frouin, R. Halthore, A. Karneli, N. T. O'Neill, C. Pietras, R. T. Pinker, K. Voss, and G. Zibordi (2001), An emerging ground-based aerosol climatology: Aerosol optical depth from AERONET, *Journal of Geophysical Research: Atmospheres*, *106*(D11), 12,067–12,097, doi:10.1029/2001JD900014.

Hubert, B., J.-C. Gérard, T. L. Killeen, Q. Wu, D. V. Bisikalo, and V. I. Shematovich (2001), Observation of anomalous temperatures in the daytime O(1D) 6300 Å thermospheric emission: A possible signature of nonthermal atoms, *Journal of Geophysical Research*, *106*(A7), 12,753, doi:10.1029/2000JA900122.

Hysell, D. L. (2007), Inverting ionospheric radio occultation measurements using maximum entropy, *Radio Science*, *42*(4), 1–8, doi:10.1029/2007RS003635.

Hysell, D. L., and J. L. Chau (2006), Optimal aperture synthesis radar imaging, *Radio Science*, *41*(2), RS2003, doi:10.1029/2005RS003383.

Hysell, D. L., M. F. Larsen, and M. P. Sulzer (2014), High time and height resolution neutral wind profile measurements across the mesosphere/lower thermosphere region using the Arecibo incoherent scatter radar, *Journal of Geophysical Research: Space Physics*, *119*(3), 2345–2358, doi:10.1002/2013JA019621.

Innis, J. L., and M. G. Conde (2002), High-latitude thermospheric vertical wind activity from Dynamics Explorer 2 Wind and Temperature Spectrometer observations: Indications of a source region for polar cap gravity waves, *Journal of Geophysical Research: Space Physics*, *107*(A8), 1–17, doi:10.1029/2001JA009130.

Innis, J. L., P. Greet, and P. Dyson (1996), Fabry-Perot spectrometer observations of the auroral oval/polar cap boundary above Mawson, Antarctica,

Journal of Atmospheric and Terrestrial Physics, *58*(16), 1973–1988, doi: 10.1016/0021-9169(96)00007-4.

Innis, J. L., P. Dyson, and P. Greet (1997), Further observations of the thermospheric vertical wind at the auroral oval/polar cap boundary, *Journal of Atmospheric and Solar-Terrestrial Physics*, *59*(16), 2009–2022, doi: 10.1016/S1364-6826(97)00034-5.

Innis, J. L., P. Greet, D. Murphy, M. G. Conde, and P. Dyson (1999), A large vertical wind in the thermosphere at the auroral oval/polar cap boundary seen simultaneously from Mawson and Davis, Antarctica, *Journal of Atmospheric and Solar-Terrestrial Physics*, *61*(14), 1047–1058, doi: 10.1016/S1364-6826(99)00060-7.

Ishii, M., S. Oyama, S. Nozawa, R. Fujii, E. Sagawa, S. Watari, and H. Shinagawa (1999), Dynamics of neutral wind in the polar region observed with two Fabry-Perot interferometers, *Earth, Planets and Space*, *51*(7-8), 833–844, doi:10.1186/BF03353242.

Ishii, M., M. G. Conde, R. W. Smith, M. Krynicki, E. Sagawa, and S. Watari (2001), Vertical wind observations with two Fabry-Perot interferometers at Poker Flat, Alaska, *Journal of Geophysical Research*, *106*(A6), 10,537, doi:10.1029/2000JA900148.

Ishimoto, M., M. R. Torr, P. G. Richards, and D. G. Torr (1986), The role of energetic O+ precipitation in a mid-latitude aurora, *Journal of Geophysical Research*, *91*(A5), 5793, doi:10.1029/JA091iA05p05793.

Ishimoto, M., G. J. Romick, and C. I. Meng (1994), Model calculation of atmospheric emission caused by energetic O+ precipitation, *Journal of Geophysical Research*, *99*(A1), 435, doi:10.1029/93JA01148.

Keika, K., L. M. Kistler, and P. C. Brandt (2013), Energization of O+ ions in the Earth's inner magnetosphere and the effects on ring current buildup: A review of previous observations and possible mechanisms, *Journal of Geophysical Research: Space Physics*, *118*(7), 4441–4464, doi:10.1002/jgra.50371.

Kerr, R. (2014), Redline Instrument Description, `http://neutralwinds.com/redlinedescribe.html`.

Kharchenko, V., A. Dalgarno, and J. L. Fox (2005), Thermospheric distribution of fast O(1 D) atoms, *Journal of Geophysical Research*, *110*(A12), A12,305, doi:10.1029/2005JA011232.

Killeen, T., and P. B. Hays (1984), Doppler line profile analysis for a multichannel Fabry-Perot interferometer, *Applied Optics*, *23*(4), 612–620.

Killeen, T. L. (1987), Energetics and dynamics of the Earth's thermosphere, *Reviews of Geophysics*, *25*(3), 433, doi:10.1029/RG025i003p00433.

Killeen, T. L., and R. G. Roble (1988), Thermosphere dynamics: Contributions from the first 5 years of the Dynamics Explorer Program, *Reviews of Geophysics*, *26*(2), 329, doi:10.1029/RG026i002p00329.

Kozyra, J. U., T. E. Cravens, and A. F. Nagy (1982), Energetic O+ precipitation, *Journal of Geophysical Research*, *87*(A4), 2481, doi:10.1029/JA087iA04p02481.

Kozyra, J. U., A. F. Nagy, and D. W. Slater (1997), High-altitude energy source(s) for stable auroral red arcs, *Reviews of Geophysics*, *35*(2), 155, doi:10.1029/96RG03194.

Kudeki, E., A. Akgiray, M. Milla, J. L. Chau, and D. L. Hysell (2007), Equatorial spread-F initiation: Post-sunset vortex, thermospheric winds, gravity waves, *Journal of Atmospheric and Solar-Terrestrial Physics*, *69*(17-18), 2416–2427, doi:10.1016/j.jastp.2007.04.012.

Larsen, M. F. (2002), Winds and shears in the mesosphere and lower thermosphere: Results from four decades of chemical release wind measurements, *Journal of Geophysical Research*, *107*(A8), 1215, doi:10.1029/2001JA000218.

Larsen, M. F., and J. W. Meriwether (2012), Vertical winds in the thermosphere, *Journal of Geophysical Research*, *117*(A9), A09,319, doi:10.1029/2012JA017843.

Link, R., and L. L. Cogger (1988), A reexamination of the O I 6300-Å nightglow, *Journal of Geophysical Research*, *93*(A9), 9883, doi:10.1029/JA093iA09p09883.

Link, R., and L. L. Cogger (1989), Correction to "A reexamination of the O I 6300-Å nightglow" by R. Link and L. L. Cogger, *Journal of Geophysical Research*, *94*(A2), 1556, doi:10.1029/JA094iA02p01556.

Liu, H.-L., J. M. McInerney, S. Santos, P. H. Lauritzen, M. A. Taylor, and N. M. Pedatella (2014), Gravity waves simulated by high-resolution whole atmosphere community climate model, *Geophysical Research Letters*, *41*(24), 9106–9112, doi:10.1002/2014GL062468.

Lummerzheim, D., and M. Galand (2001), The profile of the hydrogen H β emission line in proton aurora, *Journal of Geophysical Research*, *106*(A1), 23, doi:10.1029/2000JA002014.

Lummerzheim, D., M. Galand, J. Semeter, M. J. Mendillo, M. H. Rees, and F. J. Rich (2001), Emission of OI(630 nm) in proton aurora, *Journal of Geophysical Research*, *106*(A1), 141, doi:10.1029/2000JA002005.

Makela, J. J., J. W. Meriwether, Y. Huang, and P. J. Sherwood (2011), Simulation and analysis of a multi-order imaging Fabry-Perot interferometer for the study of thermospheric winds and temperatures., *Applied Optics*, *50*(22), 4403–16.

Makela, J. J., J. W. Meriwether, A. J. Ridley, M. Ciocca, and M. W. Castellez (2012), Large-scale measurements of thermospheric dynamics with a multisite Fabry-Perot interferometer network: Overview of plans and results from midlatitude measurements, *International Journal of Geophysics*, *2012*(3), 1–10, doi:10.1155/2012/872140.

Makela, J. J., D. J. Fisher, J. W. Meriwether, R. A. Buriti, and A. F. Medeiros (2013), Near-continual ground-based nighttime observations of thermospheric neutral winds and temperatures over equatorial Brazil from 2009 to 2012, *Journal of Atmospheric and Solar-Terrestrial Physics*, *103*, 94–102, doi:10.1016/j.jastp.2012.11.019.

Makela, J. J., B. J. Harding, J. W. Meriwether, R. Mesquita, S. Sanders, A. J. Ridley, M. W. Castellez, M. Ciocca, G. D. Earle, N. A. Frissell, D. L. Hampton, A. J. Gerrard, J. Noto, and C. R. Martinis (2014), Storm time response of the midlatitude thermosphere: Observations from a network

of Fabry-Perot interferometers, *Journal of Geophysical Research: Space Physics*, *119*(8), 6758–6773, doi:10.1002/2014JA019832.

McGranaghan, R., D. J. Knipp, T. Matsuo, and E. Cousins (2016), Optimal interpolation analysis of high-latitude ionospheric Hall and Pedersen conductivities: Application to assimilative ionospheric electrodynamics reconstruction, *Journal of Geophysical Research: Space Physics*, *121*(5), 4898–4923, doi:10.1002/2016JA022486.

Mendillo, M. (2006), Storms in the ionosphere: Patterns and processes for total electron content, *Reviews of Geophysics*, *44*(4), RG4001, doi:10.1029/2005RG000193.

Menke, W. (1989), *Geophysical Data Analysis: Discrete Inverse Theory*, International Geophysics, Elsevier Science.

Meriwether, J. W. (2006), Studies of thermospheric dynamics with a Fabry-Perot interferometer network: A review, *Journal of Atmospheric and Solar-Terrestrial Physics*, *68*(13), 1576–1589, doi:10.1016/j.jastp.2005.11.014.

Meriwether, J. W. (2008), Thermospheric dynamics at low and mid-latitudes during magnetic storm activity, in *Midlatitude Ionospheric Dynamics and Disturbances*, pp. 201–219, American Geophysical Union, doi:10.1029/181GM19.

Meriwether, J. W., M. Faivre, C. Fesen, P. Sherwood, and O. Veliz (2008), New results on equatorial thermospheric winds and the midnight temperature maximum, *Annales Geophysicae*, *26*(3), 447–466, doi:10.5194/angeo-26-447-2008.

Miladinovich, D. S., S. Datta-Barua, G. S. Bust, and J. J. Makela (2016), Assimilation of thermospheric measurements for ionosphere-thermosphere state estimation, *Radio Science*, *51*(12), 1818–1837, doi:10.1002/2016RS006098.

National Centers for Environmental Information (2017), National Oceanic and Atmospheric Administration Anonymous FTP Server, `ftp://ftp.ngdc.noaa.gov/STP/GEOMAGNETIC_DATA/INDICES/KP_AP/`.

Nicolls, M. J., R. Cosgrove, and H. Bahcivan (2014), Estimating the vector electric field using monostatic, multibeam incoherent scatter radar measurements, *Radio Science*, *49*(11), 1124–1139, doi:10.1002/2014RS005519.

Osterbrock, D. E., J. P. Fulbright, A. R. Martel, M. J. Keane, S. C. Trager, and G. Basri (1996), Night-sky high-resolution spectral atlas of OH and O2 emission lines for echelle spectrograph wavelength calibration, *Publications of the Astronomical Society of the Pacific*, *108*, 277, doi:10.1086/133722.

Picone, J. M., A. E. Hedin, D. P. Drob, and A. C. Aikin (2002), NRLMSISE-00 empirical model of the atmosphere: Statistical comparisons and scientific issues, *Journal of Geophysical Research: Space Physics*, *107*(A12), 1–16, doi:10.1029/2002JA009430.

Price, G., R. W. Smith, and G. Hernandez (1995), Simultaneous measurements of large vertical winds in the upper and lower thermosphere, *Journal of Atmospheric and Terrestrial Physics*, *57*(6), 631–643, doi:10.1016/0021-9169(94)00103-U.

Prölss, G. (2004), *Physics of the Earth's Space Environment: An Introduction*, Springer.

Prölss, G. W. (1997), Magnetic storm associated perturbations of the upper atmosphere, in *Magnetic Storms*, edited by B. T. Tsurutani, W. D. Gonzalez, Y. Kamide, and J. K. Arballo, American Geophysical Union, Washington, D. C., doi:10.1029/GM098p0227.

Raghavarao, R., W. R. Hoegy, N. W. Spencer, and L. E. Wharton (1993), Neutral temperature anomaly in the equatorial thermosphere - A source of vertical winds, *Geophysical Research Letters*, *20*(11), 1023–1026, doi:10.1029/93GL01253.

Rees, D., R. W. Smith, P. Charleton, F. McCormac, N. Lloyd, and A. Steen (1984), The generation of vertical thermospheric winds and gravity waves at auroral latitudes - I. Observations of vertical winds, *Planetary and Space Science*, *32*(6), 667–684, doi:10.1016/0032-0633(84)90092-8.

Rees, M. H. (1989), *Physics and Chemistry of the Upper Atmosphere*, Cambridge atmospheric and space science series, Cambridge University Press.

Rieger, E. (1974), Neutral air motions deduced from barium releases experiments - I. Vertical winds, *Journal of Atmospheric and Terrestrial Physics*, *36*(8), 1377–1385, doi:10.1016/0021-9169(74)90214-1.

Rishbeth, H. (1972), Thermospheric winds and the F-region: A review, *Journal of Atmospheric and Terrestrial Physics*, *34*(1972), 1–47.

Rishbeth, H. (1991), F-region storms and thermospheric dynamics, *Journal of Geomagnetism and Geoelectricity*, *43*(Supplement1), 513–524, doi:10.5636/jgg.43.Supplement1_513.

Rishbeth, H., and O. Garriott (1969), *Introduction to Ionospheric Physics*, International geophysics series, Academic Press.

Rishbeth, H., T. J. Fuller-Rowell, and A. S. Rodger (1987), F-layer storms and thermospheric composition, *Physica Scripta*, *36*(2), 327–336, doi:10.1088/0031-8949/36/2/024.

Roble, R. G. (1983), Dynamics of the Earth's thermosphere, *Reviews of Geophysics*, *21*(2), 217, doi:10.1029/RG021i002p00217.

Rochon, Y. J. (2000), The retrieval of winds, Doppler temperatures, and emission rates for the WINDII experiment, Ph.D. thesis, York University.

Ronksley, A. (2016), Optical remote sensing of mesoscale thermospheric dynamics above Svalbard and Kiruna, Ph.D. thesis, University of London.

Salah, J. E., and J. M. Holt (1974), Midlatitude thermospheric winds from incoherent scatter radar and theory, *Radio Science*, *9*(2), 301–313, doi:10.1029/RS009i002p00301.

Schmitt, G., V. Abreu, and P. Hays (1981), Non-thermal O(1D) produced by dissociative recombination of O2+: A theoretical model and observational results, *Planetary and Space Science*, *29*(10), 1095–1099, doi:10.1016/0032-0633(81)90008-8.

Schunk, R., and A. Nagy (2009), *Ionospheres: Physics, Plasma Physics, and Chemistry*, Cambridge Atmospheric and Space Science Series, Cambridge University Press.

Schunk, R., and J. Sojka (1996), Ionosphere-thermosphere space weather issues, *Journal of Atmospheric and Terrestrial Physics*, *58*(14), 1527–1574, doi:10.1016/0021-9169(96)00029-3.

Sharp, R. D., R. G. Johnson, and E. G. Shelley (1976), The morphology of energetic O + ions during two magnetic storms: Latitudinal variations, *Journal of Geophysical Research*, *81*(19), 3292–3298, doi:10.1029/JA081i019p03292.

Shelley, E. G., R. G. Johnson, and R. D. Sharp (1972), Satellite observations of energetic heavy ions during a geomagnetic storm, *Journal of Geophysical Research*, *77*(31), 6104–6110, doi:10.1029/JA077i031p06104.

Shematovich, V., J.-C. Gérard, D. V. Bisikalo, and B. Hubert (1999), Thermalization of O(1D) atoms in the thermosphere, *Journal of Geophysical Research*, *104*(A3), 4287, doi:10.1029/1998JA900154.

Shepherd, G. G., G. Thuillier, Y.-M. Cho, M.-L. Duboin, W. F. J. Evans, W. A. Gault, C. Hersom, D. J. W. Kendall, C. Lathuillère, R. P. Lowe, I. C. McDade, Y. J. Rochon, M. G. Shepherd, B. H. Solheim, D.-Y. Wang, and W. E. Ward (2012), The Wind Imaging Interferometer (WINDII) on the Upper Atmosphere Research Satellite: A 20 year perspective, *Reviews of Geophysics*, *50*(2), RG2007, doi:10.1029/2012RG000390.

Shiokawa, K., T. Kadota, M. K. Ejiri, Y. Otsuka, Y. Katoh, M. Satoh, and T. Ogawa (2001), Three-channel imaging Fabry-Perot interferometer for measurement of mid-latitude airglow, *Applied Optics*, *40*(24), 4286–4296.

Shiokawa, K., Y. Otsuka, S. Oyama, S. Nozawa, M. Satoh, Y. Katoh, Y. Hamaguchi, Y. Yamamoto, and J. W. Meriwether (2012), Development of low-cost sky-scanning Fabry-Perot interferometers for airglow and auroral studies, *Earth, Planets and Space*, *64*(11), 1033–1046, doi:10.5047/eps.2012.05.004.

Sica, R. (1984), Auroral zone thermospheric dynamics using Fabry-Perot interferometric measurements of the OI 15867 K emission, Ph.D. thesis, Univ. of Alaska, Fairbanks.

Sipler, D., M. A. Biondi, and M. Zipf (1995), Vertical winds in the mid-latitude thermosphere from Fabry-Perot interferometer measurements,

Journal of Atmospheric and Terrestrial Physics, *57*(6), 621–629, doi: 10.1016/0021-9169(94)00102-T.

Sipler, D. P., and M. A. Biondi (2003), Simulation of hot oxygen effects on ground-based Fabry-Perot determinations of thermospheric temperatures, *Journal of Geophysical Research*, *108*(A6), 1260, doi:10.1029/2003JA009911.

Smith, R. W. (1998), Vertical winds: A tutorial, *Journal of Atmospheric and Solar-Terrestrial Physics*, *60*(14), 1425–1434, doi:10.1016/S1364-6826(98)00058-3.

Smith, R. W., and G. Hernandez (1995), Vertical winds in the thermosphere within the polar cap, *Journal of Atmospheric and Terrestrial Physics*, *57*(6), 611–620, doi:10.1016/0021-9169(94)00101-S.

Sobolev, V. (1975), *Light Scattering in Planetary Atmospheres*, Pergamon Press, Oxford.

Solomon, S., and V. Abreu (1989), The 630 nm dayglow, *Journal of Geophysical Research*, *94*, 6817–6824.

Solomon, S., P. B. Hays, and V. Abreu (1988), The auroral 6300 Å emission: Observations and modeling, *Journal of Geophysical Research*, *93*(A9), 9867, doi:10.1029/JA093iA09p09867.

Spencer, N. W., R. F. Theis, L. E. Wharton, and G. R. Carignan (1976), Local vertical motions and kinetic temperature from AE-C as evidence for aurora-induced gravity waves, *Geophysical Research Letters*, *3*(6), 313–316, doi:10.1029/GL003i006p00313.

Spencer, N. W., G. R. Carignan, H. G. Mayr, H. B. Niemann, R. F. Theis, and L. E. Wharton (1979), The midnight temperature maximum in the Earth's equatorial thermosphere, *Geophysical Research Letters*, *6*(6), 444–446, doi:10.1029/GL006i006p00444.

Spencer, N. W., L. E. Wharton, G. R. Carignan, and J. C. Maurer (1982), Thermosphere zonal winds, vertical motions and temperature as measured from Dynamics Explorer, *Geophysical Research Letters*, *9*(9), 953–956, doi: 10.1029/GL009i009p00953.

Stephan, A. W., S. Chakrabarti, and D. M. Cotton (2000), Evidence of ENA precipitation in the EUV dayglow, *Geophysical Research Letters*, *27*(18), 2865–2868, doi:10.1029/2000GL000040.

Stephan, A. W., K. F. Dymond, S. A. Budzien, S. E. Thonnard, and R. P. McCoy (2004), Oxygen aurora during the recovery phase of a major geomagnetic storm, *Journal of Geophysical Research*, *109*(A9), A09,208, doi:10.1029/2004JA010557.

Stober, G., and J. L. Chau (2015), A multistatic and multifrequency novel approach for specular meteor radars to improve wind measurements in the MLT region, *Radio Science*, *50*, 1–12, doi:10.1002/2014RS005591.

Thirupathaiah, P., and V. Singh (2014), An updated model of atomic oxygen redline dayglow emission, *Advances in Space Research*, doi:10.1016/j.asr.2014.05.022.

Tinsley, B. A. (1981), Neutral atom precipitationa review, *Journal of Atmospheric and Terrestrial Physics*, *43*(5-6), 617–632, doi:10.1016/0021-9169(81)90124-0.

Torr, M., and D. Torr (1979), Energetic oxygen: A direct coupling mechanism between the magnetosphere and thermosphere, *Geophysical Research Letters*, *6*(9), 700–702.

Torr, M., D. Torr, R. G. Roble, and E. C. Ridley (1982), The dynamic response of the thermosphere to the energy influx resulting from energetic O+ ions, *Journal of Geophysical Research*, *87*, 5290–5300.

Torr, M. R., and D. G. Torr (1982), The role of metastable species in the thermosphere, *Reviews of Geophysics*, *20*(1), 91, doi:10.1029/RG020i001p00091.

Torr, M. R., J. C. G. Walker, and D. G. Torr (1974), Escape of fast oxygen from the atmosphere during geomagnetic storms, *Journal of Geophysical Research*, *79*(34), 5267–5271, doi:10.1029/JA079i034p05267.

Wardill, P., and F. Jacka (1986), Vertical motions in the thermosphere over Mawson, Antarctica, *Journal of Atmospheric and Terrestrial Physics*, *48*(3), 289–292, doi:10.1016/0021-9169(86)90104-2.

Weather Underground (2013), Weather History for Blacksburg, VA, http://www.wunderground.com/history/airport/KBCB/2013/10/2/DailyHistory.html?req_city=NA&req_state=NA&req_statename=NA&MR=1.

Wescott, E. M., H. Stenbaek-Nielsen, M. G. Conde, M. Larsen, and D. Lummerzheim (2006), The HEX experiment: Determination of the neutral wind field from 120 to 185 km altitude near a stable premidnight auroral arc by triangulating the drift of rocket-deployed chemical trails, *Journal of Geophysical Research: Space Physics*, *111*(9), 1–11, doi:10.1029/2005JA011002.

Wilson, G. R., and T. E. Moore (2005), Origins and variation of terrestrial energetic neutral atoms outflow, *Journal of Geophysical Research*, *110*(A2), A02,207, doi:10.1029/2003JA010356.

Xiong, C., H. Lühr, and B. G. Fejer (2015), Global features of the disturbance winds during storm time deduced from CHAMP observations, *Journal of Geophysical Research A: Space Physics*, *120*(6), 5137–5150, doi:10.1002/2015JA021302.

Yee, J. (1988), Non-thermal distribution of O(1D) atoms in the night-time thermosphere, *Planetary and Space Science*, *36*(1), 89–97, doi:10.1016/0032-0633(88)90149-3.

Yee, J., J. W. Meriwether, and P. B. Hays (1980), Detection of a corona of fast oxygen atoms during solar maximum, *Journal of Geophysical Research*, *85*(80), 3396–3400.

Yigit, E., and A. J. Ridley (2011), Role of variability in determining the vertical wind speeds and structure, *Journal of Geophysical Research*, *116*(A12), A12,305, doi:10.1029/2011JA016714.

Zhang, S. P., and G. G. Shepherd (2005), On the response of the atomic oxygen red line emission rates to the Sun's energy input: An empirical model deduced from WINDII/UARS global measurements, in *Proc. SPIE*, vol. 5979, edited by K. Schäfer, A. Comerón, J. R. Slusser, R. H. Picard, M. R. Carleer, and N. I. Sifakis, p. 597912, doi:10.1117/12.627150.

Zhang, Y., L. J. Paxton, J. U. Kozyra, H. Kil, and P. C. Brandt (2006), Nightside thermospheric FUV emissions due to energetic neutral atom precipitation during magnetic superstorms, *Journal of Geophysical Research*, *111*(A9), A09,307, doi:10.1029/2005JA011152.